University of Toronto, Spring 1995 Instructor: Shuhong Gao

CSC478S/2412S: Computer Algebra
Characterizations of Gröbner bases: a simple proof

Notations. Let $\mathbb{F}[X]$ denote $\mathbb{F}[x_1, x_2, \cdots, x_n]$ and T_x the set of terms: $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ for all $i_1, i_2, \cdots, i_n \geq 0$. Terms are ordered by some fixed admissible total ordering $<$. As in the text book, let hterm(p), $M(p)$, hcoeff(p) denote the head term, head monomial, head coefficient of $p \in \mathbb{F}[X]$, respectively, so that $M(p) = \text{hcoeff}(p) \cdot \text{hterm}(p)$. Let $G \subset \mathbb{F}[X]$ be a set of polynomials. For $p, q \in \mathbb{F}[X] \setminus \{0\}$, a reduction

$p \longmapsto_G q$ means that $q = p - ((\alpha t)/M(g))g$ where $p = \alpha t + r$ with $\alpha \in \mathbb{F} \setminus \{0\}$, $t \in T_x$, $g \in G$, $r \in \mathbb{F}[X]$ such that hterm$(g)|t$ and t does not appear in r.

Write $p \overset{+}{\longmapsto}_G q$ if $p = p_0 \longmapsto_G p_1 \longmapsto_G \cdots \longmapsto_G p_k = q$ for some polynomials $p_i \in \mathbb{F}[X]$, i.e. there exist a sequence of reductions that reduces p to q modulo G. Write $p \overset{*}{\longmapsto}_G q$ if $p \overset{+}{\longmapsto}_G q$ and q is irreducible (no term in q is divisible by the head term of any polynomial in G); we say that q is an irreducible successor of p modulo G.

Observe that for any monomial u, if $p \overset{+}{\longmapsto}_G q$ then $up \overset{+}{\longmapsto}_G uq$.

Lemma 1 *Let $f, g, h, h_1 \in \mathbb{F}[X]$ and $G \subset \mathbb{F}[X]$. If $f - g = h$ and $h \overset{+}{\longmapsto}_G h_1$ then there exist $f_1, g_1 \in \mathbb{F}[X]$ such that*

$$h_1 = f_1 - g_1, \quad f \overset{+}{\longmapsto}_G f_1, \quad g \overset{+}{\longmapsto}_G g_1.$$

Proof. Induction on the number k of steps in $h \overset{+}{\longmapsto}_G h_1$. Base case $k = 0$, trivial. Assume the lemma holds for $k - 1$. Suppose that $h \overset{+}{\longmapsto}_G h_2 \longmapsto_p h_1$ where $p \in G$ and the first part takes $k - 1$ steps. By induction hypothesis, there exist $f_2, g_2 \in \mathbb{F}[X]$ such that

$$h_2 = f_2 - g_2, \quad f \overset{+}{\longmapsto}_G f_2, \quad g \overset{+}{\longmapsto}_G g_2.$$

As $h_2 \longmapsto_p h_1$, we have $h_1 = h_2 - (c/b)u \cdot p$ where $b = \text{hcoeff}(p)$, $u \in T_x$, and $u \cdot \text{hterm}(p)$ is a term in h_2 with coefficient $c \neq 0$. Let c_1 and c_2 be the coefficients of the term $u \cdot \text{hterm}(p)$ in f_2 and g_2, respectively. Set

$$f_1 = f_2 - \frac{c_1}{b} u \cdot p, \quad g_1 = g_2 - \frac{c_2}{b} u \cdot p.$$

Then $f_2 \longmapsto_p f_1$ if $c_1 \neq 0$, and $g_2 \longmapsto_p g_1$ if $c_2 \neq 0$. So $f \overset{+}{\longmapsto}_G f_1$ and $g \overset{+}{\longmapsto}_G g_1$. Note that $c_1 - c_2 = c$ and $f_1 - g_1 = h_1$. The proof is complete. □

Note that when both f and g reduce to 0 modulo G, their sum $f + g$ may not reduce to 0 at all. For example, $G = \{x, x + 1\}$, $f = -x$ and $g = x + 1$. The question is when this does not happen. The following lemma answers this question partially.

Lemma 2 *Let $G \subset \mathbb{F}[X]$. If $f, g \in \mathbb{F}[X]$ always reduce to 0 in any full reductions modulo G,*

then so does $f + g$.

Proof. Suppose that $f+g \xmapsto{*}_G h$ and $h \neq 0$ is irreducible. By Lemma 1, there exist $f_1, g_1 \in \mathbb{F}[X]$ such that (note that $f+g = f-(-g)$)

$$h = f_1 + g_1, \quad f \xmapsto{+}_G f_1, \quad g \xmapsto{+}_G g_1. \tag{1}$$

Let $v = \mathrm{hterm}(h)$. Write

$$f_1 = h_1 + f_2, \quad g_1 = -h_1 + g_2, \quad h = f_2 + g_2,$$

where all the terms in h_1 are bigger than v and those in f_2 and g_2 are at most v.

By assumption, f and g always reduce to 0 in any full reductions modulo G. Now we reduce them in the following steps:

(1) $f \xmapsto{+}_G f_1, g \xmapsto{+}_G g_1$ as in (1).

(2) Reduce all the terms of $f_1 = h_1 + f_2$ and

$g_1 = -h_1 + g_2$ that are bigger than v, until the head terms become $\leq v$. In this step, no terms $\leq v$ were reduced. In particular, no terms in f_2 and g_2 were reduced. The reductions for f_1 and g_1 are the same as applied directly to h_1 (except for the minus sign for g_1). Let $h_1 \xmapsto{+}_G h_2$, where it is the first time that $\mathrm{hterm}(h_2) \leq v$. Then

$$f_1 = h_1 + f_2 \xmapsto{+}_G h_2 + f_2, \quad g_1 = h_1 + g_2 \xmapsto{+}_G -h_2 + g_2.$$

(3) Reduce $h_2 + f_2$ and $-h_2 + g_2$ to 0.

Since $h = (h_2 + f_2) + (-h_2 + g_2)$ and $v = \mathrm{hterm}(h)$, at least one of $h_2 + f_2$ and $-h_2 + g_2$ has v as its head term. But $v = \mathrm{hterm}(h)$ is irreducible modulo G by assumption, so one of them can not be reduced to 0 in step (3) (head term never cancel when reducing lower order terms). This is a contradiction. Thus h must be 0, and every full reduction of $f+g$ reduces to 0. □

Theorem 3 *If f and g have only one irreducible successor modulo G, say f_1 and g_1, respectively, then $f_1 + g_1$ is the only irreducible successor of $f+g$ modulo G.*

Proof. Since f_1 and g_1 are irreducible modulo G,

the assumption of the theorem implies that $f - f_1$ and $g - g_1$ always reduce to 0 modulo G. By Lemma 2,

$$(f+g) - (f_1 + g_1) = (f - f_1) + (g - g_1) \xmapsto{*}_G 0.$$

But $f_1 + g_1$ is irreducible, the result follows. □

Remark. A straight induction shows that Lemma 2 and Theorem 3 hold for any finite sum of polynomials.

Definition. The S-polynomial of $p, q \in \mathbb{F}[X]$ is defined as:

$$\mathrm{Spoly}(p,q) = \mathrm{lcm}(M(p), M(q)) \left[\frac{p}{M(p)} - \frac{q}{M(q)} \right].$$

Note that $\mathrm{hterm}(\mathrm{Spoly}(p,q)) < \mathrm{lcm}(M(p), M(q))$.

Theorem 4 *The following are equivalent:*

(i) *G is a Gröbner basis, i.e., $p \overset{*}{\longmapsto}_G 0$ for every $p \in < G >$ (the ideal generated by polynomials in G).*

(ii) *$Spoly(p, q) \overset{*}{\longmapsto}_G 0$ for all $p, q \in G$.*

(iii) *For any $p \in \mathbb{F}[X]$, if $p \overset{*}{\longmapsto}_G q$ and $p \overset{*}{\longmapsto}_G r$ then $q = r$.*

Proof. (i) $\Longrightarrow$ (ii) Since Spoly$(p, q) \in < G >$, it follows from the definition of Gröbner bases.

(iii) $\Longrightarrow$ (i) For $p \in < G >$, there exist $h_i \in \mathbb{F}[X]$, $g_i \in G$, $1 \leq i \leq k$, such that $p = \sum_{i=1}^{k} h_i g_i$. Since $h_i g_i$ reduces to 0 modulo g_i for $1 \leq i \leq k$, condition (iii) and Theorem 3 imply that $p \overset{*}{\longmapsto}_G 0$.

(ii) $\Longrightarrow$ (iii) Prove by induction on the head term of p. Base case: hterm$(p) = 1$. p is either irreducible if G contains no nonzero constant or reduce to 0 otherwise. Assume that (iii) holds for all p with hterm$(p) < t$ for a fixed term $t \in T_x$. We want to prove that (iii) holds for all p with hterm$(p) = t$.

If t is irreducible modulo G, then all reductions are in the lower order terms of p, i.e., $p - M(p)$ which always reduces to the same irreducible polynomial by induction hypothesis.

So assume that t is reducible modulo G. Let

$$p \overset{*}{\longmapsto}_G q, \quad p \overset{*}{\longmapsto}_G r.$$

Consider the first time t is reduced. Suppose that

$$p = M(p) + (p - M(p)) \overset{+}{\longmapsto}_G M(p) + q_1 \longmapsto_{g_1} p_1 + q_1 \overset{*}{\longmapsto}_G q,$$

and

$$p = M(p) + (p - M(p)) \overset{+}{\longmapsto}_G M(p) + r_1 \longmapsto_{g_2} p_2 + r_1 \overset{*}{\longmapsto}_G r,$$

where $g_1, g_2 \in G$ whose head terms divide $t =$ hterm(p) and

$$p_1 = M(p) - \frac{M(p)}{M(g_1)} g_1, \quad p_2 = M(p) - \frac{M(p)}{M(g_2)} g_2.$$

Since the head terms of $p - M(p), q_1, r_1, p_1, p_2, p_1 + q_1, p_2 + r_1$ are all $< t$, each of them has only one irreducible successor by induction hypothesis. Note that q_1 and r_1 have the same irreducible successor as $p - M(p)$, say p_0.

It follows from Theorem 3 that $q = \tilde{p}_1 + p_0$ and $r = \tilde{p}_2 + p_0$ where $p_1 \overset{*}{\longmapsto}_G \tilde{p}_1$ and $p_2 \overset{*}{\longmapsto}_G \tilde{p}_2$.

Now we just need to prove that $\tilde{p}_1 = \tilde{p}_2$. Note that

$$\begin{aligned} p_1 &= (p_1 - p_2) + p_2 \\ &= M(p)\left(\frac{g_2}{M(g_2)} - \frac{g_1}{M(g_1)}\right) + p_2 \\ &= \alpha u \cdot \text{Spoly}(g_2, g_1) + p_2, \end{aligned}$$

where $\alpha \in \mathbb{F} \setminus \{0\}$,

$u = M(p)/\text{lcm}(M(g_1), M(g_2)) \in T_x$. As $\text{Spoly}(g_2, g_1) \overset{*}{\longmapsto}_G 0$, we have $\alpha u \cdot \text{Spoly}(g_2, g_1) \overset{*}{\longmapsto}_G 0$. Since the head term of $\text{Spoly}(g_2, g_1)$ is less than

$\text{lcm}(M(g_1), M(g_2))$, the head term of $\alpha u \cdot \text{Spoly}(g_2, g_1)$ is less than t. By induction hypothesis, $\alpha u \cdot \text{Spoly}(g_2, g_1)$ has only one irreducible successor, i.e. 0. Therefore, by Theorem 3 again, p_1 and p_2 have the same irreducible successor, i.e., $\tilde{p}_1 = \tilde{p}_2$. This completes the induction, and thus the proof. □

Corollary 5 *G is a Gröbner basis if and only if for all $p, q \in G$ either $Spoly(p, q) \overset{*}{\longmapsto}_G 0$ or there exists $h \in G$ such that*

$$hterm(h) | lcm(M(p), M(q)), \quad Spoly(p, h) \overset{*}{\longmapsto}_G 0, \quad Spoly(h, q) \overset{*}{\longmapsto}_G 0.$$

Proof. If we replace (ii) in Theorem 4 with the above condition, we need only to ajust the proof that (ii) $\Longrightarrow$ (iii). Note that everything holds except for the proof $\tilde{p}_1 = \tilde{p}_2$ in the last paragraph.

This is easy to do under the new condition. Since $\text{hterm}(h) | \text{lcm}(M(p), M(q))$, $\text{hterm}(h) | t$. Let

$$h_1 = M(p) - \frac{M(p)}{M(h)} h.$$

Write

$$\begin{aligned} p_1 &= p_2 + (-p_2 + h_1) + (-h_1 + p_1) \\ &= p_2 + \alpha_2 u_2 \cdot \text{Spoly}(g_2, h) + \alpha_1 u_1 \cdot \text{Spoly}(h, g_1), \end{aligned}$$

where $\alpha_1, \alpha_2 \in \mathbb{F} \setminus \{0\}$, and $u_1, u_2 \in T_x$. All the polynomials have head terms $< t$. The remaining arguments apply. □

The next corollary is proved similarly.

Corollary 6 *G is a Gröbner basis if and only if for all $p, q \in G$ there exist $h_0, h_1, \cdots, h_k \in G$, where $h_0 = p$ and $h_k = q$, such that*

$$hterm(h_i) | lcm(M(p), M(q)), \quad Spoly(h_i, h_{i+1}) \overset{*}{\longmapsto}_G 0, \quad 0 \leq i \leq k-1.$$