Part IA - Groups

Lectured by J. Goedecke

Michaelmas 2014

Examples of groups
Axioms for groups. Examples from geometry: symmetry groups of regular polygons, cube, tetrahedron. Permutations on a set; the symmetric group. Subgroups and homomorphisms. Symmetry groups as subgroups of general permutation groups. The Möbius group; cross-ratios, preservation of circles, the point at infinity. Conjugation. Fixed points of Möbius maps and iteration. [4]

Lagranges theorem
Cosets. Lagranges theorem. Groups of small order (up to order 8). Quaternions. Fermat-Euler theorem from the group-theoretic point of view. [5]

Group actions
Group actions; orbits and stabilizers. Orbit-stabilizer theorem. Cayley's theorem (every group is isomorphic to a subgroup of a permutation group). Conjugacy classes. Cauchy's theorem. [4]

Quotient groups
Normal subgroups, quotient groups and the isomorphism theorem. [4]

Matrix groups
The general and special linear groups; relation with the Möbius group. The orthogonal and special orthogonal groups. Proof (in $\mathbb{R}^3$) that every element of the orthogonal group is the product of reflections and every rotation in $\mathbb{R}^3$ has an axis. Basis change as an example of conjugation. [3]

Permutations
Permutations, cycles and transpositions. The sign of a permutation. Conjugacy in S_n and in A_n. Simple groups; simplicity of A_5. [4]

Contents

1 Groups and homomorphisms

1.1 Groups

Definition (Binary operation)**.** A *(binary) operation* is a way of combining two elements to get a new element. Formally, it is a map $* : A \times A \to A$.

Definition (Group)**.** A *group* is a set G with a binary operation $*$ satisfying the following axioms:

0. (Closure) $\forall a, b \in G, a * b \in G$
1. (Identity) $\exists e \in G (\forall a \in G (a * e = e * a = a))$
2. (Inverse) $\forall a \in G (\exists a^{-1} \in G (a * a^{-1} = a^{-1} * a = e))$
3. (Associativity) $\forall a, b, c \in G ((a * b) * c = (a * (b * c)))$

Definition (Abelian group)**.** A group is *abelian* if it satisfies

4. (Commutativity) $\forall a, b \in G (a * b = b * a)$

Note: if it is clear from context, the operation $*$ is often left out. i.e. $a * b$ is written as ab.

Example. The following are abelian groups:

(i) $\mathbb{Z}$ with $+$

(ii) $\mathbb{Q}$ with $+$

(iii) $\mathbb{Z}_n$ (integers mod n) with $+_n$

(iv) $\mathbb{Q}^*$ with $\times$

(v) $\{-1, 1\}$ with $\times$

The following are non-abelian groups:

(vi) Symmetries of a regular triangle (or any n-gon) with composition. (D_{2n})

(vii) 2×2 invertible matrices with matrix multiplication ($\mathrm{GL}_2(\mathbb{R})$)

(viii) Symmetry groups of 3D objects

Proposition. Let $(G, *)$ be a group. Then

(i) The identity is unique.

(ii) Inverses are unique.

Proof.

(i) Suppose e and e' are inverses. Then we have $ee' = e'$, treating e as an inverse, and $ee' = e$, treating e' as an inverse. Thus $e = e'$.

(ii) Suppose a^{-1} and b both satisfy the identity axiom for some $a \in G$. Then $b = be = b(aa^{-1}) = (ba)a^{-1} = ea^{-1} = a^{-1}$. Thus $b = a^{-1}$.

$\square$

Proposition. Let $(G, *)$ be a group and $a, b \in G$. Then

(i) $(a^{-1})^{-1} = a$

(ii) $(ab)^{-1} = b^{-1}a^{-1}$

Proof.

(i) Given a^{-1}, both a and $(a^{-1})^{-1}$ satisfy

$$xa^{-1} = ax^{-1} = e.$$

By uniqueness of inverses, $(a^{-1})^{-1} = a$.

(ii) We have

$$\begin{aligned}(ab)(b^{-1}a^{-1}) &= a(bb^{-1})a^{-1} \\ &= aea^{-1} \\ &= aa^{-1} \\ &= e\end{aligned}$$

Similarly, $(b^{-1}a^{-1})ab = e$. So $b^{-1}a^{-1}$ is an inverse of ab. By the uniqueness of inverses, $(ab)^{-1} = b^{-1}a^{-1}$.

$\square$

Definition (Order of group)**.** The *order* of the group, denoted as $|G|$, is the number of elements in G. A group is a finite group is the order is finite.

Definition (Subgroup)**.** A *subgroup* $H \leq G$ is a subset $H \subseteq G$ such that H with the restricted operation $*$ from G is also a group.

Example. The following are subgroups:

- $(\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +) \leq (\mathbb{C}, +)$
- $(e, *) \leq (G, *)$ (trivial subgroup)
- $G \leq G$
- $(\{\pm 1\}, \times) \leq (\mathbb{Q}^*, \times)$

Lemma (Subgroup criteria I)**.** Let $(G, *)$ be a group and $H \subseteq G$. $H \leq G$ iff

(i) $e \in H$

(ii) $\forall a, b \in H (ab \in G)$

(iii) $\forall a \in H (a^{-1} \in H)$

Proof. The group axioms are satisfied as follows:

0. Closure: (ii)

1. Identity: (i). Note that H and G must have the same identity. Suppose that e_H and e_G are the identities of H and G respectively. Then $e_H e_H = e_H$. Now e_H has an inverse in G. Thus we have $e_H e_H e_H^{-1} = e_H e_H^{-1}$. Thus we have $e_H = e_G$.

2. Inverse: (iii)

3. Associativity: inherited from G.

$\square$

Lemma (Subgroup criteria II)**.** A subset $H \subseteq G$ is a subgroup of G iff:

(I) H is non-empty

(II) $\forall a, b \in H (ab^{-1} \in H)$

Proof. Proof that (I) and (II) imply (i), (ii) and (iii):

(i) H must contain at least one element a. Then $aa^{-1} = e \in H$.

(iii) $ea^{-1} = a^{-1} \in H$.

(ii) $a(b^{-1})^{-1} = ab \in H$.

Proof that (i), (ii) and (iii) imply (I) and (II): trivial. $\square$

Proposition. The subgroups of $(\mathbb{Z}, +)$ are exactly $n\mathbb{Z}$, for $n \in \mathbb{N}$. ($n\mathbb{Z}$ is the integer multiples of n)

Proof. Firstly, for any $n \in \mathbb{N}$, $n\mathbb{Z}$ is a subgroup (trivial). Now show that any subgroup must be in the form $n\mathbb{Z}$.

Let $H \leq \mathbb{Z}$. We know $0 \in H$. Pick the smallest positive integer n in H (well-ordering principle). Then $H = n\mathbb{Z}$.

Otherwise, suppose $\exists a \in H (a \nmid n)$. Let $a = pn + q$, where $0 < q < n$. Since $a - pn \in H$, $q \in H$. Yet $q < n$ but n is the smallest member of H. Contradiction. So every $a \in H$ is divisible by n, and $H = n\mathbb{Z}$. Such an n must exist unless $H = 0$, in which case $H = 0\mathbb{Z}$. $\square$

1.2 Homomorphisms

Definition (Function)**.** Given 2 sets X, Y, a *function* $f : X \to Y$ sends each $x \in X$ to a particular $f(x) \in Y$. X is called the domain and Y is the co-domain.

Example. The following are functions:

- Identity function: $1_X : X \to X$ with $1_X(x) = x$. Also written as id_X
- Inclusion map: $\iota : \mathbb{Z} \to \mathbb{Q}$: $\iota(n) = n$. Note that this differs from the identity function as the domain and codomain are different in the inclusion map.
- $f_1 : \mathbb{Z} \to \mathbb{Z}$: $f_1(x) = x + 1$.
- $f_2 : \mathbb{Z} \to \mathbb{Z}$: $f_2(x) = 2x$.
- $f_3 : \mathbb{Z} \to \mathbb{Z}$: $f_3(x) = x^2$.

- For $g : 0, 1, 2, 3, 4 \to 0, 1, 2, 3, 4$, we have:
    - $g_1(x) = x + 1$ if $x < 4$; $g_1(4) = 4$.
    - $g_2(x) = x + 1$ if $x < 4$; $g_1(4) = 0$.

Definition (Composition of functions)**.** The *composition* of two functions is a function you get by applying one after another. In particular, if $f : X \to Y$ and $G : Y \to Z$, then $g \circ f : X \to Z$ with $g \circ f(x) = g(f(x))$.

Example. $f_2 \circ f_1(x) = 2x + 2$. $f_1 \circ f_2(x) = 2x + 1$. Note that function composition is not commutative.

Definition (Injective functions)**.** A function f is *injective* if it hits everything at most once, i.e.

$$\forall x, y \in X(f(x) = f(y) \Rightarrow x = y)$$

Definition (Surjective functions)**.** A function is *surjective* if it hits everything at least once, i.e.

$$\forall y \in Y(\exists x \in X(f(x) = y))$$

Definition (Bijective functions)**.** A function is *bijective* if it is both injective and surjective. i.e. it hits everything exactly once. Note that a function has an inverse iff it is bijective.

Example. ι and f_2 are injective but not subjective. f_3 and g_1 are neither. 1_X, f_1 and g_2 are bijective.

Lemma. The composition of two bijective functions is bijective

Definition (Group homomorphism)**.** Let $(G, *)$ and $(H, \times)$ be groups. A function $f : G \to H$ is a *group homomorphism* iff

$$\forall g_1, g_2 \in G : f(g_1) \times f(g_2) = f(g_1 * g_2),$$

i.e. they "preserve group properties"

Definition (Group isomorphism)**.** *Isomorphisms* are bijective homomorphisms. 2 groups are *isomorphic* if there exists an isomorphism between them. We write $G \cong H$.

Example. The following are isomorphisms or homomorphisms:

- $1_G : G \to G$ and $f_2 : \mathbb{Z} \to 2\mathbb{Z}$ are isomorphisms. $\iota : \mathbb{Z} \to \mathbb{Q}$ and $f_2 : \mathbb{Z} \to \mathbb{Z}$ are homomorphisms.
- $\exp : (\mathbb{R}, +) \to (\mathbb{R}^+, \times)$ with $\exp(x) = e^x$ is an isomorphism.
- Take $(\mathbb{Z}_4, +)$ and $H : (\{e^{ik\pi/2} : k = 0, 1, 2, 3\}, \times)$. Then $f : \mathbb{Z}_4 \to H$ by $f(a) = e^{i\pi a/2}$ is an isomorphism.
- $f : \mathrm{GL}_2(\mathbb{R}) \to \mathbb{R}^*$ with $f(A) = \det(A)$ is a homomorphism, where $\mathrm{GL}_2(\mathbb{R})$ is the set of 2×2 invertible matrices.

Proposition. Suppose that $f : G \to H$ is a homomorphism. Then

(i) Homomorphisms send the identity to the identity, i.e.

$$f(e_G) = e_H$$

(ii) Homomorphisms send inverses to inverses, i.e.

$$f(a^{-1}) = f(a)^{-1}$$

(iii) The composite of 2 group homomorphisms is a group homomorphism.

(iv) The inverse of an isomorphism is an isomorphism.

Proof. The proofs of the statements (i), (iii) and (iv) are as follows:

(i)

$$\begin{aligned} f(e_G) &= f(e_G^2) = f(e_G)^2 \\ f(e_G)^{-1} f(e_G) &= f(e_G)^{-1} f(e_G)^2 \\ f(e_G) &= e_H \end{aligned}$$

(iii) Let $f : G_1 \to G_2$ and $g : G_2 \to G_3$. Then $g(f(ab)) = g(f(a)f(b)) = g(f(a))g(f(b))$.

(iv) Let $f : G \to H$ be an isomorphism. Then

$$\begin{aligned} f^{-1}(ab) &= f^{-1}\{f[f^{-1}(a)]f[f^{-1}(b)]\} \\ &= f^{-1}\{f[f^{-1}(a)f^{-1}(b)]\} \\ &= f^{-1}(a)f^{-1}(b) \end{aligned}$$

□

Definition (Image of homomorphism)**.** If $f : G \to H$ is a homomorphism, then the *image* of f is

$$\operatorname{im} f = f(G) = \{f(g) : g \in G\}.$$

Definition (Kernel of homomorphism)**.** The *kernel* of f, written as

$$\ker f = f^{-1}(\{e_H\}) = \{g \in G : f(g) = e_H\}.$$

Proposition. Both the image and the kernel are subgroups of the respective groups, i.e. $\operatorname{im} f \leq H$ and $\ker f \leq G$.

Proof. Since $e_H \in \operatorname{im} f$ and $e_G \in \ker f$, $\operatorname{im} f$ and $\ker f$ are non-empty. Moreover, suppose $b_1, b_2 \in \operatorname{im} f$. Now $\exists a_1, a_2 \in G$ such that $f(a_i) = b_i$. Then $b_1 b_2^{-1} = f(a_1)f(a_2^{-1}) = f(a_1 a_2^{-1}) \in \operatorname{im} f$.

Then consider $b_1, b_2 \in \ker f$. We have $f(b_1 b_2^{-1}) = f(b_1)f(b_2)^{-1} = e^2 = e$. So $b_1 b_2^{-1} \in \ker f$. □

Proposition. Given homomorphism $f : G \to H$ and some $a \in G$, for all $k \in \ker f$, $aka^{-1} \in \ker f$ (i.e. the kernel is simple)

Proof. $f(aka^{-1}) = f(a)f(k)f(a)^{-1} = f(a)ef(a)^{-1} = e$. So $aka^{-1} \in \ker f$. □

Example. Images and kernels for previously defined functions:

(i) For the identity function, $\operatorname{im} 1_G = G$ and $\ker 1_G = \{e\}$.

(ii) For the inclusion map $\iota : \mathbb{Z} \to \mathbb{Q}$, we have $\operatorname{im} \iota = \mathbb{Z}$ and $\ker \iota = \{0\}$

(iii) For $f_2 : \mathbb{Z} \to \mathbb{Z}$ and $f_2(x) = 2x$, we have $\operatorname{im} f_2 = 2\mathbb{Z}$ and $\ker f_2 = \{0\}$.

(iv) For $\det : \mathrm{GL}_2(\mathbb{R}) \to \mathbb{R}^*$, we have $\operatorname{im} \det = \mathbb{R}^*$ and $\ker \det = \{A : \det A = 1\} = \mathrm{SL}_2(\mathbb{R})$

Proposition. For all homomorphisms $f : G \to H$, f is

(i) surjective iff $\operatorname{im} f = H$

(ii) injective iff $\ker f = \{e\}$

Proof.

(i) By definition.

(ii) We know that $f(e) = e$. So if f is injective, then by definition $\ker f = \{e\}$. If $\ker f = \{e\}$, then given a, b such that $f(a) = f(b)$, $f(ab^{-1}) = f(a)f(b)^{-1} = e$. Thus $ab^{-1} \in \ker f = \{e\}$. Then $ab^{-1} = e$ and $a = b$.

□

1.3 Cyclic groups

Notation. We write $a^2 = aa$, $a^n = \underbrace{aaa \cdots a}_{n\text{copies}}$, $a^0 = e$, $a^{-n} = (a^{-1})^n$.

Definition (Cyclic group C_n)**.** A group G if *cyclic* if $\exists a \in G(\forall b \in G(\exists n \in \mathbb{Z}(b = a^n)))$, i.e. every element is some power of a. Such an a is called a generator of G.

Example. The following groups are cycle:

(i) $\mathbb{Z}$ is cyclic with generator 1 or -1. It is *the* infinite cyclic group.

(ii) $(\{+1, -1\}, \times)$ is cyclic with generator -1.

(iii) $(\mathbb{Z}_n, +)$ is cyclic with all numbers coprime with n as generators.

Notation. Given a group G and $a \in G$, we write $\langle a \rangle$ for the cyclic group generated by a, i.e. the subgroup of all powers of a. It is the smallest subgroup containing a.

Definition (Order of element)**.** The *order* of an element a is the smallest integer n such that $a^n = e$. If k doesn't exist, a has infinite order. Write $\mathrm{ord}(a)$ for the order of a.

Definition (Exponent of group)**.** The *exponent* of a group G is the smallest integer n such that $\forall a(a^n = e)$.

Lemma. For a in g, $\mathrm{ord}(a) = |\langle a \rangle|$.

Proof. If $\text{ord}(a) = \infty$, $a^n \neq a^m$ for all $n \neq m$. Otherwise $a^{m-n} = e$. Thus $|\langle a \rangle| = \infty = \text{ord}(a)$.

Otherwise, suppose $\text{ord}(a) = k$. Thus $a^k = e$. We now claim that $\langle a \rangle = \{e, a^1, a^2, \cdots a^{k-1}\}$. Note that $\langle a \rangle$ does not contain higher powers of a as $a^k = e$ and higher powers will loop back to existing elements. There are also no repeating elements in the list provided since $a^m = a^n \Rightarrow a^{m-n} = e$. $\square$

Notation. We write C_n for the cyclic group of order n. e.g. $C_n \cong (\mathbb{Z}_n, +)$.

Proposition. Cyclic groups are abelian.

1.4 Dihedral groups

Definition (Dihedral groups D_{2n})**.** Dihedral groups are the symmetries of a regular n-gon. It contains n rotations (including the identity symmetry, i.e. rotation by $0°$) and n reflections. All rotations are generated by $r = \frac{360°}{n}$. r has order n. Any reflection has order 2.

Now consider any reflection s. Then r and s generate the whole group. We have

$$\begin{aligned} D_{2n} &= \langle r, s | r^n = e = s^2, srs^{-1} = r^{-1} \rangle \\ &= \{e, r, r^2, \cdots r^{n-1}, s, rs, r^2 s, \cdots r^{n-1} s\} \end{aligned}$$

Note that we have $sr = r^{-1}s$ and $sr^k = r^{-k}s = r^{n-k}s$.

1.5 Direct products of groups

Definition (Direct product of groups)**.** Given two groups $(G_1, *_1)$ and $(G_2, *_2)$, we can define a set $G_1 \times G_2 = \{(g_1, g_2) : g_i \in G_i\}$ and an operation $(a_1, a_2) * (b_1, b_2) = (a_1 *_1 b_1, a_2 *_2 b_2)$. This forms a group.

Example.

$$\begin{aligned} C_2 \times C_2 &= \{(0,0), (0,1), (1,0), (1,1)\} \\ &= \{e, x, y, xy\} \text{ with everything order 2} \\ &= \langle x, y | x^2 = y^2 = e, xy = yx \rangle \end{aligned}$$

Proposition. $C_n \times C_m \cong C_{nm}$ iff $\text{hcf}(m, n) = 1$.

Proof. Let $C_n = \langle a \rangle$ and $C_m = \langle b \rangle$. Let k be the order of (a, b). Then $(a, b)^k = (a^k, b^k) = e$. This is possible only if $n | k$ and $m | k$, i.e. k is a common multiple n and m. Since the order is the minimum value of k that satisfies the above equation, $k = \text{lcm}(n, m) = \frac{nm}{\text{hcf}(n,m)} = nm$.

Now consider $\langle (a, b) \rangle \leq C_n \times C_m$. Since (a, b) has order nm, $\langle (a, b) \rangle$ has nm elements. Since $C_n \times C_m$ also has nm elements, $\langle (a, b) \rangle$ must be the whole of $C_n \times C_m$. And we know that $\langle (a, b) \rangle \cong C_{nm}$. So $C_n \times C_m \cong C_{nm}$. $\square$

Proposition (Direct product theorem)**.** Let $H_1, H_2 \leq G$. If

(i) $H_1 \cap H_2 = \{e\}$

(ii) $\forall a_i \in H_i (a_1 a_2 = a_2 a_1)$

(iii) $\forall a \in G(\exists a_1 \in H_1, a_2 \in H_2(a = a_1 a_2))$. (Also known as: $G = H_1 H_2$)

Then $G \cong H_1 \times H_2$.

Proof. Define $f : H_1 \times H_2 \to G$ by $f(a_1, a_2) = a_1 a_2$. Then it is a homomorphism since

$$\begin{aligned} f((a_1, a_2) * (b_1, b_2)) &= f(a_1 b_1, a_2 b_2) \\ &= a_1 b_1 a_2 b_2 \\ &= a_1 a_2 b_1 b_2 \\ &= f(a_1, a_2) f(b_1, b_2). \end{aligned}$$

Surjectivity follows from (iii). We'll show injectivity by showing that the kernel is $\{e\}$. If $f(a_1, a_2) = e$, then we know that $a_1 a_2 = e$. Then $a_1 = a_2^{-1}$. Since $a_1 \in H_1$ and $a_2^{-1} \in H_2$, we have $a_1 = a_2^{-1} \in H_1 \cap H_2 = \{e\}$. Thus $a_1 = a_2 = e$ and $\ker f = \{e\}$. $\square$

2 Symmetric group I

Definition (Permutation)**.** A *permutation* of X is a bijection from a set X to X itself. The set of all permutations on X is $\mathrm{Sym}\, X$.

Theorem. Sym X with composition forms a group.

Note: X can be an infinite set.

Proof. The groups axioms are satisfied as follows:

0. If $\sigma : X \to X$ and $\tau : X \to X$, then $\sigma \circ \tau : X \to X$. If they are both bijections, then the composite is also bijective. So if $\sigma, \tau \in \mathrm{Sym}\, X$, then $\sigma \circ \tau \in \mathrm{Sym}\, X$.
1. The identity $1_X : X \to X$ is clearly a permutation, and gives the identity of the group.
2. Every bijective function has a bijective inverse. So if $\sigma \in \mathrm{Sym}\, X$, then $\sigma^{-1} \in \mathrm{Sym}\, X$.
3. Composition of functions is associative.

$\square$

Definition (Symmetric group S_n)**.** If X is finite, say $|X| = n$ (usually use $X = \{1, 2, \cdots n\}$), we write $\mathrm{Sym}\, X = S_n$. The is THE *symmetric group* of degree n.

Notation. (Two row notation) We write $1, 2, 3, \cdots n$ on the top line and their images below, e.g.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in S_3 \text{ and } \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \end{pmatrix} \in S_5$$

In general, if $\sigma : X \to X$, we write

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}$$

Note: The order of the first row doesn't mater. When composing, we can render second element: e.g.

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

Proposition. $|S_n| = n!$

Note: degree $\neq$ order.

Example. For small n, we have

(i) When $n = 1$, $S_n = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} = \{e\} \cong C_1$.

(ii) When $n = 2$, $S_n = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\} \cong C_2$

(iii) When $n = 3$,

$$S_n = \left\{ \begin{matrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \\ \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, & \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \end{matrix} \right\} \cong D_6.$$

Note: S_3 is not abelian. Thus S_n is not abelian for $n \geq 3$ since we can always view S_3 as a subgroup of S_n by fixing $4, 5, 6, \cdots n$. *Note*: We can also view D_{2n} as a subgroup of S_n because each symmetry is a permutation of the corners. (with the restriction that adjacent corners have to remain adjacent after permutation)

Notation. (Cycle notation) e.g. in S_3, write $(1\ 2\ 3)$ for 1 goes to 2, 2 goes to 3 and 3 goes to 1. Note that $(1\ 2\ 3) = (2\ 3\ 1) = (3\ 1\ 2)$, but usually write the smallest number first. We leave out numbers that don't move, .e.g. $(1\ 2)$

In S_4, we can have $(1\ 2)(3\ 4)$.

Note: The order of each cycle is the length of the cycle, and the inverse is the cycle written the other way round, e.g. $(1\ 2\ 3)^{-1} = (3\ 2\ 1) = (1\ 3\ 2)$.

Example. Composition (right to left):

(i) $(1\ 2\ 3)(1\ 2)$. We have 1 goes to 2 and then to 3; 2 goes to 1 which goes back to 2; 3 goes to 1. Thus $(1\ 2\ 3)(1\ 2) = (1\ 3)(2)$

(ii) $(1\ 2\ 3\ 4)(1\ 4) = (1)(2\ 3\ 4) = (2\ 3\ 4)$

Definition (k-cycles and transpositions)**.** We call $(a_1\ a_2\ a_3 \cdots a_k)$ *k-cycles*. 2-cycles are called *transpositions*. Two cycles are *disjoint* if no number appears in both cycles.

Example. $(1\ 2)$ and $(3\ 4)$ are disjoint but $(1\ 2\ 3)$ and $(1\ 2)$ are not

Lemma. Disjoint cycles commute.

Proof. If $\sigma, \tau \in S_n$ are disjoint cycles. Consider any n. Show that: $\sigma(\tau(a)) = \tau(\sigma(a))$. If a is in neither of σ and τ, then $\sigma(\tau(a)) = \tau(\sigma(a)) = a$. Otherwise, wlog assume that a is in τ but not in σ. Then $\tau(a) \in \tau$ and thus $\tau(a) \notin \sigma$. Thus $\sigma(a) = a$ and $\sigma(\tau(a)) = \tau(a)$. Therefore we have $\sigma(\tau(a)) = \tau(\sigma(a)) = \tau(a)$. Therefore τ and σ commute. □

Note: Non-disjoint cycles may not commute. e.g. $(1\ 3)(2\ 3) = (1\ 3\ 2)$ while $(2\ 3)(1\ 3) = (1\ 2\ 3)$.

Theorem. Any permutation in S_n can be written (essentially) uniquely as a product of disjoint cycles. (Essentially unique means unique up to re-ordering of cycles and rotation within cycles, e.g. $(1\ 2)$ and $(2\ 1)$)

Proof. Let $\sigma \in S_n$. Start with $(1\ \sigma(1)\ \sigma^2(1)\ \sigma^3(1)\ \cdots)$. As the set $\{1, 2, 3 \cdots n\}$ is finite, for some k, we must have $\sigma^k(1)$ already in the list. If $\sigma^k(1) = \sigma^l(1)$, with $l < k$, then $\sigma^{k-l}(1) = 1$. So all $\sigma^i(1)$ are distinct until we get back to 1. Thus we have the first cycle $(1\ \sigma(1)\ \sigma^2(1)\ \sigma^3(1)\ \cdots\ \sigma^{k-1}(1))$.

Now choose the smallest number that is not yet in a cycle, say j. Repeat to obtain a cycle $(j\ \sigma(j)\ \sigma^2(j)\ \cdots\ \sigma^{l-1}(j))$. Since σ is a bijection, nothing in this cycle can be in previous cycles as well.

Repeat until all $\{1, 2, 3 \cdots n\}$ are exhausted. This is essentially unique because every number j completely determines the whole cycle it belongs to, whichever number we start with, we'll end up with the same cycle. □

Definition (Cycle type)**.** Write a permutation $\sigma \in S_n$ in disjoint cycle notation. The *cycle type* is the list of cycle lengths. This is unique up to re-ordering. We often (but not always) leave out singleton cycles.

Example. (1 2) has cycle type 2 (transposition). (1 2)(3 4) has cycle type $2, 2$ (double transposition). (1 2 3)(4 5) has cycle type $3, 2$.

Lemma. For $\sigma \in S_n$, the order of σ is the least common multiple of cycle lengths in the disjoint cycle notation. In particular, a k-cycle has order k.

Proof. As disjoint cycles commute, we can group together each cycle when we take powers. i.e. if $\sigma = \tau_1\tau_2 \cdots \tau_l$ with τ_i all disjoint cycles, then $\sigma^m = \tau_1^m\tau_2^m \cdots \tau_l^m$.

Now if cycle τ_i has length k_i, then $\tau_i^{k_i} = e$, and $\tau_i^m = e$ iff $k_i | m$. To get an m such that $\sigma^m = e$, we need all k_i to divide m. i.e. m is a common multiple of k_i. Since the order is the least possible m such that $\sigma^m = e$, the order is the least common multiple of k_i. □

Example. Any transpositions and double transpositions have order 2.

(1 2 3)(4 5) has order 6.

2.1 Sign of permutations

Proposition. Every permutation is a product of transpositions.

Proof. As each permutation is a product of disjoint cycles, it suffices to prove that each cycle is a product of transpositions. Consider a cycle $(a_1\ a_2\ a_3\ \cdots\ a_k)$. This is in fact equal to $(a_1\ a_2)(a_2\ a_3) \cdots (a_{k-1}\ a_k)$. Thus a k-cycle can be written as a product of $k - 1$ transpositions. □

Note: The product is not unique. e.g. $(1\ 2\ 3\ 4\ 5) = (1\ 2)(2\ 3)(3\ 4)(4\ 5) = (1\ 2)(2\ 3)(1\ 2)(3\ 4)(1\ 2)(4\ 5)$.

Theorem. Writing $\sigma \in S_n$ as a product of transpositions in different ways, σ is either always composed of an even number of transpositions, or always an odd number of transpositions.

Proof. Write $\#(\sigma)$ for the number of cycles in disjoint cycle notation, including singleton cycles. So $\#(e) = n$ and $\#((1\ 2)) = n - 1$. When we multiply σ by a transposition $\tau = (c\ d)$ (wlog assume $c < d$),

– If c, d are in the same σ-cycle, say, $(c\ a_2\ \cdots\ a_{k-1}\ d\ a_{k+1}\ \cdots a_{k+l})(c\ d) = (c\ a_{k+1}\ a_{k+2}\ \cdots a_{k+l})(d\ a_2\ a_3\ \cdots\ a_{k-1})$. So $\#(\sigma\tau) = \#(\sigma) + 1$.

– If c, d are in different σ-cycles, say

$$\begin{aligned}&(d\ a_2\ a_3\ \cdots\ a_{k-1})(c\ a_{k+1}\ a_{k+2}\ \cdots\ a_{k+l})(c\ d)\\&= (c\ a_2\ \cdots\ a_{k-1}\ d\ a_{k+1}\ \cdots a_{k+l})(c\ d)(c\ d)\\&= (c\ a_2\ \cdots\ a_{k-1}\ d\ a_{k+1}\ \cdots a_{k+l}) \text{ and } \#(\sigma\tau) = \#(\sigma) - 1.\end{aligned}$$

Therefore for any transposition τ, $\#(\sigma\tau) \equiv \#(\sigma) + 1 \pmod 2$.

Now suppose $\sigma = \tau_1 \cdots \tau_l = \tau'_1 \cdots \tau'_k$. Since disjoint cycle notation is unique, $\#(\sigma)$ is uniquely determined by σ.

Now we can construct σ by starting with e and multiplying the transpositions one by one. Each time we add a transposition, we increase $\#(\sigma)$ by $1 \pmod 2$. So $\#(\sigma) \equiv \#(e) + l \pmod 2$. Similarly, $\#(\sigma) \equiv \#(e) + k \pmod 2$. So $l \equiv l' \pmod 2$. □

Definition (Sign of permutation)**.** Viewing $\sigma \in S_n$ as a product of transpositions, $\sigma = \tau_1 \cdots \tau_l$, we call $\operatorname{sgn}(\sigma) = (-1)^l$. If $\operatorname{sgn}(\sigma) = 1$, we call σ an even permutation. If $\operatorname{sgn}(\sigma) = -1$, we call σ an odd permutation.

Note: While l itself is not well-defined, it is either always odd or always even, and $(-1)^l$ is well-defined.

Theorem. For $n \geq 2$, $\operatorname{sgn} : S_n \to \{\pm 1\}$ is a surjective group homomorphism.

Proof. Suppose $\sigma_1 = \tau_1 \cdots \tau_{l_1}$ and $\sigma_2 = \tau'_1 \cdots \tau_{l_2}$. Then $\operatorname{sgn}(\sigma_1\sigma_2) = (-1)^{l_1+l_2} = (-1)^{l_1}(-1)^{l_2} = \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2)$. So it is a homomorphism.

It is surjective since $\operatorname{sgn}(e) = 1$ and $\operatorname{sgn}((1\ 2)) = -1$. □

Note: The hard bit is showing that sgn is well defined. If question asks to show that sgn is a well-defined group homomorphism, you HAVE to show that it is well-defined.

Lemma. σ is an even permutation iff the number of cycles of even length is even.

Proof. A k-cycle can be written as $k - 1$ transpositions. Thus an even-length cycle is odd, vice versa.

Since sgn is a group homomorphism, writing σ in disjoint cycle notation, $\sigma = \sigma_1\sigma_2 \cdots \sigma_l$, we get $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma_1) \cdots \operatorname{sgn}(\sigma_l)$. Suppose there are m even-length cycles and n odd-length cycles, then $\operatorname{sgn}(\sigma) = (-1)^m 1^n$. This is equal to 1 iff $(-1)^m = 1$, i.e. m is even. □

Note: odd length cycles are even, even length cycles are odd.

Definition (Alternating group A_n)**.** The *alternating group* A_n is the kernel of sgn, i.e. the even permutations. Since A_n is a kernel of a group homomorphism, $A_n \leq S_n$.

Note: sgn is used in the definition of the determinant of a matrix: if $A_{n\times n}$ then

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Proposition. Any subgroup of S_n contains either no odd permutations or exactly half.

Proof. If S_n has at least one odd permutation τ, then there exists a bijection between the odd and even permutations by $\sigma \mapsto \sigma\tau$ (bijection since $\sigma \mapsto \sigma\tau^{-1}$ is a well-defined inverse). So there are as many odd permutations as even permutations. $\square$

3 Lagrange's Theorem

Definition (Cosets)**.** Let $H \leq G$ and $a \in G$. Then the set $aH = \{ah : h \in H\}$ is a *left coset* of H and $Ha = \{ha : h \in H\}$ is a *right coset* of H.

Example. Consider the following examples:

(i) Take $2\mathbb{Z} \leq Z$. Then $6 + 2\mathbb{Z} = \{\text{all even numbers}\} = 0 + 2\mathbb{Z}$. $1 + 2\mathbb{Z} = \{\text{all odd numbers}\} = 17 + 2\mathbb{Z}$.

(ii) Take $G = S_3$, let $H = \langle(1\ 2)\rangle = \{e, (1\ 2)\}$. The left cosets are

$$\begin{aligned} eH = (1\ 2)H &= \{e, (1\ 2)\} \\ (1\ 3)H = (1\ 2\ 3)H &= \{(1\ 3), (1\ 2\ 3)\} \\ (2\ 3)H = (1\ 3\ 2)H &= \{(2\ 3)(1\ 3\ 2)\} \end{aligned}$$

(iii) Take $G = D_6$ (which is isomorphic to S_3). Recall $D_6 = \langle r, s | r^3 e = s^2, rs = sr^{-1}\rangle$.Take $H = \langle s\rangle = \{e, s\}$. We have left coset $rH = \{r, re = sr^{-1}\}$ and the right coset $Hr = \{r, sr\}$. Thus $rH \neq Hr$.

Note: we can have $aH = bH$ with $a \neq b$.

Definition (Partition)**.** Let X be a set, and $X_1, \cdots X_n$ be subsets of X. The X_i are called a *partition* of X if $\bigcup X_i = X$ and $X_i \cap X_j = \emptyset$ for $i \neq j$. i.e. every element is in exactly one of X_i.

Lemma. The left cosets of a subgroup $H \leq G$ partition G, and every coset has the same size.

Proof. For each $a \in G$, $a \in aH$. Thus the union of all cosets gives all of G. Now we have to show that for all $a, b \in G$, the cosets aH and bH are either the same or disjoint.

Suppose that aH and bH are not disjoint. So $\exists h_1 \in H$ s.t. $ah_1 \in bH$ (i.e. $ah_1 \in aH \cap bH$). i.e. $ah_1 = bh_2$ for some $h_2 \in H$. So $a = bh_2h_1^{-1}$. So for any $h \in H$, $ah = b(h_2h_1^{-1}h)$. Since $h_2h_1^{-1}h \in H$ by closure, $ah \in bH$. Thus $aH \subseteq bH$. Similarly, $bH \subseteq aH$ and $aH = bH$.

To show that they each coset has the same size, note that $f : H \to aH$ with $f(h) = ah$ is invertible with inverse $f^{-1}(h) = a^{-1}h$. Thus there exists a bijection between them and they have the same size. $\square$

Theorem (Lagrange's theorem)**.** If G is a finite group and H is a subgroup of G, then $|H|$ divides $|G|$.

Note: The converse is not true. If k divides $|G|$, there is not necessarily a subgroup of order k, e.g. $|A_4| = 12$ but there is no subgroup of order 6. c.f. Cauchy's Theorem

Proof. Suppose that there are $|G : H|$ left cosets in total. Since the left cosets partition G, and each coset has size $|H|$, we have

$$|H||G : H| = |G|.$$

Thus $|H|$ divides $|G|$. $\square$

Definition (Index of a subgroup)**.** The *index* of H in G ($|G : H|$) is the number of left cosets in G.

Proposition. $aH = bH \Leftrightarrow b^{-1}a \in H$.

Proof. ($\Rightarrow$) $ah_1 = bh_2$. Then $b^{-1}a = h_2h_1^{-1} \in H$.

($\Leftarrow$). Let $b^{-1}a = h_0$. Then $a = bh_0$. Then $\forall ah \in aH$, we have $ah = b(h_0h) \in bH$. So $aH \subseteq bH$. Similarly, $bH \subseteq aH$. So $aH = bH$. □

Corollary. The order of an element divides the order of the group, i.e. for any finite group G and $a \in G$, $\text{ord}(a)$ divides $|G|$.

Proof. Consider the subgroup generated by a, which has order $\text{ord}(a)$. Then by Lagrange's theorem, $\text{ord}(a)$ divides $|G|$. □

Corollary. The exponent of a group divides the order of the group, i.e. for any finite group G and $a \in G$, $a^{|G|} = e$.

Proof. We know that $|G| = k \,\text{ord}(a)$ for some $k \in \mathbb{N}$. then $a^{|G|} = (a^{\text{ord}(a)})^k = e^k = e$. □

Corollary. Groups of prime order are cyclic and are generated by every non-identity element.

Proof. Say $|G| = p$. If $a \in G$ is not the identity, the subgroup generated by a must have order p since it has to divide p. Thus the subgroup generated by a has the same size as G and they must be equal. Then G must be cyclic since it is equal to the subgroup generated by a. □

Definition (Equivalence relation)**.** An *equivalence relation* $\sim$ is a relation that is reflexive, symmetric and transitive. i.e.

(i) Reflexive: $\forall x(x \sim x)$

(ii) Symmetric: $\forall x, y(x \sim y \Rightarrow y \sim x)$

(iii) Transitive $\forall x, y, z[(x \sim y) \wedge (y \sim z) \Rightarrow x \sim z]$

Example. The following relations are equivalence relations:

(i) Consider Z. The relation $\equiv_n$ defined as $a \equiv_n b \Leftrightarrow n|(a - b)$.

(ii) Consider the set (formally: class) of all finite groups. Then "is isomorphic to" is an equivalence relation.

Definition (Equivalence class)**.** Given an equivalence relation $\sim$ on A, the *equivalence class* of a is

$$[a]_\sim = [a] = \{b \in A | a \sim b\}$$

Proposition. The equivalence classes form a partition of A.

Proof. By reflexivity, we have $a \in [a]$. Thus the equivalence classes cover the whole set. We must now show that for all $a, b \in A$, either $[a] = [b]$ or $[a] \cap [b] = \emptyset$.

Suppose $[a] \cap [b] \neq \emptyset$. Then $\exists c \in [a] \cap [b]$. So $a \sim c, b \sim c$. By symmetry, $c \sim b$. By transitivity, we have $a \sim b$. For all $b' \in [b]$, we have $b \sim b'$. Thus by transitivity, we have $a \sim b'$. Thus $[b] \subseteq [a]$. Similarly, $[a] \subseteq [b]$ and $[a] = [b]$. □

Lemma. Given a set G and a subset H, define the equivalence relation on G with $a \sim b$ iff $b^{-1}a \in H$. The equivalence classes are the left cosets of H.

Proof. First show that it is an equivalence relation.

(i) Reflexive: Since $aa^{-1} = e \in H$, $a \sim a$.

(ii) Symmetric: $a \sim b \Rightarrow b^{-1}a \in H \Rightarrow (b^{-1}a)^{-1} = a^{-1}b \in H \Rightarrow b \sim a$.

(iii) Transitive: If $a \sim b$ and $b \sim c$, we have $b^{-1}a, c^{-1}b \in H$. So $c^{-1}bb^{-1}a = c^{-1}a \in H$. So $a \sim c$.

Then show that the equivalence classes are the cosets: $a \sim b \Leftrightarrow b^{-1}a \in H \Leftrightarrow aH = bH$. □

Example. Consider $(\mathbb{Z}, +)$, and for fixed n, take the subgroup $n\mathbb{Z}$. The cosets are $0 + H, 1 + H, \cdots (n-1) + H$. We can write these as $[0], [1], [2] \cdots [n]$. To calculate "mod n", define $[a] + [b] = [a + b]$, and $[a][b] = [ab]$. We need to check that it is well-defined, i.e. it doesn't depend on the choice of the representative of $[a]$.

If $[a_1] = [a_2]$ and $[b_1] = [b_2]$, then $a_1 = a_2 + kn$ and $b_1 = b_2 + kn$, then $a_1 + b_1 = a_2 + b_2 + n(k + l)$ and $a_1b_1 = a_2b_2 + n(kb_2 + la_2 + kln)$. So $[a_1 + b_1] = [a_2 + b_2]$ and $[a_1b_1] = [a_2b_2]$.

We have seen that $(\mathbb{Z}_n, +_n)$ is a group. What happens with multiplication? We can only take elements which have inverses (these are called units, c.f. Groups, Rings and Modules). Call the set of them $U_n = \{[a] | (a, n) = 1\}$. We'll see these are the units.

Definition (Euler totient function)**.** (Euler totient function) $\phi(n) = |U_n|$.

Example. If p is a prime, $\phi(n) = p - 1$. $\phi(4) = 2$.

Proposition. U_n is a group under multiplication mod n.

Proof. The operation is well-defined as shown above. To check the axioms:

0. Closure: if a, b are coprime to n, then $a \cdot b$ is also coprime to n. So $[a], [b] \in U_n \Rightarrow [a] \cdot [b] = [a \cdot b] \in U_n$

1. Identity: $[1]$

2. Let $[a] \in U_n$. Consider the map $U_n \to U_n$ with $[c] \mapsto [ac]$. This is injective: if $[ac_1] = [ac_2]$, then n divides $a(c_1 - c_2)$, so as a is coprime to n, n divides $c_1 - c_2$, so $[c_1] = [c_2]$. Since U_n is finite, any injection $(U_n \to U_n)$ is also a surjection. So there exists a c such that $[a][c] = [a][c] = 1$. So $[c] = [a]^{-1}$.

3. Associativity (and also commutativity): inherited from $\mathbb{Z}$.

□

Theorem. (Fermat-Euler theorem) Let $n \in N$ and $a \in \mathbb{Z}$ coprime to n. Then

$$a^{\phi(n)} \equiv 1 \pmod{n}.$$

In particular, (Fermat's Little Theorem) if $n = p$ is a prime, then for any a not a multiple of p.

$$a^{p-1} \equiv 1 \pmod{p}.$$

Proof. As a is coprime with n, $[a] \in U_n$. Then $[a]^{|U_n|} = [1]$, i.e. $a^{\phi(n)} \equiv 1 \pmod n$. $\square$

Note: : If you want to use this as a proof of Fermat-Euler in a "non-groups" setting, you must at least prove that U_n is a group.

3.1 Small groups

Example. (Using Lagrange theorem to find subgroups) To find subgroups of D_{10}, we know that the subgroups must have size 1, 2, 5 or 10:

1: $\{e\}$

2: The groups generated by the 5 reflections of order 2

5: The group must be cyclic since it has prime order 5. It is then generated by an element of order 5, i.e. r, r^2, r^3 and r^4. They generate the same group $\langle r\rangle$.

10: D_{10}

As for D_8, subgroups must have order 1, 2, 4 or 8.

1: $\{e\}$

2: 5 elements of order 2, namely 4 reflections and r^2.

4: First consider the subgroup isomorphic to C_4, which is $\langle r\rangle$. There are two other non-cyclic group.

8: D_8

Proposition. Any group of order 4 is either isomorphic to C_4 or $C_2 \times C_2$.

Proof. Let $|G| = 4$. By Lagrange theorem, possible element orders are 1 (e only), 2 and 4. If there is an element $a \in G$ of order 4, then $G = \langle a\rangle \cong C_4$.

Otherwise all non-identity elements have order 2. Then G must be abelian (For any a, b, $(ab)^2 = 1 \Rightarrow ab = (ab)^{-1} \Rightarrow ab = b^{-1}a^{-1} \Rightarrow ab = ba$). Pick 2 elements of order 2, say $b, c \in G$, then $\langle b\rangle = \{e, b\}$ and $\langle c\rangle = \{e, c\}$ so $\langle b\rangle \cap \langle c\rangle = \{e\}$. As G is abelian, $\langle b\rangle$ and $\langle c\rangle$ commute. We know that $bc = cb$ has order 2 as well, and is the only element of G left. So $G \cong \langle b\rangle \times \langle c\rangle \cong C_2 \times C_2$ by the direct product theorem $\square$

Proposition. A group of order 6 is either cyclic or dihedral (i.e. $\cong C_6$ or D_6). (See proof in next section)

3.2 Left and right cosets

As $|aH| = |H|$ and similarly $|H| = |Ha|$, left and right cosets have the same size. Are they necessarily the same?

Example. Consider the following subgroups:

(i) Take $G = (\mathbb{Z}, +)$ and $H = 2\mathbb{Z}$. We have $0 + 2\mathbb{Z} = 2\mathbb{Z} + 0 =$ even numbers and $1 + 2\mathbb{Z} = 2\mathbb{Z} + 1 =$ odd numbers. Since G is abelian, $aH = Ha$ for all $a, \in G$, $H \leq G$.

(ii) Let $G = D_6 = \langle r, s | r^3 = e = s^2, rs = sr^{-1} \rangle$. Let $K = \langle r \rangle$. Since the cosets partition G, so one must be U and the other $sU = \{s, sr = r^2 s, sr^2 = rs\} = Us$. So for all $a \in G, aU = Ua$.

(iii) Let $G = D_6$ and take $H = \langle s \rangle$. We have $H = \{e, s\}$, $rH = \{r, rs = sr^{-1}\}$ and $r^2 H = \{r^2, r^s\}$; while $H = \{e, s\}, Hr = \{r, sr\}$ and $Hr^2 = \{r^2, sr^2\}$. So the left and right subgroups do now coincide.

4 Quotient groups

4.1 Normal subgroups

Definition (Normal subgroup)**.** A subgroup K of G is a *normal subgroup* if $\forall a \in G(\forall k \in K(aka^{-1} \in K))$. We write $K \lhd G$. This is equivalent to:

(i) $\forall a \in G(aK = Ka)$, i.e. left coset = right coset

(ii) $\forall a \in G(aKa^{-1} = K$ (c.f. conjugacy classes)

From the example last time, $H = \langle s \rangle \leq D_6$ is not a normal subgroup, but $K = \langle r \rangle \lhd D_6$. We know that every group G has at least two normal subgroups $\{e\}$ and G.

Lemma.

(i) Every subgroup of index 2 is normal.

(ii) Any subgroup of an abelian group is normal.

Proof.

(i) If $K \leq G$ has index 2, then there are only two possible cosets K and $G \setminus K$. As $eK = Ke$ and cosets partition G, the other left coset and right coset must be $G \setminus K$. So all left cosets and right cosets are the same.

(ii) For all $a \in G$ and $k \in K$, we have $aka^{-1} = aa^{-1}k = k \in K$.

□

Proposition. Every kernel is a normal subgroup.

Proof. Given homomorphism $f : G \to H$ and some $a \in G$, for all $k \in \ker f$, we have $f(aka^{-1}) = f(a)f(k)f(a)^{-1} = f(a)ef(a)^{-1} = e$. Therefore $aka^{-1} \in \ker f$ by definition of the kernel. □

In fact, we will see in the next section that all normal subgroups are kernels of some homomorphism.

Example. Consider $G = D_8$. Let $K = \langle r^2 \rangle$ is normal. Check: Any element of G is either sr^ℓ or r^ℓ for some ℓ. Clearly e satisfies $aka^{-1} \in K$. Now check r^2: For the case of sr^ℓ, we have $sr^\ell r^2(sr^\ell)^{-1} = sr^\ell r^2 r^{-\ell} s^{-1} = sr^2 s = ssr^{-2} = r^2$. For the case of r^ℓ, $r^\ell r^2 r^{-\ell} = r^2$.

Proposition. A group of order 6 is either cyclic or dihedral (i.e. $\cong C_6$ or D_6).

Proof. Let $|G| = 6$. By Lagrange theorem, possible element orders are $1, 2, 3$ and 6. If there is an $a \in G$ of order 6, then $G = \langle a \rangle \cong C_6$. Otherwise, we can only have elements of orders 2 and 3 other than the identity. If G only has elements of order 2, the order must be a power of 2 by Sheet 1 Q. 8, which is not the case. So there must be an element r of order 3. So $\langle r \rangle \lhd G$ as it has index 2. Now G must also have an element s of order 2 by Sheet 1 Q. 9.

Since $\langle r \rangle$ is normal, we know that $srs^{-1} \in \langle r \rangle$. If $srs^{-1} = e$, then $r = e$, which is not true. If $srs^{-1} = r$, then $sr = rs$ and sr has order 6 (lcm of the orders of s And r, which was ruled out above. Otherwise if $srs^{-1} = r^2 = r^{-1}$, then G is dihedral by definition of the dihedral group. □

4.2 Quotient groups

Intuition: Consider $\mathbb{Z}_{12} = \{0, 1, 2 \cdots 12\}$. Consider some random equivalence classes $\{0, 3\}$, $\{1, 5\}$, $\{2, 7\}$, $\{4, 11\}$, $\{6, 9\}$ and $\{8, 10\}$. Suppose we want to add them. If we add $\{1, 5\} + \{4, 11\}$ by just randomly picking one element from each class, add them together, and see which equivalence class it lands in, we can get many different possible classes.

However, if we take cosets of group $\{0, 4, 8\}$, i.e. $\{0, 4, 8\}$, $\{1, 5, 9\}$, $\{2, 6, 10\}$, $\{3, 7, 11\}$, addition is now well-defined (e.g. adding elements from the second coset to the third coset always yields the fourth)

Proposition. Let $K \lhd G$. Then the set of (left) cosets of K in G is a group under the operation $aK * bK = (ab)K$.

Proof. First show that the operation is well-defined. If $aK = a'K$ and $bK = b'K$, we want to show that $aK * bK = a'K * b'K$. We know that $a' = ak_1$ and $b' = bk_2$ for some $k_1, k_2 \in K$. Then $a'b' = ak_1bk_2$. We know that $bk_1b^{-1} \in K$. Let $bk_1b^{-1} = k_3$. Then $k_1b = bk_3$. So $a'b' = abk_3k_2 \in (ab)K$. So picking a different representative of the coset gives the same product.

1. (Closure) If aK, bK are cosets, then $(ab)K$ is also a coset
2. (Identity) The identity is $eK = K$ (clear from definition)
3. (Inverse) The inverse of aK is $a^{-1}K$ (clear from definition)
4. (Associativity) Follows from the associativity of G.

□

Definition (Quotient group)**.** Given a group G and a normal subgroup K, the *quotient group* or *factor group* of G by K, written as G/K, is the set of (left) cosets of K in G under the operation $aK * bK = (ab)K$.

Note: Quotient groups are *not* subgroups of G. They contain different kinds of elements. It might be even not isomorphic to any subgroup of G, e.g. $\mathbb{Z}/n\mathbb{Z} \cong C_n$ are finite, but all subgroups of $\mathbb{Z}$ infinite.

Example. Consider the following:

(i) Take $G = \mathbb{Z}$ and $n\mathbb{Z}$ (which must be normal since G is abelian), the cosets are $k + n\mathbb{Z}$ for $0 \le k < n$. The quotient group is $\mathbb{Z}_n$. So we can write $\mathbb{Z}/(n\mathbb{Z}) = \mathbb{Z}_n$. In fact these are the only quotient groups of $\mathbb{Z}$ since $n\mathbb{Z}$ are the only subgroups.

(ii) Take $K = \langle r \rangle \lhd D_6$. We have two cosets K and sK. So D_6/K has order 2 and is isomorphic to C_2.

(iii) Take $K = \langle r^2 \rangle \lhd D_8$. We know that G/K should have $\frac{8}{2} = 4$ elements. We have $G/K = \{K, rK = r^3K, sK = sr^2K, srK = sr^3K\}$. We see that all elements (except K) has order 2, so $G/K \cong C_2 \times C_2$.

Note: If G is abelian, G/K is also abelian.
Note: The *set* of left cosets also exists for non-normal subgroups (abnormal subgroups?), but the group operation above is not well defined.

Example. (Non-example) Consider D_6 with $H = \langle s\rangle$. H is not a normal subgroup. We have $rH * r^2H = r^3H = H$, but $rH = rsH$ and $r^2H = srH$ (by considering the individual elements). So we have $rsH * srH = r^2H \neq H$, and the operation is not well-defined.

Lemma. Given $K \lhd G$, the *quotient map* $q : G \to G/K$ with $g \mapsto gK$ is a surjective group homomorphism.

Proof. $q(ab) = (ab)K = aKbK = q(a)q(b)$. So q is a group homomorphism. Also for all $aK \in G/K$, $q(a) = aK$. So it is surjective. □

Note: The kernel of the quotient map is K itself. So any normal subgroup is a kernel of some function.

Proposition. The quotient of a cyclic group is cyclic.

Proof. Let $G = C_n$ with $H \leq C_n$. We know that H is also cyclic. Say $C_n = \langle c\rangle$ and $H = \langle c^k\rangle \cong C_\ell$, where $k\ell = n$. We have $C_n/H = \{H, cH, c^2H, \cdots C^{k-1}H\} = \langle cH\rangle \cong C_k$. □

4.3 The Isomorphism Theorem

We saw that all kernels are normal subgroups and all normal subgroups are kernels of its quotient map. So normal subgroups are exactly the kernels of group homomorphisms. In fact,

Theorem (The Isomorphism Theorem)**.** Let $f : G \to H$ be a group homomorphism with kernel K. Then $K \lhd G$ and $G/K \cong \operatorname{im} f$.

Proof. We have proved that $K \lhd G$ before. We define a group homomorphism $\bar f : G/K \to \operatorname{im} f$ by $\bar f(aK) = f(a)$.

First check that this is well-defined: If $a_1K = a_2K$, then $a_2^{-1}a_1 \in K$. So $f(a_2)^{-1}f(a_1) = f(a_2^{-1}a_1) = e$. So $f(a_1) = f(a_2)$ and $\bar f(a_1K) = \bar f(a_2K)$.

Now we check that it is a group homomorphism: $\bar f(aKbK) = \bar f(abK) = f(ab) = f(a)f(b) = \bar f(aK)\bar f(bK)$.

To show that it is injective, we have $\bar f(aK) = \bar f(bK) \Rightarrow f(a) = f(b) \Rightarrow f(b)^{-1}f(a) = e \Rightarrow b^{-1}a \in K \Rightarrow aK = bK$.

By definition, $\bar f$ is surjective since $\operatorname{im}\bar f = \operatorname{im} f$. So $\bar f$ gives an isomorphism $G/K \cong \operatorname{im} f \leq H$. □

Note: If f is injective, then the kernel is $\{e\}$, so $G/K \cong G$ and G is isomorphic to a subgroup of H. We can think of f as an inclusion map. *Note*: If f is surjective, then $\operatorname{im} f = H$. In this case, $G/K \cong H$.

Example.

(i) Take $f : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$ with $A \mapsto \det A$, $\ker f = \mathrm{SL}_N(\mathbb{R})$. $\operatorname{im} f = \mathbb{R}*$ as for all $\lambda \in \mathbb{R}^*$, $\det \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{pmatrix} = \lambda$. So we know that $\mathrm{GL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{R}) \cong \mathbb{R}^*$.

(ii) Define $\theta : (\mathbb{R}, +) \to (\mathbb{C}*, \times)$ with $r \mapsto \exp(2\pi ir)$. This is a group homomorphism since $\theta(r + s) = \exp(2\pi i(r + s)) = \exp(2\pi ir)\exp(2\pi is) = \theta(r)\theta(s)$. We know that the kernel is $\mathbb{Z} \lhd \mathbb{R}$. Clearly the image is the unit circle $(S_1, \times)$. So $\mathbb{R}/\mathbb{Z} \cong (S_1, \times)$.

(iii) $G = (\mathbb{Z}_p^*, \times)$ for prime $p \neq 2$. We have $f : G \to G$ with $a \mapsto a^2$. This is a homomorphism since $(ab)^2 = a^2b^2$ ($\mathbb{Z}_p^*$ is abelian). The kernel is $\{\pm 1\} = \{1, p - 1\}$. We know that $\operatorname{im} f \cong G/\ker f$ with order $\frac{p-1}{2}$. This are known as quadratic residues.

Lemma. Any cyclic group is isomorphic to either $\mathbb{Z}$ or $\mathbb{Z}/(n\mathbb{Z})$ for some $n \in \mathbb{N}$.

Proof. Let $G = \langle c \rangle$. Define $f : \mathbb{Z} \to G$ with $m \mapsto c^m$. This is a group homomorphism since $c^{m_1+m_2} = c^{m_1}c^{m_2}$. f is surjective since G is by definition all c^m for all m. We know that $\ker f \lhd \mathbb{Z}$. We have three possibilities. Either

(i) $\ker f = \{e\}$, so F is an isomorphism and $G \cong \mathbb{Z}$; or

(ii) $\ker f = \mathbb{Z}$, then $G \cong \mathbb{Z}/\mathbb{Z} = \{e\} = C_1$; or

(iii) $\ker f = n\mathbb{Z}$ (since these are the only proper subgroups of $\mathbb{Z}$), then $G \cong \mathbb{Z}/(n\mathbb{Z})$

□

Definition (Simple group)**.** A group is *simple* if it has no non-trivial proper normal subgroup, i.e. only $\{e\}$ and G are normal subgroups.

Example. C_p for prime p are simple groups since it has no proper subgroups at all, let alone normal ones. A_5 is simple, which we will prove after Chapter 6.

The finite simple groups are the building blocks of all finite groups. All finite simple groups have been classified (The Atlas of Finite Groups). If we have $K \lhd G$ with $K \neq G$ or $\{e\}$, then we can "quotient out" G into G/K. If G/K is not simple, repeat. Then we can write G as an "inverse quotient" of simple groups.

5 Group actions

5.1 Group acting on sets

Recall the symmetries of an n-gon, or $\operatorname{Sym} X$, the group elements "do something" to the set of vertices or the set X.

Definition (Group action)**.** Let X be a set and G be a group. An *action* of G on X is a function $\theta : G \times X \to X$ satisfying

0. $\forall g \in G, x \in X[\theta(g, x) \in X]$.
1. $\forall x \in X[\theta(e, x) = x]$.
2. $\forall g, h \in G, x \in X[\theta(g, \theta(h, x)) = \theta(gh, x)]$

i.e. given an element $g \in G$ and an $x \in X$, g "acts on" x to give an element $\theta(g, x) \in X$ (the two conditions ensure that the group properties of G are not destroyed)

Notation. We can write $\theta(g, x) = g \cdot x = g(x)$. Given any fixed g, we can define $\theta_g : X \to X$ with $x \mapsto \theta(g, x)$.

Example. We can have the following actions:

(i) Trivial action: for any group G acting on any set X, we can have $\theta(g, x) = x$ for all g, x, i.e. G does nothing

(ii) S_n acts on $\{1, \cdots n\}$ by permutation

(iii) D_{2n} acts on the vertices of a regular n-gon (or the set $\{1, \cdots, n\}$)

(iv) The rotations of a cube act on the faces/vertices/diagonals/axes of the cube.

Note: Different groups can act on the same sets, and the same group can act on different sets.

Lemma. For each $g \in G$, $\theta_g : X \to X$ is a bijection.

Proof. $\theta_{g^{-1}}$ is its inverse. $\square$

Proposition. Let G be a group and X a set. Then $\theta : G \times X \to X$ with $\theta(g, x) = \theta_g(x)$ is an action if and only if $\varphi : G \to \operatorname{Sym} X$ with $\varphi(g) = \theta_g$ is a group homomorphism.

Proof. ($\Rightarrow$) Given θ, we know that θ_g is a bijection and thus a permutation of X. We have to check that $\varphi(gh) = \varphi(g)\varphi(h)$. We have $\varphi(gh) = \theta_{gh} = \theta_g \circ \theta_h = \varphi(g)\varphi(h)$. So φ is a group homomorphism.

($\Leftarrow$) Given a homomorphism $\varphi : G \to \operatorname{Sym} X$, we define $\theta : G \times X \to X$ by $\theta(g, x) = \varphi(g)(x)$. We have to show that the resulting θ is an action:

0. As $\varphi(g) = \theta_g \in \operatorname{Sym} X$, $\theta(g, x) = \theta_g(x) \in X$
1. Since φ is a homomorphism, we know that $\theta_e = \varphi(e) = 1_x$. So $\theta_e(x) = x$.
2. $\varphi(gh) = \varphi(g) \circ \varphi(h)$. So $\theta_{gh} = \theta_g \circ \theta_h$

□

Definition (Kernel of action)**.** The *kernel* of an action G on X is the kernel of φ, i.e. all g such that $\theta_g = 1_X$.

Note: $\ker\varphi \lhd G$ and G/K is isomorphic to a subgroup of $\operatorname{Sym} X$.

Example. Consider the following actions:

(i) D_{2n} acting on $\{1, 2 \cdots n\}$ gives $\varphi : D_{2n} \to S_n$ with kernel $\{e\}$.

(ii) Let G be the rotations of a cube and let it act on the three axes x, y, z through the faces. We have $\phi : G \to S_3$. Then any rotation by 180° doesn't change the axes, i.e. act as the identity. So the kernel of the action has at least 4 elements: e and the three 180° rotations. In fact, we'll see later that these 4 are exactly the kernel.

Definition (Faithful action)**.** An action is *faithful* if the kernel is just $\{e\}$.

5.2 Orbits and Stabilizers

Definition (Orbit of action)**.** Given an action G on X, the *orbit* of an element $x \in X$ is

$$\operatorname{orb}(x) = G(x) = \{y \in X : \exists g \in G(g(x) = y)\}.$$

Intuitively, it is the elements that x can possibly get mapped to.

Definition (Stabilizer of action)**.** The *stabilizer* of x is

$$\operatorname{stab}(x) = G_x = \{g \in G : g(x) = x\} \subseteq G.$$

Intuitively, it is the elements in G that do not change x.

Lemma. $\operatorname{stab}(x)$ is a subgroup of G.

Proof. We know that $e(x) = x$ by definition. So $\operatorname{stab}(x)$ is non-empty. Suppose $g, h \in \operatorname{stab}(x)$, then $gh^{-1}(x) = g(h^{-1}(x)) = g(x) = x$. So $gh^{-1} \in \operatorname{stab}(X)$. So $\operatorname{stab}(x)$ is a subgroup. □

Example. Consider the following:

(i) Consider D_8 acting on the corners of the square $X = \{1, 2, 3, 4\}$. Then $\operatorname{orb}(1) = X$ since 1 can go anywhere by rotations. $\operatorname{stab}(1) = \{e$, reflection in the line through $1\}$

(ii) Consider the rotations of a cube acting on the three axes x, y, z. Then $\operatorname{orb}(x)$ is everything, and $\operatorname{stab}(x)$ is e, 180° rotations and rotations about the x axis.

Definition (Transitive action)**.** An action G on X is *transitive* if $\forall x(\operatorname{orb}(x) = X)$, i.e. you can reach any element from any element.

Lemma. The orbits of an action partition X.

Proof. Firstly, $\forall x(x \in \operatorname{orb}(x))$ as $e(x) = x$. So every x is in some orbit.

Then suppose $z \in \operatorname{orb}(x)$ and $z \in \operatorname{orb}(y)$, we have to show that $\operatorname{orb}(x) = \operatorname{orb}(y)$. We know that $z = g_1(x)$ and $z = g_2(y)$ for some g_1, g_2. Then $g_1(x) = g_2(y)$ and $y = g_2^{-1}g_1(x)$.

For any $w = g_3(y) \in \operatorname{orb}(y)$, we have $w = g_3g_2^{-1}g_1(x)$. So $w \in \operatorname{orb}(x)$. Thus $\operatorname{orb}(y) \subseteq \operatorname{orb}(x)$ and similarly $\operatorname{orb}(x) \subseteq \operatorname{orb}(y)$. Therefore $\operatorname{orb}(x) = \operatorname{orb}(y)$. □

Theorem (Orbit-stabilizer theorem)**.** Let the finite group G act on X. For any $x \in X$,

$$|\operatorname{orb}(x)||\operatorname{stab}(x)| = |G|.$$

Proof. We know that $\operatorname{stab}(x)$ is a subgroup. For any $g(x) = y \in \operatorname{orb}(x)$, we want to show $\{h \in G : h(x) = y\} = g\operatorname{stab}(x)$. We have $h(x) = y = g(x) \Leftrightarrow g^{-1}h(x) = x \Leftrightarrow g^{-1}h \in \operatorname{stab} x \Leftrightarrow h \in g \operatorname{stab} x$.

For each element in the orbit, we can obtain exactly one coset as above. So $|G : \operatorname{stab}(x)| = |\operatorname{orb}(x)|$ and $|G| = |\operatorname{orb}(x)||\operatorname{stab}(x)|$ by Lagrange's theorem. □

Example. Consider the following:

(i) Suppose we want to know how big D_{2n} is. D_{2n} acts on the vertices $\{1, 2, 3, \cdots, n\}$ transitively. So $|\operatorname{orb}(1)| = n$. Also, $\operatorname{stab}(1) = \{e$, reflection in the line through $1\}$. So $|D_{2n}| = |\operatorname{orb}(1)||\operatorname{stab}(1)| = 2n$.

Note: If the action is transitive, then all orbits have size $|X|$ and thus all stabilizers have the same size.

(ii) Let $\langle(1\ 2)\rangle$ act on $\{1, 2, 3\}$. Then $\operatorname{orb}(1) = \{1, 2\}$ and $\operatorname{stab}(1) = \{e\}$. $\operatorname{orb}(3) = \{3\}$ and $\operatorname{stab}(3) = \langle(1\ 2)\rangle$.

(iii) Consider S_4 acting on $\{1, 2, 3, 4\}$. We know that $\operatorname{orb}(1) = X$ and $|S_4| = 24$. So $|\operatorname{stab}(1)| = \frac{24}{4} = 6$. That makes it easier to find $\operatorname{stab}(1)$. Clearly $S_{\{2,3,4\}} \cong S_3$ fix 1. So $S_{\{2,3,4\}} \leq \operatorname{stab}(1)$. However, $|S_3| = 6 = |\operatorname{stab}(1)|$, so this is all of the stabilizer.

5.3 Important actions

Lemma. (Left regular action) Any group G acts on itself by left multiplication. This action is faithful and transitive.

Proof. We have

1. $\forall g \in G, x \in G(g(x) = g * x \in G)$ by definition of a group.
2. $\forall x \in G(e \cdot x = x)$ by definition of a group.
3. $g(hx) = (gh)x$ by associativity.

So it is an action.

To show that it is faithful, we want to know that $[\forall x \in G(gx = x)] \Rightarrow g = e$. This follows directly from the uniqueness of identity.

To show that it is transitive, $\forall x, y \in G$, then $(yx^{-1})(x) = y$. So any x can be sent to any y. □

Theorem (Cayley's theorem)**.** Every group is isomorphic to some subgroup of some symmetric group.

Proof. Take the left regular action of G on itself. This gives a group homomorphism $\varphi : G \to \operatorname{Sym} G$ with $\ker \varphi = \{e\}$ as the action is faithful. By the isomorphism theorem, $G \cong \operatorname{im} \varphi \leq \operatorname{Sym} G$. $\square$

Lemma (Left coset action)**.** Let $H \leq G$. Then G acts on the left cosets of H by left multiplication transitively.

Proof. First show that it is an action:

0. $g(aH) = (ga)H$ is a coset of H.
1. $e(aH) = (ea)H = aH$. So e is the identity.
2. $g_1(g_2(aH)) = g_1((g_2a)H) = (g_1g_2a)H = (g_1g_2)(aH)$.

To show that it is transitive, given aH, bH, we know that $(ba^{-1})(aH) = bH$. So any aH can be mapped to bH. $\square$

Note: If $H = \{e\}$, then this is just the left regular action since $G/\{e\} \cong G$.

Definition (Conjugation of element)**.** The *conjugation* of $a \in G$ by $b \in G$ is given by $bab^{-1} \in G$.

Lemma (Conjugation action)**.** Any group G acts on itself by conjugation (i.e. $g(x) = gxg^{-1}$).

Proof. To show that this is an action, we have

0. $g(x) = gxg^{-1} \in G$ for all $g, x \in G$.
1. $e(x) = exe^{-1} = x$
2. $g(h(x)) = g(hxh^{-1}) = ghxh^{-1}g^{-1} = (gh)x(gh)^{-1} = (gh)(x)$

$\square$

Definition (Center of group)**.** The *center* of G is the elements that commute with all other elements.

$$Z(g) = \{g \in G : \forall a(gag^{-1} = a)\} = \{g \in G : \forall a(ga = ag)\}.$$

It is sometimes written as $C(G)$ instead of $Z(G)$.

Definition (Conjugacy classes and centralizers)**.** The *conjugacy classes* are the orbits of the conjugacy action.

$$\operatorname{ccl}(a) = \{b \in G : \exists g \in g(gag^{-1} = b\}.$$

The *centralizers* are the stabilizers of this action, i.e. elements that commute with a.

$$C_G(a) = \{g \in G : gag^{-1} = a\} = \{g \in G : ga = ag\}.$$

Note: $\bigcap_{g \in G} C_G(g) = Z(g)$
Note: Conjugate elements have many of the same properties, e.g. order. They are basically "the same thing looked in different ways".

Lemma. Let $K \lhd G$. Then G acts by conjugation on K.

Proof. We only have to prove closure as the other properties follow from the conjugation action. However, by definition of a normal subgroup, $\forall g \in G, k \in K(gkg^{-1} \in K)$. So it is closed. $\square$

Proposition. Normal subgroups are exactly those subgroups which are unions of conjugacy classes.

Proof. Let $K \lhd G$. If $k \in K$, so is $\forall g(gkg^{-1} \in K)$. So $\text{ccl}(k) \subseteq K$. So K is the union of the conjugacy classes of all its elements.

Conversely, if K is a union of conjugacy classes and a subgroup of G, then $\forall k \in K, g \in G(gkg^{-1} \in K)$. So K is normal. $\square$

Lemma. Let X be the set of subgroups of G. Then G acts by conjugation on X.

Proof. To show that it is an action, we have

0. If $H \leq G$, then we have to show that gHg^{-1} is also a subgroup. We know that $e \in H$ and thus $geg^{-1} = e \in gHg^{-1}$, so gHg^{-1} is non-empty. For any two elements gag^{-1} and $gbg^{-1} \in gHg^{-1}$, $(gag{-}1)(gbg^{-1})^{-1} = g(ab^{-1})g^{-1} \in gHg^{-1}$. So gHg^{-1} is a subgroup.
1. $eHe^{-1} = H$.
2. $g_1(g_2Hg_2^{-1})g_1^{-1} = (g_1g_2)H(g_1g_2)^{-1}$.

$\square$

Note: Normal subgroups have singleton orbits.
Note: Conjugate subgroups have the same size (c.f. "the same"ness)

Definition (Normalizer of subgroup)**.** The *normalizer* of a subgroup is the stabilizer of the (group) conjugation action.

$$N_G(H) = \{g \in G : gHg^{-1} = H\}.$$

Note: Clearly $H \subseteq N_G(H)$.

Proposition. $N_G(H)$ is the largest subgroup of G in which H is a normal subgroup.

There is a connection between actions in general and conjugation of subgroups.

Lemma. Stabilizers of the elements in the same orbit are conjugate. Let G act on X and let $g \in G, x \in X$. Then $\text{stab}(g(x)) = g\,\text{stab}(x)g^{-1}$

5.4 Applications

Example. Let G^+ be the rotations of a cube acting on the vertices. Let X be the set of vertices. $|X| = 8$. The action is transitive, so the orbit of element is the whole of X. The stabilizer of vertex 1 is the set of rotations through 1 and the diagonally opposite vertex, of which there are 3. Then $|G^+| = |\text{orb}(1)||\text{stab}(1)| = 8 \cdot 3 = 24$.

Example. Let G be a finite group simple group of order greater than 2, and $H \leq G$ have index $n \neq 1$. Then $n!/2 \geq |G|$.

Proof. Consider the left coset action of G on H. We get a group homomorphism $\varphi : G \to S_n$ since there are n cosets of H. Since $H \neq G$, φ is non-trivial and $\ker\varphi \neq G$. Now $\ker\varphi \lhd G$. Since G is simple, $\ker\varphi = \{e\}$. So $G \cong \operatorname{im}\varphi \subseteq S_n$ by the isomorphism theorem. So $|G| \leq |S_n| = n!$.

We can further refine this by considering $\operatorname{sgn}\circ\varphi : G \to \{\pm 1\}$. The kernel of this composite is normal in G. So $K = \ker(\operatorname{sgn}\circ\phi) = \{e\}$ or G. Since $G/K \cong \operatorname{im}(\operatorname{sgn}\circ\phi)$, we know that $|G|/|K| = 1$ or 2. For $|G| > 2$, we cannot have $K = \{e\}$. So we must have $K = G$, so $\operatorname{sgn}(\varphi(g)) = 1$ for all g and $\operatorname{im}\varphi \leq A_n$. So $|G| \geq n!/2$ □

We have seen on Sheet 1 that if $|G|$ is even, then G has an element of order 2. In fact,

Theorem (Cauchy's Theorem)**.** Let G be a finite group and prime p dividing $|G|$. Then G has an element of order p. (In fact there must be at least $p - 1$ elements of order p)
Note: By Lagrange's theorem, if p doesn't divide G, then G cannot have an element of order p. However, A_4 doesn't have an element of order 6 even though $6|12 = |A_4|$, so Cauchy's theorem only hold for primes.

Proof. Let G and p be fixed. Consider $G^p = G \times G \times \cdots \times G$, the set of p-tuples of G. Let $X \subseteq G^p$ be $X = \{(a_1, a_2, \cdots, a_p) \in G^p : a_1 a_2 \cdots a_p = e\}$.

In particular, if an element b has order p, then $(b, b, \cdots, b) \in X$. In fact, if $(b, b, \cdots, b) \in X$ and $b \neq e$, then b has order p.

Now let $H = \langle h : h^p = e\rangle \cong C_p$ be a cyclic group of order p with generator h (This h is not related to G in any way). Let H act on X by "rotation":

$$h(a_1, a_2, \cdots, a_p) = (a_2, a_3, \cdots, a_p, a_1)$$

This is an action:

0. If $a_1 \cdots a_p = e$, then $a^{-1} = a_2 \cdots a_p$. So $a_2 \cdots a_p a_1 = a_1^{-1} a_1 = e$. So $(a_2, a_3, \cdots, a_p, a_1) \in X$.
1. e acts as an identity by construction
2. The "associativity" condition also works by construction.

As orbits partition X, the sum of all orbit sizes must be $|X|$. We know that $|X| = |G|^{p-1}$ since we can freely choose the first $p-1$ entries and the last one must be the inverse of their product. Since p divides $|G|$, p also divides $|X|$. We have $|\operatorname{orb}(a_1, \cdots, a_p)||\operatorname{stab}_H(a_1, \cdots, A_p)| = |H| = p$. So all orbits have size 1 or p, and they sum to $|X| = p\times$ something. We know that there is one orbit of size 1, namely $(e, e, \cdots, e)$. So there must be at least $p - 1$ other orbits of size 1 for the sum to be divisible by p.

In order to have an orbit of size 1, they must look like $(a, a, \cdots, a)$. for some $a \in G$, which has order p. □

6 Symmetric groups II

6.1 Conjugacy classes in S_n

Recall $\sigma, \tau \in S_n$ are conjugate if $\exists \rho \in S_n$ such that $\rho\sigma\rho^{-1} = \tau$.

Proposition. If $(a_1\ a_2\ \cdots\ a_k)$ is a k-cycle and $\rho \in S_n$, then $\rho(a_1\ \cdots\ a_k)\rho^{-1}$ is the k-cycle $(\rho(a_1)\ \rho(a_2)\ \cdots \rho(a_3))$

Proof. Consider any $\rho(a_1)$ acted on by $\rho(a_1\ \cdots\ a_k)\rho^{-1}$. The three permutations send it to $\rho(a_1) \mapsto a_1 \mapsto a_2 \mapsto \rho(a_2)$ and similarly for other a_is. Since ρ is bijective, any b can be written as $\rho(a)$ for some a. □

Corollary. Two elements in S_n are conjugate iff they have the same cycle type.

Proof. Suppose $\sigma = \sigma_1\sigma_2\cdots\sigma_\ell$, where σ_i are disjoint cycles. Then $\rho\sigma\rho^{-1} = \rho\sigma_1\rho^{-1}\rho\sigma_2\rho^{-1}\cdots\rho\sigma_1\rho^{-1}$. Since the conjugation of a cycle conserves its length, $\rho\sigma\rho^{-1}$ has the same cycle type.

Conversely, if σ, τ have the same cycle type, say
$\sigma = (a_1\ a_2\ \cdots\ a_k)(a_{k+1}\ \cdots\ a_{k+\ell})$ and $\tau = (b_1\ b_2\ \cdots\ b_k)(b_{k+1}\ \cdots\ b_{k+\ell})$. If we let $\rho(a_i) = b_i$, then $\rho\sigma\rho^{-1} = \tau$. □

Example. Conjugacy classes of S_4:

Cycle type	Example element	Size of ccl	Size of centralizer	Sign
(1, 1, 1, 1)	e	1	24	+1
(2, 1, 1)	(1 2)	6	4	−1
(2, 2)	(1 2)(3 4)	3	8	+1
(3, 1)	(1 2 3)	8	3	+1
(4)	(1 2 3 4)	6	4	−1

We know that a normal subgroup is a union of conjugacy classes. We can now find all normal subgroups by finding possible union of conjugacy classes whose cardinality divides 24. Note that the normal subgroup must contain e.

(i) Order 1: $\{e\}$

(ii) Order 2: None

(iii) Order 3: None

(iv) Order 4: $\{e, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\} \cong C_2 \times C_2 = V_4$ is a possible candidate. We can check the group axioms and find that it is really a subgroup

(v) Order 6: None

(vi) Order 8: None

(vii) Order 12: A_4 (We know it is a normal subgroup since it is the kernel of the signature and/or it has index 2)

(viii) Order 24: S_4

We can also obtain the quotients of S_4: $S_4/\{e\} \cong S_4$, $S_4/V_4 \cong S_3 \cong D_6$, $S_4/A_4 \cong C_2$, $S_4/S_4 = \{e\}$.

6.2 Conjugacy classes in A_n

We have seen that $|S_n| = 2|A_n|$ and that conjugacy classes in S_n are "nice". How about in A_n?

The first thought is that we write it down:

$$\operatorname{ccl}_{S_n}(\sigma) = \{\tau \in S_n : \exists rho \in S_n(\tau = \rho\sigma\rho^{-1})\}$$
$$\operatorname{ccl}_{A_n}(\sigma) = \{\tau \in S_n : \exists rho \in A_n(\tau = \rho\sigma\rho^{-1})\}$$

Obviously $\operatorname{ccl}_{A_n}(\sigma) \subseteq \operatorname{ccl}_{S_n}(\sigma)$, but the converse need not be true since the conjugation need to map σ to τ may be odd.

Example. Consider (1 2 3) and (1 3 2). They are conjugate in S_3 by (2 3), but (2 3) $\notin A_3$. (This does not automatically entail that they are not conjugate in A_3 because there might be another even permutation that conjugate (1 2 3) and (1 3 2). In A_5, (2 3)(4 5) works (but not in A_3))

We can use the orbit-stabilizer theorem:

$$|S_n| = |\operatorname{ccl}_{S_n}(\sigma)||C_{S_n}(\sigma)|$$
$$|A_n| = |\operatorname{ccl}_{A_n}(\sigma)||C_{A_n}(\sigma)|$$

We know that A_n is half of S_n and ccl_{A_n} is contained in ccl_{S_n}. So we have two options: either $\operatorname{ccl}_{S_n}(\sigma) = \operatorname{ccl}_{A_n}(\sigma)$ and $|C_{S_n}(\sigma)| = \frac{1}{2}|C_{A_n}(\sigma)|$; or $\frac{1}{2}|\operatorname{ccl}_{S_n}(\sigma)| = |\operatorname{ccl}_{A_n}(\sigma)|$ and $C_{A_n}(\sigma) = C_{S_n}(\sigma)$.

Definition (Splitting of conjugacy classes)**.** When $|\operatorname{ccl}_{A_n}(\sigma)| = \frac{1}{2}|\operatorname{ccl}_{S_n}(\sigma)|$, we say that the conjugacy class of σ *splits* in A_n.

So the conjugacy classes are either retained or split.

Proposition. For $\sigma \in A_n$, the conjugacy class of σ splits in A_n if and only if no odd permutation commutes with σ.

Proof. We have the conjugacy classes splitting if and only if the centralizer does not. So instead we check whether the centralizer splits. Clearly $C_{A_n}(\sigma) = C_{S_n}(\sigma) \cap A_n$. So splitting of centralizer occurs if and only if an odd permutation commutes with σ. □

Example. Conjugacy classes in A_4:

Cycle type	Example	$\lvert\operatorname{ccl}_{S_4}\rvert$	Odd element in C_{S_4}?	$\lvert\operatorname{ccl}_{A_4}\rvert$
(1, 1, 1, 1)	e	1	Yes (e.g. (1 2))	1
(2, 2)	(1 2)(3 4)	3	Yes (e.g. (1 2))	3
(3, 1)	(1 2 3)	8	No	4, 4

In the (3, 1) case, by the orbit stabilizer theorem, $|C_{S_4}((1\ 2\ 3))| = 3$, which is odd and cannot split.

Example. Conjugacy classes in A_5:

Cycle type	Example	$\lvert\operatorname{ccl}_{S_5}\rvert$	Odd element in C_{S_5}?	$\lvert\operatorname{ccl}_{A_5}\rvert$
(1, 1, 1, 1, 1)	e	1	Yes (e.g. (1 2))	1
(2, 2, 1)	(1 2)(3 4)	15	Yes (1 2)	15
(3, 1, 1)	(1 2 3)	20	Yes (4 5)	20
(5)	(1 2 3 4 5)	24	No	12, 12

n Since

the centralizer of (1 2 3 4 5) has size 5, it cannot split, so its conjugacy class must split.

Lemma. $\sigma = (1\ 2\ 3\ 4\ 5) \in S_5$ has $C_{S_5}(\sigma) = \langle\sigma\rangle$.

Proof. $|\operatorname{ccl}_{S_n}(\sigma)| = 24$ and $|S_5| = 120$. So $|C_{S_5}(\sigma)| = 5$. Clearly $\langle\sigma\rangle \subseteq C_{S_5}(\sigma)$. Since they both have size 5, we know that $C_{S_5}(\sigma) = \langle\sigma\rangle$ □

Theorem. A_5 is simple.

Proof. We know that normal subgroups must be unions of the conjugacy classes, must contain e and their order must divide 60. The possible orders are 1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30. However, the conjugacy classes 1, 15, 20, 12, 12 cannot add up to any of the possible orders apart from 1 and 60. So we only have trivial normal subgroups. □

In fact, all A_n for $n \geq 5$ are simple (c.f. Group, Rings and Modules).

7 Quaternions

Definition (Quaternions)**.** The *quaternions* is the set of matrices

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \begin{pmatrix}0 & i\\ i & 0\end{pmatrix}$$
$$\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}, \begin{pmatrix}-i & 0\\ 0 & i\end{pmatrix}, \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & -i\\ -i & 0\end{pmatrix}$$

which is a subgroup of $\mathrm{GL}_2(\mathbb{C})$.

Notation. We can also write the quaternions as

$$Q_8 = \langle a, b : a^4 = e, b^2 = a^2, bab^{-1} = a^{-1}\rangle$$

Even better, we can write

$$Q_8 = \{1, -1, i, -i, j, -j, k, -k\}$$

with

(i) $(-1)^2 = 1$

(ii) $i^2 = j^2 = k^2 = -1$

(iii) $(-1)i = -i$ etc.

(iv) $ij = k$, $jk = i$, $ki = j$

(v) $ji = -k$, $kj = -i$, $ik = -j$

We have

$$1 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, i = \begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}, j = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, k = \begin{pmatrix}0 & i\\ i & 0\end{pmatrix}$$
$$-1 = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}, -i = \begin{pmatrix}-i & 0\\ 0 & i\end{pmatrix}, -j = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, -k = \begin{pmatrix}0 & -i\\ -i & 0\end{pmatrix}$$

Lemma. If G has order 8, then either G is abelian (i.e., $\cong C_8, C_4 \times C_2$ or $C_2 \times C_2 \times C_2$), or G is not abelian and isomorphic to D_8 or Q_8 (dihedral or quaternion).

Proof. Consider the different possible cases:

- If G contains an element of order 8, then $G \cong C_8$.
- If all non-identity elements have order 2, then G is abelian (Sheet 1, Q8). Let $a \neq b \in G \setminus \{e\}$. By the direct product theorem, $\langle a, b\rangle = \langle a\rangle \times \langle b\rangle$. Then take $c \notin \langle a, b\rangle$. By the direct product theorem, we obtain $\langle a, b, c\rangle = \langle a\rangle \times \langle b\rangle \times \langle c\rangle = C_2 \times C_2 \times C_2$. Since $\langle a, b, c\rangle \subseteq G$ and $|\langle a, b, c\rangle| = |G|$, $G = \langle a, b, c\rangle \cong C_2 \times C_2 \times C_2$.

– G has no element of order 8 but has an order-4 element $a \in G$. Let $H = \langle a \rangle$. Since H has index 2, it is normal in G. So $G/H \cong C_2$ since $|G/H| = 2$. This means that for any $b \notin H$, bH generates G/H. Then $(bH)^2 = b^2 H = H$. So $b^2 \in H$. Since $b^2 \in \langle a \rangle$ and $\langle a \rangle$ is a cyclic group, b^2 commutes with a.

  If $b^2 = a$ or a^3, then b has order 8. Contradiction. So $b^2 = e$ or a^2.

  We also know that H is normal, so $bab^{-1} \in H$. Let $bab^{-1} = a^\ell$. Since a and b^2 commute, we know that $a = b^2ab^{-2} = b(bab^{-1})b^{-1} = ba^\ell b^{-1} = (bab^{-1})^\ell = a^{\ell^2}$. So $\ell^2 \equiv 1 \pmod 4$. So $\ell \equiv \pm 1 \pmod 4$.

  ◦ When $l \equiv 1 \pmod 4$, $bab^{-1} = a$, i.e. $ba = ab$. So G is abelian.
    * If $b^2 = e$, then $G = \langle a, b \rangle \cong \langle a \rangle \times \langle b \rangle \cong C_4 \times C_2$.
    * If $b^2 = a^2$, then $(ba^{-1})^2 = e$. So $G = \langle a, ba^{-1} \rangle \cong C_4 \times C_2$.

  ◦ If $l \equiv -1 \pmod 4$, then $bab^{-1} = a^{-1}$.
    * If $b^2 = e$, then $G = \langle a, b : a^4 = e = b^2, bab^{-1} = a^{-1} \rangle$. So $G \cong D_8$ by definition.
    * If $b^2 = a^2$, then we have $G \cong Q_8$.

$\square$

8 Matrix groups

8.1 General and special linear groups

Consider $M_{n\times n}(F)$, i.e. the set of $n \times n$ matrices over the field $F = \mathbb{R}$ or $\mathbb{C}$. We know that matrix multiplication is associative (since they represent functions) but are, in general, not commutative. To make this a group, we want the identity matrix I to be the identity. To ensure everything has an inverse, we can only include invertible matrices.
Note: We do not necessarily need to take I as the identity of the group. We can, for example, take $e = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and obtain a group in which every matrix is of the form $\begin{pmatrix} 0 & 0 \\ 0 & a \end{pmatrix}$ for some non-zero a. This forms a group, albeit a boring one (it is simply $\cong \mathbb{R}^*$).

Definition (General linear group $\mathrm{GL}_n(F)$)**.**

$$\mathrm{GL}_n(F) = \{A \in M_{n\times n}(F) : A \text{ is invertible}\}$$

is the *general linear group.*

Note: We can alternatively define $\mathrm{GL}_n(F)$ as matrices with non-zero determinants.

Proposition. $\mathrm{GL}_n(F)$ is a group.

Proof. Identity is I, which is in $\mathrm{GL}_n(F)$ by definition (I is its self-inverse). The composition of invertible matrices is invertible, so is closed. Inverse exist by definition. Multiplication is associative. □

Proposition. $\det : \mathrm{GL}_n(F) \to F \setminus \{0\}$ is a surjective group homomorphism.

Proof. $\det AB = \det A \det B$. If A is invertible, it has non-zero determinant and $\det A \in F \setminus \{0\}$. For any $x \in F \setminus \{0\}$, then if we take the identity matrix and replace I_{11} with x, the determinant is x. So it is surjective. □

Definition (Special linear group $\mathrm{SL}_n(F)$)**.** The *special linear group* $\mathrm{SL}_n(F)$ is the kernel of the determinant, i.e.

$$SL_n(F) = \{A \in \mathrm{GL}_n(F) : \det A = 1\}.$$

So $\mathrm{SL}_n(F) \lhd \mathrm{GL}_n(F)$ as it is a kernel. Note that $Q_8 \leq \mathrm{SL}_2(\mathbb{C})$

8.2 Actions of $\mathrm{GL}_n(\mathbb{C})$

Proposition. $\mathrm{GL}_n(\mathbb{C})$ acts faithfully on $\mathbb{C}^n$ by left multiplication to the vector, with two orbits ($\mathbf{0}$ and everything else).

Proof. First show that it is a group action:

1. If $A \in \mathrm{GL}_n(\mathbb{C})$ and $\mathbf{v} \in \mathbb{C}^n$, then $A\mathbf{v} \in \mathbb{C}^n$. So it is closed.
2. $I\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in \mathbb{C}^n$.

3. $A(B\mathbf{v}) = (AB)\mathbf{v}$.

Now prove that it is faithful: a linear map is determined by what it does on a basis. Take the standard basis $\mathbf{e}_1 = (1, 0, \cdots, 0), \cdots \mathbf{e}_n = (0, \cdots, 1)$. Any matrix which maps each $\mathbf{e}_k$ to itself must be I (since the columns of a matrix are the images of the basis vectors)

To show that there are 2 orbits: Since $A\mathbf{0} = \mathbf{0}$ for all A. Also, as A is invertible, $A\mathbf{v} = \mathbf{0} \Leftrightarrow \mathbf{v} = \mathbf{0}$. So $\mathbf{0}$ forms a singleton orbit. Then given any two vectors $\mathbf{v} \neq \mathbf{w} \in \mathbb{C}^n \setminus \{0\}$, there is a matrix $A \in \mathrm{GL}_n(\mathbb{C})$ such that $A\mathbf{v} = \mathbf{w}$ (c.f. Vectors and Matrices). $\square$

Similarly, $\mathrm{GL}_n(\mathbb{R})$ acts on $\mathbb{R}^n$.

Proposition. $\mathrm{GL}_n(\mathbb{C})$ acts on $M_{n\times n}(\mathbb{C})$ by conjugation. (Proof is trivial)

This action can be thought of as a "change of basis" action. Two matrices are conjugate if they represent the same map but with respect to different bases. The P is the base change matrix.

From Vectors and Matrices, we know that there are three different types of orbits for $\mathrm{GL}_2(\mathbb{C})$: A is conjugate to a matrix of one of these forms:

(i) $\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$, with $\lambda \neq \mu$, i.e. two distinct eigenvalues

(ii) $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$, i.e. a repeated eigenvalue with 2-dimensional eigenspace

(iii) $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, i.e. a repeated eigenvalue with a 1-dimensional eigenspace

8.3 Orthogonal groups

Recall that A^T has entries $A^T_{ij} = A_{ji}$, i.e. reflect in the diagonal. They have the following properties:

(i) $(AB)^T = B^T A^T$

(ii) $(A^{-1})^T = (A^T)^{-1}$

(iii) $A^T A = I \Leftrightarrow AA^T = I \Leftrightarrow A^{-1} = A^T$. In this case A is *orthogonal*

(iv) $\det A^T = \det A$

Note: We are now in $\mathbb{R}$, because orthogonal matrices don't make sense in complex matrices

Definition (Orthogonal group $O(n)$)**.** The *orthogonal group* is

$$\mathrm{O}(n) = \mathrm{O}_n = \mathrm{O}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) : A^T A = I\},$$

i.e. the group of orthogonal matrices.

Proof. We have to check that it is a subgroup of $\mathrm{GL}_n(\mathbb{R})$: It is non-empty, since $I \in \mathrm{O}(n)$. If $A, B \in \mathrm{O}(n)$, then $(AB^{-1})(AB^{-1})^T = AB^{-1}(B^{-1})^T A^T = AB^{-1}BA^{-1} = I$, so $AB^{-1} \in \mathrm{O}(n)$ and this is indeed a subgroup. $\square$

Note that a matrix is orthogonal if the columns (or rows) form an orthonormal basis of $\mathbb{R}^n$: $AA^T = I \Leftrightarrow a_{ik}a_{jk} = \delta_{ij} \Leftrightarrow \mathbf{a_i} \cdot \mathbf{a_j} = \delta_{ij}$, where a_i is the ith column of A.

Proposition. $\det : \mathrm{O}(n) \to \{\pm 1\}$ is a surjective group homomorphism.

Proof. For $A \in \mathrm{O}(n)$, we know that $A^TA = I$. So $\det A^TA = (\det A)^2 = 1$. So $\det A = \pm 1$. Since $\det(AB) = \det A \det B$, it is a homomorphism. We have

$$\det I = 1, \quad \det \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix} = -1,$$

so it is surjective. $\square$

Definition (Special orthogonal group $SO(n)$)**.** The *special orthogonal group* is the kernel of $\det : \mathrm{O}(n) \to \{\pm 1\}$.

$$\mathrm{SO}(n) = \mathrm{SO}_n = \mathrm{SO}_n(\mathbb{R}) = \{A \in \mathrm{O}(n) : \det A = 1\}.$$

By the isomorphism theorem, $\mathrm{O}(n)/\mathrm{SO}(n) \cong C_2$.

Lemma. $\mathrm{O}(n) = \mathrm{SO}(n) \cup \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix} \mathrm{SO}(n)$

Proof. Cosets partition the group. $\square$

Lemma. (Orthogonal matrices are isometries) For $A \in \mathrm{O}(n)$ and $x, y \in \mathbb{R}^n$, we have

(i) $(A\mathbf{x}) \cdot (A\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$

(ii) $|A\mathbf{x}| = |\mathbf{x}|$

Proof. Treat the dot product as a matrix multiplication. So

$$(A\mathbf{x})^T(A\mathbf{y}) = \mathbf{x}^TA^TAy = \mathbf{x}^TI\mathbf{y} = \mathbf{x}^T\mathbf{y}$$

Then we have $|A\mathbf{x}|^2 = (A\mathbf{x}) \cdot (A\mathbf{x}) = \mathbf{x} \cdot \mathbf{x} = |\mathbf{x}|^2$. Since both are positive, we know that $|A\mathbf{x}| = |\mathbf{x}|$. $\square$

Note: All orthogonal matrices are isometries, but not isometries are orthogonal matrices since some isometries (e.g. translation) are not linear maps and are thus not matrices.

8.4 Rotations and reflections in $\mathbb{R}^2$

Lemma. $\mathrm{SO}(2)$ consists of all rotations of $\mathbb{R}^2$ around 0.

Proof. Let $A \in \mathrm{SO}(2)$. So $A^TA = I$ and $\det A = 1$. Suppose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. So $A^T = A^{-1}$ implies $ad - bc = 1$, $c = -b$, $d = a$. Combining these equations we obtain $a^2 + c^2 = 1$. Set $a = \cos\theta = d$, and $c = \sin\theta = -b$. Then these satisfies all three equations. So

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Note that A maps $(1, 0)$ to $(\cos\theta, \sin\theta)$, and maps $(0, 1) = (-\sin\theta, \cos\theta)$, which are rotations by θ counterclockwise. So A represents a rotation by θ. $\square$

Corollary. Any matrix in $\mathrm{O}(2)$ is either a rotation around 0 or a reflection in a line through 0.

Proof. If $A \in \mathrm{SO}(2)$, we've show that it is a rotation. Otherwise,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}$$

Since $\mathrm{O}(2) = \mathrm{SO}(2) \cap \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\mathrm{SO}(2)$. This has eigenvalues $1, -1$. So it is a reflection in the line of the eigenspace E_1. The line goes through $\mathbf{0}$ since the eigenspace is a subspace which must include $\mathbf{0}$. $\square$

Lemma. Every matrix in $\mathrm{SO}(3)$ is a rotation around some axis.

Proof. Let $A \in \mathrm{SO}(3)$. We know that $\det A = 1$ and A is an isometry. The eigenvalues λ must have $|\lambda| = 1$. They also multiply to $\det A = 1$. Since we are in $\mathbb{R}$, complex eigenvalues come in complex conjugate pairs. If there are complex eigenvalues λ and $\bar{\lambda}$, then $\lambda\bar{\lambda} = |\lambda|^2 = 1$. The third eigenvalue must be real and has to be $+1$.

If all eigenvalues are real. Then eigenvalues are either 1 or -1, and must multiply to 1. The possibilities are $1, 1, 1$ and $-1, -1, 1$, all of which contain an eigenvalue of 1.

So pick an eigenvector for our eigenvalue 1 as the third basis vector. Then in some orthonormal basis,

$$A = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Since the third column is the image of the third basis vector, and by orthogonality the third row is $0, 0, 1$. Now let

$$A' = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$$

with $\det A' = 1$. A' is still orthogonal, so $A' \in \mathrm{SO}(2)$. Therefore A' is a rotation and

$$A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

in some basis, and this is exactly the rotation through an axis. $\square$

Lemma. Every matrix in $\mathrm{O}(3)$ is the product of at most three reflections in planes through 0.

Note: A rotation is a product of two reflections. This lemma effectively states that every matrix in $\mathrm{O}(3)$ is a reflection, a rotation or a product of a reflection and a rotation.

Proof. Recall $O(3) = SO(3) \cup \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} SO(3)$ So if $A \in \mathrm{SO}(3)$, we know that $A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ in some basis, which is a composite of two reflections:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ \sin\theta & -\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

Then if $A \in \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} SO(3)$, then it is automatically a product of three reflections. $\square$

Note: In the last line we've shown that everything in $\mathrm{O}(3) \setminus \mathrm{SO}(3)$ can be written as a product of three reflections, but it is possible that they need only 1 reflection. However, some matrices do genuinely need 3 reflections, e.g. $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

8.5 Unitary groups

The is the complex equivalent of orthogonal groups. Recall $A^\dagger = (A^*)^T$ with $(A^\dagger)_{ij} = A^*_{ji}$, where the asterisk is the complex conjugate. We still have

(i) $(AB)^\dagger = B^\dagger A^\dagger$

(ii) $(A^{-1})^\dagger = (A^\dagger)^{-1}$

(iii) $A^\dagger A = I \Leftrightarrow AA^\dagger = I \Leftrightarrow A^\dagger A^{-1}$. We say A is a *unitary matrix*

(iv) $\det A^\dagger = (\det A)^*$

Definition (Unitary group $U(n)$)**.** The *unitary group* is $\mathrm{U}(n) = \mathrm{U}_n = \{A \in \mathrm{GL}_n(\mathbb{C}) : A^\dagger A = I\}$.

Lemma. $\det : \mathrm{U}(n) \to S^1$, where S^1 is the unit circle in the complex plane, is a surjective group homomorphism.

Proof. We know that $1 = \det I = \det A^\dagger A = |\det A|^2$. So $|\det A| = 1$. Since $\det AB = \det A \det B$, it is a group homomorphism.

Now given $\lambda \in S_1$, we have $\begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{U}(n)$. So it is surjective. $\square$

Definition (Special unitary group $SU(n)$)**.** The *special unitary group* $\mathrm{SU}(n) = \mathrm{SU}_n$ is the kernel of $\det U(n) \to S^1$.

Note: $Q_8 \leq \mathrm{SU}(2)$

Unitary matrices preserve the complex dot product: $(A\mathbf{x}) \cdot (A\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$

9 More on regular polyhedra

9.1 Symmetries of the cube

9.1.1 Rotations

Recall that there are $|G^+| = 24$ rotations of the group by the orbit-stabilizer theorem.

Proposition. $G^+ \cong S_4$, where G^+ is the group of all rotations of the cube.

Proof. Consider G^+ acting on the 4 diagonals of the cube. This gives a group homomorphism $\varphi : G^+ \to S_4$. We have $(1\ 2\ 3\ 4) \in \operatorname{im}\varphi$ by rotation around the axis through the top and bottom face. We also $(1\ 2) \in \operatorname{im}\varphi$ by rotation around the axis through the mid-point of the edge connect 1 and 2. Since $(1\ 2)$ and $(1\ 2\ 3\ 4)$ generate S_4 (Sheet 2 Q. 5d), $\operatorname{im}\varphi = S_4$, i.e. ϕ is surjective. Since $|S_4| = |G^+|$, φ must be an isomorphism. $\square$

9.1.2 All symmetries

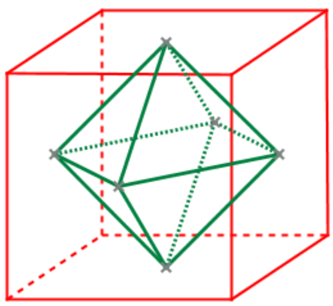
Consider the reflection in the mid-point of the cube τ, sending every point to its opposite. We can view this as $-I$ in $\mathbb{R}^3$. So it commutes with all other symmetries of the cube.

Proposition. $G \cong S_4 \times C_2$, where G is the group of all symmetries of the cube.

Proof. Let τ be "reflection in mid-point" as shown above. This commutes with everything. (Actually it is enough to check that it commutes with rotations only)

We have to show that $G = G^+\langle\tau\rangle$. This can be deduced using sizes: since G^+ and $\langle\tau\rangle$ intersect at e only, (i) and (ii) of the Direct Product Theorem gives an injective group homomorphism $G^+ \times \langle\tau\rangle \to G$. Since both sides have the same size, the homomorphism must be surjective as well. So $G \cong G^+ \times \langle\tau\rangle \cong S_4 \times C_2$. $\square$

In fact, we have also proved that the group of symmetries of an octahedron in $S_4 \times C_2$ since the octahedron is the dual of the cube. (if you join the centers of each face of the cube, you get an octahedron)



Credits: Wikimedia Commons: Birgit Lachner, CC-by-SA 3.0

9.2 Symmetries of the tetrahedron

9.2.1 Rotations

Let $1, 2, 3, 4$ be the vertices (in any order). G^+ is just the rotations. Let it act on the vertices. Then $\mathrm{orb}(1) = \{1, 2, 3, 4\}$ and $\mathrm{stab}(1) = \{$ rotations in the axis through 1 and the center of the opposite face $\} = \{e, \frac{2\pi}{3}, \frac{4\pi}{3}\}$

So $|G^+| = 4 \cdot 3 = 12$ by the orbit-stabilizer theorem.

The action gives a group homomorphism $\varphi : G^+ \to S_4$. Clearly $\ker\varphi = \{e\}$. So $G^+ \leq S_4$ and G^+ has size 12. We "guess" it is A_4 (actually it *must* be A_4 since that is the only subgroup of S_4 of order 12, but it's nice to see why that's the case).

If we rotate in an axis through 1, we get $(2\ 3\ 4), (2\ 4\ 3)$. Similarly, rotating through other axes through vertices gives all 3-cycles.

If we rotate through an axis that passes through two opposite edges, e.g. through 1-2 edge and 3-4 edge, then we have $(1\ 2)(3\ 4)$ and similarly we obtain all double transpositions. So $G^+ \cong A_4$. This shows that there is no *rotation* that fixes two vertices and swaps the other two.

9.2.2 All symmetries

Now consider the plane that goes through 1, 2 and the mid-point of 3 and 4. Reflection through this plane swaps 3 and 4, but doesn't change $1, 2$. So now $\mathrm{stab}(1) = \langle(2\ 3\ 4), (3, 4)\rangle \cong D_6$ (alternatively, if we want to fix 1, we just move 2, 3, 4 around which is the symmetries of the triangular base)

So $|G| = 4 \cdot 6 = 24$ and $G \cong S_4$ (which makes sense since we can move any of its vertices around in any way and still be a tetrahedron, so we have all permutations of vertices as the symmetry group)

10 Möbius group

We want to study maps $f : \mathbb{C} \to \mathbb{C}$ in the form $f(z) = \frac{az+b}{cz+d}$. with $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$.

We impose $ad - bc \neq 0$ or else the map will be constant: for any $z, w \in \mathbb{C}$, $f(z) - f(w) = \frac{(az+b)(cw+d)-(aw+b)(cz+d)}{(cw+d)(cz+d)} = \frac{(ad-bc)(z-w)}{(cw+d)(cz+d)}$. If $ad - bc = 0$, then f is constant and boring.

If $c \neq 0$, then $f(-\frac{d}{c})$ involves division by 0. So we add ∞ to $\mathbb{C}$ to form the extended complex plane (Riemann sphere) $\mathbb{C} \cup \{\infty\} = \mathbb{C}_\infty$ (c.f. Vectors and Matrices). Then we define $f(-\frac{d}{c}) = \infty$. We call $\mathbb{C}_\infty$ a one-point compactification of $\mathbb{C}$ (because it adds one point to $\mathbb{C}$ to make it compact, c.f. Metric and Topology).

Definition (Möbius map)**.** A *Möbius map* is a map from $\mathbb{C}_\infty \to \mathbb{C}_\infty$ of the form

$$f(z) = \frac{az + b}{cz + d},$$

where $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$, with $f(-\frac{d}{c}) = \infty$ and $f(\infty) = \frac{a}{c}$ when $c \neq 0$. (if $c = 0$, then $f(\infty) = \infty$)

Lemma. The Möbius maps are bijections $\mathbb{C}_\infty \to \mathbb{C}_\infty$.

Proof. The inverse of $f(z) = \frac{az+b}{cz+d}$ is $g(z) = \frac{dz-b}{-cz+a}$, which we can check by composition both ways. □

Proposition. The Möbius maps form a group M under function composition. (The Möbius group)

Proof. The group axioms are shown as follows:

0. If $f_1(z) = \frac{a_1z+b_1}{c_1z+d_1}$ and $f_2(z) = \frac{a_2z+b_2}{c_2z+d_2}$, then $f_2 \circ f_1(z) = \dfrac{a_2\left(\frac{a_1z+b_1}{c_1z+d_1}\right) + b_2}{c_2\left(\frac{a_1z+b_1}{c_1z+d_1}\right) + d_2} = \dfrac{(a_1a_2 + b_2c_1)z + (a_2b_1 + b_2d_1)}{(c_2a_1 + d_2c_1)z + (c_2b_1 + d_1d_2)}$. Now we have to check that $ad - bc \neq 0$: we have $(a_1a_2 + b_2c_1)(c_2b_1 + d_1d_2) - (a_2b_1 + b_2d_1)(c_2a_1 + d_2c_1) = (a_1d_1 - b_1c_1)(a_2d_2 - b_2c_2) \neq 0$.

   (This works for $z \neq \infty, -\frac{d_1}{c_1}$. We have to manually check the special cases, which is simply yet more tedious algebra)

1. The identity function is $1(z) = \frac{1z+0}{0+1}$ which satisfies $ad - bc \neq 0$.

2. We have shown above that $f^{-1}(z) = \frac{dz-b}{-cz+a}$ with $da - bc \neq 0$, which are also Möbius maps

3. Composition of functions is always associative

□

Note: M is not abelian. e.g. $f_1(z) = 2z$ and $f_2(z) = z + 1$ are not commutative: $f_1 \circ f_2(z) = 2z + 2$ and $f_2 \circ f_1(z) = 2z + 1$.
Note: The point at "infinity" is not special. Morally, ∞ is no different to any other point of the Riemann sphere. However, from the way we write down the

Möbius map, we have to check infinity specially. In this particular case, we can get quite far with conventions such as $\frac{1}{\infty} = 0$, $\frac{1}{0} = \infty$ and $\frac{a\cdot\infty}{c\cdot\infty} = \frac{a}{c}$.

Clearly $\frac{az+b}{cz+d} = \frac{\lambda az+\lambda b}{\lambda cz+\lambda d}$ for any $\lambda \neq 0$. So we do not have a unique representation of a map in terms of a, b, c, d. But a, b, c, d does uniquely determine a Möbius map.

Proposition. The map $\theta : \mathrm{GL}_2(\mathbb{C}) \to M$ sending $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \dfrac{az+b}{cz+d}$ is a surjective group homomorphism.

Proof. Firstly, since the determinant $ad - bc$ of any matrix in $\mathrm{GL}_2(\mathbb{C})$ is non-zero, it does map to a Möbius map. This also shows that θ is surjective.

We have previously calculated that

$$\theta(A_2) \circ \theta(A_1) = \frac{(a_1a_2 + b_2c_1)z + (a_2b_1 + b_2d_1)}{(c_2a_1 + d_2c_1)z + (c_2b_1 + d_1d_2)} = \theta(A_2A_1)$$

So it is a homomorphism. □

The kernel of θ is

$$\ker(\theta) = \left\{ A \in \mathrm{GL}_2(\mathbb{C}) : z = \frac{az+b}{cz+d} \forall z \right\}$$

We can try different values of z: $z = \infty \Rightarrow c = 0$; $z = 0 \Rightarrow b = 0$; $z = 1 \Rightarrow d = a$. So

$$\ker\theta = Z = \{\lambda I : \lambda \in \mathbb{C}, \lambda \neq 0\},$$

where I is the identity matrix and Z is the centre of $\mathrm{GL}_2(\mathbb{C})$.

By the isomorphism theorem, we have

$$M \cong \mathrm{GL}_2(\mathbb{C})/Z$$

Definition (Projective general linear group $\mathrm{PGL}_2(\mathbb{C})$)**.** (Non-examinable)The projective general linear group is

$$\mathrm{PGL}_2(\mathbb{C}) = \mathrm{GL}_2(\mathbb{C})/Z.$$

Note: Since $f_A = f_B$ iff $B = \lambda A$ for some $\lambda \neq 0$ (where A, B are the corresponding matrices of the maps), if we restrict θ to $\mathrm{SL}_2(\mathbb{C})$, we have $\theta|_{\mathrm{SL}_2(\mathbb{C})} : \mathrm{SL}_2(\mathbb{C}) \to M$ is also surjective. The kernel is now just $\{\pm I\}$. So

$$M \cong \mathrm{SL}_2(\mathbb{C})/\{\pm I\} = \mathrm{PSL}_2(\mathbb{C})$$

Note: Then clearly $\mathrm{PSL}_2(\mathbb{C}) \cong \mathrm{PGL}_2(\mathbb{C})$ since both are isomorphic to the Möbius group.

Proposition. Every Möbius map is a composite of maps of the following form:

(i) Dilation/rotation: $f(z) = az$, $a \neq 0$

(ii) Translation: $f(z) = z + b$

(iii) Inversion: $f(z) = \frac{1}{z}$

Proof. Let $\frac{az+b}{cz+d} \in M$.

If $c = 0$, i.e. $g(\infty) = \infty$, then $g(z) = \frac{a}{d}z + \frac{b}{d}$, i.e.

$$z \mapsto \frac{a}{d}z \mapsto \frac{a}{d}z + \frac{b}{d}.$$

If $c \neq 0$, let $g(\infty) = z_0$, Let $h(z) = \frac{1}{z - z_0}$. Then $hg(\infty) = \infty$ is of the above form. and $h^{-1}(w) = \frac{1}{w} + z_0$, and is of type (iii) followed by (ii). So $g = h^{-1}(hg)$ is a composition of maps of the three forms listed above.

Alternatively, with sufficient magic, we have

$$z \mapsto z + \frac{d}{c} \mapsto \frac{1}{z + \frac{d}{c}} \mapsto -\frac{ad + bc}{c^2(z + \frac{d}{c})} \mapsto \frac{a}{c} - \frac{ad + bc}{c^2(z + \frac{d}{c})} = \frac{az + b}{cz + d}$$

□

Note: The non-calculation method above can be transformed into another (different) composition with the same end result. So the way we compose a Möbius map from the "elementary" maps are not unique.

10.1 Fixed points of Möbius maps

Definition (Fixed point)**.** A *fixed point* of f is a z such that $f(z) = z$.

We know that any Möbius map with $c = 0$ fixes ∞. We also know that $z \to z + b$ for any $b \neq 0$ fixes ∞ only, where as $z \mapsto az$ for $a \neq 0, 1$ fixes 0 and ∞.

Proposition. Any Möbius map with at least 3 fixed points must be the identity.

Proof. Consider $f(z) = \frac{az+b}{cz+d}$. This has fixed points at those z which satisfy $\frac{az+b}{cz+d} = z \Leftrightarrow cz^2 + (d - a)z - b = 0$. A quadratic has at most two roots, unless $c = b = 0$ and $d = a$, in which the equation just says $0 = 0$.

However, if $c = b = 0$ and $d = a$, then f is just the identity. □

Proposition. Any Möbius map is conjugate to $f(z) = \nu z$ for some $\nu \neq 0$ or to $f(z) = z + 1$.

Proof. We have the surjective group homomorphism $\theta : \mathrm{GL}_2(\mathbb{C}) \to M$. The conjugacy classes of $\mathrm{GL}_2(\mathbb{C})$ are of types

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \mapsto g(z) = \frac{\lambda z + 0}{0z + \mu} = \frac{\lambda}{\mu} z$$
$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \mapsto g(z) = \frac{\lambda z + 0}{0z + \lambda} = 1z$$
$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \mapsto g(z) = \frac{\lambda z + 1}{\lambda} = z + \frac{1}{\lambda}$$

But the last one is not in the form $z + 1$. We know that the last $g(z)$ can also be represented by $\begin{pmatrix} 1 & \frac{1}{\lambda} \\ 0 & 1 \end{pmatrix}$, which is conjugate to $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (since that's its Jordan-normal form). So $z + \frac{1}{\lambda}$ is also conjugate to $z + 1$. □

Now we see easily that (for $\nu \neq 0, 1$), νz has 0 and ∞ as fixed points, $z+1$ only has ∞. Does this transfer to their conjugates?

Proposition. Every non-identity has exactly 1 or 2 fixed points.

Proof. Given $f \in M$ and $f \neq \text{id}$. So $\exists h \in M$ such that $hfh^{-1}(z) = \nu z$. Now $f(w) = w \Leftrightarrow hf(w) = h(w) \Leftrightarrow hfh^{-1}(h(w)) = h(w)$. So $h(w)$ is a fixed point of hfh^{-1}. Since h is a bijection, f and hfh^{-1} have the same number of fixed points.

So f has exactly 2 fixed points if f is conjugate to νz, and exactly 1 fixed point if f is conjugate to $z+1$. □

Intuitively, we can show that conjugation preserves fixed points because if we conjugate by h, we first move the Riemann sphere around by h, apply f (that fixes the fixed points) then restore the Riemann sphere to its original orientation. So we have simply moved the fixed point around by h.

10.2 Permutation properties of Möbius maps

We have seen that the Möbius map with three fixed points is the identity.

Proposition. Given $f, g \in M$. If $\exists z_1, z_2, z_3 \in \mathbb{C}_\infty$ such that $f(z_i) = g(z_i)$, then $f = g$. i.e. every Möbius map is uniquely determined by three points.

Proof. As Möbius maps are invertible, write $f(z_i) = g(z_i)$ as $g^{-1}f(z_i) = z_i$. So $g^{-1}f$ has three fixed points. So $g^{-1}f$ must be the identity. So $f = g$. □

Definition (Three-transitive action)**.** An action of G on X is called *three-transitive* if the induced action on $\{(x_1, x_2, x_3) \in X^3 : x_i \text{ pairwise disjoint}\}$, given by $g(x_1, x_2, x_3) = (g(x_1), g(x_2), g(x_3))$, is transitive.

This means that for any two triples x_1, x_2, x_3 and y_1, y_2, y_3 of distinct elements of X, there exists $g \in G$ such that $g(x_i) = y_i$.

If this g is always unique, then the action is called *sharply three transitive*

Proposition. The Möbius group M acts sharply three-transitively on $\mathbb{C}_\infty$.

Proof. We want to show that we can send any three points to any other three points. However, it is easier to show that we can send any three points to $0, 1, \infty$.

Suppose we want to send $z_1 \to \infty, z_2 \mapsto 0, z_3 \mapsto 1$. Then

$$f(z) = \frac{(z - z_2)(z_3 - z_1)}{(z - z_1)(z_3 - z_2)}$$

If any term z_i is ∞, we simply remove the terms with z_i, e.g. if $z_1 = \infty$, we have $f(z) = \frac{z - z_2}{z_3 - z_2}$.

So given also w_1, w_2, w_3 distinct in $\mathbb{C}_\infty$ and $g \in M$ sending $w_1 \mapsto \infty, w_2 \mapsto 0, w_3 \mapsto 1$, then we have $g^{-1}f(z_i) = w_i$.

The uniqueness of the map follows from the fact that a Möbius map is uniquely determined by 3 points. □

Note that 3 points also determines lines/circles on the plane (we can view a line in $\mathbb{C}$ as a circle on the Riemann sphere through infinity).

Lemma. The general equation of a circle or straight line in $\mathbb{C}$ is

$$Az\bar{z} + \bar{B}z + B\bar{z} + C = 0,$$

where $A, C \in \mathbb{R}$ and $|B|^2 > AC$.

Note: $A = 0$ gives a straight line. If $A \neq 0, B = 0$, we have a circle centered at the origin. If $C = 0$, the circle passes through 0.

Proof. This comes from noting that $|z - B| = r$ for $r \in \mathbb{R} > 0$ is a circle; $|z - a| = |z - b|$ with $a \neq b$ is a line. (c.f. Vectors and Matrices) □

Proposition. Möbius maps send circles/straight lines to circles/straight lines. (NOTE: it can send circles to straight lines and vice versa)

Alternatively, Möbius maps send circles on the Riemann sphere to circles on the Riemann sphere.

Proof. We can either calculate it directly using $w = \frac{az+b}{cz+d} \Leftrightarrow z = \frac{dw-b}{-cw+a}$ and substituting z into the circle equation, which gives $A'w\bar{w} + \bar{B}'w + B'\bar{w} + C' = 0$ with $A', C' \in \mathbb{R}$.

Alternatively, we know that each Möbius map is a composition of translation, dilation/rotation and inversion. We can check for each of the three types. Clearly dilation/rotation and translation maps a circle/line to a circle/line. So we simply do inversion: if $w = z^{-1}$

$$\begin{aligned} &Az\bar{z} + \bar{B}z + B\bar{z} + C = 0 \\ \Leftrightarrow\ &Cw\bar{w} + Bw + \bar{B}\bar{w} + A = 0 \end{aligned}$$

□

Example. Consider $f(z) = \frac{z-i}{z+i}$. Where does the real line go? The real line is simply a circle through $0, 1, \infty$. f maps this circle to the circle containing $f(\infty) = 1$, $f(0) = -1$ and $f(1) = -i$, which is the unit circle.

Where does the upper half plane go? We know that the Möbius map is smooth. So the upper-half plane either maps to the inside of the circle or the outside of the circle. We try the point i, which maps to 0. So the upper half plane is mapped to the inside of the circle.

10.3 Cross-ratios

Definition (Cross-ratios)**.** Given four distinct points $z_1, z_2, z_3, z_4 \in \mathbb{C}_\infty$, their *cross-ratio* is $[z_1, z_2, z_3, z_4] = g(z_4)$, with g being the unique Möbius map that maps $z_1 \mapsto \infty, z_2 \mapsto 0, z_3 \mapsto 1$. So $[\infty, 0, 1, \lambda] = \lambda$ for any $\lambda \neq \infty, 0, 1$. We have

$$[z_1, z_2, z_3, z_4] = \frac{z_4 - z_2}{z_4 - z_1} \cdot \frac{z_3 - z_1}{z_3 - z_2}$$

(with special cases as above).

We know that this exists and is uniquely defined because M acts sharply three-transitively on $\mathbb{C}_\infty$.

Note: Different authors use different permutations of $1, 2, 3, 4$, but they all lead to the same result as long as you are consistent.

Lemma. For $z_1, z_2, z_3, z_4 \in \mathbb{C}_\infty$ all distinct, then

$$[z_1, z_2, z_3, z_4] = [z_2, z_1, z_4, z_3] = [z_3, z_4, z_1, z_2] = [z_4, z_3, z_2, z_1]$$

i.e. if we perform a double transposition on the entries, the cross-ratio is retained.

Proof. By inspection of the formula. $\square$

Proposition. If $f \in M$, then $[z_1, z_2, z_3, z_4] = [f(z_1), f(z_2), f(z_3), f(z_4)]$.

Proof. Use our original definition of the cross ratio (instead of the formula). Let g be the unique Möbius map such that $[z_1, z_2, z_3, z_4] = g(z_4) = \lambda$, i.e.

$$\begin{aligned} z_1 &\overset{g}{\mapsto} \infty \\ z_2 &\mapsto 0 \\ z_3 &\mapsto 1 \\ z_4 &\mapsto \lambda \end{aligned}$$

We know that gf^{-1} sends

$$\begin{aligned} f(z_1) &\xmapsto{f^{-1}} z_1 \overset{g}{\mapsto} \infty \\ f(z_2) &\xmapsto{f^{-1}} z_2 \overset{g}{\mapsto} 0 \\ f(z_3) &\xmapsto{f^{-1}} z_3 \overset{g}{\mapsto} 1 \\ f(z_4) &\xmapsto{f^{-1}} z_4 \overset{g}{\mapsto} \lambda \end{aligned}$$

So $[f(z_1), f(z_2), f(z_3), f(z_4)] = gf^{-1}f(z_4) = g(z_4) = \lambda$ $\square$

In fact, we can see from this proof that: given z_1, z_2, z_3, z_4 all distinct and w_1, w_2, w_3, w_4 distinct in $\mathbb{C}_\infty$, then $\exists f \in M$ with $f(z_i) = w_i$ iff $[z_1, z_2, z_3, z_4] = [w_1, w_2, w_3, w_4]$

Corollary. z_1, z_2, z_3, z_4 lie on some circle/straight line iff $[z_1, z_2, z_3, z_4] \in \mathbb{R}$.

Proof. Let C be the circle/line through z_1, z_2, z_3. Let g be the unique Möbius map with $g(z_1) = \infty$, $g(z_2) = 0$, $g(z_3) = 1$. Then $g(z_4) = [z_1, z_2, z_3, z_4]$ by definition.

Since we know that Möbius maps preserve circle/lines, $z_4 \in C \Leftrightarrow g(z_4)$ is on the line through $\infty, 0, 1$, i.e. $g(z_4) \in \mathbb{R}$. $\square$

11 Projective line (non-examinable)

We have seen in matrix groups that $\mathrm{GL}_2(\mathbb{C})$ acts on $\mathbb{C}^2$, the column vectors. Instead, we can also have $\mathrm{GL}_2(\mathbb{C})$ acting on the set of 1-dimensional subspaces (i.e. lines) of $\mathbb{C}^2$.

For any $\mathbf{v} \in \mathbb{C}^2$, write the line generated by $\mathbf{v}$ as $\langle \mathbf{v} \rangle$. Then clearly $\langle \mathbf{v} \rangle = \{\lambda \mathbf{v} : \lambda \in \mathbb{C}\}$. Now for any $A \in \mathrm{GL}_2(\mathbb{C})$, define the action as $A\langle \mathbf{v} \rangle = \langle A\mathbf{v} \rangle$. Check that this is well-defined: for any $\langle \mathbf{v} \rangle = \langle \mathbf{w} \rangle$, we want to show that $\langle A\mathbf{v} \rangle = \langle A\mathbf{w} \rangle$. This is true because $\langle \mathbf{v} \rangle = \langle \mathbf{w} \rangle$ if and only if $\mathbf{w} = \lambda \mathbf{v}$ for some $\lambda \in \mathbb{C} \setminus \{0\}$, and then $\langle A\mathbf{w} \rangle = \langle A\lambda\mathbf{v} \rangle = \langle \lambda(A\mathbf{v}) \rangle = \langle A\mathbf{v} \rangle$.

What is the kernel of this action? By definition the kernel has to fix all lines. In particular, it has to fix our magic lines generated by $\binom{1}{0}, \binom{0}{1}$ and $\binom{1}{1}$. Since we want $A\langle \binom{1}{0} \rangle = \langle \binom{1}{0} \rangle$, so we must have $A\binom{1}{0} = \binom{\lambda}{0}$ for some λ. Similarly, $A\binom{0}{1} = \binom{0}{\mu}$. So we can write $A = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$. However, also need $A\langle \binom{1}{1} \rangle = \langle \binom{1}{1} \rangle$. Since A is a linear function, we know that $A\binom{1}{1} = A\binom{1}{0} + A\binom{0}{1} = \binom{\lambda}{\mu}$. For the final vector to be parallel to $\binom{1}{1}$, we must have $\lambda = \mu$. So $A = \lambda I$ for some I. Clearly any matrix of this form fixes any line. So the kernel $Z = \{\lambda I : \lambda \in \mathbb{C} \setminus \{0\}\}$.

Note that every line is uniquely determined by its slope. For any $\mathbf{v} = (v_1, v_2), \mathbf{w} = (w_1, w_2)$, we have $\langle \mathbf{v} \rangle = \langle \mathbf{w} \rangle$ iff $z_1/z_2 = w_1/w_2$. So we have a one-to-one correspondence from our lines to $\mathbb{C}_\infty$, that maps $\langle \binom{z_1}{z_2} \rangle \leftrightarrow z_1/z_2$.

Finally, for each $A \in \mathrm{GL}_2(\mathbb{C})$, given any line $\langle \binom{z}{1} \rangle$, we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \left\langle \begin{pmatrix} z \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} az + b \\ cz + d \end{pmatrix} \right\rangle \leftrightarrow \frac{az + b}{cz + d}$$

So $\mathrm{GL}_2(\mathbb{C})$ acting on the lines is just "the same" as the Möbius groups acting lines.