Lecture 7. Buchberger's criterion and algorithm

Problem: How to construct a Gröbner basis?

Definition 4.10. *Let $f, g \in k[X_1, \ldots, X_n]$ be non-zero polynomials. Let $X^\alpha = LM(f)$ with $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $X^\beta = LM(g)$ with $\beta = (\beta_1, \ldots, \beta_n)$. Define $\gamma = (\gamma_1, \ldots, \gamma_n)$ with $\gamma_i = \max\{\alpha_i, \beta_i\}$ for all i, i.e. $X^\gamma = \operatorname{lcm}(LM(f), LM(g))$. The S-polynomial of f and g is defined to be*

$$S(f,g) = \frac{X^\gamma}{LT(f)} f - \frac{X^\gamma}{LT(g)} g.$$

▶ Let $f = 4X^2Z - 7Y^2$ and $g = XYZ^2 + 3XZ^4$. Using the $\leq_{lex}$ order with $X > Y > Z$, we see that $\operatorname{lcm}(LM(f), LM(g)) = X^2YZ^2$, and

$$S(f,g) = \frac{X^2YZ^2}{4X^2Z} f - \frac{X^2YZ^2}{XYZ^2} g = \tfrac{1}{4}YZf - Xg = -3X^2Z^4 - \tfrac{7}{4}Y^3Z.$$

Here is how you may do this in Macaulay2:

```
R = QQ[x,y,z, MonomialOrder => Lex];
f = 4*x^2*z-7*y^2;
g = x*y*z^2 + 3*x*z^4;
Spoly = (f,g) -> (
    m := lcm(leadMonomial f, leadMonomial g);
    (m // leadTerm f) * f - (m // leadTerm g) * g);
Spoly(f,g)
```

▶ If $S(f,g) \neq 0$, then $LM(S(f,g)) < \operatorname{lcm}(LM(f), LM(g))$. In this sense, we can say that the S-polynomial is designed to produce cancellation of leading terms.

The name S-polynomial is short for "syzygy polynomial". You can blame Sylvester, who introduced the word into mathematics in 1850, if you don't know how to pronounce it; Hilbert's 1890 paper on syzygies made it standard. In general, syzygy is a kind of linear relation. Here it is a linear relation between polynomials in several variables with polynomial coefficients. In astronomy, a syzygy is an alignment of three celestial bodies, such as the Sun, the Earth and the Moon during an eclipse.

Theorem 4.11. *(Buchberger's criterion) Let $I \subset k[X_1, \ldots, X_n]$ be a non-zero ideal and let $G = \{g_1, \ldots, g_s\} \subset I \setminus \{0\}$ be such that $I = (g_1, \ldots, g_s)$. Then, G is a Gröbner basis for I if and only if for all $1 \leq i < j \leq s$, the remainder on division of $S(g_i, g_j)$ by $g_1, \ldots, g_s$ is 0.*

For the proof of this theorem, we will need the following technical result. The idea is that every cancellation of leading terms among polynomials of the same leading monomial comes from the cancellation that occurs for S-polynomials.

Lemma 4.12. *Suppose $g_1, \ldots, g_s \in k[X_1, \ldots, X_n]$ are non-zero, $\alpha_1, \ldots, \alpha_s \in \mathbb{N}^n$ and $c_1, \ldots, c_s \in k \setminus \{0\}$. Let*

$$f \coloneqq \sum_{i=1}^{s} c_i X^{\alpha_i} g_i \in k[X_1, \ldots, X_n].$$

Suppose that $LM(X^{\alpha_i} g_i) = X^\delta$ for all i, and that $f = 0$ or $LM(f) < X^\delta$. Then, we have

$$f = \sum_{1 \le i < j \le s} c_{ij} X^{\delta - \gamma_{ij}} S(g_i, g_j)$$

for some $c_{ij} \in k$, where $X^{\gamma_{ij}} = \operatorname{lcm}(LM(g_i), LM(g_j))$. Moreover, $LM(X^{\delta-\gamma_{ij}} S(g_i, g_j)) < X^\delta$ whenever $S(g_i, g_j) \neq 0$.

Proof. We have $X^\delta = X^{\alpha_i} LM(g_i) = X^{\alpha_j} LM(g_j)$ for all $1 \le i < j \le s$. Hence, $LM(g_i)$ and $LM(g_j)$ both divide X^δ. Thus, $X^{\gamma_{ij}} | X^\delta$. In particular, $X^{\delta - \gamma_{ij}} \in k[X_1, \ldots, X_n]$.

We first prove the last assertion. We have

$$S(g_i, g_j) = \frac{X^{\gamma_{ij}}}{LT(g_i)} g_i - \frac{X^{\gamma_{ij}}}{LT(g_j)} g_j$$

where both summands have leading term $X^{\gamma_{ij}}$. Since the leading terms cancel, $LM(S(g_i, g_j)) < X^{\gamma_{ij}}$ if $S(g_i, g_j) \neq 0$, and multiplying by $X^{\delta - \gamma_{ij}}$ gives $LM(X^{\delta-\gamma_{ij}} S(g_i, g_j)) < X^\delta$.

Replacing g_i by $g_i / LC(g_i)$ and c_i by $c_i LC(g_i)$ changes neither f nor the $S(g_i, g_j)$, so we may assume that $LC(g_i) = 1$, that is $LT(g_i) = LM(g_i)$, for all i.

We prove the first assertion by induction on s. If $s = 0$ then $f = 0$ and there is nothing to prove, and $s = 1$ cannot occur, since then $LM(f) = X^\delta$. For $s \ge 2$ we compute

$$\begin{aligned}
g &\coloneqq f - c_1 X^{\delta - \gamma_{12}} S(g_1, g_2) \\
&= c_1 X^{\alpha_1} g_1 + c_2 X^{\alpha_2} g_2 + \sum_{i=3}^{s} c_i X^{\alpha_i} g_i - c_1 X^{\delta - \gamma_{12}} \left(\frac{X^{\gamma_{12}}}{LM(g_1)} g_1 - \frac{X^{\gamma_{12}}}{LM(g_2)} g_2 \right) \\
&= c_1 \left(X^{\alpha_1} - \frac{X^\delta}{LM(g_1)} \right) g_1 + \left(c_2 X^{\alpha_2} + c_1 \frac{X^\delta}{LM(g_2)} \right) g_2 + \sum_{i=3}^{s} c_i X^{\alpha_i} g_i \\
&= (c_1 + c_2) X^{\alpha_2} g_2 + \sum_{i=3}^{s} c_i X^{\alpha_i} g_i
\end{aligned}$$

where the coefficient of g_1 vanishes because $X^{\alpha_1} LM(g_1) = X^\delta$ by assumption, and we used $X^\delta / LM(g_2) = X^{\alpha_2}$. Now g is a sum of at most $s - 1$ terms of the same form as f (we drop the first term if $c_1 + c_2 = 0$), and $g = 0$ or $LM(g) < X^\delta$, because this holds for f and for $X^{\delta - \gamma_{12}} S(g_1, g_2)$. By induction, g is a linear combination of the $X^{\delta - \gamma_{ij}} S(g_i, g_j)$, and hence so is $f = g + c_1 X^{\delta - \gamma_{12}} S(g_1, g_2)$. □

We are now ready to give the proof of the theorem.

Proof. (of Theorem 4.11) One direction is easy: suppose G is a Gröbner basis. Since $S(g_i, g_j) \in I$, division with remainder by G gives zero remainder by Theorem 4.8.

Conversely, suppose that all the remainders are zero, and let $f \in I$ be non-zero; we have to show that $LT(f) \in (LT(g_1), \ldots, LT(g_s))$. Write

$$f = \sum_{i=1}^{s} u_i g_i \text{ with } u_i \in k[X_1, \ldots, X_n].$$

Let $X^\delta = \max\{LM(u_i g_i) : 1 \leq i \leq s, u_i \neq 0\}$, and let $T = \{i : LM(u_i g_i) = X^\delta\}$ be the set of indices where the maximum is attained. We have $LM(f) \leq X^\delta$. The strategy is as follows: if $LM(f) = X^\delta$ we are done, and otherwise the leading terms of the $u_i g_i$ cancel, so by Lemma 4.12 they come from S-polynomials, and we use the vanishing remainders to rewrite f with a smaller X^δ. Below, we read an inequality $LM(h) < X^\delta$ as satisfied also when $h = 0$.

Case 1: Suppose $LM(f) = X^\delta$. Then

$$LT(f) = \sum_{i \in T} LT(u_i g_i) = \sum_{i \in T} LT(u_i) LT(g_i) \in (LT(g_1), \ldots, LT(g_s)).$$

Case 2: Suppose $LM(f) < X^\delta$. We define

$$f^* = \sum_{i \in T} LT(u_i) g_i.$$

Then $f - f^* = \sum_{i=1}^{s} v_i^* g_i$, where $v_i^* = u_i - LT(u_i)$ if $i \in T$ and $v_i^* = u_i$ otherwise, so that $LM(v_i^* g_i) < X^\delta$ for all i. In particular $LM(f^*) < X^\delta$, and applying Lemma 4.12 to f^*, with $c_i = LC(u_i)$ and $X^{\alpha_i} = LM(u_i)$ for $i \in T$, we may write

$$f^* = \sum_{i<j \text{ in } T} c_{ij} X^{\delta - \gamma_{ij}} S(g_i, g_j) \quad \text{with} \quad LM(X^{\delta - \gamma_{ij}} S(g_i, g_j)) < X^\delta.$$

Now, by assumption, for each pair $i < j$ in T division with remainder gives

$$S(g_i, g_j) = \sum_{l=1}^{s} v_l g_l \text{ with } v_l \in k[X_1, \ldots, X_n] \quad \text{and} \quad LM(v_l g_l) \leq LM(S(g_i, g_j))$$

(the v_l depend on i and j). Therefore,

$$X^{\delta - \gamma_{ij}} S(g_i, g_j) = \sum_{l=1}^{s} (X^{\delta - \gamma_{ij}} v_l) g_l \quad \text{with} \quad LM(X^{\delta - \gamma_{ij}} v_l g_l) < X^\delta.$$

It follows that there exist polynomials $u_i^* \in k[X_1, \ldots, X_n]$ such that

$$f^* = \sum_{i=1}^{s} u_i^* g_i \quad \text{with} \quad LM(u_i^* g_i) < X^\delta$$

Hence, we can write

$$f = (f - f^*) + f^* = \sum_{i=1}^{s} (u_i^* + v_i^*) g_i \quad \text{with} \quad LM((u_i^* + v_i^*) g_i) < X^\delta$$

With this new expression of f, we go back to the beginning of the proof; the new value of X^δ is strictly smaller. If we are in the first case, we are done. Otherwise, we repeat the procedure of the second case. By Lemma 4.5, this has to stop after finitely many steps. □

▶ Let $X^{\alpha_i} = LM(g_i)$, $X^{\alpha_j} = LM(g_j)$ and $X^{\gamma_{ij}} = \text{lcm}(LM(g_i), LM(g_j))$. By definition, $S(g_i, g_j) = h_i g_i + h_j g_j$ where $h_i = X^{\gamma_{ij} - \alpha_i} / LC(g_i)$ and $h_j = -X^{\gamma_{ij} - \alpha_j} / LC(g_j)$. It is a common mistake to conclude that the division with remainder of $S(g_i, g_j)$ by g_i, g_j has zero remainder. This is false in general: the expression $h_i g_i + h_j g_j$ is not usually the one produced by the division algorithm. For example, let $g_1 = X^2$ and $g_2 = XY + Y^2$ with the lex order $X > Y$. Then $S(g_1, g_2) = Y g_1 - X g_2 = -XY^2$, and division by g_1, g_2 gives $-XY^2 = 0 \cdot g_1 - Y g_2 + Y^3$ with remainder $Y^3 \neq 0$.

Finally, let us explain how the results we have covered so far provide an algorithm for constructing a Gröbner basis (hence, an algorithm for determining whether a polynomial belongs to an ideal I).

Suppose we are given $f_1, \ldots, f_t \in k[X_1, \ldots, X_n]$ such that $I := (f_1, \ldots, f_t)$ and we fix a monomial ordering $\leq$.

Initialize with $G = \{f_1, \ldots, f_t\}$. Now, suppose $G = \{g_1, \ldots, g_s\}$. For $1 \leq i < j \leq s$ divide $S(g_i, g_j)$ with remainder by $g_1, \ldots, g_s$, and let r_{ij} denote the remainder. If $r_{ij} = 0$ for all i, j, then stop: G is a Gröbner basis of I by Theorem 4.11. Otherwise, pick i, j with $r_{ij} \neq 0$, replace G by $G \cup \{r_{ij}\}$ and repeat. Note that $r_{ij} \in I$, so the new G still generates I.

Proposition 4.13. *The procedure described above stops in finitely many steps.*

Proof. By applying the procedure, we obtain a chain of sets

$$G_0 = \{f_1, \ldots, f_t\} \subset G_1 \subset G_2 \subset \ldots$$

and, writing $(LT(G))$ for the ideal generated by the leading terms of the elements of G, an ascending chain of ideals

$$(LT(G_0)) \subset (LT(G_1)) \subset (LT(G_2)) \subset \ldots$$

in $k[X_1, \ldots, X_n]$. We claim that each step of the procedure makes this chain strictly larger. Indeed, if $G_{m+1} = G_m \cup \{r\}$ with $r \neq 0$ the remainder of some $S(g, h)$ on division by the elements of G_m, then no term of r is divisible by any $LT(g_i)$, $g_i \in G_m$, so $LT(r) \notin (LT(G_m))$ by Lemma 4.4, while $LT(r) \in (LT(G_{m+1}))$. Since $k[X_1, \ldots, X_n]$ is Noetherian by the Hilbert basis theorem (Theorem 4.1), a strictly ascending chain of ideals is finite, so the procedure stops. □

▶ Gröbner bases give another approach to Hilbert's basis theorem for $k[X_1, \ldots, X_n]$. By an elementary combinatorial argument one first proves Dickson's lemma: every monomial ideal is generated by finitely many monomials. This replaces the use of the Hilbert basis theorem in Lemma 4.5, and applied to $\mathfrak{in}(I)$ it shows that every non-zero ideal I has a finite Gröbner basis G. Then $I = (G)$ by Theorem 4.8, so I is finitely generated.

◁ Compute a Gröbner basis (by hand) for the ideal $I = (X^2, XY + Y^2)$ with respect to the lex order with $X > Y$; the remark above did the first step. What happens for the lex order with $Y > X$?

▶ Here is an example of how you can compute a Gröbner basis in Macaulay2:

```
R = QQ[x,y,z, MonomialOrder => Lex];
I = ideal (x^2-z-1,z^2-y-1,x*y^2-x-1);
gens gb I
```

◁ (Elimination theory) Let $I \subset k[X_1, \ldots, X_n]$ be a non-zero ideal, G a Gröbner basis for I with respect to the $\leq_{lex}$ order with $X_1 > X_2 > \ldots > X_n$. Then $G \cap k[X_{l+1}, \ldots, X_n]$ generates the ideal $I \cap k[X_{l+1}, \ldots, X_n]$ for all $0 \leq l \leq n$.

A Gröbner basis $G = \{g_1, \ldots, g_s\}$ of an ideal I is said to be **reduced** if $LC(g_i) = 1$ for all i and for all i none of the terms of g_i is divisible by any of the $LT(g_j)$ for $j \neq i$. Starting from any Gröbner basis G, it is easy to produce a reduced Gröbner basis. First, if for some $g \in G$ we have $LT(g) \in (LT(g_i) : g_i \in G, g_i \neq g)$, then replace G with $G \setminus \{g\}$, which is still a Gröbner basis of I; repeat until no $LT(g_i)$ is divisible by another $LT(g_j)$. Next, for each i in turn, replace g_i by its remainder on division by the other elements of G. Since no $LT(g_j)$ with $j \neq i$ divides $LT(g_i)$, this does not change $LT(g_i)$, so G remains a Gröbner basis, and now no term of g_i is divisible by any $LT(g_j)$, $j \neq i$. Finally, make sure that $LC(g_i) = 1$ for all i by rescaling with an element of k. This gives a reduced Gröbner basis. An important property is that an ideal I has a unique reduced Gröbner basis with respect to a fixed monomial order. Here is the proof. Let G and G' be reduced Gröbner bases of I and let $g \in G$. By Lemma 4.4 there is $g' \in G'$ with $LT(g') \mid LT(g)$, and then $g'' \in G$ with $LT(g'') \mid LT(g')$. Since G is reduced, $g'' = g$, and so $LT(g') = LT(g)$. Applying this to every element of G and of G' shows that G and G' have the same leading terms. Now $g - g' \in I$, and each term of $g - g'$ is a non-leading term of g or of g'. Such a term is divisible by none of these leading terms: not by $LT(h)$ for $h \neq g$ because G and G' are reduced, and not by $LT(g)$ because it is smaller than $LT(g)$. If $g - g' \neq 0$, then $LT(g - g') \in \mathfrak{in}(I)$ would be divisible by some $LT(h)$, $h \in G$, by Lemma 4.4. Hence $g = g' \in G'$. This shows $G \subset G'$, and by symmetry $G = G'$.

▶ Consider the ideal $I = (X - Z^2, Y - Z^3)$ in $k[X, Y, Z]$. For the lexicographic order with $X > Y > Z$, it is easy to verify that $\{X - Z^2, Y - Z^3\}$ is the reduced Gröbner basis. On the other hand, for the graded lexicographic order with $X > Y > Z$, it is easy to verify that $\{X^2 - YZ, XZ - Y, Z^2 - X\}$ is the reduced Gröbner basis. This example shows that the number of elements in the reduced Gröbner basis for a given ideal depends on the monomial ordering.