Algebra

Sayan Das (dassayan0013@gmail.com)

May 27, 2024

Unit-1
Theory of equations: Relation between roots and coefficients, transformation of equation, Descartes rule of signs, cubic and biquadratic equation. Inequalities, weighted A.M.-G.M.-H.M. inequality, Cauchy-Schwarz inequality.

Unit-2
Definition and examples of groups including permutation groups, dihedral groups and quaternion groups. Elementary properties of groups. Subgroups and examples of subgroups, centraliser, normaliser, center of a group, product of two subgroups.

Unit-3
Properties of cyclic groups, classification of subgroups of cyclic groups. Cycle notation for permutations, properties of permutations, even and odd permutations, alternating group, properties of cosets, Lagrange's theorem and consequences including Fermat's Little theorem.

Unit-4
External direct product of a finite number of groups, normal subgroups, quotient groups, Cauchy's theorem for finite abelian groups. Group homomorphisms, properties of homomorphisms, properties of isomorphisms. First isomorphism theorem, Cayley's theorem.

Unit-5
Definition and examples of rings, properties of rings, subrings, integral domains and fields, characteristic of a ring. [50]

Contents

§1 Classical Algebra

§1.1 Theory of equations

§1.1.1 The Fundamental Theorem of Algebra

Remark. When we speak of "a polynomial" we shall mean a *univariate polynomial* (usually in x) unless stated otherwise.

The set of all polynomials in x with coefficients over a field $\mathbb{F}$ is denoted by $\mathbb{F}[x]$, and it forms a *Euclidean domain* with the degree of any $f \in \mathbb{F}[x]$ being the norm $\deg(f)$ (which we will define in the section on ring theory). In particular, if $f, g \in \mathbb{F}[x]$ then

$$\deg(f \circ g) = \deg(f) + \deg(g).$$

We consider $f \in \mathbb{R}[x]$ with $deg(f) = n$, of the form

$$f(x) = \sum_{j=0}^{n} a_j x^{n-j} = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n.$$

We say $\alpha \in \mathbb{C}$ is a root of f iff $f(\alpha) = 0$. Suppose $(\alpha_i)_{i=1}^n$ are the roots of f, then

$$f(x) = a_0 \prod_{i=1}^{n} (x - \alpha_i) = a_0 (x - \alpha_1) \cdots (x - \alpha_n).$$

Note that we haven't assumed the roots α_i to be distinct; the number of times a particular root α_i repeats in the sequence $(\alpha_i)_{i=1}^n$ is called the *multiplicity* of α_i.

Definition 1.1 (Algebraically closed)

We say a field $\mathbb{F}$ is **algebraically closed** iff every non-constant polynomial in $\mathbb{F}[x]$ has a root in $\mathbb{F}$.

Remark. We saw in Real Analysis that the field of rationals $\mathbb{Q}$ is not complete: in particular, there is no solution to the polynomial equation $x^2 - 2 = 0$ in $\mathbb{Q}$. The field of reals $\mathbb{R}$ comes about as a completion of the rationals, with the least upper bound property of the reals allowing us to solve equations such as $x^2 - 2 = 0$.

Similarly, we find that the polynomial equation $x^2 + 1 = 0$ has no solutions in $\mathbb{R}$. We thus say that $\mathbb{R}$, and certainly also $\mathbb{Q}$, are not algebraically closed. In order to solve $x^2 + 1 = 0$ we need the field of complex numbers $\mathbb{C}$ due to its following nice property.

Theorem 1.2 (Fundamental Theorem of Algebra)

$\mathbb{C}$ is algebraically closed.

Proof. This proof is due to Fefferman. To show that $\mathbb{C}$ is algebraically closed, consider an arbitrary polynomial $P \in \mathbb{C}[z] : P(z) = \sum_{j=0}^{n} a_j z^{n-j}$. Then it suffices to show that P has a zero. First we show that $|P(z)|$ attains a maximum as z varies over the entire complex plane, and next that if $|P(z_0)|$ is the minimum of $|P(z)|$, then $P(z_0) = 0$.

Since $|P(z)| = |z|^n \left| \sum_{j=0}^{n} a_j z^{-j} \right|$ $(z \neq 0)$ we can find an $M > 0$ so large that

$$|z| > M \implies |P(z)| \geq |a_n| \tag{1}$$

whilst the continuous function $|P(z)|$ attains a minimum as z varies over the compact disc $\{z \in \mathbb{C} : |z| \leq M\}$. Suppose, then, that

$$|z| \leq M \implies |P(z)| \geq |P(z_0)|. \tag{2}$$

In particular, $P(z_0) \leq P(0) = |a_n|$ so that, by (1), $|z| > M \implies |P(z_0)| \leq |P(z)|$ and using (2) we thus get that

$$|P(z)| \geq |P(z_0)| \quad (\forall z \in \mathbb{C}). \tag{3}$$

Since $P(z) = P((z - z_0) + z_0)$ we may write $P(z)$ as a sum of powers of $z - z_0$, so that for some polynomial $Q \in \mathbb{C}[z]$,

$$P(z) = Q(z - z_0). \tag{4}$$

By (3) and (4),

$$|Q(z)| \geq |Q(0)| \quad (\forall z \in \mathbb{C}). \tag{5}$$

By (4) $P(z_0) = Q(0)$ so it suffices to show that $Q(0) = 0$. Let k be the smallest nonzero exponent for which z^k has a nonzero coefficient in Q. Then we can write

$$Q(z) = c_n + c_{n-k} z^k + \sum_{j=k+1}^{n} c_{n-j} z^j \quad (c_{n-k} \neq 0)$$

$$\implies \exists R \in \mathbb{C}[z] : Q(z) = c_n + c_{n-k} z^k + z^{k+1} R(z) \quad (c_{n-k} \neq 0). \tag{6}$$

Set $-c_n / c_{n-k} = re^{i\theta}$ and $z_1 = r^{1/k} e^{i\theta/k}$, then

$$c_{n-k} z_1^k = -c_n. \tag{7}$$

Let $\varepsilon > 0$ be arbitrary, then by (6),

$$Q(\varepsilon z_1) = c_n + c_{n-k} \varepsilon^k z_1^k + \varepsilon^{k+1} z_1^{k+1} R(\varepsilon z_1). \tag{8}$$

Since polynomials are bounded on finite discs, we can find an $N > 0$ so large that, for $0 < \varepsilon < 1$, $|R(\varepsilon z_1)| \leq N$. Then, by (7) and (8) we have, for $0 < \varepsilon < 1$,

$$\begin{aligned} |Q(\varepsilon z_1)| &\leq \left| c_n + c_{n-k} \varepsilon^k z_1^k \right| + \varepsilon^{k+1} |z_1|^{k+1} |R(\varepsilon z_1)| \\ &\leq \left| c_n + \varepsilon^k (c_{n-k} z_1^k) \right| + \varepsilon^{k+1} (|z_1|^{k+1} N) \\ &= \left| c_n - c_n \varepsilon^k \right| + \varepsilon^{k+1} (|z_1|^{k+1} N) \end{aligned}$$

$$= |c_n|\,(1-\varepsilon^k) + \varepsilon^{k+1}(|z_1|^{k+1}\,N)$$
$$= |c_n| - \varepsilon^k\,|c_n| + \varepsilon^{k+1}(|z_1|^{k+1}\,N) \tag{9}$$

If $c_n \neq 0$, then take ε so small that $\varepsilon^{k+1}(|z_1|^{k+1}\,N) < \varepsilon^k\,|c_n|$. Thus, by (9)

$$|Q(\varepsilon z_1)| \leq |c_n| - \varepsilon^k\,|c_n| + \varepsilon^{k+1}(|z_1|^{k+1}\,N) < |c_n| - \varepsilon^k\,|c_n| + \varepsilon^k\,|c_n| = |c_n| = |Q(0)|$$

which contradicts (5). So $|c_n| = 0$ and thus $Q(0) = c_n = 0$. □

Theorem 1.3

The following are equivalent:

1. The field of complex numbers is algebraically closed.
2. Every non-constant polynomial with complex coefficients has a complex root.
3. Every nonzero polynomial of degree n with complex coefficients has exactly n complex roots.

Proof. (1.) $\iff$ (2.) by definition. Now to show (2.) $\iff$ (3.). That (3.) $\implies$ (2.) is obvious, so we show (2.) $\implies$ (3.). Suppose $f \in \mathbb{C}[x] : f(x) = \sum_{j=0}^{n} a_j x^{n-j}$ with $deg(f) = n$ (so $a_0 \neq 0$). Then using (2.) there exists $\alpha_1 \in \mathbb{C} : f(\alpha_1) = 0$, so by the factor theorem for polynomials $(x-\alpha_1)$ is a factor of f i.e. $f(x) = (x-\alpha_1)f_1(x)$ for some $f_1 \in \mathbb{C}[x] : deg(f_1) = n-1$ with leading coefficient a_0.

Again, using (2.) there exists $\alpha_2 \in \mathbb{C} : f_1(\alpha_2) = 0$ and again by the factor theorem $f_1(x) = (x-\alpha_2)f_2(x)$ for some $f_2 \in \mathbb{C}[x] : deg(f_2) = n-2$ with leading coefficient a_0. In this way we get $f_{k-1}(x) = (x-\alpha_k)f_k(x)$ for $k = 2, \ldots, n$ with $deg(f_k) = n-k$, $\alpha_k \in \mathbb{C}$. Then $deg(f_n) = 0 \implies f_n$ is the constant function $f_n(x) = a_0$ so that $f_{n-1}(x) = (x-\alpha_n)a_0$.

Thus, we have $f(x) = (x-\alpha_1)f_1(x)$

$$= (x-\alpha_1)(x-\alpha_2)f_2(x)$$
$$= (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)f_n(x)$$
$$= (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)a_0$$
$$\therefore f(x) = a_0(x-\alpha_1)\cdots(x-\alpha_n) = a_0\prod_{j=1}^{n}(x-\alpha_j).$$

Now suppose there exists $\beta \in \mathbb{C} : f(\beta) = 0$ and $\beta \neq \alpha_j$ for $j = 1, \ldots, n$. Then

$$f(\beta) = a_0\prod_{j=1}^{n}(\beta-\alpha_j) = 0$$
$$(a_0 \neq 0) \implies \beta = \alpha_j \quad \forall j = 1, \ldots, n.$$

An absurdity. This means that $(\alpha_j)_{j=1}^{n}$ are all the possible zeros of f, so f has exactly n zeros. □

Theorem 1.4 (Complex conjugate root theorem)
If $f \in \mathbb{R}[x] : f(\zeta) = 0$ for some $\zeta \in \mathbb{C}$, then $f(\bar{\zeta}) = 0$.

Proof. Let $f(x) = \sum_{j=0}^{n} a_{n-j}x^j$ with each $a_{n-j} \in \mathbb{R}$. Then, $f(\zeta) = 0$

$$\implies \sum_{j=0}^{n} a_{n-j}\zeta^j = 0$$

$$\implies \overline{\sum_{j=0}^{n} a_{n-j}\zeta^j} = \bar{0}$$

$$\implies \sum_{j=0}^{n} \overline{a_{n-j}\zeta^j} = 0$$

$$\implies \sum_{j=0}^{n} a_{n-j}\overline{\zeta^j} = 0$$

$$\implies \sum_{j=0}^{n} a_{n-j}\bar{\zeta}^j = f(\bar{\zeta}) = 0.$$

Thus, $f(\bar{\zeta}) = 0$. □

Remark. The above proof only worked because all the coefficient were real, so $a_{n-k} = \overline{a_{n-k}}$. Indeed, the complex conjugate root theorem is not necessarily true for $f \in \mathbb{C}[x]$.

Theorem 1.5 (Conjugate radical root theorem)
If $P \in \mathbb{Q}[x] : P(s + t\sqrt{u}) = 0$ for some $s, t, u \in \mathbb{Q}$, $\sqrt{u} \notin \mathbb{Q}$, then $P(s - t\sqrt{u}) = 0$.

Proof. Put $Q(x) = P(s+tx) = \sum b_k x^k$. Clearly, the b_k are rational, and $Q(\sqrt{u}) = 0$. We have

$$Q(\sqrt{u}) = \sum_{2|k} b_k u^{k/2} + \sqrt{u}\sum_{2\nmid k} b_k u^{(k-1)/2} = A + \sqrt{u}B.$$

Now as A, B are rationals and $\sqrt{u}$ not, we must have $A = B = 0$, and hence $Q(-\sqrt{u}) = A - \sqrt{u}B = 0$, and we are done. □

Remark. Viète found the following identities relating the roots of a polynomial with its coefficients.

Lemma 1.6 (Viète's relations)
If $f \in \mathbb{C}[x] : f(x) = a_0\prod_{i=1}^{n}(x - \alpha_i) = \sum_{j=0}^{n} a_j x^{n-j}$, $(a_0 \neq 0)$ and

$$\sigma_1 = \sum_{1\le i\le n} \alpha_i, \quad \sigma_2 = \sum_{1\le i<j\le n} \alpha_i\alpha_j, \quad \dots \quad , \quad \sigma_n = \prod_{1\le i\le n} \alpha_i$$

with

$$\sigma_k = \sum_{1\le i_1<\cdots<i_k\le n}\left(\prod_{1\le i\le k}\alpha_i\right)$$

in general, then:

$$\sigma_1 = -\frac{a_1}{a_0},\quad \dots\quad ,\quad \sigma_k = (-1)^k\frac{a_k}{a_0},\quad \dots\quad ,\quad \sigma_n = (-1)^n\frac{a_n}{a_0}$$

Proof Sketch. Observe that

$$f(x) = \sum_{j=0}^{n} a_j x^{n-j} = 0$$

$$\iff x^n + \sum_{j=1}^{n}\frac{a_j}{a_0}x^{n-j} = 0$$

and

$$f(x) = a_0\prod_{i=1}^{n}(x-\alpha_i) = 0$$

$$\iff \prod_{i=1}^{n}(x-\alpha_i) = 0$$

$$\iff x^n-(\alpha_1+\cdots+\alpha_n)x^{n-1}+(\alpha_1\alpha_2+\cdots+\alpha_{n-1}\alpha_n)x^{n-2}+\cdots+(-1)^n(\alpha_1\cdots\alpha_n) = 0$$

$$\iff x^n - \sigma_1 x^{n-1} + \sigma_2 x^{n-2} + \cdots + (-1)^k\sigma_k x^{n-k} + \cdots + (-1)^n\sigma_n$$

$$= x^n + \sum_{k=1}^{n}(-1)^k\sigma_k x^{n-k} = 0$$

thus $\sigma_k = (-1)^k\frac{a_k}{a_0}$. □

Remark. Observe that the above relations are all polynomials in the roots $(\alpha_i)_{i=1}^n$ of the polynomial $f(x)$; moreover they are symmetric with respect to the roots $(\alpha_i)_{i=1}^n$, i.e., the relations remain invariant under any permutation of the roots. We call these $\sigma_i's$ *elementary symmetric polynomials*. We will study symmetries in general in group theory.

Before concluding this section we note that it is often easier to work with monic polynomials (polynomials having leading coefficient of 1). So we often divide by the leading coefficient:

$$a_0x^n + \cdots + a_{n-1}x + a_n = 0 \iff x^n + \cdots + \frac{a_{n-1}}{a_0}x + \frac{a_n}{a_0} = 0 \quad (a_0\neq 0).$$

$$\iff x^n + \cdots + b_1x + b_0 = 0$$

where $b_i = \frac{a_{n-i}}{a_0}$.

§1.1.2 Descartes' rule of signs

Fact 1.7

We consider $f \in \mathbb{R}[x] : f(x) = \sum_{j=0}^{n} a_j x^{n-j}$.

The number of positive roots of $f(x) = 0$ does not exceed the number of variations signs in the sequence $(\text{sgn}(a_j))_{j=0}^{n}$ of the signs of the coefficients of $f(x)$, and if less it is less by an even number.

Consequently, we also have that the number of negative roots of $f(x) = 0$ does not exceed the number of variations of signs in the sequence

$$(\text{sgn}(b_j))_{j=0}^{n} = (\text{sgn}((-1)^j a_j))_{j=0}^{n}$$

of the signs of the coefficients of $f(-x)$, and if less it is less by an even number.

Example 1.8

Let $f(x) = 5x^6 - 7x^4 + x^2 - 7x + 8$, then $f(-x) = 5x^6 - 7x^4 + x^2 + 7x + 8$.

The sequence of signs of the coefficients of $f(x)$ is $(+1, -1, +1, -1, +1)$. There are 4 variations so the no. of positive roots of $f(x) = 0$ is $0, 2$ or 4.

The sequence of signs of the coefficients of $f(-x)$ is $(+1, -1, +1, +1, +1)$. There are 2 variations so the no. of negative roots of $f(x) = 0$ is 0 or 2.

$\deg(f) = 6$ means that it has 6 complex roots by the Fundamental Theorem of Algebra. So the number of non-real complex roots can be $0, 2, 4$, or 6 by the complex conjugate root theorem. So we can summarise the nature of the roots of $f(x)$ as follows:

Positive real	Negative real	Non-real complex
4	2	0
4	0	2
2	2	2
2	0	4
0	2	4
0	0	6

§1.1.3 Transformation of Equations

Remark. Given a polynomial equation it is possible, without knowing the roots, to obtain a new equation whose roots are connected with those of the original equation by some assigned relation. The method of finding this new equation is called a transformation. Such a transformation is occasionally useful for studying the nature of the roots of the given polynomial which might have proved difficult otherwise.

In general, given a polynomial equation $f(x) = 0 : f \in \mathbb{F}[x]$, we are to obtain another polynomial equation $\varphi(y) = 0 : \varphi \in \mathbb{F}[x]$ whose roots are connected with the roots of $f(x)$ by some relation $\psi(x, y) = 0$.

We obtain $\varphi(y) = 0$ by eliminating x between $f(x) = 0$ and $\psi(x, y) = 0$.

§1.1.4 Cubics

Remark. We already know how to solve quadratic equations of the form $ax^2+bx+c=0$ by completing the square:

$$ax^2+bx+c = a(x+bH)^2 - ab^2H^2 + c$$
$$\implies ax^2+bx+c = ax^2 + 2abHx + ab^2H^2 - (ab^2H^2 - c)$$
$$\therefore H = \frac{1}{2a},\ a(x+\frac{b}{2a})^2 - \frac{b^2-4ac}{4a} = 0$$

and consequently

$$\boxed{x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}.}$$

A corollary of the quadratic formula is:

Lemma 1.9

Given any $M,\ N \in \mathbb{C}$, there exist $g,\ h \in \mathbb{C} : g+h = M$ and $gh = N$; moreover, g and h are the roots of $x^2 - Mx + N$.

Proof. The quadratic formula provides roots g and h of $x^2 - Mx + N$. Now,

$$x^2 - Mx + N = (x-g)(x-h) = x^2 - (g+h)x + gh$$

and so $g+h = M$ and $gh = N$. □

Remark. Arising from a tradition of public mathematics contests in Venice and Pisa, methods to solve equations of degree 3 (cubics) and 4 (quartics/biquadratics) were found in the early 1500s by del Ferro, Tartaglia, Ferrari and Cardano.

We now derive the general formula for the roots of a cubic. The change of variable $X = x - \frac{b}{3a}$ transforms the cubic $f(X) = aX^3 + bX^2 + cX + d$ into a simpler cubic polynomial $f(x)$ with no quadratic terms:

$$F\left(x - \frac{b}{3a}\right) = f(x) = ax^3 + \frac{(3ac-b^2)}{3a}x + \frac{2b^3 - 9abc + 27a^2d}{27a^2}$$
$$= x^3 + \underbrace{\frac{(3ac-b^2)}{3a^2}}_{q}\ x + \underbrace{\frac{2b^3 - 9abc + 27a^2d}{27a^3}}_{r}$$
$$= x^3 + qx + r$$

So, $F\left(x - \frac{b}{3a}\right) = f(x) = x^3 + qx + r$.

Theorem 1.10 (Cubic Formula)

The roots of $f \in \mathbb{R}[x] : f(x) = x^3 + qx + r$ are

$$\boxed{\alpha_1 = g+h,\quad \alpha_2 = \omega g + \omega^2 h,\quad \alpha_3 = \omega^2 g + \omega h,}$$

where $g^3 = \frac{1}{2}\left(-r + \sqrt{R}\right)$, $h = -q/3g$, $R = r^2 + \frac{4}{27}q^3$ and $\omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ is a

primitive cube root of unity. Moreover,

$$\begin{cases} R > 0 \implies \text{one real root, two complex conjugate roots} \\ R = 0 \implies \text{three real roots, at least two equal} \\ R < 0 \implies \text{three distinct real roots} \end{cases}$$

Proof. Write a root u of $f(x) = x^3 + qx + r$ as

$$u = g + h,$$

where g and h are to be chosen, and substitute:

$$\begin{aligned} 0 = f(u) &= f(g+h) \\ &= (g+h)^3 + q(g+h) + r \\ &= g^3 + h^3 + 3gh(g+h) + q(g+h) + r \\ &= g^3 + h^3 + (3gh+q)u + r. \end{aligned}$$

If $3gh + q = 0$ then $gh = -q/3$. By Lemma (1.9), given u, $-q/3 \in \mathbb{C}$ there exist g, $h \in \mathbb{C} : g + h = u$ and $gh = -q/3$; this choice forces $3gh + q = 0$, so that $g^3 + h^3 = -r$. Cubing both sides of $gh = -q/3$ we get

$$\begin{cases} g^3 + h^3 = -r, \\ g^3 h^3 = -q^3/27. \end{cases}$$

By Lemma (1.9), there is a quadratic in g^3:

$$g^6 + rg^3 - q^3/27 = 0.$$

The quadratic formula gives

$$g^3 = \frac{1}{2}\left(-r + \sqrt{r^2 + \frac{4}{27}q^3}\right) = \frac{1}{2}\left(-r + \sqrt{R}\right)$$

and $h^3 = -r - g^3 = \frac{1}{2}\left(-r - \sqrt{R}\right)$ is also a root of this quadratic. So $g^3 - h^3 = \sqrt{R}$. There are three cube roots of g^3: g, ωg, and $\omega^2 g$. Due to the constraint $gh = -q/3$, each of these has a "mate": g and $h = -q/(3g)$; ωg and $\omega^2 h = -q/(3\omega g)$; $\omega^2 g$ and $\omega h = -q/(3\omega^2 g)$ (for $\omega^3 = 1$).

When $R < 0$, we have $r^2 + \frac{4}{27}q^3 = -k^2$ so that

$$g^3 = \frac{1}{2}\left(-r + \sqrt{R}\right) = \frac{1}{2}\left(-r + i\sqrt{k}\right), \quad h^3 = \frac{1}{2}\left(-r - \sqrt{R}\right) = \frac{1}{2}\left(-r - i\sqrt{k}\right).$$

Set $-\frac{r}{2} = \rho\cos(\theta)$, $\frac{k}{2} = \rho\sin(\theta)$ where $\theta \in (-\pi, \pi]$. Then

$$g^3 = \rho(\cos(\theta) + i\sin(\theta)) \text{ and } \rho^2 = -\frac{4}{27}q^3$$

so using de Moivre's theorem we get three values of $g =$

$$\sqrt[3]{\rho}\left(\cos\left(\frac{\theta}{3}\right)+i\sin\left(\frac{\theta}{3}\right)\right),\quad \sqrt[3]{\rho}\left(\cos\left(\frac{2\pi+\theta}{3}\right)+i\sin\left(\frac{2\pi+\theta}{3}\right)\right),$$

$$\sqrt[3]{\rho}\left(\cos\left(\frac{4\pi+\theta}{3}\right)+i\sin\left(\frac{4\pi+\theta}{3}\right)\right)$$

and as $gh = -q/3$, corresponding values of h will be $\Re(g) - \Im(g)$; thus in $u = g + h$ the imaginary parts cancel out and we get real roots (as $q < 0$ when $R < 0$)

$$\boxed{2\sqrt{-\frac{4^{1/3}}{3}q}\cos\left(\frac{\theta}{3}\right),\quad 2\sqrt{-\frac{4^{1/3}}{3}q}\cos\left(\frac{2\pi+\theta}{3}\right),\quad 2\sqrt{-\frac{4^{1/3}}{3}q}\cos\left(\frac{4\pi+\theta}{3}\right).}$$

□

Example 1.11

If $f(x) = x^3 - 15x - 126$, then $q = -15$, $r = -126$ and

$$R = r^2 + \frac{4}{27}q^3 = 15876 - 500 = 15376 > 0.$$

Thus, $g^3 = \frac{1}{2}(126 + 124) = 125 \implies g = 5,\ h = 1$.

So the roots are $x = 6,\quad 5\omega + \omega^2 = -3 + 2i\sqrt{3},\quad 5\omega^2 + \omega = -3 - 2i\sqrt{3}$.

Alternatively, having found one root to be 6, the other two roots can be found as the roots of the quadratic $f(x)/(x-6) = x^2 + 6x + 21$.

Example 1.12

If $f(x) = x^3 - 7x + 6$, then $q = -7$, $r = 6$, and

$$R = r^2 + \frac{4}{27}q^3 = \frac{972 - 1372}{27} = -\frac{400}{27} < 0$$

$$\text{then } g + h = \sqrt[3]{\frac{1}{2}\left(-6 + i\frac{20\sqrt{3}}{9}\right)} + \sqrt[3]{\frac{1}{2}\left(-6 - i\frac{20\sqrt{3}}{9}\right)}$$

$$= \sqrt[3]{\left(-3 + i\frac{10\sqrt{3}}{9}\right)} + \sqrt[3]{\left(-3 - i\frac{10\sqrt{3}}{9}\right)}$$

$$= \sqrt{\frac{7}{3}}\left(\cos\left(\frac{\pi - \arctan\left(\frac{10\sqrt{3}}{27}\right)}{3}\right) + i\sin\left(\frac{\pi - \arctan\left(\frac{10\sqrt{3}}{27}\right)}{3}\right)\right)$$

$$+\sqrt{\frac{7}{3}}\left(\cos\left(\frac{\pi - \arctan\left(\frac{10\sqrt{3}}{27}\right)}{3}\right) - i\sin\left(\frac{\pi - \arctan\left(\frac{10\sqrt{3}}{27}\right)}{3}\right)\right)$$

$$= 2\sqrt{\frac{7}{3}}\left(\cos\left(\frac{\pi - \arctan\left(\frac{10\sqrt{3}}{27}\right)}{3}\right)\right) = 2\sqrt{\frac{7}{3}}\left(\sqrt{\frac{3}{7}}\right) = 2.$$

The other two roots are then

$$2\sqrt{\frac{7}{3}}\left(\cos\left(\frac{2\pi + \pi - \arctan\left(\frac{10\sqrt{3}}{27}\right)}{3}\right)\right) = 2\sqrt{\frac{7}{3}}\left(\frac{-3}{2}\sqrt{\frac{3}{7}}\right) = -3,$$

and

$$2\sqrt{\frac{7}{3}}\left(\cos\left(\frac{4\pi + \pi - \arctan\left(\frac{10\sqrt{3}}{27}\right)}{3}\right)\right) = 2\sqrt{\frac{7}{3}}\left(\frac{1}{2}\sqrt{\frac{3}{7}}\right) = 1.$$

Thus, $f(x) = (x-1)(x-2)(x+3)$.

§1.1.5 Quartics

Remark. We conclude this chapter with a discussion of quartic polynomials.

Consider the quartic $F(X) = X^4 + bX^3 + cX^2 + dX + e$ (if it isn't monic we can always transform it into a polynomial that is monic). As before, we do a change of variable $X = x - \frac{1}{4}b$ to get a simpler polynomial

$$F\left(x - \frac{1}{4}b\right) = f(x) = x^4 + qx^2 + rx + s$$

whose roots yield the roots of $F(X)$: if $f(u) = 0$ then $F\left(u - \frac{1}{4}b\right) = 0$. The quartic formula was found by Ferrari in the 1540s, but the version we discuss is from the work done by Descartes in 1637. Factorise $f(x)$ into two quadratic terms,

$$f(x) = x^4 + qx^2 + rx + s = (x^2 + jx + \ell)(x^2 - jx + m)$$

Exercise 1.13. Prove that the roots of the following equations are all real.

1. $\displaystyle\sum_{i=1}^{n} \frac{1}{x + a_i} = \frac{1}{x}, \quad a_i \in \mathbb{R}^+.$
2. $\displaystyle\sum_{i=1}^{n} \frac{1}{x + a_i} = \frac{1}{x}, \quad a_i \in \mathbb{R}^-.$
3. $\displaystyle\sum_{i=1}^{n} \frac{1}{x + a_i} = \frac{1}{x + b}, \quad b, a_i \in \mathbb{R}^+,\ b \geq a_i.$
4. $\displaystyle\sum_{i=1}^{n} \frac{1}{x + a_i} = \frac{1}{x + b}, \quad b, a_i \in \mathbb{R},\ b < a_i.$
5. $\displaystyle\sum_{i=1}^{n} \frac{A_i}{x + a_i} = x + b, \quad b, a_i, A_i \in \mathbb{R},\ A_i > 0.$

Exercise 1.14. The roots of the equation $x^3 + px^2 + qx + r = 0$, $(r \neq 0)$, are α, β, γ. Find the equation whose roots are:

1. $\frac{1}{\alpha}+\frac{1}{\beta}-\frac{1}{\gamma}, \frac{1}{\beta}+\frac{1}{\gamma}-\frac{1}{\alpha}, \frac{1}{\gamma}+\frac{1}{\alpha}-\frac{1}{\beta}$,
2. $\alpha\beta+\frac{1}{\gamma}, \beta\gamma+\frac{1}{\alpha}, \gamma\alpha+\frac{1}{\beta}$,
3. $\alpha-\frac{\beta\gamma}{\alpha}, \beta-\frac{\gamma\alpha}{\beta}, \gamma-\frac{\alpha\beta}{\gamma}$,
4. $\frac{\alpha+\beta}{\gamma}, \frac{\beta+\gamma}{\alpha}, \frac{\gamma+\alpha}{\beta}$.

§1.2 Inequalities

Theorem 1.15 (Triangle Inequality)
If $x, y, z \in \mathbb{R}$, then $\|x+y\|+\|y+z\| \geq \|x+z\|$.

Theorem 1.16 (Arithmetic Mean $\geq$ Geometric Mean $\geq$ Harmonic Mean Inequality)
If $a_1, \ldots, a_n$ are arbitrary elements of $\mathbb{R}$, then

$$\left(\frac{1}{n}\sum_{j=1}^{n} a_j\right) \geq \left(\prod_{j=1}^{n} a_j\right)^{\frac{1}{n}} \geq \left(\frac{n}{\sum_{j=1}^{n} a_j}\right).$$

Theorem 1.17 (Cauchy-Schwarz Inequality)
If $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ are arbitrary elements of $\mathbb{R}$, then

$$\left(\sum_{j=1}^{n} a_j^2\right)\left(\sum_{j=1}^{n} b_j^2\right) \geq \left(\sum_{j=1}^{n} a_j b_j\right)^2.$$

Moreover, if some $a_i \neq 0$ equality holds iff there is a $\lambda \in \mathbb{F}$ such that $a_j\lambda + b_j = 0$ for all $j = 1, \ldots, n$.

Theorem 1.18 (Bernoulli's Inequality)
If $x \in \mathbb{R}$ such that $x \geq -1$, then for every positive integer n

$$(1+x)^n \geq 1+nx.$$

Moreover, if $x > -1$ and $x \neq 0$, then $(1+x)^n > 1+nx$ for all $n \geq 2$.

Definition 1.19 (Convexity)
A function $f : D_f \to \mathbb{R}, D_f \subseteq \mathbb{R}$ *is* ***convex*** *iff* $\forall t \in (0,1)$ and $r, s \in D_f$ we have:

$$f(tr+(1-t)s) \leq tf(r)+(1-t)f(s).$$

Also, f *is* ***concave*** *iff* $-f$ *is convex.*

Theorem 1.20 (Jensen's Inequality)

Let $f : D_f \to \mathbb{R}, D_f \subseteq \mathbb{R}$ and $\{x_j\}_{j=1}^n \subseteq D_f$ with $a_1, \ldots, a_n$ being arbitrary positive reals. If

1. *f is **convex** then*

$$\frac{\sum_{j=1}^n a_j f(x_j)}{\sum_{j=1}^n a_j} \geq f\left(\frac{\sum_{j=1}^n a_j x_j}{\sum_{j=1}^n a_j}\right).$$

2. *f is **concave** then*

$$\frac{\sum_{j=1}^n a_j f(x_j)}{\sum_{j=1}^n a_j} \leq f\left(\frac{\sum_{j=1}^n a_j x_j}{\sum_{j=1}^n a_j}\right).$$

Theorem 1.21 (Minkowski's Inequality)

If $p \geq 1$ and $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are arbitrary elements of $\mathbb{R}$, then

$$\left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^n |y_k|^p\right)^{\frac{1}{p}} \geq \left(\sum_{k=1}^n |x_k + y_k|^p\right)^{\frac{1}{p}}.$$

Theorem 1.22 (Hölder's Inequality)

If $p, q \geq 1 : 1/p + 1/q = 1$ and $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are arbitrary elements of $\mathbb{R}$, then

$$\left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^n |y_k|^q\right)^{\frac{1}{q}} \geq \left(\sum_{k=1}^n |x_k + y_k|\right).$$

Theorem 1.23 (Tschebyscheff's Inequality)

If $(a_k)_{k=1}^n, (b_k)_{k=1}^n$ are either both monotonically increasing or both monotonically decreasing sequences in $\mathbb{R}$, then

$$n\left(\sum_{k=1}^n a_k b_k\right) \geq \left(\sum_{k=1}^n a_k\right)\left(\sum_{k=1}^n b_k\right).$$

Theorem 1.24 (Rearrangement Inequality)

If $b_1, \ldots, b_n$ is any rearrangement of the positive reals $a_1, \ldots, a_n$, then:

$$\sum_{i=1}^n \frac{a_i}{b_i} \geq n.$$

Theorem 1.25 (Weierstrass's Inequalities)

If $\sum_{k=1}^{n} a_k < 1 : a_k \in (0,1)$ for some arbitrary positive reals $a_1, \ldots, a_n$, then:

$$\frac{1}{1 \mp \sum_{k=1}^{n} a_k} < \prod_{k=1}^{n} (1 \pm a_k) < 1 \pm \sum_{k=1}^{n} a_k.$$

§2 Groups and Monoids

§2.1 Definitions and motivation

Remark. In classical algebra we sought formulas for the roots of a polynomial $f(x)$, involving only radicals and elementary arithmetic operations on the coefficients of $f(x)$ (if such a formula exists we say that $f(x)$ is **solvable by radicals**). We already know the quadratic formula, and have also seen Cardano's and Ferrari's general solutions for the cubic and quartic cases. Naturally the question arises: is there such a formula for the quintic case ? Moreover, is there a formula for the roots of polynomials which generalises the quadratic, cubic and quartic formulas - a formula for the roots of any polynomial of degree n ?

But first, we must define what a group is.

Definition 2.1 (Concrete group)

Let X be a set. A **group** G_X is the set of symmetries of X.

Remark. A symmetry (or *permutation*) of a set is a bijection from a set onto itself. Group theory is the study of symmetries.

The order of a symmetry is how many times we have to repeat it to get back to the original configuration. If σ denotes any symmetry in G_X, then its order is $o(\sigma) = m$ where m is the least positive integer such that $\sigma^m = id_S$, with id_S being the identity map on $G_X : \sigma \mapsto \sigma$.

Why should we study permutations ? What could they have to do with formulas for roots ? The key idea is that formulas involving radicals are necessarily ambiguous. After all, if s is an n^{th} root of a number r i.e. $s^n = r$, then ωs is also an n^{th} root of r (ω being any n^{th} root of unity), for $(\omega s)^n = \omega^n s^n = s^n = r$. Recall also Viète's relations, relating the roots of $f(x)$ in terms of *elementary symmetric polynomials* in its coefficients. So we know that the coefficients of $f(x)$ are *symmetric*, i.e., they are unchanged by permuting the roots of $f(x)$.

In 1799, Ruffini claimed that the general quintic was in fact *unsolvable by radicals*. His proof wasn't accepted, however, as although his general ideas were, in fact, correct, his proof had gaps in it.

In 1815, Cauchy introduced the multiplication of permutations and proved basic properties of what is known as the *symmetric group* S_n.

In 1824, Abel filled the gaps in Ruffini's proof by building on Cauchy's work and constructing permutations of the roots of a quintic, using certain rational functions introduced by Lagrange. We now know the result that there is no general quintic formula as the *Abel-Ruffini Theorem*.

In 1830 Évariste Galois, who met a tragic but nevertheless romantic end at an early age due to his dueling tendencies, realised the importance of what he called *groups* (subsets of S_n which are closed under composition, which we call *subgroups*) towards understanding which polynomials of any degree are solvable by radicals. He associated each polynomial $f(x)$ with a group, now called the *Galois group* of $f(x)$. He recognised conjugation, normal subgroups, quotient groups, and simple groups, and he proved that any polynomial over a field of characteristic 0 is solvable by radicals iff its Galois group is a *solvable group* (solvability being a property generalising commutativity).

We will not cover everything that Galois did just yet in this course (there will be a second course, Group Theory 2). However, we note that since Galois' time, groups have arisen in many areas within and beyond mathematics outside of the study of roots of polynomials, for they are a precise way to describe the notion of symmetry.

Example 2.2

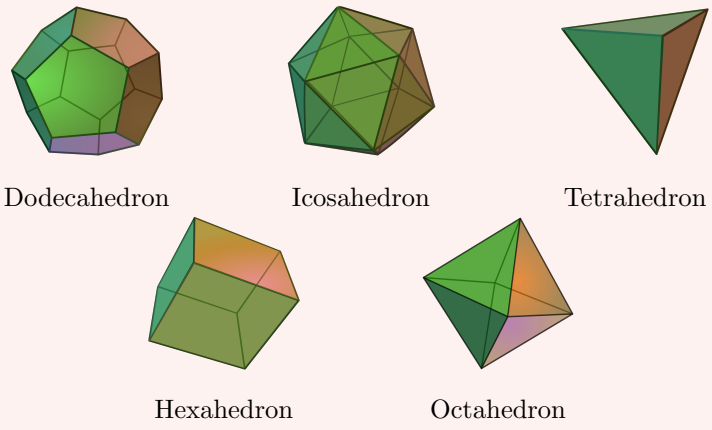
Consider a rectangle, then it has the following symmetries:

1. we do nothing
2. we reflect it horizontally
3. we reflect it vertically
4. we rotate it by π radians

We'll return to the geometric interpretation later, noting that the rectangle has a symmetric group of order 4.

Example 2.3

Consider the five regular Platonic solids. The dodecahedron has symmetries of order 5, 3, 2 and 1. It has 12 faces, so if we pick one face and put it at the bottom, we'd have 5 ways to rotate it about its top-bottom axis. So the total number of symmetries is $5 \times 12 = 60$, which is the order of its symmetric group. If we were to count reflections as well, its symmetric group would be of order 120.



The 5 regular Platonic solids.

We'll return to the notion of symmetric group and alternating group later, noting that the the dual polyhedron of the dodecahedron, the icosahedron, also has a symmetric group of order 60, the tetrahedron of order 12, and the hexahedron (cube) and the octahedron of order 24.

Remark. Note that a symmetry can be *noncommutative*: consider the transformations of a hexahedron for example.

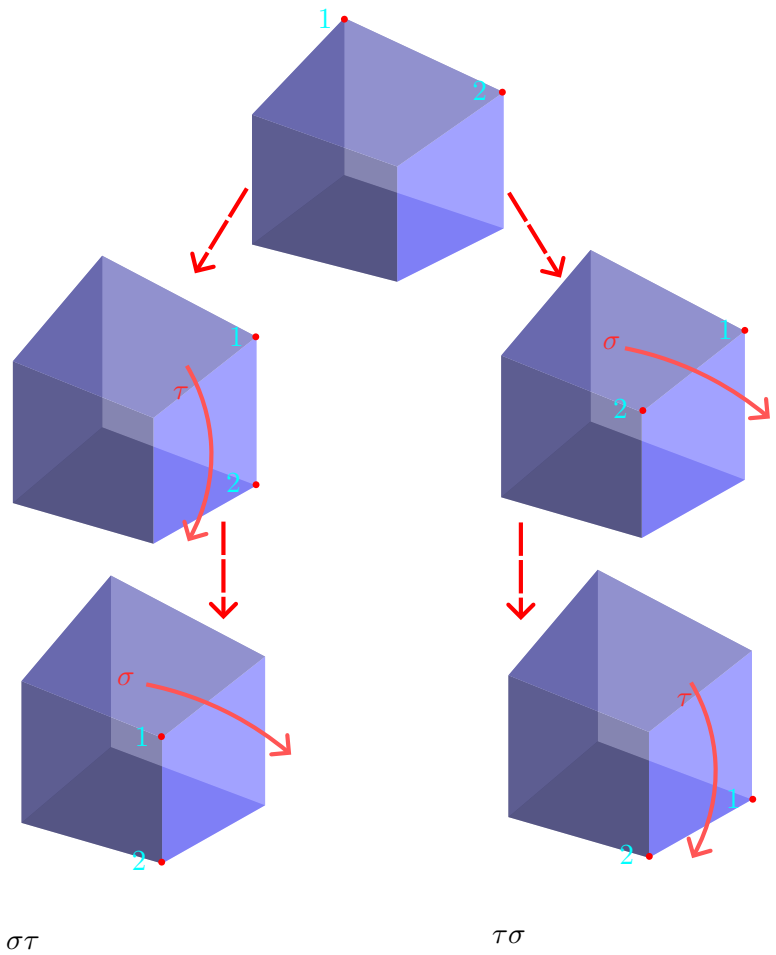


Figure 1: $\sigma\tau \neq \tau\sigma$

So it is natural that we should want some operation in which we consider *ordered pairs*.

Definition 2.4

A **binary operation** on a set R is a function $* : R \times R \to R$, denoted by $(r, r') \mapsto r * r'$.

Remark. As $*$ is a function, it is single-valued; i.e., the **law of substitution** holds: if $r = r'$ and $s = s'$, then $r * s = r' * s'$.

The above shows that $*$ is **well-defined**: the definition of $*$ assigns a unique value $r * s = q \in R$ to every r, $s \in R$; the same $(r, s) \in R \times R$ cannot have multiple different q, q_1, …, $q_n \in R$ assigned to it (although the same $q = r * s$ can coresspond to multiple different ordered pairs in $R \times R$, so a binary operation need not be injective).

Also note that $r, s \in R \implies r * s \in R$ by definition; we say that R is **closed** under $*$.

We will now make the notion of *permutation* more precise.

§2.2 Permutations

Definition 2.5 (Permutation)

A **permutation** of a set X is a bijection from X to itself.

Remark. A permutation of a finite set X can be viewed as a rearrangement, i.e., as a list with no repetitions of all the elements of X.

Example 2.6

1. The rearrangements of $X = \{1, 2, 3\}$ are :

$$123; \quad 132; \quad 213; \quad 231; \quad 312; \quad 321.$$

2. Now let $X = \{1, \ldots, n\}$. Then there are exactly $n!$ rearrangements of the n-element set X.

Remark. A rearrangement $i_1, i_2, \ldots, i_n$ of X determines a function $\alpha : X \to X$, namely, $\alpha(1) = i_1, \quad \alpha(2) = i_2, \quad \ldots, \quad \alpha(n) = i_n$.

Example 2.7

The rearrangement 213 determines a function α with $\alpha(1) = 2$, $\alpha(2) = 1$, and $\alpha(3) = 3$.

Notation. We use a two-rowed notation to denote the function corresponding a rearrangement; if $\alpha(j)$ is the j^{th} item on the list, then

$$\alpha = \begin{pmatrix} 1 & 2 & \ldots & j & \ldots & n \\ \alpha(1) & \alpha(2) & \ldots & \alpha(j) & \ldots & \alpha(n) \end{pmatrix}.$$

Remark. That a list contains all the elements of X says that the correspondence function α is surjective, for the bottom row is $\operatorname{im}\alpha$. That there are no repetitions on the list says that distinct points have distinct values, i.e., α is injective. Thus, every list determines a bijection $\alpha : X \to X$; i.e., each rearrangement determines a permutation. Conversely, each permutation determines a rearrangement, namely, the list $\alpha(1), \alpha(2), \ldots, \alpha(n)$ displayed as the bottom row. Rearrangements are permutations and permutations are rearrangements. The advantage of viewing permutations as functions, however, is that they can be composed.

Notation. We denote the family of all permutations of a set X by

$$S_X,$$

but when $X = \{1, \ldots, n\}$ we denote S_X by

$$S_n.$$

The identity permutation is usually denoted (1).

Remark. Composition is a binary operation on S_X: the composite of two permutations is a permutation. Composition in S_3 is *noncommutative*: consider permutations α, β of S_3 such that

$$\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Then,

$$\alpha\beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \beta\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Thus $\alpha\beta \neq \beta\alpha$. It follows that composition is noncommutative in S_n, $n \geq 3$. We'll now introduce some special permutations:

Definition 2.8

Let $f : X \to X$. If $x \in X$ then f **fixes** x iff $f(x) = x$, and f **moves** x iff $f(x) \neq x$. Let $i_1,\ i_2,\ \ldots,\ i_r$ be distinct integers in $X = \{1,\ 2,\ \ldots,\ n\}$. If $\alpha \in S_n$ fixes the other integers in X (if any) and if

$$\alpha(i_1) = i_2,\ \alpha(i_2) = i_3,\ \ldots,\ \alpha(i_{r-1}) = i_r,\ \alpha(i_r) = i_1,$$

then α is called an ***r*-cycle**, or a *cycle of length* r, and we denote it by

$$\alpha = \begin{pmatrix} i_1 & i_2 & \ldots & i_r \end{pmatrix}.$$

A 2-cycle $\begin{pmatrix} i_1 & i_2 \end{pmatrix}$ interchanges i_1 and i_2 and fixes the rest, and so a 2-cycle is also called a **transposition**.

Remark. Cycle comes from the Greek word for circle. The cycle $\alpha = \begin{pmatrix} i_1 & i_2 & \ldots & i_r \end{pmatrix}$ can be visualised as a clockwise rotation of the circle:

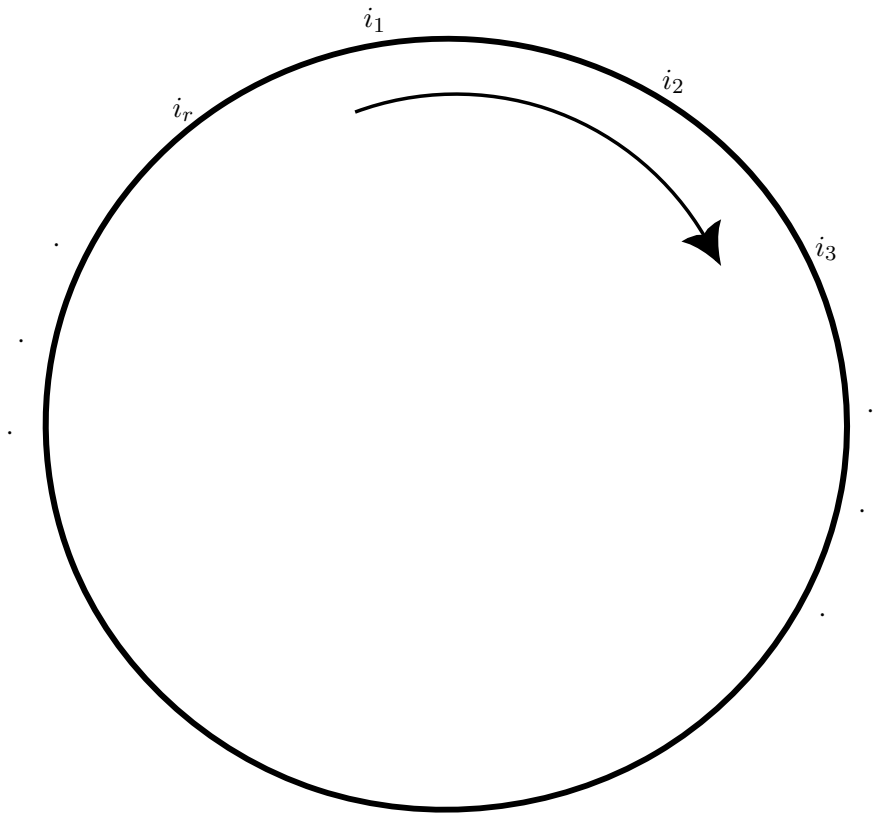


Figure 2: Cycle $\alpha = \begin{pmatrix} i_1 & i_2 & \ldots & i_r \end{pmatrix}$.

We extend the cycle notation to 1-cycles, writing $(i) = (1)$ for all i, as a 1-cycle is the identity permutation: sending i to i and fixing the rest i.e. fixing everything.

There are r different cycle notations for any r-cycle α, since any i_j can be taken as its "starting point":

$$\alpha = \begin{pmatrix} i_1 & i_2 & \dots & i_r \end{pmatrix} = \begin{pmatrix} i_2 & i_3 & \dots & i_r & i_1 \end{pmatrix} = \cdots = \begin{pmatrix} i_r & i_1 & i_2 & \dots & i_{r-1} \end{pmatrix}.$$

Definition 2.9

Two permutations $\alpha, \beta \in \S_n$ are **disjoint** iff every i moved by one is fixed by the other: if $\alpha(i) \neq i$ then $\beta(i) = i$, and if $\beta(j) \neq j$ then $\alpha(j) = j$. A family $(\beta_k)_{k=1}^t$ of permutations is *disjoint* iff each pair of them is disjoint.

§2.3 Abstract groups and monoids

Definition 2.10 (Abstract group and monoid)

A **group** is a set G with a binary operation $G \times G \to G$ (usually written $(a, b) \mapsto a + b, a \times b, a \cdot b, a \circ b$, or just ab) such that

1. There is an **identity element** in G (denoted $e, 1$, or 0) such that $ea = a = ae$ for every a in G.
2. For every $a \in G$ there exists an **inverse element** $a^{-1} \in G$ such that $aa^{-1} = e = a^{-1}a$.
3. The operation is **associative**: $(ab)c = a(bc)$ $\forall a, b, c \in G$.

If the operation is $(a, b) \mapsto a + b$, G is an additive group. If the operation is $(a, b) \mapsto ab$, G is a multiplicative group. By default, we write the operation the multiplicative way. If we do not require the second axiom (*existence of inverses*), then we have a **monoid**. If we do not require the first axiom (*existence of identity*) and second axiom, then we have a **semigroup**. If we only require closure, then we have a **groupoid** (or **magma**).

Axiom introduced	Algebraic structure
Closure	Groupoid (or magma)
Associativity	Semigroup
Identity	Monoid
Inverse	Group

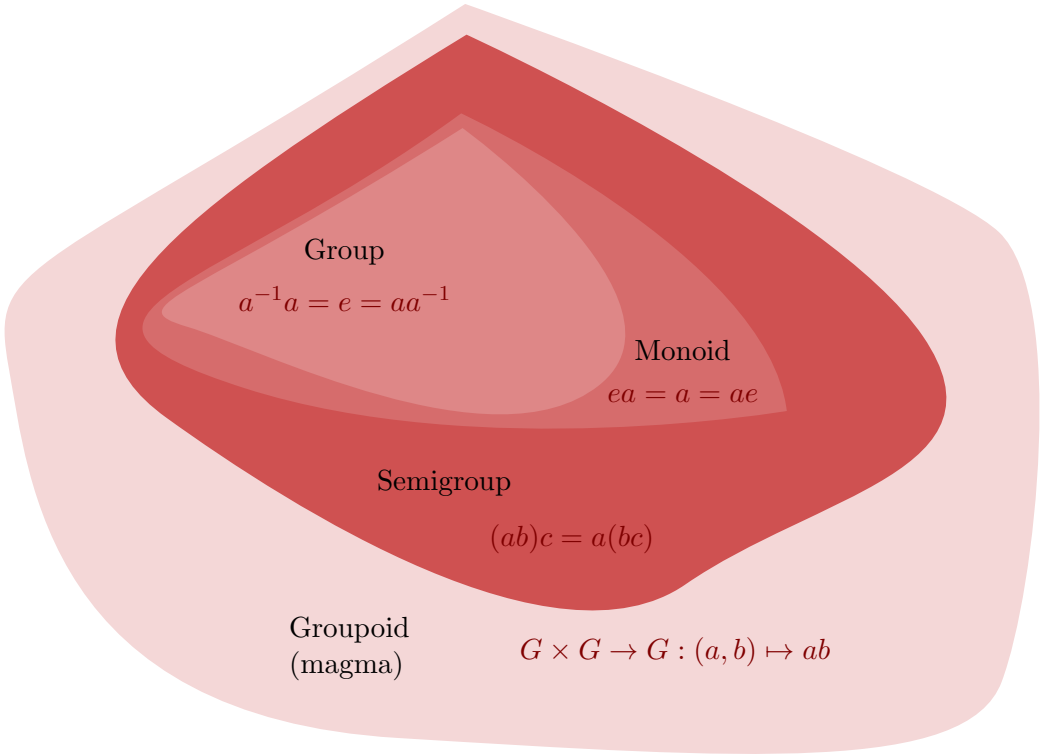


Figure 3: Some abstract algebraic structures

Remark. It is clear that if we take the composition of symmetries as our binary relation,

then the concrete notion of a group can be translated to the abstract notion. It is a subtle and important point that the converse is also true, which is what Cayley's Theorem says as we shall see in the next section.

Also, it is clear from the examples of the symmetries of a cube, the composition of permutations, subtraction of numbers and product of matrices (from linear algebra) why we want ordered pairs in the binary operation: as ab and ba can be different. Nevertheless, there are examples of groups where the commutative law $ab = ba$ holds, such as S_2. In fact, Abel proved that if the Galois group of a polynomial is commutative, then f is solvable by radicals. As a result,

Definition 2.11

A group G is called **abelian** iff it satisfies the **commutative law**:

$$ab = ba$$

for every $a, b \in G$.

There are many examples of groups.

Example 2.12 1. The set S_X of all permutations of a set X, with composition as binary operation and $1_X = (1)$ as the identity, is a group, called the **symmetric group** on X. For a finite set X with $|X| = n$ the symmetric group is denoted as S_n. The groups S_n for $n \geq 3$ are nonabelian because $\begin{pmatrix}1 & 3 & 2\end{pmatrix}$ and $\begin{pmatrix}1 & 3\end{pmatrix}$ do not commute:

$$\begin{pmatrix}1 & 3 & 2\end{pmatrix}\begin{pmatrix}1 & 3\end{pmatrix} = \begin{pmatrix}1 & 2\end{pmatrix} \neq \begin{pmatrix}3 & 2\end{pmatrix} = \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}1 & 3 & 2\end{pmatrix}.$$

2. Consider $A, B, C \in M_n(\mathbb{F})$ (the set of square matrices of order n over the field $\mathbb{F}$) with addition operation $+$. Then $A + B \in M_n(\mathbb{F})$ so it is closed under addition. Also $A + (B + C) = (A + B) + C$ and $A + (-1)(A) = A - A = O$ where O is the $n \times n$ zero matrix and (-1) is the inverse of the identity element of $\mathbb{F}$. In fact we also have $A + B = B + A$. Thus, $(M_n(\mathbb{F}), +)$ is an abelian group.

3. Consider the same set as above, but with matrix product $\cdot : M_n(\mathbb{F}) \times M_n(\mathbb{F}) \to M_n(\mathbb{F})$, $(A, B) \mapsto P$ defined by

$$p_{ij} = \sum_{1 \leq k \leq n} a_{ik} b_{kj}.$$

So,

$$AB = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \end{pmatrix} \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} & \dots & \begin{pmatrix} a_{11} & \dots & a_{1n} \end{pmatrix} \begin{pmatrix} b_{1n} \\ \vdots \\ b_{nn} \end{pmatrix} \\ \vdots & \ddots & \vdots \\ \begin{pmatrix} a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} & \dots & \begin{pmatrix} a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{1n} \\ \vdots \\ b_{nn} \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} p_{11} & \dots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \dots & p_{nn} \end{pmatrix}$$

which is clearly in $M_n(\mathbb{F})$. However, it may so happen that $\det(A) = 0$, i.e., that A is **singular** $\iff$ there is no $B \in M_n(\mathbb{F})$ such that $AB = I$ where I is the $n \times n$ identity matrix. So A may not be invertible, and this set therefore can't be a group under matrix product unless subjected to certain constraints as in the following examples.

4. For a field $\mathbb{F}$ and positive integer n consider the set

$$GL_n(\mathbb{F}) = \{A \in M_n(\mathbb{F}) : \overbrace{det(A) \neq 0}^{\text{nonsingular}}\}$$

with the operation of matrix multiplication. If $A, B \in GL_n(\mathbb{F})$ then A^{-1}, B^{-1} exist. Now, $(AB)^{-1} = B^{-1}A^{-1}$. Now

$$\det(B^{-1}A^{-1}) = \det(B^{-1}\det(A^{-1})) = \frac{1}{\det(A)\det(B)} \neq 0.$$

So AB is nonsingular, thus the operation is closed. Matrix product is associative, I is the identity, and every element, being nonsingular, has an inverse by definition. Thus, $GL_n(\mathbb{F})$ forms a nonabelian group, called the **general linear group**.

5. For a field $\mathbb{F}$ and positive integer n the set

$$SL_n(\mathbb{F}) = \{A \in M_n(\mathbb{F}) : \det(A) = 1\}$$

is also a nonabelian group called the **special linear group**.

Example 2.13

A field $\mathbb{F}$ is a group under addition and $\mathbb{F}^* = \mathbb{F} \setminus \{0\}$ is a group under multiplication.

Example 2.14

Let X be a set and let 2^X denote the power set of X. Define addition over 2^X to be the symmetric difference $A+B = (A \cup B) \setminus (A \cap B)$, and multiplication over 2^X to be intersection $A \cap B$. The **Boolean group** $\mathcal{B}(X)$ is the additive group $(2^X, +)$, with $\varnothing$ as the zero element and with every element being its own inverse. Also, $(2^X, \cap)$ is a semigroup.

Example 2.15

The set $\mathbb{Z}$ of all integers is an additive abelian group under ordinary addition $(a, b) \mapsto a + b$, with identity 0 and $-n$ being the additive inverse of each $n \in \mathbb{Z}$. However, $\mathbb{Z}^*$ is not a group under multiplication; aside from ± 1 none of the elements in $\mathbb{Z}^*$ have a multiplicative inverse.

The situation changes when we consider the integers modulo m for some positive integer m.

Lemma 2.16

Let G be a group.

(i) **Cancellation Law**: If either $xa = xb$ or $ax = bx$, then $a = b$.

(ii) The identity element $e \in G$ is the unique element with $ea = a = ae$ for any $a \in G$.

(iii) Every $a \in G$ has a unique inverse: there is only one element $a^{-1} \in G$ such that $aa^{-1} = e = a^{-1}a$.

(iv) $(a^{-1})^{-1} = a$ for any $a \in G$.

Proof. (iii) Let $xa = e$, $xb = e$. Then $e = e \Rightarrow xa = xb \Rightarrow a(xa) = a(xb)$

$$\Rightarrow ae = (ax)b \Rightarrow a = (xa)b = eb = b.$$

□

§2.4 Isomorphism and Cayley's theorem

Theorem 2.17

Every group G is isomorphic to a subgroup of the symmetric group S_G.

Theorem 2.18 (Generalised Cayley's theorem)

Let G be a group and $H \subseteq G$ be a subgroup of G. Let $X = G/H$ be the set of left cosets of H in G. Let N be the **normal core** of H in G, defined to be the intersection of the conjugates of H in G. Then the quotient group G/N is isomorphic to a subgroup of $\mathrm{Sym}(X)$.

Remark. We'll prove the generalised Cayley's theorem in Group Theory 2.

§2.5 Submonoids, subgroups and cyclic groups

§2.6 Orbits and cosets

§2.7 Congruences, quotient monoids and groups

§2.8 Homomorphisms

Problem 2.19. Show that $(\mathbb{Q}^+, \cdot) \cong (\mathbb{Z}[x], +)$.

Problem 2.20. Show that $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn} \iff \gcd(m,n) = 1$.

Problem 2.21. Let G be an abelian subgroup of the symmetric group S_n and $p_1, \cdots, p_k$ be all prime divisors of $|G|$. Prove that $p_1 + \cdots + p_k \leq n$.

§3 Rings