Alan Hahn
8500 Project

REFINEMENTS OF THE GB CRITERION AND IMPROVEMENTS OF BUCHBERGER'S ALGORITHM

0.1. **Introduction.** Recall that there are many equivalent definitions of a Gröbner basis; in order to talk about one of the definitions of interest used in this paper, it is useful to remind the reader of the definition of a syzygy polynomial:

Definition 1. The syzygy polynomial of f and g, $S(f,g)$ is defined as

$$S(f,g) = \frac{lcm(LM(f),LM(g))}{LT(f)} \cdot f - \frac{lcm(LM(f),LM(g))}{LT(g)} \cdot g.$$

With this in mind, recall this equivalent definition of a Gröbner basis:

Definition 2. (GB Criterion 1) A basis $G = \{g_1,...,g_s\}$ for an ideal $I \subseteq k[x_1,...,x_n]$ is a Gröbner basis for I iff for all pairs $i \neq j$, $\text{rem}_G(S(g_i,g_j)) = 0$, where $\text{rem}_G(S(g_i,g_j))$ is the remainder of $S(g_i,g_j)$ under division by the ordered set $G = \{g_1,...,g_s\}$.

This definition is of interest as it lends itself to an algorithmic construction of a Gröbner basis, given by Buchberger's Algorithm:

Definition 3. (Buchberger's Algorithm) Let $0 \neq I = \langle f_1,...,f_s\rangle$ be an ideal in $k[x_1,...,x_n]$. A Gröbner basis for I may be constructed as follows:
$G := \{f_1,...,f_s\}$
For all $i \neq j$, if $\text{rem}_G(S(f_i,f_j)) = r \neq 0$, then $G := G \cup \{r\}$. Restart with this new G, and do until all remainders are zero.

As 3 is an algorithm it is guaranteed to terminate; however, Buchberger's algorithm is relatively slow, with a major pitfall being that computing remainders is computationally expensive, so reducing the number of remainders $\text{rem}_G(S(f_i,f_j))$ which need to be checked would be advantageous. An easy observation to be made is that once a remainder for a pair $\text{rem}_G(S(f_i,f_j))$ has been seen to be zero, there is no reason to check again, as the remainder is still zero after adjoining new elements to G. Indeed, if new generators f_j are added one at a time, the only remainders that need to be checked are $\text{rem}_G(S(f_i,f_j))$, where $i \leq j-1$. This improvement is somewhat superficial, and the goal of this paper is to describe some deeper improvements which may be made.

0.2. **Improvements to Buchberger's Algorithm.** The first improvement is fairly immediate following a definition and a bit of discussion. First, a definition:

Definition 4. Fix a monomial order and let $G = \{g_1,...,g_t\} \subseteq k[x_1,...,x_n]$. Given $f \in k[x_1,...,x_n]$, we say that f **reduces to zero modulo** G, written

$$f \to_G 0,$$

if f has a **standard representation**

$$f = A_1g_1 + ... + A_tg_t,\ A_i \in k[x_1,...,x_n],$$

such that whenever $A_ig_i \neq 0$, then

$$\text{multideg}(f) \geq \text{multideg}(A_ig_i).$$

From this definition immediately follows a lemma:

Lemma 5. Let $G = (g_1, ..., g_t)$ be an ordered set of elements of $k[x_1, ..., x_n]$ and fix $f \in k[x_1, ..., x_n]$. Then $\text{rem}_G(f) = 0 \implies f \to_G 0$. The converse is not true in general.

Proof. $\text{rem}_G(f) = 0 \implies f = q_1 g_1 + ... + q_t g_t + 0$ with $\text{multideg}(f) \geq \text{multideg}(q_i g_i)$ whenever $q_i g_i \neq 0$. □

The utility of this lemma in the context of this paper is that it gives rise to the following equivalent definition for a Gröbner basis:

Proposition 6. (GB Criterion 2) A basis $G = \{g_1, ..., g_t\}$ for an ideal I is a Gröbner basis iff $S(g_i, g_j) \to_G 0$ for all $i \neq j$.

Proof. The forward implication is given by 5. The other implication is due to noting that in the proof of 2, all that was used was that $S(g_i, g_j)$ has a standard representation which satisfies the conditions of definition 4. □

For the details of the proof of 2, see (Cox, Little, & O'Shea, 2015). With this discussion, an improvement to Buchberger's algorithm may be introduced:

Proposition 7. (Improvement 1) Given a finite set $G \subseteq k[x_1, ..., x_n]$, suppose that for $f, g \in G$, $\text{LM}(f)$ and $\text{LM}(g)$ are relatively prime. Then $S(f, g) \to_G 0$.

Proof. For simplicity, assume f, g have been multiplied by appropriate constants to make $\text{LC}(f) = \text{LC}(g) = 1$. Write $f = \text{LM}(f) + p$, $g = \text{LM}(g) + q$. As $\text{LM}(f)$, $\text{LM}(g)$ are relatively prime, then $\text{lcm}(\text{LM}(f),\text{LM}(g)) = \text{LM}(f)\text{LM}(g)$. Thus the S-polynomial $\text{S}(f, g)$ may be written

$$S(f, g) = LM(g)f - LM(f)g = (g - q)f - (f - p)g = gf - qf - fg + pg = pg - qf.$$

It remains to note that $pg - qf$ is a standard representation which satisfies 4. □

This is an improvement because for pairs f_i, f_j of elements in a potential Gröbner basis for a polynomial ideal, if $\text{LM}(f_i)$ and $\text{LM}(f_j)$ are relatively prime, then $S(f_i, f_j) \to_G 0$ and it is not necessary to do any polynomial divisions for this pair, thus reducing a series of polynomial divisions to a simple check of whether two monomials are relatively prime.

The next improvement requires a bit of theory; first a definition is in order.

Definition 8. Let $F = (f_1, ..., f_s)$. A **syzygy** on the leading terms $LT(f_1), ..., LT(f_s)$ of F is an s-tuple of polynomials $S = (h_1, ..., h_s) \in (k[x_1, ..., x_n])^s$ such that

$$\sum_{i=1}^{s} h_i \cdot LT(f_i) = 0.$$

Denote by $S(F)$ the subset of $(k[x_1, ..., x_n])^s$ consisting of all syzygies on the leading terms of F.

As an example of a syzygy, consider the following:

Example 9. Consider $F = (x, x^2 + z, y + z)$ with the lex ordering. Then $S = (-x + y, 1, -x)$ defines a syzygy in $S(F)$:

$$(-x + y) \cdot LT(x) + 1 \cdot LT(x^2 + z) + -x \cdot LT(y + z) = -x^2 + xy + x^2 - xy = 0.$$

Note that $S(F)$ is closed under addition and multiplication in the sense that for syzygies $S = (s_1, ..., s_t)$, $U = (u_1, ..., u_t) \in S(F)$, then $S + U := (s_1 + u_1, ..., s_t + u_t) \in S(F)$, and for $g \in k[x_1, ..., x_n]$, then $gS := (gs_1, ..., gs_t) \in S(F)$. In light of this it is not hard to see that, taking $e_i = (0, ..., 0, 1, 0, ..., 0) \in (k[x_1, ..., x_n])^s$ where 1 is in the ith place, then a syzygy $S \in S(F)$ may be written $S = \sum_{i=1}^{s} h_i e_i$.

From the discussion above, $\{e_i\}$ may be thought of as a basis for S(F) in some sense. While the $\{e_i\}$ are useful, for this paper it is useful to also look at some other elements of $S(F)$, namely homogeneous elements:

Definition 10. An element $S \in S(F)$ is **homogeneous of multidegree** α, where $\alpha \in \mathbb{Z}_{\geq 0}$, provided that

$$S = (c_1 x^{\alpha(1)}, ..., c_s x^{\alpha(s)}),$$

where $c_i \in k$ and $\alpha(i) + multideg(f_i) = \alpha$ whenever $c_i \neq 0$.

As an example of a homogeneous element of $S(F)$, consider:

Definition 11.

$$S_{ij} := \frac{lcm(LM(f_i), LM(f_j))}{LT(f_i)} \cdot e_i - \frac{lcm(LM(f_i), LM(f_j))}{LT(f_j)} \cdot e_j.$$

Note S_{ij} is homogeneous of degree multideg($lcm(LT(f_i), LT(f_j))$).

.

Definition 11 may remind the reader of the syzygy polynomial of f_i, f_j, and for good reason. With the following definition, the connection between the two may be made.

Definition 12. For $S = (H_1, ..., H_s) \in S(F)$,

$$S \cdot F := \sum_{i=1}^{t} H_i f_i.$$

With definition 12 in mind, it is clear to see that $S_{ij} \cdot F = S(f_i, f_j)$. An example may be helpful:

Example 13. Consider $G = (x^2y^2 + z, xy^2 - y, x^2y + yz)$. Then
$S_{1,2} = 1(1, 0, 0) - x(0, 1, 0) = (1, -x, 0)$.
Note $S_{ij} \cdot G = 1(x^2y^2 + z) - x(xy^2 - y) = x^2y^2 + z - x^2y^2 + xy = xy + z = S(g_i, g_j)$.

Now, some main results:

Lemma 14. Every element of $S(F)$ can be written uniquely as a sum of homogeneous elements of $S(F)$.

Proof. Fix $\alpha \in \mathbb{Z}_{\geq 0}$, and let $h_{i\alpha}$ be the term of h_i such that $h_{i\alpha} f_i$ has multidegree α, if such term exists. Then $\sum_{i=1}^{s} h_{i\alpha} LT(f_i) = 0$ as $h_{i\alpha} LT(f_i)$ are the terms of multidegree α in the sum $\sum_{i=1}^{s} h_i LT(f_i) = 0$.
Thus $S_\alpha = (h_{1\alpha}, ..., h_{s\alpha})$ is a homogeneous element of $S(F)$ of degree α and $S = \sum_\alpha S_\alpha$. □

Thus $S(F)$ has a basis of homogeneous elements. In particular, $\{S_{ij}\}$ is such a homogeneous basis:

Proposition 15. Given $F = (f_1, ..., f_s)$, every syzygy $S \in S(F)$ can be written as

$$S = \sum_{i<j} u_{ij} S_{ij},$$

where $u_{ij} \in k[x_1, ..., x_n]$.

Proposition 15 leads to another equivalent definition of a Gröbner basis, which in turn is what is needed in order to state the second improvement to Buchberger's algorithm:

Proposition 16. (GB Criterion 3) A basis $G = (g_1, ..., g_t)$ for an ideal I is a Gröbner basis iff for every element $S = (H_1, ..., H_t)$ in a homogeneous basis for the syzygies $S(G)$, $S \cdot G \to_G 0$.

For proofs of 16 and 15, see (Cox et al., 2015). Now, the second improvement to Buchberger's algorithm may be stated:

Proposition 17. (Improvement 2) Given $G = (g_1, ..., g_t)$, suppose that $S \subseteq \{S_{ij} \mid 1 \leq i < j \leq t\}$ is a basis of $S(G)$. In addition, suppose we have distinct elements $g_i, g_j, g_l \in G$ such that $LT(g_l)$ divides $lcm(LT(g_i), LT(g_j))$. If $S_{il}, S_{jl} \in S$, then $S \setminus \{S_{ij}\}$ is also a basis of S(G).

Proof. For simplicity, suppose $i < j < l$. Set $x^{\gamma_{ij}} = lcm(LM(g_i), LM(g_j))$, and let $x^{\gamma_{il}}$ and $x^{\gamma_{jl}}$ be defined similarly. Then by assumption, $x^{\gamma_{il}}$, $x^{\gamma_{jl}}$ both divide $x^{\gamma_{ij}}$. It remains to note that

$$S_{ij} = \frac{x^{\gamma_{ij}}}{x^{\gamma_{il}}} S_{il} - \frac{x^{\gamma_{ij}}}{x^{\gamma_{jl}}} S_{jl}.$$

□

Similar to proposition 7, proposition 17 is an improvement to Buchberger's algorithm by further reducing the number of computations which need to be done. Taking $\{S_{ij}\}$ as a basis for the syzygies $S(G)$ for a potential Gröbner basis G, then by proposition 16, $S_{ij} \cdot G \to_G 0$ must be checked for each S_{ij}; proposition 17 allows the removal of some unneeded elements in the basis so that they won't needlessly be checked.

0.3. **Summary and Conclusion.** Recall that the first definition for a Gröbner basis given in this paper, 2, lends itself to an algorithm, namely Buchberger's algorithm, 3. Also, recall that for Buchberger's algorithm, $\text{rem}_G(S(f_i, f_j))$ is computed for each pair f_i, f_j in a potential Gröbner basis and that polynomial division is computationally expensive. In trying to reduce the number of polynomial divisions executed, some theory was developed which circumvents the need to calculate a remainder for every pair. The first result, 7 is based on 6, a second, equivalent definition for a Gröbner basis where the condition that $\text{rem}_G(S(f_i, f_j)) = 0$ for all f_i, f_j was updated to $S(f_i, f_j) \to_G 0$. This first result is used to reduce the number of polynomial divisions by instead checking a condition about the elements of a potential Gröbner basis, namely, checking whether LM(f_i) and LM(f_j) are relatively prime, as this then implies that $S(f_i, f_j) \to_G 0$. The second result, 17 is based on a third equivalent definition of a Gröbner basis, 16 where the condition $S(f_i, f_j) \to_G 0$ for all f_i, f_j was updated to $S \cdot G \to_G 0$ for every element S of a homogeneous basis for $S(G)$. This second result is used to reduce the number of polynomial divisions by instead checking a condition about elements of a homogeneous basis for $S(G)$ for a potential Gröbner basis G, to see whether any elements may be excluded from consideration. To see how these two results may be used together, it is useful to note that the third equivalent definition for a Gröbner basis, proposition 16, is a generalization of the second, proposition 6, as taking $\{S_{ij}\}$ as the homogeneous basis in 16 gives that the condition to check is $S_{ij} \cdot G \to_G 0$, which is exactly the condition to check in 6. Putting all of this together gives the main result of this paper, a more efficient version of Buchberger's algorithm:

Theorem 18. Let I = $\langle f_1, ..., f_s\rangle$ be a polynomial ideal. Then a Gröbner basis of I can be constructed in a finite number of steps by the following algorithm:
Input: $F = (f_1, ..., f_s)$
Output: a Gröbner basis G for $I = \langle f_1, ..., f_s\rangle$

```
$B := \{(i,j) \mid 1 \leq i < j \leq s\}$
$G := F$
$t := s$
While $B \neq \emptyset$ Do
          Select $(i,j) \in B$
          If $lcm(LT(f_i), LT(f_j)) \neq LT(f_i)LT(f_j)$ And Criterion$(f_i, f_j, B)$ = false Then
              $r := \text{rem}_G(S(f_i, f_j))$
              If $r \neq 0$ Then
                   $t := t + 1; f_t := r$
                   $G := G \cup \{f_t\}$
                   $B := B \cup \{(i,t) \mid 1 \leq i \leq t - 1\}$
          $B := B \setminus \{(i,j)\}$
Return $G$
```

Criterion(f_i, f_j, B) is true provided that there is some $l \notin \{i,j\}$ for which the pairs (i,l) and (j,l) are not in B and $LT(f_l)$ divides $lcm(LT(f_i), LT(f_j))$. This is based on proposition 17.

B is the set of indices (i,j) for which $S_{ij} \cdot G \to_G 0$ must be checked. It may be seen above that if $lcm(LT(f_i), LT(f_j)) = LT(f_i)LT(f_j)$, i.e. if the conditions of 7 hold, then that set of indices is removed from consideration as $S_{ij} \cdot G \to_G 0$ is then known to be true. If those conditions fail to hold, then next the conditions for 17 are checked, and again if those conditions hold, then that set of indices is removed from consideration as again $S_{ij} \cdot G \to_G 0$ is known to be true. If neither of the conditions of 7 or 17 hold, then the algorithm above reverts to algorithm 3 and continues with computing $\text{rem}_G(S(f_i, f_j))$.

References

Cox, D., Little, J., & O'Shea, D. (2015). *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra.* Springer.