Introduction

Quadratics (Babylonians):

$$X^2 + bX = c = (X + \frac{1}{2}b)^2 + c - \frac{b^2}{4}$$

$$= (X - x_1)(X - x_2) \implies x_1 x_2 = c, x_1 + x_2 = -b$$

$$x_1 = \frac{1}{2}\left[(x_1 + x_2) + (x_1 - x_2)\right] = \frac{1}{2}\left[-b + \sqrt{b^2 - 4c}\right]$$

Cubics (Italy, 16th Century):

$$X^3 + aX^2 + bX + c = (X - x_1)(X - x_2)(X - x_3)$$

$$\implies x_1 + x_2 + x_3 = -a, x_1x_2 + x_1x_3 + x_2x_3 = b, x_1x_2x_3 = -c$$

WLOG $X \to X - a/3$ and $a = 0$

$$x_1 = \frac{1}{3}\left[(x_1 + x_2 + x_3) + \underbrace{(x_1 + \omega x_2 + \omega^2 x_3)}_{=u} + \underbrace{(x_1 + \omega^2 x_2 + \omega x_3)}_{=v}\right]$$

where $\omega = e^{2\pi i/3}$ so $\omega^2 + \omega + 1 = 0$. Cyclic permutation of x_1, x_2, x_3 gives $u \to \omega u \to \omega^2 u$ and $v \to \omega v \to \omega^2 v$ which implies u^3 and v^3 are invariant under cyclic permutations of the roots.

Also $u \leftrightarrow v$ under $x_2 \leftrightarrow x_3$. So $u^3 + v^3$, $u^3 v^3$ are invariant under permutations of roots.

In fact,

$$u^3 + v^3 = 27x_1x_2x_3 = -27c$$

$$u^3 v^3 = -27b^2$$

So u^3, v^3 are roots of $Y^2 + 27cY - 27b^2$. This gives a formula for x_1 (Cardano's formula).

Can follow a similar method for quartics - auxilliary cubic equation. Unfortunately it doesn't work for quintics - the reason being group theory.

1 Polynomials

In this course, all rings are commutative and non-zero. Let R be a ring, then $R[X]$ denotes the ring of polynomials $\sum_{i=0}^{n} a_i X^i$, $a_i \in R$. A polynomial $f \in R[X]$ determines a function $R \to R$, $r \mapsto f(r)$.

The polynomial is not in general determined by this function, e.g let $R = \mathbb{Z}/p\mathbb{Z}$ (p prime). Then for all $a \in R$, $a^p = a$ so the polynomials X^p and X represent the same function.

In the case when $R = K$ (a field), $K[X]$ is a Euclidean domain. The "division algorithm" says that if $f, g \in K[X]$, $g \neq 0$ then there exists unique $q, r \in K[X]$ such that $f = gq + r$ and $\deg r < \deg g$ (define $\deg(0) = -\infty$).

In particular, if $g = X - a$ is linear then $f = (X - a)q + f(a)$ ("remainder theorem"). So $K[X]$ is also a PID and a UFD - every polynomial is a product of irreducible polynomials, and there are GCD's, computable via Euclids algorithm in the usual way.

Proposition 1.1. *If K is a field, $0 \neq f \in K[X]$, then f has at most $\deg f$ roots in K.*

Proof. If f has no roots then we are done. Otherwise, suppose $f(a) = 0$ for $a \in K$. Then
$$f = (X - a)g$$
for some $g \in K[X]$ and $\deg g = \deg f - 1$. If $b \in K$ is a root of f then either $b = a$ or $g(b) = 0$ so the number of roots of f is at most one more than the number of roots of g. Now done by induction. $\square$

2 Symmetric polynomials

Let R be a ring, consider $R[X_1, \ldots, X_n]$ for $n \geq 1$.

Definition. A polynomial $f \in R[X_1, \ldots, X_n]$ is *symmetric* if for every $\sigma \in S_n$, $f(X_{\sigma(1)}, \ldots, X_{\sigma(n)}) = f$.

The set of symmetric polynomials is a subring of $R[X_1, \ldots, X_n]$.

Example. $X_1 + \ldots + X_n$, or more generally, $p_k = X_1^k + \ldots + X_n^k = \sum_{i=1}^n X_i^k$.

Alternative definition: if $f \in R[X_1, \ldots, X_n]$, define $f\sigma = f(X_{\sigma(1)}, \ldots, X_{\sigma(n)})$. This is an action (on the right) of S_n on $R[X_1, \ldots, X_n]$. A polynomial f is symmetric if and only if it is fixed by this action.

Definition. *The elementary symmetric polynomials* are

$$s_r(X_1, \ldots, X_n) = \sum_{1 \leq i_1 < \ldots < i_r \leq n} X_{i_1} X_{i_2} \ldots X_{i_r}$$

Example. When $n = 3$ we have

$$\begin{aligned} s_1 &= X_1 + X_2 + X_3 \\ s_2 &= X_1X_2 + X_1X_3 + X_2X_3 \\ s_3 &= X_1X_2X_3 \end{aligned}$$

Theorem 2.1.

(i) *Every symmetric polynomial over R can be expressed as a polynomial in $\{s_r : 1 \leq r \leq n\}$, with coefficients in R.*

(ii) *There are no non-trivial relations between $s_1, \ldots, s_n$.*

Remark:

(a) Consider the ring homomorphism

$$\theta : R[Y_1, \ldots, Y_n] \to R[X_1, \ldots, X_n],\ Y_r \mapsto s_r$$

then (i) says the image of θ is the set of symmetric polynomials. (ii) says that θ is injective.

(b) Equivalent definition of the s_r's is

$$\prod_{i=1}^{n}(T + X_i) = T^n + s_1 T^{n-1} + \ldots + s_{n-1}T + s_n$$

If we need to specify the number of variables, write $s_{r,n}$ instead of s_r.

Proof. Terminology:

- A *monomial* is some $X_I = X_1^{i_1} \ldots X_n^{i_n}$ for $I \in \mathbb{N}^n = \{0, 1, 2, \ldots\}^n$. Its (total) degree is $\sum_\alpha i_\alpha$.
- A *term* is some cX_I, for $0 \neq c \in R$. So a polynomial is uniquely a sum of terms.
- *Total degree* of f is the maximum degree over its terms

Lexicographical ordering on monomials X_I: write $X_I > X_J$ if either $i_1 > j_1$ or, for some $1 \leq r < n$, $i_1 = j_1, \ldots, i_r = j_r$ and $i_{r+1} > j_{r+1}$.

This is a total ordering: for each pair $I \neq J$, exactly one of $X_I > X_J$ or $X_J > X_I$ holds.

First we prove (ii):

Let d be the total degree of some symmetric polynomial f, and let X_I be the largest (in lexicographical order) monomial which occurs in f, with coefficient $c \in R$. As f is symmetric, we must have $i_1 \geq i_2 \geq \ldots \geq i_n$ (otherwise we could exchange variables to get a larger monomial).

So

$$X_I = X_1^{i_1 - i_2} (X_1 X_2)^{i_2 - i_3} \ldots (X_1, \ldots X_n)^{i_n}$$

consider

$$g = s_1^{i_1 - i_2} s_2^{i_2 - i_3} \ldots s_{n-1}^{i_{n-1} - i_n} s_n^{i_n}$$

the leading monomial (i.e largest in lexicographical order) of g is X_I, and g is symmetric. So $f - cg$ is symmetric of total degree $\leq d$, and its leading monomial term is smaller (lexicographical) than X_I. As the set of monomials of degree at most d is finite, this process terminates.

To prove (ii): induct on n. Suppose we have $G \in R[Y_1, \dots, Y_n]$ with $G(s_{n,1}, \dots, s_{n,n}) = 0$. We want to show $G = 0$. If $n = 1$, this is trivial ($s_{1,1} = X_1$). If $G = Y_n^k H$, with $Y_n \nmid H$, then $s_{n,n}^k H(s_{n,1}, \dots, s_{n,n}) = 0$. As $s_{n,n} = X_1 \dots X_n$, $s_{n,n}$ is not a zero divisor in $R[X_1, \dots, X_n]$ so $H(s_{n,1}, \dots, s_{n,n}) = 0$.

So we may assume G is not divisible by Y_n. Replace X_n by 0. Then

$$s_{n,r}(X_1, \dots, X_{n-1}, 0) = \begin{cases} s_{n-1,r}(X_1, \dots, X_{n-1}) & \text{if } r < n \\ 0 & \text{if } r = n \end{cases}$$

and so $G(s_{n-1,1}, \dots, s_{n-1,n-1}, 0) = 0$. So by induction, $G(Y_1, \dots, Y_{n-1}, 0) = 0$, i.e $Y_n \mid G$, a contradiction.

□

Example. $f = \sum_{i \neq j} X_i^2 X_j$ for $n \geq 3$. The leading term is $X_1^2 X_2 = X_1(X_1 X_2)$. Then compute

$$s_1 s_2 = \sum_i \sum_{j<k} X_i X_j X_k = \sum_{i \neq j} X_i^2 X_j + 3 \sum_{i<j<k} X_i X_j X_k$$

so $f = s_1 s_2 - 3 s_3$.

Computing say $\sum X_i^5$ by hand is tedious. But there are alternative formulae.

Recall $p_k = \sum_{i=1}^n X_i^k$ for $k \geq 1$.

Theorem 2.2 (Newton's formulae)**.** *Let $n \geq 1$. Then for all $k \geq 1$*

$$p_k - s_1 p_{k-1} + \dots + (-1)^{k-1} s_{k-1} p_1 + (-1)^k k s_k = 0$$

by convention, $s_0 = 1$, and $s_r = 0$ if $r > n$.

Proof. We may assume $R = \mathbb{Z}$ (or $\mathbb{R}$). Generating function

$$F(T) = \prod_{i=1}^n (1 - X_i T) = \sum_{r=0}^n (-1)^r s_r T^r$$

Take logarithmic derivative with respect to T:

$$\frac{F'(T)}{F(T)} = \sum_{i=1}^n \frac{-X_i}{1 - X_i T} = -\frac{1}{T} \sum_{i=1}^n \sum_{r=1}^\infty X_i^r T^r = -\frac{1}{T} \sum_{r=1}^\infty p_r T^r$$

So

$$-TF'(T) = s_1 T - 2 s_2 T^2 + \dots + (-1)^{n-1} n s_n T^n$$

$$= F(T) \sum_{r=1}^\infty p_r T^r = (s_0 - s_1 T + \dots + (-1)^n s_n T^n) \left(p_1 T + p_2 T^2 + \dots\right)$$

comparing coefficients of T^k gives the result. □

Definition. The *discriminant polynomial* is

$$D(X_1, \ldots, X_n) = \Delta(X_1, \ldots, X_n)^2$$

where $\Delta = \prod_{i<j}(X_i - X_j)$. (Recall from IA Groups that applying $\sigma \in S_n$ to Δ multiplies Δ by $\mathrm{sgn}(\sigma)$, so D is symmetric.)

So $D(X_1, \ldots, X_n) = d(s_1, \ldots, s_n)$ for some polynomial d ($\mathbb{Z}$-coefficients). For example, when $n = 2$, $D = (X_1 - X_2)^2 = s_1^2 - 4s_2$.

Definition. Let $f = T^n + \sum_{i=0}^{n-1} a_{n-i}T^i \in R[T]$. Its *discriminant* is $\mathrm{Disc}(f) = d(-a_1, a_2, -a_3, \ldots, (-1)^n a_n) \in R$.

Observe that if $f = \prod_{i=1}^{n}(T - x_i)$, $x_i \in R$, then $a_r = (-1)^r s_r(x_1, \ldots, x_n)$, so

$$\mathrm{Disc}(f) = \prod_{i<j}(x_i - x_j)^2 = D(x_1, \ldots, x_n)$$

If moreover $R = K$ is a field, then $\mathrm{Disc}(f) = 0$ iff f has a repeated root (i.e $x_i = x_j$ for some $i \neq j$). E.g when $n = 2$, $\mathrm{Disc}(T^2 + bT + c) = b^2 - 4c$.

3 Fields

Recall:

Definition. A *field* is a ring K (commutative with a 1) in which every non-zero element has a multiplicative inverse. The set of non-zero elements of K is a group under multiplication, written $K^\times$ or K^*, called the *multiplicative group of* K.

Definition. The *characteristic of a field* K is the least positive integer p (if it exists) such that $p \cdot 1_K = 0_K$, or is said to be 0 if no such p exists.

Example. $\mathbb{Q}$ has characteristic 0 and $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ has characteristic p (p prime).

The characteristic $\mathrm{char}(K)$ of K is either 0 or a prime. Inside K, there is a smaller subfield, called the *prime subfield* of K. It is either isomorphic to $\mathbb{Q}$ (if characteristic is 0), or to $\mathbb{F}_p$ (if $\mathrm{char}(K) = p$).

Proposition 3.1. *Let $\varphi : K \to L$ be a homomorphism of fields. Then φ is an injection.*

Proof. $\varphi(1_K) = 1_L \neq 0$, so $\mathrm{Ker}(\varphi) \subsetneq K$ is a proper ideal of K, so $\mathrm{Ker}(\varphi) = (0)$ □

Definition. Let $K \subseteq L$ be fields (where the field operations on K are the same as those on L). We say K is a *subfield of* L, and L is *an extension of* K, denoted L/K.

Remarks:

(i) The notation L/K has nothing to do with the quotient (some write $L \mid K$)

(ii) It is useful to be more general - if $i : K \to L$ is a homomorphism of fields, then Proposition 3.1 says that K is isomorphic to its image $i(K) \subseteq L$. In this situation, also say L is an extension of K.

Example. Some extensions include

- $\mathbb{C}/\mathbb{R}$
- $\mathbb{R}/\mathbb{Q}$
- $\mathbb{Q}(i) = \{a + bi : a, b \in \mathbb{Q}\}/\mathbb{Q}$

Definition. $K \subseteq L$, $x \in L$. Define $K[x] = \{p(x) : p \in K[T]\}$ (a subring of L). Define $K(x) = \{\frac{p(x)}{q(x)} : p, q \in K[T], q(x) \neq 0\}$ (a subfield of L) "K adjoin x". For $x_1, \ldots, x_n \in L$, define

$$K(x_1, \ldots, x_n) = \left\{\frac{p(x_1, \ldots, x_n)}{q(x_1, \ldots, x_n)} : p, q \in K[T_1, \ldots, T_n], q(x_1, \ldots, x_n) \neq 0\right\}$$

(Easy to check $K(x_1, \ldots, x_{n-1})(x_n) = K(x_1, \ldots, x_n)$). Likewise $K[x_1, \ldots, x_n]$ is defined analagously.

Definition. Suppose L/K is a field extension. Then L is naturally a vector space over its subfield K (forget multiplication by elements of L). We can ask if it is a finite-dimensional vector space, if so we say that L/K is a *finite extension* and write $[L : K] = \dim_K(L)$ for the dimension. The dimension is called the *degree of the extension L over K*. If the dimension is infinite write $[L : K] = \infty$.

$\dim_K$ denotes the dimension as a K-vector space. Of course L has dimension 1 over itself. As a K-vector space, $L \cong K^{[L:K]}$.

Example.

(i) $\mathbb{C}/\mathbb{R}$, $[\mathbb{C} : \mathbb{R}] = 2$

(ii) For any field K, $K(X)$ = field of rational functions in X = field of fractions of polynomial ring $K[X] = \{\frac{p}{q} : p, q \in K[X], q \neq 0\}$. Then $[K(X) : K] = \infty$ since $1, X, X^2, \ldots$ are linearly independent.

(iii) $\mathbb{R}/\mathbb{Q}$, $[\mathbb{R} : \mathbb{Q}] = \infty$. This follows from countability - every finite dimensional vector space over $\mathbb{Q}$ is countable.

This course is largely about properties (and symmetries) of finite extensions of fields.

Definition. We say an extension L/K is *quadratic* (*cubic*,...) if $[L : K] = 2(, 3, \ldots)$

Proposition 3.2. *Suppose K is a finite field (necessarily of characteristic $p > 0$). Then $|K|$ is a power of p.*

Proof. Certainly $K/\mathbb{F}_p$ is finite, so $K \cong (\mathbb{F}_p)^n$ (as a vector space), where $n = [K : \mathbb{F}_p]$, so $|K| = p^n$. □

Later on we will see that every prime power $q = p^n$ admits a field $\mathbb{F}_q$ with q elements.

Here is a simple but powerful fact:

Theorem 3.3 ("Tower Law")**.** *Suppose M/L and L/K are field extensions. Then M/K is a finite extension if and only if both M/L and L/K are finite. If so, then $[M : K] = [M : L][L : K]$.*

In fact, a slightly more general statement holds:

Theorem 3.4. *Let L/K be an extension, V an L-vector space. Then $\dim_K(V) = [L : K] \dim_L(V)$ (and obvious conclusions if any quantities are infinite).*

Example. If $V = \mathbb{C}^n$ then $V \cong \mathbb{R}^{2n}$.

Proof. Let $\dim_L(V) = d < \infty$. Then $V \cong L \oplus \ldots \oplus L = L^d$ as an L-vector space, so also as a K-vector space. If $[L : K] = n < \infty$, then $L \cong K^n$ as a K-vector space, so

$$V \cong \underbrace{K^n \oplus \ldots \oplus K^n}_{d \text{ times}} = K^{nd}$$

so $\dim_K(V) = [L : K] \dim_L(V)$. If V is finite-dimensional over K, then a K-basis for V certainly spans V over L. So if $\dim_L(V) = \infty$ then $\dim_K(V) = \infty$. Likewise, if $[L : K] = \infty$ and $V \neq \{0\}$, then V has an infinite linearly independent subset, so $\dim_K(V) = \infty$. □

Another important fact:

Proposition 3.5.

(i) Let K be a field, $G \subseteq K^\times$ a finite subgroup. Then G is cyclic

(ii) If K is finite, then $K^\times$ is cyclic

Proof. We prove (i) ((ii) follows immediately): (recall from IB GRM) we can write

$$G \cong \frac{\mathbb{Z}}{m_1\mathbb{Z}} \oplus \ldots \oplus \frac{\mathbb{Z}}{m_k\mathbb{Z}}$$

where $1 < m_1 \mid m_2 \mid \ldots \mid m_k = m$. So for all $x \in G$, $x^m = 1$. As K is a field, the polynomial $T^m - 1$ has at most m roots. So $|G| < m$. Hence $k = 1$ and G is cyclic.

□

Remark: Let $K = F = \mathbb{Z}/p\mathbb{Z}$. The above says there exists $a \in \{1, \ldots, p-1\}$ such that $\mathbb{Z}/pZ = \{0\} \cup \{a, a^2, \ldots, a^{p-1}\}$. a is called a primitive root modulo p.