Chapter 2

Algorithms for Algebraic Geometry

Outline:

2.1 Gröbner basics

Gröbner bases are a foundation for many algorithms to represent and manipulate varieties on a computer. While these algorithms are important in applications, we shall see that Gröbner bases are also a useful theoretical tool.

A motivating problem is that of recognizing when a polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ lies in an ideal I. When the ideal I is radical and $\mathbb{K}$ is algebraically closed, this is equivalent to asking whether or not f vanishes on $\mathcal{V}(I)$. For example, we may ask which of the polynomials $x^3z - xz^3$, $x^2yz - y^2z^2 - x^2y^2$, and/or $x^2y - x^2z + y^2z$ lies in the ideal

$$\langle x^2y - xz^2 + y^2z,\ y^2 - xz + yz\rangle \text{ ?}$$

This *ideal membership problem* is easy for univariate polynomials. Suppose that $I = \langle f(x), g(x), \ldots, h(x)\rangle$ is an ideal and $F(x)$ is a polynomial in $\mathbb{K}[x]$, the ring of polynomials in a single variable x. We determine if $F(x) \in I$ via a two-step process.

1. Use the Euclidean Algorithm to compute $\varphi(x) := \gcd(f(x), g(x), \ldots, h(x))$.

2. Use the Division Algorithm to determine if $\varphi(x)$ divides $F(x)$.

This is valid, as $I = \langle \varphi(x) \rangle$. The first step is a simplification, where we find a simpler (lower-degree) polynomial which generates I, while the second step is a reduction, where we compute F modulo I. Both steps proceed systematically, operating on the terms of the polynomials involving the highest power of x. A good description for I is a prerequisite for solving our ideal membership problem.

We shall see how Gröbner bases give algorithms which extend this procedure to multivariate polynomials. In particular, a Gröbner basis of an ideal I gives a sufficiently good description of I to solve the ideal membership problem. Gröbner bases are also the foundation of algorithms that solve many other problems.

Monomial ideals are central to what follows. A *monomial* is a product of powers of the variables $x_1, \ldots, x_n$ with nonnegative integer exponents. The *exponent* of a monomial $x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ is a vector $\alpha \in \mathbb{N}^n$. If we identify monomials with their exponent vectors, the multiplication of monomials corresponds to addition of vectors, and divisibility to the partial order on $\mathbb{N}^n$ of componentwise comparison.

Definition 2.1.1. A *monomial ideal* $I \subset \mathbb{K}[x_1, \ldots, x_n]$ is an ideal which satisfies the following two equivalent conditions.

(*i*) I is generated by monomials.

(*ii*) If $f \in I$, then every monomial of f lies in I.

One advantage of monomial ideals is that they are essentially combinatorial objects. By Condition (*ii*), a monomial ideal is determined by the set of monomials which it contains. Under the correspondence between monomials and their exponents, divisibility of monomials corresponds to componentwise comparison of vectors.

$$x^\alpha | x^\beta \iff \alpha_i \leq \beta_i\,,\ i = 1, \ldots, n \iff \alpha \leq \beta\,,$$

which defines a partial order on $\mathbb{N}^n$. Thus

$$(1,1,1) \leq (3,1,2) \qquad \text{but} \qquad (3,1,2) \not\leq (2,3,1)\,.$$

The set $O(I)$ of exponent vectors of monomials in a monomial ideal I has the property that if $\alpha \leq \beta$ with $\alpha \in O(I)$, then $\beta \in O(I)$. Thus $O(I)$ is an (upper) *order ideal* of the *poset* (*p*artially *o*rdered *set*) $\mathbb{N}^n$.

A set of monomials $G \subset I$ generates I if and only if every monomial in I is divisible by at least one monomial of G. A monomial ideal I has a unique minimal set of generators—these are the monomials x^α in I which are not divisible by any other monomial in I.

Let us look at some examples. When $n = 1$, monomials have the form x^d for some natural number $d \geq 0$. If d is the minimal exponent of a monomial in I, then $I = \langle x^d \rangle$. Thus all monomial ideals have the form $\langle x^d \rangle$ for some $d \geq 0$.

When $n = 2$, we may plot the exponents in the order ideal associated to a monomial ideal. For example, the lattice points in the shaded region of Figure 2.1 represent the

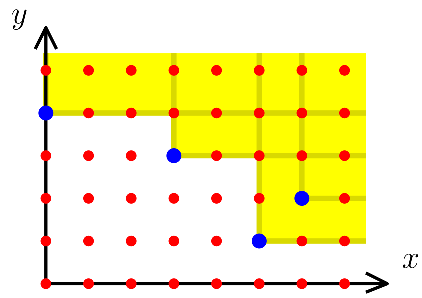
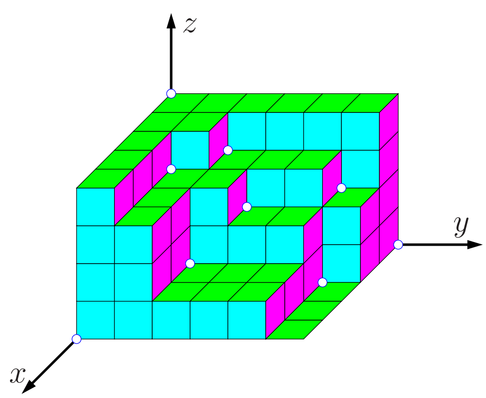


Figure 2.1: Exponents of monomials in the ideal $\langle y^4,\, x^3y^3,\, x^5y,\, x^6y^2\rangle$.

monomials in the ideal $I := \langle y^4,\, x^3y^3,\, x^5y,\, x^6y^2\rangle$, with the generators marked. From this picture we see that I is minimally generated by y^4, x^3y^3, and x^5y.

Since $x^ay^b \in I$ implies that $x^{a+\alpha}y^{b+\beta} \in I$ for any $(\alpha,\beta) \in \mathbb{N}^2$, a monomial ideal $I \subset \mathbb{K}[x,y]$ is the union of the shifted positive quadrants $(a,b)+\mathbb{N}^2$ for every monomial $x^ay^b \in I$. It follows that the monomials in I are those above the staircase shape that is the boundary of the shaded region. The monomials not in I lie under the staircase, and they form a vector space basis for the quotient ring $\mathbb{K}[x,y]/I$.

This notion of staircase for two variables makes sense when there are more variables. The *staircase* of an ideal consists of the monomials which are on the boundary of the ideal, in that they are visible from the origin of $\mathbb{N}^n$. For example, here is the staircase for the ideal $\langle x^5,\, x^2y^5,\, y^6,\, x^3y^2z,\, x^2y^3z^2,\, xy^5z^2,\, x^2yz^3,\, xy^2z^3,\, z^4\rangle$.



We offer a purely combinatorial proof that monomial ideals are finitely generated, which is independent of the Hilbert Basis Theorem.

Lemma 2.1.2 (Dickson's Lemma)**.** *Monomial ideals are finitely generated.*

Proof. We prove this by induction on n. The case $n = 1$ was covered in the preceding examples.

Let $I \subset \mathbb{K}[x_1,\ldots,x_n,y]$ be a monomial ideal. For each $d \in \mathbb{N}$, observe that the set of monomials

$$\{x^\alpha \mid x^\alpha y^d \in I\}\,,$$

generates a monomial ideal I_d of $\mathbb{K}[x_1,\ldots,x_n]$, and the union of all such monomials,

$$\{x^\alpha \mid x^\alpha y^d \in I \text{ for some } d \geq 0\}\,,$$

generates a monomial ideal I_∞ of $\mathbb{K}[x_1,\ldots,x_n]$. By our inductive hypothesis, I_d has a finite generating set G_d, for each $d = 0, 1, \ldots, \infty$.

Note that $I_0 \subset I_1 \subset \cdots \subset I_\infty$. We must have $I_\infty = I_d$ for some $d < \infty$. Indeed, each generator $x^\alpha \in G_\infty$ of I_∞ comes from a monomial $x^\alpha y^b$ in I, and we may let d be the maximum of the numbers b which occur. Since $I_\infty = I_d$, we have $I_b = I_d$ for any $b > d$. Note that if $b > d$, then we may assume that $G_b = G_d$ as $I_b = I_d$.

We claim that the finite set

$$G \;=\; \bigcup_{b=0}^{d} \{x^\alpha y^b \mid x^\alpha \in G_b\}$$

generates I. Indeed, let $x^\alpha y^b$ be a monomial in I. We find a monomial in G which divides $x^\alpha y^b$. Since $x^\alpha \in I_b$, there is a generator $x^\gamma \in G_b$ which divides x^α. If $b \leq d$, then $x^\gamma y^b \in G$ is a monomial dividing $x^\alpha y^b$. If $b > d$, then $x^\gamma y^d \in G$ as $G_b = G_d$ and $x^\gamma y^d$ divides $x^\alpha y^b$. □

A simple consequence of Dickson's Lemma is that any strictly increasing chain of monomial ideals is finite. Suppose that

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

is an increasing chain of monomial ideals. Let I_∞ be their union, which is another monomial ideal. Since I_∞ is finitely generated, there must be some ideal I_d which contains all generators of I_∞, and so $I_d = I_{d+1} = \cdots = I_\infty$. We used this fact crucially in our proof of Dickson's lemma.

The key idea behind Gröbner bases is to determine what is meant by 'term of highest power' in a polynomial having two or more variables. There is no canonical way to do this, so we must make a choice, which is encoded in the notion of a term or monomial order. An order $\succ$ on monomials in $\mathbb{K}[x_1,\ldots,x_n]$ is *total* if for monomials x^α and x^β exactly one of the following holds

$$x^\alpha \;\succ\; x^\beta \qquad \text{or} \qquad x^\alpha \;=\; x^\beta \qquad \text{or} \qquad x^\alpha \;\prec\; x^\beta\,.$$

Definition 2.1.3. A *monomial order* on $\mathbb{K}[x_1,\ldots,x_n]$ is a total order $\succ$ on the monomials in $\mathbb{K}[x_1,\ldots,x_n]$ such that

(*i*) 1 is the minimal element under $\succ$.

(*ii*) $\succ$ respects multiplication by monomials: If $x^\alpha \succ x^\beta$ then $x^\alpha \cdot x^\gamma \succ x^\beta \cdot x^\gamma$, for any monomial x^γ.

Conditions (i) and (ii) in Definition 2.1.3 imply that if x^α is divisible by x^β, then $x^\alpha \succ x^\beta$. A *well-ordering* is a total order with no infinite descending chain, equivalently, one in which every subset has a minimal element.

Lemma 2.1.4. *Monomial orders are exactly the well-orderings $\succ$ on monomials that satisfy Condition (ii) of Definition* 2.1.3.

Proof. Let $\succ$ be a well-ordering on monomials which satisfies Condition (ii) of Definition 2.1.3. Suppose that $\succ$ is not a monomial order, then there is some monomial x^a with $1 \succ x^a$. By Condition (ii), we have $1 \succ x^a \succ x^{2a} \succ x^{3a} \succ \cdots$, which contradicts $\succ$ being a well-order, and 1 is the $\succ$-minimal monomial.

Let $\succ$ be a monomial order and M be any set of monomials. Set I to be the ideal generated by M. By Dickson's Lemma, I is generated by a finite set G of monomials. We may assume that $G \subset M$, for if $x^\alpha \in G \smallsetminus M$, then as M generates I, there is some $x^\beta \in M$ that divides x^α, and so we may replace x^α by x^β in G. After finitely many such replacements, we will have that $G \subset M$. Since G is finite, let x^γ be the minimal monomial in G under $\succ$. We claim that x^γ is the minimal monomial in M.

Let $x^\alpha \in M$. Since G generates I and $M \subset I$, there is some $x^\beta \in G$ which divides x^α and thus $x^\alpha \succ x^\beta$. But x^γ is the minimal monomial in G, so $x^\alpha \succ x^\beta \succ x^\gamma$. $\square$

The well-ordering property of monomials orders is key to what follows, as many proofs use induction on $\succ$, which is only possible as $\succ$ is a well-ordering.

Example 2.1.5. The (*total*) *degree,* $\deg(x^\alpha)$, of a monomial $x^\alpha = x_1^{\alpha_1}\cdots x_n^{\alpha_n}$ is $\alpha_1 + \cdots + \alpha_n$. We describe four important monomial orders.

1. The *lexicographic order* $\succ_{lex}$ on $\mathbb{K}[x_1,\ldots,x_n]$ is defined by

$$x^\alpha \succ_{\text{lex}} x^\beta \iff \left\{\begin{matrix}\text{The first non-zero entry of the}\\ \text{vector } \alpha-\beta \text{ in } \mathbb{Z}^n \text{ is positive.}\end{matrix}\right\}$$

2. The *degree lexicographic order* $\succ_{dlx}$ on $\mathbb{K}[x_1,\ldots,x_n]$ is defined by

$$x^\alpha \succ_{\text{dlx}} x^\beta \iff \begin{cases} \deg(x^\alpha) > \deg(x^\beta) & \text{or}\,, \\ \deg(x^\alpha) = \deg(x^\beta) & \text{and } \ x^\alpha \succ_{\text{lex}} x^\beta\,. \end{cases}$$

3. The *degree reverse lexicographic order* $\succ_{drl}$ $\mathbb{K}[x_1,\ldots,x_n]$ is defined by

$$x^\alpha \succ_{\text{drl}} x^\beta \iff \begin{cases} \deg(x^\alpha) > \deg(x^\beta) & \text{or}\,, \\ \deg(x^\alpha) = \deg(x^\beta) & \text{and the last non-zero entry of the} \\ & \text{vector } \alpha-\beta \text{ in } \mathbb{Z}^n \text{ is negative}\,. \end{cases}$$

4. More generally, we have *weighted orders*. Let $\omega \in \mathbb{R}^n$ be a vector with non-negative components, called a weight. This defines a partial order $\succ_\omega$ on monomials

$$x^\alpha \;\succ_\omega\; x^\beta \iff \omega \cdot \alpha > \omega \cdot \beta\,.$$

If all components of ω are positive, then $\succ_\omega$ satisfies the two conditions of Definition 2.1.3. Its only failure to be a monomial order is that it may not be a total order on monomials. (For example, consider $\omega = (1, 1, \ldots, 1)$, then $\omega \cdot \alpha$ is the total degree of x^α.) This may be remedied by picking a monomial order to break ties. For example, if we use $\succ_{\text{lex}}$, then we get a monomial order

$$x^\alpha \;\succ_{\omega,\text{lex}}\; x^\beta \iff \begin{cases} \omega \cdot \alpha > \omega \cdot \beta & \text{or}\,, \\ \omega \cdot \alpha = \omega \cdot \beta & \text{and } x^\alpha \succ_{\text{lex}} x^\beta \end{cases}$$

Another way to do this is to break the ties with a different monomial order, or a different weight, and this may be done recursively.

You are asked to prove these are monomial orders in Exercise 7.

Remark 2.1.6. We compare these three orders on monomials of degrees 1 and 2 in $\mathbb{K}[x, y, z]$ where the variables are ordered $x \succ y \succ z$.

$$\begin{gathered}
x^2 \succ_{\text{lex}} xy \succ_{\text{lex}} xz \succ_{\text{lex}} x \succ_{\text{lex}} y^2 \succ_{\text{lex}} yz \succ_{\text{lex}} y \succ_{\text{lex}} z^2 \succ_{\text{lex}} z \\
x^2 \succ_{\text{dlx}} xy \succ_{\text{dlx}} xz \succ_{\text{dlx}} y^2 \succ_{\text{dlx}} yz \succ_{\text{dlx}} z^2 \succ_{\text{dlx}} x \succ_{\text{dlx}} y \succ_{\text{dlx}} z \\
x^2 \succ_{\text{drl}} xy \succ_{\text{drl}} y^2 \succ_{\text{drl}} xz \succ_{\text{drl}} yz \succ_{\text{drl}} z^2 \succ_{\text{drl}} x \succ_{\text{drl}} y \succ_{\text{drl}} z
\end{gathered}$$

For the remainder of this section, $\succ$ will denote a fixed, but arbitrary monomial order on $\mathbb{K}[x_1, \ldots, x_n]$. A *term* is a product ax^α of a scalar $a \in \mathbb{K}$ with a monomial x^α. We may extend any monomial order $\succ$ to an order on terms by setting $ax^\alpha \succ bx^\beta$ if $x^\alpha \succ x^\beta$ and $ab \neq 0$. Such a term order is no longer a partial order as different terms with the same monomial are incomparable. For example $3x^2y$ and $5x^2y$ are incomparable. Term orders are however well-founded in that they have no infinite strictly decreasing chains.

The *initial term* $\text{in}_\succ(f)$ of a polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ is the term of f that is maximal with respect to $\succ$ among all terms of f. For example, if $\succ$ is lexicographic order with $x \succ y$, then

$$\text{in}_\succ(3x^3y - 7xy^{10} + 13y^{30}) \;=\; 3x^3y\,.$$

When $\succ$ is understood, we may write $\text{in}(f)$. As $\succ$ is a total order that respects the multiplication of monomials, taking initial terms is multiplicative, $\text{in}_\succ(fg) = \text{in}_\succ(f)\text{in}_\succ(g)$, for $f, g \in \mathbb{K}[x_1, \ldots, x_n]$. The *initial ideal* $\text{in}_\succ(I)$ (or $\text{in}(I)$) of an ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ is the ideal generated by the initial terms of polynomials in I,

$$\text{in}_\succ(I) \;=\; \langle \text{in}_\succ(f) \mid f \in I \rangle\,.$$

We make the most important definition of this section.

Definition 2.1.7. Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal and $\succ$ a monomial order. A set $G \subset I$ is a Gröbner basis for I with respect to the monomial order $\succ$ if the initial ideal $\mathrm{in}_\succ(I)$ is generated by the initial terms of polynomials in G, that is, if

$$\mathrm{in}_\succ(I) \;=\; \langle \mathrm{in}_\succ(g) \mid g \in G\rangle\,.$$

Notice that if G is a Gröbner basis and $G \subset G'$, then G' is also a Gröbner basis. Note also that I is a Gröbner basis for I, and every Gröbner basis contains a finite subset that is also a Gröbner basis, by Dickson's Lemma.

We justify our use of the term 'basis' in 'Gröbner basis'.

Lemma 2.1.8. *If G is a Gröbner basis for I with respect to a monomial order $\succ$, then G generates I.*

Proof. We begin with a computation and a definition. Let $f \in I$. Since $\{\mathrm{in}(g) \mid g \in G\}$ generates $\mathrm{in}(I)$, there is a polynomial $g \in G$ whose initial term $\mathrm{in}(g)$ divides the initial term $\mathrm{in}(f)$ of f. Thus there is some term ax^α so that

$$\mathrm{in}(f) \;=\; ax^\alpha \mathrm{in}(g) \;=\; \mathrm{in}(ax^\alpha g)\,,$$

as $\succ$ respects multiplication. If we set $f_1 := f - cx^\alpha g$, then $\mathrm{in}(f) \succ \mathrm{in}(f_1)$.

We will prove the lemma by induction on $\mathrm{in}(f)$ for $f \in I$. Suppose first that $f \in I$ is a polynomial whose initial term $\mathrm{in}(f)$ is the $\succ$-minimal monomial in $\mathrm{in}(I)$. Then $f_1 = 0$ and so $f \in \langle G\rangle$. In fact, up to a scalar multiple, $f \in G$. Suppose now that $I \neq \langle G\rangle$, and let $f \in I$ be a polynomial with $\mathrm{in}(f)$ is $\succ$-minimal among all $f \in I \smallsetminus \langle G\rangle$. But then $f_1 = f - cx^\alpha g \in I$ and as $\mathrm{in}(f) \succ \mathrm{in}(f_1)$, we must have that $f_1 \in \langle G\rangle$, which implies that $f \in \langle G\rangle$, a contradiction. $\square$

An immediate consequence of Dickson's Lemma and Lemma 2.1.8 is the following Gröbner basis version of the Hilbert Basis Theorem.

Theorem 2.1.9 (Hilbert Basis Theorem)**.** *Every ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ has a finite Gröbner basis with respect to any given monomial order.*

Example 2.1.10. Different monomial orderings give different Gröbner bases, and the sizes of the Gröbner bases can vary. Consider the ideal generated by the three polynomials

$$xy^3 + xz^3 + x - 1,\quad yz^3 + yx^3 + y - 1,\quad zx^3 + zy^3 + z - 1$$

In the degree reverse lexicographic order, where $x \succ y \succ z$, this has a Gröbner basis

$x^3z + y^3z + z - 1,$
$xy^3 + xz^3 + x - 1,$
$x^3y + yz^3 + y - 1,$
$y^4z - yz^4 - y + z,$
$2xyz^4 + xyz + xy - xz - yz,$

$$
\begin{array}{l}
2y^3z^3 - x^3 + y^3 + z^3 + x^2 - y^2 - z^2, \\
y^6 - z^6 - y^5 + y^3z^2 - 2x^2z^3 - y^2z^3 + z^5 + y^3 - z^3 - x^2 - y^2 + z^2 + x, \\
x^6 - z^6 - x^5 - y^3z^2 - x^2z^3 - 2y^2z^3 + z^5 + x^3 - z^3 - x^2 - y^2 + y + z, \\
2z^7 + 4x^2z^4 + 4y^2z^4 - 2z^6 + 3z^4 - x^3 - y^3 + 3x^2z + 3y^2z - 2z^3 + x^2 + y^2 - 2xz - 2yz - z^2 + z - 1, \\
2yz^6 + y^4 + 2yz^3 + x^2y - y^3 + yz^2 - 2z^3 + y - 1, \\
2xz^6 + x^4 + 2xz^3 - x^3 + xy^2 + xz^2 - 2z^3 + x - 1,
\end{array}
$$

consisting of 11 polynomials with largest coefficient 4 and degree 7. If we consider instead the lexicographic monomial order, then this ideal has a Gröbner basis

$$
\begin{array}{l}
64z^{34} - 64z^{33} + 384z^{31} - 192z^{30} - 192z^{29} + 1008z^{28} + 48z^{27} - 816z^{26} + 1408z^{25} + 976z^{24} \\
-1296z^{23} + 916z^{22} + 1964z^{21} - 792z^{20} - 36z^{19} + 1944z^{18} + 372z^{17} - 405z^{16} + 1003z^{15} \\
+879z^{14} - 183z^{13} + 192z^{12} + 498z^{11} + 7z^{10} - 94z^9 + 78z^8 + 27z^7 - 47z^6 - 31z^5 + 4z^3 \\
-3z^2 - 4z - 1,
\end{array}
$$

$$
\begin{array}{l}
64yz^{21} + 288yz^{18} + 96yz^{17} + 528yz^{15} + 384yz^{14} + 48yz^{13} + 504yz^{12} + 600yz^{11} + 168yz^{10} \\
+200yz^9 + 456yz^8 + 216yz^7 + 120yz^5 + 120yz^4 - 8yz^2 + 16yz + 8y - 64z^{33} + 128z^{32} \\
-128z^{31} - 320z^{30} + 576z^{29} - 384z^{28} - 976z^{27} + 1120z^{26} - 144z^{25} - 2096z^{24} + 1152z^{23} \\
+784z^{22} - 2772z^{21} + 232z^{20} + 1520z^{19} - 2248z^{18} - 900z^{17} + 1128z^{16} - 1073z^{15} - 1274z^{14} \\
+229z^{13} - 294z^{12} - 966z^{11} - 88z^{10} - 81z^9 - 463z^8 - 69z^7 + 26z^6 - 141z^5 - 32z^4 + 24z^3 \\
-12z^2 - 11z + 1
\end{array}
$$

$$
\begin{array}{l}
589311934509212912y^2 - 11786238690184258240yz^{20} - 9428990952147406592yz^{19} \\
-23572477380368851648yz^{18} - 48323578629755458784yz^{17} - 48323578629755458784yz^{16} \\
-20036605773313239008yz^{15} - 81914358896780594768yz^{14} - 97825781128529343392yz^{13} \\
-53038074105829162080yz^{12} - 78673143256979923752yz^{11} - 99888372899311588584yz^{10} \\
-63645688926994994496yz^9 - 37126651874080413456yz^8 - 43903739120936361944yz^7 \\
-34474748168788955352yz^6 - 9134334984892800136yz^5 - 5893119345092129120yz^4 \\
-4125183541564490384yz^3 - 1178623869018425824yz^2 - 2062591770782245192yz \\
-1178623869018425824y + 4666564515534984336z^{33} - 5256138633033865068z^{32} \\
+2519587235202032992z^{31} + 28156769162372952723z^{30} - 19392177430724378694z^{29} \\
-22383823960598695936z^{28} + 81706533724600969099z^{27} - 16308104685758723524z^{26} \\
-42770559036883403033z^{25} + 139057816837182085380z^{24} + 39000434368484674580z^{23} \\
-98032219788785598166z^{22} + 134542511722129797387z^{21} + 128795606593903673167z^{20} \\
-95338316228249822884z^{19} + 63120234731058122985z^{18} + 170430196786922739602z^{17} \\
-155208567786555149988z^{16} - 167640668622573965 05z^{15} + 125747540327715070096z^{14} \\
+52668596890136716959z^{13} - 164751530000556264880z^{12} + 49124953163927565405z^{11} \\
+45712630887118688230z^{10} - 87008396189513562747z^9 + 15803768907185828750z^8 \\
+13932068156394410127z^7 - 17355919586383317961z^6 - 50777365233910819054z^5 \\
-4630862847055988750z^4 + 8085080238139562826z^3 + 1366850803924776890z^2 \\
-3824545208919673161z - 2755936363893486164,
\end{array}
$$

$$
\begin{array}{l}
589311934509212912x + 589311934509212912y - 87966378396509318592z^{33} \\
+13338340253167146649 6z^{32} - 59115312141727767552z^{31} - 50692680764859328012 8z^{30} \\
+52214177181017233427 2z^{29} + 48286434009450032640z^{28} - 143472598833873638875 2z^{27} \\
+62997181176686959171 2z^{26} + 91798600277439166526 4z^{25} - 238987119897484320513 6z^{24}
\end{array}
$$

$$-2469823148310669418888z^{23}+2038968926105271519536z^{22}-2174896389643343086620z^{21}$$
$$-1758138782546221156976z^{20}+2025390185406562798552z^{19}-774542641420363828364z^{18}$$
$$-2365390641451278278484z^{17}+627824835559363304992z^{16}+398484633232859115907z^{15}$$
$$-1548683110130934220322z^{14}-500192666710091510419z^{13}+551921427998474758510z^{12}$$
$$-490368794345102286410z^{11}-480504004841899057384z^{10}+220514007454401175615z^{9}$$
$$+385159849019800047305z^{8}-136644301635686684609z^{7}+174107126941325207 94z^{6}$$
$$+58724552354094225803z^{5}+15702341971895307356z^{4}-7440058907697789332z^{3}$$
$$-1398341089468668912z^{2}+391320563053161 2397z+268914524400616 8857,$$

consisting of 4 polynomials with largest degree 34 and significantly larger coefficients.

Exercises

1. Prove the equivalence of conditions (i) and (ii) in Definition 2.1.1.

2. Show that a monomial ideal is radical if and only if it is square-free. (Square-free means that it has generators in which no variable occurs to a power greater than 1.)

3. Show that the elements of a monomial ideal I which are minimal with respect to division form a minimal set of generators of I in that they generate I and are a subset of any generating set of I.

4. Which of the polynomials x^3z-xz^3, $x^2yz-y^2z^2-x^2y^2$, and/or $x^2y-x^2z+y^2z$ lies in the ideal
$$\langle x^2y-xz^2+y^2z,\ y^2-xz+yz\rangle\ ?$$

5. Using Definition 2.1.1, show that a monomial order is a linear extension of the divisibility partial order on monomials.

6. Show that if an ideal I has a square-free initial ideal, then I is radical. Give an example to show that the converse of this statement is false.

7. Show that each of the order relations $\succ_{\mathrm{lex}}$, $\succ_{\mathrm{dlx}}$, and $\succ_{\mathrm{drl}}$, are monomial orders. Show that if the coordinates of $\omega\in\mathbb{R}^n_>$ are linearly independent over $\mathbb{Q}$, then $\succ_\omega$ is a monomial order. Show that each of $\succ_{\mathrm{lex}}$, $\succ_{\mathrm{dlx}}$, and $\succ_{\mathrm{drl}}$ are weighted orders.

8. Suppose that $\succ$ is a term order. Prove that for any two nonzero polynomials f,g, we have $\mathrm{in}_\succ(fg)=\mathrm{in}_\succ(f)\mathrm{in}_\succ(g)$.

9. Show that for a monomial order $\succ$, $\mathrm{in}(I)\mathrm{in}(J)\subseteq\mathrm{in}(IJ)$ for any two ideals I and J. Find I and J such that the inclusion is proper.

2.2 Algorithmic applications of Gröbner bases

Many practical algorithms to study and manipulate ideals and varieties are based on Gröbner bases. The foundation of algorithms involving Gröbner bases is the multivariate division algorithm. The subject began with Buchberger's thesis which contained his algorithm to compute Gröbner bases.

Both steps in the algorithm for ideal membership in one variable relied on the same elementary procedure: using a polynomial of low degree to simplify a polynomial of higher degree. This same procedure was also used in the proof of Lemma 2.1.8. This leads to the *multivariate division algorithm*, which is a cornerstone of the theory of Gröbner bases.

Algorithm 2.2.1 (Multivariate division algorithm).
INPUT: Polynomials $g_1, \ldots, g_m, f$ in $\mathbb{K}[x_1, \ldots, x_n]$ and a monomial order $\succ$.
OUTPUT: Polynomials $q_1, \ldots, q_m$ and r such that

$$f \;=\; q_1 g_1 + q_2 g_2 + \cdots + q_m g_m + r\,, \tag{2.1}$$

where no term of r is divisible by an initial term of any polynomial g_i and we also have $\text{in}(f) \succeq \text{in}(r)$, and $\text{in}(f) \succeq \text{in}(q_i g_i)$, for each $i = 1, \ldots, m$.
INITIALIZE: Set $r := f$ and $q_1 := 0, \ldots, q_m := 0$. Perform the following steps.

(1) If no term of r is divisible by an initial term of some g_i, then exit.

(2) Otherwise, let ax^α be the largest (with respect to $\succ$) term of r divisible by some $\text{in}(g_i)$. Choose j minimal such that $\text{in}(g_j)$ divides x^α and suppose that $ax^\alpha = bx^\beta \cdot \text{in}(g_j)$. Replace r by $r - bx^\beta g_j$ and q_j by $q_j + bx^\beta$, and return to step (1).

Proof of correctness. Each iteration of (2) is a *reduction* of r by the polynomials $g_1, \ldots, g_m$. With each reduction, the largest term in r divisible by some $\text{in}(g_i)$ decreases with respect to $\succ$. Since the term order $\succ$ is well-founded, this algorithm must terminate after a finite number of steps. Every time the algorithm executes step (1), condition (2.1) holds. We also always have $\text{in}(f) \succeq \text{in}(r)$ because it holds initially, and with every reduction any new terms of r are less than the term that was canceled. Lastly, $\text{in}(f) \succeq \text{in}(q_i g_i)$ always holds, because it held initially, and the initial terms of the $q_i g_i$ are always terms of r. $\square$

Given a list $G = (g_1, \ldots, g_m)$ of polynomials and a polynomial f, let r be the remainder obtained by the multivariate division algorithm applied to G and f. Since $f - r$ lies in the ideal generated by G, we write *f mod G* for this remainder r. While it is clear (and expected) that f mod G depends on the monomial order $\succ$, in general it will also depend upon the order of the polynomials $(g_1, \ldots, g_m)$. For example, in the degree lexicographic order

$$\begin{aligned} x^2y \bmod (x^2,\, xy + y^2) &= 0\,, \qquad \text{but} \\ x^2y \bmod (xy + y^2,\, x^2) &= y^3\,. \end{aligned}$$

This example shows that we cannot reliably use the multivariate division algorithm to test when f is in the ideal generated by G. However, this does not occur when G is a Gröbner basis.

Lemma 2.2.2 (Ideal membership test)**.** *Let G be a finite Gröbner basis for an ideal I with respect to a monomial order $\succ$. Then a polynomial $f \in I$ if and only if $f \bmod G = 0$.*

Proof. Set $r := f \bmod G$. If $r = 0$, then $f \in I$. Suppose $r \neq 0$. Since no term of r is divisible any initial term of a polynomial in G, its initial term $\mathrm{in}(r)$ is not in the initial ideal of I, as G is a Gröbner basis for I. But then $r \notin I$, and so $f \notin I$. □

When G is a Gröbner basis for an ideal I and $f \in \mathbb{K}[x_1, \ldots, x_n]$, no term of the remainder $f \bmod G$ lies in the initial ideal of I. A monomial x^α is *standard* if $x^\alpha \notin \mathrm{in}(I)$. The images of standard monomials in the ring $\mathbb{K}[x_1, \ldots, x_n]/\mathrm{in}(I)$ form a vector space basis. Much more interesting is the following theorem of Macaulay.

Theorem 2.2.3. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal and $\succ$ a monomial order. Then the images of standard monomials in $\mathbb{K}[x_1, \ldots, x_n]/I$ form a vector space basis.*

Proof. Let G be a finite Gröbner basis for I with respect to $\succ$. Given a polynomial f, both f and $f \bmod G$ represent the same element in $\mathbb{K}[x_1, \ldots, x_n]/I$. Since $f \bmod G$ is a linear combination of standard monomials, the standard monomials span $\mathbb{K}[x_1, \ldots, x_n]/I$.

A linear combination f of standard monomials is zero in $\mathbb{K}[x_1, \ldots, x_n]/I$ only if $f \in I$. But then $\mathrm{in}(f)$ is both standard and lies in $\mathrm{in}(I)$, and so we conclude that $f = 0$. Thus the standard monomials are linearly independent in $\mathbb{K}[x_1, \ldots, x_n]/I$. □

Because of Macaulay's Theorem, if we have a monomial order $\succ$ and an ideal I, then for every polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$, there is a unique polynomial $\overline{f}$ which involves only standard monomials such that f and $\overline{f}$ have the same image in the quotient ring $\mathbb{K}[x_1, \ldots, x_m]/I$. Moreover, this polynomial $\overline{f}$ equals $f \bmod G$, where G is any finite Gröbner basis of I with respect to the monomial order $\succ$, and thus may be computed from f and G using the division algorithm. This unique representative $\overline{f}$ of f is called the *normal form* of f modulo I and the division algorithm called with a Gröbner basis for I is often called *normal form reduction*.

Macaulay's Theorem shows that a Gröbner basis allows us to compute in the quotient ring $\mathbb{K}[x_1, \ldots, x_n]/I$ using the operations of the polynomial ring and ordinary linear algebra. Indeed, suppose that G is a finite Gröbner basis for an ideal I with respect to a given monomial order $\succ$ and that $f, g \in \mathbb{K}[x_1, \ldots, x_n]/I$ are in normal form, expressed as a linear combination of standard monomials. Then $f + g$ is a linear combination of standard monomials and we can compute the product fg in the quotient ring as $fg \bmod G$, where this product is taken in the polynomial ring.

Theorem 2.1.9, which asserted the existence of a finite Gröbner basis, was purely existential. To use Gröbner bases, we need methods to detect and generate them. Such methods were given by Bruno Buchberger in his 1965 Ph.D. thesis.

A given set of generators for an ideal will fail to be a Gröbner basis if the initial terms of the generators fail to generate the initial ideal. That is, if there are polynomials in the ideal whose initial terms are not divisible by the initial terms of our generators. A necessary step towards generating a Gröbner basis is to generate polynomials in the ideal with 'new' initial terms. This is the *raison d'etre* for the following definition.

Definition 2.2.4. The *least common multiple*, $\mathrm{lcm}\{ax^\alpha, bx^\beta\}$ of two terms ax^α and bx^β is the minimal monomial x^γ divisible by both x^α and x^β. Here, the exponent vector γ is the componentwise maximum of α and β.

Let $0 \neq f, g \in \mathbb{K}[x_1, \ldots, x_n]$ and suppose $\succ$ is a monomial order. The *S-polynomial* of f and g, $\mathrm{Spol}(f, g)$, is the polynomial linear combination of f and g,

$$\mathrm{Spol}(f,g) \;:=\; \frac{\mathrm{lcm}\{\mathrm{in}(f), \mathrm{in}(g)\}}{\mathrm{in}(f)} f \;-\; \frac{\mathrm{lcm}\{\mathrm{in}(f), \mathrm{in}(g)\}}{\mathrm{in}(g)} g\,.$$

Note that both terms in this expression have initial term equal to $\mathrm{lcm}\{\mathrm{in}(f), \mathrm{in}(g)\}$.

Buchberger gave the following simple criterion to detect when a set G of polynomials is a Gröbner basis for the ideal it generates.

Theorem 2.2.5 (Buchberger's Criterion). *A set G of polynomials is a Gröbner basis for the ideal it generates with respect to a monomial order $\succ$ if and only if for for all pairs $f, g \in G$,*

$$\mathrm{Spol}(f,g) \bmod G \;=\; 0\,.$$

Proof. Suppose first that G is a Gröbner basis for I with respect to $\succ$. Then, for $f, g \in G$, their S-polynomial $\mathrm{Spol}(f, g)$ lies in I and the ideal membership test implies that $\mathrm{Spol}(f, g) \bmod G = 0$.

Now suppose that $G = \{g_1, \ldots, g_m\}$ satisfies Buchberger's criterion and let I be the ideal generated by G. Let $f \in I$. We will show that $\mathrm{in}(f)$ is divisible by $\mathrm{in}(g)$, for some $g \in G$. This implies that G is a Gröbner basis for I.

Given a list $h = (h_1, \ldots, h_m)$ of polynomials in $\mathbb{K}[x_1, \ldots, x_n]$ let $\mathrm{mm}(h)$ be the largest monomial appearing in one of $h_1g_1, \ldots, h_mg_m$. This will necessarily be the monomial in at least one of the initial terms $\mathrm{in}(h_1g_1), \ldots, \mathrm{in}(h_mg_m)$. Let $j(h)$ be the minimum index i for which $\mathrm{mm}(h)$ is the monomial of $\mathrm{in}(h_ig_i)$.

Consider lists $h = (h_1, \ldots, h_m)$ of polynomials with

$$f \;=\; h_1g_1 + \cdots + h_mg_m \tag{2.2}$$

for which $\mathrm{mm}(h)$ minimal among all lists satisfying (2.2). Of these, let h be a list with $j := j(h)$ maximal. We claim that $\mathrm{mm}(h)$ is the monomial of $\mathrm{in}(f)$, which implies that $\mathrm{in}(g_j)$ divides $\mathrm{in}(f)$.

Otherwise, $\mathrm{mm}(h) \succ \mathrm{in}(f)$, and so the initial term $\mathrm{in}(h_jg_j)$ must be canceled in the sum (2.2). Thus there is some index k such that $\mathrm{mm}(h)$ is the monomial of $\mathrm{in}(h_kg_k)$. By our assumption on j, we have $k > j$. Let $x^\beta := \mathrm{lcm}\{\mathrm{in}(g_j), \mathrm{in}(g_k)\}$, the monomial which is canceled in $\mathrm{Spol}(g_j, g_k)$. Since $\mathrm{in}(g_j)$ and $\mathrm{in}(g_k)$ both divide $\mathrm{mm}(h)$, both divide $\mathrm{in}(h_j g_j)$, and there is some term ax^α such that $ax^\alpha x^\beta = \mathrm{in}(h_j g_j)$. Set $cx^\gamma := \mathrm{in}(h_j g_j)/\mathrm{in}(g_k)$. Then

$$ax^\alpha \mathrm{Spol}(g_j, g_k) \;=\; ax^\alpha \frac{x^\beta}{\mathrm{in}(g_j)} g_j \;-\; ax^\alpha \frac{x^\beta}{\mathrm{in}(g_k)} g_k \;=\; \mathrm{in}(h_j) g_j \;-\; cx^\gamma g_k\,.$$

As Buchberger's criterion is satisfied, there are polynomials $q_1, \ldots, q_m$ with

$$\mathrm{Spol}(g_j, g_k) \;=\; q_1 g_1 + \cdots + q_m g_m\,,$$

and we may assume that $\mathrm{in}(q_i g_i) \preceq \mathrm{in}(\mathrm{Spol}(g_j, g_k)) \prec x^\beta$, by the division algorithm and the construction of $\mathrm{Spol}(g_j, g_k)$.

Define a new list h' of polynomials,

$$h' \;=\; (h_1 + ax^\alpha q_1, \ldots, h_j - \mathrm{in}(h_j) + ax^\alpha q_j, \ldots, h_k + cx^\gamma + ax^\alpha q_k, \ldots, h_m + ax^\alpha q_m)\,,$$

and consider the sum $\sum h'_i g_i$, which is

$$\sum_i h_i g_i + ax^\alpha \sum_i q_i g_i - \mathrm{in}(h_j) g_j + cx^\gamma g_k$$
$$= \; f + ax^\alpha \mathrm{Spol}(g_j, g_k) - ax^\alpha \mathrm{Spol}(g_j, g_k) \;=\; f\,,$$

so h' is a list satisfying (2.2).

We have $\mathrm{in}(q_i g_i) \preceq \mathrm{in}(\mathrm{Spol}(g_j, g_k))$, so $\mathrm{in}(ax^\alpha q_i g_i) \prec x^\alpha x^\beta = \mathrm{mm}(h)$. But then $\mathrm{mm}(h') \preceq \mathrm{mm}(h)$. By the minimality of $\mathrm{mm}(h)$, we have $\mathrm{mm}(h') = \mathrm{mm}(h)$. Since $\mathrm{in}(h_j - \mathrm{in}(h_j)) \prec \mathrm{in}(h_j)$, we have $j(h') > j = j(h)$, which contradicts our choice of h. □

Buchberger's algorithm to compute a Gröbner basis begins with a list of polynomials and augments that list by adding reductions of S-polynomials. It halts when the list of polynomials satisfies Buchberger's Criterion.

Algorithm 2.2.6 (Buchberger's Algorithm)**.** Let $G = (g_1, \ldots, g_m)$ be generators for an ideal I and $\succ$ a monomial order. For each $1 \leq i < j \leq m$, let $h_{ij} := \mathrm{Spol}(g_i, g_j) \bmod G$. If each reduction vanishes, then by Buchberger's Criterion, G is a Gröbner basis for I with respect to $\succ$. Otherwise append all the non-zero h_{ij} to the list G and repeat this process.

This algorithm terminates after finitely many steps, because the initial terms of polynomials in G after each step generate a strictly larger monomial ideal and Dickson's Lemma implies that any increasing chain of monomial ideals is finite. Since the manipulations in Buchberger's algorithm involve only algebraic operations using the coefficients of the input polynomials, we deduce the following corollary, which is important when studying real varieties. Let $\Bbbk$ be any subfield of $\mathbb{K}$.

Corollary 2.2.7. *Let $f_1, \ldots, f_m \in \Bbbk[x_1, \ldots, x_n]$ be polynomials and $\succ$ a monomial order. Then there is a Gröbner basis $G \subset \Bbbk[x_1, \ldots, x_n]$ for the ideal $\langle f_1, \ldots, f_m \rangle$ in $\mathbb{K}[x_1, \ldots, x_n]$ with respect to the monomial order $\succ$.*

Example 2.2.8. Consider applying the Buchberger algorithm to $G = (x^2, xy + y^2)$ with any monomial order where $x \succ y$. First

$$\mathrm{Spol}(x^2, xy + y^2) \;=\; y \cdot x^2 - x(xy + y^2) \;=\; -xy^2\,.$$

Then

$$-xy^2 \bmod (x^2, xy + y^2) \;=\; -xy^2 + y(xy + y^2) \;=\; y^3\,.$$

Since all S-polynomials of $(x^2, xy + y^2, y^3)$ reduce to zero, this is a Gröbner basis. ⋄

Among the polynomials h_{ij} computed at each stage of the Buchberger algorithm are those where one of $\mathrm{in}(g_i)$ or $\mathrm{in}(g_j)$ divides the other. Suppose that $\mathrm{in}(g_i)$ divides $\mathrm{in}(g_j)$ with $i \neq j$. Then $\mathrm{Spol}(g_i, g_j) = g_j - ax^\alpha g_i$, where ax^α is some term. This has strictly smaller initial term than does g_j and so we never use g_j to compute $h_{ij} := \mathrm{Spol}(g_i, g_j) \bmod G$. It follows that $g_j - h_{ij}$ lies in the ideal generated by $G \smallsetminus \{g_j\}$ (and *vice-versa*), and so we may replace g_j by h_{ij} in G without changing the ideal generated by G, and only possibly increasing the ideal generated by the initial terms of polynomials in G.

This gives the following elementary improvement to the Buchberger algorithm:

> In each step, initially compute h_{ij} for those $i \neq j$ where $\mathrm{in}(g_i)$ divides $\mathrm{in}(g_j)$, and replace g_j by h_{ij}. (2.3)

In some important cases, this step computes the Gröbner basis. Another improvement, which identifies some S-polynomials that reduce to zero and therefore need not be computed, is given in Exercise 3.

A Gröbner basis G is *reduced* if the initial terms of polynomials in G are monomials with coefficient 1 and if for each $g \in G$, no monomial of g is divisible by an initial term of another Gröbner basis element. A reduced Gröbner basis for an ideal is uniquely determined by the monomial order. Reduced Gröbner bases are the multivariate analog of unique monic polynomial generators of ideals of $\mathbb{K}[x]$. Elements f of a reduced Gröbner basis have a special form,

$$x^\alpha - \sum_{\beta \in B} c_\beta x^\beta\,, \tag{2.4}$$

where $x^\alpha = \mathrm{in}(f)$ is the initial term and B consists of exponent vectors of standard monomials. This rewrites the nonstandard initial monomial as a linear combination of standard monomials. In this way a Gröbner basis may be thought of as system of rewriting rules for polynomials. The reduced Gröbner basis has one generator for every generator of the initial ideal.

Example 2.2.9. Let M be a $m \times n$ matrix, which we consider to be the matrix of coefficients of m linear forms $g_1, \ldots, g_m$ in $\mathbb{K}[x_1, \ldots, x_n]$, and suppose that $x_1 \succ x_2 \succ \cdots \succ x_n$. We can apply (2.3) to two forms g_i and g_j when their initial terms have the same variable. Then the S-polynomial and subsequent reductions are equivalent to the steps in the algorithm of Gaussian elimination applied to the matrix M. If we iterate our applications of (2.3) until the initial terms of the forms g_i have distinct variables, then the forms $g_1, \ldots, g_m$ are a Gröbner basis for the ideal they generate.

If the forms g_i are a reduced Gröbner basis and are sorted in decreasing order according to their initial terms, then the resulting matrix $\overline{M}$ of their coefficients is an *echelon matrix*: The initial non-zero entry in each row is 1 and is the only non-zero entry in its column and these columns increase with row number.

Gaussian elimination produces the same echelon matrix from M. Thus the Buchberger algorithm is a generalization of Gaussian elimination to non-linear polynomials. $\diamond$

The form (2.4) of elements in a reduced Gröbner basis G for an ideal I with respect to a given monomial order $\succ$ implies that G depends on the monomial ideal $\mathrm{in}_\succ(I)$, and thus only indirectly on $\succ$. That is, if $\succ'$ is a second monomial order with $\mathrm{in}_{\succ'}(I) = \mathrm{in}_\succ(I)$, then G is also a Göber basis for I with respect to $\succ'$. It turns out that while there are uncountably many monomial orders, any given ideal has only finitely many initial ideals.

Theorem 2.2.10. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal. Then its set of initial ideals,*

$$\mathrm{In}(I) \ := \ \{\mathrm{in}_\succ(I) \mid \succ \ \textit{is a monomial order}\}$$

is finite.

Proof. For each initial ideal M in $\mathrm{In}(I)$, choose a monomial order $\succ_M$ such that $M = \mathrm{in}_{\succ_M}(I)$. Let

$$T \ := \ \{\succ_M \mid M \in \mathrm{In}(I)\}$$

be this set of monomial orders, one for each initial ideal of I.

Suppose that $\mathrm{In}(I)$ and hence T is infinite and let $g_1, \ldots, g_m \in \mathbb{K}[x_1, \ldots, x_n]$ be generators for I. Since each polynomial g_i has only finitely many terms, there is an infinite subset T_1 of T with the property that any two monomial orders $\succ, \succ'$ in T_1 will select the same initial terms from each of the g_i,

$$\mathrm{in}_\succ(g_i) \ = \ \mathrm{in}_{\succ'}(g_i) \qquad \text{for } i = 1, \ldots, m\,.$$

Set $M_1 := \langle \mathrm{in}_\succ(g_1), \ldots, \mathrm{in}_\succ(g_m)\rangle$, where $\succ$ is any monomial order in T_1. Either $(g_1, \ldots, g_m)$ is a Gröbner basis for I with respect to $\succ$ or else there is a some polynomial g_{m+1} in I whose initial term does not lie in M_1. Replacing g_{m+1} by $g_{m+1} \bmod (g_1, \ldots, g_m)$, we may assume that g_{m+1} has no term in M_1.

Then there is an infinite subset T_2 of T_1 such that any two monomial orders $\succ, \succ'$ in T_2 will select the same initial term of g_{m+1}, $\mathrm{in}_\succ(g_{m+1}) = \mathrm{in}_{\succ'}(g_{m+1})$. Let M_2 be the monomial ideal generated by M_1 and $\mathrm{in}_\succ(g_{m+1})$ for some monomial order $\succ$ in T_2. As before, either $(g_1,\dots,g_m,g_{m+1})$ is a Gröbner basis for I with resepct to $\succ$, or else there is an element g_{m+2} of I having no term in M_2.

Continuing in this fashion, we construct an increasing chain $M_1 \subsetneq M_2 \subsetneq \cdots$ of monomial ideals in $\mathbb{K}[x_1,\dots,x_n]$. By Dickson's Lemma, this process must terminate, at which point we will have an infinite subset T_r of T and polynomials $g_1,\dots,g_{m+r}$ that form a Gröbner basis for I with respect to a monomial order $\succ$ in T_r, and these have the property that for any other monomial order $\succ'$ in T_r, we have

$$\mathrm{in}_{\succ}(g_i) \;=\; \mathrm{in}_{\succ'}(g_i) \qquad \text{for } i=1,\dots,m{+}r\,.$$

But this implies that $\mathrm{in}_{\succ}(I) = \mathrm{in}_{\succ'}(I)$ is an initial ideal for two distinct monomial orders in $T_r \subset T$, which contradicts the construction of the set T. □

Definition 2.2.11. A consequence of Theorem 2.2.10 that an ideal I has only finitely many initial ideals is that it has only finitely many reduced Gröbner bases. The union of this finite set of reduced Gröbner bases is a finite generating set for I that is a Gröbner basis for I with resepct to any monomial order. Such a generating set is called a *universal Gröbner basis* for the ideal I. The existence of universal Gröbner bases has a number of useful consequences.

Exercises

1. Describe how Buchberger's algorithm behaves when it computes a Gröbner basis from a list of monomials. What if we use the elementary improvement (2.3)?

2. Use Buchberger's algorithm to compute by hand the reduced Gröbner basis of $\langle y^2 - xz + yz,\; x^2y - xz^2 + y^2z\rangle$ in the degree reverse lexicographic order where $x \succ y \succ z$.

3. Let $f,g \in \mathbb{K}[x_1,\dots,x_m]$ be polynomials with relatively prime initial terms, and suppose that their leading coefficients are 1.

   (a) Show that
   $$\mathrm{Spol}(f,g) \;=\; -(g-\mathrm{in}(g))f + (f-\mathrm{in}(f))g\,.$$
   Deduce that the leading monomial of $\mathrm{Spol}(f,g)$ is a multiple of either the leading monomial of f or the leading monomial of g.

   (b) Analyze the steps of the reduction computing $\mathrm{Spol}(f,g)$ mod (f,g) using the division algorithm to show that this is zero.

   This illustrates another improvement on Buchberger's algorithm: avoid computing and reducing (to zero) those S-polynomials of polynomials with relatively prime initial terms.

4. Let U be a universal Gröbner basis for an ideal I in $\mathbb{K}[x_1, \ldots, x_n]$. Show that for every subset $Y \subset \{x_1, \ldots, x_n\}$ the *elimination ideal* $I \cap \mathbb{K}[Y]$ is generated by $U \cap \mathbb{K}[Y]$.

5. Let I be a ideal generated by homogeneous linear polynomials. We call a nonzero linear form f in I a *circuit* of I if f has minimal support (with respect to inclusion) among all polynomials in I. Prove that the set of all circuits of I is a universal Gröbner basis of I.

6. Let $I := \langle x^2 + y^2, x^3 + y^3 \rangle \subset \mathbb{Q}[x, y]$ and suppose that the monomial order $\succ$ is the lexicographic order with $x \succ y$.
   (a) Show that $y^4 \in I$.
   (b) Show that the reduced Gröbner basis for I is $\{y^4, xy^2 - y^3, x^2 + y^2\}$.
   (c) Show that $\{x^2 + y^2, x^3 + y^3\}$ cannot be a Gröbner basis for I for any monomial ordering.

7. (a) Prove that the ideal $\langle x, y \rangle \subset \mathbb{Q}[x, y]$ is not a principal ideal.
   (b) Is $\langle x^2 + y, x + y \rangle$ already a Gröbner basis with respect to some term ordering?
   (c) Use Buchberger's algorithm to compute by hand a Gröbner basis of the ideal $I = \langle y - z^2, z - x^3 \rangle \in \mathbb{Q}[x, y, z]$ with lexicographic and the degree reverse lexicographic monomial orders.

8. This exercise illustrates an algorithm to compute the saturation of ideals. Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal, and fix $f \in \mathbb{K}[x_1, \ldots, x_n]$. Then the *saturation* of I with respect to f is the set

$$(I : f^\infty) \;=\; \{g \in \mathbb{K}[x_1, \ldots, x_n] \mid f^m g \in I \text{ for some } m > 0\}\,.$$

   (a) Prove that $(I : f^\infty)$ is an ideal.
   (b) Prove that we have an ascending chain of ideals

$$(I : f) \;\subset\; (I : f^2) \;\subset\; (I : f^3) \;\subset \cdots$$

   (c) Prove that there exists a nonnegative integer N such that $(I : f^\infty) = (I : f^N)$.
   (d) Prove that $(I : f^\infty) = (I : f^m)$ if and only if $(I : f^m) = (I : f^{m+1})$.

   When the ideal I is homogeneous and $f = x_n$ then one can use the following strategy to compute the saturation. Fix the degree reverse lexicographic order $\succ$ where $x_1 \succ x_2 \succ \cdots \succ x_n$ and let G be a reduced Gröbner basis of a homogeneous ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$.

(e) Show that the set

$$G' \;=\; \{f \in G \mid x_n \text{ does not divide } f\} \bigcup \{f/x_n \mid f \in G \text{ and } x_n \text{ divides } f\}$$

is a Gröbner basis of $(I : x_n)$.

(f) Show that a Gröbner basis of $(I : x_n^\infty)$ is obtained by dividing each element $f \in G$ by the highest power of x_n that divides f.

9. Suppose that $\prec$ is the lexicographic order with $x \prec y \prec z$.

   (a) Apply Buchberger's algorithm to the ideal $\langle x + y, xy\rangle$.

   (b) Apply Buchberger's algorithm to the ideal $\langle x + y + z, xy + xz + yz, xyz\rangle$.

   (c) Define the *elementary symmetric polynomials $e_i(x_1, \ldots, x_n)$* by

$$\sum_{i=0}^{n} t^{n-i} e_i(x_1, \ldots, x_n) \;=\; \prod_{i=1}^{n} (t + x_i)\,,$$

   that is, $e_0 = 1$ and if $i > 0$, then

$$e_i(x_1, \ldots, x_n) \;:=\; e_i(x_1, \ldots, x_{n-1}) \;+\; x_n e_{i-1}(x_1, \ldots, x_{n-1})\,.$$

   Alternatively, $e_i(x_1, \ldots, x_n)$ is also the sum of all square-free monomials of total degree i in $x_1, \ldots, x_n$.

   The symmetric ideal is $\langle e_i(x_1, \ldots, x_n) \mid 1 \le i \le n\rangle$. Describe its Gröbner basis and the set of standard monomials with respect to lexicographic order when $x_1 \prec x_2 \prec \cdots \prec x_n$.

   What is its Gröbner basis with respect to degree reverse lexicographic order? How about an order with $x_1 \prec x_2 \prec \cdots \prec x_n$?

   (d) Describe a universal Gröbner basis for the symmetric ideal.

2.3 Resultants and Bézout's Theorem

Algorithms based on Gröbner bases are universal in that their input may be any list of polynomials. This comes at a price as Gröbner basis algorithms may have poor performance and the output is quite sensitive to the input. An alternative foundation for some algorithms is provided by resultants. These are are special polynomials having determinantal formulas which were introduced in the 19th century. A drawback is that they are not universal—different inputs require different algorithms, and for many inputs, there are no formulas for resultants.

The key algorithmic step in the Euclidean algorithm for the greatest common divisor (gcd) of two univariate polynomials f and g in $\mathbb{K}[x]$ with $n = \deg(g) \geq \deg(f) = m$,

$$\begin{aligned} f &= f_0x^m + f_1x^{m-1} + \cdots + f_{m-1}x + f_m \\ g &= g_0x^n + g_1x^{n-1} + \cdots + g_{n-1}x + g_n\,, \end{aligned} \tag{2.5}$$

is to replace g by

$$g - \frac{g_0}{f_0}x^{n-m} \cdot f\,,$$

which has degree at most $n-1$. (Note that $f_0 \cdot g_0 \neq 0$.) In some cases (for example, when $\mathbb{K}$ is a function field), we will want to avoid division. Resultants give a way to detect common factors without using division. We will use them for much more than this.

Let $\mathbb{K}$ be any field, not necessarily algebraically closed or even infinite. Let $\mathbb{K}_\ell[x]$ be the set of polynomials in $\mathbb{K}[x]$ of degree at most ℓ. This is a vector space over $\mathbb{K}$ of dimension $\ell+1$ with a canonical ordered basis of monomials $x^\ell, \ldots, x, 1$. Given f and g as in (2.5), consider the linear map

$$\begin{aligned} L_{f,g} \,:\, \mathbb{K}_{n-1}[x] \times \mathbb{K}_{m-1}[x] &\longrightarrow \mathbb{K}_{m+n-1}[x] \\ (h(x), k(x)) &\longmapsto f \cdot h + g \cdot k\,. \end{aligned}$$

The domain and range of $L_{f,g}$ each have dimension $m + n$.

Lemma 2.3.1. *The polynomials f and g have a nonconstant common divisor if and only if* $\ker L_{f,g} \neq \{(0,0)\}$.

Proof. Suppose first that f and g have a nonconstant common divisor, p. Then there are polynomials h and k with $f = pk$ and $g = ph$. As p is nonconstant, $\deg(k) < \deg(f) = m$ and $\deg(h) < \deg(g) = n$ so that $(h, -k) \in \mathbb{K}_{n-1}[x] \times \mathbb{K}_{m-1}[x]$. Since

$$fh - gk \;=\; pkh - phk \;=\; 0\,,$$

we see that $(h, -k)$ is a nonzero element of the kernel of $L_{f,g}$.

Suppose that f and g are relatively prime and let $(h, k) \in \ker L_{f,g}$. Since $\langle f, g\rangle = \mathbb{K}[x]$, there exist polynomials p and q with $1 = gp + fq$. Using $0 = fh + gk$ we obtain

$$k \;=\; k \cdot 1 \;=\; k(gp + fq) \;=\; gkp + fkq \;=\; -fhp + fkq \;=\; f(kq - hp)\,.$$

This implies that $k = 0$ for otherwise $m-1 \geq \deg(k) > \deg(f) = m$, which is a contradiction. We similarly have $h = 0$, and so $\ker L_{f,g} = \{(0,0)\}$. $\square$

The matrix of the linear map $L_{f,g}$ in the ordered bases of monomials for $\mathbb{K}_{m-1}[x] \times \mathbb{K}_{n-1}[x]$ and $\mathbb{K}_{m+n-1}[x]$ is called the *Sylvester matrix*. When f and g have the form (2.5), it is

$$\mathrm{Syl}(f,g;x) \;=\; \mathrm{Syl}(f,g) \;:=\; \left(\begin{array}{cccc|ccc} f_0 & & & & g_0 & & 0 \\ f_1 & f_0 & & 0 & g_1 & \ddots & \\ \vdots & \vdots & \ddots & & \vdots & & g_0 \\ f_m & \vdots & & \ddots & \vdots & & \vdots \\ & f_m & & f_0 & g_{n-1} & & \vdots \\ & & \ddots & \vdots & g_n & & \vdots \\ & 0 & & \ddots\ \vdots & & \ddots & \vdots \\ & & & f_m & 0 & & g_n \end{array}\right) . \tag{2.6}$$

Note that the sequence $f_0, \ldots, f_0, g_n, \ldots, g_n$ lies along the main diagonal and the left side of the matrix has n columns while the right side has m columns.

Often, we will treat the coefficients $f_0, \ldots, f_m, g_0, \ldots, g_m$ of f and g as variables. That is, we will regard them as algebraically independent over $\mathbb{Q}$ of $\mathbb{Z}$. Any formulas proven under this assumption will remain valid when the coefficients of f and g lie in any field or ring.

The (*Sylvester*) *resultant* $\mathrm{Res}(f,g)$ is the determinant of the Sylvester matrix. To emphasize that the Sylvester matrix represents the map $L_{f,g}$ in the basis of monomials in x, we also write $\mathrm{Res}(f,g;x)$ for $\mathrm{Res}(f,g)$. We summarize some properties of resultants, which follow from its formula as the determinant of the Sylvester matrix (2.6) and from Lemma 2.3.1.

Theorem 2.3.2. *The resultant of two nonconstant polynomials $f, g \in \mathbb{K}[x]$ is an integer polynomial in the coefficients of f and g. The resultant vanishes if and only if f and g have a nonconstant common factor.*

We give another expression for the resultant in terms of the roots of f and g.

Lemma 2.3.3. *Suppose that $\mathbb{K}$ contains all the roots of the polynomials f and g so that*

$$f(x) \;=\; f_0 \prod_{i=1}^{m}(x - \alpha_i) \qquad \textit{and} \qquad g(x) \;=\; g_0 \prod_{i=1}^{n}(x - \beta_i)\,,$$

where $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ are the roots of f and $\beta_1, \ldots, \beta_n \in \mathbb{K}$ are the roots of g. Then

$$\mathrm{Res}(f,g;x) \;=\; (-1)^{mn} f_0^n g_0^m \prod_{i=1}^{m}\prod_{j=1}^{n}(\alpha_i - \beta_j)\,. \tag{2.7}$$

This implies the Poisson formula,

$$\mathrm{Res}(f,g;x) \;=\; (-1)^{mn} f_0^m \prod_{i=1}^m g(\alpha_i) \;=\; g_0^n \prod_{i=1}^n f(\beta_i)\,.$$

Proof. Consider these formulas as expressions in $\mathbb{Z}[f_0, g_0, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n]$. Recall that the coefficients of f and g are essentially the elementary symmetric polynomials in their roots,

$$f_i \;=\; (-1)^i f_0 e_i(\alpha_1,\ldots,\alpha_m) \qquad \text{and} \qquad g_i \;=\; (-1)^i g_0 e_i(\beta_1,\ldots,\beta_n)\,.$$

We claim that both sides of (2.7) are homogeneous polynomials of degree mn in the variables $\alpha_1, \ldots, \beta_n$. This is straightforward for the right hand side. For the resultant, we extend our notation, setting $f_i := 0$ when $i < 0$ or $i > m$ and $g_i := 0$ when $i < 0$ or $i > n$. Then the entry in row i and column j of the Sylvester matrix is

$$\mathrm{Syl}(f,g;x)_{i,j} \;=\; \begin{cases} f_{i-j} & \text{if } j \le n\,, \\ g_{n+i-j} & \text{if } n < j \le m+n\,. \end{cases}$$

The determinant is a signed sum over permutations w of $\{1, \ldots, m{+}n\}$ of terms

$$\prod_{j=1}^n f_{w(j)-j} \;\cdot\; \prod_{j=n+1}^{m+n} g_{n+w(j)-j}\,.$$

Since f_i and g_i are each homogeneous of degree i in the variables $\alpha_1, \ldots, \beta_n$, this term is homogeneous of degree

$$\sum_{j=1}^m w(j){-}j \quad + \quad \sum_{j=n+1}^{m+n} n{+}w(j){-}j \;=\; mn + \sum_{j=1}^{m+n} w(j){-}j \;=\; mn\,,$$

which proves the claim.

Both sides of (2.7) vanish exactly when some $\alpha_i = \beta_j$. Since they have the same degree, they are proportional. This will now be done in Chapter 1 as a consequence of the Nullstellensatz. We compute this constant of proportionality. The term in $\mathrm{Res}(f,g)$ which is the product of diagonal entries of the Sylvester matrix is

$$f_0^n g_n^m \;=\; f_0^n g_0^m e_n(\beta_1,\ldots,\beta_n)^m \;=\; f_0^n g_0^m \beta_1^m \cdots \beta_n^m\,.$$

This is the only term of $\mathrm{Res}(f,g)$ involving the monomial $\beta_1^m \cdots \beta_n^m$. The corresponding term on the right hand side of (2.7) is

$$(-1)^{mn} f_0^n g_0^m (-\beta_1)^m \cdots (-\beta_n)^m \;=\; f_0^n g_0^m \beta_1^m \cdots \beta_n^m\,,$$

which completes the proof. ∎

Remark 3.2.8 uses geometric arguments to show that the resultant is irreducible and gives another characterization of resultants, which we give below.

Theorem 2.3.4. *The resultant polynomial is irreducible. It is the unique (up to sign) irreducible integer polynomial in the coefficients of f and g that vanishes on the set of pairs of polynomials (f, g) which have a common root.*

When both f and g have the same degree n, there is an alternative formula for their resultant as the determinant of a $n \times n$ matrix. (Sylvester's formula is as the determinant of a $2n \times 2n$ matrix.) The *Bezoutian polynomial* of f and g is the bivariate polynomial

$$\Delta_{f,g}(y,z) \;:=\; \frac{f(y)g(z) - f(z)g(y)}{y-z} \;=\; \sum_{i,j=0}^{n-1} b_{i,j} y^i z^j \;.$$

The $n \times n$ matrix $\mathrm{Bez}(f,g)$ whose entries are the coefficients $(b_{i,j})$ of the Bezoutian polynomial is called the *Bezoutian matrix* of f and g. Each entry of the Bezoutian matrix $\mathrm{Bez}(f,g)$ is a linear combination of the brackets $[ij] := f_i g_j - f_j g_i$. For example, when $n = 2$ and $n = 3$, the Bezoutian matrices are

$$\begin{pmatrix} [02] & [12] \\ [01] & [02] \end{pmatrix} \qquad \begin{pmatrix} [03] & [13] & [23] \\ [02] & [03]+[12] & [13] \\ [01] & [02] & [03] \end{pmatrix} .$$

Theorem 2.3.5. *When f and g both have degree n,* $\mathrm{Res}(f,g) = (-1)^{\binom{n}{2}} \det(\mathrm{Bez}(f,g))$.

Proof. Suppose that $\mathbb{K}$ is algebraically closed. Let B be the determinant of the Bezoutian matrix and Res the resultant of the polynomials f and g, both of which lie in the ring $\mathbb{K}[f_0, \ldots, f_n, g_0, \ldots, g_n]$. Then B is a homogeneous polynomial of degree $2n$, as is the resultant. Suppose that f and g are polynomials having a common root, $a \in \mathbb{K}$ with $f(a) = g(a) = 0$. Then the Bezoutian polynomial $\Delta_{f,g}(y,z)$ vanishes when $z = a$,

$$\Delta_{f,g}(y,a) \;=\; \frac{f(y)g(a) - f(a)g(y)}{y-a} \;=\; 0\,.$$

Thus

$$0 \;=\; \sum_{i,j=0}^{n-1} b_{i,j} y^i a^j \;=\; \sum_{i=0}^{n-1} \Big(\sum_{j=0}^{n-1} b_{i,j} a^j \Big) y^i\,.$$

Since every coefficient of this polynomial in y must vanish, the vector $(1, a, a^2, \ldots, a^{d-1})^T$ lies in the kernel of the Bezoutian matrix, and so the determinant $B(f,g)$ of the Bezoutian matrix vanishes.

Since the resultant generates the ideal of the pairs (f,g) of polynomial that are not relatively prime, Res divides B. As they have the same degree B is a constant multiple of Res. In Exercise 5 you are asked to show this constant is $(-1)^{\binom{n}{2}}$. $\square$

Example 2.3.6. We give an application of resultants. A polynomial $f \in \mathbb{K}[x]$ of degree n has fewer than n distinct roots in the algebraic closure of $\mathbb{K}$ when it has a factor in $\mathbb{K}[x]$ of multiplicity greater than 1, and in that case f and its derivative f' have a factor in common. The *discriminant* of f is a polynomial in the coefficients of f which vanishes precisely when f has a repeated factor. It is defined to be

$$\mathrm{disc}(f) \ := \ \frac{(-1)^{\binom{n}{2}}}{f_0}\mathrm{Res}(f, f')\,.$$

The discriminant is a polynomial of degree $2n-2$ in the coefficients $f_0, f_1, \ldots, f_n$.

Resultants do much more than detect the existence of common factors in two polynomials. One of their most important uses is to eliminate variables from multivariate equations. The first step towards this is another interesting formula involving the Sylvester resultant. Not only is it a polynomial in the coefficients, but it has a canonical expression as a polynomial linear combination of f and g.

Lemma 2.3.7. *Given univariate polynomials $f, g \in \mathbb{K}[x]$, there are polynomials $h, k \in \mathbb{K}[x]$ whose coefficients are universal integer polynomials in the coefficients of f and g such that*

$$f(x)h(x) + g(x)k(x) \ = \ \mathrm{Res}(f, g)\,. \tag{2.8}$$

Proof. Set $\mathbb{K} := \mathbb{Q}(f_0, \ldots, f_m, g_0, \ldots, g_n)$, the field of rational functions (quotients of integer polynomials) in the variables $f_0, \ldots, f_m, g_0, \ldots, g_n$ and let $f, g \in \mathbb{K}[x]$ be univariate polynomials as in (2.5). Then $\gcd(f, g) = 1$ and so the map $L_{f,g}$ is invertible.

Set $(h, k) := L_{f,g}^{-1}(\mathrm{Res}(f, g))$ so that

$$f(x)h(x) \ + \ g(x)k(x) \ = \ \mathrm{Res}(f, g)\,,$$

with $h, k \in \mathbb{K}[x]$ where $h \in \mathbb{K}_{n-1}[x]$ and $k \in \mathbb{K}_{m-1}[x]$.

Recall the adjoint formula for the inverse of a $n \times n$ matrix A,

$$\det(A) \cdot A^{-1} \ = \ \mathrm{ad}(A)\,. \tag{2.9}$$

Here $\mathrm{ad}(A)$ is the *adjoint* of the matrix A. Its (i, j)-entry is $(-1)^{i+j} \cdot \det A_{i,j}$, where $A_{i,j}$ is the $(n-1) \times (n-1)$ matrix obtained from A by deleting its ith column and jth row.

Since $\det(L_{f,g}) = \mathrm{Res}(f, g) \in \mathbb{K}$, we have

$$L_{f,g}^{-1}(\mathrm{Res}(f, g)) \ = \ \det(L_{f,g}) \cdot L_{f,g}^{-1}(1) \ = \ \mathrm{ad}(\mathrm{Syl}(f, g))(1)\,.$$

In the monomial basis of $\mathbb{K}_{m+n-1}[x]$ the polynomial 1 is the vector $(0, \ldots, 0, 1)^T$. Thus, the coefficients of $L_{f,g}^{-1}(\mathrm{Res}(f, g))$ are the entries of the last column of $\mathrm{ad}(\mathrm{Syl}(f, g))$, which are $\pm$ the minors of the Sylvester matrix $\mathrm{Syl}(f, g)$ with its last row removed. In particular, these are integer polynomials in the variables $f_0, \ldots, g_n$. ∎

This proof shows that $h, k \in \mathbb{Z}[f_0, \ldots, f_m, g_0, \ldots, g_m][x]$ and that (2.8) holds as an expression in this polynomial ring with $m+n+3$ variables. It leads to a method to eliminate variables. Suppose that $f, g \in \mathbb{K}[x_1, \ldots, x_n]$ are multivariate polynomials. We may consider them as polynomials in the variable x_n whose coefficients are polynomials in the other variables, that is as polynomials in $\mathbb{K}(x_1, \ldots, x_{n-1})[x_n]$. Then the resultant $\mathrm{Res}(f, g; x_n)$ both lies in the ideal generated by f and g and in the subring $\mathbb{K}[x_1, \ldots, x_{n-1}]$. We examine the geometry of this elimination of variables.

Suppose that $1 \leq m < n$ and let $\pi\colon \mathbb{K}^n \to \mathbb{K}^m$ be the coordinate projection

$$\pi \;:\; (a_1, \ldots, a_n) \;\longmapsto\; (a_1, \ldots, a_m)\,.$$

Also, for $I \subset \mathbb{K}[x_1, \ldots, x_n]$ set $I_m := I \cap \mathbb{K}[x_1, \ldots, x_m]$.

Lemma 2.3.8. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal. Then $\pi(\mathcal{V}(I)) \subset \mathcal{V}(I_m)$. When $\mathbb{K}$ is algebraically closed $\mathcal{V}(I_m)$ is the smallest variety in $\mathbb{K}^m$ containing $\pi(\mathcal{V}(I))$.*

Proof. Let us set $X := \mathcal{V}(I)$. For the first statement, suppose that $a = (a_1, \ldots, a_n) \in X$. If $f \in I_m = I \cap \mathbb{K}[x_1, \ldots, x_m]$, then

$$0 \;=\; f(a) \;=\; f(a_1, \ldots, a_m) \;=\; f(\pi(a))\,,$$

which establishes the inclusion $\pi(X) \subset \mathcal{V}(I_m)$. (For this we view f as a polynomial in either $x_1, \ldots, x_n$ or in $x_1, \ldots, x_m$.) This implies that $\mathcal{V}(\mathcal{I}(\pi(X))) \subset \mathcal{V}(I_m)$.

Now suppose that $\mathbb{K}$ is algebraically closed. Let $f \in \mathcal{I}(\pi(X))$. Then $f \in \mathbb{K}[x_1, \ldots, x_m]$ has the property that $f(a_1, \ldots, a_m) = 0$ for all $(a_1, \ldots, a_m) \in \pi(X)$. Viewing f as an element of $\mathbb{K}[x_1, \ldots, x_n]$ shows that f vanishes on $X = \mathcal{V}(I)$.

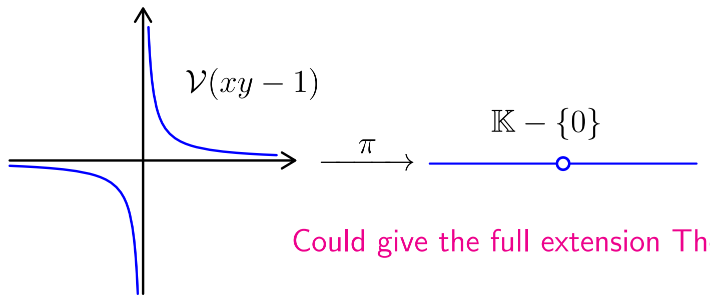
By the Nullstellensatz, there is a positive integer N such that $f^N \in I$ (as elements of the ring $K[x_1, \ldots, x_n]$). But then $f^N \in I \cap \mathbb{K}[x_1, \ldots, x_m]$, which implies that $f \in \sqrt{I_m}$. Thus $\mathcal{I}(\pi(X)) \subset \sqrt{I_m}$, so that

$$\mathcal{V}(\mathcal{I}(\pi(X))) \;\supset\; \mathcal{V}(\sqrt{I_m}) \;=\; \mathcal{V}(I_m)\,,$$

which completes the proof. □

The ideal $I_m = I \cap \mathbb{K}[x_1, \ldots, x_m]$ is called an *elimination ideal* as the variables $x_{m+1}, \ldots, x_n$ have been eliminated from the ideal I. By Lemma 2.3.8, elimination is the algebraic counterpart to projection, but the correspondence is not exact. For example, the inclusion $\pi(\mathcal{V}(I)) \subset \mathcal{V}(I \cap \mathbb{K}[x_1, \ldots, x_m])$ may be strict. Let $\pi\colon \mathbb{K}^2 \to \mathbb{K}$ be the map which forgets the second coordinate. Then $\pi(\mathcal{V}(xy-1)) = \mathbb{K} - \{0\} \subsetneq \mathbb{K} = V(0)$ and

$\{0\} = \langle xy-1\rangle \cap F[x]$.



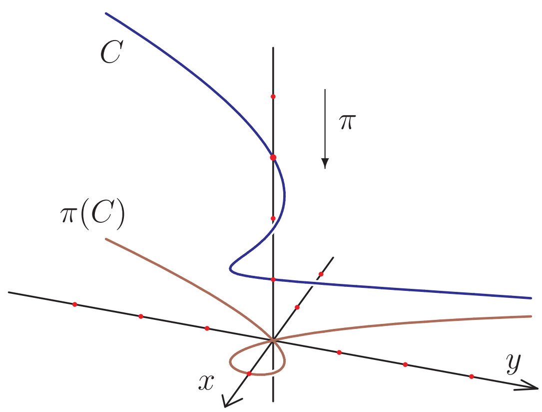
Could give the full extension Theorem?

The missing point, $\{0\}$ corresponds to the coefficient x of the highest power of y.

We may solve the implicitization problem for plane curves using elimination. For example, consider the parametric plane curve

$$x \;=\; t^2-1, \qquad y \;=\; t^3-t\,. \tag{2.10}$$

This is the image of the space curve $C := \mathcal{V}(t^2-1-x, t^3-t-y)$ under the projection $(x,y,t) \mapsto (x,y)$. We display this with the t-axis vertical and the xy-plane at $t=-2$.



By Lemma 2.3.8, the plane curve is defined by $\langle t^2-x-x, t^3-t-y\rangle \cap \mathbb{K}[x,y]$. If we set

$$f(t) \;:=\; t^2-1-x \qquad \text{and} \qquad g(t) \;:=\; t^3-t-y\,,$$

then the Sylvester resultant $\mathrm{Res}(f,g;t)$ is

$$\det\left(\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0 & 1\\ -x-1 & 0 & 1 & -1 & 0\\ 0 & -x-1 & 0 & -y & -1\\ 0 & 0 & -x-1 & 0 & -y \end{array}\right) \;=\; y^2+x^2-x^3\,,$$

which is the implicit equation of the parameterized cubic $\pi(C)$ (2.10).

We use resultants to study the variety $\mathcal{V}(f,g) \subset \mathbb{K}^2$ for $f,g \in \mathbb{K}[x,y]$. A by-product will be a form of Bézout's Theorem bounding the number of points in the variety $\mathcal{V}(f,g)$.

The ring $\mathbb{K}[x,y]$ of bivariate polynomials is a subring of the ring $\mathbb{K}(x)[y]$ of polynomials in y whose coefficients are rational functions in x. Suppose that $f,g \in \mathbb{K}[x,y]$. Considering f and g as elements of $\mathbb{K}(x)[y]$, the resultant $\mathrm{Res}(f,g;y)$ is the determinant of their Sylvester matrix expressed in the basis of monomials in y. By Theorem 2.3.2, $\mathrm{Res}(f,g;y)$ is a univariate polynomial in x which vanishes if and only if f and g have a common factor in $\mathbb{K}(x)[y]$. In fact it vanishes if and only if $f(x,y)$ and $g(x,y)$ have a common factor in $\mathbb{K}[x,y]$ with positive degree in y, by the following version of Gauss's lemma for $\mathbb{K}[x,y]$.

Lemma 2.3.9. *Polynomials f and g in $\mathbb{K}[x,y]$ have a common factor of positive degree in y if and only if they have a common factor in $\mathbb{K}(x)[y]$.*

Proof. The forward direction is clear. For the reverse, suppose that

$$f \;=\; h\cdot\overline{f} \qquad\text{and}\qquad g \;=\; h\cdot\overline{g} \tag{2.11}$$

is a factorization in $\mathbb{K}(x)[y]$ where h has positive degree in y.

There is a polynomial $d \in \mathbb{K}[x]$ which is divisible by every denominator of a coefficient of h, $\overline{f}$, and $\overline{g}$. Multiplying the expressions (2.11) by d^2 gives

$$d^2 f \;=\; (dh)\cdot(d\overline{f}) \qquad\text{and}\qquad d^2 g \;=\; (dh)\cdot(d\overline{g})\,,$$

where dh, $d\overline{f}$, and $d\overline{g}$ are polynomials in $\mathbb{K}[x,y]$. Let $p(x,y) \in \mathbb{K}[x,y]$ be an irreducible polynomial factor of dh having positive degree in y. Then p divides both d^2f and d^2g. However, p cannot divide d as $d \in \mathbb{K}[x]$ and p has positive degree in y. Therefore $p(x,y)$ is the desired common polynomial factor of f and g. $\square$

Let $\pi\colon \mathbb{K}^2 \to \mathbb{K}$ be the projection which forgets the last coordinate, $\pi(x,y) = x$. Set $I := \langle f,g\rangle \cap \mathbb{K}[x]$. By Lemma 2.3.7, the resultant $\mathrm{Res}(f,g;y)$ lies in I. Combining this with Lemma 2.3.8 gives the chain of inclusions

$$\pi(\mathcal{V}(f,g)) \;\subset\; \mathcal{V}(I) \;\subset\; \mathcal{V}(\mathrm{Res}(f,g;y))\,,$$

with the first inclusion an equality if $\mathbb{K}$ is algebraically closed and $\pi(\mathcal{V}(f,g))$ is a variety, which occurs, for example, when $\mathcal{V}(f,g)$ is a finite set.[†].

We now suppose that $\mathbb{K}$ is algebraically closed. Let $f,g \in \mathbb{K}[x,y]$ and write them as polynomials in y with coefficients in $\mathbb{K}[x]$,

$$\begin{aligned} f &= f_0(x)y^m + f_1(x)y^{m-1} + \cdots + f_{m-1}(x)y + f_m(x)\\ g &= g_0(x)y^n + g_1(x)y^{n-1} + \cdots + g_{n-1}(x)y + g_n(x)\,, \end{aligned}$$

where neither $f_0(x)$ nor $g_0(x)$ is the zero polynomial.

[†]Where is this proven?

Theorem 2.3.10 (Extension Theorem). *If $a \in \mathcal{V}(I) \smallsetminus \mathcal{V}(f_0(x), g_0(x))$, then there is some $b \in \mathbb{K}$ with $(a, b) \in \mathcal{V}(f, g)$.*

This establishes the chain of inclusions of subvarieties of $\mathbb{K}$

$$\mathcal{V}(I) \smallsetminus \mathcal{V}(f_0, g_0) \ \subset\ \pi(\mathcal{V}(f,g)) \ \subset\ \mathcal{V}(I) \subset\ \mathcal{V}(\mathrm{Res}(f,g;x))\,.$$

Observe that if either of f_0 and g_0 are constants, or if $\gcd(f,g) = 1$, then $\mathcal{V}(I) = \mathcal{V}(\mathrm{Res}(f,g;y)$.

Parts of this treatment of elimination and extension would be better to do in generality, rather in two variables.

Proof. Let $a \in \mathcal{V}(I) \smallsetminus \mathcal{V}(f_0, g_0)$. Suppose first that $f_0(a) \cdot g_0(a) \neq 0$. Then $f(a,y)$ and $g(a,y)$ are polynomials in y of degrees m and n, respectively. It follows that the Sylvester matrix $\mathrm{Syl}(f(a,y), g(a,y))$ has the same format (2.6) as the Sylvester matrix $\mathrm{Syl}(f,g;y)$, and it is in fact obtained from $\mathrm{Syl}(f,g;y)$ by the substitution $x = a$.

This implies that $\mathrm{Res}(f(a,y), g(a,y))$ is the evaluation of the resultant $\mathrm{Res}(f,g;y)$ at $x = a$. Since $\mathrm{Res}(f,g;y) \in I$ and $a \in \mathcal{V}(I)$, this evaluation is 0. By Theorem 2.3.2, $f(a,y)$ and $g(a,y)$ have a nonconstant common factor. As $\mathbb{K}$ is algebraically closed, they have a common root, say b. But then $(a,b) \in \mathcal{V}(f,g)$, and so $a \in \pi(\mathcal{V}(f,g))$.

Now suppose that $f_0(a) \neq 0$ but $g_0(a) = 0$. Since $\langle f, g\rangle = \langle f, g + y^\ell f\rangle$, if we replace g by $g + y^\ell f$ where $\ell + m > n$, then we are in the previous case. $\square$

Example 2.3.11. Suppose that $f, g \in \mathbb{C}[x,y]$ are the polynomials,

$$\begin{aligned} f &= (5 - 10x + 5x^2)y^2 \ +\ (-14 + 42x - 24x^2)y \ +\ (5 - 28x + 19x^2) \\ g &= (5 - 10x + 5x^2)y^2 \ +\ (-16 + 46x - 26x^2)y \ +\ (19 - 36x + 21x^2) \end{aligned}$$

Figure 2.2 shows the curves $\mathcal{V}(f)$ and $\mathcal{V}(G)$, which meet in three points,

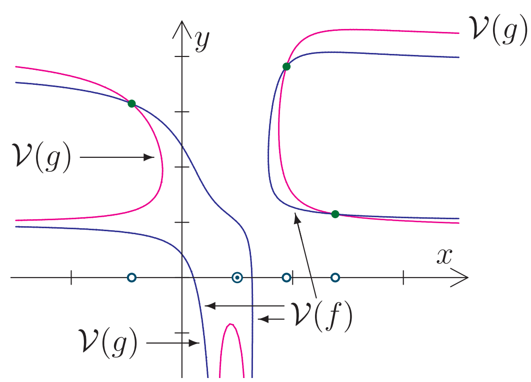


Figure 2.2: Comparing resultants to elimination.

$$\mathcal{V}(f,g) \ =\ \{\,(-0.9081601, 3.146707)\,,\ (1.888332, 3.817437)\,,\ (2.769828, 1.146967)\,\}\,.$$

Thus $\pi(\mathcal{V}(f,g))$ consists of three points which are roots of $h = 4x^3 - 15x^2 + 4x + 19$, where $\langle h\rangle = \langle f,g\rangle \cap \mathbb{K}[x]$. However, the resultant is

$$\mathrm{Res}(f,g;y) \;=\; 160(4x^3 - 15x^2 + 4x + 19)(x-1)^4\,,$$

whose roots are shown on the x-axis, including the point $x = 1$ with multiplicity four. $\diamond$

Corollary 2.3.12. *If the coefficients of the highest powers of y in f and g do not involve x, then $\mathcal{V}(I) = \mathcal{V}(\mathrm{Res}(f,g;x))$.* Not true if $\gcd(f,g) \neq 1$.

Lemma 2.3.13. *When $\mathbb{K}$ is algebraically closed, the system of bivariate polynomials*

$$f(x,y) \;=\; g(x,y) \;=\; 0$$

has finitely many solutions in $\mathbb{K}^2$ if and only if f and g have no common factor.

Proof. We instead show that $\mathcal{V}(f,g)$ is infinite if and only if f and g do have a common factor. If f and g have a common factor $h(x,y)$ then their common zeroes $\mathcal{V}(f,g)$ include $\mathcal{V}(h)$ which is infinite as h is nonconstant and $\mathbb{K}$ is algebraically closed. We need to prove this in Chapter 1

Now suppose that $\mathcal{V}(f,g)$ is infinite. Then its projection to at least one of the two coordinate axes is infinite. Suppose that the projection π onto the x-axis is infinite. Set $I := \langle f,g\rangle \cap \mathbb{K}[x]$, the elimination ideal. By the Extension Theorem 2.3.10, we have $\pi(\mathcal{V}(f,g)) \subset \mathcal{V}(I) \subset \mathcal{V}(\mathrm{Res}(f,g;y))$. Since $\pi(\mathcal{V}(f,g))$ is infinite, $\mathcal{V}(\mathrm{Res}(f,g;y)) = \mathbb{K}$, which implies that $\mathrm{Res}(f,g;y)$ is the zero polynomial. By Theorem 2.3.2 and Lemma 2.3.9, f and g have a common factor. $\square$

Let $f,g \in \mathbb{K}[x,y]$ and suppose that neither $\mathrm{Res}(f,g;x)$ nor $\mathrm{Res}(f,g;y)$ vanishes so that f and g have no common factor. Then $\mathcal{V}(f,g)$ consists of finitely many points. The Extension Theorem gives the following algorithm to compute $\mathcal{V}(f,g)$.

Algorithm 2.3.14 (Elimination Algorithm)**.** INPUT: Polynomials $f,g \in \mathbb{K}[x,y]$ with $\gcd(f,g) = 1$.
OUTPUT: $\mathcal{V}(f,g)$.

First, compute the resultant $\mathrm{Res}(f,g;x)$, which is not the zero polynomial. Then, for every root a of $\mathrm{Res}(f,g;y)$, find all common roots b of $f(a,y)$ and $g(a,y)$. The finitely many pairs (a,b) computed are the points of $\mathcal{V}(f,g)$.

The Elimination Algorithm reduces the problem of solving a bivariate system

$$f(x,y) \;=\; g(x,y) \;=\; 0\,, \tag{2.12}$$

to that of finding the roots of univariate polynomials.

Often we only want to count the number of solutions to a system (2.12), or give a realistic bound for this number which is attained when f and g are generic polynomials. The most basic of such bounds was given by Etienne Bézout in 1779. Our first step toward establishing Bézout's Theorem is an exercise in algebra and some book-keeping. The monomials in a polynomial of degree n in the variables x, y are indexed by the set

$$n\Delta \;:=\; \{(i,j) \in \mathbb{N}^2 \mid i+j \leq n\}\,.$$

Let $F := \{f_{i,j} \mid (i,j) \in m\Delta\}$ and $G := \{g_{i,j} \mid (i,j) \in n\Delta\}$ be variables and consider generic polynomials f and g of respective degrees m and n in $\mathbb{K}[F,G][x,y]$,

$$f(x,y) \;:=\; \sum_{(i,j)\in m\Delta} f_{i,j}x^iy^j \qquad \text{and} \qquad g(x,y) \;:=\; \sum_{(i,j)\in n\Delta} g_{i,j}x^iy^j\;.$$

Lemma 2.3.15. *The generic resultant* $\mathrm{Res}(f,g;y)$ *is a polynomial in* x *of degree* mn.

Proof. Write

$$f \;:=\; \sum_{j=0}^{m} f_j(x)y^{m-j} \qquad \text{and} \qquad g \;:=\; \sum_{j=0}^{n} g_j(x)y^{n-j}\,,$$

where the coefficients are univariate polynomials in x,

$$f_j(x) \;:=\; \sum_{i=0}^{j} f_{i,j}x^i \qquad \text{and} \qquad g_j(x) \;:=\; \sum_{i=0}^{j} g_{i,j}x^i\,.$$

Then the Sylvester matrix $\mathrm{Syl}(f,g;y)$ has the form

$$\mathrm{Syl}(f,g;y) \;:=\; \left(\begin{array}{cccc|cccc}
f_0(x) & & 0 & & g_0(x) & & 0 & \\
\vdots & \ddots & & & \vdots & \ddots & & \\
\vdots & & \ddots & & \vdots & & g_0(x) & \\
f_m(x) & & & f_0(x) & g_{n-1}(x) & & \vdots & \\
 & \ddots & & \vdots & g_n(x) & & \vdots & \\
 & & \ddots & \vdots & & \ddots & \vdots & \\
0 & & & f_m(x) & 0 & & g_n(x) &
\end{array}\right)\,,$$

and so the resultant $\mathrm{Res}(f,g;y) = \det(\mathrm{Syl}(f,g;y))$ is a univariate polynomial in x.

As in the proof of Lemma 2.3.3, if we set $f_j := 0$ when $j<0$ or $j>m$ and $g_j := 0$ when $j<0$ or $j>n$, then the entry in row i and column j of the Sylvester matrix is

$$\mathrm{Syl}(f,g;y)_{i,j} \;=\; \begin{cases} f_{i-j}(x) & \text{if } j \leq n \\ g_{n+i-j}(x) & \text{if } n < j \leq m+n \end{cases}$$

The determinant is a signed sum over permutations w of $\{1,\ldots,m{+}n\}$ of terms

$$\prod_{j=1}^{n} f_{w(j)-j}(x) \;\cdot\; \prod_{j=n+1}^{m+n} g_{n+w(j)-j}(x)\,.$$

This is a polynomial of degree at most

$$\sum_{j=1}^{m} w(j){-}j \quad + \quad \sum_{j=n+1}^{m+n} n{+}w(j){-}j \;=\; mn + \sum_{j=1}^{m+n} w(j){-}j \;=\; mn\,.$$

Thus $\mathrm{Res}(f,g;y)$ is a polynomial of degree at most mn.

We complete the proof by showing that the resultant does indeed have degree mn. The product $f_0(x)^n \cdot g_n(x)^m$ of the entries along the main diagonal of the Sylvester matrix has constant term $f_{0,0}^n \cdot g_{0,n}^m$, and the coefficient of x^{mn} in this product is $f_{0,0}^n \cdot g_{n,n}^m$, and these are the only terms in the expansion of the determinant of the Sylvester matrix involving either of these monomials in the coefficients $f_{i,j}, g_{k,l}$. □

We now state and prove Bézout's Theorem, which bounds the number of points in the variety $\mathcal{V}(f,g)$ in $\mathbb{K}^2$.

Theorem 2.3.16 (Bézout's Theorem)**.** *Two polynomials $f,g \in \mathbb{K}[x,y]$ either have a common factor or else $|\mathcal{V}(f,g)| \leq \deg(f)\cdot\deg(g)$.*

When $|\mathbb{K}|$ is at least $\max\{\deg(f),\deg(g)\}$, this inequality is sharp in that the bound is attained. When $\mathbb{K}$ is algebraically closed, the bound is attained when f and g are general polynomials of the given degrees.

Proof. Suppose that $m := \deg(f)$ and $n = \deg(g)$. By Lemma 2.3.13, if f and g are relatively prime, then $\mathcal{V}(f,g)$ is finite. Let us extend $\mathbb{K}$ to its algebraic closure $\overline{\mathbb{K}}$, which in infinite. Changing coordinates, replacing f by $f(A(x,y))$ and g by $g(A(x,y))$, where A is an invertible affine transformation,

$$A(x,y) \;=\; (ax+by+c,\ \alpha x+\beta y+\gamma)\,, \tag{2.13}$$

with $a,b,c,\alpha,\beta,\gamma \in \overline{\mathbb{K}}$ with $a\beta-\alpha b \neq 0$. As $\overline{\mathbb{K}}$ is infinite, we can choose these parameters so that the constant terms and terms with highest power of x in each of f and g are nonzero. By Lemma 2.3.15, this implies that the resultant $\mathrm{Res}(f,g;y)$ has degree at most mn and thus at most mn zeroes. If we set $I := \langle f,g\rangle \cap \overline{\mathbb{K}}[x]$, then this also implies that $\mathcal{V}(I) = \mathcal{V}(\mathrm{Res}(f,g;x))$, by Corollary 2.3.12.

We can furthermore choose the parameters in A so that the projection $\pi\colon (x,y) \mapsto x$ is 1-1 on $\mathcal{V}(f,g)$, as $\mathcal{V}(f,g)$ is a finite set. Thus

$$\pi(\mathcal{V}(f,g)) \;=\; \mathcal{V}(I) \;=\; \mathcal{V}(\mathrm{Res}(f,g;x))\,,$$

which implies the inequality of the theorem as $|\mathcal{V}(\mathrm{Res}(f,g;y))| \leq mn$.

To see that the bound is sharp when $|\mathbb{K}|$ is large enough, let $a_1,\ldots,a_m$ and $b_1,\ldots,b_n$ be distinct elements of $\mathbb{K}$. Note that the system

$$f \;:=\; \prod_{i=1}^{m}(x-a_i) \;=\; 0 \qquad \text{and} \qquad g \;:=\; \prod_{i=1}^{n}(y-b_i) \;=\; 0 \tag{2.14}$$

has mn solutions $\{(a_i, b_j) \mid 1 \leq i \leq m, 1 \leq j \leq n\}$, so the inequality is sharp.

Suppose now that $\mathbb{K}$ is algebraically closed. If the resultant $\mathrm{Res}(f, g; y)$ has fewer than mn distinct roots, then either it has degree strictly less than mn or else it has a multiple root. In the first case, its leading coefficient vanishes and in the second case, its discriminant vanishes. But the leading coefficient and the discriminant of $\mathrm{Res}(f, g; y)$ are polynomials in the $\binom{m+2}{2} + \binom{n+2}{2}$ coefficients of f and g. Neither is the zero polynomial, as they do not vanish when evaluated at the coefficients of the polynomials (2.14). Thus the set of pairs of polynomials (f, g) with $\mathcal{V}(f, g)$ consisting of mn points in $\mathbb{K}^2$ is a nonempty generic set in $\mathbb{K}^{\binom{m+2}{2}+\binom{n+2}{2}}$. $\square$

Exercises

1. Give some finger exercises related to solving using resultants.

2. Using the formula (2.7) deduce the Poisson formula for the resultant of univariate polynomials f and g,

$$\mathrm{Res}(f, g; x) \;=\; (-1)^{mn} f_0^n \prod_{i=1}^{m} g(\alpha_i)\,,$$

   where $\alpha_1, \ldots, \alpha_m$ are the roots of f.

3. Suppose that the polynomial $g = g_1 \cdot g_2$ factors. Show that the resultant also factors, $\mathrm{Res}(f, g; x) = \mathrm{Res}(f, g_1; x) \cdot \mathrm{Res}(f, g_2; x)$.

4. Compute the Bezoutian matrix when $n = 4$. Give a general formula for the entries of the Bezoutian matrix.

5. Compute the constant in the proof of Theorem 2.3.5 by computing the resultant and Bezoutian polynomials when $f(x) := x^m$ and $g(x) = x^n + 1$. Why does this computation suffice?

6. Compute the discriminant of a general cubic $x^3 + ax^2 + bx + c$ by taking the determinant of a 5×5 matrix. Show that the discriminant of the depressed quartic $x^4 + ax^2 + bx + c$ is

$$16a^4c - 4a^3b^2 - 128a^2c^2 + 144ab^2c - 27b^4 + 256c^3\,.$$

7. Show that the discriminant of a polynomial f of degree n may also be expressed as

$$\prod_{i \neq j} (\alpha_i - \alpha_j)^2\,,$$

   where $\alpha_1, \ldots, \alpha_n$ are the roots of f.

2.4 Solving equations with Gröbner bases

Algorithm 2.3.14 reduced the problem of solving two equations in two variables to that of solving univariate polynomials, using resultants to eliminate a variable. For an ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ whose variety $\mathcal{V}(I)$ consists of finitely many points, this same idea leads to an algorithm to compute $\mathcal{V}(I)$, provided we can compute the elimination ideals $I \cap \mathbb{K}[x_i, x_1, \ldots, x_m]$. Gröbner bases provide a universal algorithm for computing elimination ideals. More generally, ideas from the theory of Gröbner bases can help to understand solutions to systems of equations.

Suppose that we have N polynomial equations in n variables $(x_1, \ldots, x_n)$

$$f_1(x_1, \ldots, x_n) \;=\; \cdots \;=\; f_N(x_1, \ldots, x_n) \;=\; 0\,, \tag{2.15}$$

and we want to understand the solutions to this system. By understand, we mean answering (any of) the following questions.

(i) Does (2.15) have finitely many solutions?

(ii) If not, can we understand the isolated solutions of (2.15)?

(iii) Can we count them, or give (good) upper bounds on their number?

(iv) Can we *solve* the system (2.15) and find all complex solutions?

(v) When the polynomials have real coefficients, can we count (or bound) the number of real solutions to (2.15)? Or simply find them?

We describe symbolic algorithms based upon Gröbner bases that begin to address these questions.

The solutions to (2.15) in $\mathbb{K}^n$ constitute the affine variety $\mathcal{V}(I)$, where I is the ideal generated by the polynomials $f_1, \ldots, f_N$. Algorithms based on Gröbner bases to address Questions (i)-(v) involve studying the ideal I. An ideal I is *zero-dimensional* if, over the algebraic closure of $\mathbb{K}$, $\mathcal{V}(I)$ is finite. Thus I is zero-dimensional if and only if its radical $\sqrt{I}$ is zero-dimensional.

Theorem 2.4.1. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal. Then I is zero-dimensional if and only if $\mathbb{K}[x_1, \ldots, x_n]/I$ is a finite-dimensional $\mathbb{K}$-vector space, if and only if $\mathcal{V}(I)$ is a finite set in $\overline{K}^n$.*

When an ideal I is zero-dimensional, we will call the points of $\mathcal{V}(I)$ the *roots of I*.

Proof. We may assume the $\mathbb{K}$ is algebraically closed, as this does not change the dimension of quotient rings.

Suppose first that I is radical, so that $I = \mathcal{I}(\mathcal{V}(I))$, by the Nullstennensatz. Then $\mathbb{K}[x_1, \ldots, x_n]/I$ is the coordinate ring $\mathbb{K}[X]$ of $X := \mathcal{V}(I)$, consisting consists of all functions obtained by restricting polynomials to $\mathcal{V}(I)$, and is therefore a subring of the ring of functions on X. If X is finite, then $\mathbb{K}[X]$ is finite-dimensional as the space of functions on X has dimension equal to the number of points in X. Suppose that X is infinite. Then there is some coordinate, say x_1, such that the projection of X to the x_1-axis is infinite. In particular, no polynomial in x_1, except the zero polynomial, vanishes on X. † Restriction of polynomials in x_1 to X is therefore an injective map from $\mathbb{K}[x_1]$ to $\mathbb{K}[X]$ which shows that $\mathbb{K}[X]$ is infinite-dimensional.

Now suppose that I is any ideal. If $\mathbb{K}[x_1,\dots,x_n]/I$ is finite-dimensional, then so is $\mathbb{K}[x_1,\dots,x_n]/\sqrt{I}$ as $I \subset \sqrt{I}$. For the other direction, we suppose that $\mathbb{K}[x_1,\dots,x_n]/\sqrt{I}$ is finite-dimensional. For each variable x_i, there is some linear combination of $1, x_i, x_i^2, \dots$ which is zero in $\mathbb{K}[x_1,\dots,x_n]/\sqrt{I}$ and hence lies in $\sqrt{I}$. But this is a univariate polynomial $g_i(x_i) \in \sqrt{I}$, so there is some power $g_i(x_i)^{M_i}$ of g_i which lies in I. But then we have $\langle g_1(x_1)^{M_1},\dots,g_n(x_n)^{M_n}\rangle \subset I$, and so the map

$$\mathbb{K}[x_1,\dots,x_n]/\langle g_1(x_1)^{M_1},\dots,g_n(x_n)^{M_n}\rangle \quad\longrightarrow\quad \mathbb{K}[x_1,\dots,x_n]/I$$

is a surjection. But $\mathbb{K}[x_1,\dots,x_n]/\langle g_1(x_1)^{M_1},\dots,g_n(x_n)^{M_n}\rangle$ has dimension $\prod_i M_i \deg(g_i)$, which implies that $\mathbb{K}[x_1,\dots,x_n]/I$ is finite-dimensional. □

A consequence of this proof is the following criterion for an ideal to be zero-dimensional.

Corollary 2.4.2. *An ideal $I \subset \mathbb{K}[x_1,\dots,x_n]$ is zero-dimensional if and only if for every variable x_i, there is a univariate polynomial $g_i(x_i)$ which lies in I.*

Together with Macaulay's Theorem 2.2.3, Theorem 2.4.1 leads to a Gröbner basis criterion/algorithm to solve Question (i).

Corollary 2.4.3. *An ideal $I \subset \mathbb{K}[x_1,\dots,x_n]$ is zero-dimensional if and only if for any monomial order $\succ$, the initial ideal $\mathrm{in}_\succ I$ of I contains some power of every variable.*

Thus we can determine if I is zero-dimensional and thereby answer Question (i) by computing a Gröbner basis for I and checking that the leading terms of elements of the Gröbner basis include pure powers of all variables.

When I is zero-dimensional, its degree is the dimension of $\mathbb{K}[x_1,\dots,x_n]/I$ as a $\mathbb{K}$-vector space, which is the number of standard monomials, by Macaulay's Theorem 2.2.3. A Gröbner basis for I gives generators of the initial ideal which we can use to count the number of standard monomials to determine the degree of an ideal.

When I is a zero-dimensional radical ideal and $\mathbb{K}$ is algebraically closed, the degree of I equals the number of points in $\mathcal{V}(I) \subset \mathbb{K}^n$ (see Exercise 3) and thus we obtain an answer to Question (iii).

Theorem 2.4.4. *Let I be the ideal generated by the polynomials f_i of* (2.15)*. If I is zero-dimensional, then the number of solutions to the system* (2.15) *is bounded by the degree of I. When $\mathbb{K}$ is algebraically closed, the number of solutions is equal to this degree if and only if I is radical.*

---

† Make sure the injectivity of restricting polynomial functions to $\mathbb{K}^n$ is mentioned in Chapter 1.

In many important cases, there are sharp upper bounds for the number of isolated solutions to the system (2.15) which do not require a Gröbner basis. For example, Theorem 2.3.16 (Bézout's Theorem in the plane) gives such bounds when $N = n = 2$. Suppose that $N = n$ so that the number of equations equals the number of variables. This is called a *square system*. Bézout's Theorem in the plane has a natural extension in this case, which we will prove in Section 3.6. A common solution x to a square system of equations is *nondegenerate* if the differentials of the equations are linearly independent at x.

Theorem 2.4.5 (Bézout's Theorem)**.** *Given polynomials $f_1, \ldots, f_n \in \mathbb{K}[x_1, \ldots, x_n]$ with $d_i = \deg(f_i)$, the number of nondegenerate solutions to the system*

$$f_1(x_1, \ldots, x_n) \;=\; \cdots \;=\; f_n(x_1, \ldots, x_n) \;=\; 0$$

in $\mathbb{K}^n$ is at most $d_1 \cdots d_n$. When $\mathbb{K}$ is algebraically closed, this is a bound for the number of isolated solutions, and it is attained for generic choices of the polynomials f_i.

This product of degrees $d_1 \cdots d_n$ is called the *Bézout bound* for such a system. While this bound is sharp for generic square systems, few practical problems involve generic systems and other bounds are often needed (see Exercise 4). We discuss such bounds in Chapter 9, where we establish the polyhedral bounds of Kushnirenko's and Bernsteins's Theorems.

We discuss a symbolic method to solve systems of polynomial equations (2.15) based upon elimination theory and the Shape Lemma, which describes the form of a Gröbner basis of a zero-dimensional ideal I with respect to a lexicographic monomial order. Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal. A univariate polynomial $g(x_i)$ is an *eliminant for I* if g generates the elimination ideal $I \cap \mathbb{K}[x_i]$.

Theorem 2.4.6. *Suppose that $g(x_i)$ is an eliminant for an ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$. Then $g(a_i) = 0$ for every $a = (a_1, \ldots, a_n) \in \mathcal{V}(I) \in \mathbb{K}^n$. When $\mathbb{K}$ is algebraically closed, every root of g occurs in this way.*

Proof. We have $g(a_i) = 0$ as this is the value of g at the point a. Suppose that $\mathbb{K}$ is algebraically closed and that ξ is a root of $g(x_i)$ but there is no point $a \in \mathcal{V}(I)$ whose ith coordinate is ξ. Let $h(x_i)$ be a polynomial whose roots are the other roots of g. Then h vanishes on $\mathcal{V}(I)$ and so $h \in \sqrt{I}$. But then some power, h^N, of h lies in I. Thus $h^N \in I \cap \mathbb{K}[x_i] = \langle g \rangle$. But this is a contradiction as $h(\xi) \neq 0$ while $g(\xi) = 0$. $\square$

Theorem 2.4.7. *If $g(x_i)$ is a monic eliminant for an ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$, then g lies in the reduced Gröbner basis for I with respect to any monomial order in which the pure powers x_i^m of x_i preceed variables x_j with $j \neq i$.*

Proof. Suppose that $\succ$ is such a monomial order. Then its minimal monomials are $1, x_i, x_i^2, \ldots$. Since g generates the elimination ideal $I \cap \mathbb{K}[x_i]$, it is the lowest degree monic polynomial in x_i lying in I. As $g \in I$, we have that $x_i^{\deg(g)} \in \mathrm{in}_\prec(I)$. Let x_i^m be the generator of $\text{in}_\prec(I) \cap \mathbb{K}[x_i]$. Then $m \leq \deg(g)$. Let f be the polynomial in the reduced Gröbner basis of I with respect to $\prec$ whose leading term is x_i^m. Then its remaining terms involve smaller standard monomials and are thus pure powers of x_i. We conclude that $f \in I \cap \mathbb{K}[x_i] = \langle g \rangle$, and so g divides f, so $m = \deg(g)$. As $f{-}g$ is a polynomial in x_i which lies in I but has degree less than $\deg(g)$, the minimality of f and g implies that $f{-}g = 0$. This proves that g lies in the reduced Gröbner basis. □

You will prove the following theorem relating Gröbner bases and elimination ideals in the exercises.

Theorem 2.4.8. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal and let $\prec$ be the lexicographic monomial order with $x_1 \prec x_2 \prec \cdots \prec x_n$ and G a Gröbner basis for I with respect to $\prec$. Then, for each $m = 1, \ldots, n$, the polynomials in G that lie in $\mathbb{K}[x_1, \ldots, x_m]$ form a Gröbner basis for the elimination ideal $I_m = I \cap \mathbb{K}[x_1, \ldots, x_n]$.*

Theorem 2.4.7 gives an algorithm to compute eliminants—simply compute a lexicographic Gröbner basis. This is not recommended, as lexicographic Gröbner bases appear to be the most expensive to compute. If we only need to compute a univariate eliminant $g(x_i)$, we may use an elimination order, which is a monomial order $\prec$ where any pure power x_i^d of x_i is smaller than any monomial involving any other variable x_j for $j \neq i$. A Gröbner basis with respect to some elimination order is still expensive to compute. We instead offer the following algorithm.

Algorithm 2.4.9.
INPUT: Ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ and a variable x_i.
OUTPUT: Either a univariate eliminant $g(x_i) \in I$ or else a certificate that one does not exist.

(1) Compute a Gröbner basis G for I with respect to any monomial order.

(2) If no initial term of any element of G is a pure power of x_i, then halt and declare that I does not contain a univariate eliminant for x_i.

(3) Otherwise, compute the sequence $1 \bmod G$, $x_i \bmod G$, $x_i^2 \bmod G$, …, until a linear dependence is found,

$$\sum_{j=0}^{m} a_j (x_i^j \bmod G) \;=\; 0\,, \tag{2.16}$$

where m is minimal. Then

$$g(x_i) \;=\; \sum_{j=0}^{m} a_j x_i^j$$

is a univariate eliminant.

Proof of correctness. If I does not have an eliminant in x_i, then $I \cap \mathbb{K}[x_i] = \{0\}$. Then all monomials $1, x_i, x_i^2, \ldots$ are standard and no Gröbner basis contains a polynomial with initial monomial a pure power of x_i. This shows that the algorithm correctly identifies when no eliminant exists.

Suppose now that I does have an eliminant $g(x_i)$. Since $g \bmod G = 0$, the Gröbner basis G must contain a polynomial whose initial monomial divides that of g and is hence a pure power of x_i. If $g = \sum b_j x_i^j$ and has degree N, then

$$0 \;=\; g \bmod G \;=\; \Bigl(\sum_{j=0}^{N} b_j x_i^j\Bigr) \bmod G \;=\; \sum_{j=0}^{N} b_j (x_i^j \bmod G)\,,$$

which is a linear dependence among the elements of the sequence $1 \bmod G$, $x_i \bmod G$, $x_i^2 \bmod G, \ldots$. Thus the algorithm halts when it is in Step (3). The minimality of the degree of g implies that $N = m$ and the uniqueness of such minimal linear combinations implies that the coefficients b_j and a_j are proportional, which shows that the algorithm computes a scalar multiple of g, which is also an eliminant. $\square$

Elimination using Gröbner bases leads to an algorithm for Question (v). The first step is to understand the optimal form of a Gröbner basis of a zero-dimensional ideal.

Lemma 2.4.10 (Shape Lemma)**.** *Suppose g is an eliminant of a zero-dimensional ideal I with $\deg(g) = \deg(I)$. When $\mathbb{K}$ is algebraically closed, I is radical if and only if g has no multiple factors.*

Suppose that $g = g(x_1)$, then in the lexicographic monomial order with $x_1 \prec x_2 \prec \cdots \prec x_n$, the ideal I has a Gröbner basis of the form:

$$g(x_1)\,, \quad x_2 - g_2(x_1)\,, \quad \ldots\,, \quad x_n - g_n(x_1)\,, \tag{2.17}$$

where $\deg(g) > \deg(g_i)$ for $i = 2, \ldots, n$.

If I is generated by polynomials with coefficients in a subfield $\Bbbk$, then the number of points of $\mathcal{V}(I)$ in $\Bbbk^n$ equals the number of roots of g in $\Bbbk$.

This is a simplified version of the Shape Lemma, which describes the form of a reduced Gröbner basis for any zero-dimensional ideal in the lexicographic order. For a zero-dimensional ideal which does not satisfy the hypotheses of Lemma 2.4.10, see Example 2.1.10.

Proof. Replacing $\mathbb{K}$ by its algebraic closure does not affect these algebraic statements, as the polynomials g and g_i coefficients in the original field, by Corollary 2.2.7. Replace $\mathbb{K}$ by its algebraic closure. We have

$$\#\text{roots of } g \;\leq\; \#\mathcal{V}(I) \;\leq\; \deg(I) \;=\; \deg(g)\,,$$

the first inequality is by Theorem 2.4.6 and the second by Theorem 2.4.4. If the roots of g are distinct, then their number is $\deg(g)$ and so these inequalities are equalities. This implies that I is radical, by Theorem 2.4.4. Conversely, if $g = g(x_i)$ has multiple roots, then there is a polynomial h with the same roots as g but with smaller degree. (We may select h to be the square-free part of g.) Since $\langle g\rangle = I \cap \mathbb{K}[x_i]$, we have that $h \notin I$, but since $h^{\deg(g)}$ is divisible by g, $h^{\deg(g)} \in I$, so I is not radical.

To prove the second statement, let d be the degree of the eliminant $g(x_1)$. Then each of $1, x_1, \ldots, x_1^{d-1}$ is a standard monomial, and since $\deg(g) = \deg(I)$, there are no others. Thus the initial ideal in the lexicographic monomial order is $\langle x_1^d, x_2, \ldots, x_n\rangle$. Each element of the reduced Gröbner basis for I expresses a generator of the initial ideal as a $\mathbb{K}$-linear combination of standard monomials. It follows that the reduced Gröbner basis has the form claimed.

For the last statement, observe that the common zeroes of the polynomials (2.17) are

$$\{(a_1, \ldots, a_n) \mid g(a_1) = 0 \text{ and } a_i = g_i(a_1)\,,\ i = 2, \ldots, n\}\,.$$

By Corollary 2.2.7, the polynomials $g, g_2, \ldots, g_n$ all have coefficients from $\Bbbk$, and so a component a_i lies in $\Bbbk$ if the root a_1 of $g(x_1)$ lies in $\Bbbk$. $\square$

Not all ideals I can have an eliminant g with $\deg(g) = \deg(I)$. For example, let $\mathfrak{m}_0 := \langle x, y\rangle$ be the maximal ideal corresponding to the origin $\{(0,0\} \in \mathbb{K}^2$. Then its square $\mathfrak{m}_0^2 = \langle x^2, xy, y^2\rangle$ has degree three, but any eliminant has degree two. For example x^2 is its elimination to the x-axis.

The failure of the condition $\deg(g) = \deg(I)$ in the Shape Lemma is easier to understand when I is radical. Indeed, when I is radical, $\deg(g(x_i)) = \deg(I)$ if and only if the projection map π_i to the coordinate x_i-axis is one-to-one. For example, if I is generated by the three polynomials

$$\begin{aligned}
f &:= 1574y^2 - 625yx - 1234y + 334x^4 - 4317x^3 + 19471x^2\\
&\qquad - 34708x + 19764 + 45x^2y - 244y^3\,,\\
g &:= 45x^2y - 305yx - 2034y - 244y^3 - 95x^2 + 655x + 264 + 1414y^2\,, \quad \text{and}\\
h &:= -33x^2y + 197yx + 2274y + 38x^4 - 497x^3 + 2361x^2 - 4754x\\
&\qquad + 1956 + 244y^3 - 1414y^2\,,
\end{aligned}$$

then $\mathcal{V}(I)$ is the seven nondegenerate points of Figure 2.3. There are only five points in the projection to the x-axis and four in the projection to the y-axis. The corresponding eliminants have degrees five and four,

$$2x^5 - 29x^4 + 157x^3 - 391x^2 + 441x - 180 \qquad 2y^4 - 13y^3 + 28y^2 - 23y + 6$$

Nevertheless, the key condition on the eliminant g, that $\deg(g) = \deg(I)$, often holds after a generic change of coordinates, as in the above example just as in the proof of Bézout's Theorem in the plane (Theorem 2.3.16). This gives the symbolic algorithm to count the number of real solutions to a system of equations whose ideal satisfies the hypotheses of the Shape Lemma.

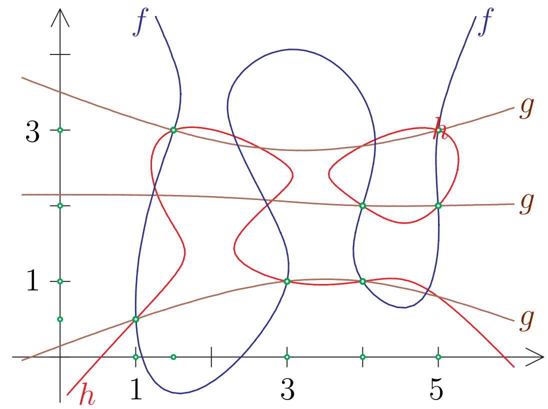


Figure 2.3: The seven points of $\mathcal{V}(f, g, h)$ and their projections.

Algorithm 2.4.11 (Counting real roots)**.**
INPUT: An ideal $I \subset \mathbb{R}[x_1, \ldots, x_n]$.
OUTPUT: The number of real points in $\mathcal{V}(I)$, if I satisfies the hypotheses of the Shape Lemma, or else "I does not satisfy the hypotheses of the Shape Lemma".

Compute $\dim(I)$ and $\deg(I)$. If I does not have dimension 0, then exit with "I is not zero-dimensional", else set $i := 1$.

1. Compute an eliminant $g(x_i)$ for I. If $deg(g) = \deg(I)$ and $\gcd(g, g') = 1$, then output the number of real roots of g. Else if $i < n$, set $i := i + 1$ and return to (1).

2. If no eliminant has been computed and $i = n$, then output "I does not satisfy the hypotheses of the Shape Lemma".

While this algorithm will not successfully compute the number of real ponts in $\mathcal{V}(I)$ (it would fail for the ideal of Figure 2.3), it may be combined with more sophisticated methods to accomplish that important task.

While the Shape Lemma describes an optimal form of a Gröbner basis for a zero-dimensional ideal, it is typically not optimal to compute such a Gröbner basis directly. An alternative to direct computation of a lexicographic Gröbner basis is the *FGLM algorithm* of Faugère, Gianni, Lazard, and Mora for Gröbner basis conversion. That is, given a Gröbner basis for a zero-dimensional ideal with respect to one monomial order $\lhd$ and a different monomial order $\succ$, the FGLM algorithm computes a Gröbner basis for the ideal with respect to $\succ$.

Algorithm 2.4.12 (FGLM)**.**
INPUT: A Gröbner basis G for a zero-dimensional ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ with respect to a monomial order $\lhd$, and a different monomial order $\succ$.
OUTPUT: A Gröbner basis H for I with respect to $\succ$.
INITIALIZE: Set $H := \{\}$, $x^\alpha := 1$, and $S := \{\}$.

(1) Compute $\overline{x^\alpha} := x^\alpha \mod G$. (or should we use $[x^\alpha]$?).

(2) If $\overline{x^\alpha}$ does not lie in the linear span of S, then set $S := S \cup \{\overline{x^\alpha}\}$.

Otherwise, there is a (unique) linear combination of elements of S such that

$$\overline{x^\alpha} \;=\; \sum_{\overline{x^\beta}\in S} c_\beta \overline{x^\beta}\,.$$

Set $H := H \cup \{x^\alpha - \sum_\beta c_\beta x^\beta\}$.

(3) If

$$\{x^\gamma \mid x^\gamma \succ x^\alpha\} \;\subset\; \mathrm{in}_\succ(H) := \langle \mathrm{in}_\succ(h) \mid h \in H\rangle\,,$$

then halt and output H. Otherwise, set x^α to be the $\succ$-minimal monomial in the set $\{x^\gamma \not\in \mathrm{in}_\succ(H) \mid x^\gamma \succ x^\alpha\}$ and return to (1).

Proof of correctness. By construction, H always consists of elements of I, and elements of S are linearly independent in the quotient ring $\mathbb{K}[x_1,\dots,x_n]/I$. Thus $\mathrm{in}_\succ(H)$ is a subset of the initial ideal $\mathrm{in}_\succ I$, and we always have the inequalities

$$|S| \;\leq\; \dim_{\mathbb{K}}(\mathbb{K}[x_1,\dots,x_n]/I) \qquad \text{and} \qquad \mathrm{in}_\succ(H) \;\subset\; \mathrm{in}_\succ I\,.$$

Every time we return to (1) either the set S or the set H (and also $\mathrm{in}_\succ(H)$) increases. Since the cardinality of S is bounded by $\deg(I)$ and the monomial ideals $\mathrm{in}_\succ(H)$ form a strictly increasing chain, the algorithm must halt.

When the algorithm halts, every monomial is either in the set $\mathrm{SM} := \{\mathrm{x}^\beta \mid \overline{\mathrm{x}^\beta} \in \mathrm{S}\}$ or else in the monomial ideal $\mathrm{in}_\succ(H)$. By our choice of x^α in (3), these two sets are disjoint, so that SM is the set of standard monomials for $\mathrm{in}_\succ(H)$. Since

$$\mathrm{in}_\succ(H) \;\subset\; \mathrm{in}_\succ\langle H\rangle \;\subset\; \mathrm{in}_\succ I\,,$$

and elements of S are linearly independent in $\mathbb{K}[x_1,\dots,x_n]/I$, we have

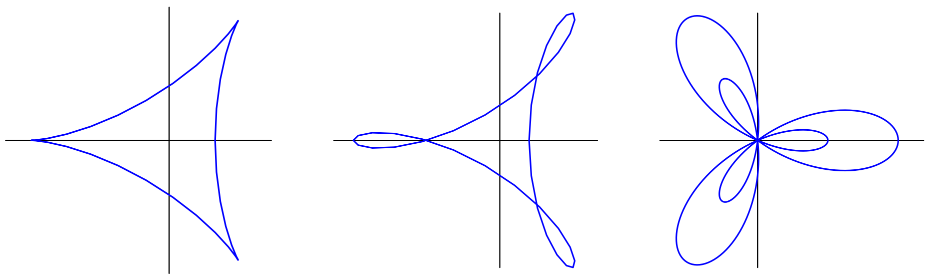
$$|S| \;\leq\; \dim_{\mathbb{K}}(\mathbb{K}[x_1,\dots,x_n]/I) \;=\; \dim_{\mathbb{K}}(\mathbb{K}[x_1,\dots,x_n]/\mathrm{in}_\succ I) \;\leq\; \dim_{\mathbb{K}}(\mathbb{K}[x]/\mathrm{in}_\succ(H)) \;=\; |S|\,.$$

Thus $\mathrm{in}_\succ I = \mathrm{in}_\succ(H)$, which proves that H is a Gröbner basis for I with respect to the monomial order $\succ$. By the form of the elements of H, it is the reduced Gröbner basis. □

Exercises

Need more exercises

1. The trigonometric curves parametrized by $(\cos(\theta) - \cos(2\theta)/2, \sin(\theta) + \sin(2\theta)/2)$, $(\cos(\theta) - 2\cos(2\theta)/3, \sin(\theta) + 2\sin(2\theta)/3)$, and the polar curve $r = 1 + 3\cos(3\theta)$ (for $\theta \in [0, 2\pi]$) are the cuspidal and trinodal plane quartics, and the flower with three petals, respectively.

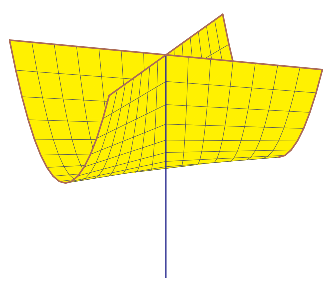


   Find their implicit equations by writing each as the projection to the (x, y)-plane of an algebraic variety in $\mathbb{K}^4$. Hint: These are images of the circle $c^2 + s^2 = 1$ under maps to the (x, y) plane, where the variables (c, s) correspond to $(\cos(\theta), \sin(\theta))$. The graph of the first is given by the three polynomials

$$c^2 + s^2 - 1\,,\ x - (c - (c^2 - s^2)/2)\,,\ y - (s + sc)\,,$$

   using the identities $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$ and $\sin(2\theta) = 2\sin(\theta)\cos(\theta)$.

2. The Whitney umbrella is the image in $\mathbb{K}^3$ of the map $(u, v) \mapsto (uv, u, v^2)$. Use elimination to find an implicit equation for the Whitney umbrella.



   Which points in $\mathbb{K}^2$ give the handle of the Whitney umbrella?

3. Suppose $I \subset \mathbb{K}[x_1, \ldots, x_n]$ is radical, $\mathbb{K}$ is algebraically closed, and $\mathcal{V}(I) \subset \mathbb{K}^n$ consists of finitely many points. Show that the coordinate ring $\mathbb{K}[x_1, \ldots, x_n]/I$ of restrictions of polynomial functions to $\mathcal{V}(I)$ has dimension as a $\mathbb{K}$-vector space equal to the number of points in $\mathcal{V}(I)$.

4. Compute the number of solutions to the system of polynomials

$$1 + 2x + 3y + 5xy \;=\; 7 + 11xy + 13xy^2 + 17x^2y \;=\; 0\,.$$

   Show that each is nondegenerate and compare this to the Bézout bound for this system. How many solutions are real?

5. In this and subsequent exercises, you are asked to use computer experimentation to study the number of solutions to certain structured polynomial systems. This is a good opportunity to become acquainted with symbolic software.

   For several small values of n and d, generate n random polynomials in n variables of degree d, and compute their numbers of isolated solutions. Does your answer agree with Bézout's Theorem?

6. A polynomial is *multilinear* if no variable occurs to a power greater than 1. For example,
$$3xyz - 7xy + 13xz - 19yz + 29x - 37y + 43z - 53$$
is a multilinear polynomial in the variables x, y, z. For several small values of n generate n random multilinear polynomials and compute their numbers of common zeroes, Does your answer agree with Bézout's Theorem?

7. Let $\mathcal{A} \subset \mathbb{N}^n$ be a finite set of integer vectors, which we regard as exponents of monomials in $\mathbb{K}[x_1, \ldots, x_n]$. A polynomial with support $\mathcal{A}$ is a linear combination of monomials whose exponents are from $\mathcal{A}$. For example
$$1 + 3x + 9x^2 + 27y + 81xy + 243xy^2$$
is a polynomial whose support is the column vectors of $\mathcal{A} = \left(\begin{smallmatrix} 0 & 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{smallmatrix}\right)$.

   For $n = 2, 3$ and many $\mathcal{A}$ with $|\mathcal{A}| > n$ and $0 \in \mathcal{A}$, generate random systems of polynomials with support $\mathcal{A}$ and determine their numbers of isolated solutions. Try to formulate a conjecture about this number of solutions as a function of $\mathcal{A}$.

8. Fix $m, p \geq 2$. For $\alpha\colon 1 \leq \alpha_1 < \cdots < \alpha_p \leq m{+}p$, let E_α be a $p \times (m{+}p)$ matrix whose entries in the columns indexed by α form the identity matrix, and the entries in position i, j are either variables if $j < \alpha_i$ or 0 if $\alpha_i < j$. For example, when $m = p = 3$, here are E_{245} and E_{356},
$$E_{245} = \begin{pmatrix} a & 1 & 0 & 0 & 0 & 0 \\ b & 0 & c & 1 & 0 & 0 \\ d & 0 & e & 0 & 1 & 0 \end{pmatrix} \qquad E_{356} = \begin{pmatrix} a & b & 1 & 0 & 0 & 0 \\ c & d & 0 & e & 1 & 0 \\ f & g & 0 & h & 0 & 1 \end{pmatrix}.$$
Set $|\alpha| := \alpha_1 - 1 + \alpha_2 - 2 + \cdots + \alpha_p - p$ be the number of variables in E_α. For all small m, p, and α, generate $|\alpha|$ random $m \times (m{+}p)$ matrices $M_1, \ldots, M_{|\alpha|}$ and determine the number of isolated solutions to the system of equations
$$\det\begin{pmatrix} E_\alpha \\ M_1 \end{pmatrix} = \det\begin{pmatrix} E_\alpha \\ M_2 \end{pmatrix} = \cdots = \det\begin{pmatrix} E_\alpha \\ M_{|\alpha|} \end{pmatrix} = 0\,.$$
Try to formulate a conjecture for the number of solutions as a function of m, p, and α.

2.5 Eigenvalue techniques

We discuss a connection between the solutions to systems of polynomial systems and eigenvalues of linear algebra. This leads to further methods to compute and analyze the roots of a zero-dimensional ideal. The techniques are based on classical results, but their computational aspects have only been developed systematically fairly recently.

Suppose that $\mathbb{K}$ is algebraically closed and $J \subset \mathbb{K}[x_1, \ldots, x_n]$ is a zero-dimensional ideal. Our goal is to interpret the coordinates of points in $\mathcal{V}(J)$ in terms of eigenvalues of suitable matrices. This is efficient as numerical linear algebra provides efficient methods to numerically determine the eigenvalues of a complex matrix, and the matrices we use are readily computed using algorithms based on Gröbner bases.

It is instructive to start with univariate polynomials. Given a monic univariate polynomial $p = c_0 + c_1x + \cdots + c_{d-1}x^{d-1} + x^d \in \mathbb{K}[x]$, the matrix

$$C_p \;=\; \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{d-1} \end{pmatrix} \;\in\; \mathbb{K}^{d\times d}\,.$$

is the *companion matrix* of p.

For any given matrix $A \in \mathbb{K}^{d\times d}$, the eigenvalues of A are the roots of the characteristic polynomial $\chi_A(x) \;=\; \det(A - xI_d)$. Thus the following statement tells us that the roots of p coincide with the eigenvalues of the companion matrix C_p.

Theorem 2.5.1. *Let $p = c_0 + \cdots + c_{d-1}x^{d-1} + x^d \in \mathbb{K}[x]$ be a monic univariate polynomial of degree $d \geq 1$. The characteristic polynomial of its companion matrix C_p is*

$$\det(xI_d - C_p) \;=\; (-1)^d p(x)\,.$$

Its companion matrix expresses multiplication by x in the ring $\mathbb{K}[x]/\langle p\rangle$ in the basis $1, x, \ldots, x^{d-1}$ of standard monomials.

Proof. For $d = 1$, the statement is clear, and for $d > 1$ expanding the determinant along the first row of $C_p - xI_d$ yields

$$\det(xI_d - C_p) \;=\; x\det(xI_{d-1} - C_q) + (-1)^{d+1}(-1)^{d-1}c_0\,,$$

where C_q is the companion matrix of the polynomial

$$q \;:=\; c_1 + c_2x + \cdots + c_{d-1}x^{d-2} + x^{d-1} \;=\; (p - c_0)/x\,.$$

Applying the induction hypothesis gives the result.

The claim that the matrix C_p exresses multiplication by x in $\mathbb{K}[x]/\langle p\rangle$ in the basis $1, x, \ldots, x^{d-1}$ of standard monomials is Exercise 2 below. $\square$

Let $I \subset \mathbb{K}[x_1,\ldots,x_n]$ be a zero-dimensional ideal. By Theorems 2.4.1 and 2.4.4, the $\mathbb{K}$-vector space $\mathbb{K}[x_1,\ldots,x_n]/I$ is finite-dimensional, and the cardinality of the variety $\mathcal{V}(I)$ is bounded from above by the dimension of $\mathbb{K}[x_1,\ldots,x_n]/I$. Given a polynomial $f \in \mathbb{K}[x_1,\ldots,x_n]$, write $\overline{f}$ for its residue class in the quotient ring $\mathbb{K}[x_1,\ldots,x_n]/I$.

For any $i = 1,\ldots,n$, multiplication of an element in $\mathbb{K}[x_1,\ldots,x_n]/I$ with the residue class $\overline{x_i}$ of a variable x_i defines an endomorphism m_i,

$$\begin{array}{rcl} m_i \,:\, \mathbb{K}[x_1,\ldots,x_n]/I & \longrightarrow & \mathbb{K}[x_1,\ldots,x_n]/I\,, \\ \overline{f} & \longmapsto & \overline{x_i}\cdot\overline{f} \;=\; \overline{x_i f}\,. \end{array}$$

Lemma 2.5.2. *The map $x_i \mapsto m_i$ induces an injection $K[x_1,\ldots,x_n]/I \hookrightarrow \mathrm{End}(K[x_1,\ldots,x_n]/I)$.*

Proof. For the second, the map $x_i \mapsto M_i$ induces a map φ from $\mathbb{K}[x_1,\ldots,x_n]$ to the endmorphism ring. For a polynomial $p, f \in \mathbb{K}[x_1,\ldots,x_n]$, the value of $p(m_1,\ldots,m_n)(\overline{f})$ is $\overline{p(x_1,\ldots,x_n)f}$. This implies that $I \subset \ker(\varphi)$. Setting $f = 1$ shows the other inclusion. □

This map $K[x_1,\ldots,x_n]/I \hookrightarrow \mathrm{End}(K[x_1,\ldots,x_n]/I)$ is called the regular representation of $K[x_1,\ldots,x_n]/I$. We will use it to study the variety $\mathcal{V}(I)$. Since the vector space $\mathbb{K}[x_1,\ldots,x_n]/I$ is finite-dimensional, we may represent each linear multiplication map m_i as a matrix with repsect to a fixed basis of $\mathbb{K}[x_1,\ldots,x_n]/I$. For this, the basis of standard monomials is not only convenient, but given a Gröbner basis, this representation is particularly easy to compute.

Let $\mathcal{B}$ be the set of standard monomials for I with respect a monomial order $\prec$. Let G be a Gröbner basis for I with respect to $\prec$. For each $i = 1,\ldots,n$, let $M_i \in \mathrm{Mat}_{\mathcal{B}\times\mathcal{B}}(K)$ be the matrix representing the endomorphism m_i of multiplication by the variable x_i with repsect to the basis $\mathcal{B}$, which we call the *i-th companion matrix* of the ideal I with respect to $\mathcal{B}$. The rows and the columns of the companion matrix M_i are indexed by the monomials in $\mathcal{B}$. For a pair of monimials $x^\alpha, x^\beta \in \mathcal{B}$, the entry of M_i in the row corresponding to x^α and column corresponding to x^β is the coefficient of x^α in $x_i \cdot x^\beta \bmod G$, the normal form of $x_i \cdot x^\beta$. It follows that there is an easy Gröbner basis algorithm for compute M_i.

Lemma 2.5.3. *The companion matrices commute,*

$$M_i \cdot M_j \;=\; M_j \cdot M_i \quad \textit{for } 1 \le i < j \le n\,.$$

Proof. The matrices M_iM_j and M_jM_i represent the compositions $m_i \circ m_j$ and $m_j \circ m_i$, respectively. The first statement follows as multiplication in $\mathbb{K}[x_1,\ldots,x_n]/I$ is commutative. □

By Lemma 2.5.2, the companion matrices $M_1,\ldots,M_n$ generate a subalgebra of $\mathrm{Mat}_{\mathcal{B}\times\mathcal{B}}(\mathbb{K})$ isomorphic to $\mathbb{K}[x_1,\ldots,x_n]/I$. As the quotient ring $\mathbb{K}[x_1,\ldots,x_n]/I$ is commutative, when $\mathbb{K}$ is algebraically closed this subalgebra has a collection of common eigenvectors whose eigenvalues are characters (homomorphisms to $\mathbb{K}$) of $\mathbb{K}[x_1,\ldots,x_n]/I$. The following fundamental result will allow us to identify the eigenvectors with the points of $a \in \mathcal{V}(I)$ with corresponding eigenvalue the evaluation of a element of $\mathbb{K}[x_1,\ldots,x_n]/I$ at the point a.

Theorem 2.5.4 (Stickelberger's Theorem)**.** *Suppose that $\mathbb{K}$ is algebraically closed and $I \subset \mathbb{K}[x_1, \ldots, x_n]$ is a zero-dimensional ideal. For each $i = 1, \ldots, n$ and any $\lambda \in \mathbb{K}$, the value λ is an eigenvalue of the endomorphism m_i if and only if there exists a point $a \in \mathcal{V}(I)$ with $a_i = \lambda$.*

Corollary 2.5.5. *Let $R \subset \mathrm{End}(\mathbb{K}[x_1, \ldots, x_n]/I)$ be the commutative subalgebra generated by the endomorphisms $m_1, \ldots, m_n$. The joint eigenvectors of R correspond to points of $\mathcal{V}(I)$. For $p \in \mathbb{K}[x_1, \ldots, x_n]$ and $a \in \mathcal{V}(I)$, the eigenvalue of $p(m_1, \ldots, m_n)$ on the eigenvector corresponding to a is $p(a)$.*

For the proof of this Stickelberger's Theorem, we we recall some facts from linear algebra related to the Cayley-Hamilton Theorem.

Definition 2.5.6. Let V be a vector space over $\mathbb{K}$ and ϕ an endomorphism on V. For any polynomial $p = \sum_{i=0}^{d} c_i t^i \in \mathbb{K}[t]$, set $p(\phi) := \sum_{i=0}^{d} c_i \phi^i \in \mathrm{End}(V)$, where ϕ^i is the i-fold composition of the endomorphism ϕ with itself. The ideal $I_\phi := \{p \in \mathbb{K}[t] \mid p(\phi) = 0\}$ is the kernel of the homomorphism $\mathbb{K}[t] \to \mathrm{End}(V)$ defined by $t \mapsto \phi$. Its unique monic generator h_ϕ is the *minimal polynomial of ϕ*.

The eigenvalues and the minimal polynomial of an endomorphism are related.

Lemma 2.5.7. *Let V be a finite-dimensional vector space over an algebraically closed field $\mathbb{K}$ and ϕ be an endomorphism of V. Then an element $\lambda \in \mathbb{K}$ is an eigenvalue of ϕ if and only if λ is a zero of the minimal polynomial h_ϕ.*

Proof. The eigenvalues of ϕ are the roots of its characteristic polynomial χ_ϕ. By the Cayley-Hamilton Theorem, the characteristic polynomial vanishes on ϕ, $\chi_\phi(\phi) = 0$. Thus $\chi_\phi \in I_\phi$ and h_ϕ divides χ_ϕ.

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of ϕ, which are the roots of χ_ϕ. Suppose there is some eigen value, say λ_1, for which $h_\phi(\lambda_1) \neq 0$. That is, the roots of h_ϕ are a proper subset of the eigenvalues, and we may write

$$h_\phi(t) \ = \ (t - \lambda_2)^{d_2}(t - \lambda_3)^{d_3} \cdots (t - \lambda_m)^{d_m} .$$

Let $v \in V$ be an eigenvector of ϕ with eigenvalue λ_1. For any other eigenvalue $\lambda_i \neq \lambda_1$, we have $(\phi - \lambda_i id).v = (\lambda_1 - \lambda_i).v \neq 0$, and so

$$h_\phi(\phi).v \ = \ (\phi - \lambda_2)^{d_2} \cdots (\phi - \lambda_m)^{d_m}.v \ = \ (\lambda_1 - \lambda_2)^{d_2} \cdots (\lambda_1 - \lambda_m)^{d_m} v \ \neq \ 0 \,,$$

which contradics h_ϕ being the minimal polynomial of ϕ. $\square$

We can now prove Stickelberger's Theorem 2.5.4.

Proof of Theorem 2.5.4. Let λ be an eigenvalue of the multiplication endomorphism m_i on $\mathbb{K}[x_1, \ldots, x_n]/I$ with corresponding eigenvalue $\overline{v}$. That is, $\overline{x_i v} = \lambda \overline{v}$ and thus $\overline{(x_i - \lambda) \cdot v} =$

0 in the vector space $\mathbb{K}[x_1,\ldots,x_n]/I$ so that $(x_i-\lambda)v \in I$. Let us assume by way of contradiction that there is no point $a \in \mathcal{V}(I)$ with ith coordinate λ.

This implies that $x_i-\lambda$ vanishes at no point of $\mathcal{V}(I)$. We will use this to show that $\overline{x_i-\lambda}$ is invertible in $\mathbb{K}[x_1,\ldots,x_n]/I$. Multiplying the equation $\overline{(x_i-\lambda)\cdot v}=0$ by this inverse imples that $\overline{v}=0$, which is a contradiction as eigenvectors are nonzero.

In Exercise 5 of Section 1.3 you are asked to show that the map $\mathbb{K}[x_1,\ldots,x_n] \to \mathbb{K}^{\mathcal{V}(I)}$ is surjective, where $\mathbb{K}^{\mathcal{V}(I)}$ is the ring of functions on the finite set $\mathcal{V}(I)$, which is $\mathbb{K}[x_1,\ldots,x_n]/\sqrt{I}$. It follows that there exists a polynomial $f \in \mathbb{K}[x_1,\ldots,x_n]$ with image

$$\overline{f} \;=\; \sum_{a\in\mathcal{V}(I)} \frac{1}{a_i-\lambda}\delta_a$$

in $\mathbb{K}^{\mathcal{V}(I)} = \mathbb{K}[x_1,\ldots,x_n]/\sqrt{I}$, where δ_a is the Kronecker delta function, whose value at a point b is zero unless $b=a$, and then its value is 1. Then $f(a)=1/(a_i-\lambda)$ for $a\in\mathcal{V}(I)$, from which we obtain

$$(1 \;-\; (x_i-\lambda(f)) \; in \; \mathcal{I}(\mathcal{V}(I)) \;=\; \sqrt{I}\,.$$

By Hilbert's Nullstellensatz, there is a positive integer N such that $(1-(x_i-\lambda(f))^N \in I$. Expanding this, we obtain

$$1 \;-\; N(x_i-\lambda)f \;+\; \tbinom{N}{2}(x_i-\lambda)^2f^2 \;-\; \cdots \;\in\; I\,,$$

and so there exists a polynomial g such that $1-(x_i-\lambda)g \in I$. Then $\overline{g}$ is the desired inverse to $\overline{x_i-\lambda}$ in $\mathbb{K}[x_1,\ldots,x_n]/I$.

Conversely, let $a\in\mathcal{V}(I)$ with $a_i=\lambda$. Let h_i be the minimal polynomial of m_i. By Lemma 2.5.7 we need only show that $h_i(\lambda)=0$. By the definition of minimal polynomial, the function $h_i(m_i)$ is the zero endomorphism on $\mathbb{K}[x_1,\ldots,x_n]/I$. In particular, $h_i(\overline{x_i}) = h_i(m_i)(\overline{1}) = 0$ in $\mathbb{K}[x_1,\ldots,x_n]/I$, which implies that the polynomial $h_i(x_i)\in\mathbb{K}[x_1,\ldots,x_n]$ lies in I. Evaluating this polynomial at a point $a\in\mathcal{V}(I)$ gives $0=h(a)=h(a_i)=h(\lambda)$. □

Example 2.5.8. Let $I=\langle x^2y+1, y^2-1\rangle$. A Gröbner basis of I with respect to the lexicographic ordering is given by $\{x^4-1, y+x^2\}$, hence a basis of $\mathbb{K}[x,y]/I$ is $\{1,x,x^2,x^3\}$. With respect to this basis, the representing matrices of the endomorphisms m_x and m_y are

$$M_x \;=\; \begin{pmatrix} 0&0&0&1\\ 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0 \end{pmatrix} \quad\text{and}\quad M_y \;=\; \begin{pmatrix} 0&0&-1&0\\ 0&0&0&-1\\ -1&0&0&0\\ 0&-1&0&0 \end{pmatrix}.$$

The eigenvalues of M_x are $-1,1,-i,i$ and the eigenvalues of M_y are -1 (twice) and 1 (twice). Indeed, we have $\mathcal{V}(I)=\{(i,1),(-i,1),(1,-1),(-1,-1)\}$.

Have we already addressed this? From a computational point of view, Theorem 2.5.4 requires that we know a basis of the coordinate ring $\mathbb{K}[x_1,\dots,x_n]/I$ and the companion matrices in this basis. Given these data, the computational complexity depends on the dimension, d, of $\mathbb{K}[x_1,\dots,x_n]/I$.

These methods simplify when there exists a joint basis of eigenvectors. That is, if there exists a matrix $S\in\mathbb{K}^{d\times d}$ and diagonal matrices $D_i\in\mathbb{K}^{d\times d}$ for $i=1,\dots,n$ with

$$M_iS \;=\; SD_i\,,\quad \text{for } i=1,\dots,n 1\leq i\leq n\,.$$

When this occurs, we say that the companion matricesM_i are simultaneously diagonalizable.

Theorem 2.5.9. *The companion matrices $M_1,\dots,M_n$ are simultaneously diagonalizable if I is radical.* What about the converse?

Proof. Let $a=(a_1,\dots,a_n)$ be a point in $\mathcal{V}(I)$. As in the proof of Theorem 2.5.4, there exists a polynomial g with $g(a)=1$ and $g(b)=0$ for all $b\in\mathcal{V}(I)\smallsetminus\{a\}$. Hence, the polynomial $(x_i-a_i)g$ vanishes on $\mathcal{V}(I)$. Hilbert's Nullstensatz then implies $(x_i-a_i)[g]\in\sqrt{I}=I$, and thus $[g]$ is a joint eigenvector of $M_1,\dots,M_n$. □

Stickelberger's Theorem 2.5.4 not only connects classical linear algebra to the problem of finding the common zeroes of a zero-dimensional ideal, but it leads to another method to compute eliminants.

Corollary 2.5.10. *Suppose that $I\subset\mathbb{K}[x_1,\dots,x_n]$ is a zero-dimensional ideal. The eliminant $g(x_i)$ is the minimal polynomial of the operator m_i of multiplication by x_i on $\mathbb{K}[x_1,\dots,x_n]/I$. It is a factor of the characteristic polynomial χ_{m_i} of m_i that contains all the roots of χ_{m_i}.* can say more, perhaps.

This leads to an algorithm to compute the eliminant $g(x_i)$ of the radical

Algorithm 2.5.11.
Input: A zero-dimensional ideal $I\subset\mathbb{K}[x_1,\dots,x_n]$ and an index i.
Output: The eliminant $g(x_i)$ of the radical of I.

Compute a Gröbner basis G for I with respect to any monomial order $\prec$. If $\dim I\neq 0$, then exit, else let $\mathcal{B}$ be the corresponding finite set of standard monomials.

Construct M_i, the matrix in $\mathrm{Mat}_{\mathcal{B}\times\mathcal{B}}(\mathbb{K})$ representing multiplication by x_i on $\mathbb{K}[x_1,\dots,x_n]/I$ in the basis of standard monomials. Let χ_{m_i} be the characteristic polynomial of M_i.

Proof of correctness. □

Are there any other results to put in? What about I not radical if some M_i is not semisimple?

Exercises

1. Let $G := \{yx - 1, z - x\}$ and $I := \langle G \rangle$ be an ideal in $\mathbb{C}[x, y]$. Show that G is a Gröbner basis of I for the lexicographic order $x \prec y \prec z$, determine the set of standard monomials of $\mathbb{C}[x, y]/I$ and compute the multiplication matrices M_x and M_y. do a a meatier example.

2. Let $p = c_0 + \cdots + c_{d-1}x^{d-1} + x^d$ be a monic, univariate polynomial and set $I := \langle p \rangle$. Show that the matrix M_x representing the endomorphism $m_x : R/I \to R/I$, $[f] \mapsto [xf]$ with respect to a natural basis coincides with the companion matrix C_p.

3. Let $f \in \mathbb{K}[x_1, \ldots, x_n]$. Show that $m_f \colon \mathbb{K}[x_1, \ldots, x_n]/I \to \mathbb{K}[x_1, \ldots, x_n]/I$, where $[g] \mapsto [f] \cdot [g]$ is an endomorphism. State a little more here.

4. In a computer algebra system, use the method of Stickelberger's Theorem to determine the common complex zeroes of $x^2 + 3xy + y^2 - 1$ and $x^2 + 2xy + y + 3$.

5. If two endormorphisms f and g on a finite-dimensional vector space V are diagonalizable and $f \circ g = g \circ f$, then they are jointly diagonalizable. Conclude that for Stickelberger's Theorem for the ring $\mathbb{K}[x, y]$ with only two variables, there always exist a basis of joint eigenvectors.

6. Perform the following compuational experiment. Compute eliminants using all three methods given in the text.

2.6 Notes

Resultants were developed in the nineteenth century by Sylvester, were part of the computational toolkit of algebra from that century, and have remained a fundamental symbolic tool in algebra and its applicaations. Even more classical is Bézout's Theorem, stated by Etienne Bézout in his 1779 treatise *Théorie Générale des Équations Algébriques* [10, 11].

The subject ot Gröbner bases began with Buchberger's 1965 Ph.D. thesis which contained his algorithm to compute Gröbner bases [15, 17]. The term "Gröbner basis" honors Buchberger's doctoral advisor Wolfgang Gröbner. Key ideas about Gröbner bases had appeared earlier in work of Gordan and of Macaulay, and in Hironaka's resolution of singularities [39]. Hironaka called Gröbner bases "standard bases", a term which persists. For example, in the computer algebra package `Singular` [31] the command `std(I);` computes the Gröbner basis of an ideal `I`. Despite these precedents, the theory of Gröbner bases rightly begins with these Buchberger's contributions.

Theorem 2.2.3 was proven by Macaulay [55], who the Gröbner basis package `Macaulay 2` [**?**] was named after.

There are additional improvements in Buchberger's algorithm (see Ch. 2.9 in [20] for a discussion), and even a series of completely different algorithms due to Jean-Charles Faugère [27] based on linear algebra with vastly improved performance.

An alternative to direct computation of a lexicographic Gröbner basis is the *FGLM algorithm* of Faugère, Gianni, Lazard, and Mora [28], which is an algorithm for Gröbner basis conversion.

For further information on technqiues for solving systems of polynomial equations see the books of Cox, Little, and O'Shea [21, 20], Sturmfels [90] as well as Emiris and Dickenstein [23].

For numerical methods concerning the simultaneous diagonalization of matrices we refer the reader to Bunse-Gerstner, Byers, and Mehrmann [18]. In Section 5.2, a further refinement of the eigenvalue techniques will be used to study real roots.