Chapter 1

Introduction

Algebraic geometry is the study of geometric objects defined by polynomial equations, using algebraic means. Its roots go back to Descartes' introduction of coordinates to describe points in Euclidean space and his idea of describing curves and surfaces by algebraic equations. Over the long history of the subject, both powerful general theories and detailed knowledge of many specific examples have been developed. Recently, with the development of computer algebra systems and the discovery (or rediscovery) of algorithmic approaches to many of the basic computations, the techniques of algebraic geometry have also found significant applications, for example in geometric design, combinatorics, integer programming, coding theory, and robotics. Our goal in *Using Algebraic Geometry* is to survey these algorithmic approaches and many of their applications.

For the convenience of the reader, in this introductory chapter we will first recall the basic algebraic structure of *ideals* in polynomial rings. In §2 and §3 we will present a rapid summary of the *Gröbner basis algorithms* developed by Buchberger for computations in polynomial rings, with several worked out examples. Finally, in §4 we will recall the geometric notion of an *affine algebraic variety*, the simplest type of geometric object defined by polynomial equations. The topics in §1, §2, and §3 are the common prerequisites for all of the following chapters. §4 gives the geometric context for the algebra from the earlier sections. We will make use of this language at many points. If these topics are familiar, you may wish to proceed directly to the later material and refer back to this introduction as needed.

§1 Polynomials and Ideals

To begin, we will recall some terminology. A *monomial* in a collection of variables $x_1, \ldots, x_n$ is a product

$$x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \tag{1.1}$$

where the α_i are non-negative integers. To abbreviate, we will sometimes rewrite (1.1) as x^α where $\alpha = (\alpha_1, \ldots, \alpha_n)$ is the vector of exponents in the monomial. The *total degree* of a monomial x^α is the sum of the exponents: $\alpha_1 + \cdots + \alpha_n$. We will often denote the total degree of the monomial x^α by $|\alpha|$. For instance $x_1^3 x_2^2 x_4$ is a monomial of total degree 6 in the variables x_1, x_2, x_3, x_4, since $\alpha = (3, 2, 0, 1)$ and $|\alpha| = 6$.

If k is any field, we can form finite linear combinations of monomials with coefficients in k. The resulting objects are known as *polynomials* in $x_1, \ldots, x_n$. We will also use the word *term* on occasion to refer to a product of a nonzero element of k and a monomial appearing in a polynomial. Thus, a general polynomial in the variables $x_1, \ldots, x_n$ with coefficients in k has the form

$$f = \sum_\alpha c_\alpha x^\alpha,$$

where $c_\alpha \in k$ for each α, and there are only finitely many terms $c_\alpha x^\alpha$ in the sum. For example, taking k to be the field $\mathbb{Q}$ of rational numbers, and denoting the variables by x, y, z rather than using subscripts,

$$p = x^2 + \tfrac{1}{2} y^2 z - z - 1 \tag{1.2}$$

is a polynomial containing four terms.

In most of our examples, the field of coefficients will be either $\mathbb{Q}$, the field of real numbers, $\mathbb{R}$, or the field of complex numbers, $\mathbb{C}$. Polynomials over finite fields will also be introduced in Chapter 9. We will denote by $k[x_1, \ldots, x_n]$ the collection of all polynomials in $x_1, \ldots, x_n$ with coefficients in k. Polynomials in $k[x_1, \ldots, x_n]$ can be added and multiplied as usual, so $k[x_1, \ldots, x_n]$ has the structure of a *commutative ring* (with identity). However, only nonzero constant polynomials have multiplicative inverses in $k[x_1, \ldots, x_n]$, so $k[x_1, \ldots, x_n]$ is not a field. However, the set of *rational functions* $\{f/g : f, g \in k[x_1, \ldots, x_n], g \neq 0\}$ is a field, denoted $k(x_1, \ldots, x_n)$.

A polynomial f is said to be *homogeneous* if all the monomials appearing in it with nonzero coefficients have *the same* total degree. For instance, $f = 4x^3 + 5xy^2 - z^3$ is a homogeneous polynomial of total degree 3 in $\mathbb{Q}[x, y, z]$, while $g = 4x^3 + 5xy^2 - z^6$ is not homogeneous. When we study resultants in Chapter 3, homogeneous polynomials will play an important role.

Given a collection of polynomials, $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$, we can consider all polynomials which can be built up from these by multiplication by arbitrary polynomials and by taking sums.

(1.3) Definition. Let $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$. We let $\langle f_1, \ldots, f_s \rangle$ denote the collection

$$\langle f_1, \ldots, f_s \rangle = \{p_1 f_1 + \cdots + p_s f_s : p_i \in k[x_1, \ldots, x_n] \text{ for } i = 1, \ldots, s\}.$$

For example, consider the polynomial p from (1.2) above and the two polynomials

$$f_1 = x^2 + z^2 - 1$$
$$f_2 = x^2 + y^2 + (z-1)^2 - 4.$$

We have

$$\begin{aligned}(1.4)\qquad p &= x^2 + \tfrac{1}{2}y^2z - z - 1\\ &= (-\tfrac{1}{2}z + 1)(x^2 + z^2 - 1) + (\tfrac{1}{2}z)(x^2 + y^2 + (z-1)^2 - 4).\end{aligned}$$

This shows $p \in \langle f_1, f_2\rangle$.

Exercise 1.
a. Show that $x^2 \in \langle x - y^2, xy\rangle$ in $k[x,y]$ (k any field).
b. Show that $\langle x - y^2, xy, y^2\rangle = \langle x, y^2\rangle$.
c. Is $\langle x - y^2, xy\rangle = \langle x^2, xy\rangle$? Why or why not?

Exercise 2. Show that $\langle f_1, \dots, f_s\rangle$ is closed under sums in $k[x_1, \dots, x_n]$. Also show that if $f \in \langle f_1, \dots, f_s\rangle$, and $p \in k[x_1, \dots, x_n]$ is an arbitrary polynomial, then $p \cdot f \in \langle f_1, \dots, f_s\rangle$.

The two properties in Exercise 2 are the defining properties of *ideals* in the ring $k[x_1, \dots, x_n]$.

(1.5) Definition. Let $I \subset k[x_1, \dots, x_n]$ be a non-empty subset. I is said to be an *ideal* if
a. $f + g \in I$ whenever $f \in I$ and $g \in I$, and
b. $pf \in I$ whenever $f \in I$, and $p \in k[x_1, \dots, x_n]$ is an arbitrary polynomial.

Thus $\langle f_1, \dots, f_s\rangle$ is an ideal by Exercise 2. We will call it the *ideal generated by* $f_1, \dots, f_s$ because it has the following property.

Exercise 3. Show that $\langle f_1, \dots, f_s\rangle$ is the *smallest* ideal in $k[x_1, \dots, x_n]$ containing $f_1, \dots, f_s$, in the sense that if J is any ideal containing $f_1, \dots, f_s$, then $\langle f_1, \dots, f_s\rangle \subset J$.

Exercise 4. Using Exercise 3, formulate and prove a general criterion for equality of ideals $I = \langle f_1, \dots, f_s\rangle$ and $J = \langle g_1, \dots, g_t\rangle$ in $k[x_1, \dots, x_n]$. How does your statement relate to what you did in part b of Exercise 1?

Given an ideal, or several ideals, in $k[x_1, \dots, x_n]$, there are a number of algebraic constructions that yield other ideals. One of the most important of these for geometry is the following.

(1.6) Definition. Let $I \subset k[x_1, \ldots, x_n]$ be an ideal. The *radical of* I is the set

$$\sqrt{I} = \{g \in k[x_1, \ldots, x_n] : g^m \in I \text{ for some } m \geq 1\}.$$

An ideal I is said to be a *radical ideal* if $\sqrt{I} = I$.

For instance,

$$x + y \in \sqrt{\langle x^2 + 3xy, 3xy + y^2\rangle}$$

in $\mathbb{Q}[x, y]$ since

$$(x + y)^3 = x(x^2 + 3xy) + y(3xy + y^2) \in \langle x^2 + 3xy, 3xy + y^2\rangle.$$

Since each of the generators of the ideal $\langle x^2 + 3xy, 3xy + y^2\rangle$ is homogeneous of degree 2, it is clear that $x + y \notin \langle x^2 + 3xy, 3xy + y^2\rangle$. It follows that $\langle x^2 + 3xy, 3xy + y^2\rangle$ is *not* a radical ideal.

Although it is not obvious from the definition, we have the following property of the radical.

- (Radical Ideal Property) For every ideal $I \subset k[x_1, \ldots, x_n]$, $\sqrt{I}$ is an ideal containing I.

See [CLO], Chapter 4, §2, for example. We will consider a number of other operations on ideals in the exercises.

One of the most important general facts about ideals in $k[x_1, \ldots, x_n]$ is known as the Hilbert Basis Theorem. In this context, a *basis* is another name for a generating set for an ideal.

- (Hilbert Basis Theorem) Every ideal I in $k[x_1, \ldots, x_n]$ has a *finite* generating set. In other words, given an ideal I, there exists a finite collection of polynomials $\{f_1, \ldots, f_s\} \subset k[x_1, \ldots, x_n]$ such that $I = \langle f_1, \ldots, f_s\rangle$.

For polynomials in one variable, this is a standard consequence of the one-variable polynomial division algorithm.

- (Division Algorithm in $k[x]$) Given two polynomials $f, g \in k[x]$, we can divide f by g, producing a unique quotient q and remainder r such that

  $$f = qg + r,$$

  and either $r = 0$, or r has degree strictly smaller than the degree of g.

See, for instance, [CLO], Chapter 1, §5. The consequences of this result for ideals in $k[x]$ are discussed in Exercise 6 below. For polynomials in several variables, the Hilbert Basis Theorem can be proved either as a byproduct of the theory of Gröbner bases to be reviewed in the next section (see [CLO], Chapter 2, §5), or inductively by showing that if every ideal in a ring R is finitely generated, then the same is true in the ring $R[x]$ (see [AL], Chapter 1, §1, or [BW], Chapter 4, §1).

Additional Exercises for §1

Exercise 5. Show that $\langle y - x^2, z - x^3 \rangle = \langle z - xy, y - x^2 \rangle$ in $\mathbb{Q}[x, y, z]$.

Exercise 6. Let k be any field, and consider the polynomial ring in one variable, $k[x]$. In this exercise, you will give one proof that every ideal in $k[x]$ is finitely generated. In fact, every ideal $I \subset k[x]$ is generated by a single polynomial: $I = \langle g \rangle$ for some g. We may assume $I \neq \{0\}$ for there is nothing to prove in that case. Let g be a nonzero element in I of minimal degree. Show using the division algorithm that every f in I is divisible by g. Deduce that $I = \langle g \rangle$.

Exercise 7.
a. Let k be any field, and let n be any positive integer. Show that in $k[x]$, $\sqrt{\langle x^n \rangle} = \langle x \rangle$.
b. More generally, suppose that

$$p(x) = (x - a_1)^{e_1} \cdots (x - a_m)^{e_m}.$$

What is $\sqrt{\langle p(x) \rangle}$?
c. Let $k = \mathbb{C}$, so that *every* polynomial in one variable factors as in b. What are the radical ideals in $\mathbb{C}[x]$?

Exercise 8. An ideal $I \subset k[x_1, \ldots, x_n]$ is said to be *prime* if whenever a product fg belongs to I, either $f \in I$, or $g \in I$ (or both).
a. Show that a prime ideal is radical.
b. What are the prime ideals in $\mathbb{C}[x]$? What about the prime ideals in $\mathbb{R}[x]$ or $\mathbb{Q}[x]$?

Exercise 9. An ideal $I \subset k[x_1, \ldots, x_n]$ is said to be *maximal* if there are no ideals J satisfying $I \subset J \subset k[x_1, \ldots, x_n]$ other than $J = I$ and $J = k[x_1, \ldots, x_n]$.
a. Show that $\langle x_1, x_2, \ldots, x_n \rangle$ is a maximal ideal in $k[x_1, \ldots, x_n]$.
b. More generally show that if $(a_1, \ldots, a_n)$ is any point in k^n, then the ideal $\langle x_1 - a_1, \ldots, x_n - a_n \rangle \subset k[x_1, \ldots, x_n]$ is maximal.
c. Show that $I = \langle x^2 + 1 \rangle$ is a maximal ideal in $\mathbb{R}[x]$. Is I maximal considered as an ideal in $\mathbb{C}[x]$?

Exercise 10. Let I be an ideal in $k[x_1, \ldots, x_n]$, let $\ell \geq 1$ be an integer, and let I_ℓ consist of the elements in I that do not depend on the first ℓ variables:

$$I_\ell = I \cap k[x_{\ell+1}, \ldots, x_n].$$

I_ℓ is called the ℓth *elimination ideal* of I.
a. For $I = \langle x^2 + y^2, x^2 - z^3 \rangle \subset k[x, y, z]$, show that $y^2 + z^3$ is in the first elimination ideal I_1.

b. Prove that I_ℓ is an ideal in the ring $k[x_{\ell+1}, \ldots, x_n]$.

Exercise 11. Let I, J be ideals in $k[x_1, \ldots, x_n]$, and define

$$I + J = \{f + g : f \in I, g \in J\}.$$

a. Show that $I + J$ is an ideal in $k[x_1, \ldots, x_n]$.
b. Show that $I + J$ is the smallest ideal containing $I \cup J$.
c. If $I = \langle f_1, \ldots, f_s \rangle$ and $J = \langle g_1, \ldots, g_t \rangle$, what is a finite generating set for $I + J$?

Exercise 12. Let I, J be ideals in $k[x_1, \ldots, x_n]$.
a. Show that $I \cap J$ is also an ideal in $k[x_1, \ldots, x_n]$.
b. Define IJ to be the smallest ideal containing all the products fg where $f \in I$, and $g \in J$. Show that $IJ \subset I \cap J$. Give an example where $IJ \neq I \cap J$.

Exercise 13. Let I, J be ideals in $k[x_1, \ldots, x_n]$, and define $I\colon J$ (called the *quotient ideal* of I by J) by

$$I\colon J = \{f \in k[x_1, \ldots, x_n] : fg \in I \text{ for all } g \in J\}.$$

a. Show that $I\colon J$ is an ideal in $k[x_1, \ldots, x_n]$.
b. Show that if $I \cap \langle h \rangle = \langle g_1, \ldots, g_t \rangle$ (so each g_i is divisible by h), then a basis for $I\colon \langle h \rangle$ is obtained by cancelling the factor of h from each g_i:

$$I\colon \langle h \rangle = \langle g_1/h, \ldots, g_t/h \rangle.$$

§2 Monomial Orders and Polynomial Division

The examples of ideals that we considered in §1 were artificially simple. In general, it can be difficult to determine by inspection or by trial and error whether a given polynomial $f \in k[x_1, \ldots, x_n]$ is an element of a given ideal $I = \langle f_1, \ldots, f_s \rangle$, or whether two ideals $I = \langle f_1, \ldots, f_s \rangle$ and $J = \langle g_1, \ldots, g_t \rangle$ are equal. In this section and the next one, we will consider a collection of algorithms that can be used to solve problems such as deciding ideal membership, deciding ideal equality, computing ideal intersections and quotients, and computing elimination ideals. See the exercises at the end of §3 for some examples.

The starting point for these algorithms is, in a sense, the polynomial division algorithm in $k[x]$ introduced at the end of §1. In Exercise 6 of §1, we saw that the division algorithm implies that every ideal $I \subset k[x]$ has the form $I = \langle g \rangle$ for some g. Hence, if $f \in k[x]$, we can also use division to determine whether $f \in I$.

Exercise 1. Let $I = \langle g \rangle$ in $k[x]$ and let $f \in k[x]$ be any polynomial. Let q, r be the unique quotient and remainder in the expression $f = qg + r$ produced by polynomial division. Show that $f \in I$ if and only if $r = 0$.

Exercise 2. Formulate and prove a criterion for equality of ideals $I_1 = \langle g_1 \rangle$ and $I_2 = \langle g_2 \rangle$ in $k[x]$ based on division.

Given the usefulness of division for polynomials in one variable, we may ask: Is there a corresponding notion for polynomials in several variables? The answer is *yes*, and to describe it, we need to begin by considering different ways to *order* the monomials appearing within a polynomial.

(2.1) Definition. A *monomial order* on $k[x_1, \ldots, x_n]$ is any relation $>$ on the set of monomials x^α in $k[x_1, \ldots, x_n]$ (or equivalently on the exponent vectors $\alpha \in \mathbb{Z}^n_{\geq 0}$) satisfying:

a. $>$ is a *total (linear) ordering* relation;
b. $>$ is *compatible with multiplication* in $k[x_1, \ldots, x_n]$, in the sense that if $x^\alpha > x^\beta$ and x^γ is any monomial, then $x^\alpha x^\gamma = x^{\alpha+\gamma} > x^{\beta+\gamma} = x^\beta x^\gamma$;
c. $>$ is a *well-ordering*. That is, every nonempty collection of monomials has a smallest element under $>$.

Condition a implies that the terms appearing within any polynomial f can be uniquely listed in increasing or decreasing order under $>$. Then condition b shows that that ordering does not change if we multiply f by a monomial x^γ. Finally, condition c is used to ensure that processes that work on collections of monomials, e.g., the collection of all monomials less than some fixed monomial x^α, will terminate in a finite number of steps.

The division algorithm in $k[x]$ makes use of a monomial order *implicitly*: when we divide g into f by hand, we always compare the leading term (the term of highest degree) in g with the leading term of the intermediate dividend. In fact there is no choice in the matter in this case.

Exercise 3. Show that the *only* monomial order on $k[x]$ is the *degree order* on monomials, given by

$$\cdots > x^{n+1} > x^n > \cdots > x^3 > x^2 > x > 1.$$

For polynomial rings in several variables, there are many choices of monomial orders. In writing the exponent vectors α and β in monomials x^α and x^β as ordered n-tuples, we implicitly set up an ordering on the variables x_i in $k[x_1, \ldots, x_n]$:

$$x_1 > x_2 > \cdots > x_n.$$

With this choice, there are still many ways to define monomial orders. Some of the most important are given in the following definitions.

(2.2) Definition (Lexicographic Order). Let x^α and x^β be monomials in $k[x_1, \ldots, x_n]$. We say $x^\alpha >_{lex} x^\beta$ if in the difference $\alpha - \beta \in \mathbb{Z}^n$, the leftmost nonzero entry is positive.

Lexicographic order is analogous to the ordering of words used in dictionaries.

(2.3) Definition (Graded Lexicographic Order). Let x^α and x^β be monomials in $k[x_1, \ldots, x_n]$. We say $x^\alpha >_{grlex} x^\beta$ if $\sum_{i=1}^n \alpha_i > \sum_{i=1}^n \beta_i$, or if $\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i$, and $x^\alpha >_{lex} x^\beta$.

(2.4) Definition (Graded Reverse Lexicographic Order). Let x^α and x^β be monomials in $k[x_1, \ldots, x_n]$. We say $x^\alpha >_{grevlex} x^\beta$ if $\sum_{i=1}^n \alpha_i > \sum_{i=1}^n \beta_i$, or if $\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i$, and in the difference $\alpha - \beta \in \mathbb{Z}^n$, the rightmost nonzero entry is negative.

For instance, in $k[x, y, z]$, with $x > y > z$, we have

$$x^3y^2z >_{lex} x^2y^6z^{12} \tag{2.5}$$

since when we compute the difference of the exponent vectors:

$$(3, 2, 1) - (2, 6, 12) = (1, -4, -11),$$

the leftmost nonzero entry is positive. Similarly,

$$x^3y^6 >_{lex} x^3y^4z$$

since in $(3, 6, 0) - (3, 4, 1) = (0, 2, -1)$, the leftmost nonzero entry is positive. Comparing the *lex* and *grevlex* orders shows that the results can be quite different. For instance, it is true that

$$x^2y^6z^{12} >_{grevlex} x^3y^2z.$$

Compare this with (2.5), which contains the same monomials. Indeed, *lex* and *grevlex* are *different* orderings even on the monomials of the same total degree in three or more variables, as we can see by considering pairs of monomials such as $x^2y^2z^2$ and xy^4z. Since $(2, 2, 2) - (1, 4, 1) = (1, -2, 1)$,

$$x^2y^2z^2 >_{lex} xy^4z.$$

On the other hand by Definition (2.4),

$$xy^4z >_{grevlex} x^2y^2z^2.$$

Exercise 4. Show that $>_{lex}$, $>_{grlex}$, and $>_{grevlex}$ are monomial orders in $k[x_1, \ldots, x_n]$ according to Definition (2.1).

Exercise 5. Show that the monomials of a *fixed* total degree d in *two* variables $x > y$ are ordered in the same sequence by $>_{lex}$ and $>_{grevlex}$. Are these orderings the same on all of $k[x, y]$ though? Why or why not?

For future reference, we next discuss a general method for specifying monomial orders on $k[x_1, \ldots, x_n]$. We start from any $m \times n$ real matrix M and write the rows of M as $\mathbf{w}_1, \ldots, \mathbf{w}_m$. Then we can compare monomials x^α and x^β by first comparing their $\mathbf{w}_1$-weights $\alpha \cdot \mathbf{w}_1$ and $\alpha \cdot \mathbf{w}_1$. If $\alpha \cdot \mathbf{w}_1 > \beta \cdot \mathbf{w}_1$ or $\beta \cdot \mathbf{w}_1 > \alpha \cdot \mathbf{w}_1$, then we order the monomials accordingly. If $\alpha \cdot \mathbf{w}_1 = \beta \cdot \mathbf{w}_1$, then we continue to the later rows, breaking ties successively with the $\mathbf{w}_2$-weights, the $\mathbf{w}_3$-weights, and so on through the $\mathbf{w}_m$-weights. This process defines an order relation $>_M$. In symbols: $x^\alpha >_M x^\beta$ if there is an $\ell \leq m$ such that $\alpha \cdot \mathbf{w}_i = \beta \cdot \mathbf{w}_i$ for $i = 1, \ldots, \ell - 1$, but $\alpha \cdot \mathbf{w}_\ell > \beta \cdot \mathbf{w}_\ell$.

To obtain a total order by this construction, it must be true that $\ker(M) \cap \mathbb{Z}^n = \{0\}$. If the entries of M are rational numbers, then this property implies that $m \geq n$, and M has full rank n. The same construction also works for M with irrational entries, but there is a small subtlety concerning what notion of rank is appropriate in that case. See Exercise 9 below. To guarantee the well-ordering property of monomial orders, it is sufficient (although not necessary) to require that M have all entries nonnegative.

Exercise 6. All the monomial orders we have seen can be specified as $>_M$ orders for appropriate matrices M.

a. Show that the *lex* order with $x > y > z$ is defined by the identity matrix

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and similarly in $k[x_1, \ldots, x_n]$ for all $n \geq 1$.

b. Show that the *grevlex* order with $x > y > z$ is defined by either the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

or the matrix

$$M' = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

and similarly in $k[x_1, \ldots, x_n]$ for all $n \geq 1$. This example shows that matrices with negative entries can also define monomial orders.

c. The *grlex* order compares monomials first by total degree (weight vector $\mathbf{w}_1 = (1, 1, 1)$), then breaks ties by the *lex order*. This, together with part a, shows $>_{grlex} = >_M$ for the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Show that we could also use

$$M' = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

That is, show that the last row in M is actually superfluous. (Hint: Making comparisons, when would we ever need to use the last row?)

d. One very common way to define a monomial order is to compare weights with respect to one vector first, then break ties with another standard order such as *grevlex*. We denote such an order by $>_{\mathbf{w},grevlex}$. These weight orders are studied, for instance, in [CLO], Chapter 2, §4, Exercise 12. Suppose $\mathbf{w} = (2, 4, 7)$ and ties are broken by *grevlex* with $x > y > z$. To define this order, it is most natural to use

$$M = \begin{pmatrix} 2 & 4 & 7 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

However, some computer algebra systems (e.g., Maple V, Release 5 and later versions with the `Groebner` package) require square weight matrices. Consider the two matrices obtained from M by deleting a row:

$$M' = \begin{pmatrix} 2 & 4 & 7 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad M'' = \begin{pmatrix} 2 & 4 & 7 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Both have rank 3 so the condition $\ker(M) \cap \mathbb{Z}^3 = \{0\}$ is satisfied. Which matrix defines the $>_{\mathbf{w},grevlex}$ order?

e. Let $m > n$. Given an $m \times n$ matrix M defining a monomial order $>_M$, describe a general method for picking an $n \times n$ submatrix M' of M to define the same order.

In Exercise 8 below, you will prove that $>_M$ defines a monomial order for any suitable matrix M. In fact, by a result of Robbiano (see [Rob]), the $>_M$ construction gives all monomial orders on $k[x_1, \ldots, x_n]$.

We will use monomial orders in the following way. The natural generalization of the leading term (term of highest degree) in a polynomial in $k[x]$ is defined as follows. Picking any particular monomial order $>$ on $k[x_1, \ldots, x_n]$, we consider the terms in $f = \sum_\alpha c_\alpha x^\alpha$. Then the *leading* *term* of f (with respect to $>$) is the product $c_\alpha x^\alpha$ where x^α is the *largest* monomial appearing in f in the ordering $>$. We will use the notation $\text{LT}_>(f)$ for the leading term, or just $\text{LT}(f)$ if there is no chance of confusion about which monomial order is being used. Furthermore, if $\text{LT}(f) = cx^\alpha$, then $\text{LC}(f) = c$ is the *leading coefficient* of f and $\text{LM}(f) = x^\alpha$ is the *leading monomial*. Note that $\text{LT}(0)$, $\text{LC}(0)$, and $\text{LM}(0)$ are undefined.

For example, consider $f = 3x^3y^2 + x^2yz^3$ in $\mathbb{Q}[x, y, z]$ (with variables ordered $x > y > z$ as usual). We have

$$\text{LT}_{>_{lex}}(f) = 3x^3y^2$$

since $x^3y^2 >_{lex} x^2yz^3$. On the other hand

$$\text{LT}_{>_{grevlex}}(f) = x^2yz^3$$

since the total degree of the second term is 6 and the total degree of the first is 5.

Monomial orders are used in a generalized division algorithm.

- (Division Algorithm in $k[x_1, \ldots, x_n]$) Fix any monomial order $>$ in $k[x_1, \ldots, x_n]$, and let $F = (f_1, \ldots, f_s)$ be an ordered s-tuple of polynomials in $k[x_1, \ldots, x_n]$. Then every $f \in k[x_1, \ldots, x_n]$ can be written as

  $$f = a_1 f_1 + \cdots + a_s f_s + r, \tag{2.6}$$

  where $a_i, r \in k[x_1, \ldots, x_n]$, for each i, $a_i f_i = 0$ or $\text{LT}_>(f) \geq \text{LT}_>(a_i f_i)$, and either $r = 0$, or r is a linear combination of monomials, none of which is divisible by any of $\text{LT}_>(f_1), \ldots, \text{LT}_>(f_s)$. We will call r a *remainder* of f on division by F.

In the particular algorithmic form of the division process given in [CLO], Chapter 2, §3, and [AL], Chapter 1, §5, the intermediate dividend is reduced at each step using the divisor f_i with the *smallest possible* i such that $\text{LT}(f_i)$ divides the leading term of the intermediate dividend. A characterization of the expression (2.6) that is produced by this version of division can be found in Exercise 11 of Chapter 2, §3 of [CLO]. More general forms of division or polynomial reduction procedures are considered in [AL] and [BW], Chapter 5, §1.

You should note two differences between this statement and the division algorithm in $k[x]$. First, we are allowing the possibility of dividing f by an s-tuple of polynomials with $s > 1$. The reason for this is that we will usually want to think of the divisors f_i as generators for some particular ideal I, and ideals in $k[x_1, \ldots, x_n]$ for $n \geq 2$ might not be generated by any single polynomial. Second, although any algorithmic version of division, such as the one presented in Chapter 2 of [CLO], produces one particular expression of the form (2.6) for each ordered s-tuple F and each f, there are always *different* expressions of this form for a given f as well. Reordering

F or changing the monomial order can produce different a_i and r in some cases. See Exercise 7 below for some examples.

We will sometimes use the notation

$$r = \overline{f}^F$$

for a remainder on division by F.

Most computer algebra systems that have Gröbner basis packages provide implementations of some form of the division algorithm. However, in most cases the output of the division command is just the remainder $\overline{f}^F$, the quotients a_i are not saved or displayed, and an algorithm different from the one described in [CLO], Chapter 2, §3 may be used. For instance, the Maple `Groebner` package contains a function `normalf` which computes a remainder on division of a polynomial by any collection of polynomials. To use it, one must start by loading the `Groebner` package (just once in a session) with

```
with(Groebner);
```

The format for the `normalf` command is

```
normalf(f, F, torder);
```

where `f` is the dividend polynomial, `F` is the ordered list of divisors (in square brackets, separated by commas), and `torder` specifies the monomial order. For instance, to use the $>_{lex}$ order, enter `plex`, then in parentheses, separated by commas, list the variables in descending order. Similarly, to use the $>_{grevlex}$ order, enter `tdeg`, then in parentheses, separated by commas, list the variables in descending order. Let us consider dividing $f_1 = x^2y^2 - x$ and $f_2 = xy^3 + y$ into $f = x^3y^2 + 2xy^4$ using the *lex* order on $\mathbb{Q}[x, y]$ with $x > y$. The Maple commands

(2.7)
```
f := x^3*y^2 + 2*x*y^4;
F := [x^2*y^2 - x, x*y^3 + y];
normalf(f,F,plex(x,y));
```

will produce as output

$$x^2 - 2y^2. \tag{2.8}$$

Thus the remainder is $\overline{f}^F = x^2 - 2y^2$. The `normalf` procedure uses the algorithmic form of division presented, for instance, in [CLO], Chapter 2, §3.

The `Groebner` package contains several additional ways to specify monomial orders, including one to construct $>_M$ for a square matrix M with positive integer entries. Hence it can be used to work with general monomial orders on $k[x_1, \ldots, x_n]$. We will present a number of examples in later chapters.

Additional Exercises for §2

Exercise 7.

a. Verify by hand that the remainder from (2.8) occurs in an expression
$$f = a_1 f_1 + a_2 f_2 + x^2 - 2y^2,$$
where $a_1 = x$, $a_2 = 2y$, and f_i are as in the discussion before (2.7).

b. Show that reordering the variables and changing the monomial order to `tdeg(x,y)` has no effect in (2.8).

c. What happens if you change `F` in (2.7) to
$$F = [x^2y^2 - x^4, xy^3 - y^4]$$
and take $f = x^2y^6$? Does changing the order of the variables make a difference now?

d. Now change `F` to
$$F = [x^2y^2 - z^4, xy^3 - y^4],$$
take $f = x^2y^6 + z^5$, and change the monomial order to `plex(x,y,z)`. Also try *lex* orders with the variables permuted and other monomial orders.

Exercise 8. Let M be an $m \times n$ real matrix with nonnegative entries. Assume that $\ker(M) \cap \mathbb{Z}^n = \{0\}$. Show that $>_M$ is a monomial order on $k[x_1, \ldots, x_n]$.

Exercise 9. Given $\mathbf{w} \in (\mathbb{R}^n)^+$ define $x^\alpha >_{\mathbf{w}} x^\beta$ if $\alpha \cdot \mathbf{w} > \beta \cdot \mathbf{w}$.

a. Give an example to show that $>_{\mathbf{w}}$ is not necessarily a monomial order on $k[x_1, \ldots, x_n]$.

b. With $n = 2$, let $\mathbf{w} = (1, \sqrt{2})$. Show that $>_{\mathbf{w}}$ is a monomial order on $k[x_1, x_2]$ in this case.

c. What property of the components of the vector $\mathbf{w} \in (\mathbb{R}^n)^+$ guarantees that $>_{\mathbf{w}}$ *does* define a monomial order on $k[x_1, \ldots, x_n]$? Prove your assertion. (Hint: See Exercise 11 of Chapter 2, §4 of [CLO].)

§3 Gröbner Bases

Since we now have a division algorithm in $k[x_1, \ldots, x_n]$ that seems to have many of the same features as the one-variable version, it is natural to ask if deciding whether a given $f \in k[x_1, \ldots, x_n]$ is a member of a given ideal $I = \langle f_1, \ldots, f_s \rangle$ can be done along the lines of Exercise 1 in §2, by computing the remainder on division. One direction is easy. Namely, from (2.6) it follows that if $r = \overline{f}^F = 0$ on dividing by $F = (f_1, \ldots, f_s)$, then $f = a_1 f_1 + \cdots + a_s f_s$. By definition then, $f \in \langle f_1, \ldots, f_s \rangle$. On the other hand, the following exercise shows that we are not guaranteed to get $\overline{f}^F = 0$ for every $f \in \langle f_1, \ldots, f_s \rangle$ if we use an arbitrary basis F for I.

Exercise 1. Recall from (1.4) that $p = x^2 + \frac{1}{2}y^2z - z - 1$ *is* an element of the ideal $I = \langle x^2 + z^2 - 1, x^2 + y^2 + (z-1)^2 - 4 \rangle$. Show, however, that the remainder on division of p by this generating set F is not zero. For instance, using $>_{lex}$, we get a remainder

$$\overline{p}^F = \tfrac{1}{2}y^2z - z - z^2.$$

What went wrong here? From (2.6) and the fact that $f \in I$ in this case, it follows that the remainder is *also an element of* I. However, $\overline{p}^F$ is not zero because it contains terms that cannot be removed by division by these particular generators for I. The leading terms of $f_1 = x^2 + z^2 - 1$ and $f_2 = x^2 + y^2 + (z-1)^2 - 4$ do not divide the leading term of $\overline{p}^F$. In order for division to produce zero remainders for all elements of I, we need to be able to remove *all* leading terms of elements of I using the leading terms of the divisors. That is the motivation for the following definition.

(3.1) Definition. Fix a monomial order $>$ on $k[x_1, \ldots, x_n]$, and let $I \subset k[x_1, \ldots, x_n]$ be an ideal. A *Gröbner basis* for I (with respect to $>$) is a finite collection of polynomials $G = \{g_1, \ldots, g_t\} \subset I$ with the property that for every nonzero $f \in I$, $\text{LT}(f)$ is divisible by $\text{LT}(g_i)$ for some i.

We will see in a moment (Exercise 3) that a Gröbner basis for I is indeed a basis for I, i.e., $I = \langle g_1, \ldots, g_t \rangle$. Of course, it must be proved that Gröbner bases *exist* for all I in $k[x_1, \ldots, x_n]$. This can be done in a non-constructive way by considering the ideal $\langle \text{LT}(I) \rangle$ generated by the leading terms of all the elements in I (a *monomial ideal*). By a direct argument (Dickson's Lemma: see [CLO], Chapter 2, §4, or [BW], Chapter 4, §3, or [AL], Chapter 1, §4), or by the Hilbert Basis Theorem, the ideal $\langle \text{LT}(I) \rangle$ has a finite generating set consisting of monomials $x^{\alpha(i)}$ for $i = 1, \ldots, t$. By the definition of $\langle \text{LT}(I) \rangle$, there is an element $g_i \in I$ such that $\text{LT}(g_i) = x^{\alpha(i)}$ for each $i = 1, \ldots, t$.

Exercise 2. Show that if $\langle \text{LT}(I) \rangle = \langle x^{\alpha(1)}, \ldots, x^{\alpha(t)} \rangle$, and if $g_i \in I$ are polynomials such that $\text{LT}(g_i) = x^{\alpha(i)}$ for each $i = 1, \ldots, t$, then $G = \{g_1, \ldots, g_t\}$ is a Gröbner basis for I.

Remainders computed by division with respect to a Gröbner basis are much better behaved than those computed with respect to arbitrary sets of divisors. For instance, we have the following results.

Exercise 3.

a. Show that if G is a Gröbner basis for I, then for any $f \in I$, the remainder on division of f by G (listed in any order) is zero.

b. Deduce that $I = \langle g_1, \ldots, g_t \rangle$ if $G = \{g_1, \ldots, g_t\}$ is a Gröbner basis for I. (If $I = \langle 0 \rangle$, then $G = \emptyset$ and we make the convention that $\langle \emptyset \rangle = \{0\}$.)

Exercise 4. If G is a Gröbner basis for an ideal I, and f is an arbitrary polynomial, show that if the algorithm of [CLO], Chapter 2, §3 is used, the remainder on division of f by G is independent of the ordering of G. Hint: If two different orderings of G are used, producing remainders r_1 and r_2, consider the difference $r_1 - r_2$.

Generalizing the result of Exercise 4, we also have the following important statement.

- (Uniqueness of Remainders) Fix a monomial order $>$ and let $I \subset k[x_1, \ldots, x_n]$ be an ideal. Division of $f \in k[x_1, \ldots, x_n]$ by a Gröbner basis for I produces an expression $f = g + r$ where $g \in I$ and no term in r is divisible by any element of $\textsc{lt}(I)$. If $f = g' + r'$ is any other such expression, then $r = r'$.

See [CLO], Chapter 2, §6, [AL], Chapter 1, §6, or [BW], Chapter 5, §2. In other words, the remainder on division of f by a Gröbner basis for I is a uniquely determined *normal form* for f modulo I depending only on the choice of monomial order and not on the way the division is performed. Indeed, uniqueness of remainders gives another characterization of Gröbner bases.

More useful for many purposes than the existence proof for Gröbner bases above is an *algorithm*, due to Buchberger, that takes an arbitrary generating set $\{f_1, \ldots, f_s\}$ for I and produces a Gröbner basis G for I from it. This algorithm works by forming new elements of I using expressions guaranteed to cancel leading terms and uncover other possible leading terms, according to the following recipe.

(3.2) Definition. Let $f, g \in k[x_1, \ldots, x_n]$ be nonzero. Fix a monomial order and let

$$\textsc{lt}(f) = cx^\alpha \qquad \text{and} \qquad \textsc{lt}(g) = dx^\beta,$$

where $c, d \in k$. Let x^γ be the least common multiple of x^α and x^β. The *S-polynomial* of f and g, denoted $S(f,g)$, is the polynomial

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

Note that by definition $S(f,g) \in \langle f, g \rangle$. For example, with $f = x^3y - 2x^2y^2 + x$ and $g = 3x^4 - y$ in $\mathbb{Q}[x,y]$, and using $>_{lex}$, we have $x^\gamma = x^4y$, and

$$S(f,g) = xf - (y/3)g = -2x^3y^2 + x^2 + y^2/3.$$

In this case, the leading term of the S-polynomial is divisible by the leading term of f. We might consider taking the remainder on division by $F = (f, g)$ to uncover possible new leading terms of elements in $\langle f, g\rangle$. And indeed in this case we find that the remainder is

$$\overline{S(f, g)}^F = -4x^2y^3 + x^2 + 2xy + y^2/3 \tag{3.3}$$

and $\mathrm{LT}(\overline{S(f, g)}^F) = -4x^2y^3$ is divisible by neither $\mathrm{LT}(f)$ nor $\mathrm{LT}(g)$. An important result about this process of forming S-polynomial remainders is the following statement.

- (Buchberger's Criterion) A finite set $G = \{g_1, \ldots, g_t\}$ is a Gröbner basis of $I = \langle g_1, \ldots, g_t\rangle$ if and only if $\overline{S(g_i, g_j)}^G = 0$ for all pairs $i \neq j$.

See [CLO], Chapter 2, §7, [BW], Chapter 5, §3, or [AL], Chapter 1, §7. Using this criterion above, we obtain a very rudimentary procedure for producing a Gröbner basis of a given ideal.

- (Buchberger's Algorithm)

  Input: $F = (f_1, \ldots, f_s)$

  Output: a Gröbner basis $G = \{g_1, \ldots, g_t\}$ for $I = \langle F\rangle$, with $F \subset G$

  $G := F$

  REPEAT

  $G' := G$

  FOR each pair $p \neq q$ in G' DO

  $S := \overline{S(p, q)}^{G'}$

  IF $S \neq 0$ THEN $G := G \cup \{S\}$

  UNTIL $G = G'$

See [CLO], Chapter 2, §6, [BW], Chapter 5, §3, or [AL], Chapter 1, §7. For instance, in the example above we would adjoin $h = \overline{S(f, g)}^F$ from (3.3) to our set of polynomials. There are two new S-polynomials to consider now: $S(f, h)$ and $S(g, h)$. Their remainders on division by (f, g, h) would be computed and adjoined to the collection if they are nonzero. Then we would continue, forming new S-polynomials and remainders to determine whether further polynomials must be included.

Exercise 5. Carry out Buchberger's Algorithm on the example above, continuing from (3.3). (You may want to use a computer algebra system for this.)

In Maple, there is an implementation of a more sophisticated version of Buchberger's algorithm in the `Groebner` package. The relevant command is called `gbasis`, and the format is

```
gbasis(F,torder);
```

Here `F` is a list of polynomials and `torder` specifies the monomial order. See the description of the `normalf` command in §2 for more details. For instance, the commands

```
F := [x^3*y - 2*x^2*y^2 + x,3*x^4 - y];
gbasis(F,plex(x,y));
```

will compute a *lex* Gröbner basis for the ideal from Exercise 4. The output is

$$[-9y + 48y^{10} - 49y^7 + 6y^4, 252x - 624y^7 + 493y^4 - 3y] \tag{3.4}$$

(possibly up to the ordering of the terms, which can vary). This is not the same as the result of the rudimentary form of Buchberger's algorithm given before. For instance, notice that neither of the polynomials in F actually appears in the output. The reason is that the `gbasis` function actually computes what we will refer to as a *reduced* Gröbner basis for the ideal generated by the list F.

(3.5) Definition. A *reduced Gröbner basis* for an ideal $I \subset k[x_1, \ldots, x_n]$ is a Gröbner basis G for I such that for all distinct $p, q \in G$, no monomial appearing in p is a multiple of $\mathrm{LT}(q)$. A *monic Gröbner basis* is a reduced Gröbner basis in which the leading coefficient of every polynomial is 1, or $\emptyset$ if $I = \langle 0 \rangle$.

Exercise 6. Verify that (3.4) is a reduced Gröbner basis according to this definition.

Exercise 7. Compute a Gröbner basis G for the ideal I from Exercise 1 of this section. Verify that $\overline{p}^G = 0$ now, in agreement with the result of Exercise 3.

A comment is in order concerning (3.5). Many authors include the condition that the leading coefficient of each element in G is 1 in the definition of a reduced Gröbner basis. However, many computer algebra systems (including Maple, see (3.4)) do not perform that extra normalization because it often increases the amount of storage space needed for the Gröbner basis elements when the coefficient field is $\mathbb{Q}$. The reason that condition is often included, however, is the following statement.

- (Uniqueness of Monic Gröbner Bases) Fix a monomial order $>$ on $k[x_1, \ldots, x_n]$. Each ideal I in $k[x_1, \ldots, x_n]$ has a *unique* monic Gröbner basis with respect to $>$.

See [CLO], Chapter 2, §7, [AL], Chapter 1, §8, or [BW], Chapter 5, §2. Of course, varying the monomial order can change the reduced Gröbner basis guaranteed by this result, and one reason different monomial orders are considered is that the corresponding Gröbner bases can have different, useful properties. One interesting feature of (3.4), for instance, is that the second polynomial in the basis does not depend on x. In other words, it is an element of the elimination ideal $I \cap \mathbb{Q}[y]$. In fact, *lex* Gröbner bases systematically eliminate variables. This is the content of the Elimination Theorem from [CLO], Chapter 3, §1. Also see Chapter 2, §1 of this book for further discussion and applications of this remark. On the other hand, the *grevlex* order often minimizes the amount of computation needed to produce a Gröbner basis, so if no other special properties are required, it can be the best choice of monomial order. Other *product orders* and *weight orders* are used in many applications to produce Gröbner bases with special properties. See Chapter 8 for some examples.

Additional Exercises for §3

Exercise 8. Consider the ideal $I = \langle x^2y^2 - x, xy^3 + y\rangle$ from (2.7).

a. Using $>_{lex}$ in $\mathbb{Q}[x, y]$, compute a Gröbner basis G for I.
b. Verify that each basis element g you obtain is in I, by exhibiting equations $g = A(x^2y^2 - x) + B(xy^3 + y)$ for suitable $A, B \in \mathbb{Q}[x, y]$.
c. Let $f = x^3y^2 + 2xy^4$. What is $\overline{f}^G$? How does this compare with the result in (2.7)?

Exercise 9. What monomials can appear in *remainders* with respect to the Gröbner basis G in (3.4)? What monomials appear in leading terms of elements of the ideal generated by G?

Exercise 10. Let G be a Gröbner basis for an ideal $I \subset k[x_1, \ldots, x_n]$ and suppose there exist distinct $p, q \in G$ such that $\mathrm{LT}(p)$ is divisible by $\mathrm{LT}(q)$. Show that $G \setminus \{p\}$ is also a Gröbner basis for I. Use this observation, together with division, to propose an algorithm for producing a reduced Gröbner basis for I given G as input.

Exercise 11. This exercise will sketch a Gröbner basis method for computing the intersection of two ideals. It relies on the Elimination Theorem for *lex* Gröbner bases, as stated in [CLO], Chapter 3, §1. Let $I = \langle f_1, \ldots, f_s\rangle \subset k[x_1, \ldots, x_n]$ be an ideal. Given $f(t)$, an arbitrary polynomial in $k[t]$, consider the ideal

$$f(t)I = \langle f(t)f_1, \ldots, f(t)f_s\rangle \subset k[x_1, \ldots, x_n, t].$$

a. Let I, J be ideals in $k[x_1, \ldots, x_n]$. Show that

$$I \cap J = (tI + (1 - t)J) \cap k[x_1, \ldots, x_n].$$

b. Using the Elimination Theorem, deduce that a Gröbner basis G for $I \cap J$ can be found by first computing a Gröbner basis H for $tI + (1-t)J$ using a *lex* order on $k[x_1, \ldots, x_n, t]$ with the variables ordered $t > x_i$ for all i, and then letting $G = H \cap k[x_1, \ldots, x_n]$.

Exercise 12. Using the result of Exercise 11, derive a Gröbner basis method for computing the quotient ideal $I\colon\langle h\rangle$. Hint: Exercise 13 of §1 shows that if $I \cap \langle h\rangle$ is generated by $g_1, \ldots, g_t$, then $I\colon\langle h\rangle$ is generated by $g_1/h, \ldots, g_t/h$.

§4 Affine Varieties

We will call the set $k^n = \{(a_1, \ldots, a_n) : a_1, \ldots, a_n \in k\}$ the *affine n-dimensional space* over k. With $k = \mathbb{R}$, for example, we have the usual coordinatized Euclidean space $\mathbb{R}^n$. Each polynomial $f \in k[x_1, \ldots, x_n]$ defines a function $f : k^n \to k$. The value of f at $(a_1, \ldots, a_n) \in k^n$ is obtained by substituting $x_i = a_i$, and evaluating the resulting expression in k. More precisely, if we write $f = \sum_\alpha c_\alpha x^\alpha$ for $c_\alpha \in k$, then $f(a_1, \ldots, a_n) = \sum_\alpha c_\alpha a^\alpha \in k$, where

$$a^\alpha = a_1^{\alpha_1} \cdots a_n^{\alpha_n}.$$

We recall the following basic fact.

- (Zero Function) If k is an *infinite* field, then $f : k^n \to k$ is the zero function if and only if $f = 0 \in k[x_1, \ldots, x_n]$.

See, for example, [CLO], Chapter 1, §1. As a consequence, when k is infinite, two polynomials define the same function on k^n if and only if they are equal in $k[x_1, \ldots, x_n]$.

The simplest geometric objects studied in algebraic geometry are the subsets of affine space defined by one or more polynomial equations. For instance, in $\mathbb{R}^3$, consider the set of (x, y, z) satisfying the equation

$$x^2 + z^2 - 1 = 0,$$

a circular cylinder of radius 1 along the y-axis (see Fig. 1.1).

Note that any equation $p = q$, where $p, q \in k[x_1, \ldots, x_n]$, can be rewritten as $p - q = 0$, so it is customary to write all equations in the form $f = 0$ and we will always do this. More generally, we could consider the simultaneous solutions of a system of polynomial equations.

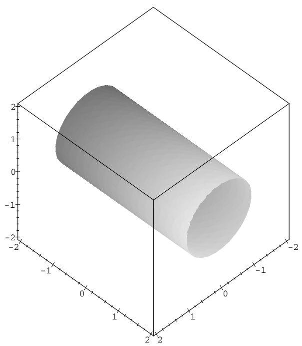

FIGURE 1.1. Circular cylinder

(4.1) Definition. The set of all simultaneous solutions $(a_1, \ldots, a_n) \in k^n$ of a system of equations

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0 \\ f_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ f_s(x_1, \ldots, x_n) &= 0 \end{aligned}$$

is known as the *affine variety* defined by $f_1, \ldots, f_s$, and is denoted by $\mathbf{V}(f_1, \ldots, f_s)$. A subset $V \subset k^n$ is said to be an *affine variety* if $V = \mathbf{V}(f_1, \ldots, f_s)$ for some collection of polynomials $f_i \in k[x_1, \ldots, x_n]$.

In later chapters we will also introduce projective varieties. For now, though, we will often say simply "variety" for "affine variety." For example, $\mathbf{V}(x^2 + z^2 - 1)$ in $\mathbb{R}^3$ is the cylinder pictured above. The picture was generated using the Maple command

```
implicitplot3d(x^2+z^2-1,x=-2..2,y=-2..2,z=-2..2,
    grid=[20,20,20]);
```

The variety $\mathbf{V}(x^2 + y^2 + (z - 1)^2 - 4)$ in $\mathbb{R}^3$ is the sphere of radius 2 centered at $(0, 0, 1)$ (see Fig. 1.2).

If there is more than one defining equation, the resulting variety can be considered as an *intersection* of other varieties. For example, the variety $\mathbf{V}(x^2 + z^2 - 1, x^2 + y^2 + (z - 1)^2 - 4)$ is the curve of intersection of the

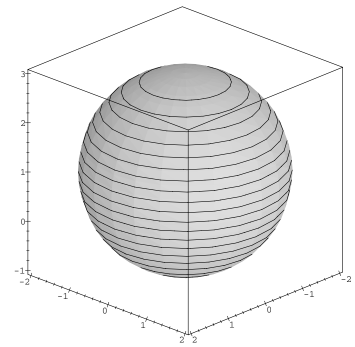

FIGURE 1.2. Sphere

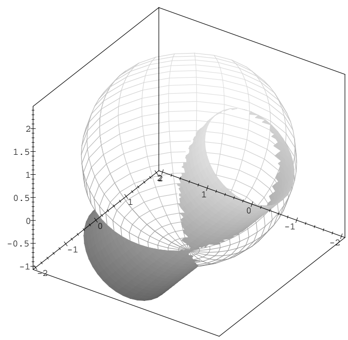

FIGURE 1.3. Cylinder-sphere intersection

cylinder and the sphere pictured above. This is shown, from a viewpoint below the xy-plane, in Fig. 1.3.

The *union* of the sphere and the cylinder is also a variety, namely $\mathbf{V}((x^2 + z^2 - 1)(x^2 + y^2 + (z-1)^2 - 4))$. Generalizing examples like these, we have:

Exercise 1.

a. Show that any finite intersection of affine varieties is also an affine variety.

b. Show that any finite union of affine varieties is also an affine variety. Hint: If $V = \mathbf{V}(f_1, \ldots, f_s)$ and $W = \mathbf{V}(g_1, \ldots, g_t)$, then what is $\mathbf{V}(f_i g_j : 1 \leq i \leq s, 1 \leq j \leq t)$?

c. Show that any finite subset of k^n, $n \geq 1$, is an affine variety.

On the other hand, consider the set $S = \mathbb{R} \setminus \{0, 1, 2\}$, a subset of $\mathbb{R}$. We claim S is *not* an affine variety. Indeed, if f is any polynomial in $\mathbb{R}[x]$ that vanishes at every point of S, then f has infinitely many roots. By standard properties of polynomials in one variable, this implies that f must be the zero polynomial. (This is the one-variable case of the Zero Function property given above; it is easily proved in $k[x]$ using the division algorithm.) Hence the smallest variety in $\mathbb{R}$ containing S is the whole real line itself.

An affine variety $V \subset k^n$ can be described by many different systems of equations. Note that if $g = p_1 f_1 + p_2 f_2 + \cdots + p_s f_s$, where $p_i \in k[x_1, \ldots, x_n]$ are any polynomials, then $g(a_1, \ldots, a_n) = 0$ at each $(a_1, \ldots, a_n) \in \mathbf{V}(f_1, \ldots, f_s)$. So given any set of equations defining a variety, we can always produce infinitely many additional polynomials that also vanish on the variety. In the language of §1 of this chapter, the g as above are just the elements of the ideal $\langle f_1, \ldots, f_s \rangle$. Some collections of these new polynomials can define the same variety as the $f_1, \ldots, f_s$.

Exercise 2. Consider the polynomial p from (1.2). In (1.4) we saw that $p \in \langle x^2 + z^2 - 1, x^2 + y^2 + (z-1)^2 - 4 \rangle$. Show that

$$\langle x^2 + z^2 - 1, x^2 + y^2 + (z-1)^2 - 4 \rangle = \langle x^2 + z^2 - 1, y^2 - 2z - 2 \rangle$$

in $\mathbb{Q}[x, y, z]$. Deduce that

$$\mathbf{V}(x^2 + z^2 - 1, x^2 + y^2 + (z-1)^2 - 4) = \mathbf{V}(x^2 + z^2 - 1, y^2 - 2z - 2).$$

Generalizing Exercise 2 above, it is easy to see that

- (Equal Ideals Have Equal Varieties) If $\langle f_1, \ldots, f_s \rangle = \langle g_1, \ldots, g_t \rangle$ in $k[x_1, \ldots, x_n]$, then $\mathbf{V}(f_1, \ldots, f_s) = \mathbf{V}(g_1, \ldots, g_t)$.

See [CLO], Chapter 1, §4. By this result, together with the Hilbert Basis Theorem from §1, it also makes sense to think of a variety as being defined by an *ideal* in $k[x_1, \ldots, x_n]$, rather than by a specific system of equations. If we want to think of a variety in this way, we will write $V = \mathbf{V}(I)$ where $I \subset k[x_1, \ldots, x_n]$ is the ideal under consideration.

Now, given a variety $V \subset k^n$, we can also try to turn the construction of V from an ideal around, by considering the entire collection of polynomials that vanish at every point of V.

(4.2) Definition. Let $V \subset k^n$ be a variety. We denote by $\mathbf{I}(V)$ the set

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

We call $\mathbf{I}(V)$ the *ideal of* V for the following reason.

Exercise 3. Show that $\mathbf{I}(V)$ is an ideal in $k[x_1, \ldots, x_n]$ by verifying that the two properties in Definition (1.5) hold.

If $V = \mathbf{V}(I)$, is it always true that $\mathbf{I}(V) = I$? The answer is *no*, as the following simple example demonstrates. Consider $V = \mathbf{V}(x^2)$ in $\mathbb{R}^2$. The ideal $I = \langle x^2 \rangle$ in $\mathbb{R}[x, y]$ consists of all polynomials divisible by x^2. These polynomials are certainly contained in $\mathbf{I}(V)$, since the corresponding variety V consists of all points of the form $(0, b)$, $b \in \mathbb{R}$ (the y-axis). Note that $p(x, y) = x \in \mathbf{I}(V)$, but $x \notin I$. In this case, $\mathbf{I}(\mathbf{V}(I))$ is strictly larger than I.

Exercise 4. Show that the following inclusions are always valid:

$$I \subset \sqrt{I} \subset \mathbf{I}(\mathbf{V}(I)),$$

where $\sqrt{I}$ is the *radical* of I from Definition (1.6).

It is also true that the properties of the field k influence the relation between $\mathbf{I}(\mathbf{V}(I))$ and I. For instance, over $\mathbb{R}$, we have $\mathbf{V}(x^2 + 1) = \emptyset$ and $\mathbf{I}(\mathbf{V}(x^2 + 1)) = \mathbb{R}[x]$. On the other hand, if we take $k = \mathbb{C}$, then every polynomial in $\mathbb{C}[x]$ factors completely by the Fundamental Theorem of Algebra. We find that $\mathbf{V}(x^2 + 1)$ consists of the two points $\pm i \in \mathbb{C}$, and $\mathbf{I}(\mathbf{V}(x^2 + 1)) = \langle x^2 + 1 \rangle$.

Exercise 5. Verify the claims made in the preceding paragraph. You may want to start out by showing that if $a \in \mathbb{C}$, then $\mathbf{I}(\{a\}) = \langle x - a \rangle$.

The first key relationships between ideals and varieties are summarized in the following theorems.

- (Strong Nullstellensatz) If k is an *algebraically closed* field (such as $\mathbb{C}$) and I is an ideal in $k[x_1, \ldots, x_n]$, then

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

- (Ideal-Variety Correspondence) Let k be an arbitrary field. The maps

$$\text{affine varieties} \xrightarrow{\mathbf{I}} \text{ideals}$$

and

$$\text{ideals} \xrightarrow{\mathbf{V}} \text{affine varieties}$$

are inclusion-reversing, and $\mathbf{V}(\mathbf{I}(V)) = V$ for all affine varieties V. If k is algebraically closed, then

$$\text{affine varieties} \xrightarrow{\mathbf{I}} \text{radical ideals}$$

and

$$\text{radical ideals} \xrightarrow{\mathbf{V}} \text{affine varieties}$$

are inclusion-reversing bijections, and inverses of each other.

See, for instance [CLO], Chapter 4, §2, or [AL], Chapter 2, §2. We consider how the operations on ideals introduced in §1 relate to operations on varieties in the following exercises.

Additional Exercises for §4

Exercise 6. In §1, we saw that the polynomial $p = x^2 + \frac{1}{2}y^2z - z - 1$ is in the ideal $I = \langle x^2 + z^2 - 1, x^2 + y^2 + (z-1)^2 - 4\rangle \subset \mathbb{R}[x, y, z]$.

a. What does this fact imply about the varieties $\mathbf{V}(p)$ and $\mathbf{V}(I)$ in $\mathbb{R}^3$? ($\mathbf{V}(I)$ is the curve of intersection of the cylinder and the sphere pictured in the text.)
b. Using a 3-dimensional graphing program (e.g. Maple's `implicitplot3d` function from the `plots` package) or otherwise, generate a picture of the variety $\mathbf{V}(p)$.
c. Show that $\mathbf{V}(p)$ contains the variety $W = \mathbf{V}(x^2 - 1, y^2 - 2)$. Describe W geometrically.
d. If we solve the equation

$$x^2 + \tfrac{1}{2}y^2z - z - 1 = 0$$

for z, we obtain

$$z = \frac{x^2 - 1}{1 - \frac{1}{2}y^2}. \tag{4.3}$$

The right-hand side $r(x, y)$ of (4.3) is a quotient of polynomials or, in the terminology of §1, a rational function in x, y, and (4.3) is the equation of the *graph* of $r(x, y)$. Exactly how does this graph relate to the variety $\mathbf{V}(x^2 + \frac{1}{2}y^2z - z - 1)$ in $\mathbb{R}^3$? (Are they the same? Is one a subset of the other? What is the domain of $r(x, y)$ as a function from $\mathbb{R}^2$ to $\mathbb{R}$?)

Exercise 7. Show that for any ideal $I \subset k[x_1, \ldots, x_n]$, $\sqrt{\sqrt{I}} = \sqrt{I}$. Hence $\sqrt{I}$ is automatically a radical ideal.

Exercise 8. Assume k is an algebraically closed field. Show that in the Ideal-Variety Correspondence, sums of ideals (see Exercise 11 of §1) correspond to intersections of the corresponding varieties:

$$\mathbf{V}(I + J) = \mathbf{V}(I) \cap \mathbf{V}(J).$$

Also show that if V and W are any varieties,

$$\mathbf{I}(V \cap W) = \sqrt{\mathbf{I}(V) + \mathbf{I}(W)}.$$

Exercise 9.

a. Show that the intersection of two radical ideals is also a radical ideal.
b. Show that in the Ideal-Variety Correspondence above, intersections of ideals (see Exercise 12 from §1) correspond to unions of the corresponding varieties:

$$\mathbf{V}(I \cap J) = \mathbf{V}(I) \cup \mathbf{V}(J).$$

Also show that if V and W are any varieties,

$$\mathbf{I}(V \cup W) = \mathbf{I}(V) \cap \mathbf{I}(W).$$

c. Show that products of ideals (see Exercise 12 from §1) also correspond to unions of varieties:

$$\mathbf{V}(IJ) = \mathbf{V}(I) \cup \mathbf{V}(J).$$

Assuming k is algebraically closed, how is the product $\mathbf{I}(V)\mathbf{I}(W)$ related to $\mathbf{I}(V \cup W)$?

Exercise 10. A variety V is said to be *irreducible* if in every expression of V as a union of other varieties, $V = V_1 \cup V_2$, either $V_1 = V$ or $V_2 = V$. Show that an affine variety V is irreducible if and only if $\mathbf{I}(V)$ is a prime ideal (see Exercise 8 from §1).

Exercise 11. Let k be algebraically closed.

a. Show by example that the set difference of two affine varieties:

$$V \setminus W = \{p \in V : p \notin W\}$$

need not be an affine variety. Hint: For instance, consider $k[x]$ and let $V = k = \mathbf{V}(0)$ and $W = \{0\} = \mathbf{V}(x)$.
b. Show that for any ideals I, J in $k[x_1, \ldots, x_n]$, $\mathbf{V}(I\colon J)$ contains $\mathbf{V}(I) \setminus \mathbf{V}(J)$, but that we may not have equality. (Here $I\colon J$ is the quotient ideal introduced in Exercise 13 from §1.)
c. If I is a radical ideal, show that $\mathbf{V}(I) \setminus \mathbf{V}(J) \subset \mathbf{V}(I : J)$ and that any variety containing $\mathbf{V}(I) \setminus \mathbf{V}(J)$ must contain $\mathbf{V}(I\colon J)$. Thus $\mathbf{V}(I\colon J)$ is the *smallest* variety containing the difference $\mathbf{V}(I) \setminus \mathbf{V}(J)$; it is called the *Zariski closure* of $\mathbf{V}(I) \setminus \mathbf{V}(J)$. See [CLO], Chapter 4, §4.
d. Show that if I is a radical ideal and J is any ideal, then $I\colon J$ is also a radical ideal. Deduce that $\mathbf{I}(V)\colon\mathbf{I}(W)$ is the radical ideal corresponding to the Zariski closure of $V \setminus W$ in the Ideal-Variety Correspondence.