6 Gröbner Bases

We are all familiar with the notion of simultaneous linear equations such as

$$a_{11}x_1 + a_{21}x_2 = b_1$$
$$a_{12}x_1 + a_{22}x_2 = b_2$$

where the a_{ij} represent elements of some field (e.g., $\mathbb{R}$). There are well known simple conditions on the a_{ij} for these equations to have a solution for x_1, x_2. In fact we can state straightforward conditions for the existence of solutions to systems of m equations in n unknowns. Our interest here is to look at a more wide reaching generalization.

6.1 Basics of Algebraic Geometry

Consider polynomials $p_1(x_1, \ldots, x_n), \ldots, p_m(x_1, \ldots, x_n)$ in the indeterminates $X = \{x_1, \ldots, x_n\}$ with coefficients from some field k. The *variety* corresponding to these polynomials is the set of their common zeros[5]. To be more precise it is the subset $\mathbf{V}(p_1, \ldots, p_m)$ of k^n defined by

$$\mathbf{V}(p_1, \ldots, p_m) = \{\, (a_1, \ldots, a_n) \in k^n \mid p_i(a_1, \ldots, a_n) = 0, \text{ for } 1 \leq i \leq n \,\}.$$

For example if $k = \mathbb{R}$ and $n = 2$ then $\mathbf{V}(x^2+y^2-1)$ is a circle of radius 1 and centre at the origin in ordinary 2-dimensional space. Similarly $\mathbf{V}(x^2 + y^2 - 1, x - y)$ consists of the points of intersection of the circle and the line $y = x$. Note that $\mathbf{V}(x^2 + y^2 + 1)$ is empty for $k = \mathbb{R}$ but on the other hand $\mathbf{V}(x^2 + y^2 + 1)$ has infinitely many points if $k = \mathbb{C}$.

The definition of a variety can be extended to arbitrary (possibly infinite) subsets S of $k[X]$:

$$\mathbf{V}(S) = \{\, (a_1, \ldots, a_n) \in k^n \mid p(a_1, \ldots, a_n) = 0, \text{ for all } p \in S \,\}.$$

Suppose that $p_1, \ldots, p_s \in S$ and $q_1, \ldots, q_s \in k[X]$. Put

$$q = q_1p_1 + \cdots + q_sp_s.$$

Then it is clear that for any $(a_1, \ldots, a_n) \in \mathbf{V}(S)$ we have

$$q(a_1, \ldots, a_n) = 0.$$

Thus

$$\mathbf{V}(S \cup \{q\}) = \mathbf{V}(S).$$

Indeed we can add any set of polynomials such as q to S without changing the variety we obtain. This observation motivates the notion of an *ideal* generated by a set S. This is denoted by (S) and is defined by

$$(S) = \{\, q_1p_1 + \cdots + q_sp_s \mid s \geq 1, q_i \in k[X], p_i \in S, \text{ for } 1 \leq i \leq s \,\}.$$

This definition is made clearer by seeing it in operational terms: the elements of the ideal (S) are built by taking one or more elements of S, an equal number of *arbitrary* polynomials from $k[X]$ then, multiplying and adding up as shown. For example if S contains the polynomials $p_1 = x^2y + x - 1$

[5] Some authors use the term *algebraic set* and reserve the term variety for a special kind of algebraic set.

and $p_2 = xy^2 + y - 1$ then the ideal (S) has amongst its (infinitely many) elements the polynomials $x - y$ and $3x^2y^3 + 2x^3y + 3xy^2 + 2x^2 - 3y^2 - 2x$ since the first of these is $yp_1 - xp_2$ while the second is $(2x + 3y^2)p_1$. Note that if S is a finite set, say $S = \{p_1, p_2, \ldots, p_m\}$, then we can simplify the definition to

$$(S) = \{\, q_1p_1 + \cdots + q_mp_m \mid q_i \in k[X], \text{ for } 1 \leq i \leq m \,\}.$$

(Make sure you understand why this is so.) Furthermore it is clear that if $f_1, f_2 \in (S)$ then $f_1g_1 + f_2g_2 \in (S)$ for all $g_1, g_2 \in k[X]$. The preceding discussion shows that

$$\mathbf{V}\big(S\big) = \mathbf{V}\big((S)\big).$$

The abstract definition of ideals is: a subset I of $k[X]$ is an ideal if it is not empty and for all $q \in k[X]$ and $p_1, p_2 \in I$ we have $p_1q \in I$ and $p_1 - p_2 \in I$.

Note that although we have defined ideals in $k[X]$ exactly the same definition applies to arbitrary commutative rings. Thus it makes sense to talk about ideals in $\mathbb{Z}$. In fact $\mathbb{Z}$ is an example of a *prinicipal ideal domain*, these are integral domains for which every ideal is generated by a single element. Another example of a principal ideal domain is $k[x]$, i.e., polynomials in one indeterminate with coefficients from a field. An interesting property of these domains is that any ideal $(a_1, \ldots, a_n)$ is the same as (d) where d is any gcd of $a_1, \ldots, a_n$.

Exercise 6.1 *Show that:*

1. $S \subseteq (S)$ *for all* $S \subseteq k[X]$.
2. *If an ideal* I *contains a non-zero constant (i.e., a non-zero element of* k*) then* $I = k[X]$.
3. 0 *is an element of every ideal.*
4. $f \in (g)$ *if and only if* $g \mid f$.
5. $f - g \in I$ *if and only if* $f \in I$ *where* I *is an ideal of* $k[X]$, $g \in I$ *and* $f \in k[X]$.
6. *In the abstract definition of ideals we can replace the condition* $p_1 - p_2 \in I$ *by* $p_1 + p_2 \in I$ *without causing any change to what is meant by an ideal.*
7. *The abstract definition of ideals agrees with the version we introduced first, i.e., if* I *is an abstract ideal then there is an* S *such that* $I = (S)$. *Conversely* (S) *satisfies the abstract definition of ideals.*

Exercise 6.2 *Let* I *be an ideal of* $k[x]$*, note the single indeterminate here. Show that for all* $p_1, p_2 \in I$ *we have* $\gcd(p_1, p_2) \in I$. *Deduce that there is a single polynomial* p *such that* $I = (p)$ *(in technical terms this says that* $k[x]$ *is a* principal ideal domain*).*

(An alternative way to prove that $I = (p)$ *is to choose* p *to be an element of* I *of least possible degree then prove that* p *divides every other member of* I*.)*

Exercise 6.3 *Let* D *be a principal ideal domain (see above for the definition) and let* $a_1, \ldots, a_n \in D$.

- *Choose* d *such that* $(a_1, \ldots, a_n) = (d)$. *Prove that* d *is a gcd of* $a_1, \ldots, a_n$*, where the definition of gcd is the obvious generalisation of the one given for pairs of elements.*

- *Suppose d is a gcd of $a_1, \ldots, a_n$. Prove that $(a_1, \ldots, a_n) = (d)$.*

At first sight the move from a finite set of polynomials to an infinite set might seem a little retrograde. In order to give a hint of the power of the idea consider:

$$\begin{aligned} p_1 &= x + y - 2z - 1, \\ p_2 &= 2x - 3y - z + 2, \\ p_3 &= x - y + z, \end{aligned}$$

where the coefficients are from $\mathbb{Q}$ and let $I = (p_1, p_2, p_3)$. Now

$$p_4 = p_2 - 2p_1 = -5y + 3z + 4,$$

is a member of I. Therefore

$$p_5 = p_3 - p_1 - 2/5p_4 = 9/5z - 3/5,$$

is also a member of I. Thus

$$(p_1, p_4, p_5) \subseteq I.$$

In fact we can easily express p_2, p_3 as linear combinations of p_1, p_4, p_5 so that

$$I = (p_1, p_4, p_5).$$

Thus

$$\begin{aligned} V(I) &= V(p_1, p_2, p_3) \\ &= V(x + y - 2z - 1, -5y + 3z + 4, 9/5z - 3/5) \end{aligned}$$

and of course the final set of equations is very easy to solve since it is in triangular form.

As another example consider

$$S = \{x^3yz + x + 1, xy^2z + 1, x^2y^2 + z^2\} \subset \mathbb{C}[x, y, z].$$

It is not immediately obvious that these equations even have any simultaneous solutions let alone what such solutions look like. However we can show that

$$(S) = (x - z^3, y + z^{15} - z^{12} + z^9 - z^6 + z^5 + z^2, z^{16} + z^6 + 2z^3 + 1)$$

We can now solve for z using the last polynomial, each value of z fixes the value of x and y.

Finally consider the polynomials

$$\begin{aligned} f_1 &= x^2 + y^2 - 1, \\ f_2 &= xy - 1, \\ f_3 &= x^2 - y^2 - 1, \end{aligned}$$

from $\mathbb{C}[x, y]$. Now the polynomial

$$f = (-xy + y^2 + 2)f_1 + (-2xy + 2x^2 - 4)f_2 + (y^2 - xy)f_3,$$

is a member of (f_1, f_2, f_3) so that $\mathbf{V}(f_1, f_2, f_3) = \mathbf{V}(f, f_1, f_2, f_3)$. However $f = 2$ and so it follows that $\mathbf{V}(f, f_1, f_2, f_3) = \emptyset$.

The preceding examples show that a suitable change in the generating set of an ideal can give us a great deal of insight into the the variety that it defines. In terms of linear equations a version of this process is very familiar under the guise of Gaussian elimination. Can we find a systematic approach to the general situation? The aim of this section is to show that we can indeed do this.

Given an arbitrary ideal I we say that S is a *basis* of I if $I = (S)$. One of the major questions of nineteenth century algebra was: does every ideal have a *finite* basis?, i.e., given any ideal I are there finitely many polynomials $p_1, \ldots, p_m$ such that $I = (p_1 \ldots, p_m)$? The significance of this question is easier to appreciate if we phrase it geometrically. Suppose that we define figures in n dimensional space by means of infinitely many polynomial equations. Are there finitely many equations which give us precisely the same figures?

For $n = 1$ it is easy to provide an affirmative answer to the question (we rely on our old friend the Euclidean algorithm—see Exercise 6.2). The case $n = 2$ was settled by Gordan, also in the affirmative. Unfortunately his proof involved very lengthy algebraic calculations. Despite much effort by many mathematicians all attempts to generalize Gordan's proof got bogged down in excessive calculations. It was therefore a major surprise when Hilbert published a very short note which settled the question for all values of n:

Theorem 6.1 (Hilbert's Basis Theorem, 1888) *Every ideal of $k[X]$ has a finite basis.*

Hilbert's method of proof was so revolutionary that many mathematicians doubted that it was even mathematics. Indeed Gordan's reaction was: 'Das ist nicht Mathematik. Das ist Theologie'. This type of reaction is explained by the fact that for most nineteenth century mathematicians proving the existence of an object meant the demonstration of a method for finding that object. Hilbert on the other hand simply demonstrated the existence of finite bases without showing how to find them. Later on he produced a constructive proof based on his earlier non-constructive one. Hilbert's new approach was so powerful that it produced a major revolution in mathematical thinking. After such a build-up the laws of show business[6] demand that we give the

Proof of Hilbert's Basis Theorem: We show that if R is a commutative ring with identity with the property that every ideal of R is finitely generated then the same holds for the polynomial ring $R[x]$. (Hilbert's basis theorem as stated above then follows from induction on the number of indeterminates and the fact that fields have only two ideals, viz. (0) and (1) both of which are finitely generated.)

Let I be an ideal of $R[x]$. For each $n \in \mathbb{N}$ put

$$I_n = \{\, a \in R \mid \exists\, ax^n + a_{n-1}x^{n-1} + \cdots + a_0 \in I \,\}.$$

It is easy to see that I_n is an ideal of R and $I_n \subseteq I_{n+1}$. Now $\cup_{n\geq 0} I_n$ is also an ideal of R and so it must be finitely generated. Using this fact we easily deduce that for some $N \geq 0$ we have

$$I_N = I_{N+1} = \cdots$$

(see Exercise 6.3). Of course each I_n for $n \leq N$ is also finitely generated by $a_{n1}, a_{n2}, \ldots, a_{nm_n} \in R$, say. For each such generator we pick a polynomial $f_{ns} = a_{ns}x^n + \cdots \in J$ where n is the degree of f_{ns} (such a polynomial exists by the definition of I_n). Now for the *coup de grâce*. We claim that

$$\{f_{ns} \mid 0 \leq n \leq N, 1 \leq s \leq m_i\}$$

[6] Don't bring a cannon on the stage unless you intend to fire it.

generates I. To see this take any $g \in I$ and suppose $\deg(g) = d$ (we can assume w.l.o.g. that $g \neq 0$). Then $g = bx^d +$ (terms of lower degree). From what we know about I_d we can write $b = \sum c_{d'j} a_{d'j}$ where $d' = d$ if $d' \leq N$ otherwise $d' = N$. Consider $g_1 = g - x^{(d-d')} \sum c_{d'j} f_{d'j}$. By construction the coefficient of x^d in g_1 is 0 so that either $g_1 = 0$ or $\deg(g_1) < \deg(g)$. If $g_1 = 0$ we already have g as a linear combination of the f_{ns} otherwise we are done by induction on d. □

Exercise 6.4 *A ring R is called* Noetherian *if every ascending chain of ideals terminates, i.e., whenever we have $I_1 \subseteq I_2 \subseteq \cdots$ where the I_j are ideals of R then there is an s such that $I_s = I_{s+1} = \cdots$. Use Hilbert's Basis Theorem to deduce that $k[X]$ is Noetherian. Conversely deduce Hilbert's Basis Theorem from the assumption that $k[X]$ is Noetherian.*

(The terminology is in honour of Emmy Noether who studied such rings very extensively.)

Exercise 6.5 *A monomial ideal I is an ideal of form (S) where $S \subset k[X]$ is a set of power products. Assume that $I = (S)$ is such an ideal.*

1. *Let s be a power product. Prove that $s \in I$ if and only if s is divisible by a power product from S.*
2. *Prove that that a polynomial f is in I if and only if each power product of f is in I.*
3. *Deduce that every monomial ideal has a finite basis consisting of power products.*

We can view $\mathbf{V}$ as a function

$$\text{Ideals} \to \text{Varieties}.$$

There is an obvious function I in the opposite direction:

$$\text{Varieties} \to \text{Ideals}$$

which assigns to a variety V the ideal

$$\mathrm{I}(V) = \{\, p \mid p \in k[X] \;\&\; p(a_1, \ldots, a_n) = 0, \text{ for all } (a_1, \ldots, a_n) \in V \,\}.$$

(It is easy to verify that the set is indeed an ideal—try it.) For example a single point $P = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in k^n$ is a variety, e.g., it is $\mathbf{V}(x_1 - \alpha_1, x_2 - \alpha_2, \ldots, x_n - \alpha_n)$. Let us determine its ideal, i.e., $\mathrm{I}(P)$. First of all it is clear that P is a zero of any polynomial of the form

$$(x_1 - \alpha_1)g_1 + (x_2 - \alpha_2)g_2 + \cdots + (x_n - \alpha_n)g_n,$$

where $g_1, g_2, \ldots, g_n \in k[x_1, x_2, \ldots, x_n]$ are arbitrary. Since the set of all such polynomials is precisely the ideal $(x_1 - \alpha_1, x_2 - \alpha_2, \ldots, x_n - \alpha_n)$ we conclude that this ideal is contained in $\mathrm{I}(P)$. On the other hand we know from the remainder theorem of §4.7.7 that for every $f \in k[x_1, x_2, \ldots, x_n]$ we have

$$f(x_1, x_2, \ldots, x_n) - f(\alpha_1, \alpha_2, \ldots, \alpha_n) = (x_1 - \alpha_1)g_1 + (x_2 - \alpha_2)g_2 + \cdots + (x_n - \alpha_n)g_n,$$

for some $g_1, g_2, \ldots, g_n \in k[x_1, x_2, \ldots, x_n]$. Now if $f \in \mathrm{I}(P)$ then by definition $f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0$ so that

$$f = (x_1 - \alpha_1)g_1 + (x_2 - \alpha_2)g_2 + \cdots + (x_n - \alpha_n)g_n,$$

which means that $f \in (x_1 - \alpha_1, x_2 - \alpha_2, \ldots, x_n - \alpha_n)$. Thus we have

$$\mathrm{I}(P) = (x_1 - \alpha_1, x_2 - \alpha_2, \ldots, x_n - \alpha_n).$$

Unfortunately, this example is not typical since determining the ideal of a general variety is by no means an easy process.

Two obvious questions now arise:

1. is $I = \mathrm{I}\,\mathbf{V}(I)$ for an arbitrary ideal I of $k[X]$?
2. is $V = \mathbf{V}\,\mathrm{I}(V)$ for an arbitrary variety V of k^n?

In fact it is very easy to see that

1. $I \subseteq \mathrm{I}\,\mathbf{V}(I)$ for all ideals I of $k[X]$,
2. $V \subseteq \mathbf{V}\,\mathrm{I}(V)$ for all varieties V of k^n,

and it can be shown that equality always holds in the second case but in general it does not hold in the first case. To see that equality can fail in the first case consider the ideal generated by a single polynomial of the form p^2, where $p \neq 0$. It is clear that p is not a member of this ideal but of course p vanishes at all points at which p^2 vanishes. (Clearly we can replace the exponent 2 by any integer $s > 1$.)

A remarkable theorem (again due to Hilbert) gives us a complete description of $\mathrm{I}\,\mathbf{V}(I)$ provided we assume that k is *algebraically closed* (i.e., every non-constant polynomial in one indeterminate with coefficients from k has a root in k). The best known example of such a field is $\mathbb{C}$, the field of complex numbers. (It can be shown that every field k can be extended to a field $\overline{k}$ which is algebraically closed—this is what we obtain in moving from $\mathbb{R}$ to $\mathbb{C}$.) From now on we shall assume that k is algebraically closed, unless we state otherwise.

Theorem 6.2 (Hilbert's Nullstellensatz, 1893) *Let k be algebraically closed, I an ideal of $k[X]$ and q a polynomial of $k[X]$ which is zero at all points of $\mathbf{V}(I)$, i.e., $q \in \mathrm{I}\,\mathbf{V}(I)$. Then $q^s \in I$ for some integer $s > 0$.*

In concrete terms the theorem says that if q, $p_1, \ldots, p_m$ are polynomials in $k[X]$ and q vanishes whenever $p_1, \ldots, p_m$ do then there exist $s > 0$ and polynomials $q_1, \ldots, q_m$ such that

$$q^s = q_1 p_1 + \cdots + q_m p_m.$$

There are various equivalent forms of Hilbert's Nullstellensatz one of which is: $\mathbf{V}(I) = \emptyset$ if and only if $1 \in I$ (which is the same as saying that $I = k[X]$). The concrete form of this version is that a simultaneous system of polynomial equations:

$$\begin{aligned} p_1(x_1, \ldots, x_n) &= 0 \\ p_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ p_m(x_1, \ldots, x_n) &= 0 \end{aligned} \tag{10}$$

does *not* have a simultaneous solution if and only if

$$1 = q_1 p_1 + \cdots + q_m p_m$$

for some polynomials $q_1, \ldots, q_m$. This theorem does not hold if k is not algebraically closed: to see this consider $\mathbb{R}$ and the polynomial $p = x^2 + 1$. Clearly p does not have a root in $\mathbb{R}$ but $pq \neq 1$ for any polynomial q.

For a modern proof of the Nullstellensatz see Reid [51]. This book is a very readable and entertaining introduction Algebraic Geometry. For an equally entertaining read and more information (at least on certain topics) see Abhyankar [1]. (N.B. The best forms of entertainment are hard work for all the participants.) For a very down to earth proof of the Nullsetellensatz see van der Waerden [62, **Vol II**].

Exercise 6.6 *Let $p, p_1, \ldots, p_m \in k[X]$ and let z be a new indeterminate. Show that p vanishes whenever $p_1, \ldots, p_m$ do, i.e., $p \in \mathbf{IV}(p_1, \ldots, p_m)$, if and only if*

$$1 \in (p_1, \ldots, p_m, pz - 1).$$

6.2 Gröbner Bases

The preceding discussion shows that the following problem is very natural and important:

Polynomial Ideal Membership: Given polynomials q, $p_1, \ldots, p_m$ of $k[X]$ is $q \in (p_1, \ldots, p_m)$?

One algorithmic answer to this question lies in constructing certain special bases for ideals. The construction of such bases was discovered by Buchberger in the mid 1960's and he called them *Gröbner bases* in honour of his supervisor. The fact that the algorithm terminates is a consequence of Hilbert's Basis Theorem. The Nullstellensatz and the nature of Gröbner Bases then tell us that a simultaneous system of equations such as (10) has a solution if and only if the basis of the polynomials does not contain a non-zero constant.

There are now quite a few survey articles on Gröbner bases, e.g., Buchberger [12], Pauer and Pfeifhofer [48] or Mishra and Yap [45].

Before diving into the technical details we take a closer look at the problem and a possible approach to its solution. It will be convenient to let $[X]$ denote the set of all power products in the indeterminates of X. We multiply elements of $[X]$ in the obvious way. (In technical terms $[X]$ is a *monoid*, i.e., it satisfies the axioms of a group except for the requirement that each element must have an inverse. The identity of $[X]$ is $x_1^0 \cdots x_n^0$ which we denote by 1.) Let $F = \{ p_1, \ldots, p_m \}$. We want to know if $q \in (F)$ and this is so if and only if

$$q = p_1 q_1 + \cdots + p_m q_m,$$

for some $q_1, \ldots, q_m \in k[X]$. If we expand each product $p_i q_i$ by multiplying p_i with each monomial of q_i then the preceding equation is equivalent to

$$q = c_1 f_1 s_1 + \cdots + c_r f_r s_r,$$

where $c_i \in k$, $f_i \in F$, $s_i \in [X]$ for $1 \leq i \leq r$ (note that the f_i and s_i need not all be different). Let us write $g \to_F h$ to mean that $h = g - cfs$ for some $c \in k$, $f \in F$ and $s \in [X]$ (we say that g *reduces* to h). Then $q \in (F)$ if and only if there is a sequence of reductions

$$q = q_1 \to_F q_2 \to_F \cdots \to_F q_r = 0.$$

Unfortunately this is far from being an algorithm since at each step there are infinitely many choices for c and s. One way to restrict the choices is to introduce a suitable order on the power products

and then reason that as we are trying to reduce q to 0 we should avoid reductions which introduce power products which are greater than the ones we have met so far. The idea is to try to 'squeeze' q down to 0. In trying to find a reduction

$$q_i \to_F q_{i+1}$$

we therefore aim to kill at least one power product of q_i possibly at the expense of introducing some smaller ones into q_{i+1}. We therefore pick a victim power product from q_i, call it v (in fact it would be enough to choose the largest power of product of q_i). The only way to kill v *and* avoid introducing larger power products[7] is to find some $f \in F$ whose leading power product u divides v, i.e., we have $v = ut$ for some $t \in [X]$. If this is so then

$$q_i - \frac{\text{coeff}(v, q_i)}{\text{lc}(f)} ft$$

is a good reduction of q_i. Let us call a sequence of such reductions *restricted* and one of the previous type *unrestricted*. The process described so far is much more reasonable from a computational point of view: given q and F there are only finitely many possible reductions at each stage so that we have a tree. Also it turns out that we can ensure there are no infinite chains of reductions (this is *not* obvious, it is equivalent to Hilbert's Basis Theorem). So we construct the finite tree. If any leaf holds 0 then we deduce that $q \in (F)$ and otherwise $q \notin (F)$.

Let us look at two simple examples. First take $p_1 = y - 1$, $p_2 = x$ and $q = xy + x$. Then

$$\begin{aligned} xy + x &\to_{p_1} (xy + x) - x(y - 1) \\ &= 2x \\ &\to_{p_2} 2x - 2x \\ &= 0 \end{aligned}$$

so that $xy + x \in (y - 1, x)$. For the second example we take $p_1 = y + 1$, $p_2 = xy$ and $q = x$. For our order we can take any reasonable one (e.g., lexicographic with $x <_L y$), the leading power products will be y and xy respectively. It is clear that no reductions apply to q and so we deduce that $q \notin (p_1, p_2)$. Alas $q = xp_1 - p_2$ and so $q \in (p_1, p_2)$. The error is easy to find: the discussion above really does prove that if there is a restricted sequence of reductions from q to 0 then $q \in (F)$. However we did not prove the converse, i.e., that if $q \in (F)$ then there is a restricted sequence of reductions from q to 0. Of course the converse does hold if we consider the original unrestricted reductions. The reason the restricted version fails is because during the course of unrestricted reductions we could introduce power products of any order which are canceled later on. This motivates the idea of changing basis for the ideal so that we can avoid the cancellation difficulty: we want a new finite basis G of (F) with the property that if there is an unrestricted sequence of reductions (now w.r.t. G rather than F) from q to 0 then there is also a restricted sequence of reductions.

Is there any hope here? Let us first observe some simple properties of reductions.

Lemma 6.1 *Suppose $f \to_F g$ then $f \in (F)$ if and only if $g \in (F)$.*

[7] This actually assumes a property of the order relation—indeed this requirement motivates the introduction of the property.

Proof Since $f \to_F g$ we have $g = f - chs$ for some $c \in k$ and $h \in F$. Note that $shs \in (F)$. So if $f \in (F)$ then g is the difference of two members of (F) and so is in (F). Conversely $h = g + chs$, the rest is similar. □

Corollary 6.1 *Suppose* $f_1 \to_F f_2 \to_F \cdots \to_F f_n$ *then* $f_1 \in (F)$ *if and only if* $f_n \in (F)$.

Proof Induction on n and the preceding lemma. □

Note that as $0 \in (F)$ it follows that if we start with f_1 and reduce it to 0 using members of F then $f_1 \in (F)$. When can the restricted reductions idea go wrong? Let us examine the possibilities (using restricted reductions).

1. $f_1 \notin (F)$: all reduction sequences will stop at a non-zero polynomial, by the preceding observation. This is correct.
2. $f_1 \in (F)$:
   i. Reduce to 0, this is fine it proves that $f_1 \in (F)$.
   ii. Reductions stop with some $g \neq 0$; this is wrong, it implies that $f_1 \notin (F)$ in our algorithm.

Suppose we can find a basis G of (F) with the property that

$$f \in (F) \ \& \ f \neq 0 \Rightarrow \mathrm{lpp}(p) \mid \mathrm{lpp}(f), \quad \text{for some } p \in G.$$

Then case 2.ii cannot happen and the other cases stay as before; i.e., the proposed algorithm works.

Does such a G exist? Yes it certainly does: consider the ideal L generated by all the leading power products of all the elements of the ideal (F). By Hilbert's Basis Theorem, L has a finite basis M, which, by Exercise 6.5, can be assumed to consist of power products. Now take a finite subset G of (F) whose leading power products include all those of M (convince yourself that G exists—this is a simple exercise in remembering definitions). G is the required basis (this follows by a simple argument; each polynomial in the ideal reduces to 0 by members of G). The only problem that remains is to find a way of computing G from F. Roughly speaking we start with $G := F$ and systematically try to create elements of (F) whose leading power products are not divisible by the leading power product of some member of G. Whenever such an element is discovered it is put into G. We think of such an element as a counterexample to G being an appropriate basis.

How can we produce a polynomial with such a leading power product? First of all any such polynomial is a member of (G) and so must be expressible as

$$c_1 f_1 s_1 + \cdots + c_r f_r s_r,$$

where $c_i \in k$, $f_i \in G$, $s_i \in [X]$ for $1 \leq i \leq r$. If all the f_i are equal to f, say, then we do not have a counterexample since the sum is just a product of f and this reduces to 0. So at the very least we must consider expressions of the form

$$c_1 f_1 s_1 + c_2 f_2 s_2$$

where $f_1 \neq f_2$. If the leading power product of this is a product of $u_1 = \mathrm{lpp}(f_1)$ or of $u_2 = \mathrm{lpp}(f_2)$ then again we do not have a counterexample. Thus we must have $u_1 s_1 = u_2 s_2$ and $c_1 + c_2 = 0$.

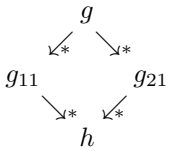
Moreover it makes sense to pick s_1, s_2 as small as possible (why?). This motivates the choice $s_1 = \operatorname{lcm}(u_1, u_2)/u_1$, $s_2 = \operatorname{lcm}(u_1, u_2)/u_2$, $c_1 = 1/\operatorname{lc}(f_1)$ and $c_2 = -1/\operatorname{lc}(f_2)$, i.e., we consider

$$\frac{1}{\operatorname{lc}(f_1)} \cdot \frac{\operatorname{lcm}(\operatorname{lpp}(f_1), \operatorname{lpp}(f_2))}{\operatorname{lpp}(f_1)} \cdot f_1 - \frac{1}{\operatorname{lc}(f_2)} \cdot \frac{\operatorname{lcm}(\operatorname{lpp}(f_1), \operatorname{lpp}(f_2))}{\operatorname{lpp}(f_2)} \cdot f_2. \qquad (\dagger)$$

This polynomial might still be reducible by f_1 or f_2 or by other members of G. So we reduce it as far as possible to produce h, say. Now if $h \neq 0$ then we have produced a new polynomial of the desired form and we put it into G. We start again by considering a new pair of distinct polynomials f_1, f_2 from F. The process is stopped when every polynomial (†) reduces to 0.

It is easy to see that this process has to stop: let $L_0, L_1, \ldots$ be the ideals generated by the leading power products of the elements of G at each stage of the process (stage 0 corresponds to the initialization $G := F$). Now we have $L_0 \subset L_1 \subset \ldots$ where the containments are strict. But Hilbert's Basis Theorem tells us that such a chain must be finite, i.e., the process has only finitely many stages. The one thing we have not shown is that it suffices to consider only expressions of the form (†), this is left to the formal part of the presentation. (The decision to take $r = 2$ seems reasonable if we think of (G) as being generated in stages. At each stage we generate all elements of the form $c_1 g_1 s_1 + c_2 g_2 s_2$ where $c_i \in k$, the g_i were generated in some previous stage and $s_i \in [X]$. It is easy to see that each element of (G) is generated by this process. The only difference from above is that we do not necessarily have $g_1, g_2 \in G$.)

There is another property that would be extremely useful. As described so far we have to consider the whole tree of possible reductions even w.r.t. G. From a computational point of view it would be much better if we could just follow one reduction to its bitter end. This amounts to requiring the following:



i.e., although we may start two different reduction sequences of g they eventually converge to the same result (this is the *Church–Rosser* property). This is indeed the case as can be seen from the discussion.

As an antidote to constructivists, and at the risk of some repetition, let us summarize the rôle that Hilbert's Basis Theorem plays in Gröbner bases. It gives us two invaluable facts. First of all it assures us that Gröbner bases exist. Secondly it assures us that the algorithm for computing them always halts. The nature of the proof of the theorem is irrelevant to these two pieces of information. Indeed the (non-constructive) proof is a rare example of depth of insight combined with extreme simplicity.

Exercise 6.7 *Consider polynomials in the single indeterminate x. A natural order on monomials is $1 < x < x^2 < x^3 < \ldots$. Given a finite set F of polynomials in $k[x]$ what does the process described above do? (Start with F having two polynomials then three, the pattern will become clear.)*

6.3 Definition and Characterization of Gröbner Bases

Throughout let

1. k be a field,

2. $k[X] = k[x_1, \ldots, x_n]$,

3. $F \subseteq k[X]$.

Hilbert's Basis Theorem tells us that $k[X]$ is Noetherian, i.e., every ideal of $k[X]$ has a finite basis (or equivalently every increasing chain of ideals of $k[X]$ terminates).

Recall that $[X]$ is the monoid of all power products in the x_i and 1 denotes the identity (ie. $x_1^0 \cdots x_n^0$). An *admissible order* on $[X]$ is one which satisfies

1. $1 < t$ for all $t \in [X] - \{1\}$,

2. if $s < t$ then $su < tu$ for all $s, t, u \in [X]$.

$([X], <)$ is an ordered monoid. Three important examples are as follows, in each case and we assume that the indeterminates have been ordered as

$$x_1 > x_2 > \cdots > x_n,$$

and set

$$s = x_1^{i_1} \cdots x_n^{i_n},$$
$$t = x_1^{j_1} \cdots x_n^{j_n},$$

Lexicographic Order (lex): $s >_{\text{lex}} t$ if and only if there is an r such that $i_l = j_l$ for $1 \leq l < r$ and $i_r > j_r$. Equivalently, if we examine $(i_1 - j_1, \ldots, i_n - j_n)$ from left to right the first non-zero entry is positive. We will normally write $s >_L t$ instead of $s >_{\text{lex}} t$.

Graded Lexicographic Order (grlex): $s >_{\text{grlex}} t$ if and only if $\deg(s) > \deg(t)$ or $\deg(s) = \deg(t)$ and $s >_L t$. We have the same equivalent condition on $(i_1 - j_1, \ldots, i_n - j_n)$ as above for ordering power products of the same degree.

Graded Reverse Lexicographic Order (grevlex): for this we have $s >_{\text{grevlex}} t$ if and only if $\deg(s) > \deg(t)$ or $\deg(s) = \deg(t)$ and there is an r such that $i_{n-l} = j_{n-l}$ for $0 \leq l < r$ but $i_{n-r} < j_{n-r}$. Equivalently, for power products of the same degree, if we examine $(i_1 - j_1, \ldots, i_n - j_n)$ from right to left the first non-zero entry is negaitive.

The second ordering is also called *total degree then lexicographic* similarly for the third one. As an example suppose that $x > y > z$ (so $n = 3$ with $x_1 = x$, $x_2 = y$ and $x_3 = z$) then the power products under the lexicographic order are ordered as:

$$1 <_L z <_L z^2 <_L \cdots <_L y <_L yz <_L yz^2 <_L \cdots$$
$$<_L y^2 <_L y^2 z <_L \cdots <_L x <_L xz <_L \cdots <_L xy <_L \cdots$$

We illustrate the second and third orders for degree 3 power products. For the graded lexicographic order we have

$$z^3 <_{\text{grlex}} yz^2 <_{\text{grlex}} y^2 z <_{\text{grlex}} y^3 <_{\text{grlex}} xyz < x^2 z <_{\text{grlex}} x^2 y <_{\text{grlex}} x^3.$$

For the graded reverse lexicographic order we have

$$z^3 <_{\text{grevlex}} yz^2 <_{\text{grevlex}} y^2 z <_{\text{grevlex}} xyz <_{\text{grevlex}} x^2 z <_{\text{grevlex}} y^3 <_{\text{grevlex}} x^2 y <_{\text{grevlex}} x^3.$$

This example shows that the graded reverse lexicographic order is *not* the same as the graded lexicographic order but with the lexicographic ordering of the power products of each degree reversed.

Lemma 6.2 *Let $\leq$ be any admissible order. Then*

1. *For all $s, t \in [X]$ we have $s \mid t \Rightarrow s \leq t$.*
2. *There are no infinite decreasing sequences (Noetherianity).*

Proof The first part is easy. The second part follows from Hilbert's Basis Theorem: suppose (for a contradiction) that $s_1 > s_2 > \cdots$ is an infinite descending sequence. Consider the ascending chain of ideals

$$(s_1) \subseteq (s_1, s_2) \subseteq \ldots.$$

We claim that in fact each containment is strict. We prove this by showing that $s_{i+1} \notin (s_1, \ldots, s_i)$ for each $i \in \mathbb{N}$. By Exercise 6.5 $s_{i+1} \in (s_1, \ldots, s_i)$ if and only if $s_j \mid s_{i+1}$ for some j with $1 \leq j \leq i$. But by the first part of this lemma this means that $s_{j+1} < s_i$ which is a contradiciton to the fact that $s_1 > s_2 > \cdots$ is a descending sequence. Thus we have the infinite strictly ascending sequence of ideals

$$(s_1) \subset (s_1, s_2) \subset \ldots.$$

This contradicts Exercise 6.3, whose proof depends on Hilbert's Basis Theorem. □

At first sight the second part of this lemma seems to contradict the example from above where we have $y >_L \cdots >_L z^2 >_L z >_L 1$. However this is not a sequence, the first element is y but what is the second element? We must choose one and this can only be of the form z^e for some $e \in \mathbb{N}$. But now the sequence is finite, namely $y >_L z^e >_L z^{e-1} \cdots >_L z^2 >_L z >_L 1$.

We introduce some more notation. Let $f \in k[X] - \{0\}$ and fix an admissible ordering. The *leading power product* of f is the highest power product (w.r.t $<$) which occurs in f with a non-zero coefficient:

$$\mathrm{lpp}(f) = \max_<\{\, t \in [X] \mid \mathrm{coeff}(t, f) \neq 0 \,\}.$$

The *leading coefficient* of f is the coefficient of the leading power product:

$$\mathrm{lc}(f) = \mathrm{coeff}(\mathrm{lpp}(f), f).$$

The *initial term* (or *initial monomial*) of f is

$$\mathrm{in}(f) = \mathrm{lc}(f)\ \mathrm{lpp}(f).$$

For $F \subseteq k[X] - \{0\}$ we put

$$\mathrm{lpp}(F) = \{\, \mathrm{lpp}(f) \mid f \in F \,\},$$
$$\mathrm{in}(F) = \{\, \mathrm{in}(f) \mid f \in F \,\}.$$

Let J be a non-zero ideal of $k[X]$. We say that a finite subset G of $J - \{0\}$ is a *Gröbner basis* for J if

$$\big(\mathrm{in}(G)\big) = \big(\mathrm{in}(J)\big),$$

(recall that by (H) we denote the ideal generated by a subset H of a ring R). We proceed to make some remarks and give some examples.

1. A finite subset G of $J - \{0\}$ is a Gröbner basis for J if and only if $\mathrm{lpp}(J) = \mathrm{lpp}(G) \cdot [X]$. This means that for every $f \in J - \{0\}$ there is a $g \in G$ such that $\mathrm{lpp}(g) \mid \mathrm{lpp}(f)$. This is perhaps the best way to think of Gröbner bases since it tells us directly that every non-zero polynomial in J is reducible by at least one polyniomial from the basis.

2. Every non-zero ideal J of $k[X]$ has a Gröbner basis. Take $M \subseteq \mathrm{in}(J)$ where M is a finite basis of $\big(\mathrm{in}(J)\big)$. Now take a finite subset G of J such that $\mathrm{in}(G) \supseteq M$. Then it is easy to see that G is a Gröbner basis for J.

3. If G is a Gröbner basis for J and $f \in J$ then $G \cup \{f\}$ is also a Gröbner basis for J.

4. G is a Gröbner basis for J if and only if $(\mathrm{lpp}(G)) = (\mathrm{lpp}(J))$. This is a consequence of the fact that k is a field. (We can define Gröbner bases when the coefficients come from certain types of rings and then the original definition is necessary.)

5. Not every basis for an ideal is a Gröbner basis. Let
$$f_1 = y + 1,$$
$$f_2 = xy,$$
and put
$$J = (f_1, f_2) \subseteq \mathbb{Q}[x, y].$$
Take $<$ to be the lexicographic order (with $x <_L y$). Now
$$x = xf_1 - f_2 \in J$$
but $x \notin \big(\mathrm{in}(f_1), \mathrm{in}(f_2)\big) = (x, xy)$. Thus $\{f_1, f_2\}$ is not a Gröbner basis for J. We saw this example on p. 66.

6. A set of monomials $\{c_1 t_1, \ldots, c_m t_m\}$ is always a Gröbner basis for the ideal which it generates.

Now we introduce the reduction relation based on a set of polynomials F (this is what we called restricted reduction above). Let $g, h \in k[X]$. We say that $g \to_F h$ if and only if there is some $f \in F$ and $s, t \in [X]$ such that $\mathrm{coeff}(s, g) \neq 0$, $s = \mathrm{lpp}(f)t$ and $h = g - \big(\mathrm{coeff}(s, g)/\mathrm{lc}(f)\big)tf$. In such a situation we say that g can be reduced to h w.r.t. F (usually F is understood from the context). Note that in the definition s is a power product of g chosen as a victim to be killed off by the reduction. In fact almost everything we do works if we take $s = \mathrm{lpp}(g)$—the only place where we need the general definition is Theorem 6.6. In examples we can simplify matters by choosing leading power products as victims, the only price we pay is that we give up the uniqueness property stated by Theorem 6.6 (in fact we can recover uniqueness by some extra computation).

For an example let
$$F = \{f_1, f_2\} \subseteq \mathbb{Q}[x, y]$$
where
$$f_1 = x^2y^2 + y - 1,$$
$$f_2 = x^2y + x.$$
For the ordering we use the lexicographic one with $x >_L y$. Now
$$\begin{aligned} 2\underbrace{x^2y^3}_{s} + x^2y + 1 &\to_{f_1} (2x^2y^3 + x^2y + 1) - 2\underbrace{y}_{t} f_1 \\ &= -2y^2 + \underbrace{x^2y}_{s} + 2y + 1 \\ &\to_{f_2} (-2y^2 + x^2y + 2y + 1) - 1 \cdot 1 \cdot f_2 \\ &= -2y^2 + 2y - x + 1, \end{aligned}$$

where the last polynomial cannot be reduced any further (w.r.t. F).

Lemma 6.3 *For every $F \subseteq k[X] - \{0\}$ the reduction relation $\to_F$ always halts, i.e., every chain of reductions $f_1 \to_F f_2 \to_F \cdots$ leads to a polynomial which cannot be reduced. (This situation usually summed up by saying that $\to_F$ is* Noetherian.*)*

Proof If $\mathrm{lpp}(f_1) = 1$ the claim is trivial for we are either stuck or reduce to 0 in one step. Let $\bar{t} > 1$ be a power product. We use induction on $<$ with the hypothesis that all polynomials with $\mathrm{lpp} < \bar{t}$ lead to finite sequences. Now consider any $f = c\bar{t} + \cdots$. If we reduce lower terms the induction hypothesis implies that we eventually stop. If we kill $\bar{t}$ then we obtain only terms which are $< \bar{t}$ and again we are done. □

We write $g \to_F^* h$ to mean that by applying zero or more reductions starting with g we eventually reach h. We write $g \to_F^* \underline{h}$ to mean that $g \to_F^* h$ and h cannot be reduced.

Lemma 6.4 *1. If $g \to_F^* h$ then*

$$g - h = \sum_{f_i \in F} p_i f_i.$$

2. For every g there is an h such that $g \to_F^* \underline{h}$.

Proof The first part follows by induction on the number of reductions. The second part is simply another way of saying that every sequence of reductions eventually halts. □

Note that $g \to_F^* 0 \Rightarrow g \in J$ but the converse need not hold (remember that F is just a basis for an ideal not necessarily a Gröbner basis).

Theorem 6.3 *Let J be an ideal of $k[X]$ and $F \subseteq J - \{0\}$. Then the following are equivalent.*

1. *F is a Gröbner basis for J.*
2. *For all $g, h \in k[X]$ if $g \in J$ and $g \to_F^* \underline{h}$ then $h = 0$.*
3. *For all $g \in J$ we have $g \to_F^* 0$.*

Proof $(1 \Rightarrow 2)$: $h \in J$ and $h = 0$ or $\mathrm{in}(h) \notin \big(\mathrm{in}(F)\big)$. But the second possibility cannot happen with a Gröbner basis and so $h = 0$.

$(2 \Rightarrow 3)$: trivial.

$(3 \Rightarrow 1)$: For $g \in J - \{0\}$, do we have $\mathrm{in}(g) \in \big(\mathrm{in}(F)\big)$? This is indeed so since by the third condition g can be reduced. Thus $\big(\mathrm{in}(J)\big) = \big(\mathrm{in}(F)\big)$. □

Corollary 6.2 *Let F be a Gröbner basis for an ideal J of $k[X]$. Then*

1. *F generates J.*
2. *$g \in J \iff g \to_F^* 0$.*

We now give a central definition in the theory of Gröbner bases. Let $f, g \in k[X] - \{0\}$. The *S-polynomial* of f and g is defined as

$$\mathrm{spol}(f,g) = \tfrac{1}{\mathrm{lc}(f)} \cdot \frac{\mathrm{lcm}(\mathrm{lpp}(f), \mathrm{lpp}(g))}{\mathrm{lpp}(f)} \cdot f - \tfrac{1}{\mathrm{lc}(g)} \cdot \frac{\mathrm{lcm}(\mathrm{lpp}(f), \mathrm{lpp}(g))}{\mathrm{lpp}(g)} \cdot g.$$

(The 'S' in S-polynomial stands for 'subtraction'.)

As an example consider the lexicographic ordering with $x <_L y$ and let

$$f = 2x^2y + 3x^2 + 1,$$
$$g = 3xy^2 - 2x.$$

Then

$$\begin{array}{ccc} & x^2y^2 & \\ \swarrow_f & & \searrow_g \\ (1/2)y(3x^2+1) & & (1/3)x(-2x) \end{array}$$

and the difference of the two new polynomials is $\mathrm{spol}(f,g)$. The two polynomials $(1/2)y(3x^2+1)$ and $(1/3)x(-2x)$ are called the *critical pair* of f and g.

Lemma 6.5 *Suppose that $f_1, \ldots, f_l \in k[X] - \{0\}$, $c_1, \ldots, c_l \in k$, $\mathrm{lpp}(f_1) = \ldots = \mathrm{lpp}(f_l) = t$ and $\mathrm{lpp}\left(\sum_{i=1}^{l} c_i f_i\right) < t$. Then $\sum c_i f_i$ is a linear combination of S-polynomials of the form $\mathrm{spol}(f_i, f_j)$ for $1 \le i, j \le l$.*

Proof Put $d_i = \mathrm{lc}(f_i)$, $f_i' = f_i/d_i$. Then

$$\sum_{i=1}^{l} c_i f_i = c_1 d_1 (f_1' - f_2') + (c_1 d_1 + c_2 d_2)(f_2' - f_3') + \cdots + \left(\sum_{i=1}^{l-1} c_i d_i\right)(f_{l-1}' - f_l') + \left(\sum_{i=1}^{l} c_i d_i\right) f_l.$$

This completes the proof since $f_{i+1}' - f_i' = \mathrm{spol}(f_i, f_{i+1})$ and $\sum_{i=1}^{l} c_i d_i = 0$. □

6.4 Computation of Gröbner Bases

Here is the really ingenious bit:

Theorem 6.4 (Buchberger's Theorem) *Let J be an ideal of $k[X]$ which is generated by a finite subset F of $k[X] - \{0\}$. Then the following are equivalent:*

1. *F is a Gröbner basis for J.*
2. *For all $f, g \in F$ we have $\mathrm{spol}(f,g) \to_F^* 0$.*

Proof $(1 \Rightarrow 2)$: immediate.

$(2 \Rightarrow 1)$: let $g \in J - \{0\}$ we show that $\mathrm{in}(g) \in \big(\mathrm{in}(F)\big)$. Put

$$g = \sum_{i=1}^{m} c_i t_i f_i, \quad f_i \in F,\ t_i \in [X], \quad c_i \in K - \{0\}. \qquad (*)$$

Consider the highest power products t_i in the sum. Put

$$u = \max_{<}\{\, t_i \operatorname{lpp}(f_i) \mid 1 \le i \le m \,\},$$

$$\bar{g} = \sum_{t_i \operatorname{lpp}(f_i) = u} c_i t_i f_i.$$

Now choose a representation $(*)$ of g which makes u minimal.

Suppose that $\operatorname{lpp}(\bar{g}) < u$. By the preceding lemma we have

$$\bar{g} = \sum d_{ij} \operatorname{spol}(t_i f_i, t_j f_j).$$

But all the power products in each $\operatorname{spol}(t_i f_i, t_j f_j)$ are $< u$ and this is a contradiction to the minimality of u. Thus $\operatorname{lpp}(\bar{g}) = u$ and $\operatorname{in}(g) = \operatorname{in}(\bar{g}) \in \big(\operatorname{in}(F)\big)$. It now follows that F is a Gröbner basis for J. □

The basic algorithm is as follows (here the phrase 'compute a normal form of g w.r.t. G' means that we reduce g w.r.t G as much as possible—it doesn't matter which of the many choices we make at each reduction step).

Algorithm: *GRÖBNER_BASIS*$(F) \mapsto G$

(F and G are finite sets of polynomials, $(F) = (G)$ and G is a Gröbner basis for (F).)

$G := F$;

while not all S-polynomials of G have been considered **do**

choose a new $\operatorname{spol}(f, g)$;

compute a normal form h of it w.r.t G;

if $h \neq 0$ **then** $G := G \cup \{h\}$ **fi**

od

This may be elaborated as:

Algorithm: *GRÖBNER_BASIS*$(F) \mapsto G$

(F and G are finite sets of polynomials, $(F) = (G)$ and G is a Gröbner basis for (F).)

let $F = \{\, f_1, f_2, \ldots, f_m \,\}$;

$P := \{(f_i, f_j) \mid 1 \le i < j \le m\}$;

$G := F$;

while $P \neq \emptyset$ **do**

remove the first pair (f, g) from P;

compute a normal form h of $\operatorname{spol}(f, g)$ w.r.t. G;

if $h \neq 0$ **then**

$P := P \cup \{(h, p) \mid p \in G\}$;

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

fi

od

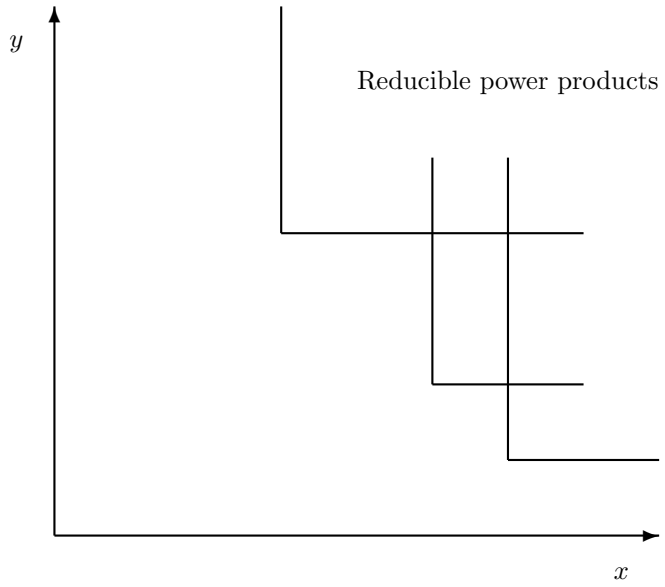


Figure 1: Progress of *GRÖBNER_BASIS* with indeterminates x, y.

An indication of the algorithm's progress can be seen from the diagram of figure 1 for the case of two indeterminates x and y. As more and more h's are reduced we cut off more regions and so we eventually stop (the region left over could be infinite).

For an example let

$$f_1 = x^2y^2 + y - 1,$$
$$f_2 = x^2y + x,$$

and take the lexicographic order with $x >_L y$. Thus initially we have $G = \{f_1, f_2\}$. The following describes one way in which the algorithm could execute (recall that we have a free choice for spol(f, g) in the loop):

1. spol$(f_1, f_2) = f_1 - yf_2 = -xy + y - 1$. This does not reduce further so put

$$f_3 = -xy + y - 1,$$
$$G = \{f_1, f_2, f_3\}.$$

2. spol$(f_2, f_3) = f_2 + xf_3 = xy \to_{f_3} y - 1$. We put

$$f_4 = y - 1,$$
$$G = \{f_1, f_2, f_3, f_4\}.$$

3. spol$(f_3, f_4) = f_3 + xf_4 = -x + y - 1 \to_{f_4} -x$. We put

$$f_5 = -x,$$
$$G = \{f_1, f_2, f_3, f_4, f_5\}.$$

4. $\text{spol}(f_1, f_3) = f_1 + xyf_3 = xy^2 - xy + y - 1 \rightarrow_G 0$.

5. $\text{spol}(f_2, f_4) = f_2 - x^2 f_4 = x^2 + x \rightarrow_G 0$.

The algorithm is now finished with

$$G = \{x^2y^2 + y - 1,\, x^2y + x,\, -xy + y - 1,\, y - 1,\, -x\}.$$

In fact $\{x,\, y - 1\}$ is a Gröbner basis for $\{f_1, f_2\}$.

Theorem 6.5 *Let G be a Gröbner basis for an ideal I of $k[X]$. Let $g, h \in G$ with $g \neq h$. Then*

1. *If $\text{lpp}(g) \mid \text{lpp}(h)$ then $G' = G - \{h\}$ is also a Gröbner basis for I.*

2. *If $h \rightarrow_G h'$ then $G' = (G - \{h\}) \cup \{h'\}$ is also a Gröbner basis for I.*

Proof For the first part we note that $(G') \subseteq I$. For $f \in I$ we have $f \rightarrow_G^* 0$ but in fact we have $f \rightarrow_{G'}^* 0$. (N.B. we cannot throw polynomials out *during* the computation of a Gröbner basis.)

For the second part note that $(G') = (G)$. If $\text{lpp}(h)$ is reduced then the result follows from the first part. If some other power product is reduced then $\big(\text{lpp}(G')\big) = \big(\text{lpp}(G)\big) = \big(\text{lpp}(I)\big)$. □

Suppose that G is a Gröbner basis. We say that

1. G is *minimal* if and only if $\text{lpp}(g) \not| \text{lpp}(h)$ for all $g, h \in G$ with $g \neq h$.

2. G is *reduced* if and only if for all $h, g \in G$ with $h \neq g$, h cannot be reduced by g.

3. G is a *normed* basis if $\text{lc}(f) = 1$ for all $f \in G$.

Theorem 6.6 *A normed reduced basis for an ideal I is unique.*

Proof Suppose G, G' are normed reduced bases for I. Let

$$G = \{g_1, \ldots, g_m\},$$
$$G' = \{g'_1, \ldots, g'_{m'}\}.$$

Now $g_1 \rightarrow_{G'}^* 0$, in particular $\text{lpp}(g_1)$ can be reduced by G' and so w.l.o.g. $\text{lpp}(g'_1) \mid \text{lpp}(g_1)$. Also $g'_1 \rightarrow_G^* 0$ and again $\text{lpp}(g_k) \mid \text{lpp}(g'_1)$ for some k. But $\text{lpp}(g'_1) \mid \text{lpp}(g_1)$ and so $k = 1$. Thus $\text{lpp}(g_1) = \text{lpp}(g'_1)$. In his way we obtain $m = m'$ and $\text{lpp}(g_i) = \text{lpp}(g'_i)$ for $1 \leq i \leq m$.

Now consider any g_i. We have $g_i \rightarrow_{G'}^* 0$, suppose that $g'_i \neq g_i$. The only way to kill $\text{lpp}(g_i)$ is to use g'_i. Now $g_i - g'_i \neq 0$, but none of the power products in this polynomial can be reduced w.r.t. G' and so $g_i \not\rightarrow_{G'}^* 0$ which is a contradiction. □

We make a couple of observations:

1. The size of a Gröbner basis depends on the ordering that we use.

2. For a field k and ideal I of $k[X]$ we have that $k[X]/I$ is a vector space over k. ($k[X]/I$ is called a quotient ring; roughly speaking we use polynomials as always but now we deem two polynomials to be equal whenever their difference in in I. You can find a description of this construction in any introductory book on ring theory). We can compute with this by working with a Gröbner basis G and using irreducible elements w.r.t. G as representatives of equivalence classes. To obtain a basis we just take the irreducible power products. Thus $\dim_k k[X]/I$ is the number of such irreducible power products and so this number is independent of the Gröbner basis used.

6.5 Applications of Gröbner Bases

Many applications are given by Buchberger [12]. Here we look at just two applications.

Ideal Membership

Given an ideal $I = (f_1, \ldots, f_s)$ and a polynomial f is $f \in I$? Compute a Gröbner basis G for I. Then

$$f \in I \iff f \to_G^* 0.$$

Solution of Equations by Gröbner Bases

We know from Hilbert's Nullstellensatz that a system of polynomial equations (over an algebraically closed field such as $\mathbb{C}$) is inconsistent (i.e., has no solution) if and only if 1 is in the ideal that they generate. But it is obvious that 1 is a member of an ideal if and only if its normed Gröbner basis is just $\{1\}$. Alternatively we can just compute a Gröbner basis which is not necessarily normed or reduced and see if it contains a non-zero constant.

A system of polynomial equations over some field k might have only finitely many or infinitely many solutions. The following result gives us a method of deciding between the two possibilities (when k is algebraically closed).

Theorem 6.7 *Let G be a Gröbner basis for I where I is an ideal of $k[X]$ and k is algebraically closed. Then $\mathbf{V}(I)$ is finite if and only if for each x_i, $1 \leq i \leq n$, there is an $e_i \geq 0$ such that $x_i^{e_i} \in \mathrm{lpp}(G)$.*

For example consider

$$f = y^3 + x^2 + 2xy,$$
$$g = x^2 + y^2 - 1.$$

If we use Axiom's `groebner` with total degree then reverse lexicographic order with $x > y$ (type `HDMP([y,x],INT))`[8]) then a Gröbner basis for (f, g) is:

$$T = \{y^3 - y^2 + 2xy + 1, y^2 + x^2 - 1\},$$

(this is normed and reduced but the Theorem does not require the extra property). Now

$$\mathrm{lpp}(T) = \{y^3, x^2\}$$

which shows that f, g have only finitely many common zeros (over $\mathbb{C}$, say). Note however that T is not particularly useful if we want to *find* the common zeros. For this it is much better to use the lexicographic order since then

$$L = \{2x - y^5 + 2y^4 - 5y^3 - y^2 + 5y, y^6 - 2y^5 + 5y^4 + 2y^3 - 6y^2 + 1\},$$

is the normed reduced Gröbner basis for (f, g) (assuming this time that $x >_L y$). Note that we have one polynomial in y alone. We can solve this (either approximately or using more sophisticated symbolic techniques) and then substitute each solution into the other polynomial to find the corresponding value of x.

[8] `HomogeneousDistributedMultivariatePolynomial([x,y],Integer))` in full!

Let us examine the general situation. Suppose that we have a system:

$$\begin{aligned} p_1(x_1, \ldots, x_n) &= 0, \\ p_2(x_1, \ldots, x_n) &= 0, \\ &\vdots \\ p_m(x_1, \ldots, x_n) &= 0, \end{aligned}$$

which we know (from other considerations) has only finitely many solutions, say N of them. We might as well assume that $N > 0$ (i.e., there is at least one solution). If we diagonalize the system then we are in a much better position to solve it, i.e., we want to replace it with an equivalent system that looks like:

$$\begin{aligned} q_1(x_1, x_2, x_3 \ldots, x_n) &= 0, \\ q_2(x_2, x_3 \ldots, x_n) &= 0, \\ q_3(x_3 \ldots, x_n) &= 0, \\ &\vdots \\ q_n(x_n) &= 0, \end{aligned}$$

where $\mathrm{lpp}(q_i) = x_i^{e_i}$ for some $e_i > 0$. (In practice we might have some extra equations which serve to rule out certain solutions, so it would be more accurate to say that part of the new system looks like the above.) We can solve the diagonalized system since the last equation gives finitely many values for x_n. We can substitute each of these values in the penultimate equation to obtain finitely many corresponding values for x_{n-1} (the fact that $\mathrm{lpp}(q_{n-1})$ involves only x_{n-1} means that even after substituting a value for x_n we are left with a non-constant polynomial in x_{n-1}). Iterating this process gives us all the solutions to the system.

Exercise 6.8 *Give an upper bound on the number of solutions to a diagonalized system in terms of the degrees of the polynomials that occur in it.*

How can we produce a diagonalized system? First of all suppose that we have a set F of polynomials, as above, with only finitely many solutions:

$$P_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in}),$$

where $1 \leq i \leq N$. Note that each of the polynomials

$$f_j = (x_j - \alpha_{1j})(x_j - \alpha_{2j}) \ldots (x_j - \alpha_{Nj}),$$

for $1 \leq j \leq n$, vanishes at all the solutions of the system F. Assuming that k is algebraically closed, it follows from the Nullstellensatz that some power of f_j is in the ideal (F). We conclude from this that for each x_i the ideal (F) contains non-constant polynomials which involve no other indeterminates. Let us order $[X]$ lexicographically with $x_1 > x_2 > \cdots > x_n$. Now let G be a Gröbner basis for (F) w.r.t. the lexicographic ordering. We know that (G) has a polynomial $h_n(x_n) \in k[x_n]$. By the defining property of Gröbner bases there must be a $g \in G$ such that

$$\mathrm{lpp}(g) \mid \mathrm{lpp}(h_n). \qquad (*)$$

Now if g involves any indeterminate other than x_n then so does $\mathrm{lpp}(g)$ (because of the lexicographic ordering) in which case $(*)$ cannot hold (since $\mathrm{lpp}(h_n)$ cannot involve any indeterminate other than x_n). Thus G has a polynomial from $k[x_n]$ which is not in k (the polynomial cannot be in k for then the system has no solutions at all). Of course the leading power of this polynomial is of the form $x_n^{e_n}$ for some $e_n > 0$.

We can repeat the preceding argument with x_{n-1} in place of x_n and h_{n-1} in place of h_n. In this case the conclusion is that g cannot involve any indeterminates other than x_{n-1} and x_n (the occurrence of x_n does not prevent $\mathrm{lpp}(g)$ from involving only x_{n-1}). We therefore conclude that G also has a polynomial g which is in $k[x_{n-1}, x_n]$ but not $k[x_n]$ (g cannot be in $k[x_n]$ for then it is either a constant, in which case the system has no solutions, or $\mathrm{lpp}(g)$ involves x_n and so cannot divide a power product of the form x_{n-1}^m). Clearly we must have $\mathrm{lpp}(g) = x_{n-1}^{e_{n-1}}$ for some $e_{n-1} > 0$. Obviously we can continue this argument down to considering x_1 in place of x_n.

What we have given is an outline proof of:

Theorem 6.8 *A system of polynomial equations has finitely many solutions over an algebraically closed field (e.g., $\mathbb{C}$) if and only if each indeterminate appears in the form $x_i^{e_i}$, where $e_i \geq 0$, as the initial term of one of the members of the normed reduced Gröbner basis of the polynomials where the basis is computed w.r.t. a lexicographic ordering. (If the system has at least one solution then $e_i > 0$ for each i.) Moreover such a basis includes a diagonalized set of polynomials.*

Note that the situation when there are no solutions is covered because then the normed reduced Gröbner basis is just $\{1\}$. The leading power product of 1 is just 1 and this is x^0 for any indeterminate x.

Exercise 6.9 *Show that the Theorem need not hold if k is not algebraically closed.*

We can phrase most of the preceding discussion purely in terms of ideals, without any assumption about the corresponding variety, as follows (also we do *not* assume that k is algebraically closed because we do not use the Nullstellensatz). Let I be an ideal of $k[X]$ and set

$$I_j = I \cap k[x_j, x_{j+1}, \ldots, x_n],$$

for $1 \leq j \leq n$. Note that I_j is an ideal of $k[x_j, x_{j+1}, \ldots, x_n]$, it is called the jth *elimination ideal* of I. Observe that $I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n$ and so if $I_r = 0$ then $I_{r+1} = \ldots = I_n = 0$. Moreover every non-zero member of I_n is a polynomial in x_n alone, every non-zero member of I_{n-1} is a polynomial in x_{n-1}, x_n etc. Now let G be a Gröbner basis for I w.r.t. the lexicographic ordering with $x_1 > x_2 > \cdots > x_n$. Then, as above, we see that whenever $I_j \neq 0$ there must be a member h of G which is in $k[x_j, x_{j+1}, \ldots, x_n]$. Also if $I_j \not\subseteq k[x_{j+1}, \ldots, x_n]$ (i.e., I_j has elements which really involve x_j) then $h \notin k[x_{j+1}, \ldots, x_n]$ provided that $I \neq k[X]$. Moreover if we set

$$G_j = G \cap k[x_j, x_{j+1}, \ldots, x_n],$$

for $1 \leq j \leq n$, then it is a simple matter to see that G_j is a Gröbner basis for I_j whenever $I_j \neq 0$. It now follows that if $I_j \neq 0$ then

$$I_j = (G \cap k[x_j, x_{j+1}, \ldots, x_n]),$$

and we can remove the assumption on I_j provided we agree that the empty set generates the zero ideal. In summary a Gröbner basis computed w.r.t. a lexicographic ordering produces a generating set for the ideal which is as diagonalized as possible.

6.6 Improvements of the Basic Gröbner Basis Algorithm

The basic algorithm can be optimized in various ways. The reduction of S-polynomials is a very costly part of the algorithm so it is worthwhile avoiding as many of these as possible. Buchberger [10] shows that if the basis constructed so far has an element h such that $\mathrm{lpp}(h) \mid \mathrm{lcm}(\mathrm{lpp}(f), \mathrm{lpp}(g))$ and if we have already considered $\mathrm{spol}(f,h)$ and $\mathrm{spol}(h,g)$ then we need not consider $\mathrm{spol}(f,g)$ because this will reduce to zero. We say that f, g are *connected*. (The situation can be generalized further. There are also other criteria, e.g., if $\mathrm{lcm}(\mathrm{lpp}(f_1), \mathrm{lpp}(f_2)) = \mathrm{lpp}(f_1)\,\mathrm{lpp}(f_2)$ then $\mathrm{spol}(f_1, f_2) \to_F^* 0$.)

We can now give an improved version of the Gröbner basis algorithm.

Algorithm: $rGB(F) \mapsto G$

$G := F$;
$C := \{(\{g_1, g_2\}, p) \mid g_1, g_2 \in G, g_1 \neq g_2, p = \mathrm{lcm}(\mathrm{lpp}(g_1), \mathrm{lpp}(g_2))\}$;
while $C \neq 0$ **do**
  $(\{g_1, g_2\}, p) :=$ an element from C;
  if (g_1, g_2 are not connected) **then**
    $h :=$ a normal form of $\mathrm{spol}(g_1, g_2)$ w.r.t. G;
    if $h \neq 0$ **then**
      $C := C \cup$ (new critical pairs from h);
      $G := G \cup \{h\}$
    fi
  fi
od

Let us look at an example from $\mathbb{Z}_2[x,y,z]$. Put

$$F = \{f_1, f_2, f_3\}$$

where

$$\begin{aligned} f_1 &= x^3yz + xz^2, \\ f_2 &= xy^2z + xyz, \\ f_3 &= x^2y^2 + z^2. \end{aligned}$$

Initially

$$C := \{(\{f_1, f_2\}, x^2y^2z), (\{f_1, f_3\}, x^3y^2z), (\{f_2, f_3\}, x^2y^2z)\}.$$

We start with the least p, i.e., x^2y^2z and obtain

$$f_4 = x^2yz + z^3.$$

We add f_4 to the basis, remove $(\{f_2, f_3\}, x^2y^2z)$ from C and add new pairs to C. Next we obtain from f_1, f_4

$$f_5 = xz^3 + xz^2.$$

Consideration of f_2, f_4 yields

$$f_6 = yz^3 + z^3.$$

Now consider $\{f_3, f_4\}$:

$$\begin{array}{ccc} & x^2y^2z & \\ \swarrow_{f_3} & \downarrow_{f_2} & \searrow_{f_4} \\ \leftrightarrow^* & & \leftrightarrow^* \end{array}$$

We have already considered f_2, f_3 and f_2, f_4 and the lcm's 'fit' so there is no need to reduce $\{f_3, f_4\}$.

Completion leads to 9 polynomials. The basic algorithm considers $\binom{9}{2} = 36$ reductions of S-polynomials. As observed above it is these reductions which tend to be the most expensive part of the algorithm.

6.7 Complexity of Computing Gröbner Bases

The method of Gröbner bases is clearly a very powerful technique and so it is important to have implementations which are as efficient as possible. This raises the question of the inherent complexity of computing Gröbner bases. Unfortunately the worst case *space complexity* is double exponential in the number of indeterminates (although the behaviour is much better in practice).

The ingredients for a Gröbner basis are a finite set G of multivariate polynomials (usually with coefficients from $\mathbb{Q}$) and a suitable total order on the power products. It is a well known observation that, for a given set G, the runtime for a total degree order (i.e., power products are sorted first according to degree and then by some other criterion, especially reverse lexicographic) is usually better than for a lexicographic one and frequently it is dramatically better. In this section we show that, for a class of examples, this behaviour is explained by a conjecture in Complexity Theory. See Bayer and Stillman [4] or Eisenbud [23] for special properties of the reverse lexicographic order (however these references are quite advanced). We will also produce examples of polynomials such that any Gröbner basis for them under a lexicographic order is exponentially larger than the input polynomials.

It is an easy exercise to encode NP-complete problems in terms of Gröbner bases. For example, given an instance of Satisfiability we can produce a set of equations such that the given formula is satisfiable if and only if the equations have a common zero in some algebraically closed field (in fact the encoding ensures that any solution will have components from $\{0, 1\}$). We test the last condition by computing a Gröbner basis for the polynomials and checking to see if it has a nonzero constant. Although this gives us a hint that in the worst case Gröbner bases will be hard to compute, such an approach does not help to explain the difference in runtimes between total degree and lexicographic orders. This suggests that we should consider problems for which solutions are known to exist but for which some other property is believed to be intractable. In this note we focus on #P-complete problems; see Papadimitriou [50] for background. Although this class includes the counting versions of NP-complete problems (e.g., Satisfiability) it also includes very restricted versions of #Monotone Satisfiability (prefixing a decision problem with # indicates that we are considering its counting version). We show how to encode efficiently one of these problems in such a way that finding the Gröbner basis under a total degree order costs no more than the encoding while even partial information about the Gröbner basis under a lexicographic order would amount to solving the #P-complete problem. In fact we also show that, without any assumptions on #P, it is quite easy to produce examples where the difference in runtimes is exponential in the size of the input (the second basis is exponentially larger than the first).

Becker and Weispfenning [5] include a brief discussion on complexity issues in an Appendix. Here we note that Möller and Mora [47] and Huynh [33] show that in the worst case Gröbner bases have polynomials whose degree is $\Omega(d^{2^n})$ where d is the maximum of the degrees of the inputs and n is the number of indeterminates; their proofs are based on work of Mayr and Meyer [43]. Moreover Huynh [33] proves that the same holds for the cardinality of Gröbner bases. For zero-dimensional ideals (i.e., ones with finitely many zeros) the situation is not so bad. Lakshman [42] shows that for polynomials with rational coefficients with finitely many common zeros the cost of computing their Gröbner basis under any admissible ordering is bounded by a polynomial in d^n where d, n are as above. The ideals we use are all zero dimensional and indeed it is easy to see how to obtain their Gröbner basis in $O(2^n)$ time for the orderings under consideration.

6.7.1 Algebraic Preliminaries

Throughout k will be a field and $X = \{ x_1, \ldots, x_n \}$ a nonempty set of distinct indeterminates over k. For each $f \in k[X]$ we set $\overline{f} = 1 - f$ and note that $\overline{\overline{f}} = f$. We also set

$$\mathcal{R} = \{y_1 y_2 \cdots y_n | y_i \in \{ x_i, \overline{x_i} \}, \text{for } 1 \leq i \leq n\},$$
$$S = \{ x_1^2 - x_1, x_2^2 - x_2, \ldots, x_n^2 - x_n \}.$$

Let I be an ideal of $k[X]$ that contains S as a subset. Since $x_i^2 - x_i = \overline{x}_i^2 - \overline{x}_i = x_i \overline{x}_i$ it is clear that for all $m_1, m_2 \in \mathcal{R}$ we have

$$m_1 m_2 \equiv \begin{cases} m_1 \pmod{I}, & \text{if } m_1 = m_2; \\ 0 \pmod{I}, & \text{if } m_1 \neq m_2. \end{cases}$$

(The notation $a \equiv b \pmod{I}$ simply means that $a - b \in I$.) It is now easy to see that the members of $\mathcal{R}$ are linearly independent over k (consider their images in $k[X]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$). Moreover every power product $x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n}$ when considered modulo I can be written as a unique linear combination of the members of $\mathcal{R}$ (if $e_i > 1$ then $x_i^{e_i} \equiv x_i \pmod{I}$, while if $e_i = 0$ then multiply by $x_i + \overline{x}_i$ and expand).

Lemma 6.6 *Let I be an ideal of $k[X]$ that contains S as a subset. Let $f \in k[X]$ and set $f \equiv \sum_{i=1}^{r} a_i m_i \pmod{I}$ where each $a_i \in k^*$ and $m_i \in \mathcal{R}$. Then $f \in I$ if and only if $m_i \in I$ for $1 \leq i \leq n$.*

Proof Suppose that $f \in I$. Then for each i we have $a_i^{-1} m_i f \in I$. However $a_i^{-1} m_i f \equiv m_i \pmod{I}$. The converse is immediate. □

Lemma 6.7 *Let I be an ideal of $k[X]$ that contains S as a subset. Then I is a radical ideal, i.e., $f^s \in I$ for some $s > 0$ if and only if $f \in I$.*

Proof Set $f \equiv \sum_{i=1}^{r} a_i m_i \pmod{I}$ where each $a_i \in k^*$ and $m_i \in \mathcal{R}$. Then, from the observations made above, we have $f^s \equiv \sum_{i=1}^{r} a_i^s m_i \pmod{I}$ and the result follows from the preceding lemma. □

Suppose now that k has characteristic 0 so that it contains $\mathbb{Q}$ as a subfield. Let α be the endomorphism of $k[X]$ induced by $x_1 \mapsto x_1 - 2x_2 - 2^2 x_3 - \cdots - 2^{n-1} x_n$ and $x_i \mapsto x_i$, for $2 \leq i \leq n$. Clearly α is an automorphism of $k[X]$.

Lemma 6.8 *Let I be an ideal of $k[X]$ that contains S as a subset and assume that k has characteristic 0. Then $\alpha(I)$ is a radical ideal of $k[X]$. Furthermore if $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ are zeros of $\alpha(I)$ with $a_1 = a_2$ then $a_i = b_i$ for $2 \leq i \leq n$.*

Proof The fact that α is an automorphism implies that $\alpha(I)$ is an ideal. Now $f^s \in \alpha(I)$ if and only if $\alpha^{-1}(f)^s \in I$ and so $\alpha(I)$ is radical by Lemma 6.7.

The last part follows from the observation that $(c_1, c_2, \ldots, c_n) \in \{0, 1\}^n$ is a zero of I if and only if $(c_1 + 2c_2 + \cdots + 2^{n-1}c_n, c_2, \ldots, c_n)$ is a zero of $\alpha(I)$. □

In the following we will use $\mathbf{V}(I)$ to denote the set of common zeros of an ideal I. The next lemma draws together some well known facts about radical zero-dimensional ideals, e.g., see Becker and Weispfenning [5], Chapter 8.

Lemma 6.9 *Let I be an ideal of $k[X]$ that contains S as a subset and assume that k has characteristic 0. Then $\alpha(I) \cap k[x_1] \neq 0$. Furthermore let p_1 be a nonzero monic element of $\alpha(I) \cap k[x_1]$ of minimal degree and set $d = \deg(p)$. Then*

1. *$|\mathbf{V}(I)| = d$.*
2. *$p_1 = (x_1 - \xi_1)(x_1 - \xi_2)\cdots(x_1 - \xi_d)$ where $\xi_1, \xi_2, \ldots, \xi_d$ are the x_1-coordinates of all the elements of $\mathbf{V}(\alpha(I))$. In particular p_1 has integer coefficients.*
3. *There are polynomials $p_2, \ldots, p_n \in \mathbb{Q}[x_1]$ such that $x_2 - p_2, \ldots, x_n - p_n \in \alpha(I)$ and either $p_i = 0$ or $\deg(p_i) < d$ for $2 \leq i \leq n$.*

Proof Note that $\mathbf{V}(I) \subseteq \{0, 1\}^n$ and is unchanged if we enlarge k and so we may assume that k is algebraically closed. Clearly $|\mathbf{V}(I)| = |\mathbf{V}(\alpha(I))|$. Let $\xi_1, \xi_2, \ldots, \xi_r$ be the x_1-coordinates of all the elements of $\mathbf{V}(\alpha(I))$. By Lemma 6.8, we have $r = |\mathbf{V}(I)|$. Now the polynomial $q = (x_1 - \xi_1)(x_1 - \xi_2)\cdots(x_1 - \xi_r)$ vanishes on $\mathbf{V}(\alpha(I))$ and so by Hilbert's Nullstellensatz it lies in the radical of $\alpha(I)$. Hence $q \in \alpha(I)$ since $\alpha(I)$ is radical by Lemma 6.8. Thus, by the definition of d, we have $d \leq r$. Since p_1 vanishes on $\mathbf{V}(\alpha(I))$ we have $x_i - \xi_i \mid p_1$, for $1 \leq i \leq r$; therefore $d \geq r$ and so $d = r$. Thus $p_1 = q$.

Let $(a_{1i}, a_{2i}, \ldots, a_{ni})$ for $1 \leq i \leq d$ be the zeros of $\alpha(I)$. By interpolation we can find, for each j with $2 \leq j \leq n$, a polynomial $p_i \in \mathbb{Q}[x_1]$ such that $p_j(a_{1i}) = a_{ji}$ and either $p_j = 0$ or $\deg(p_j) < d$. Now $x_j - p_j$ vanishes on $\mathbf{V}(\alpha(I))$ and so $x_j - p_j \in \alpha(I)$. □

This lemma provides a simple proof of the fact that the size of a Gröbner basis can be exponentially larger than the size of the input. For example consider the set S. This has size $O(n)$ (we assume the use of a sparse distributed representation as the data structure for multivariate polynomials, although this is not critical for our example). Let B be a Gröbner basis of $\alpha(S)$ under a lexicographic order in which x_1 is least. Then, by the elimination property of Gröbner bases using lexicographic order, B contains a non-zero constant multiple of the polynomial p_1 of Lemma 6.9. The roots of p_1 are precisely $0, 1, \ldots, 2^{n-1}$ so that p_1 has degree 2^n; however this does not mean that the size of p_1 is exponential in n since it is possible that when expanded it is sparse. In fact when expanded all coefficients of p_1 except for the constant term are nonzero. This follows from the fact that the coefficients are obtained by evaluating the elementary symmetric functions in n variables at the roots of p_1 and these roots are all strictly positive except for one which is 0.

Fom now on let G be a subset of $k[X]$ such that $S \subseteq G$. Define β to be the endomorphism of $k[X]$ induced by $x_1 \mapsto x_1^2$ and $x_i \mapsto x_i$, for $2 \leq i \leq n$. Note that β is a injective but not surjective. Set

$$I = (G), \quad J = \alpha(I), \quad K = (\beta(J)).$$

Lemma 6.10 *Let $I = (G)$ as above and assume that k has characteristic 0. Then $K \cap k[x_1] \neq 0$. Furthermore, let p_1 be a nonzero monic element of $K \cap k[x_1]$ of minimal degree and set $d = \deg(p_1)$. Then*

1. *$|\mathbf{V}(I)| = d/2$.*
2. *$p_1 = (x_1^2 - \xi_1)(x_1^2 - \xi_2) \cdots (x_1^2 - \xi_r)$ where $\xi_1, \xi_2, \ldots, \xi_r$ are the x_1-coordinates of all the elements of $\mathbf{V}(J)$.*
3. *There are polynomials $p_2, \ldots, p_n \in \mathbb{Q}[x_1^2]$ such that $x_2 - p_2, \ldots, x_n - p_n \in K$ and either $p_i = 0$ or $\deg(p_i) < d - 1$ for $2 \leq i \leq n$.*

Proof As in Lemma 6.9, we may assume that k is algebraically closed. Let $\xi_1, \xi_2, \ldots, \xi_r$ be the x_1-coordinates of all the elements of $\mathbf{V}(J)$ and set $q(x_1) = (x_1 - \xi_1)(x_2 - \xi_2) \cdots (x_1 - \xi_r)$. By Lemma 6.9, $q(x_1) \in J$ and so $q(x_1^2) \in K$. Thus $d \leq 2r$.

We claim that $x_1^2 - \xi_i \mid p_1$, for $1 \leq i \leq r$. Consider $(a_1, a_2, \ldots, a_n) \in \{0, 1\}^n$ and set $A = \sum_{i=1}^n a_i 2^{i-1}$. Now $(A, a_2, \ldots, a_n)$ is a zero of J if and only if $(B, a_2, \ldots, a_n)$ is a zero of K for all B such that $B^2 = A$. Since k has characteristic 0 and is algebraically closed it follows that each nonzero element of k has exactly two distinct square roots in k. Since p_1 vanishes at each zero of K it follows that $x_1^2 - \xi_i \mid p_1$ whenever $\xi_i \neq 0$. Now if $(0, 0, \ldots, 0)$ is not zero of J then $\xi_i \neq 0$, for $1 \leq i \leq n$, and the proof of the claim is complete. On the other hand if $(0, 0, \ldots, 0)$ is a zero of J then exactly one ξ_i is equal to 0. However we have $x_1^2 \mid p_1$ since no element of G can have a nonzero constant term and so sending x_i to 0, for $2 \leq i \leq n$, leaves p_1 fixed and sends K to $(x_1^4 - x_1^2)$ or to $(x_1^4 - x_1^2, x_1^2) = (x_1^2)$. Thus $d \geq 2r$ and $p_1 = q(x_1^2)$ as claimed.

The final part follows from the last part of Lemma 6.9. □

Note that K need not be radical, however it is 'nearly' so. The next lemma clarifies the situation (we will not need this result subsequently).

Lemma 6.11 *Suppose that k has characteristic 0. Then*

1. *K is radical if and only if G contains a polynomial with a nonzero constant term.*
2. *Suppose that $f(x_1^2, x_2, \ldots, x_n)^s \in K$ for some $s > 0$. Then $f(x_1^2, x_2, \ldots, x_n) \in K$.*

Proof For the first part we use a result given by Becker and Weispfenning [5] as Proposition 8.14 (based on a Lemma of Seidenberg): If k is perfect, then a zero-dimensional ideal is radical if and only if it contains a univariate squarefree polynomial in each indeterminate. In our case K contains $x_i^2 - x_i$, for $2 \leq i \leq n$, which are squarefree. Thus K is radical if and only if the generator of $K \cap k[x_1]$ is squarefree, i.e., if and only if the polynomial $p(x_1)$ of Lemma 6.10 is squarefree. This is so if and only if $(0, 0, \ldots, 0)$ is not a zero of J and this is equivalent to the stated condition on G.

For the second part, if $f(x_1^2, x_2, \ldots, x_n)^s \in K$ for some $s > 0$ then there are polynomials $f_1, f_2, \ldots, f_r \in k[X]$ and $g_1, g_2, \ldots, g_r \in G$ such that $f(x_1^2, x_2, \ldots, x_n)^s = f_1\beta(g_1) + \cdots + f_r\beta(g_r)$.

Set $f_i = f_{i0} + f_{i1}$, for $1 \leq i \leq r$, where each term of f_{i0} has even degree in x_1 and each term of f_{i1} has odd degree in x_1. Since each term in $\beta(g_i)$, for $1 \leq i \leq r$, has even degree in x_1 it follows that $f(x_1^2, x_2, \ldots, x_n)^s = f_{10}\beta(g_1) + \cdots + f_{r0}\beta(g_r)$. Now replacing x_1 by $x_1^{1/2}$ we see that $f(x_1, x_2, \ldots, x_n)^s \in J$ and so $f(x_1, x_2, \ldots, x_n) \in J$ since J is radical, by Lemma 6.7. Thus $f(x_1^2, x_2, \ldots, x_n) \in K$ as claimed. □

Lemma 6.12 *Let L be an ideal of $k[X]$ and suppose that there are $p_1, p_2, \ldots, p_n \in k[x_1]$ such that $p_1, x_2 - p_2, \ldots, x_n - p_n \in L$. Then, provided that p_1 is of minimal degree amongst all members of $L \cap k[x_1]$, we have $L = (p_1, x_2 - p_2, \ldots, x_n - p_n)$.*

Proof Set $L' = (p_1, x_2 - p_2, \ldots, x_n - p_n)$ and choose $f \in L$. Then $f \equiv q \pmod{L'}$ for some $q \in L \cap k[x_1]$. It follows that $p_1 \mid q$ since $L \cap k[x_1] = (p_1)$, by the assumption on the degree of p_1. Thus $f \equiv 0 \pmod{L'}$ and so $f \in L'$ which means that $L \subseteq L'$. The result follows since $L' \subseteq L$ by assumption. □

6.7.2 Counting

Let $y_1, y_2, \ldots, y_N$ be boolean variables and consider the following problem:

#Monotone 2-Sat

Input: A boolean formula $\phi = c_1 \wedge c_2 \wedge \cdots c_s$ where each c_i is of the form $y_i \vee y_j$ for some i, j with $1 \leq i, j \leq N$.

Output: The number of satisfying assignments to the given formula.

Valiant (1979) shows that this problem is #P-complete. Note that we may assume that $i \neq j$ for each clause $y_i \vee y_j$ of ϕ and we make this assumption from now on (this just makes the encoding given below a little simpler; see the remark after Lemma 6.13). Let k, X be as in the preceding section where the cardinality n of X is set to $N + 1$. Given a boolean formula ϕ as above, we can encode it as a set of polynomials G_ϕ in $k[X]$ as follows:

$$\begin{aligned} \textbf{true} &\mapsto 0, \\ \textbf{false} &\mapsto 1, \\ y_i \vee y_j &\mapsto x_{i+1}x_{j+1}. \end{aligned}$$

G_ϕ consists of the encoded clauses of ϕ together with x_1 and the set S but with $x_1^2 - x_1$ omitted. It is clear that there is a 1–1 correspondence between the satisfying assignments of ϕ and the zeros of $I = (G_\phi)$. Set $J = \alpha(I)$ and $K = (\beta(J))$, as in §6.7.1.

Lemma 6.13 *Consider any total degree order on the power products of $k[X]$. Then $\beta(\alpha(G_\phi))$ is a Gröbner basis of K.*

Proof There is a subset P of $\{2, 3, \ldots, n\}^2$ such that the members of $\beta(\alpha(G_\phi))$ are precisely

$$\begin{aligned} l &= x_1^2 - \sum_{i=2}^{n} 2^{i-1}x_i, \\ f_{ij} &= x_i x_j, \quad \text{for } (i,j) \in P \\ s_i &= x_i^2 - x_i, \quad \text{for } 2 \leq i \leq n. \end{aligned}$$

This set is a Gröbner basis if and only if every S-polynomial of each pair of its elements reduces to 0. We make use of Buchberger's first criterion: if the leading power product of g is coprime to that of h then $S(f,g)$ reduces to 0 using g and h. This means that we do not need to consider l since its leading power product is x_1^2 and x_1 does not appear in any other polynomial of our set. Likewise we do not need to consider $S(s_i, s_j)$. It is also clear that $S(x_i x_j, x_r x_s) = 0$ (this is so for arbitrary power products; S-polynomials are *designed* to do this). Finally we need only consider $S(x_i x_j, x_i^2 - x_i)$ for $(i,j) \in P$. Recall that $i \neq j$ so that $S(x_i x_j, x_i^2 - x_i) = x_i x_j$ and this reduces to 0 since $x_i x_j$ is in our set. □

We note that if we allow clauses in ϕ of the form $y_i \vee y_i$ then everything works provided such a clause is encoded as x_{i+1} rather than x_{i+1}^2 (we can also omit $x_i^2 - x_i$).

Lemma 6.14 *Assume that k has characteristic 0 and consider a lexicographic order on the power products of $k[X]$ in which x_1 is the smallest indeterminate. Let $p_1, p_2, \ldots, p_n$ be as in Lemma 6.10 with $K = (\beta(\alpha(G_\phi)))$. Then*

1. *Every Gröbner basis of K using the preceding order includes a nonzero constant multiple of p_1.*
2. *$p_1, x_2 - p_2, \ldots, x_n - p_n$ form a Gröbner basis for K.*

Proof By Lemma 6.10, $p_1 \in K$ and so this must reduce to 0 under any Gröbner basis for K. This means that the basis must have a member $q \in k[x_1]$ such that $q \mid p_1$ (since we are using a lexicographic order in which x_1 is the smallest indeterminate). But p_1 has minimal degree amongst all nonzero members of $K \cap k[x_1]$. The first part now follows.

For the second part we note that $p_1, x_2 - p_2, \ldots, x_n - p_n$ are certainly a Gröbner basis (under the stated order) for the ideal that they generate; see the remarks in the proof of Lemma 6.13. By Lemma 6.10 and Lemma 6.12 this ideal is K. □

We can now see one way in which the well known differences in runtime for computing Gröbner bases under a total degree order as opposed to a lexicographic one can be linked with Complexity Theory. On the one hand the Gröbner basis of K under a total degree order is as cheap to compute as possible; it is the same as the input! On the other hand if we use a lexicographic order with x_1 as the smallest indeterminate then even finding the degree of p_1 amounts to solving a #P-complete problem. Moreover if we fix a ϕ with exponentially many satisfying assignments then the polynomial p_1 in the Gröbner basis of K under a lexicographic order with x_1 as the least indeterminate has exponentially many terms (see the remarks after Lemma 6.9). Thus the runtime of any algorithm to compute this Gröbner basis is exponential while the Gröbner basis under a total degree order can be computed in linear time. This provides simple examples for which a change of ordering method such as that of Faugère, Gianni, Lazard and Mora [24] does not help.

6.8 The Case of Two Indeterminates

For polynomials in two indeterminates it can be shown that if all the inputs to the Gröbner basis algorithm have total degree bounded by d then the total degrees of the elements of a reduced basis are always bounded by d^2. (If a total degree order is used then the bound can be improved to $2d - 1$.) Moreover the number of of polynomials in the reduced basis is bounded by $m + 1$ where m is the minimum of the total degrees of the leading power products of the inputs. These facts

$$
\begin{aligned}
h_1 = {} & \langle 777 \rangle y + \langle 769 \rangle x^{48} - \langle 770 \rangle x^{47} - \langle 770 \rangle x^{46} - \langle 771 \rangle x^{45} + \langle 773 \rangle x^{44} \\
& - \langle 774 \rangle x^{43} - \langle 774 \rangle x^{42} - \langle 773 \rangle x^{41} - \langle 774 \rangle x^{40} - \langle 775 \rangle x^{39} - \langle 775 \rangle x^{38} \\
& - \langle 775 \rangle x^{37} - \langle 775 \rangle x^{36} - \langle 775 \rangle x^{35} - \langle 775 \rangle x^{34} - \langle 776 \rangle x^{33} + \langle 775 \rangle x^{32} \\
& + \langle 776 \rangle x^{31} - \langle 776 \rangle x^{30} - \langle 776 \rangle x^{29} + \langle 776 \rangle x^{28} - \langle 776 \rangle x^{27} + \langle 777 \rangle x^{26} \\
& + \langle 776 \rangle x^{25} - \langle 777 \rangle x^{24} + \langle 776 \rangle x^{23} + \langle 777 \rangle x^{22} - \langle 777 \rangle x^{21} + \langle 776 \rangle x^{20} \\
& + \langle 776 \rangle x^{19} - \langle 776 \rangle x^{18} + \langle 777 \rangle x^{17} + \langle 776 \rangle x^{16} - \langle 778 \rangle x^{15} + \langle 777 \rangle x^{14} \\
& + \langle 777 \rangle x^{13} - \langle 777 \rangle x^{12} + \langle 777 \rangle x^{11} - \langle 776 \rangle x^{10} - \langle 777 \rangle x^{9} + \langle 778 \rangle x^{8} \\
& + \langle 775 \rangle x^{7} - \langle 777 \rangle x^{6} + \langle 777 \rangle x^{5} - \langle 777 \rangle x^{4} - \langle 777 \rangle x^{3} + \langle 777 \rangle x^{2} \\
& + \langle 777 \rangle y + \langle 776 \rangle x - \langle 776 \rangle, \\
h_2 = {} & \langle 16 \rangle x^{49} - \langle 17 \rangle x^{48} + \langle 17 \rangle x^{47} - \langle 19 \rangle x^{46} + \langle 20 \rangle x^{45} - \langle 21 \rangle x^{44} \\
& - \langle 21 \rangle x^{43} + \langle 21 \rangle x^{42} - \langle 22 \rangle x^{41} - \langle 22 \rangle x^{40} + \langle 22 \rangle x^{39} - \langle 22 \rangle x^{38} \\
& - \langle 22 \rangle x^{37} - \langle 22 \rangle x^{36} + \langle 22 \rangle x^{35} - \langle 23 \rangle x^{34} + \langle 24 \rangle x^{33} - \langle 23 \rangle x^{32} \\
& - \langle 23 \rangle x^{31} + \langle 23 \rangle x^{30} + \langle 23 \rangle x^{29} - \langle 24 \rangle x^{28} + \langle 24 \rangle x^{27} - \langle 24 \rangle x^{26} \\
& + \langle 24 \rangle x^{25} + \langle 24 \rangle x^{24} - \langle 24 \rangle x^{23} - \langle 24 \rangle x^{22} + \langle 24 \rangle x^{21} - \langle 25 \rangle x^{20} \\
& + \langle 24 \rangle x^{19} + \langle 24 \rangle x^{18} - \langle 25 \rangle x^{17} + \langle 24 \rangle x^{16} + \langle 25 \rangle x^{15} - \langle 25 \rangle x^{14} \\
& + \langle 24 \rangle x^{13} + \langle 24 \rangle x^{12} - \langle 25 \rangle x^{11} + \langle 25 \rangle x^{10} + \langle 23 \rangle x^{9} - \langle 24 \rangle x^{8} \\
& + \langle 25 \rangle x^{7} - \langle 23 \rangle x^{6} - \langle 24 \rangle x^{5} + \langle 24 \rangle x^{4} - \langle 24 \rangle x^{3} - \langle 24 \rangle x^{2} \\
& - \langle 24 \rangle x - \langle 24 \rangle
\end{aligned}
$$

Figure 2: The form of the Gröbner basis for f_1, g_1.

are occasionally cited as proof that computing the Gröbner basis of such polynomials is an efficient process. This of course is highly suspect since the bounds do not tell us anything about the growth of the coefficients. In this connection it is worthwhile considering the following examples.

$$
\begin{aligned}
f_1 &= 8y^7 + y^6x - 3y^5x + 64y^5 - 35y^3x^3 + 79y^2x, \\
g_1 &= 92y^6 - 95y^5x^2 + 56y^3x - 87y - 46x^7 + 96.
\end{aligned}
$$

Let B be the normed reduced Gröbner basis of f_1, g_1 w.r.t. the lexicographic order where $x <_L y$. If we turn the elements of B into primitive polynomials with integer coefficients (by clearing denominators) then we obtain two polynomials h_1, h_2 of the form shown in Figure 2, where $\langle n \rangle$ denotes a positive integer of n decimal digits. This basis took 20 hours, 5 minutes and 13 seconds to compute using Maple's package `grobner`. It is interesting to observe that if we perform the same computation but with $x >_L y$ then the basis is considerably smaller and the time taken is only 15 minutes and 31 seconds. The machine used was a Sun SparcServer 1000 with six 60Mhz SuperSPARC cpus and 384Mb of memory (this applies to all the runtimes given below).

If we repeat the experiment with $f_2(x, y) = f_1(x, y) + f_1(y, x)$ and $g_2(x, y) = g_1(x, y) + g_1(y, x)$ the situation is *much* worse. The polynomials of the Gröbner basis have the format shown in Figure 3. The cpu time for this computation was 42 hours, 19 minutes and 39 seconds.

It is well known that computing a Gröbner basis w.r.t. a lexicographic order is usually more

$$
\begin{aligned}
h_1 = {} & \langle\, 953 \,\rangle y + \langle\, 947 \,\rangle x^{47} + \langle\, 948 \,\rangle x^{46} + \langle\, 948 \,\rangle x^{45} + \langle\, 949 \,\rangle x^{44} - \langle\, 950 \,\rangle x^{43} \\
& - \langle\, 951 \,\rangle x^{42} + \langle\, 951 \,\rangle x^{41} - \langle\, 951 \,\rangle x^{40} - \langle\, 951 \,\rangle x^{39} + \langle\, 952 \,\rangle x^{38} + \langle\, 953 \,\rangle x^{37} \\
& + \langle\, 953 \,\rangle x^{36} + \langle\, 953 \,\rangle x^{35} - \langle\, 952 \,\rangle x^{34} - \langle\, 953 \,\rangle x^{33} + \langle\, 953 \,\rangle x^{32} + \langle\, 953 \,\rangle x^{31} \\
& - \langle\, 952 \,\rangle x^{30} + \langle\, 953 \,\rangle x^{29} - \langle\, 953 \,\rangle x^{28} - \langle\, 954 \,\rangle x^{27} + \langle\, 954 \,\rangle x^{26} + \langle\, 954 \,\rangle x^{25} \\
& - \langle\, 953 \,\rangle x^{24} + \langle\, 954 \,\rangle x^{23} + \langle\, 953 \,\rangle x^{22} - \langle\, 954 \,\rangle x^{21} + \langle\, 954 \,\rangle x^{20} + \langle\, 954 \,\rangle x^{19} \\
& - \langle\, 954 \,\rangle x^{18} + \langle\, 954 \,\rangle x^{17} + \langle\, 954 \,\rangle x^{16} - \langle\, 954 \,\rangle x^{15} + \langle\, 953 \,\rangle x^{14} + \langle\, 954 \,\rangle x^{13} \\
& - \langle\, 954 \,\rangle x^{12} + \langle\, 954 \,\rangle x^{11} + \langle\, 954 \,\rangle x^{10} - \langle\, 954 \,\rangle x^{9} - \langle\, 954 \,\rangle x^{8} + \langle\, 954 \,\rangle x^{7} \\
& - \langle\, 954 \,\rangle x^{6} - \langle\, 953 \,\rangle x^{5} + \langle\, 953 \,\rangle x^{4} + \langle\, 953 \,\rangle x^{3} - \langle\, 953 \,\rangle x^{2} + 9\langle\, 53 \,\rangle y \\
& + \langle\, 953 \,\rangle x - \langle\, 953 \,\rangle \\
h_2 = {} & \langle\, 22 \,\rangle x^{48} + \langle\, 22 \,\rangle x^{47} - \langle\, 21 \,\rangle x^{46} + \langle\, 23 \,\rangle x^{45} - \langle\, 24 \,\rangle x^{44} - \langle\, 25 \,\rangle x^{43} \\
& + \langle\, 25 \,\rangle x^{42} - \langle\, 26 \,\rangle x^{41} - \langle\, 25 \,\rangle x^{40} + \langle\, 27 \,\rangle x^{39} + \langle\, 26 \,\rangle x^{38} + \langle\, 26 \,\rangle x^{37} \\
& + \langle\, 27 \,\rangle x^{36} - \langle\, 27 \,\rangle x^{35} - \langle\, 27 \,\rangle x^{34} + \langle\, 27 \,\rangle x^{33} + \langle\, 27 \,\rangle x^{32} - \langle\, 27 \,\rangle x^{31} \\
& + \langle\, 28 \,\rangle x^{30} - \langle\, 28 \,\rangle x^{29} - \langle\, 28 \,\rangle x^{28} + \langle\, 28 \,\rangle x^{27} + \langle\, 28 \,\rangle x^{26} - \langle\, 28 \,\rangle x^{25} \\
& + \langle\, 28 \,\rangle x^{24} - \langle\, 28 \,\rangle x^{23} - \langle\, 29 \,\rangle x^{22} + \langle\, 29 \,\rangle x^{21} + \langle\, 27 \,\rangle x^{20} - \langle\, 28 \,\rangle x^{19} \\
& + \langle\, 29 \,\rangle x^{18} - \langle\, 28 \,\rangle x^{17} - \langle\, 29 \,\rangle x^{16} + \langle\, 29 \,\rangle x^{15} + \langle\, 28 \,\rangle x^{14} - \langle\, 29 \,\rangle x^{13} \\
& + \langle\, 29 \,\rangle x^{12} - \langle\, 28 \,\rangle x^{11} - \langle\, 28 \,\rangle x^{10} + \langle\, 27 \,\rangle x^{9} + \langle\, 28 \,\rangle x^{8} - \langle\, 28 \,\rangle x^{7} \\
& + \langle\, 28 \,\rangle x^{6} - \langle\, 27 \,\rangle x^{5} + \langle\, 27 \,\rangle x^{4} - \langle\, 28 \,\rangle x^{3} + \langle\, 28 \,\rangle x^{2} - \langle\, 28 \,\rangle x + \langle\, 27 \,\rangle.
\end{aligned}
$$

Figure 3: The form of the Gröbner basis for f_2, g_2.

expensive than w.r.t. a total degree order. The polynomials given in this section provide a striking example of this phenomenon. Computing the Gröbner basis of f_1, g_1 using Maple's `tdeg` order (total degree then reverse lexicographic) with $x <_L y$ took only 3.9 seconds and the result was quite small. The same experiment with f_2, g_2 took only 7.6 seconds and again the result was of a modest size. These observations suggest that the basis conversion method of Faugère, Gianni, Lazard and Mora [24] would lead to a reasonable improvement. (See also Trinks [59] on the question of coefficient size.) Even so the resulting bases would still be as shown above and therefore quite large.

Exercise 6.10 *Find, by hand calculations only, the normed Gröbner basis of*

$$\begin{aligned} g_1 &= x^3yz - xz^2, \\ g_2 &= xy^2z - xyz, \\ g_3 &= x^2y^2 - z, \end{aligned}$$

using the lexicographic order with $x < y$. Use Maple to check your answer.

Exercise 6.11 *Let G be a set of linear polynomials in the indeterminates $x_1, \ldots, x_n$ (i.e., polynomials of the form $a_1x_1 + \cdots + a_nx_n$ where $a_i \in k$ for $1 \leq i \leq n$). Which familiar algorithm does the Gröbner basis algorithm resemble when applied to G with an admissible ordering?*

Exercise 6.12 *Let $f(x, y)$ be a non-zero polynomial with complex coefficients. The* algebraic curve *defined by $f(x, y)$ is just $\mathbf{V}(f)$, i.e., all the solutions in $\mathbb{C}^2$ to $f(x, y) = 0$. The* singular *points of the curve are those points on the curve at which both partial derivatives vanish, i.e., they are the solutions to the simultaneous system*

$$\begin{aligned} f(x, y) &= 0, \\ f_x(x, y) &= 0, \\ f_y(x, y) &= 0, \end{aligned}$$

where f_x is the partial derivative w.r.t. x and similarly for f_y. The curve is said to have no multiple components if $f(x, y)$ is square free.

1. *Give two different proofs of the fact that every algebraic curve with no multiple components has only finitely many singularities. (For one proof use Gröbner bases and for the other use resultants.)*

   Can you put an upper bound on the number of singularities in terms of the degree of f?

2. *Let*

   $$f(x, y) = (x^2 + y^2 - 1)((x - 1)^2 + y^2 - 1)((x + 1)^2 + y^2 - 1)(x^2 + (y - 1)^2 - 1).$$

   Find all the singularities of $f(x, y) = 0$ exactly (use Axiom's `groebner` *function). Can you give a geometrical interpretation of the singularities?*

Exercise 6.13 *Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite alphabet. Consider a term rewriting system over X^* (the set of finite strings over X) in which the rules are bidirectional:*

$$\begin{aligned} t_1 &\leftrightarrow t_1', \\ t_2 &\leftrightarrow t_2', \\ &\vdots \\ t_s &\leftrightarrow t_s', \end{aligned}$$

where $t_i, t_i' \in X^*$ for $1 \leq i \leq s$. Thus if $T = At_iB$ and $T' = At_i'B$, where $A, B \in X^*$, then T' may be derived from T in a single step and vice versa. (In general, a word $W' \in X^*$ is derivable from a word $W \in X^*$ if $W = W'$ or there is a finite sequence of single-step derivations which starts with W and ends with W'.)

Consider the elements of X^* to be monomials of $\mathbb{C}[X]$ and define the polynomials

$$p_i = t_i - t_i', \quad \text{for } 1 \leq i \leq s.$$

Let W, W' be arbitrary elements of X^*. It can be shown that W' is derivable from W if and only if

$$W - W' \in (p_1, p_2, \ldots, p_s). \qquad (*)$$

1. *Give an algorithm which solves the derivability problem.*

2. *Prove the claim made about $(*)$, i.e., prove that W' is derivable from W if and only if $(*)$ holds.*

   [Hint: The 'only if' part is easier so do this first. For the 'if' part write an expression of the form $q_1p_1 + \cdots + q_sp_s$ as a finite sum $\sum \epsilon_m m p_m$, where each m is a monomial and $\epsilon_m = \pm 1$. Now use induction on the number of summands.]