

# INTRODUCTION TO GROBNER BASES

MARCH 1 2006  
SIVAKUMAR RATHINAM

## 1. INTRODUCTION

Grobner Bases was first introduced by Bruno Buchberger in his Ph.D. dissertation work (1965). They are named after Buchberger's advisor, Wolfgang Grobner.

## 2. MOTIVATION

We have powerful tools to solve a set of linear equations in linear algebra. For example, say we have a set of three equations  $x + y - z = 1$ ,  $x + 3y - 8z = -1$  and  $x - z = 2$ . We can pick an order on the variables  $x, y, z$  and then use the Gauss-Jordan elimination scheme to solve them. The basic method is to find a simpler set of equations from the given original set that has the same solution space. The main idea behind the Gauss-Jordan elimination scheme is to express equations (coefficients) in terms of other equations (coefficients). Now the question is as follows: Can we extend or apply this idea to solving a set of polynomial equations in one variable or multiple variables? Grobner bases tries to answer this question for the multiple variable case.

## 3. DEFINITIONS

A monomial is the product of non-negative integer powers of a set of variables. It is of the form  $x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$  where  $\alpha_i$  are non-negative integers. A polynomial is a formal expression representing a function  $f : \mathcal{K}^n \rightarrow \mathcal{K}$  such that  $f(x) = \sum_{\alpha} a_{\alpha} x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$  where  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  and  $a_{\alpha}$  belongs to the field  $\mathcal{K}$ . The ideas presented here are applicable to any polynomial ring  $\mathcal{K}[x_1, \dots, x_n]$ . An affine variety  $V(f_1, f_2, \dots, f_s)$  is the set of common zeros of a collection of polynomials  $\{f_1, f_2, \dots, f_s\}$ . That is  $V(f_1, f_2, \dots, f_s) = \{x \in \mathcal{K}^n : f_1(x) = 0, f_2(x) = 0, \dots, f_s(x) = 0\}$ .

An ideal  $\mathcal{I}$  is a subset of elements of  $\mathcal{K}[x_1, x_2, \dots, x_n]$  that has the following properties:

- $0 \in \mathcal{I}$
- If  $f \in \mathcal{I}$  and  $g \in \mathcal{I}$  then  $f + g \in \mathcal{I}$ .
- $f \in \mathcal{I}$  and  $h \in \mathcal{K}[x_1, x_2, \dots, x_n]$  implies the product  $h * f \in \mathcal{I}$ .

The ideal generated by functions  $f_1, f_2, \dots, f_s$ , denoted by  $\langle f_1, f_2, \dots, f_s \rangle$ , is defined as the set  $\{f : f = 0 \text{ at all points of } V(f_1, f_2, \dots, f_s)\}$ .

## 4. QUESTIONS THAT WE ARE INTERESTED IN:

- (1) Is the given function  $f$  in the ideal  $\langle f_1, f_2, \dots, f_s \rangle$ ?
- (2) Solve the set of polynomial equations given by  $f_1, f_2, \dots, f_s$ .

## 5. POLYNOMIALS OF SINGLE VARIABLE

In this section we deal with ideals over  $\mathcal{K}[x]$ . The polynomials are of the form  $f(x) = a_0x^m + a_1x^{m-1} + \dots + a_m$ . Here the degree of  $f$ , denoted by  $\deg(f)$ , is  $m$ . The leading term of  $f$ , denoted by  $LT(f)$ , is  $a_0x^m$ . There are two main theorems that answer the questions posed in section 4 for the single variable case.

**Theorem 5.1.** *Every ideal  $I$  of  $\mathcal{K}[x]$  is of the form  $\langle g \rangle$  for some  $g \in \mathcal{K}[x]$ .*

*Proof.* Let  $g$  be a polynomial of minimum degree in  $I$ . Consider  $f \in I$ . Given two polynomials  $f$  and  $g$ , we can write  $f$  uniquely as  $f = qg + r$  where  $r = 0$  or  $\deg(r) < \deg(g)$ . Here  $q$  is called the quotient and  $r$ , the remainder. This implies that  $r = f - qg$ . Hence  $r \in I$ . By the choice of  $g$ , the degree of polynomial  $r$  cannot be less than  $g$ . Therefore  $r = 0$ . In other words, every polynomial in the ideal can be generated from  $g$ .  $\square$

This theorem converts the problem of solving a set of polynomial equations into solving a single equation in  $g$ . The following theorem states that this equation  $g$  is nothing but the greatest common divisor of the given set of polynomials. Recall that  $g = GCD(f_1, f_2, \dots, f_s)$  is the greatest common divisor of polynomials  $f_1, f_2, \dots, f_s$  provided the following conditions hold:

- $g$  divides  $f_i$  for all  $i \in 1..s$
- any other polynomial  $p$  divides  $f_i$ , for all  $i \in 1..s$ , then  $p$  must divide  $g$ .

**Theorem 5.2.** *If  $f_1, f_2, \dots, f_s$  are polynomials in  $\mathcal{K}[x]$ , then  $\langle f_1, f_2, \dots, f_s \rangle = \langle GCD(f_1, f_2, \dots, f_s) \rangle$ .*

*Proof.* Let the function  $g$  be the polynomial with minimum degree in  $\langle f_1, f_2, \dots, f_s \rangle$  as chosen in theorem 5.1. Clearly  $g$  divides each of the polynomials  $f_i$  for all  $i \in 1..s$ . Now, we also know since  $g \in \langle f_1, f_2, \dots, f_s \rangle$ , we can write  $g = \sum_{i \in \{1..s\}} h_i f_i$ . If any other polynomial  $p$  divides  $f_i \forall i \in 1..s$ , then one can write  $f_i = l_i p$  for all  $i \in 1..s$ . Substituting for  $f$  we have  $g = \sum_{i \in \{1..s\}} h_i l_i p$ . This implies that the polynomial  $p$  divides polynomial  $g$ . Hence proved.  $\square$

Now, using the fact that  $GCD(f_1, f_2, \dots, f_s) = GCD(GCD(f_1, f_2), f_3, \dots, f_s)$ , we only need a method to compute the GCD of two polynomials.

5.1. GCD of two polynomials  $f_1$  and  $f_2$ .

- Let  $f := f_1, g := f_2$ .
- while  $g \neq 0$ , do the following:  
    Compute the remainder  $r$  in  $f = qg + r$ . Let  $f := g; g := r$ .
- return  $f, g$ .

$GCD(f_1, f_2) = \frac{1}{lc(f)} f$ , where  $lc(f)$  is the coefficient of the leading term in  $f$ .

## 5.2. Division algorithm to compute the remainder.

- Let  $q := 0, r := f$ .
- while  $r \neq 0$  and  $LT(g)$  divides  $LT(r)$ , do the following:  
     $r := r - g \frac{LT(r)}{LT(g)}; q := q + \frac{LT(r)}{LT(g)}$
- return  $q, r$ .

$$y^2 - 1 \left| \begin{array}{r} 1 \\ \hline x + y^2 + y \\ y^2 - 1 \\ \hline x + y + 1 \end{array} \right.$$

FIGURE 1

## 6. POLYNOMIALS OF MULTIPLE VARIABLES

In polynomials with one variable the leading term for division is based on the degree of the variable. In the multi variate case, one needs to define an ordering on the variables to define the leading term. We present here an ordering called the lexicographic ordering.

Let  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  and  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ .  $\alpha > \beta$  if  $\exists i$  such that  $\alpha_i > \beta_i$  and  $\alpha_j = \beta_j$  for  $j < i$ . For example  $x^2y + xy^2 + y + 3$  is arranged according to this ordering with  $x^2y$  being the leading term.

Now, let's formulate the division rules that we desire by looking at examples. Refer to figure 1. Note that one cannot stop the division by just looking at the first term. That is, we must divide  $LT(g)$  into the lexicographically first term in  $f$  that it divides. Let  $f = \sum h_i f_i + r$ . The desired property we want while finding  $h_i$  and  $r$  is that no term in  $r$  should be a multiple of any of the leading terms of any one of the  $f_i$ . We would like  $r$  to be independent of the order in which we do the divisions. Refer to figures 2 and 3. In figure 2, the function  $f = x^2 + xy^2 + y^2$  is first divided by  $xy - 1$  and then by  $y^2 - 1$ . In figure 3, the function  $f = x^2 + xy^2 + y^2$  is first divided by  $y^2 - 1$  and then by  $xy - 1$ . Note that the remainders are different. Grobner bases essentially avoids this problem.

UNIVERSITY OF CALIFORNIA, BERKELEY - 94720

*E-mail address:* [rsiva@berkeley.edu](mailto:rsiva@berkeley.edu)

$$\begin{array}{r}
 xy-1 \overline{) \begin{array}{l} x+y \\ x^2y+xy^2+y^2 \\ x^2y-x \\ \hline xy^2+y^2+x \\ xy^2-y \\ \hline y^2+x+y \end{array} }
 \end{array}$$

$$\begin{array}{r}
 y^2-1 \overline{) \begin{array}{l} 1 \\ y^2+x+y \\ y^2-1 \\ \hline x+y+1 \end{array} }
 \end{array}$$

FIGURE 2. Divide first using  $xy - 1$  and then using  $y^2 - 1$ 

$$\begin{array}{r}
 y^2-1 \overline{) \begin{array}{l} x+1 \\ x^2y+xy^2+y^2 \\ xy^2-x \\ \hline x^2y+y^2+x \\ y^2-1 \\ \hline x^2y+x+1 \end{array} }
 \end{array}$$

$$\begin{array}{r}
 xy-1 \overline{) \begin{array}{l} x \\ x^2y+x+1 \\ x^2y-x \\ \hline 2x+1 \end{array} }
 \end{array}$$

FIGURE 3. Divide first using  $y^2 - 1$  and then using  $xy - 1$