Lecture LVI: Towards Gröbner bases I

§56.1 Roots of Polynomials and their multiplicity:

In Lecture 54, we discussed irreducibility of a polynomial in one variable with coefficients from a UFD. We also used the following fact in our examples from Lecture 55:

Lemma 1: Let K be a field (eg. $K = \mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}/p\mathbb{Z}$ for $p \geq 2$ prime) & let $f \in K[x]$ and $\alpha \in K$. Then:

$$(x-\alpha) \text{ divides } f(x) \text{ in } K[x] \iff f(\alpha) = 0$$

Proof: ($\Rightarrow$) If $f(x) = (x-\alpha)\, g(x)$ for some $g(x) \in K[x]$, then $f(\alpha) = 0 \cdot g(\alpha) = 0$

($\Leftarrow$) Use the Euclidean Algorithm to write $f(x) = (x-\alpha)\, g(x) + r(x)$

so that $r(x) = 0$ or $r(x) \neq 0$ and $\deg(r(x)) < \deg(x-\alpha) = 1$, ie $r \in K$

Now, $\left.\begin{aligned} f(\alpha) &= 0 \cdot g(\alpha) + r = r \\ f(\alpha) &= 0 \end{aligned}\right\} \Rightarrow r = 0$ so $f(x) = (x-\alpha)\, g(x)$. □

Similarly, we can prove a more general result.

Lemma 2: Let K be a field. Let $f \in K[x]$, $\alpha \in K$ and $N \geq 1$. Then:

$$(x-\alpha)^N \text{ divides } f(x) \text{ in } K[x] \iff f(\alpha) = \underbrace{f'(\alpha) = f''(\alpha) = \cdots = f^{(N-1)}(\alpha)}_{\text{derivatives of } f, f', \ldots} = 0$$

Definition: If $f = a_n x^n + \cdots + a_1 x + a_0$, then $f'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + 2a_2 x + a_1$

Proof: ($\Rightarrow$) Write $f(x) = (x-\alpha)^N g(x)$ for some $g \in K[x]$

Then $f'(x) = N(x-\alpha)^{N-1} g(x) + (x-\alpha)^N g'(x) = (x-\alpha)^{N-1} \underbrace{(N g(x) + (x-\alpha) g'(x))}_{=: h(x)}$

We proceed by induction on N:

Base Case: $N = 1$ is the content of Lemma 1.

Inductive Step: $(x-\alpha)^N$ divides f, so $(x-\alpha)$ divides f. By Lemma 1, $f(\alpha) = 0$

Since $f'(x) = (x-\alpha)^{N-1} h(x)$, the (IH) says $f'(\alpha) = (f')'(\alpha) = \cdots = (f')^{(N-2)}(\alpha) = 0$

Thus we get $f(\alpha) = f'(\alpha) = f''(\alpha) = \cdots = f^{(N)}(\alpha) = 0$.

($\Leftarrow$) Using the Euclidean Algorithm, we write

$$f(x) = (x-\alpha)^N g(x) + r(x) \qquad (*)$$

with $r(x) \equiv 0$ or ($r(x) \neq 0$ and $\deg r(x) < \deg (x-\alpha)^N = N$).

We take derivative of (*) on both sides:

$$f'(x) = N(x-\alpha)^{N-1} g(x) + (x-\alpha)^N g'(x) + r'(x)$$
$$= (x-\alpha)^{N-1} \underbrace{\left(N g(x) + (x-\alpha) g' \right)}_{=: g(x) \in K[x]} + r'(x)$$

We proceed by induction on N:

Base Case: $N=1$ is the statement of Lemma 1

Inductive Step: Evaluating both sides of (*) at $x=\alpha$ we get

$$0 = f(\alpha) = 0^N g(\alpha) + r(\alpha) = 0$$

So $r(\alpha)=0$.

Taking the derivative of (*) yields $f'(x) = (x-\alpha)^{N-1} g(x) + r'(x)$ (**)

If $r'(x) \neq 0$, then $\deg(r'(x)) = \deg(r(x)) - 1 < N-1$

By (IH) applied to f': $f'(\alpha) = f''(\alpha) = \cdots = f^{(N-1)}(\alpha) = 0 \Rightarrow (x-\alpha)^{N-1} \mid f'$.
($(f')'(\alpha)$ $\quad$ $(f')^{(N-2)}(\alpha)$)

Thus: $f'(x) = (x-\alpha)^{N-1} h(x)$ for some $h(x) \in K[x]$

Comparing this with (**) we get $(x-\alpha)^{N-1} \mid r'$

This cannot happen if $r' \neq 0$ since $\deg r' < N-1$ but $N-1 \leq r'$ from $(x-\alpha)^{N-1} \mid r'$.

Thus, $r' \equiv 0$, so $r \in K$.

Since $r(\alpha) = 0$, we conclude $r(x) \equiv 0$ so $(x-\alpha)^N \mid f(x)$ in $K[x]$ □

Definition: Let $f \in K[x]$ & $\alpha \in K$ with $f(\alpha)=0$. We say the multiplicity of α as a root of $f(x)$ is N if $f(\alpha) = f'(\alpha) = \cdots = f^{(N-1)}(\alpha) = 0$ but $f^{(N)}(\alpha) \neq 0$

Equivalently, $(x-\alpha)^N \mid f(x)$ in $K[x]$ but $(x-\alpha)^{N+1} \nmid f(x)$ in $K[x]$.

Corollary: Assume K is a field and $f(x) \in K[x]$. If $f(x)$ has m distinct roots in K, then $m \leq \deg(f)$.

Side note: We actually proved and use this result in the proof of Theorem 1 §25.2, to conclude that $\mathrm{Aut}_{Gp}(\mathbb{Z}/p\mathbb{Z}) = \mathbb{F}_p^\times$ is cyclic for $p \geq 2$ prime.

Example: Let $f(x) = x^2+x+1 \in \mathbb{F}_2[x]$. Then $\alpha \in \mathbb{F}_2$ can only be 0 or 1.

$f(0)=1$, $f(1)=3 \equiv 1 \bmod 2$, so $\forall \alpha \in \mathbb{F}_2$ $f(\alpha) \neq 0$. Thus, $f(x)$ cannot be written as a product of two linear factors in $\mathbb{F}_2[x]$, hence $f(x)$ is irreducible for degree reasons.

§56.2 The ring $K[x_1, \dots, x_n]$ for $n \geq 1$

Fix K a field and let $R = K[x_1, \dots, x_n]$ be the ring of polynomials in n variables with coefficients from K.

Let us quickly review the proof of Hilbert Basis Theorem, which implies that R is Noetherian (ie, every ideal in R is finitely generated).

$$R = K[x_1, \dots, x_n] = \underbrace{K[x_1, \dots, x_{n-1}]}_{\text{call it } A}[x_n]$$

(x_n: call this variable u for now)

$$\Rightarrow \quad R = A[u]$$

$\forall$

I ideal $\rightsquigarrow$ Take $L(I) \subseteq A$ ideal in A generated by the leading coefficients of those $p(u) \in I \subsetneq A[u]$

Step 1: We use the fact that A is Noetherian to get $(a_1, \dots, a_k) = L(I)$ in A

Pick $p_1(u), \dots, p_k(u) \in A[u]$ so that Leading coefficient of $p_j(u) = a_j$

Using the divisim algorithm, modulo $p_1(u), \dots, p_k(u)$, the degree of any $q(u) \in A[u]$ can be brought down to be (strictly) less than $\max_{1 \leq j \leq k} \{ \deg(p_j(u)) \} =: D$

Step 2: Pick "generators" of $I_{<D} = \{ p(u) \in I : \deg(p(u)) < D \}$ (finitely many)

Note: (1) By writing $u = x_n$ we made a choice: x_n is "better" than $x_1, \dots, x_{n-1}$ (NOT democratic)

(2) Step 2 is not as "algorithmic" as Step 1.

Gröbner basis will allow us to solve these issues.

§56.3 Monomials, orderings and related notions:

Definition: A **monomial** is a polynomial with only one term

Example: $x_1^2 x_2 \in K[x_1, x_2]$.

We can write a polynomial in $K[x_1, \dots, x_n]$ as a finite sum of monomials.

Example: A typical polynomial in $K[x_1, x_2]$ is of the form $\sum_{\substack{k,\ell \geq 0 \\ \text{finite}}} c_{k\ell}\, x_1^k x_2^\ell$

We will use a convenient short hand notation:

- $\underline{x}$ to denote $x_1, \dots, x_n$
- $\underline{\alpha} = \alpha_1, \dots, \alpha_n$ to denote "exponents" $(\alpha_1, \dots, \alpha_n \in \mathbb{Z}_{\geq 0})$
- $\underline{x}^{\underline{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ to denote a typical monomial

$\sum_{\substack{\underline{\alpha} \in (\mathbb{Z}_{\geq 0})^n \\ \text{finite sum}}} c_{(\underline{\alpha})}\, \underline{x}^{\underline{\alpha}}$ to denote a typical polynomial.

} in $K[x_1, \dots, x_n]$

Definition: A **monomial order** is a total order on the set of monomials $\{\underline{x}^{\underline{\alpha}} : \underline{\alpha} \in (\mathbb{Z}_{\geq 0})^n\}$ s.t.

- $\underline{x}^{\underline{\alpha}} < \underline{x}^{\underline{\beta}} \Rightarrow \underline{x}^{\underline{\alpha}+\underline{\gamma}} < \underline{x}^{\underline{\beta}+\underline{\gamma}} \quad \forall\, \underline{\gamma} \in (\mathbb{Z}_{\geq 0})^n$.
- $<$ is a well-ordering (every non-empty subset of monomials has a smallest element).

§56.4 Leading terms:

Let us assume we fixed a monomial order for monomials in $K[x_1, \dots, x_n]$

Definition: Given $f(\underline{x}) = \sum_{\substack{\underline{\alpha} \\ \text{finite}}} c_{(\underline{\alpha})}\, \underline{x}^{\underline{\alpha}} \in K[x_1, \dots, x_n] \setminus \{0\}$, we define the **leading term** of f (wrt $<$) as $LT(f) = c_{(\underline{\alpha}_0)}\, \underline{x}^{\underline{\alpha}_0}$ where $\underline{\alpha}_0$ is such that

(1) $c_{(\underline{\alpha}_0)} \neq 0$ and (2) $c_{(\underline{\alpha})} \neq 0 \Rightarrow \underline{x}^{\underline{\alpha}} \leq \underline{x}^{\underline{\alpha}_0}$

Convention: $LT(0) = 0$

Note: Since $<$ is a total order $LT(f)$ is well-defined and consists of a single monomial.