June 28, 2003

INTRODUCTORY GRÖBNER BASES

JASON PRESZLER

Abstract. This paper is a very basic introduction to Gröbner Bases and is focused on undergraduates with some experience in linear algebra, but should also be of interest to those with prior knowledge of ideals and polynomial rings. Gröbner Bases are an abstraction of Gauss-Jordan elimination for systems of multivariate polynomials.

1. History, Notation and Background

Gröbner Bases were first introduced in 1965 in Bruno Buchberger's Ph.D. dissertation and are named after his advisor Wolfgang Gröbner.[2] Their importance wasn't fully realized for several years, as with most ideas in mathematics.

The ideas presented here are applicable to any polynomial ring $k[x_1, \ldots, x_n]$ where k is a field, and I is an ideal of the ring $k[x_1, \ldots, x_n]$. The ring that will be used in examples is $\mathbb{Q}[x, y, z]$, which is the set of all possible polynomials in the indeterminants x, y, and z with rational coefficients. The symbols f, g, and h will be used to express elements in this set unless otherwise noted. Precise definitions of a ring, a field, and an ideal follow.

Definition 1 (Ring)**.** A *ring* is a nonempty set, R, with two associative binary operations, usually $+$ and $*$, called addition and multiplication respectively, satisfying the following $\forall\ a, b, c \in R$:

- $a + b \in R$,
- $a + b = b + a$,
- $(a + b) + c = a + (b + c)$,
- $\exists\ 0 \in R$ such that $a + 0 = a$,
- $\exists\ -a \in R$ such that $a + (-a) = 0$,
- $a * b \in R$,

- $a * (b * c) = (a * b) * c$,
- $a * (b + c) = a * b + a * c$ and $(b + c) * a = b * a + c * a$.

Furthermore, all rings we will deal with will have a multiplicative identity ($\exists\ 1 \in R$ such that $1 * a = a * 1 = a$), and will be commutative under multiplication ($a * b = b * a$). Such rings are formally called a *commutative ring with unity*, this is what will be meant by the term ring.

Definition 2 (Field)**.** A *field* is a commutative ring with unity with the additional property that every nonzero element, a, has a multiplicative inverse, a^{-1}. $(a * a^{-1} = 1)$

Definition 3 (Ideal)**.** An *ideal* I is a subset of a ring R such that $\forall\ a, b \in I,\ a - b \in I$ and $\forall\ a \in I$ and $r \in R$ the product $a * r \in I$. Let $I = \langle A \rangle$ denote the ideal I generated by the set A, this means any $a \in I$ is of the form $a = a_1 r + \cdots + a_n r$ where each $a_i \in A$ and $r \in R$. If A is finite then I is a finitely generated ideal and if A is a singleton then I is called a *principal ideal*. A ring in which every ideal can be expressed as a principal ideal is called a principal ideal domain, or PID.

The following terms and notation are present in the literature of Gröbner bases and will be useful later on.

Definition 4 (degree, leading term, leading coefficient, and power product)**.** A *power product* is a product of indeterminants $\{x_1^{\beta_1} \cdots x_n^{\beta_n} : \beta_i \in \mathbb{N}, 1 \leq i \leq n\}$. The *degree* of a term of a polynomial f is the sum of the exponents of the term's power product. The *degree* of a polynomial f, denoted $\deg(f)$, is the greatest degree of the terms in f. The *leading term* of f, denoted $lt(f)$, is the term with the greatest degree. The *leading coefficient* of f is the coefficient of the leading term in f while the power product of the leading term is the *leading power product*, $lp(f)$.[1]

Before we can describe what Gröbner Bases are, it is useful to see what they are an abstraction of. Thus, brief examples of linear systems of equations and single variable polynomials follow to motivate the work of Buchberger.

2. Linear Systems of Equations

Linear Algebra has powerful tools for solving systems of linear equations such as:

$$x + y - z = 1 \tag{1}$$

$$7z - y + 2x = 8 \tag{2}$$

$$2y - x - 5z = -5 \tag{3}$$

Notice that there is no explicit order imposed on the variables. We can pick any order we want when we solve the system, we just need an order, this will become an important point later on. To solve the above system the coefficients are put into a matrix with respect to our imposed ordering and then Gaussian (or Gauss-Jordan) elimination is used to obtain the solution.

$$\begin{bmatrix} 1 & 1 & -1 & 1 \\ 2 & -1 & 7 & 8 \\ -1 & 2 & -5 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & \frac{5}{3} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{2}{3} \end{bmatrix} \tag{4}$$

The matrix on the right corresponds to the new system below, which is much easier to solve.

$$x = \frac{5}{3} \tag{5}$$

$$y = 0 \tag{6}$$

$$z = \frac{2}{3} \tag{7}$$

Linear algebra has methods of forming an equivalent system out of the original system and then solving the simpler, equivalent system; which has the same solution set. The crucial idea contained in the process of Gaussian elimination is that equations or the coefficients were expressed in terms of the other equations or coefficients. How can a similar idea be applied to systems of polynomials in one variable?

3. Single Variable Systems of Polynomials

First we must recall that the Euclidean Algorithm can be applied to polynomials in one variable. The process of polynomial division should be familiar, to refresh your memory try dividing $f = x^3 - 2x^2 + 2x + 8$ by $g = 2x^2 + 3x + 1$; the result is $q = \frac{1}{2}x - \frac{7}{4}$ and $r = \frac{27}{4}x + \frac{39}{4}$ where q is the quotient and r is the remainder (thus $f = qg + r$). Analogous to the Euclidean Algorithm for division in $\mathbb{Q}$, the quotient and remainder are unique because $\mathbb{Q}[x]$ is a unique factorization domain.[1] A pseudo-code version of a single variable polynomial division algorithm is given in [1]. An important consequence of such an algorithm is the concept of polynomial reduction.

Definition 5 (Reduction)**.** A polynomial h is a *reduction* of f by g if $h = f - \frac{lt(f)}{lt(g)}g$ is the remainder after the first step of dividing f by g. This is denoted

$$f \xrightarrow{g} h.$$

Repeated reduction steps, such as

$$f \xrightarrow{g} h \xrightarrow{g} r$$

will be denoted as

$$f \xrightarrow{g+} r.$$

This is influenced by modular arithmetic in the integers and for convenience we will say f is congruent to r modulo g.

With a division algorithm in hand we can now address the issue of finding a *greatest common divisor*, or gcd, of two polynomials. This is what the Euclidean algorithm actually yields, its crux being the single variable division algorithm. The following is algorithm 1.3.2 from [1].

[1]Unique factorization domains are very interesting algebraic structures, readers unfamiliar with them should consult [3], but knowledge of them is not crucial to this paper.

Algorithm 1. INPUT: $f_1, f_2 \in \mathbb{Q}[x]$ with at least one not being the zero polynomial.

OUTPUT: $f = \gcd(f_1, f_2)$

INITIALIZATION: $f = f_1,\ g = f_2$

WHILE $g \neq 0$

$$f \xrightarrow{g+} r$$

$$f = g;$$

$$g = r;$$

$f = \frac{1}{lc(f)} f;$

With the ability to compute the gcd of several polynomials, we are in a position to address two important questions, one concerning ideal membership and the other concerning an ideal's generating set.

As previously noted, ideals are usually generated by a finite set of elements, and the most preferred type of ideal is a principal ideal. With the gcd algorithm above we can prove that $k[x]$ is a principal ideal domain, but of more immediate interest is the following theorem.

Theorem 1. *Let $f_1, \ldots, f_2$ be non-zero polynomials in $k[x]$, then*

$$\langle f_1, \ldots, f_2 \rangle = \langle \gcd(f_1, \ldots, f_2) \rangle.$$

The reader is encouraged to prove this theorem on their own, all that is needed is the definition of gcd and the definition of an ideal.

The problem of ideal membership asks when is a given polynomial f in an ideal I? Since any ideal can be generated by one element, this question is answered in the one variable case by the following theorem.

Theorem 2. *A polynomial f is a member of the ideal I generated by g iff*

$$f \xrightarrow{g+} 0.$$

The theorem is almost trivial, but it is important to note that with the single variable division algorithm and the Euclidean algorithm, the issue of ideal membership can be easily solved given any polynomial and the generating set of an ideal.[2]

It's useful to draw some connections between linear algebra and solving systems of equations. Let S be a set of polynomials, the ideal generated by S is the span of S in the language of linear algebra and the solution set of S is the *variety of* S, denoted $V(S)$, and is referred to as the null space or kernel of S in linear algebra. Now recall the method of Gaussian elimination above, where a system was transformed into a system with the same solution set, but whose solutions were much easier to find. Gaussian elimination is the division algorithm in disguise since the ideal generated by S is the same ideal as that generated by $g = \gcd(S)$, and it's easier to find $V(g)$ then $V(S)$, additionally the mechanics of both processes are very similar. Thus the problem of solving systems of single variable polynomials has been put to rest, but before we can generalize the above results to multivariate polynomials we must take brief detour into term orders.

4. Term Orders

Earlier it was mentioned that an arbitrary ordering on the terms of each equation is important to have. In the linear case this ordering is as simple to come by as putting the coefficients of the same variable in the same column of a matrix and in the single variable polynomial case we used the natural ordering based on the degree.[3] With multivariate polynomials these two natural orders are combined to to form the degree lexicographic, or deglex, and degree reverse lexicographic, or degrevlex. For simplicities sake I will only make use of the degrevlex ordering, see [1] or [2] for details on deglex.

[2]For the algebrists in the audience it may be of interest to consider the coset representatives of $f + I$ in the quotient ring $\frac{k[x]}{I}$. Also one should determine a basis for the vector space over $\frac{k[x]}{I}$. This is discussed in [1] and elsewhere.

[3]This orderings are formally referred to as lexicographic, or lex, and degree, or deg, respectively.

Definition 6 (Degree Reverse Lexicographic Ordering)**.** Let $x > y > z$ be a lex ordering and $\vec{x}^{\alpha} = x^{\alpha_1}y^{\alpha_2}z^{\alpha_3}$. Then $\vec{x}^{\alpha} < \vec{x}^{\beta}$ iff one of the following is true.

(a) $\alpha_1 + \alpha_2 + \alpha_3 < \beta_1 + \beta_2 + \beta_3$

(b) $\alpha_1 + \alpha_2 + \alpha_3 = \beta_1 + \beta_2 + \beta_3$ and the first coordinates α_i and β_i from the right which are different satisfy $\alpha_i > \beta_i$.[4]

A few examples to illustrate the above definition are useful. Consider the polynomial $f = 3x^4z - 2x^3y^4 + 7x^2y^2z - 8xy^3z^3 + 2$. Ordering the terms of f with respect to the degrevlex term order produces $x^3y^4 > xy^3z^3 > x^4z > x^2y^2z > 2$. Having a well-defined term order (which degrevlex is) allows us to use the leading coefficient, power product, and term (lc, lp, lt) when addressing the issue of divisibility and exercise 1.4.9 of [1] also demonstrates the use of the degrevlex term order to solve a special case of the problem of ideal membership.

5. Multivariate Polynomial Division

The goal of this section is to combine the ideas and methods of the linear and single variable polynomial cases so we can determine a better generating set of a multivariate polynomial ideal. For the remainder of this section we will assume that a term order is fixed on the polynomial ring $k[x_1, \ldots, x_n]$.

Given any two polynomials $f, g \in k[x_1, \ldots, x_n]$ we can define reduction of f modulo g as in the one variable case. As in the one-variable case we can also reduce a polynomial by a set of polynomials. This is indeed the multivariate division algorithm we desire and is very similar to the single variable division algorithm.

Algorithm 2. INPUT: $f, f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$ where $f_i \neq 0 \ \forall \ 1 \leq i \leq s$.
OUTPUT: $u_1, \ldots, u_s, r$ such that $f = u_1f_1 + \cdots + u_sf_s + r$ and r is reduced with respect to $\{f_1, \ldots, f_s\}$ and $\max(lp(u_1)lp(f_1), \ldots, lp(u_s)lp(f_s), r) = lp(f)$.

[4]Notice that the rightmost variable is the smallest, so a smaller variable with a larger power make the entire power product less than another with the same degree.

INITIALIZATION: $u_i = r = 0$ and $h = f$

WHILE $h \neq 0$ **IF** $\exists\ i$ such that $lp(f_i)|lp(h)$

THEN choose the least such i

$u_i = u_i + \frac{lt(h)}{lt(f_i)}$

$h = h - \frac{lt(h)}{lt(f_i)} f_i$

ELSE

$r = r + lt(h)$

$h = h - lt(h)$

6. Gröbner Bases

With our multivariate division algorithm in hand we are prepared to finally introduce a Gröbner bases.

Definition 7. A set of non-zero polynomials $G = \{g_1, \ldots, g_s\}$ contained in an ideal I is called a Gröbner basis for I iff $\forall\ f \in I$ such that $f \neq 0$, $\exists\ i \in \{1, \ldots, s\}$ such that $lp(g_i) \mid lp(f)$.

From this definition several questions immediately arise. First is that of existence, and second is the issue of uniqueness. For any ideal, we are must show that a Gröbner basis must exist and we would like this basis to be unique. We will prove existence, but we must stregthen the idea of a Gröbner basis before we can get a uniqueness result.

As with most mathematical objects, it will be useful to have several characterizations of a Gröbner basis. We will need the idea of a leading term ideal to do this. Let $S \subseteq k[x_1, \ldots, x_n]$, then the leading term ideal of S is

$$Lt(S) = \langle lt(s) : s \in S \rangle. \tag{8}$$

Theorem 3. *Let I be a non-zero ideal of $k[x_1, \ldots, x_n]$. The following are equivalent for a set of non-zero polynomials $G = \{g_1, \ldots, g_s\} \subseteq I$:*

(1) *G is a Gröbner basis for I.*

(2) $f \in I$ *iff* $f \equiv 0 \mod (G)$.

(3) $Lt(G) = Lt(I)$.

(4) $f \in I$ *iff* $f = \sum_{i=1}^{s} h_i g_i$ *with* $lp(f) = \max_{1 \leq i \leq s}(lp(h_i)lp(g_i))$.

From this theorem it is obvious that a Gröbner basis for I is a generating set of I. Slightly more complicated results follow.

Theorem 4. *Let $I = \langle S \rangle$ be an ideal generated by non-zero terms and let $f \in k[x_1, \ldots, x_n]$. Then $f \in I$ iff for every term X appearing in f there exists $Y \in S$ such that $Y|X$. Moreover, there exists a finite subset S_0 of S such that $I = \langle S_0 \rangle$.*

Corollary 1. *Every non-zero ideal I of $k[x_1, \ldots, x_n]$ has a Gröbner basis.*

Proof. By (4) $Lt(I)$ has a finite generating set (this also follows from $k[x_1, \ldots, x_n]$ being a Noetherian ring) which can be assumed to be of the form $\{lt(g_1), \ldots, lt(g_s)\}$ with $g_i \in I$. Letting $G = \{g_1, \ldots, g_s\}$ we have $Lt(G) = Lt(I)$ so G is a Gröbner basis for I. □

We have now settled the issue of existence. At this point we need to address uniqueness and the related issue of constructing a Gröbner basis. The following theorem is another useful characterization of a Gröbner basis that is related to a different issue of uniqueness.

Theorem 5. *Let $G = \{g_1, \ldots, g_s\}$ be a set of non-zero polynomials of $k[x_1, \ldots, x_n]$. Then G is a Gröbner basis iff $\forall\ f \in k[x_1, \ldots, x_n]$, the remainder of the division of f by G is unique.*

We must note that a Gröbner basis with respect to a fixed term-order is not necessarily a Gröbner basis with respect to a different term-order. Thus all of our actions are taking place with a term-order already fixed.

Next we will determine how to compute a Gröbner basis since this will shed light on the uniqueness of the computed basis. The following algorithm was first given by Buchberger, while most of our earlier results were already known at his time but they are clearly not very useful with out a means of computing a Gröbner basis. Before we present Buchberger's theorem and the associated algorithm, we need to introduce S-polynomials.

Definition 8. Let $0 \neq f, g \in k[x_1, \ldots, x_n]$ and let $L = lcm(lp(f), lp(g))$. The polynomial

$$S(f,g) = \frac{L}{lt(f)}f - \frac{L}{lt(g)}g \tag{9}$$

is called the S-polynomial of f and g.

S-polynomials introduce a way of canceling the leading terms and thereby remove some of the ambiguity from the division algorithm. Buchberger proved that S-polynomials actually remove all the ambiguity.

Theorem 6 (Buchberger's Theorem)**.** *Let $G = \{g_1, \ldots, g_s\}$ be a set of non-zero polynomials in $k[x_1, \ldots, x_n]$. Then G is a Gröbner basis for the ideal $I = \langle G \rangle$ iff for all $i \neq j$*

$$f \xrightarrow{G_+} 0.$$

This gives rise to Buchberger's algorithm for computing a Gröbner basis.

Algorithm 3. INPUT: $F = \{f_1, \ldots, f_s\} \subseteq k[x_1, \ldots, x_n]$ with $f_i \neq 0$.

OUTPUT: $G = \{g_1, \ldots, g_t\}$ a Gröbner basis for $I = \langle F \rangle$.

INITIALIZATION: $G = F$, $\mathfrak{G} = \{\{f_i, f_j\} : f_i \neq f_j \in G\}$.

WHILE $\mathfrak{G} \neq \emptyset$

Choose any $\{f, g\} \in \mathfrak{G}$

$\mathfrak{G} = \mathfrak{G} - \{\{f, g\}\}$

$S(f,g) \xrightarrow{G_+} h$

IF h $\neq 0$ **THEN:**

$\mathfrak{G} = \mathfrak{G} \cup \{\{u, h\} : \forall\ u \in G\}$

$G = G \cup h$

Now that we can obtain a Gröbner basis, we would like to obtain a unique basis. This is done by putting restrictions on the polynomials in our basis, and is called a reduced Gröbner basis. From Buchberger's algorithm we can obtain different a different basis by changing the order polynomials are input and the random choice of polynomials out of $\mathfrak{G}$. First let

us impose restrictions to obtain a minimal Gröbner basis, then we will proceed to a reduced basis.

Definition 9. A Gröbner basis $G = \{g_1, \ldots, g_s\}$ is minimal if for all i, $lc(g_i) = 1$ and for $i \neq j$, $lp(g_i)$ does not divide $lp(g_j)$.

The following result shows that a minimal Gröbner basis is minimal in the sense of having the least number of polynomials.

Theorem 7. *Let $G = \{g_1, \ldots, g_n\}$ be a Gröbner basis for an ideal I. If $lp(g_2)|lp(g_1)$ then $\{g_2, \ldots, g_n\}$ is also a Gröbner basis for I.*

Proof. Let f be a polynomial in I such that $lp(g_1)|lp(f)$, then $lp(g_2)|lp(f)$ so the set $G \setminus \{g_1\}$ still satisfies the definition of a Gröbner basis. □

Clearly this theorem applies to any polynomial in the basis, so a minimal basis is obtained by dividing each polynomial by its leading coefficient and then removing all the polynomials whose leading power product is divisible by the leading power product of another polynomial in the basis. A minimal basis in not yet unique though, because we will often have a choice as to what elements to remove from the basis. For example if our basis is $\{y^2 + yx + x^2, y + x, y, x^2, x\}$ for the ideal $\langle y^2 + yx + x^2, y + x, y\}$ with the lex $y > x$ term order then both $\{y, x\}$ and $\{y + x, x\}$ are minimal Gröbner bases. However, a minimal Gröbner basis is closer to being unique as the next theorem shows.

Theorem 8. *Let $G = \{g_1, \ldots, g_n\}$ and $F = \{f_1, \ldots, f_t\}$ be minimal Gröbner bases for an ideal I. Then $n = t$ and after possible renumbering $lt(f_i) = lt(g_i)$ for all $1 \leq i \leq n$.*

The proof is left as an exercise to the reader.

Definition 10. A Gröbner basis $G = \{g_1, \ldots, g_n\}$ is reduced if for all i $lc(g_i) = 1$ and g_i is reduced with respect to $G \setminus \{g_i\}$.

A reduced basis is different from a minimal basis in the sense that in a minimal basis we looked only at the leading power product. In a reduced basis no non-zero term of g_i is

divisible by $lp(g_j)$ for all $j \neq i$. From a minimal basis we obtain a reduced basis by reducing g_1 by $G \setminus \{g_1\}$ to get h_1 and then replacing g_1 by h_1 in G. Then repeat for all remaining polynomials in G. Buchberger first proved the following important result concerning reduced Gröbner bases.

Theorem 9 (Buchberger)**.** *Fix a term order. Then every non-zero ideal I has a unique reduced Gróbner basis with respect to this term order.*

The existence of a reduced basis follows easily from earlier existence results and our method of obtaining the reduced basis. The uniqueness portion is also easily obtained and is left to the reader.

Thus we have completed our task of introducing the idea of a Gröbner basis. We have shown that a multivariate polynomial ring with a fixed term order has a unique reduce basis for any ideal, which can be easily computed once the ideal has been given in terms of a finite generating set.

There are many applications of Gröbner bases, primarily to commutative algebra and algebraic geometry, but are also useful in probability theory. It should be mentioned that similar results also exist for polynomial rings over a general ring as opposed to a field. In the case of a commutative Noetherian ring little changes, but in an arbitrary ring some interesting difficulties arise.

References

1. William Adams and Philippe Loustaunau, *An introduction to gröbner bases*, first ed., American Mathematical Society, Rhode Island, 1996.
2. Bruno Buchberger and Franz Winkler (eds.), *Gröbner bases and applications*, Cambridge, UK, Cambridge University Press and the London Mathematical Society, 1998, London Mathematical Society Lecture Note Series 251.
3. I. N. Herstein, *Topics in algebra*, Blaisdell, London, 1964.

E-mail address: jpreszler@ups.edu