Gröbner Bases

Gröbner bases are a tool for doing explicit algorithmic calculations in a polynomial ring over a field or a homomorphic image of a polynomial ring over a field. We assume that arithmetic operations on the elements of the field can be performed algorithmically. Throughout, K is a field, and $x_1, \ldots, x_n$ are indeterminates over K.

While Gröbner bases are tools for calculation, they can also be used to prove substantial theorems, such as the Hilbert basis theorem (ideals in R are finitely generatded) and the Hilbert syzygy theorem (discussed below). Moreover, not surprisingly, the systematic study of Gröbner bases introduces many new theoretical problems.

One of the problems we want to solve is this: given generators for an ideal of the ring, how do we tell whether a given element of the polynomial ring is in the ideal?

Here is another problem: If we have finitely many generators for an ideal

$$I \subseteq K[x_1, \ldots, x_n] = R,$$

how can we find finitely many generators for $I \cap K[x_s + 1, \ldots, x_n]$, $1 \leq s \leq n-1$?

This problem is intimately connected with the problem of solving the equations obtained by setting the generators of the ideal equal to 0. Suppose that K is algebraically closed and that we know that there are only finitely many solutions.

By intersecting I with $K[x_n]$ we get a principal ideal generated by one polynomial. It has only finitely many roots, say $\lambda_1, \ldots, \lambda_r$. If the elements of the solution set have last coordinates $\lambda_1, \ldots, \lambda_d$, these will be the same as the roots of that single polynomial, although there may be multiplicities. (The polynomial $(x_n - \lambda_1) \cdots (x_n - \lambda_d)$ vanishes on the solution set: by Hilbert's Nullstellensatz, it has a power in I. Conversely, if $f(x_n) \in I$, the first coordinate of any point in the solution set obviously satisfies f.)

For each of these roots λ_i we can substitute $x_n = \lambda_i$ in all of the polynomials. We have now replaced the original problem by finitely many, each one involving fewer variables than in the original.

This idea reduces the problem of solving systems of polynomial equations in several variables with finitely many solutions to the problem of solving one equation in one variable.

Another use of Gröbner bases is this: given elements $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$, find generators for all the relations on those elements, i.e., for the module of m-tuples of polynomials $g_1, \ldots, g_m$ such that $\sum_{j=1}^{m} g_j f_j = 0$. In fact, one can require insteasd that the g_i satisfy several equations like this, i.e., a system

$$\sum_{i=1}^{m} g_j f_{i,j} = 0, \quad 1 \leq i \leq r.$$

This is equivalent to finding the relations on the r columns of the $m \times r$ matrix $\big(f_{i,j}\big)$. Consider the R-submodule M of R^m spanned by these columns. The module of relations on the columns is called a *first module of syzygies* of M. More generally, whenever we have a short exact sequence of finitely generated R-modules $0 \to M' \to R^r \to M \to 0$, M' is called a *first module of syzygies* of M. A first module of syzygies of a kth module of syzygies is called a $(k+1)$ *st module of syzygies*: when N is a nth module of syzygies of M there is an exact sequence

$$0 \to N \to R^{b_{n-1}} \to \cdots \to R^{b_0} \to M \to 0$$

of finitely generated R-modules.

Gröbner bases can be used to prove the famous Hilbert syzygy theorem, that every finitely generated module over $K[x_1, \dots, x_n]$ has an nth module of syzygies that is free. (Equivalently, that every finitely generated R-module has a free resolution of length at most n.) Beyond that, they can be used to compute the resolution.

In the graded case, this means that Gröbner bases can be used to find finite free resolutions of graded R-modules M. If we let $H_M(t) = \dim_K(M_t)$, where M_t is the tth graded piece of M, then $H_M(t)$, the *Hilbert function* of M, agrees with a polynomial in t for all $n \gg 0$. The degree of this polynomial is one less than the Krull dimension of M (which is the same as the Krull dimension of the ring $R/\mathrm{Ann}_R M$.) One finds a graded free resolution of M in which a typical term is a direct sum of copies of $R(-s)$, s a variable integer, where $R(-s)$ is free on one generator but is graded so that its generator lives in degree s (that is, the tth graded piece of $R(-s)$ is R_{t-s}). The Hilbert function of $K[x_1, \dots, x_n]$ itself is $\binom{t+n-1}{n-1}$ (and it is zero for $t < 0$). One can then immediately compute the Hilbert function of any graded module M from a finite graded free resolution: one simply takes the alternating sum of the Hilbert functions of the free modules in the resolution, each of which is sum of copies of $H_{R(-s)}(t) = H_R(t-s)$ with s varying. Thus, Gröbner bases can be used to find Hilbert functions.

Let $\mathcal{M}$ denote the multiplicative semigroup of all monomials in $x_1, \dots, x_n$: a typical element is $x_1^{k_1} \cdots x_n^{k_n}$ for $(k_1, \dots, k_n) \in \mathbb{N}^n$, where $\mathbb{N}$ is the set of non-negative integers.

Throughout, free modules are assumed to be finitely generated. If F is a finitely generated free $K[x_1, \dots, x_n]$-module with a given free basis $e_1, \dots, e_r$ by a *monomial* in F we mean an element of the form μe_i, where $\mu \subseteq \mathcal{M}$.

In order to define the notion of *Gröbner* basis, one first needs to fix a *monomial* order. This is a total ordering $>$ of $\mathcal{M}$ that is compatible with multiplication, and such that 1 is least. That is, if λ, μ, and ν are monomials, and $\mu \geq \nu$, then $\lambda\mu \geq \lambda\nu$ (and $\lambda\nu \geq \lambda$ follows since $\nu \geq 1$). We shall make the convention that the variables always have the order $x_1 > \cdots > x_n$: this can always be achieved by renumbering.

Given a free R-module with free basis $e_1, \dots, e_r$, we extend the monomial order to the monomials of the free R-module by requiring that $e_1 > \cdots > e_r$ and then letting $\mu e_i > \nu e_j$ if $\mu > \nu$ or $\mu = \nu$ and $e_i > e_j$. Such monomial orders are referred to by the acronym TOP (term over position). Other monomial orders on free modules are possible, but these will suffice for our purposes here. (In the general case, one requires that for any $\mu e_i \geq \nu e_j$, one has $\lambda \mu e_i \geq \lambda \nu e_j \geq \nu e_j$.)

Here are some examples of monomial orders on $\mathcal{M}$. One is *lexicographic order*, or *lex* order, which means that $x_1^{k_1} \cdots x_n^{k_n} > x_1^{h_1} \cdots x_n^{h_n}$ precisely if for some s, $1 \leq s \leq n$, $k_i = h_i$ for $i < s$ while $k_s > h_s$.

A variant is *homogeneous lexicographic* or *hlex* order, where $\mu >_{\text{hlex}} \nu$ means either that μ has larger degree than ν or they have the same degree and $\mu >_{\text{lex}} \nu$.

A very important monomial order is *reverse lexicographic order.* Here $\mu >_{\text{revlex}} \nu$ means that μ has larger degree, or that the degrees are equal and for some s, $1 \leq s \leq n$, $k_s < h_s$ while $k_i = h_i$ for $i > s$. Among monomials of a fixed degree, this is the opposite of the lexicographic order obtained by numbering the variables backwards, so that a double reversal is involved. Note that $x_1 x_3 >_{\text{hlex}} x_2^2$ but $x_2^2 >_{\text{revlex}} x_1 x_3$. Very roughly speaking, larger monomials in lex order involve more of the earlier variables, while in revlex order they involve fewer of the later variables. This "tends" to be the same thing, but not always.

In two or more variables there are always uncountably many monomial orders! E.g., let $\theta_1, \ldots, \theta_n$ be positive real numbers linearly independent over the rationals, and map the monomials to the reals by letting W send $x_1^{k_1} \cdots x_n^{k_n}$ to $\sum_{i=1}^n k_i \theta_i$. Order the monomials by the rule $\mu > \nu$ precisely when $W(\mu) > W(\nu)$. Even in the case of two variables there are uncountably many orders of this type.

Now fix a monomial order, both on $\mathcal{M}$ and, if we are working with a free module with free basis $e_1, \ldots, e_r$, on the monomials of the free module. By a *term* in a polynomial we mean $c\mu$, where c is a nonzero scalar, that occurs. Likewise, every element of F is a linear combination of monomials, and we define a *term* to mean $c\mu e_i$ occurring in the element, where c is a nonzero scalar. The terms occurring in a given element are linearly ordered by the monomial order if we ignore the scalars. Therefore, every element u of R or F has a largest term, called the *initial term*, $\text{in}(u)$. Given an ideal of $I \subseteq R$ or a submodule M of $F \cong R^r$, we define $\text{in}(I)$ (respectively, $\text{in}(M)$ to be the ideal (respectively, submodule of F) spanned by all the initial terms of elements of I (respectively M). This will be a monomial ideal (respectively, module), i.e., one generated by monomials. Monomial ideals and modules are much easier to work with than others: e.g., the intersection of two such is spanned by the monomials in the intersection.

We now come to a key definition. Fix a monomial order for R (respectively, for F). A *Gröbner basis* for an ideal I or for submodule M of F is a set of elements of I (respectively M) whose initial terms generate $\text{in}(I)$ (respectively, $\text{in}(M)$). It is then easy to show that a Gröbner basis actually does generate I (respectively, M).

In fact, one has that if $N \subseteq M \subseteq F$ then $\text{in}(N) = \text{in}(M)$ implies that $N = M$.

A Gröbner basis is called *minimal* if no terms can be omitted. This means that no initial term of an element in the basis divides any other. It is called *reduced* if no initial term of an element divides any term in any other element and the scalar coefficient of every initial term is 1. We shall see that reduced Gröbner bases exist, can be found algorithmically, and are unique.

Next note that every monomial order has DCC: the reason is that in any set of monomials, finitely many are minimal under the partial ordering by divisibility, and one of these must be least in any total ordering that refines it.

The following fact is of great importance in the theory of Gröbner bases. From this point on we shall state several results for the module case: the case of ideals is included.

Theorem (division). *Let F be a free R-module, with $R = K[x_1, \dots, x_n]$, and suppose that a monomial order $>$ has been fixed. If $f, g_1, \dots, g_h \in F$ are given then one can write*

$$f = \sum_{i=1}^{h} f_i g_i + \rho$$

with the $f_i \in R$, $\rho \in F$, where none of the monomials occurring in ρ is in

$$(in(g_1), \dots, in(g_h))$$

and $in(f) \geq in(f_i g_i)$ for $1 \leq i \leq h$. An expression satsifying these conditions is called standard, and ρ is called a remainder of f with respect to $g_1, \dots, g_h$.

Here is both how one proves this theorem and how one carries out the "division" algorithmically. Choose the maximal term in f that is divisible by some $\text{in}(g_i)$. Let m_1 be the quotient of that term by the relevant $\text{in}(g_i)$ (choose i as small as possible if it is desired to make the process choice-free). Let i_1 be the value of i used. Now look at $f_1 = f - m_1 g_{i_1}$ and repeat the process. This generates a sequence, $f_0 = f, f_1, f_2$, etc. Eventually, either f_s is 0, or else it has no term that is divisible by any $\text{in}(g_j)$. In the latter case, $f_s = \rho$.

It is trivial that if $g_1, \dots, g_h$ are a Gröbner basis and g is in their span, then any remainder in the division algorithm must be zero. (The remainder is in the span of the $g_1, \dots, g_h$, but its initial term is not divisible by any $\text{in}(g_i)$.)

This shows that one has an effective membership test once one has any Gröbner basis.

Second, one can always modify a Gröbner basis until it is reduced: consider the largest monomial in any g_j that is divisible by some $\text{in}(g_i)$ for $i \neq j$ Subtracting a suitable multiple of g_i from g_j decreases the number of occurrences of that largest monomial in the Gröbner basis. It follows that one eventually reaches a reduced Gröbner basis (one also needs to adjust the coefficients of the initial terms at the end, but that is a trivial step). We leave it as an exercise to show that two reduced Gröbner bases for the same submodule of F must be identical.

We next give a criterion and algorithm due to Buchberger for testing whether one has a Gröbner basis and, if not, enlarging it. Eventually, this produces a Gröbner basis. Suppose that we have a free module F over $R = K[x_1, \dots, x_n]$ and monomial order as usual. Let $g_1, \dots, g_h$ be elements of F. For every pair of indices i, j such that $\text{in}(g_i)$ and $\text{in}(g_j)$ involve the same basis element e_k of F, let $m_{ij} = \text{in}(g_i)/GCD(\text{in}(g_i), \text{in}(g_j)) \in S$. For each such i, j we can choose a standard expression for $m_{ji}g_i - m_{ij}g_j$: call the reminder ρ_{ij}. (This means that we have expressions:

$$m_{ji}g_i - m_{ij}g_j = \sum_{t=1}^{h} f_{ijt}g_t + \rho_{ij}$$

where for all i, j, $\mathrm{in}(m_{ji}g_i - m_{ij}g_j) \geq \mathrm{in}(f_{ijt}g_t)$ for all t, and none of the monomials occurring in any ρ_{ij} is in $(\mathrm{in}(g_1), \ldots, \mathrm{in}(g_h))$.

This is actually very simple: in writing down $m_{ji}g_i - m_{ij}g_j$ we are trying to get the "obvious" candidate for the initial term of a linear combination of the g_i, in fact, of just two of them, to turn out to be zero, so that we might get a new element of $\in (M)$ from it, where M is the span of the g_i. But we first perform division on the difference: if the remainder is 0 we are not actually getting anything new.

Buchberger's criterion asserts that the g_i are a Gröbner basis for their span if and only if all the $\rho_{ij} = 0$. The proof is not difficult, and is given, for example, in D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, GTM **150** Springer-Verlag, New York, 1995, §15.4. In further references we refer to Eisenbud's book as [E].

Buchberger's algorithm is an immediate consequence: if the g_i are not a Gröbner basis, enlarge them with the ρ_{ij}. Repeating will eventually produce a Gröbner basis, since the module spanned by the initial terms cannot go on increasing forever.

The problem of computing syzygies (or relations) on a module is also solved easily now. We first note that it does not affect the problem to increase the set of generators by one element that is already a linear combination of them (or, inductively, by several). If we know the relations on $g_1, \ldots, g_h$ and $g_{h+1} = \sum_{i=1}^{h} r_i g_i$, then this single new relation together with the ones on $g_1, \ldots, g_h$ spans all of the relations on $g_1, \ldots, g_h, g_{h+1}$. On the other hand, from any relation on $g_1, \ldots, g_h, g_{h+1}$ one gets a relation on $g_1, \ldots, g_h$ by substituting $g_{h+1} = \sum_{i=1}^{h} r_i g_i$, and in this way generators for the relations on $g_1, \ldots, g_h, g_{h+1}$ give rise to generators for the relations on $g_1, \ldots, g_h$. Finally, given two different sets of generators for M, one can compare each to the union, and so it does not matter which set of generators one uses in computing relations.

Therefore, one may assume that one is working with a Gröbner basis for the module in doing this: call it $g_1, \ldots, g_h$. With the m_{ij} defined as in the Buchberger criteron, we have certain equations, coming from the division algorithm, but now, every $\rho_{ij} = 0$. But then, after we move the two terms on the left to the right, every equation

$$m_{ji}g_i - m_{ij}g_j = \sum_{t=1}^{h} f_{ijt}g_t$$

gives a relation on $g_1, \ldots, g_h$. It is not difficult to show that these relations span all relations on $g_1, \ldots, g_h$! This method is due to Schreyer. Cf. [E], §15.5. In fact, for a suitable monomial order on the free module containing the relations, these relations already are a Gröbner basis for all the relations. (Map the free module with free basis $e'_1, \ldots, e'_h$ into the free module containing the Gröbner basis $g_1, \ldots, g_h$ by sending e'_i to g_i. Put a monomial order on the new free module by $\mu e'_i > \nu e'_j$ if $\mathrm{in}(\mu g_i) > \mathrm{in}(\nu g_j)$ with respect to the monomial order on F, or they are equal up to multiplication by a nonzero scalar and $i < j$. The relations $\tau_{ij} = m_{ji}e'_i - m_{ij}e'_j - \sum_{t=1}^{h} f_{ijt}e'_t$ give a Gröbner basis for the module of relations, and τ_{ij} has initial term $m_{ji}e'_i$.)

We conclude with several exercises on Gröbner bases. The results in 1., 2., 5., and 6. are significant applications, while 3. and 4. are calculations of Gröbner bases and syzygies.

We first recall that $y_1, \ldots, y_k$ is a regular sequence on N if $N \neq (y_1, \ldots, y_k)N$, y_1 is not a zerodivisor on N, y_2 is not a zerodivisor on N/y_1N, and, for every $t < k$, y_{t+1} is not a zerodivisor on $N/(y_1, \ldots, y_t)N$.

Exercise 1 (elimination theory). Use lexicographic order on R, with $x_1 > \cdots > x_n$. If $g_1, \ldots, g_h$ is a reduced Gröbner basis for I, then those elements that happen to involve only $x_s, \ldots, x_n$ is a reduced Gröbner basis for $I \cap K[xs, \ldots, x_n]$. Is that easy or what?! (The proof is *very* easy. Cf. D. Cox, J. Little, and D O'Shea, *Ideals, Varieties, and Algorithms*, Springer-Verlag, New York, 1992, p. 114, Theorem 2, or [E], pp. 357–359.)

Exercise 2 (regular sequence test). Let F be free and use reverse lexicographic order. Let M be a graded submodule. Then $x_1, \ldots, x_k$ is a regular sequence on F/M if and only if it is a regular sequence on $F/\mathrm{in}(M)$. (This is a theorem of Bayer and Stillman. Cf. [E], Thoerem 15.13.)

Exercise 3. Find a Gröbner basis for xy, $xy + y^2$ for lexicographic order with $x > y$. (Cf. [E], pp. 334–5.)

Exercise 4. Find a Gröbner basis for $x^2, y^2, xy + yz$ using reverse lexicographic order with $x > y > z$, and then find the relations (or syzygies) on the elements of the Gröbner basis by Schreyer's method.

Exercise 5 (Hilbert basis theorem). Give a proof by combinatorial methods that every monomial ideal in R is finitely generated. It follows that any ideal I has a finite Gröbner basis. Therefore, every ideal of R is finitely generated: this is a Gröbner basis proof of the Hilbert basis theorem.

Exercise 6 (Hilbert syzygy theorem). With notation as in the discussion of finding syzygies above, arrange the Gröbner basis so that whenever $\mathrm{in}(g_i)$ and $\mathrm{in}(g_j)$ involve the same e_k, say $\mathrm{in}(g_i) = \nu_i e_k$ and $\mathrm{in}(g_j) = \nu_k e_k$, then if $i < j$ we have that $\nu_i > \nu_j$ in lex order. Show that if the variables $x_1, \ldots, x_s$ are missing from the initial terms of the g_i then the variables $x_1, \ldots, x_{s+1}$ are missing from the initial terms of the τ_{ij}. This easily implies the Hilbert syzygy theorem.