MAG Lecture 3 : Monomial ordering

We have now seen the beginning of the dictionary that relates <u>geometry</u> (in the form of affine varieties) to <u>algebra</u> (in the form of ideals). Given a field k, polynomial ring $R = k[x_1, \dots, x_n]$ and affine space $\mathbb{A}^n = k^n$ we have

- For $I = \langle f_1, \dots, f_s \rangle$, $\mathbb{V}(I) = \mathbb{V}(f_1, \dots, f_s) \subseteq \mathbb{A}^n$, an <u>affine variety</u>

- For $V \subseteq \mathbb{A}^n$ an affine variety, $\mathbb{I}(V) \subseteq R$ an ideal (finitely generated?)

- $\langle f_1, \dots, f_s \rangle \subseteq \mathbb{I}(\mathbb{V}(f_1, \dots, f_s))$ (Lemma CLO 1.4.7)

- if V, W are affine varieties $V \subseteq W$ iff. $\mathbb{I}(V) \supseteq \mathbb{I}(W)$

We have looked at the division algorithm, and you have been "getting your hands dirty" in exercises working with polynomials. But is all that algebra really <u>geometry</u>? Isn't it just shuffling coefficients around? Yes and Yes: <u>the soul of geometry is in the algebraic manipulations</u> and not the pictures, which in any case will become close to useless as soon as we move beyond three variables. It will take some time before you are convinced of this (maybe the Division Algorithm is the most geometric thing in Euclid).

<u>Lemma 1</u> Let $f \in k[x]$. Then for $a \in k$, $f(a) = 0$ if and only if $x - a \mid f(x)$.

<u>Proof</u> By Euclidean Division we can write $f = q(x-a) + r$ where $\deg(r) < \deg(x-a) = 1$ (r is zero or) so $r \in k$. Then $f(a) = r$ so it is clear that if $f(a) = 0$ then $x - a \mid f$. The converse is easier to see. $\square$

<u>Lemma 2</u> Let $P=(P_1,\dots,P_n)\in\mathbb{A}^n$. Then $\mathbb{I}(\{P\}) = \langle x_1-P_1,\dots,x_n-P_n\rangle$.

<u>Proof</u> The inclusion $\supseteq$ is clear. For the reverse inclusion suppose $f\in\mathbb{I}(P)$ or what is the same $f(P)=0$. Suppose we collect terms to write

$$f = \sum_{i\geq 0} g_i(x_2,\dots,x_n)x_1^i$$

and run the division algorithm on f "treating the x_i for $i\geq 2$ as scalars", with x_1-P_1 as the divisor, i.e. if

$$f = g_N x_1^N + g_{N-1}x_1^{N-1} + \cdots$$

we subtract $g_N x_1^{N-1}(x_1-P_1)$ to obtain

$$\begin{aligned} f - g_N x_1^{N-1}(x_1-P_1) &= f - g_N x_1^N + g_N P_1 x_1^{N-1} \\ &= (g_{N-1}+g_N P_1)x_1^{N-1} + \cdots \end{aligned} \tag{2.1}$$

now subtracting $(g_{N-1}+g_N P_1)x_1^{N-2}(x_1-P_1)$ and continuing in this fashion we eventually obtain $f - q_1(x_1-P_1) = r$ where r is a polynomial in the variables $x_2,\dots,x_n$. Hence

$$f = q_1(x_1-P_1) + r(x_2,\dots,x_n)$$

Now apply the same algorithm to divide x_2-P_2 into r, and so on, obtaining

$$f = \sum_{i=1}^n q_i(x_i-P_i) + \lambda$$

with $\lambda\in k$. By substitution $f(P)=\lambda$. Hence if $f(P)=0$ then f is in $\langle x_1-P_1,\dots,x_n-P_n\rangle$. $\square$

In the proof there was no reason we couldn't have divided by the $x_i - P_i$ in some <u>other order</u>. The order we chose was $x_1 > x_2 > \cdots$ meaning that we prioritised terms with large x_1-degree, then terms with large x_2-degree, and so on. What made this work was that once we were "done" dividing by $x_1 - P_1, \ldots, x_i - P_i$ no $x_1, \ldots, x_i$'s were re-introduced into our dividend r by subsequent divisions by $x_{i+1} - P_{i+1}, \ldots, x_n - P_n$. Why was that? In the first step the original

$$f = g_N x_1^N + g_{N-1} x_1^{N-1} + \cdots$$

becomes the "first remainder" or <u>dividend</u>

$$r = f - g_N x_1^{N-1}(x_1 - P_1) = (g_{N-1} + g_N P_1) x_1^{N-1} + g_{N-2} x_1^{N-2} + \cdots$$

By def^N, $g_N, g_{N-1} \in k[x_2, \ldots, x_n]$, and so $g_{N-1} + g_N P_1 \in k[x_2, \ldots, x_n]$. and the coefficients of the other powers of x_1 are unchanged. Of course $P_1 \in k$, but note that even if P_1 were a <u>polynomial</u> in the $x_2, \ldots, x_n$ the logic would survive, and at each step the x_1-degree decreases until eventually our dividend R is in $k[x_2, \ldots, x_n]$. Suppose we now divide by $x_2 - P_2$, with $R = h_M(x_3, \ldots, x_n) x_2^M + h_{M-1}(x_3, \ldots, x_n) x_2^{M-1} + \cdots$

$$R' = R - h_M x_2^{M-1}(x_2 - P_2) = (h_{M-1} + h_M P_2) x_2^{M-1} + \cdots$$

Again this will work out just fine if $P_2 \in k[x_3, \ldots, x_n]$, in the sense that we can continue dividing by $x_2 - P_2$ until a remainder in $k[x_3, \ldots, x_n]$. But if P_2 contains x_2's it will potentially <u>stop the x_2-degree from decreasing</u>, and if P_2 contains x_1's then these may be introduced into R' and we're <u>back to the beginning again</u>!

If you're astute you'll notice the first problem is not a big deal. Suppose we replace $x_2 - P_2$ by $x_2^k - x_2^l P_2'$ with $P_2' \in k[x_3, \ldots, x_n]$. Then our division looks like $(l < k)$

$$R' = R - h_M x_2^{M-k}(x_2^k - x_2^l P_2') = (h_{M-1} + h_M x_2^l P_2') x_2^{M-k} + \cdots$$

$$= P_2' h_M x_2^{\overbrace{M-k+l}^{<M}} + h_{M-1} x_2^{M-k} + \cdots$$

but since $M-k+l < M$ <u>we're still making progress</u>. This is of course just the familiar fact that we match up <u>leading terms</u> in the polynomial division algorithm in one variable. So division will make progress as long as $P_2 = x_2^l P_2'$ is "smaller" than x_2^k in two senses: it shouldn't involve x_1 (which counts as "bigger" than any power of x_2) or powers of x_2 above k.
or equal to

<u>Example 1</u> Let $f_1 = y^2 - xz$, $f_2 = z - x^2$ We claim if $Z = \{(t^2, t^3, t^4) \mid t \in \mathbb{R}\}$ that $\mathbb{I}(Z) = \langle f_1, f_2 \rangle$. It is easy to check $\mathbb{V}(f_1, f_2) = Z$ so $\langle f_1, f_2 \rangle \subseteq \mathbb{I}(\mathbb{V}(f_1, f_2)) = \mathbb{I}(Z)$. Now suppose $f \in \mathbb{I}(Z)$. Dividing f_1 into f gives $f = q_1 f_1 + r_1(x,z) + r_2(x,z) y$

Hence for $t \in \mathbb{R}$

$$0 = f(t^2, t^3, t^4) = r_1(t^2, t^4) + r_2(t^2, t^4) t^3$$

substituting $-t$ gives $0 = r_1(t^2, t^4) - r_2(t^2, t^4) t^3$ so $r_1(t^2, t^4) = 0$ and $r_2(t^2, t^4) t^3 = 0$ for all t. Hence for $t \neq 0$, $r_2(t^2, t^4) = 0$. The polynomial $r_2(t^2, t^4)$ in $\mathbb{R}[t]$ has infinitely many roots and is therefore zero. We have reduced to proving $\mathbb{I}(\{(t^2, t^4) \mid t \in \mathbb{R}\}) = \langle f_2 \rangle$ in $\mathbb{R}[x,z]$. Suppose $g(t^2, t^4) = 0$ for all t, and divide g by $z - x^2$ treating z as the "primary variable" so we obtain $g = q(z - x^2) + R(x)$. Then $0 = g(t^2, t^4) = R(t^2)$ for all t, so $R = 0$ and $g \in \langle f_2 \rangle$, completing the proof that $f \in \langle f_1, f_2 \rangle$.

Note the order $y > z > x$ implicitly used here

In this Example we could solve the problem easily because we chose the right ordering $y > z > x$ and tailored our division process to this ordering in order that the remainders became always "smaller". We now make these ideas precise.

Monomial orderings

Let k be a field. We explained how monomials $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ in $k[x_1, \ldots, x_n]$ are in bijection with tuples $\alpha \in \mathbb{N}^n$ (for us $\mathbb{N} = \mathbb{Z}_{\geq 0}$) and we freely interchange them. We write $e_i = (0, \ldots, \overset{i}{1}, \ldots, 0)$ so that $x_i = x^{e_i}$.

A total order $<$ is a relation on a set S which is irreflexive ($\forall s \in S$ not $s < s$), transitive ($\forall s, t, u \in S$ if $s < t$ and $t < u$ then $s < u$) and total ($\forall s, t \in S$ $s < t$ or $s = t$ or $t < s$). We write $s > t$ for $t < s$, and $s \leq t$ for $s = t$ or $s < t$, similarly $s \geq t$.

Def^n A monomial ordering $<$ on $k[x_1, \ldots, x_n]$ is a relation on $\mathbb{N}^n$ (or $\{x^\alpha\}_{\alpha \in \mathbb{N}^n}$) satisfying

(i) $<$ is a total order

(ii) if $\alpha > \beta$ and $\gamma \in \mathbb{N}^n$ then $\alpha + \gamma > \beta + \gamma$ (i.e. $x^\alpha x^\gamma > x^\beta x^\gamma$).

(iii) $<$ is a well-ordering, that is, every nonempty subset $S \subseteq \mathbb{N}^n$ has a smallest element (i.e. $\exists s \in S \ \forall t \in S \ s \leq t$).

We do not here (and never will) define an order on general polynomials; we order only monomials.

Def^n Given $\alpha, \beta \in \mathbb{N}^n$ we define $\alpha >_{lex} \beta$ if the leftmost nonzero entry of $\alpha - \beta = (\alpha_1 - \beta_1, \ldots, \alpha_n - \beta_n)$ is positive. This is called lexicographic order, or lex.

Example 2 $(1, 0, \ldots, 0) >_{lex} (0, 1, 0, \ldots) >_{lex} \cdots >_{lex} (0, \ldots, 0, 1)$ so $x_1 >_{lex} x_2 >_{lex} \cdots >_{lex} x_n$. Note $x_1 >_{lex} x_2^{100}$.

<u>Warning</u> As we flagged in Lecture 1, "x_1" is really a name for x^{e_1} and sometimes we use another name. The monomial order, viewed as an order on $\mathbb{N}^n$, does not care; but if we say something like $x >_{lex} y >_{lex} z$ what we mean is (since <u>always</u> $x^{e_1} >_{lex} x^{e_2} >_{lex} x^{e_3}$) that we are using "$x$", "$y$", "$z$" to denote resp. $x^{e_1}, x^{e_2}, x^{e_3}$.

<u>Lemma CLO 2.2.2</u> A total order on $\mathbb{N}^n$ is a well-ordering iff. every strictly decreasing sequence $\alpha(1) > \alpha(2) > \cdots$ in $\mathbb{N}^n$ is finite.

<u>Proof</u> Suppose $(\mathbb{N}^n, <)$ is a well-ordering and $\{\alpha(i)\}_{i=1}^{\infty}$ is a sequence with $\alpha(i) \geq \alpha(i+1)$ for all i. Then the set $\{\alpha(i)\}_i$ has a least element $\alpha(N)$, and clearly $\alpha(i) = \alpha(i+1)$ for $i \geq N$. If every strictly decreasing sequence is finite and $S \subseteq \mathbb{N}^n$ is nonempty, let $\mathcal{C}$ be the set of maximal chains $\alpha(1) > \cdots > \alpha(n)$ in S of finite length. This is nonempty since we may choose any $s \in S$, and if it is not minimal choose $s > t$, and by hypothesis this terminates with a finite sequence. If $\alpha(1) > \cdots > \alpha(n)$, $\beta(1) > \cdots > \beta(m)$ are in $\mathcal{C}$ and $\alpha(n) < \beta(m)$ or $\alpha(n) > \beta(m)$ we have a contradiction, hence $\alpha(n) = \beta(m)$. This common final entry in every sequence of $\mathcal{C}$ is a least element of S. $\square$

<u>Proposition CLO 2.2.4</u> $<_{lex}$ is a monomial order.

<u>Proof</u> (i) $<_{lex}$ is clearly irreflexive and transitive.

(ii) if $\alpha >_{lex} \beta$ and $\gamma \in \mathbb{N}^n$ then $\alpha + \gamma >_{lex} \beta + \gamma$ since $(\alpha+\gamma) - (\beta+\gamma) = \alpha - \beta$.

(iii) We use the previous lemma. Suppose $\alpha(1) >_{lex} \alpha(2) >_{lex} \alpha(3) > \cdots$ then we claim there exists N_1 such that for $i \geq N_1$, $\alpha(i)_1 = \alpha(i+1)_1$. This is because $\alpha(i) >_{lex} \alpha(i+1)$ means either $\alpha(i)_1 = \alpha(i+1)_1$ or $\alpha(i)_1 > \alpha(i+1)_1$ and there are finitely many non-negative integers less than $\alpha(1)_1$. There must then be N_2 such that $\alpha(i)_2 = \alpha(i+1)_2$ for $i \geq N_2 \geq N_1$, and by induction for some N, $\alpha(i) = \alpha(i+1)$ for $i \geq N$, as required. $\square$

<u>Def^n</u> Let $f = \sum_\alpha a_\alpha x^\alpha$ be a nonzero polynomial, and $<$ a monomial order.

- The <u>multidegree</u> of f is

$$\text{multideg}(f) = \max\{\alpha \in \mathbb{N}^n \mid a_\alpha \neq 0\}$$

where the max is w.r.t. $<$.

- If $\alpha = \text{multideg}(f)$ then the <u>leading coefficient</u> of f is

$$LC(f) = a_\alpha,$$

the <u>leading monomial</u> of f is

$$LM(f) = x^\alpha$$

and the <u>leading term</u> of f is

$$LT(f) = a_\alpha x^\alpha.$$

<u>Example 3</u> If $x >_{lex} y >_{lex} z$ and $f = 3x^2 + y^7 z + y^6 z^8$ then $x^2 >_{lex} y^7 z >_{lex} y^6 z^8$, $LT(f) = 3x^2$, $LC(f) = 3$, $LM(f) = x^2$.

<u>Lemma CLO 2.2.8</u> Let $f, g \in k[x_1, \dots, x_n]$ be nonzero. Then

(i) $\text{multideg}(f) + \text{multideg}(g) = \text{multideg}(fg)$

(ii) If $f + g \neq 0$ then $\text{multideg}(f+g) \leq \max\{\text{multideg}(f), \text{multideg}(g)\}$.
If in addition $\text{multideg}(f) \neq \text{multideg}(g)$ this is an equality.

<u>Remark</u> Let $<$ be a monomial order on $k[x_1, \dots, x_n]$. The axioms say if $x^\alpha < x^\beta$ then $x^\alpha x^\gamma < x^\beta x^\gamma$ but the converse also holds. If $x^{\alpha+\gamma} < x^{\beta+\gamma}$ then by totality we have either $x^\alpha < x^\beta$, $x^\alpha = x^\beta$ or $x^\alpha > x^\beta$. If $x^\alpha = x^\beta$ (i.e. $\alpha = \beta$) then $\alpha + \gamma = \beta + \gamma$ a contradiction, and if $x^\alpha > x^\beta$ then $x^{\alpha+\gamma} > x^{\beta+\gamma}$ a contradiction. Hence $x^\alpha < x^\beta$.

<u>Question</u> Can we have $1 > x^\alpha$ for some $\alpha \neq \underline{0}$ in $k[x_1, \dots, x_n]$?

<u>Question</u> What are the possible monomial orders on $k[x]$?