M2PM2 Notes

By popular request, here are some notes on the M2PM2 lectures. They should not be used as a substitute for going to lectures: the notes will just contain the results, proofs and a few examples. The lectures will hopefully have much more discussion of the proofs, and many more examples, as well as fine artwork.....

Like M1P2 last year, this will be a course of two halves:

(A) Group theory; (B) Linear Algebra.

1 Revision from M1P2

Would be a good idea to refresh your memory on the following topics from group theory.

(a) *Group axioms*: closure, associativity, identity, inverses

(b) *Examples of groups*:

$(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{Q}^*,\times)$, $(\mathbb{C}^*,\times)$, etc

$GL(n,\mathbb{R})$, the group of all invertible $n\times n$ matrices over $\mathbb{R}$, under matrix multiplication

S_n, the symmetric group, the set of all permutations of $\{1,2,\dots,n\}$, under composition. Recall the cycle notation for permutations – every permutation can be expressed as a product of disjoint cycles.

For p prime $\mathbb{Z}_p^* = \{[1],[2],\dots,[p-1]\}$ is a group under multiplication modulo p.

$C_n = \{x\in\mathbb{C} : x^n = 1\} = \{1,\omega,\omega^2,\dots,\omega^{n-1}\}$ is a cyclic group of size n, where $\omega = e^{2\pi i/n}$.

(c) *Some theory*:

Criterion for subgroups: H is a subgroup of G iff (1) $e\in H$; (2) $x,y\in H\Rightarrow xy\in H$, and (3) $x\in H\Rightarrow x^{-1}\in H$.

For $a\in G$, we define the cyclic subgroup $\langle a\rangle = \{a^n : n\in\mathbb{Z}\}$. The size of $\langle a\rangle$ is equal to $o(a)$, the *order* of a, which is defined to be the smallest positive integer k such that $a^k = e$.

Lagrange: if H is a subgroup of a finite group G then $|H|$ divides $|G|$.

Consequences: (1) For any element $a\in G$, $o(a)$ divides $|G|$.

(2) If $|G| = n$ then $x^n = e$ for all $x \in G$

(3) If $|G|$ is prime then G is a cyclic group.

2 More examples: symmetry groups

For any object in the plane $\mathbb{R}^2$ (later $\mathbb{R}^3$) we'll show how to define a group called the symmetry group of the object. This group will consist of functions called *isometries*, which we now define. Recall for $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R}^2$, the distance

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

We define an *isometry* of $\mathbb{R}^2$ to be a bijection $f : \mathbb{R}^2 \to \mathbb{R}^2$ which preserves distance, i.e. for all $x, y \in \mathbb{R}^2$,

$$d(f(x), f(y)) = d(x, y).$$

There are many familiar examples of isometries:

(1) Rotations: let $\rho_{P,\theta}$ be the function $\mathbb{R}^2 \to \mathbb{R}^2$ which rotates every point about P through angle θ. This is an isometry.

(2) Reflections: if l is a line, let σ_l be the function which sends every point to its reflection in l. This is an isometry.

(3) Translations: for $a \in \mathbb{R}^2$, let τ_a be the translation sending $x \to x + a$ for all $x \in \mathbb{R}^2$. This is an isometry.

Not every isometry is one of these three types – for example a glide-reflection (i.e. a function of the form $\sigma_l \circ \tau_a$) is not a rotation, reflection or translation.

Define $I(\mathbb{R}^2)$ to be the set of all isometries of $\mathbb{R}^2$. For isometries f, g, we have the usual composition function $f \circ g$ defined by $f \circ g(x) = f(g(x))$.

Proposition 2.1 *$I(\mathbb{R}^2)$ is a group under composition.*

Proof *Closure*: Let $f, g \in I(\mathbb{R}^2)$. We must show $f \circ g$ is an isometry. It is a bijection as f, g are bijections (recall M1F). And it preserves distance as

$$\begin{aligned} d(f \circ g(x),\, f \circ g(y)) &= d(f(g(x)),\, f(g(y))) \\ &= d(g(x), g(y) \text{ (as } f \text{ is isometry)} \\ &= d(x, y) \text{ (as } g \text{ is isometry).} \end{aligned}$$

Assoc: this is always true for composition of functions (since $f \circ (g \circ h)(x) = (f \circ g) \circ h(x) = f(g(h(x))))$.

Identity is the identity function e defined by $e(x) = x$ for all $x \in \mathbb{R}^2$, which is obviously an isometry.
Inverses: let $f \in I(\mathbb{R}^2)$. Then f^{-1} exists as f is a bijection, and f^{-1} preserves distance since

$$d(f^{-1}(x), f^{-1}(y)) = d(f(f^{-1}(x)), f(f^{-1}(y))) = d(x, y).$$

So we've checked all the axioms and $I(\mathbb{R}^2)$ is a group. □

Now let Π be a subset of $\mathbb{R}^2$. For a function $g : \mathbb{R}^2 \to \mathbb{R}^2$,

$$g(\Pi) = \{g(x) \mid x \in \Pi\}$$

Example: Π =square with centre in the origin and aligned with axes, $g = \rho_{\pi/4}$. Then $g(\Pi)$ is the original square rotated by $\pi/4$.

Definition The *symmetry group* of Π is $G(\Pi)$ – the set of isometries g such that $g(\Pi) = \Pi$, i.e.

$$G(\Pi) = \left\{g \in I(\mathbb{R}^2) \mid g(\Pi) = \Pi\right\}.$$

Example: For the square from the previous example, $G(\Pi)$ contains $\rho_{\pi/2}$, $\sigma_x \dots$

Proposition 2.2 *$G(\Pi)$ is a subgroup of $I(\mathbb{R}^2)$.*

Proof We check the subgroup criteria:

(1) $e \in G(\Pi)$ as $e(\Pi) = \Pi$.

(2) Let $f, g \in G(\Pi)$, so $f(\Pi) = g(\Pi) = \Pi$. So

$$\begin{aligned} f \circ g(\Pi) &= f(g(\Pi)) && (1) \\ &= f(\Pi) && (2) \\ &= \Pi. && (3) \end{aligned}$$

So $f \circ g \in G(\Pi)$.

(3) Let $f \in G(\Pi)$, so

$$f(\Pi) = \Pi.$$

Apply f^{-1} to get

$$f^{-1}(f(\Pi)) = f^{-1}(\Pi) \quad (4)$$
$$\Pi = f^{-1}(\Pi) \quad (5)$$

and $f^{-1} \in G(\Pi)$. □

So we have a vast collection of new examples of groups $G(\Pi)$.

Examples

1. **Equilateral triangle** $(= \Pi)$
   Here $G(\Pi)$ contains

   3 *rotations*: $e = \rho_0$, $\rho = \rho_{2\pi/3}$, $\rho^2 = \rho_{4\pi/3}$,
   3 *reflections*: $\sigma_1 = \sigma_{l_1}$, $\sigma_2 = \sigma_{l_2}$, $\sigma_3 = \sigma_{l_3}$.

   Each of these corresponds to a permutation of the corners 1, 2, 3:

$$e \sim e, \quad (6)$$
$$\rho \sim (1\ 2\ 3), \quad (7)$$
$$\rho^2 \sim (1\ 3\ 2), \quad (8)$$
$$\sigma_1 \sim (2\ 3), \quad (9)$$
$$\sigma_2 \sim (1\ 3), \quad (10)$$
$$\sigma_3 \sim (1\ 2). \quad (11)$$

   Any isometry in $G(\Pi)$ permutes the corners. Since all the permutations of the corners are already present, there can't be any more isometries in $G(\Pi)$. So the Symmetry group of equilateral triangle is

$$\{e, \rho, \rho^2, \sigma_1, \sigma_2, \sigma_3\},$$

   called the *dihedral group* D_6.

   Note that it is easy to work out products in D_6: e.g.

$$\rho\sigma_3 \sim (1\ 2\ 3)(1\ 2) = (1\ 3) \quad (12)$$
$$\sim \sigma_2. \quad (13)$$

2. **The square**
   Here $G = G(\Pi)$ contains

4 *rotations*: e, ρ, ρ^2, ρ^3 where $\rho = \rho_{\pi/2}$,

4 *reflections*: $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ where $\sigma_i = \sigma_{l_i}$.

So $|G| \geq 8$. We claim that $|G| = 8$: Any $g \in G$ permutes the corners $1, 2, 3, 4$ (as g preserves distance). So g sends

$1 \to i$, (4 choices of i)

$2 \to j$, neighbour of i, (2 choices for j)

$3 \to$ opposite of i,

$4 \to$ opposite of j.

So $|G| \leq$ (num. of choices for i) $\times$ (for j) $= 4 \times 2 = 8$. So $|G| = 8$. Symmetry group of the square is

$$\left\{e, \rho, \rho^2, \rho^3, \sigma_1, \sigma_2, \sigma_3, \sigma_4\right\},$$

called the *dihedral group* D_8.

Can work out products using the corresponding permutations of the corners.

$$\begin{aligned}
e &\sim e, && (14)\\
\rho &\sim (1\ 2\ 3\ 4), && (15)\\
\rho^2 &\sim (1\ 3)(2\ 4), && (16)\\
\rho^3 &\sim (1\ 4\ 3\ 2), && (17)\\
\sigma_1 &\sim (1\ 4)(2\ 3), && (18)\\
\sigma_2 &\sim (1\ 3), && (19)\\
\sigma_3 &\sim (1\ 2)(3\ 4), && (20)\\
\sigma_4 &\sim (2\ 4). && (21)
\end{aligned}$$

For example

$$\begin{aligned}
\rho^3\sigma_1 &\to (1\ 4\ 3\ 2)(1\ 4)(2\ 3) = (1\ 3) && (22)\\
&\to \sigma_2. && (23)
\end{aligned}$$

Note that *not* all permutations of the corners are present in D_8, e.g. $(1\ 2)$.

More on D_8: Define H to be the cyclic subgroup of D_8 generated by ρ, so

$$H = \langle\rho\rangle = \left\{e, \rho, \rho^2, \rho^3\right\}.$$

Write $\sigma = \sigma_1$. The right coset

$$H\sigma = \left\{\sigma, \rho\sigma, \rho^2\sigma, \rho^3\sigma\right\}$$

is different from H.

H	$H\sigma$

So the two distinct right cosets of H in D_8 are H and $H\sigma$, and

$$D_8 = H \cup H\sigma.$$

Hence

$$\begin{aligned} H\sigma &= \left\{\rho, \rho\sigma, \rho^2\sigma, \rho^3\sigma\right\} && (24) \\ &= \left\{\sigma_1, \sigma_2, \sigma_3, \sigma_4\right\}. && (25) \end{aligned}$$

So the elements of D_8 are

$$e, \rho, \rho^2, \rho^3, \sigma, \rho\sigma, \rho^2\sigma, \rho^3\sigma.$$

To work out products, use the "magic equation" (see Sheet 1, Question 2)

$$\sigma\rho = \rho^{-1}\sigma.$$

3. **Regular n-gon**

Let Π be the regular polygon with n sides. Symmetry group $G = G(\Pi)$ contains

> n *rotations*: $e, \rho, \rho^2, \ldots, \rho^{n-1}$ where $\rho = \rho_{2\pi/n}$,
>
> n *reflections* $\sigma_1, \sigma_2, \ldots, \sigma_n$ where $\sigma_i = \sigma_{l_i}$.

So $|G| \geq 2n$. We claim that $|G| = 2n$.

Any $g \in G$ sends corners to corners, say

$1 \to i$, (n choices for i)

$2 \to j$ neighbour of i. (2 choices for j)

Then g sends n to the other neighbour of i and $n-1$ to the remaining neighbour of $g(n)$ and so on. So once i, j are known, there is only one possibility for g. Hence

$$|G| \leq \text{number of choices for } i, j = 2n.$$

Therefore $|G| = 2n$.

Symmetry group of regular n-gon is

$$D_{2n} = \left\{e, \rho, \rho^2, \ldots, \rho^n, \sigma_1, \ldots, \sigma_n\right\},$$

the *dihedral group* of size $2n$.

Again can work in D_{2n} using permutations

$$\rho \rightarrow (1\ 2\ 3\ \cdots\ n) \quad (26)$$
$$\sigma_1 \rightarrow (2\ n)(3\ n-1)\cdots \quad (27)$$

4. **Benzene molecule**
   C_6H_6. Symmetry group is D_12.

5. **Infinite strip of `F`'s**

$$\begin{array}{ccccc} \ldots & \mathtt{F} & \mathtt{F} & \mathtt{F} & \ldots \\ & -1 & 0 & 1 & \end{array}$$

   What is symmetry group $G(\Pi)$?

   $G(\Pi)$ contains translation

$$\tau_{(1,0)} : v \mapsto v + (1, 0).$$

   Write $\tau = \tau_{(1,0)}$. Then $G(\Pi)$ contains all translations $\tau^n = \tau_{(n,0)}$. Note $G(\Pi)$ is infinite. We claim that

$$G(\Pi) = \{\tau^n \mid n \in \mathbb{Z}\} \quad (28)$$
$$= \langle\tau\rangle, \quad (29)$$

   infinite cyclic group.

   Let $g \in G(\Pi)$. Must show that $g = \tau^n$ for some n. Say g sends F at 0 to F at n. Note that τ^{-n} sends `F` at n to `F` at 0. So $\tau^{-n}g$ sends `F` at 0 to `F` at 0. So $\tau^{-n}g$ is a symmetry of the `F` at 0. It is easy to observe that `F` has only symmetry e. Hence

$$\tau^{-n}g = e \quad (30)$$
$$\tau^n\tau^{-n}g = \tau^n \quad (31)$$
$$g = \tau^n. \quad (32)$$

Note Various other figures have more interesting symmetry groups, e.g. infinite strip of `E`'s, square tiling of a plane, octagons and squares tiling of the plane, 3 dimensions – platonic solids…later.

3 Isomorphism

Let $G = C_2 = \{1, -1\}$, $H = S_2 = \{e, a\}$ (where $a = (1\,2)$). Multiplication tables:

Of G:	1	-1
1	1	-1
-1	-1	1

Of H:	e	a
e	e	a
a	a	e

These are the same, except that the elements have different labels ($1 \sim e$, $-1 \sim a$).

Similarly for $G = C_3 = \{1, \omega, \omega^2\}$, $H = \langle a \rangle = \{e, a, a^2\}$ (where $a = (1\,2\,3) \in S_3$):

Of G:	1	ω	ω^2
1	1	ω	ω^2
ω	ω	ω^2	1
ω^2	ω^2	1	ω

Of H:	e	a	a^2
e	e	a	a^2
a	a	a^2	e
a^2	a^2	e	a

Again, these are same groups with relabelling

$$\begin{aligned} 1 &\sim e, \\ \omega &\sim a, \\ \omega^2 &\sim a^2. \end{aligned}$$

In these examples, there is a "relabelling" function $\phi : G \to H$ such that if

$$\begin{aligned} g_1 &\mapsto h_1, \\ g_2 &\mapsto h_2, \end{aligned}$$

then

$$g_1 g_2 \mapsto h_1 h_2.$$

Definition G, H groups. A function $\phi : G \to H$ is an *isomorphism* if

(1) ϕ is a bijection,

(2) $\phi(g_1)\phi(g_2) = \phi(g_1 g_2)$ for all $g_1, g_2 \in G$.

If there exists an isomorphism $\phi : G \to H$, we say G is *isomorphic* to H and write $G \cong H$.

Notes 1. If $G \cong H$ then $|G| = |H|$ (as ϕ is a bijection).
2. The relation $\cong$ is an equivalence relation, i.e.

- $G \cong G$,
- $G \cong H \Rightarrow H \cong G$,
- $G \cong H, H \cong K \Rightarrow G \cong K$.

Example Which pairs of the following groups are isomorphic?

$$\begin{aligned} G_1 &= C_4 = \langle i \rangle = \{1, -1, i, -i\}, \\ G_2 &= \text{symmetry group of a rectangle} = \{e, \rho_\pi, \sigma_1, \sigma_2\}, \\ G_3 &= \text{cyclic subgroup of } D_8\ \langle \rho \rangle = \left\{e, \rho, \rho^2, \rho^3\right\}. \end{aligned}$$

1. $G_1 \cong G_3$? To prove this, define $\phi : G_1 \to G_2$

$$\begin{aligned} i &\mapsto \rho, \\ -1 &\mapsto \rho^2, \\ -i &\mapsto \rho^3, \\ 1 &\mapsto e, \end{aligned}$$

i.e. $\phi : i^n \mapsto \rho^n$. To check that ϕ is an isomorphism

(1) ϕ is a bijection,

(2) for $m, n \in \mathbb{Z}$

$$\begin{aligned} \phi(i^m i^n) &= \phi(i^{m+n}) \\ &= \rho^{m+n} \\ &= \rho^m \rho^n \\ &= \phi(i^m)\phi(i^n). \end{aligned}$$

So ϕ is an isomorphism and $G_1 \cong G_3$.

Note that there exist many bijections $G_1 \to G_3$ which are not isomorphisms.

2. $G_2 \cong G_3$ or $G_2 \cong G_1$? Answer: $G_2 \ncong G_1$. By contradiction. Assume there exists an isomorphism $\phi : G_1 \to G_2$. Say $\phi(i) = x \in G_2$, $\phi(1) = y \in G_2$. Then

$$\phi(-1) = \phi(i^2) = \phi(i \cdot i) = \phi(i)\phi(i) = x^2 = e$$

as $g^2 = e$ for all $g \in G_2$. Similarly $\phi(1) = \phi(1 \cdot 1) = \phi(1)\phi(1) = y^2 = e$. So $\phi(-1) = \phi(1)$, a contradiction as ϕ is a bijection.

In general, to decide whether two groups G, H are isomorphic:

- If you think $G \cong H$, try to define an isomorphism $\phi : G \to H$.
- If you think $G \ncong H$, try to use the following proposition.

Proposition 3.1 *Let G, H be groups.*

(1) If $|G| \neq |H|$ then $G \ncong H$.

(2) If G is abelian and H is not abelian, then $G \ncong H$.

(3) If there is an integer k such that G and H have different number of elements of order k, then $G \ncong H$.

Proof (1) Obvious.

(2) We show that if G is abelian and $G \cong H$, then H is abelian (this gives (2)). Suppose G is abelian and $\phi : G \to H$ is an isomorphism. Let $h_1, h_2 \in H$. As ϕ is a bijection, there exist $g_1, g_2 \in G$ such that $h_1 = \phi(g_1)$ and $h_2 = \phi(g_2)$. So

$$\begin{aligned} h_2 h_1 &= \phi(g_2)\phi(g_1) \\ &= \phi(g_2 g_1) \\ &= \phi(g_1)\phi(g_2) \\ &= h_1 h_2. \end{aligned}$$

(3) Let

$$\begin{aligned} G_k &= \{g \in G \mid o(g) = k\}, \\ H_k &= \{h \in H \mid o(h) = k\}. \end{aligned}$$

We show that $G \cong H$ implies $|G_k| = |H_k|$ for all k (this gives (3)).

Suppose $G \cong H$ and let $\phi : G \to H$ be an isomorphism. We show that ϕ sends G_k to H_k: Let $g \in G_k$, so $o(g) = k$, i.e.

$$g^k = e_G, \text{ and } g^i \neq e_G \text{ for } 1 \leq i \leq k - 1.$$

Now $\phi(e_G) = e_H$, since

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

Also

$$\begin{aligned}\phi(g^i) &= \phi(gg\cdots g) \ \ (i \text{ times}) \\ &= \phi(g)\phi(g)\cdots\phi(g) \\ &= \phi(g)^i.\end{aligned}$$

Hence

$$\begin{aligned}\phi(g)^k &= \phi(e_G) = e_H, \\ \phi(g)^i &\neq e_H \text{ for } 1 \leq i \leq k-1.\end{aligned}$$

In other words, $\phi(g)$ has order k, so $\phi(g) \in H_k$. So ϕ sends G_k to H_k. As ϕ is 1-1, this implies $|G_k| \leq |H_k|$.

Also $\phi^{-1} : H \to G$ is an isomorphism and similarly sends H_k to G_k, hence $|H_k| \leq |G_k|$. Therefore $|G_k| = |H_k|$. □

Examples 1. Let $G = S_4$, $H = D_8$. Then $|G| = 24$, $|H| = 8$, so $G \not\cong H$.

2. Let $G = S_3$, $H = C_6$. Then G is non-abelian, H is abelian, so $G \not\cong H$.

3. Let $G = C_4$, $H =$ symmetry group of the rectangle $= \{e, \rho_\pi, \sigma_1, \sigma_2\}$. Then G has 1 element of order 2, H has 3 elements of order 2, so $G \not\cong H$.

4. Question: $(\mathbb{R}, +) \cong (\mathbb{R}^*, \times)$? Answer: No, since $(\mathbb{R}, +)$ has 0 elements of order 2, $(\mathbb{R}^*, \times)$ has 1 element of order 2.

Cyclic groups

Proposition 3.2 *(1) If G is a cyclic group of size n, then $G \cong C_n$.*

(2) If G is an infinite cyclic group, then $G \cong (\mathbb{Z}, +)$.

Proof (1) Let $G = \langle x \rangle$, $|G| = n$, so $o(x) = n$ and therefore

$$G = \left\{e, x, x^2, \ldots, x^{n-1}\right\}.$$

Recall

$$C_n = \left\{1, \omega, \omega^2, \ldots, \omega^{n-1}\right\},$$

where $\omega = e^{2\pi i/n}$. Define $\phi : G \to G$ by $\phi(x^r) = \omega^r$ for all r. Then ϕ is a bijection, and

$$\begin{aligned}\phi(x^r x^s) &= \phi(x^{r+s}) \\ &= \omega^{r+s} \\ &= \omega^r \omega^s \\ &= \phi(x^r)\phi(x^s).\end{aligned}$$

So ϕ is an isomorphism, and $G \cong C_n$.

(2) Let $G = \langle x \rangle$ be infinite cyclic, so $o(x) = \infty$ and

$$G = \{\ldots, x^{-2}, x^{-1}, e, x, x^2, x^3, \ldots\},$$

all distinct. Define $\phi : G \to (\mathbb{Z}, +)$ by $\phi(x^r) = r$ for all r. Then ϕ is an isomorphism, so $G \cong (\mathbb{Z}, +)$. $\square$

This proposition says that if we think of isomorphic groups as being "the same", then there is only *one* cyclic group of each size. We say: "up to isomorphism", the only cyclic groups are C_n and $(\mathbb{Z}, +)$.

Example Cyclic subgroup $\langle 3 \rangle$ of $(\mathbb{Z}, +)$ is $\{3n \mid n \in \mathbb{Z}\}$, infinite, so by the proposition $\langle 3 \rangle \cong (\mathbb{Z}, +)$.

4 Even and odd permutations

We'll classify each permutation in S_n as either "even" or "odd" (reason given later).

Example For $n = 3$. Consider the expression

$$\Delta = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3),$$

a polynomial in 3 variables x_1, x_2, x_3. Take each permutation in S_3 to permute x_1, x_2, x_3 in the same way it permutes $1, 2, 3$. Then each $g \in S_3$ sends Δ to $\pm\Delta$. For example

for $e, (1\ 2\ 3), (1\ 3\ 2) : \Delta \mapsto +\Delta$,

for $(1\ 2), (1\ 3), (2\ 3) : \Delta \mapsto -\Delta$.

Generalizing this: for arbitrary $n \geq 2$, define

$$\Delta = \prod_{i<j} (x_i - x_j),$$

a polynomial in n variables $x_1, \ldots, x_n$.

If we let each permutation $g \in S_n$ permute the variables $x_1, \ldots, x_n$ just as it permutes $1, \ldots, n$ then g sends Δ to $\pm\Delta$.

Definition For $g \in S_n$, define the *signature* $\text{sgn}(g)$ to be $+1$ if $g(\Delta) = \Delta$ and -1 if $g(\Delta) = -\Delta$. So

$$g(\Delta) = \text{sgn}(g)\Delta.$$

The function sgn : $S_n \to \{+1, -1\}$ is the *signature function* on S_n. Call g an *even* permutation if $\mathrm{sgn}(g) = 1$, and *odd* permutation if $\mathrm{sgn}(g) = -1$.

Example In S_3 $e, (1\ 2\ 3), (1\ 3\ 2)$ are even and $(1\ 2), (1\ 3), (2\ 3)$ are odd.

Given $(1\ 2\ 3\ 5)(6\ 7\ 9)(8\ 4\ 10) \in S_{10}$, what's its signature ? Our next aim is to be able answer such questions instantaneously. This is the key:

Proposition 4.1 *(a)* $\mathrm{sgn}(xy) = \mathrm{sgn}(x)\mathrm{sgn}(y)$ *for all* $x, y \in S_n$

(b) $\mathrm{sgn}(e) = 1$*,* $\mathrm{sgn}(x^{-1}) = \mathrm{sgn}(x)$*.*

(c) If $t = (i\ j)$ *is a 2-cycle then* $\mathrm{sgn}(t) = -1$*.*

Proof (a) By definition

$$\begin{aligned} x(\Delta) &= \mathrm{sgn}(x)\Delta, \\ y(\Delta) &= \mathrm{sgn}(y)\Delta. \end{aligned}$$

So

$$\begin{aligned} xy(\Delta) &= x(y(\Delta)) \\ &= x(\mathrm{sgn}(y)\Delta) \\ &= \mathrm{sgn}(y)x(\Delta) = \mathrm{sgn}(y)\mathrm{sgn}(x)\Delta. \end{aligned}$$

Hence

$$\mathrm{sgn}(xy) = \mathrm{sgn}(x)\mathrm{sgn}(y).$$

(b) We have $e(\Delta) = \Delta$, so $\mathrm{sgn}(e) = 1$. So

$$\begin{aligned} 1 &= \mathrm{sgn}(e) = \mathrm{sgn}(xx^{-1}) \\ &= \mathrm{sgn}(x)\mathrm{sgn}(x^{-1}) \text{ (by (a))} \end{aligned}$$

and hence $\mathrm{sgn}(x) = \mathrm{sgn}(x^{-1})$.

(c) Let $t = (i\ j)$, $i < j$. We count the number of brackets in Δ that are sent to brackets $(x_r - x_s)$, $r > s$. These are

$$\begin{aligned} &(x_i - x_j), \\ &(x_i - x_{i+1}), \ldots, (x_i - x_{j-1}), \\ &(x_{i+1} - x_j), \ldots, (x_{j-1} - x_j). \end{aligned}$$

Total number of these is $2(j - i - 1) + 1$, an odd number. Hence $t(\Delta) = -\Delta$ and $\mathrm{sgn}(t) = -1$. □

To work out $\mathrm{sgn}(x)$, $x \in S_n$ here's what we shall do:

- express x as a product of 2-cycles
- use proposition 4.1

Proposition 4.2 *Let $c = (a_1 a_2 \dots a_r)$, an r-cycle. Then c can be expressed as a product of $(r-1)$ 2-cycles.*

Proof Consider the product

$$(a_1 a_r)(a_1 a_{r-1}) \cdots (a_1 a_3)(a_1 a_2).$$

This product sends

$$a_1 \mapsto a_2 \mapsto a_3 \mapsto \cdots \mapsto a_{r-1} \mapsto a_1.$$

Hence the product is equal to c. □

Corollary 4.3 *The signature of an r-cycle is $(-1)^{r-1}$.*

Proof Follows from previous two props. □

Corollary 4.4 *Every $x \in S_n$ can be expressed as a product of 2-cycles.*

Proof From first year, we know that

$$x = c_1 \cdots c_m,$$

a product of disjoint cycles c_i. Each c_i is a product of 2-cycles by 4.2. Hence so is x. □

Proposition 4.5 *Let $x = c_1 \cdots c_m$ a product of disjoint cycles $c_1, \dots, c_m$ of lengths $r_1, \dots, r_m$. Then*

$$\operatorname{sgn}(x) = (-1)^{r_1 - 1} \cdots (-1)^{r_m - 1}.$$

Proof We have

$$\begin{aligned} \operatorname{sgn}(x) &= \operatorname{sgn}(c_1) \cdots \operatorname{sgn}(c_m) \text{ by 4.1(a)} \\ &= (-1)^{r_1-1} \cdots (-1)^{r_m-1} \text{ by 4.3.} \end{aligned}$$

Example (1 2 5 7)(3 4 6)(8 9)(10 12 83)(79 11 26 15) has sgn $= -1$.

Importance of signature

1. We'll use it to define a new family of groups below.
2. Fundamental in the theory of determinants (later).

Definition Define

$$A_n = \{x \in S_n \mid \text{sgn}(x) = 1\},$$

the set of even permutations in S_n. Call A_n the *alternating group* (after showing that it is a group).

Theorem 4.6 *A_n is a subgroup of S_n, of size $\frac{1}{2}n!$.*

Proof (a) A_n is a subgroup:

(1) $e \in A_n$ as $\text{sgn}(e) = 1$.

(2) for $x, y \in A_n$,

$$\begin{aligned} \text{sgn}(x) &= \text{sgn}(y) = 1, \\ \text{sgn}(xy) &= \text{sgn}(x)\text{sgn}(y) = 1, \end{aligned}$$

so $xy \in A_n$,

(3) for $x \in A_n$, we have $\text{sgn}(x) = 1$, so by 4.1(b), $\text{sgn}(x^{-1}) = 1$, i.e. $x^{-1} \in A_n$.

(b) $|A_n| = \frac{1}{2}n!$: Recall that there are right cosets of A_n,

$$A_n = A_n e, A_n(1\ 2) = \{x(1\ 2) \mid x \in A_n\}.$$

These cosets are distinct (as $(1\ 2) \in A_n(1\ 2)$ but $(1\ 2) \notin A_n$), and have equal size (i.e. $|A_n| = |A_n(1\ 2)|$). We show that $S_n = A_n \cup A_n(1\ 2)$: Let $g \in S_n$. If g is even, then $g \in A_n$. If g is odd, then $g(1\ 2)$ is even (as $\text{sgn}(g(1\ 2)) = \text{sgn}(g)\text{sgn}(1\ 2) = 1$), so $g(1\ 2) = x \in A_n$. Then $g = x(1\ 2) \in A_n(1\ 2)$.

So $|A_n| = \frac{1}{2}|S_n| = \frac{1}{2}n!$. □

Examples

1. $A_3 = \{e, (1\ 2\ 3), (1\ 3\ 2)\}$, size $3 = \frac{1}{2}3!$.
2. A_4:

cycle shape	e	(2)	(3)	(4)	$(2,2)$
in A_4?	yes	no	yes	no	yes
no.	1		8		3

Total $|A_4| = 12 = \frac{1}{2}4!$.

3. A_5:

cycle shape	e	(2)	(3)	(4)	(5)	$(2,2)$	$(3,2)$
in A_5?	yes	no	yes	no	yes	yes	no
no.	1		20		24	15	

Total $|A_5| = 60 = \frac{1}{2}5!$.

5 Direct Products

So far, we've seen the following examples of finite groups: C_n, D_{2n}, S_n, A_n. We'll get many more using the following construction.

Recall: if $T_1, T_2, \ldots, T_n$ are sets, the *Cartesian product* $T_1 \times T_2 \times \cdots \times T_n$ is the set consisting of all *n-tuples* $(t_1, t_2, \ldots, t_n)$ with $t_i \in T_i$.

Now let $G_1, G_2, \ldots, G_n$ be groups. Form the Cartesian product $G_1 \times G_2 \times \cdots \times G_n$ and define multiplication on this set by

$$(x_1, \ldots, x_n)(y_1, \ldots, y_n) = (x_1 y_1, \ldots, x_n y_n)$$

for $x_i, y_i \in G_i$.

Definition Call $G_1 \times \cdots \times G_n$ the *direct product* of the groups $G_1, \ldots, G_n$.

Proposition 5.1 *Under above defined multiplication,* $G_1 \times \cdots \times G_n$ *is a group.*

Proof

- *Closure* True by closure in each G_i.
- *Associativity* Using associativity in each G_i,

$$\begin{aligned}
\left[(x_1, \ldots, x_n)(y_1, \ldots, y_n)\right](z_1, \ldots, z_n) &= (x_1 y_1, \ldots, x_n y_n)(z_1, \ldots, z_n) \\
&= ((x_1 y_1) z_1, \ldots, (x_n y_n) z_n) \\
&= (x_1 (y_1 z_1), \ldots, x_n (y_n z_n)) \\
&= (x_1, \ldots, x_n)(y_1 z_1, \ldots, y_n z_n) \\
&= (x_1, \ldots, x_n)\left[(y_1, \ldots, y_n)(z_1, \ldots, z_n)\right].
\end{aligned}$$

- *Identity* is $(e_1, \dots, e_n)$, where e_i is the identity of G_i.
- *Inverse* of $(x_1, \dots, x_n)$ is $(x_1^{-1}, \dots, x_n^{-1})$.

Examples

1. Some new groups: $C_2 \times C_2, C_2 \times C_2 \times C_2, S_4 \times D_{36}, A_5 \times A_6 \times S_{297}, \dots, \mathbb{Z} \times \mathbb{Q} \times S_{13}, \dots$.
2. Consider $C_2 \times C_2$. Elements are $\{(1,1),(1,-1),(-1,1),(-1,-1)\}$. Calling these e, a, b, ab, mult table is

	e	a	b	ab
e	e	a	b	ab
a	a	e	ab	b
b	b	ab	e	a
ab	ab	b	a	e

   $G = C_2 \times C_2$ is abelian and $x^2 = e$ for all $x \in G$.

3. Similarly $C_2 \times C_2 \times C_2$ has elements $(\pm 1, \pm 1, \pm 1)$, size 8, abelian, $x^2 = e$ for all x.

Proposition 5.2 *(a) Size of $G_1 \times \cdots \times G_n$ is $|G_1||G_2| \cdots |G_n|$.*

(b) If all G_i are abelian so is $G_1 \times \cdots \times G_n$.

(c) If $x = (x_1, \dots, x_n) \in G_1 \times \cdots \times G_n$, then order of x is the least common multiple of $o(x_1), \dots, o(x_n)$.

Proof (a) Clear.

(b) Suppose all G_i are abelian. Then

$$\begin{aligned}(x_1, \dots, x_n)(y_1, \dots, y_n) &= (x_1 y_1, \dots, x_n y_n) \\ &= (y_1 x_1, \dots, y_n x_n) \\ &= (y_1, \dots, y_n)(x_1, \dots, x_n).\end{aligned}$$

(c) Let $r_i = o(x_i)$. Recall from M1P2 that $x_i^k = e$ iff $r_i | k$. Let $r = \text{lcm}(r_1, \dots, r_n)$. Then

$$\begin{aligned}x^r &= (x_1^r, \dots, x_n^r) \\ &= (e_1, \dots, e_n) = e.\end{aligned}$$

For $1 \leq s < r$, $r_i \not| s$ for some i. So $x_i^s \neq e$. So

$$x^s = (\dots, x_i^s, \dots) \neq (e_1, \dots, e_n).$$

Hence $r = o(x)$. □

Examples

1. Since cyclic groups C_r are abelian, so are all direct products

$$C_{r_1} \times C_{r_2} \times \cdots \times C_{r_k}.$$

2. $C_4 \times C_2$ and $C_2 \times C_2 \times C_2$ are abelian of size 8. Are they isomorphic?

   Claim: NO.

   Proof Count the number of elements of order 2 :

   In $C_4 \times C_2$ these are $(\pm 1, \pm 1)$ except for $(1, 1)$, so there are 3.

   In $C_2 \times C_2 \times C_2$, all the elements except e have order 2, so there are 7.

   So $C_4 \times C_2 \ncong C_2 \times C_2 \times C_2$.

Proposition 5.3 *If* $\text{hcf}(m, n) = 1$, *then* $C_m \times C_n \cong C_{mn}$.

Proof Let $C_m = \langle \alpha \rangle$, $C_n = \langle \beta \rangle$. So $o(\alpha) = m$, $o(\beta) = n$. Consider

$$x = (\alpha, \beta) \in C_m \times C_n.$$

By 5.2(c), $o(x) = \text{lcm}(m, n) = mn$. Hence cyclic subgroup $\langle x \rangle$ of $C_m \times C_n$ has size mn, so is whole of $C_m \times C_n$. So $C_m \times C_n = \langle x \rangle$ is cyclic and hence $C_m \times C_n \cong C_{mn}$ by 2.2. □

Direct products are fundamental to the theory of abelian groups:

Theorem 5.4 *Every finite abelian group is isomorphic to a direct product of cyclic groups.*

Won't give a proof here. Reference: [Allenby, p. 254].

Examples

1. Abelian groups of size 6: by theorem 5.4, possibilities are $C_6, C_3 \times C_2$. By 5.3, these are isomorphic, so there is only one abelian group of size 6 (up to isomorphism).

2. By 5.4, the abelian groups of size 8 are: C_8, $C_4 \times C_2, C_2 \times C_2 \times C_2$.

   Claim : No two of these are isomorphic.

   Proof

Group	$C_2 \times C_2 \times C_2$	$C_4 \times C_2$	C_8
$\|\{x \mid o(x) = 2\}\|$	7	3	1

   So up to isomorphism, there are 3 abelian groups of size 8.

6 Groups of small size

We'll find *all* groups of size ≤ 7 (up to isomorphism). Useful results:

Proposition 6.1 *If $|G| = p$, a prime, then $G \cong C_p$.*

Proof By corollary of Lagrange, G is cyclic. Hence $G \cong C_p$ by 2.2.

Proposition 6.2 *If $|G|$ is even, then G contains an element of order 2.*

Proof Suppose $|G|$ is even and G has no element of order 2. List the elements of G as follows:

$$e, x_1, x_1^{-1}, x_2, x_2^{-1}, \ldots, x_k, x_k^{-1}.$$

Note that $x_i \neq x_i^{-1}$ since $o(x_i) \neq 2$. Hence $|G| = 2k + 1$, a contradiction. □

Groups of size 1, 2, 3, 5, 7

By 6.1, only such groups are C_1, C_2, C_3, C_5, C_7.

Groups of size 4

Proposition 6.3 *The only groups of size 4 are C_4 and $C_2 \times C_2$.*

Proof Let $|G| = 4$. By Lagrange, every element of G has order 1, 2 or 4. If there exists $x \in G$ of order 4, then $\langle x \rangle$ is cyclic, so $G \cong C_4$. Now suppose $o(x) = 2$ for all $x \neq e$, $x \in G$. So $x^2 = e$ for all $x \in G$.

Let e, x, y be 3 distinct elements of G. If $xy = e$ then $y = x^{-1} = x$, a contradiction; if $xy = x$ then $y = e$, a contradiction; similarly $xy \neq y$. It follows that

$$G = \{e, x, y, xy\}.$$

As above, $yx \neq e, x, y$ hence $yx = xy$. So multiplication table of G is

	e	x	y	xy
e	e	x	y	xy
x	x	e	xy	y
y	y	xy	e	x
xy	xy	y	x	e

This is the same as the table for $C_2 \times C_2$, so $G \cong C_2 \times C_2$. □

Groups of size 6

We know the following groups of size 6: C_6, D_6, S_3. Recall D_6 is the symmetry group of the equilateral triangle and has elements

$$e, \rho, \rho^2, \sigma, \rho\sigma, \rho^2\sigma.$$

satisfying the following equations:

$$\begin{aligned} \rho^3 &= e, \\ \sigma^2 &= e \\ \sigma\rho &= \rho^2\sigma. \end{aligned}$$

The whole multiplication table of D_6 can be worked out using these equations. e.g.

$$\sigma \cdot (\rho\sigma) = \rho^2\sigma\sigma = \rho^2.$$

Proposition 6.4 *Up to isomorphism, the only groups of size 6 are C_6 and D_6.*

Proof Let G be a group with $|G| = 6$. By Lagrange, every element of G has order 1, 2, 3 or 6. If there exists $x \in G$ of order 6, then $G = \langle x \rangle$ is cyclic and therefore $G \cong C_6$ by 2.2. So assume G has no elements of order 6. Then every $x \in G$, $(x \neq e)$ has order 2 or 3. If all have order 2 then $x^2 = e$ for all $x \in G$. So by Sheet 2 Q5, $|G|$ is divisible by 4, a contradiction. We conclude that there exists $x \in G$ with $o(x) = 3$. Also by 6.2, there is an element y of order 2.

Let $H = \langle x \rangle = \{e, x, x^2\}$. Then $y \notin H$ so $Hy \neq H$ and

$$G = H \cup Hy = \{e, x, x^2, y, xy, x^2y\}.$$

What is yx? Well,

$$\left.\begin{array}{lllllll} yx & = & e & \Rightarrow & y & = & x^{-1} \\ yx & = & x & \Rightarrow & y & = & e \\ yx & = & x^2 & \Rightarrow & y & = & x \\ yx & = & y & \Rightarrow & x & = & e \end{array}\right\} \text{ a contradiction.}$$

If $yx = xy$, let's consider the order of xy:

$$(xy)^2 = xyxy = xxyy \ \ (\text{as } yx = xy) \ \ = x^2y^2 = x^2.$$

Similarly

$$(xy)^3 = x^3y^3 = y \neq e.$$

So xy does not have order 2 or 3, a contradiction. Hence $yx \neq xy$. We conclude that $yx = x^2y$.

At this point we know the following:

- $G = \{e, x, x^2, y, xy, x^2y\}$,
- $x^3 = e$, $x^2 = e$, $yx = x^2y$.

In exactly the same way as for D_6, can work out the whole multiplication table for G using these equations. It will be the same as the table for D_6 (with x, y instead of ρ, σ). So $G \cong D_6$. $\square$

Remark Note that $|S_3| = 6$, and $S_3 \cong D_6$.

Summary

Proposition 6.5 *Up to isomorphism, the groups of size ≤ 7 are*

Size	*Groups*
1	C_1
2	C_2
3	C_3
4	C_4, $C_2 \times C_2$
5	C_5
6	C_6, D_6
7	C_7

Remarks on larger sizes

Size 8: here are the groups we know:

Abelian $C_8, C_4 \times C_2, C_2 \times C_2 \times C_2$,

Non-abelian D_8.

Any others? Yes, the *quaternion* group Q_8:

Define matrices

$$A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Check equations:

$$A^4 = I, \quad B^4 = I, \quad A^2 = B^2, BA = A^4 B.$$

Define

$$\begin{aligned} Q_8 &= \{A^r B^s \mid r, s \in \mathbb{Z}\} \\ &= \{A^m B^n \mid 0 \le m \le 3,\ 0 \le n \le 1\}. \end{aligned}$$

Sheet 3 Q5: $|Q_8| = 8$. Q_8 is a subgroup of $GL(2, \mathbb{C})$ and is not abelian and $Q_8 \not\cong D_8$. Call Q_8 the *quaternion group.* Sheet 3 Q7: The only non-abelian groups of size 8 are D_8 and Q_8. Yet more info:

Size	Groups
9	only abelian (Sh3 Q4)
10	C_{10}, D_{10}
11	C_{11}
12	abelian, D_{12}, A_4 + one more
13	C_{13}
14	C_{14}, D_{14}
15	C_{15}
16	14 groups

7 Homomorphisms, normal subgroups and factor groups

Homomorphisms are functions between groups which "preserve multiplication".

Definition Let G, H be groups. A function $\phi : G \to H$ is a *homomorphism* if $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in G$.

Note that an isomorphism is a homomorphism which is a *bijection.*

Examples

1. G, H any groups. Define $\phi : G \to H$ by
$$\phi(x) = e_H \forall x \in G$$
Then ϕ is a homomorphism since $\phi(xy) = e_H = e_H e_H = \phi(x)\phi(y)$.

2. Recall the signature function sgn : $S_n \to C_2$. By 4.1(a), $\mathrm{sgn}(xy) = \mathrm{sgn}(x)\mathrm{sgn}(y)$, so sgn is a homomorphism.

3. Define $\phi : (\mathbb{R}, +) \to (\mathbb{C}^*, \times)$ by
$$\phi(x) = e^{2\pi i x} \forall x \in \mathbb{R}.$$
Then $\phi(x + y) = e^{2\pi i(x+y)} = e^{2\pi i x} e^{2\pi i y} = \phi(x)\phi(y)$, so ϕ is a homomorphism.

4. Define $\phi : D_{2n} \to C_2$ (writing $D_{2n} = \left\{e, \rho, \ldots, \rho^{n-1}, \sigma, \rho\sigma, \ldots, \rho^{n-1}\sigma\right\}$) by
$$\phi(\rho^r \sigma^s) = (-1)^s.$$
(so ϕ sends rotations to $+1$ and reflections to -1). Then ϕ is a homomorphism since:
$$\begin{aligned} \phi\left((\rho^r\sigma^s)(\rho^t\sigma^u)\right) &= \phi(\rho^{r \pm t}\sigma^{s+u}) \\ &= (-1)^{s+u} = \phi(\rho^r\sigma^s)\phi(\rho^r\sigma^u). \end{aligned}$$

Proposition 7.1 *Let $\phi : G \to H$ be a homomorphism*

(a) $\phi(e_G) = e_H$

(b) $\phi(x^{-1}) = \phi(x)^{-1}$ *for all $x \in G$.*

(c) $o(\phi(x))$ *divides* $o(x)$ *for all $x \in G$.*

Proof (a) Note that $\phi(e_G) = \phi(e_G e_G) = \phi(e_G)\phi(e_G)$. Multiply by $\phi(e_G)^{-1}$ to get $e_H = \phi(e_G)$.

(b) By (a), $e_H = \phi(e_G) = \phi(xx^{-1}) = \phi(x)\phi(x^{-1})$. So $\phi(x^{-1}) = \phi(x)^{-1}$.

(c) Let $r = o(x)$. Then

$$\phi(x)^r = \phi(x) \cdots \phi(x) = \phi(x \cdots x) = \phi(x^r) = \phi(e_G) = e_H.$$

Hence $o(\phi(x))$ divides r. □

Definition Let $\phi : G \to H$ be homomorphism. The *image* of ϕ is

$$\mathrm{Im}\phi = \phi(\mathrm{G}) = \{\phi(\mathrm{x}) \mid \mathrm{x} \in \mathrm{G}\} \subseteq \mathrm{H}.$$

Proposition 7.2 *If $\phi : G \to H$ is a homomorphism, then* $\mathrm{Im}\phi$ *is a subgroup of H.*

Proof

(1) $e_H \in \mathrm{Im}\phi$ since $e_H = \phi(e_G)$.

(2) Let $g, h \in \mathrm{Im}\phi$. Then $g = \phi(x)$ and $h = \phi(y)$ for some $x, y \in G$, so $gh = \phi(x)\phi(y) = \phi(xy) \in \mathrm{Im}\phi$.

(3) Let $g \in \mathrm{Im}\phi$. Then $g = \phi(x)$ for some $x \in G$. So $g^{-1} = \phi(x)^{-1} = \phi(x^{-1}) \in \mathrm{Im}\phi$.

Hence $\mathrm{Im}\phi$ is a subgroup of H. □

Examples

1. Is there a homomorphism $\phi : S_3 \to C_3$? Yes, $\phi(x) = 1$ for all $x \in S_3$. For this homomorphism, $\mathrm{Im}\phi = \{1\}$.

2. Is there a homomorphism $\phi : S_3 \to C_3$ such that $\mathrm{Im}\phi = \mathrm{C}_3$?

   To answer this, suppose $\phi : S_3 \to C_3$ is a homomorphism. Consider $\phi(1\ 2)$. By 7.1(c), $\phi(1\ 2)$ has order dividing $o(1\ 2) = 2$. As $\phi(1\ 2) \in C_3$, this implies that $\phi(1\ 2) = 1$. Similarly $\phi(1\ 3) = \phi(2\ 3) = 1$. Hence

   $$\phi(1\ 2\ 3) = \phi\left((1\ 3)(1\ 2)\right) = \phi(1\ 3)\phi(1\ 2) = 1$$

   and similarly $\phi(1\ 3\ 2) = 1$. We've shown that

   $$\phi(x) = 1 \forall x \in S_3.$$

   So there is no surjective homomorphism $\phi : S_3 \to C_3$.

Kernels

Definition Let $\phi : G \to H$ be a homomorphism. Then *kernel* of ϕ is

$$\mathrm{Ker}\phi = \{\mathrm{x} \in \mathrm{G} \mid \phi(\mathrm{x}) = \mathrm{e_H}\}.$$

Examples

1. If $\phi : G \to H$ is $\phi(x) = e_H$ for all $x \in G$, then $\mathrm{Ker}\phi = \mathrm{G}$.
2. For $\mathrm{sgn} : S_n \to C_2$,

$$\mathrm{Ker(sgn)} = \{\mathrm{x} \in \mathrm{S_n} \mid \mathrm{sgn(x)} = 1\} = \mathrm{A_n}, \text{ the alternating group.}$$

3. If $\phi : (\mathbb{R}, +) \to (\mathbb{C}^*, \times)$ is $\phi(x) = e^{2\pi i x}$ for all $x \in \mathbb{R}$, then

$$\mathrm{Ker}\phi = \left\{\mathrm{x} \in \mathbb{R} \mid \mathrm{e}^{2\pi \mathrm{ix}} = 1\right\} = \mathbb{Z}.$$

4. Let $\phi : D_{2n} \to C_2$ be given by $\phi(\rho^r \sigma^s) = (-1)^s$. Then $\mathrm{Ker}\phi = \langle \rho \rangle$.

Proposition 7.3 *If $\phi : G \to H$ is a homomorphism, then* $\mathrm{Ker}\phi$ *is a subgroup of G.*

Proof

(1) $e_G \in \mathrm{Ker}\phi$ as $\phi(e_G) = e_H$ by 7.1.

(2) $x, y \in \mathrm{Ker}\phi$ then $\phi(x) = \phi(y) = e_H$, so $\phi(xy) = \phi(x)\phi(y) = e_H$; i.e. $xy \in \mathrm{Ker}\phi$.

(3) $x \in \mathrm{Ker}\phi$ then $\phi(x) = e_H$, so $\phi(x)^{-1} = \phi(x^{-1}) = e_H$, so $x^{-1} \in \mathrm{Ker}\phi$. □

In fact, $\mathrm{Ker}\phi$ is a very special type of subgroup of G known as a *normal* subgroup.

Normal subgroups

Definition Let G be a group, and $N \subseteq G$. We say N is a *normal subgroup* of G if

(1) N is a subgroup of G,

(2) $g^{-1}Ng = N$ for all $g \in G$, where $g^{-1}Ng = \{g^{-1}ng \mid n \in N\}$.

If N is a normal subgroup of G, write $N \lhd G$.

Examples

1. G any group. Subgroup $\langle e \rangle = \{e\} \lhd G$ as $g^{-1}eg = e$ for all $g \in G$. Also subgroup G itself is normal, i.e. $G \lhd G$, as $g^{-1}Gg = G$ for all $g \in G$.

Next lemma makes condition (2) a bit easier to check.

Lemma 7.4 *Let N be a subgroup of G. Then $N \lhd G$ if and only if $g^{-1}Ng \subseteq N$ for all $g \in G$.*

Proof

$\Rightarrow$ Clear.

$\Leftarrow$ Suppose $g^{-1}Ng \subseteq N$ for all $g \in G$. Let $g \in G$. Then

$$g^{-1}Ng \subseteq N.$$

Using g^{-1} instead, we get $(g^{-1})^{-1}Ng^{-1} \subseteq N$, hence

$$gNg^{-1} \subseteq N.$$

Hence $N \subseteq g^{-1}Ng$. Therefore $g^{-1}Ng = N$. □

Examples (1) We show that $A_n \lhd S_n$. Need to show that

$$g^{-1}A_ng \subseteq A_n \forall g \in S_n$$

(this will show $A_n \lhd S_n$ by 7.4).

For $x \in A_n$, using 4.1 we have

$$\text{sgn}(g^{-1}xg) = \text{sgn}(g^{-1})\text{sgn}(x)\text{sgn}(g) = \text{sgn}(g^{-1}) \cdot 1 \cdot \text{sgn}(g) = 1.$$

So $g^{-1}xg \in A_n$ for all $x \in A_n$. Hence

$$g^{-1}A_ng \subseteq A_n.$$

So $A_n \lhd S_n$.

(2) Let $G = S_3$, $N = \langle(1\ 2)\rangle = \{e, (1\ 2)\}$. Is $N \lhd G$? Well,

$$(1\ 3)^{-1}(1\ 2)(1\ 3) = (1\ 3)(1\ 2)(1\ 3) = (2\ 3) \notin N.$$

So $(1\ 3)^{-1}N(1\ 3) \neq N$ and $N \not\lhd S_3$.

(3) If G is abelian, then *all* subgroups N of G are normal since for $g \in G$, $n \in N$,

$$g^{-1}ng = g^{-1}gn = n,$$

and hence $g^{-1}Ng = N$.

(4) Let $D_{2n} = \{e, \rho, \dots, \rho^{n-1}, \sigma, \rho\sigma, \dots, \rho^{n-1}\sigma\}$. Fix an integer r. Then

$$\langle \rho^r \rangle \lhd D_{2n}.$$

Proof – sheet 4. (key: magic equation $\sigma\rho = \rho^{-1}\sigma, \dots, \sigma\rho^n = \rho^{-n}\sigma$).

Proposition 7.5 *If $\phi : G \to H$ is a homomorphism, then* Ker$\phi \lhd$ G.

Proof Let $K = \text{Ker}\phi$. By 7.3 K is a subgroup of G. Let $g \in G$, $x \in K$. Then

$$\phi(g^{-1}xg) = \phi(g^{-1})\phi(x)\phi(g) = \phi(g)^{-1}e_H\phi(g) = e_H.$$

So $g^{-1}xg \in \text{Ker}\phi = \text{K}$. This shows $g^{-1}Kg \subseteq K$. So $K \lhd G$. □

Examples

1. We know that sgn : $S_n \to C_2$ is a homomorphism, with kernel A_n. So $A_n \lhd S_n$ by 7.5.

2. Know $\phi : D_{2n} \to C_2$ defined by $\phi(\rho^r\sigma^s) = (-1)^s$ is a homomorphism with kernel $\langle\rho\rangle$. So $\langle\rho\rangle \lhd D_{2n}$.

3. Here's a different homomorphism $\alpha : D_8 \to C_2$ where

$$\alpha(\rho^r\sigma^s) = (-1)^r.$$

This is a homomorphism, as

$$\begin{aligned} \alpha((\rho^r\sigma^s)(\rho^t\sigma^u)) &= \alpha(\rho^{r\pm t}\sigma^{s+u}) \\ &= (-1)^{r\pm t} = (-1)^r \cdot (-1)^t \\ &= \alpha(\rho^r\sigma^s)\alpha(\rho^t\sigma^u). \end{aligned}$$

The kernel of α is

$$\text{Ker}\alpha = \{\rho^{\text{r}}\sigma^{\text{s}} \mid \text{r even}\} = \left\{\text{e}, \rho^2, \sigma, \rho^2\sigma\right\}.$$

Hence

$$\left\{e, \rho^2, \sigma, \rho^2\sigma\right\} \lhd D_8.$$

Factor groups

Let G be a group, N a subgroup of G. Recall that there are exactly $\frac{|G|}{|N|}$ different right cosets Nx $(x \in G)$. Say

$$Nx_1, Nx_2, \ldots, Nx_r$$

where $r = \frac{|G|}{|N|}$. Aim is to make this set of right cosets into a group in a natural way. Here is a "natural" definition of multiplication of these cosets:

$$(Nx)(Ny) = N(xy). \tag{33}$$

Does this definition make sense? To make sense, we need:

$$\left.\begin{array}{l} Nx = Nx' \\ Ny = Ny' \end{array}\right\} \Rightarrow Nxy = Nx'y'$$

for all $x, y, x', y' \in G$. This property may or may not hold.

Example $G = S_3$, $N = \langle(1\ 2)\rangle = \{e, (1\ 2)\}$. The 3 right cosets of N in G are

$$N = Ne, N(1\ 2\ 3), N(1\ 3\ 2).$$

Also

$$\begin{array}{lcl} N & = & N(1\ 2) \\ N(1\ 2\ 3) & = & N(1\ 2)(1\ 2\ 3) = N(2\ 3) \\ N(1\ 3\ 2) & = & N(1\ 2)(1\ 3\ 2) = N(1\ 3). \end{array}$$

According to (33),

$$\left(N(1\ 2\ 3)\right)\left(N(1\ 2\ 3)\right) = N(1\ 2\ 3)(1\ 2\ 3) = N(1\ 3\ 2).$$

But (33) also says that

$$\left(N(2\ 3)\right)\left(N(2\ 3)\right) = N(2\ 3)(2\ 3) = Ne.$$

So (33) makes no sense in this example.

How do we make (33) make sense? The condition is that $N \lhd G$. Key is to prove the following:

Proposition 7.6 *Let $N \lhd G$. Then for $x_1, x_2, y_1, y_2 \in G$*

$$\left.\begin{array}{l} Nx_1 = Nx_2 \\ Ny_1 = Ny_2 \end{array}\right\} \Rightarrow Nx_1y_1 = Nx_2y_2.$$

(Hence definition of multiplication of cosets in (33) makes sense when $N \lhd G$.)

To prove this we need a definition and a lemma: for H a subgroup of G and $x \in G$ define the *left coset*

$$xH = \{xh : h \in H\}.$$

Lemma 7.7 *Suppose $N \lhd G$. Then $xH = Hx$ for all $x \in G$.*

Proof Let $h \in H$. As $H \lhd G$, $xHx^{-1} = H$, and so $xhx^{-1} = h' \in H$. Then $xh = h'x \in Hx$. This shows that $xH \subseteq Hx$. Similarly we see that $Hx \subseteq xH$, hence $xH = Hx$. □

Proof of Prop 7.6

Let $N \lhd G$. Suppose $Nx_1 = Nx_2$ and $Ny_1 = Ny_2$. Then

$$\begin{array}{lll} Nx_1y_1 & = Nx_2y_1 & \text{as } Nx_1 = Nx_2 \\ & = x_2Ny_1 & \text{by Prop 7.7} \\ & = x_2Ny_2 & \text{as } Ny_1 = Ny_2 \\ & = Nx_2y_2 & \text{by Prop 7.7.}\square \end{array}$$

So we have established that when $N \lhd G$, the definition of multiplication of cosets

$$(Nx)(Ny) = Nxy$$

for $x, y \in G$ makes sense.

Theorem 7.8 *Let $N \lhd G$. Define G/N to be the set of all right cosets Nx $(x \in G)$. Define multiplication on G/N by*

$$(Nx)(Ny) = Nxy.$$

Then G/N is a group under this multiplication.

Proof

Closure obvious.

Associativity Using associativity in G

$$\begin{aligned}(NxNy)Nz &= (Nxy)Nz \\ &= N(xy)z \\ &= Nx(yz) \\ &= (Nx)(Nyz) \\ &= Nx(NyNz).\end{aligned}$$

Identity is $Ne = N$, since $NxNe = Nxe = Nx$ and $NeNx = Nex = Nx$.

Inverse of Nx is Nx^{-1}, as $NxNx^{-1} = Nxx^{-1} = Ne$, the identity.

Definition The group G/N is called the *factor group* of G by N.

Note that

$$|G/N| = \frac{|G|}{|N|}.$$

Examples

1. $A_n \lhd S_n$. Since $\frac{|S_n|}{|A_n|} = 2$, the factor group S_n/A_n has 2 elements

$$A_n, A_n(1\ 2).$$

So $S_n/A_n \cong C_2$. Note: in the group S_n/A_n the identity is the coset A_n and the non identity element $A_n(1\ 2)$ has order 2 as

$$(A_n(1\ 2))^2 = A_n(1\ 2)A_n(1\ 2) = A_n(1\ 2)(1\ 2) = A_n.$$

2. G any group. We know that $G \lhd G$. What is the factor group G/G? Ans: G/G has 1 element, the identity coset G. So $G/G \cong C_1$.

Also $\langle e \rangle = \{e\} \lhd G$. What is $G/\langle e \rangle$? Coset $\langle e \rangle g = \{g\}$, and multiplication

$$(\langle e \rangle g)(\langle e \rangle h) = \langle e \rangle gh.$$

So $G/\langle e \rangle \cong G$ (isomorphism $g \mapsto \langle e \rangle g$).

3. $G = D_{12} = \{e, \rho, \ldots, \rho^5, \sigma, \sigma\rho, \ldots, \sigma\rho^5\}$ where $\rho^6 = \sigma^2 = e$, $\sigma\rho = \rho^{-1}\sigma$.

(a) Know that $\langle\rho\rangle \lhd D_{12}$. Factor group $D_{12}/\langle\rho\rangle$ has 2 elements $\langle\rho\rangle, \langle\rho\rangle\sigma$ so $D_{12}/\langle\rho\rangle \cong C_2$.

(b) Know also that $\langle\rho^2\rangle = \{e, \rho^2, \rho^4\} \lhd D_{12}$. So $D_{12}/\langle\rho^2\rangle$ has 4 elements, so

$$D_{12}/\langle\rho^2\rangle \cong C_4 \text{ or } C_2 \times C_2.$$

Which? Well, let $N = \langle\rho^2\rangle$. The 4 elements of D_{12}/N are

$$N, N\rho, N\sigma, N\rho\sigma.$$

We work out the order of each of these elements of D_{12}/N:

$$\begin{aligned} (N\rho)^2 &= N\rho N\rho = N\rho^2 \\ &= N, \\ (N\sigma)^2 &= N\sigma N\sigma = N\sigma^2 \\ &= N, \\ (N\rho\sigma)^2 &= N(\rho\sigma)^2 \\ &= N. \end{aligned}$$

So all non-identity elements of D_{12}/N have order 2, hence $D_{12}/\langle\rho\rangle \cong C_2 \times C_2$.

(c) Also $\langle\rho^3\rangle = \{e, \rho^3\} \lhd D_{12}$. Factor group $D_{12}\langle\rho^3\rangle$ has 6 elements so is $\cong C_6$ or D_6. Which? Let $M = \langle\rho^3\rangle$. The 6 elements of D_{12}/M are

$$M, M\rho, M\rho^2, M\sigma, M\rho\sigma, M\rho^2\sigma.$$

Let $x = M\rho$ and $y = M\sigma$. Then

$$\begin{aligned} x^3 &= (M\rho)^3 = M\rho M\rho M\rho = M\rho^3 \\ &= M, \\ y^2 &= (M\sigma)^2 = M\sigma^2 \\ &= M, \\ yx &= M\sigma M\rho = M\sigma\rho = M\rho^{-1}\sigma = M\rho^{-1}M\sigma \\ &= x^{-1}y. \end{aligned}$$

So $D_{12}/M = \{\text{identity}, x, x^2, y, xy, x^2y\}$ and $x^3 = y^2 =$ identity,$yx = x^{-1}y$. So $D_{12}/\langle\rho^3\rangle \cong D_6$.

Here's a result tying all these topics together:

Theorem 7.9 (First Isomorphism Theorem) *Let $\phi : G \to H$ be a homomorphism. Then*

$$G/\text{Ker}\phi \cong \text{Im}\phi.$$

Proof Let $K = \text{Ker}\phi$. So G/K is the group consisting of the cosets Kx $(x \in G)$ with multiplication $(Kx)(Ky) = Kxy$. We want to define a "natural" function $G/K \to \text{Im}\phi$. Obvious choice is the function $Kx \mapsto \phi(x)$ for $x \in G$. To show this is a function, need to prove:

Claim 1. If $Kx = Ky$, then $\phi(x) = \phi(y)$.

To prove this, suppose $Kx = Ky$. Then $xy^{-1} \in K$ (as $x \in Kx \Rightarrow x = ky$ for some $k \in K \Rightarrow xy^{-1} = k \in K$). Hence $xy^{-1} \in K = \text{Ker}\phi$, so

$$\begin{aligned} & \phi(xy^{-1}) = e \\ \Rightarrow\ & \phi(x)\phi(y^{-1}) = e \\ \Rightarrow\ & \phi(x)\phi(y)^{-1} = e \\ \Rightarrow\ & \phi(x) = \phi(y). \end{aligned}$$

By Claim 1, we can define a function $\alpha : G/K \to \text{Im}\phi$ by

$$\alpha(Kx) = \phi(x)$$

for all $x \in G$.

Claim 2. α is an isomorphism.

Here is a proof of this claim.

(1) α is surjective: for if $\phi(x) \in \text{Im}\phi$ then $\phi(x) = \alpha(Kx)$.

(2) α is injective:

$$\begin{aligned} & \alpha(Kx) = \alpha(Ky) \\ \Rightarrow\ & \phi(x) = \phi(y) \\ \Rightarrow\ & \phi(x)\phi(y)^{-1} = e \\ \Rightarrow\ & \phi(xy^{-1}) = e, \end{aligned}$$

so $xy^{-1} \in \text{Ker}\phi = \text{K}$ and so $Kx = Ky$.

(3) Finally

$$\begin{aligned} \alpha((Kx)(Ky)) &= \alpha(Kxy) \\ &= \phi(xy) \\ &= \phi(x)\phi(y) \\ &= \alpha(Kx)\alpha(Ky). \end{aligned}$$

Hence α is an isomorphism.

This completes the proof that $G/K \cong \text{Im}\phi$. □

Corollary 7.10 *If $\phi : G \to H$ is a homomorphism, then*

$$|G| = |\text{Ker}\phi| \cdot |\text{Im}\phi|.$$

One can think of this as the group theoretic version of the rank-nullity theorem.

Examples

1. Homomorphism sgn : $S_n \to C_2$. By 7.9

$$S_n/\mathrm{Ker}(\mathrm{sgn}) \cong \mathrm{Im}(\mathrm{sgn}),$$

so

$$S_n/A_n \cong C_2.$$

2. Homomorphism $\phi : (\mathbb{R}, +) \to (\mathbb{C}^*, \times)$

$$\phi(x) = e^{2\pi i x}.$$

Here

$$\begin{aligned} \mathrm{Ker}\phi &= \left\{x \in \mathbb{R} \mid e^{2\pi i x} = 1\right\} \\ &= \mathbb{Z}, \\ \mathrm{Im}\phi &= \left\{e^{2\pi i x} \mid x \in \mathbb{R}\right\} \\ &= T \text{ the unit circle.} \end{aligned}$$

So $\mathbb{R}/\mathbb{Z} \cong T$.

3. Is there a surjective homomorphism ϕ from S_3 onto C_3? Shown previously – No.

   Here's a better way to see this: suppose there exist such ϕ. Then $\mathrm{Im}\phi = \mathrm{C}_3$, so by 7.9, $S_3/\mathrm{Ker}\phi \cong \mathrm{C}_3$. So $\mathrm{Ker}\phi$ is a normal subgroup of S_3 of size 2. But S_3 has no normal subgroups of size 2 (they are $\langle(1\ 2)\rangle$, $\langle(1\ 3)\rangle$, $\langle(2\ 3)\rangle$).

Given a homomorphism $\phi : G \to H$, we know $\mathrm{Ker}\phi \lhd \mathrm{G}$. Converse question: Given a normal subgroup $N \lhd G$, does there exist a homomorphism with kernel N? Answer is YES:

Proposition 7.11 *Let G be a group and $N \lhd G$. Define $H = G/N$. Let $\phi : G \to H$ be defined by*

$$\phi(x) = Nx$$

for all $x \in G$. Then ϕ is a homomorphism and $\mathrm{Ker}\phi = \mathrm{N}$.

Proof First, $\phi(xy) = Nxy = (Nx)(Ny) = \phi(x)\phi(y)$, so ϕ is a homomorphism. Also

$$x \in \text{Ker}\phi \Leftrightarrow \phi(\text{x}) = \text{e}_\text{H} \Leftrightarrow \text{Nx} = \text{N} \Leftrightarrow \text{x} \in \text{N}.$$

Hence $\text{Ker}\phi = \text{N}$. $\square$

Example From a previous example, we know $\left\langle \rho^2 \right\rangle = \left\{e, \rho^2, \rho^4\right\} \lhd D_{12}$. We showed that $D_{12} \left\langle \rho^2 \right\rangle \cong C_2 \times C_2$. So by 7.11, the function $\phi(x) = \left\langle \rho^2 \right\rangle x$ ($x \in D_{12}$) is a homomorphism $D_{12} \to C_2 \times C_2$ which is surjective, with kernel $\left\langle \rho^2 \right\rangle$.

Summary

There is a correspondence

$$\{\text{normal subgroups of } G\} \leftrightarrow \{\text{homomorphisms of } G\}.$$

For $N \lhd G$ there is a homomorphism $\phi : G \to G/N$ with $\text{Ker}\phi = \text{N}$. For a homomorphism ϕ, $\text{Ker}\phi$ is a normal subgroup of G.

Given G, to find all H such that there exist a surjective homomorphism $G \to H$:

(1) Find all normal subgroups of G.

(2) The possible H are the factor groups G/N for $N \lhd G$.

Example: $G = S_3$.

(1) Normal subgroups of G are

$$\langle e \rangle, G, A_3 = \langle (1\ 2\ 3) \rangle$$

(cyclic subgroups of size 2 $\langle (i\ j) \rangle$ are not normal).

(2) Factor groups:

$$S_3/\langle e \rangle \cong S_3, \quad S_3/S_3 \cong C_1, \quad S_3/A_3 \cong C_2$$

8 Symmetry groups in 3 dimensions

These are defined similarly to symmetry groups in 2 dimensions, see chapter 2. An *isometry* of $\mathbb{R}^3$ is a bijection $f : \mathbb{R}^3 \to \mathbb{R}^3$ such that $d(x,y) = d(f(x), f(y))$ for all $x, y \in \mathbb{R}^3$.

Examples of isometries are: rotation about an axis, reflection in a plane, translation.

As in 2.1, the set of all isometries of $\mathbb{R}^3$, under composition, forms a group $I(\mathbb{R}^3)$. For $\Pi \subseteq \mathbb{R}^3$, the *symmetry group* of Π is $G(\Pi) = \{g \in I(\mathbb{R}^3) \mid g(\Pi) = \Pi\}$. There exist many interesting symmetry groups in $\mathbb{R}^3$. Some of the most interesting are the symmetry groups of the Platonic solids: tetrahedron, cube, octahedron, icosahedron, dodecahedron.

Example: *The regular tetrahedron*

Let Π be regular tetrahedron in $\mathbb{R}^3$, and let $G = G(\Pi)$.

- *Rotations in* G: Let R be the set of rotations in G. Some elements of R:

  (1) e,

  (2) rotations of order 3 fixing one corner: these are

  $$\rho_1, \rho_1^2, \rho_2, \rho_2^2, \rho_3, \rho_3^2, \rho_4, \rho_4^2$$

  (where ρ_i fixes corner i),

  (3) rotations of order 2 about an axis joining the mid-points of opposite sides

  $$\rho_{12,34}, \rho_{13,24}, \rho_{14,23}.$$

  So $|R| \geq 12$. Also $|R| \leq 12$: can rotate to get any face i on bottom (4 choices). If i is on the bottom, only 3 possible configurations. Hence $|R| \leq 4 \cdot 3 = 12$. Hence $|R| = 12$.

  Claim 1: $R \cong A_4$.

  To see this, observe that each rotation $r \in R$ gives a permutation of the corners $1, 2, 3, 4$, call it π_r:

  $$\begin{array}{lll} e & \to & \pi_e = \text{ identity permutation} \\ \rho_i, \rho_i^2 & \to & \text{all 8 3-cycles in } S_4 \ (1\ 2\ 3), (1\ 3\ 2), \ldots \\ \rho_{12,34} & \to & (1\ 2)(3\ 4) \\ \rho_{13,24} & \to & (1\ 3)(2\ 4) \\ \rho_{14,23} & \to & (1\ 4)(2\ 3). \end{array}$$

Notice that $\{\pi_r \mid r \in R\}$ consists of all the 12 *even* permutations in S_4, i.e. A_4. The map $r \mapsto \pi_r$ is an isomorphism $R \to A_4$. So $R \cong A_4$.

Claim 2: The symmetry group G is S_4.

Obviously G contains a reflection σ with corresponding permutation $\pi_\sigma = (1\ 2)$. So G contains

$$R \cup R\sigma.$$

So $|G| \geq |R| + |R\sigma| = 24$. On the other hand, each $g \in G$ gives a unique permutation $\pi_g \in S_4$, so $|G| \leq |S_4| = 24$. So $|G| = 24$ and the map $g \mapsto \pi_g$ is an isomorphism $G \to S_4$.

9 Counting using groups

Consider the following problem. Colour edges of an equilateral triangle with 2 colours R, B. How many distinguishable colourings are there?

Answer: There are 8 colourings altogether:

(1) all the edges red – RRR,

(2) all the edges blue – BBB,

(3) two reds and a blue – RRB,RBR,BRR,

(4) two blues and a red – BBR,BRB,RBB.

Clearly there are 4 distinguishable colourings. Point: Two colourings are not distinguishable iff there exists a symmetry of the triangle sending one to the other.

To bring groups into the picture: call C the set of all 8 colorings. So

$$C = \{RRR, \ldots, RBB\}.$$

Let G be the symmetry group of the equilateral triangle, $D_6 = \left\{e, \rho, \rho^2, \sigma, \rho\sigma, \rho^2\sigma\right\}$. Each element of D_6 gives a permutation of C, e.g. ρ gives the permutation $(RRR)\,(BBB)\,(RRB\ RBR\ BRR)\,(BBR\ BRB\ RBB)$.

Divide the set C into subsets called *orbits* of G: two colourings c, d are in the same orbit if there exists $g \in D_6$ sending c to d. The orbits are the sets (1) - (4) above. The number of distinguishable colourings is equal to the number of orbits of G.

General situation

Suppose we have a set S and a group G consisting of some permutations of S (e.g. $S = C$, $G = D_6$ above). Partition S into *orbits* of G, by saying that two elements $s, t \in S$ are in the same orbit iff there exists a $g \in G$ such that $g(s) = t$. How many orbits are there?

Lemma 9.1 (Burnside's Counting Lemma) *For $g \in G$, define*

$$\begin{aligned} \text{fix}(g) &= \textit{number of elements of } S \textit{ fixed by } g \\ &= |\{s \in S \mid g(s) = s\}|. \end{aligned}$$

Then

$$\textit{number of orbits of } G = \frac{1}{|G|} \sum_{g \in G} \text{fix}(g).$$

I won't give a proof. Look it up in the recommended book by Fraleigh if you are interested.

Examples

(1) C = set of 8 colourings of the equilateral triangle. $G = D_6$. Here are the values of fix(g):

g	e	ρ	ρ^2	σ	$\rho\sigma$	$\rho^2\sigma$
fix(g)	8	2	2	4	4	4

By 9.1, number of orbits is $\frac{1}{6}(8 + 2 + 2 + 4 + 4 + 4) = 4$.

(2) 6 beads coloured R, R, W, W, Y, Y are strung on a necklace. How many distinguishable necklaces are there?

Each necklace is a colouring of a regular hexagon. Two colourings are indistinguishable if there is a rotation or reflection sending one to the other (a reflection is achieved by turning the hexagon upside down). Let D be the set of colourings of the hexagon and $G = D_{12}$.

g	e	ρ	ρ^2	ρ^3	ρ^4	ρ^5
fix(g)	$\binom{6}{2} \times \binom{4}{2}$	0	0	6	0	0

g	σ	$\rho\sigma$	$\rho^2\sigma$	$\rho^3\sigma$	$\rho^4\sigma$	$\rho^5\sigma$
fix(g)	6	6	6	6	6	6

So by 9.1

$$\text{number of orbits } = \frac{1}{12}(90+42) = 11.$$

So the number of distinguishable necklaces is 11.

(3) Make a tetrahedral die by putting 1, 2, 3, 4 on the faces. How many distinguishable dice are there?

Each die is a colouring (colours 1, 2, 3, 4) of a regular tetrahedron. Two such colourings are indistinguishable if there exists a *rotation* of the tetrahedron sending one to the other. Let E be the set of colourings, and G = rotation group of tetrahedron (so $|G| = 12$, $G \cong A_4$ by Chapter 8). Here for $g \in G$

$$\text{fix}(g) = \begin{cases} 24 & \text{if } g = e, \\ 0 & \text{if } g \neq e. \end{cases}$$

So by 9.1, number of orbits is $\frac{1}{12}(24) = 2$. So there are 2 distinguishable tetrahedral dice.

Part(B): Linear Algebra

Revision from M1GLA:

Matrices, linear equations; Row operations; echelon form; Gaussian elimination; Finding inverses; 2×2, 3×3 determinants; eigenvalues and eigenvectors; diagonalization.

From M1P2:

Vector spaces; subspaces; spanning sets; linear independence; basis, dimension; rank, col-rank = row-rank; linear transformations; kernel, image, rank-nullity theorem; matrix $[T]_B$ of a linear transformation with respect to a basis B; diagonalization, change of basis .

10 Determinants

In M1GLA, we defined determinants of 2×2 and 3×3 matrices. Recall the definition of 3×3 determinant:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{23} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

This expression has 6 terms. Each term

(1) is a product of 3 entries, one from each column,

(2) has a sign $\pm$.

Property (1) gives for each term a *permutation* of $\{1, 2, 3\}$, sending $i \mapsto j$ if a_{ij} is present.

Term	Permutation	Sign
$a_{11}a_{22}a_{33}$	e	$+$
$a_{11}a_{23}a_{32}$	(2 3)	$-$
$a_{12}a_{21}a_{33}$	(1 2)	$-$
$a_{12}a_{23}a_{31}$	(1 2 3)	$+$
$a_{13}a_{21}a_{32}$	(1 3 2)	$+$
$a_{13}a_{22}a_{31}$	(1 3)	$-$

Notice:

- the sign is sgn(permutation),

- all 6 permutations in S_3 are present.

So

$$|A| = \sum_{\pi \in S_3} \text{sgn}(\pi) \cdot a_{1,\pi(1)} a_{2,\pi(2)} a_{3,\pi(3)}.$$

Here's a general definition:

Definition Let $A = (a_{ij})$ be $n \times n$. Then the *determinant* of A is

$$\det(A) = |A| = \sum_{\pi \in S_n} \text{sgn}(\pi) \cdot a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}.$$

Example

For $n = 1$, $A = (a_{11})$ and $S_1 = \{e\}$, so $\det(A) = a_{11}$.

For $n = 2$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $S_2 = \{e, (1\ 2)\}$. So $|A| = a_{11}a_{22} - a_{12}a_{21}$.

The new definition agrees with M1GLA.

Aim: to prove basic properties of determinants. These are:

(1) to see the effects of row operations on the determinant,

(2) to prove multiplicative property of the determinant:

$$\det(AB) = \det(A)\det(B).$$

Basic properties

Let $A = (a_{ij})$ be $n \times n$. Recall the *transpose* of A is $A^T = (a_{ji})$.

Proposition 10.1 $|A^T| = |A|$.

Proof Let $A^T = (b_{ij})$, so $b_{ij} = a_{ji}$. Then

$$\begin{aligned} |A^T| &= \textstyle\sum_{\pi \in S_n} \text{sgn}(\pi) b_{1,\pi(1)} \cdots b_{n,\pi(n)} \\ &= \textstyle\sum_{\pi \in S_n} \text{sgn}(\pi) a_{\pi(1),1} \cdots a_{\pi(n),n}. \end{aligned}$$

Let $\sigma = \pi^{-1}$. Then

$$a_{\pi(1),1} \cdots a_{\pi(n),n} = a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Also observe $\text{sgn}(\pi) = \text{sgn}(\sigma)$ by 4.1. So

$$|A^T| = \sum_{\pi \in S_n} \text{sgn}(\sigma) \cdot a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

As π runs through all permutations in S_n, so does $\sigma = \pi^{-1}$. Hence $|A^T| = |A|$. □

So any result about determinants concerning rows will have an analogous result concerning columns.

Proposition 10.2 *Suppose B is obtained from A by swapping two rows (or two columns). Then $|B| = -|A|$.*

Proof We prove this for columns (follows for rows using 10.1). Say columns numbered r and s are swapped. Let $\tau = (r\ s)$, 2-cycle in S_n. Then if $B = (b_{ij})$, $b_{ij} = a_{i,\tau(j)}$. So

$$\begin{aligned} |B| &= \textstyle\sum_{\pi \in S_n} \text{sgn}(\pi) b_{1,\pi(1)} \cdots b_{n,\pi(n)} \\ &= \textstyle\sum_{\pi \in S_n} \text{sgn}(\pi) a_{1,\tau\pi(1)}, \cdots a_{n,\tau\pi(n)}. \end{aligned}$$

Now $\text{sgn}(\tau\pi) = \text{sgn}(\tau)\text{sgn}(\pi) = -\text{sgn}(\pi)$ by 4.1. So

$$|B| = \sum_{\pi \in S_n} -\text{sgn}(\tau\pi) \cdot a_{1,\tau\pi(1)}, \cdots a_{n,\tau\pi(n)}.$$

As π runs through all elements of S_n so does $\tau\pi$. So $|B| = -|A|$. □

Proposition 10.3 *(1) If A has a row (or column) of 0's then $|A| = 0$.*

(2) If A has two identical rows (or columns) then $|A| = 0$.

(3) If A is triangular (upper or lower) then $|A| = a_{11}a_{22} \cdots a_{nn}$.

Proof (1) Each term in $|A|$ has an entry from every row, so is 0.

(2) If we swap the identical rows, we get A again, so by 10.2 $|A| = -|A|$. Hence $|A| = 0$.

(3) The only nonzero term in $|A|$ is $a_{11}a_{22} \cdots a_{nn}$. □

For example, by (3), $|I| = 1$.

We can now find the effect of doing row operations on $|A|$.

Theorem 10.4 *Suppose B is obtained from A by using an elementary row operation.*

(1) If two rows are swapped to get B, then $|B| = -|A|$.

(2) If a row of A is multiplied by a nonzero scalar k to get B, then $|B| = k|A|$.

(3) If a scalar multiple of one row of A is added to another row to get B, then $|B| = |A|$.

(4) If $|A| = 0$, then $|B| = 0$ and if $|A| \neq 0$ then $|B| \neq 0$.

Proof (1) is 10.2.

(2) Every term in $|A|$ has exactly one entry from the row in question, so is multiplied by k. Hence $|B| = k|A|$.

(3) Suppose $c \times$ row k is added to row j. So

$$\begin{aligned} |B| &= \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{ji} + ca_{k1} & \cdots & a_{jn} + ca_{kn} \\ & \vdots & \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{ji} & \cdots & a_{jn} \\ & \vdots & \end{vmatrix} + c \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{k1} & \cdots & a_{kn} \\ & \vdots & \\ a_{k1} & \cdots & a_{kn} \end{vmatrix} \\ &= |A| + 0 \end{aligned}$$

by 10.3(2). Hence $|B| = |A|$.

(4) is clear from (1), (2), (3). □

Expansions of determinants

As in M1GLA, recall that if $A = (a_{ij})$ is $n \times n$, the *ij-minor* A_{ij} is the $(n-1) \times (n-1)$ matrix obtained by deleting row i and column j from A.

Proposition 10.5 (Laplace expansion by rows) *Let A be $n \times n$.*

(1) Expansion by 1^{st} row:

$$|A| = a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}| - \cdots + (-1)^{n-1}a_{1n}|A_{1n}|.$$

(2) Expansion by i^{th} row:

$$(-1)^{i-1}|A| = a_{i1}|A_{i1}| - a_{i2}|A_{i2}| + a_{i3}|A_{i3}| - \cdots + (-1)^{n-1}a_{in}|A_{in}|.$$

Note that using 10.1 we can get similar expansions by columns.

Proof (1) For the first row: Consider

$$|A| = \sum_{\pi \in S_n} (\text{sgn}\pi) a_{1,\pi(1)} \cdots a_{n,\pi(n)}.$$

Terms with a_{11} are

$$\sum_{\pi \in S_n, \pi(1)=1} \text{sgn}(n) a_{11} a_{2,\pi(2)} \cdots a_{n,\pi(n)} = a_{11}|A_{11}|.$$

To calculate terms with a_{12}, swap columns 1 and 2 of A to get

$$B = \begin{pmatrix} a_{12} & a_{11} & a_{13} & \cdots \\ a_{22} & a_{21} & a_{23} & \cdots \\ \vdots & \vdots & \vdots & \\ a_{n2} & a_{n1} & a_{n3} & \cdots \end{pmatrix}.$$

Then $|B| = -|A|$ by 10.2. Terms in $|B|$ with a_{12} add to $a_{12}|A_{12}$. So terms in $|A|$ with a_{12} add to $-a_{12}|A_{12}|$. For terms with a_{13}, swap columns 2 and 3 of A, then swap columns 1 and 2 to get

$$B' = \begin{pmatrix} a_{13} & a_{11} & a_{12} & \cdots \\ a_{23} & a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \vdots & \\ a_{n3} & a_{n1} & a_{n2} & \cdots \end{pmatrix}.$$

Then $|B'| = |A|$ and a_{13} terms add to $a_{13}|A_{13}|$.

Continuing like this, see that $|A| = a_{11}|A_{11}| - a_{12}|A_{12} + \cdots$ which is expansion by the first row.

(2) For expansion by i^{th} row, do $i-1$ row swaps in A to get

$$B'' = \begin{pmatrix} a_{i1} & \cdots & a_{in} \\ a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ & \vdots & \end{pmatrix}.$$

Then $|B''| = (-1)^{i-1}|A|$. Now use expansion of B'' by 1^{st} row. □

Major properties of determinants

Two major results. First was proved in M1GLA for 2×2 and 3×3 cases:

Theorem 10.6 *Let A be $n \times n$. The following statements are equivalent.*

(1) $|A| \neq 0$.

(2) A is invertible.

(3) The system $Ax = 0$ ($x \in \mathbb{R}^n$) has only solution $x = \underline{0}$.

(4) A can be reduced to I_n by elementary row operations.

Proof We proved (2) ⇔ (3) ⇔ (4) in M1GLA (7.5).

(1) ⇒ (4): Suppose $|A| \neq 0$. Reduce A to echelon form A' by elementary row operations. Then $|A'| \neq 0$ by 10.4(4). So A' does not have a zero row. Therefore A' is upper triangular with 1's on diagonal and hence can be reduced further to I_n by row operations.

(4) ⇒ (1): Suppose A can be reduced to I_n by row operations. We know that $|I_n| = 1$. So $|A| \neq 0$ by 10.4(4). □

Corollary 10.7 *Let A be $n \times n$. If the system $Ax = 0$ has a nonzero solution $x \neq 0$ then $|A| = 0$.*

Second major result on determinants:

Theorem 10.8 *If A, B are $n \times n$ then*

$$\det(AB) = \det(A)\det(B).$$

To prove this need to study

Elementary matrices

These are $n \times n$ of the following types:

$$A_i(r) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & r & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \qquad r \neq 0,$$

$$B_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \qquad I_n \text{ with rows } i, j \text{ swapped},$$

$$C_{ij}(r) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & r & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}. \qquad r \text{ is the } ij\text{-th entry, } i \neq j.$$

The elementary matrices correspond to elementary row operations:

Proposition 10.9 *Let A be $n \times n$. An elementary row operation on A changes it to EA, where E is an elementary matrix.*

Proof Let the rows of A be $v_1, \ldots, v_n$.
(1) Row operation $v_i \mapsto rv_i$ sends A to $A_i(r)A$.
(2) Row operation $v_i \leftrightarrow v_j$ sends A to $B_{ij}A$.
(3) Row operation $v_i \mapsto v_i + rv_j$ sends A to $C_{ij}(r)A$. □

Proposition 10.10 *(1) The determinant of an elementary matrix is nonzero and*

$$|A_i(r)| = r, \ |B_{ij}| = -1, \ |C_{ij}(r)| = 1.$$

(2) The inverse of an elementary matrix is also an elementary matrix:

$$A_i(r)^{-1} = A_i(r^{-1}),\ B_{ij}^{-1} = B_{ij},\ C_{ij}(r)^{-1} = C_{ij}(-r).$$

Proposition 10.11 *Let A be $n \times n$, and suppose A is invertible. Then A is equal to a product of elementary matrices, i.e. $A = E_1 \cdots E_k$ where each E_i is an elementary matrix.*

Proof By 10.6, A can be reduced to I by elementary row operations. By 10.9 first row operations changes A to $E_1 A$ with E_1 elementary matrix. Second changes $E_1 A$ to $E_2 E_1 A$, E_2 elementary matrix ... and so on, until we end up with I. Hence

$$I = E_k E_{k-1} \cdots E_1 A,$$

where each E_i is elementary. Multiply both sides on left by $E_1^{-1} \cdots E_{k-1}^{-1} E_k^{-1}$ to get

$$E_1^{-1} \cdots E_k^{-1} = A.$$

Each E_i^{-1} is elementary by 10.10(2). □

Towards Theorem 10.8:

Proposition 10.12 *If E is an elementary $n \times n$ matrix, and A is $n \times n$, then* $\det(EA) = \det(E)\det(A)$.

Proof Let the rows of A be $v_1, \ldots, v_n$.

(1) If $E = A_i(r)$, then EA has rows $v_1, \ldots, rv_i, \ldots v_n$, so $|EA| = r|A|$ by 10.4 and therefore $|EA| = |E||A|$ by 10.10.

(2) If $E = B_{ij}$, then EA is obtained by swapping rows i and j of A, so $|EA| = -|A|$ by 10.4 and so $|EA| = |E||A|$ by 10.10.

(3) If $E = C_{ij}(r)$ then EA has rows $v_1, \ldots, v_i + rv_j, \ldots v_n$, so $|EA| = |E||A|$ by 10.4 and 10.10. □

Corollary 10.13 *If $A = E_1 \ldots E_k$, where each E_i is elementary, then $|A| = |E_1| \cdots |E_k|$.*

Proof

$$\begin{aligned} |A| &= |E_1 \cdots E_k| \\ &= |E_1||E_2 \cdots E_k| && \text{by 10.12} \\ &\ \ \cdots \\ &= |E_1||E_2| \cdots |E_k|. \end{aligned}$$

Proof of Theorem 10.8

(1) If $|A| = 0$ or $|B| = 0$, then $|AB| = 0$ by Sheet 6, Q7.

(2) Now assume that $|A| \neq 0$ and $|B| \neq 0$. Then A, B are invertible by 10.6. So by 10.11,

$$A = E_1 \cdots E_k, \qquad B = F_1 \cdots F_l$$

where all E_i, F_i are elementary matrices. By 10.13,

$$|A| = |E_1| \cdots |E_k|, \quad |B| = |F_1| \cdots |F_k|.$$

Also $AB = E_1 \cdots E_k F_1 \cdots F_l$, so by 10.13

$$|AB| = |E_1| \cdots |E_n||F_1| \cdots |F_k| = |A||B|.$$

Immediate consequence:

Proposition 10.14 *Let P be an invertible $n \times n$ matrix.*

(1) $\det(P^{-1}) = \frac{1}{\det(P)}$*,*

(2) $\det(P^{-1}AP) = \det(A)$ *for all $n \times n$ matrices A.*

Proof (1) $\det(P)\det(P^{-1}) = \det PP^{-1} = \det I = 1$ by 10.8.

(2) $\det(P^{-1}AP) = \det(P^{-1})\det A \det P = \det A$ by 10.8 and (1). □

11 Matrices and linear transformations

Recall from M1P2:

Let V be a finite dimensional vector space and $T : V \to V$ a linear transformation. If $B = \{v_1, \ldots, v_n\}$ is a basis of V, write

$$\begin{aligned} T(v_1) &= a_{11}v_1 + \ldots + a_{n1}v_n, \\ &\vdots \\ T(v_n) &= a_{1n}v_1 + \ldots + a_{nn}v_n. \end{aligned}$$

The *matrix of T with respect to B* is

$$[T]_B = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

A result from M1P2:

Proposition 11.1 *Let $S : V \to V$ and $T : V \to V$ be linear transformations and let B be a basis of V. Then*

$$[ST]_B = [S]_B[T]_B,$$

where ST is the composition of S and T.

Consequences of 11.1:

As in 11.1, let V be n-dimensional over $F = \mathbb{R}$ or $\mathbb{C}$, basis B. The map $T \mapsto [T]_B$ gives a correspondence

$$\{\text{linear transformations } V \to V\} \leftrightarrow \{n \times n \text{ matrices over } F\}.$$

This has many nice properties:

1. If $[T]_B = A$ then $[T^2]_B = A^2$ and similarly $\left[T^k\right]_B = A^k$.

For a polynomial $q(x) = a_r x^r + \cdots + a_1 x + a_0$ ($a_i \in \mathbb{C}$), define

$$q(A) = a_r A^r + \cdots + a_1 A + a_0 I$$

and

$$q(T) = a_r T^r + \cdots + a_1 T + a_0 1_V$$

where $1_V : V \to V$ is the identity map. Then 11.1 implies that

$$[q(T)]_B = q(A).$$

Example Let V = polynomials of degree ≤ 2, $T(p(x)) = p'(x)$. Then $(T^2 - T)(p(x)) = p''(x) - p'(x)$ and

$$[T^2 - T]_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}^2 - \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

2. Define $GL(V)$ to be the set of all invertible linear transformations $V \to V$. Then $GL(V)$ is a group under composition, and $T \mapsto [T]_B$ is an isomorphism from $GL(V)$ to $GL(n, F)$ (recall that $GL(n, F)$ is the group of all $n \times n$ invertible matrices under matrix multiplication).

Change of basis

Let V be n-dimensional, with bases $E = \{e_1, \ldots, e_n\}$, $F = \{f_1, \ldots, f_n\}$. Write

$$\begin{aligned} f_1 &= p_{11}e_1 + \cdots + p_{n1}e_n, \\ &\vdots \\ f_n &= p_{1n}e_1 + \cdots + p_{nn}e_n. \end{aligned}$$

and define P to be the $n \times n$ matrix (p_{ij}). Recall from M1P2 that P is the *change of basis matrix* from E to F. Here's another basic result from M1P2:

Proposition 11.2 *(1) P is invertible.*

(2) If $T : V \to V$ is a linear transformation, then $[T]_F = P^{-1}[T]_E P$.

Determinant of a linear transformation

Definition Let A, B be $n \times n$ matrices. We say A is *similar* to B if there exists an invertible $n \times n$ matrix P such that $B = P^{-1}AP$.

Note that the relation $\sim$ defined by

$$A \sim B \Leftrightarrow A \text{ is similar to } B$$

is an equivalence relation (Sheet 7, Q6).

Proposition 11.3 *(1) If A, B are similar then $|A| = |B|$.*

(2) Let $T : V \to V$ be linear transformations and let E, F be two bases of V. Then the matrices $[T]_E$ and $[T]_F$ are similar.

Proof (1) is 10.14, and (2) is 12.2(2). □

Definition Let $T : V \to V$ be a linear transformation. By 11.3, for any two bases E, F of V, the matrices $[T]_E$ and $[T]_F$ have same determinant. Call $\det[T]_E$ the *determinant of* T, written $\det T$.

Example Let V = polynomials of degree ≤ 2 and $T(p(x)) = p(2x + 1)$. Take $B = \{1, x, x^2\}$, so

$$[T]_B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 4 \end{pmatrix}.$$

So $\det T = 8$.

12 Characteristic polynomials

Recall from M1P2: let $T : V \to V$ be a linear transformation. We say $v \in V$ is an *eigenvector* of T if

(1) $v \neq 0$, and

(2) $T(v) = \lambda v$ where λ is a scalar.

The scalar λ is an *eigenvalue* of T.

Definition The *characteristic polynomial* of $T : V \to V$ is the polynomial $\det(xI - T)$, where $I : V \to V$ is the identity linear transformation.

By the definition of determinant, this polynomial is equal to $\det(xI - [T]_B)$ for any basis B.

Example V = polynomials of degree ≤ 2, $T(p(x)) = p(1-x)$, $B = \{1, x, x^2\}$. The characteristic polynomial of T is

$$\det\left(xI - \begin{pmatrix} 1 & 1 & \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} x-1 & -1 & -1 \\ 0 & x+1 & 2 \\ 0 & 0 & x-1 \end{pmatrix} = (x-1)^2(x+1).$$

From M1P2:

Proposition 12.1 *(1) The eigenvalues of T are the roots of the characteristic polynomial of T.*

(2) If λ is an eigenvalue of T, the eigenvectors corresponding to λ are the nonzero vectors in

$$E_\lambda = \{v \in V \mid (\lambda I - T)(v) = 0\} = \ker(\lambda I - T).$$

(3) The matrix $[T]_B$ is a diagonal matrix iff B consists of eigenvectors of T.

Note that $E_\lambda = \ker(\lambda I - T)$ is a subspace of V, called the λ-*eigenspace* of T.

Example In previous example, eigenvalues of T are $1, -1$. Eigenspace E_1 is $\ker(I - T)$. Solve

$$\left(\begin{array}{ccc|c} 0 & -1 & -1 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Solutions are vectors $\begin{pmatrix} a \\ b \\ -b \end{pmatrix}$. So $E_1 = \{a + bx - bx^2 \mid a, b \in F\}$.

Eigenspace E_{-1}. Solve

$$\left(\begin{array}{ccc|c} 2 & -1 & -1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -2 & 0 \end{array}\right).$$

Solutions are vectors $\begin{pmatrix} c \\ -2c \\ 0 \end{pmatrix}$. So $E_{-1} = \{c - 2cx \mid c \in F\}$.

Basis of E_1 is $1, x - x^2$. Basis of E_{-1} is $1 - 2x$. Putting these together, get basis

$$B = \{1, x - x^2, 1 - 2x\}$$

of V consisting of eigenvectors of T, and

$$[T]_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Proposition 12.2 *Let V a finite-dimensional vector space over $\mathbb{C}$. Let $T : V \to V$ be a linear transformation. Then T has an eigenvalue $\lambda \in \mathbb{C}$.*

Proof The characteristic polynomial of T has a root $\lambda \in \mathbb{C}$ by the Fundamental theorem of Algebra. □

Note that Proposition 12.2 is not necessarily true for vector spaces over $\mathbb{R}$. For example $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x_1, x_2) = (x_2, -x_1)$ has characteristic polynomial $x^2 + 1$, which has no real roots.

Diagonalisation

Basic question is: How to tell if there exists a basis B such that $[T]_B$ is diagonal? Useful result:

Proposition 12.3 *Let $T : V \to V$ be a linear transformation. Suppose $v_1, \ldots, v_k$ are eigenvectors of T corresponding to* distinct *eigenvalues $\lambda_1, \ldots, \lambda_k$. Then $v_1, \ldots, v_k$ are linearly independent.*

Proof By induction on k. Let $P(k)$ be the statement of the proposition. $P(1)$ is true, since $v_1 \neq 0$, so v_1 is linearly independent. Assume $P(k-1)$ is true, so $v_1, \ldots, v_{k-1}$ are linearly independent. We show $v_1, \ldots, v_k$ are linearly independent. Suppose

$$r_1 v_1 + \cdots + r_k v_k = 0. \tag{34}$$

Apply T to get

$$\lambda_1 r_1 v_1 + \cdots + \lambda_k r_k v_k = 0 \tag{35}$$

Then (35)-$\lambda_k \times$(34) gives

$$r_1(\lambda_1 - \lambda_k)v_1 + \cdots + r_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

As $v_1, \ldots, v_{k-1}$ are linearly independent, all coefficients are 0. So

$$r_1(\lambda_1 - \lambda_k) = \ldots = r_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

As the λ_i are distinct, $\lambda_1 - \lambda_k, \ldots, \lambda_{k-1} - \lambda_k \neq 0$. Hence

$$r_1 = \ldots = r_{k-1} = 0.$$

Then (34) gives $r_k v_k = 0$, so $r_k = 0$. Hence $v_1, \ldots, v_k$ are linearly independent, completing the proof by induction. □

Corollary 12.4 *Let* $\dim V = n$ *and* $T : V \to V$ *be a linear transformation. Suppose the characteristic polynomial of* T *has* n *distinct roots. Then* V *has a basis* B *consisting of eigenvectors of* T *(i.e* $[T]_B$ *is diagonal).*

Proof Let $\lambda_1, \ldots, \lambda_n$ be the (distinct) roots, so these are the eigenvalues of T. Let $v_1, \ldots, v_n$ be corresponding eigenvectors. By 12.3, $v_1, \ldots, v_n$ are linearly independent, hence form a basis of V since $\dim V = n$. □

Example Let

$$A = \begin{pmatrix} \lambda_1 & & & \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

be triangular, with diagonal entries $\lambda_1, \ldots, \lambda_n$, all distinct. The characteristic polynomial of A is

$$|xI - A| = \prod_{i=1}^{n} (x - \lambda_i)$$

which has roots $\lambda_1, \ldots, \lambda_n$. Hence by 12.4, A can be diagonalized, i.e. there exists P such that $P^{-1}AP$ is diagonal.

Note that this is not necessarily true if the diagonal entries are not distinct, e.g. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ cannot be diagonalized.

Algebraic and geometric multiplicities

Let $T : V \to V$ be a linear transformation with characteristic polynomial $p(x) = \det(xI - T)$. Let λ be an eigenvalue of T, i.e. a root of $p(x)$. Write

$$p(x) = (x - \lambda)^{a(\lambda)} q(x),$$

where λ is not a root of $q(x)$. Call $a(\lambda)$ the *algebraic multiplicity* of λ.

The *geometric multiplicity* of λ is defined to be

$$g(\lambda) = \dim E_\lambda,$$

where $E_\lambda = \ker(\lambda I - T)$, the λ-eigenspace of T.

We adopt similar definitions for $n \times n$ matrices.

Example For $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, we have

$$a(1) = g(1) = 1, \ \ a(2) = g(2) = 1.$$

And for $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, we have

$$a(1) = 2, g(1) = 1.$$

Proposition 12.5 *If λ is an eigenvalue of $T : V \to V$, then $g(\lambda) \leq a(\lambda)$.*

Proof Let $r = g(\lambda) = \dim E_\lambda$ and let $v_1, \ldots, v_r$ be a basis of E_λ. Extend to a basis of V:

$$B = \{v_1, \ldots, v_r, w_1, \ldots, w_s\}.$$

We work out $[T]_B$:

$$\begin{aligned} T(v_1) &= \lambda v_1, \\ &\vdots \\ T(v_r) &= \lambda v_r, \\ T(w_1) &= a_{11}v_1 + \cdots + a_{r1}v_r + b_{11}w_1 + \cdots + b_{s1}w_s, \\ &\vdots \\ T(w_s) &= a_{1s}v_1 + \cdots + a_{rs}v_r + b_{1s}w_1 + \cdots + b_{ss}w_s. \end{aligned}$$

So

$$[T]_B = \left(\begin{array}{cccc|ccc} \lambda & 0 & \cdots & 0 & a_{11} & \cdots & a_{1s} \\ 0 & \lambda & \cdots & 0 & \vdots & & \vdots \\ \vdots & \vdots & \ddots & & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda & a_{r1} & \cdots & a_{rs} \\ \hline 0 & \cdots & \cdots & 0 & b_{11} & \cdots & b_{1s} \\ \vdots & & & \vdots & \vdots & & \vdots \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & b_{s1} & \cdots & b_{ss} \end{array}\right).$$

Clearly the characteristic polynomial of this is

$$p(x) = \det \left(\begin{array}{c|c} (x-\lambda)I_r & -A \\ \hline 0 & xI_s - B \end{array}\right).$$

By Sheet 7 Q5, this is

$$p(x) = \det((x-\lambda)I_r)\det(xI_s - B) = (x-\lambda)^r q(x).$$

Hence the algebraic multiplicity $a(\lambda) \geq r = g(\lambda)$. □

Here is a basic criterion for diagonalisation:

Theorem 12.6 *Let* $\dim V = n$, $T : V \to V$ *be a linear transformation, let* $\lambda_1, \ldots, \lambda_r$ *be the distinct eigenvalues of* T*, and the characteristic polynomial of* T *be*

$$p(x) = \prod_{i=1}^{r} (x - \lambda_i)^{a(\lambda_i)}$$

(so $\sum_{i=1}^r a(\lambda_i) = n$*). The following statements are equivalent:*

(1) V *has a basis* B *consiting of eigenvectors of* T *(i.e.* $[T]_B$ *is diagonal).*

(2) $\sum_{i=1}^r g(\lambda_i) = \sum_{i=1}^r \dim E_{\lambda_i} = n$.

(3) $g(\lambda_i) = a(\lambda_i)$ *for all* i.

Proof To prove$(1) \Rightarrow (2), (3)$: Suppose (1) holds. Each vector in B is in some E_{λ_i}, so

$$\sum_{i=1}^{r} \dim E_{\lambda_i} \geq |B| = n.$$

By 12.5
$$\sum_{i=1}^{r} \dim E_{\lambda_i} = \sum_{i=1}^{r} g(\lambda_i) \leq \sum_{i=1}^{r} a(\lambda_i) = n.$$
Hence $\sum_{i=1}^{r} \dim E_{\lambda_i} = n$ and $g(\lambda_i) = a(\lambda_i)$ for all i.

Evidently (2) $\Leftrightarrow$ (3), so it is enough to show that (2) $\Rightarrow$ (1). Suppose $\sum_{i=1}^{r} \dim E_{\lambda_i} = n$. Let B_i be a basis of E_{λ_i} and let $B = \bigcup_{i=1}^{r} B_i$, so $|B| = n$ (the B_i's are disjoint as they consist of eigenvectors for different eigenvalues). We claim B is a basis of V, hence (1) holds:
It's enough to show that B is linearly independent (since $|B| = n = \dim V$). Suppose there is a linear relation
$$\sum_{v \in B_1} \alpha_v v + \cdots + \sum_{z \in B_r} \alpha_z z = 0.$$
Write
$$\begin{array}{rcl} v_1 & = & \sum_{v \in B_1} \alpha_v v, \\ & \vdots & \\ v_r & = & \sum_{z \in B_r} \alpha_z z, \end{array}$$
so $v_i \in E_{\lambda_i}$ and $v_1 + \cdots + v_r = 0$. As $\lambda_1, \ldots, \lambda_r$ are distinct, the set of nonzero v_i's is linearly independent by 12.3. Hence $v_i = 0$ for all i. So
$$v_i = \sum_{v \in B_i} \alpha_v v = 0.$$
As B_i is linearly independent (basis of E_{λ_i}) this forces $\alpha_v = 0$ for all $v \in B_i$. This completes the proof that B is linearly independent, hence a basis of V.
□

Using 12.6 we get an algorithm to check whether a given $n \times n$ matrix or linear transformation is diagonalizable:

1. Find the characteristic polynomial, factorise it as
$$\prod (x - \lambda_i)^{a(\lambda_i)}.$$
2. Calculate each $g(\lambda_i) = \dim E_{\lambda_i}$.
3. If $g(\lambda_i) = a(\lambda_i)$ for all i, YES.
   If $g(\lambda_i) < a(\lambda_i)$ for some i, NO.

Example Let $A = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}$. Check that

(1) Characteristic polynomial is $(x+2)^2(x-4)$.

(2) For eigenvalue 4: $a(4) = 1, g(4) = 1$ (as it is $\leq a(4)$).
For eigenvalue -2: $a(-2) = 2$, $g(-2) = \dim E_{-2} = 1$.

So A is not diagonalizable by 12.6.

13 The Cayley-Hamilton theorem

Recall that if $T : V \to V$ is a linear transformation and $p(x) = a_k x^k + \cdots + a_1 x + a_0$ is a polynomial, then $p(T) : V \to V$ is defined by

$$p(T) = a_k T^k + a_{k-1} T^k + \cdots + a_1 T + a_0 1_V.$$

Likewise if A is $n \times n$ matrix,

$$p(A) = a_k A^k + \cdots a_1 A + a_0 I.$$

Theorem 13.1 (Cayley-Hamilton Theorem) *Let V be finite-dimensional vector space, and $T : V \to V$ a linear transformation with characteristic polynomial $p(x)$. Then $p(T) = 0$, the zero linear transformation.*

Proof later.

Corollary 13.2 *If A is a $n \times n$ matrix with characteristic polynomial $p(x)$, then $p(A) = 0$.*

This can easily be deduced from Theorem 13.1: simply apply 13.1 to the linear transformation $T : F^n \to F^n$ ($F = \mathbb{R}$ or $\mathbb{C}$) given by $T(v) = Av$.

Examples 1. 13.2 is obvious for diagonal matrices

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

This is because the λ_i are the roots of $p(x)$, so

$$p(A) = \begin{pmatrix} p(\lambda_1) & & \\ & \ddots & \\ & & p(\lambda_n) \end{pmatrix} = 0.$$

Corollary 13.2 is also quite easy to prove for *diagonalisable* matrices (Sheet 8 Q3).

2. For 2×2 matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the characteristic polynomial is

$$p(x) = \begin{vmatrix} x-a & -b \\ -c & x-d \end{vmatrix} = x^2 - (a+d)x + ad - bc.$$

So 13.2 tells us that

$$A^2 - (a+d)A + (ad-bc)I = 0.$$

Could verify this directly. For 3×3, $\ldots$, $n \times n$ need a better idea.

Proof of Cayley-Hamilton

Let $T : V \to V$ be a linear transformation with characteristic polynomial $p(x)$.

Aim: for $v \in V$, show that $p(T)(v) = 0$.

Strategy: Study the subspace

$$\begin{aligned} v^T &= \text{Span}(v, T(v), T^2(v), \ldots) \\ &= \text{Span}(T^i(v) \mid i \geq 0). \end{aligned}$$

Definition A subspace W of V is *T-invariant* if $T(W) \subseteq W$, i.e. $T(w) \in W$ for all $w \in W$.

Proposition 13.3 *Pick $v \in V$ and let*

$$W = v^T = \text{Span}(T^i(v) \mid i \geq 0).$$

Then W is T-invariant.

Proof Let $w \in W$, and write

$$w = a_1 T^{i_1}(v) + \cdots + a_r T^{i_r}(v).$$

Then

$$T(w) = a_1 T^{i_1+1}(v) + \cdots + a_r T^{i_r+1}(v),$$

so $T(w) \in W$. $\square$

Example V = polynomials of deg ≤ 2, $T(p(x)) = p(x+1)$. Then

$$\begin{aligned} x^T &= \text{Span}(x, T(x), T^2(x), \ldots) \\ &= \text{Span}(x, x+1) = \text{subspace of polynomials of deg} \leq 1. \end{aligned}$$

Clearly this is T-invariant.

Definition Let W be a T-invariant subspace of V. Define $T_W : W \to W$ by

$$T_W(w) = T(w)$$

for all $w \in W$. Then T_W is a linear transformation, the *restriction* of T to W.

Proposition 13.4 *If W is a T-invariant subspace of V, then the characteristic polynomial of T_W divides the characteristic polynomial of T.*

Proof Let

$$B_W = \{w_1, \ldots, w_k\}$$

be a basis of W and extend it to a basis

$$B = \{w_1, \ldots, w_k, x_1, \ldots, x_l\}$$

of V. As W is T-invariant,

$$\begin{aligned} T(w_1) &= a_{11}w_1 + \cdots + a_{k1}w_k, \\ &\vdots \\ T(w_k) &= a_{1k}w_1 + \cdots + a_{kk}w_k. \end{aligned}$$

Then

$$[T_W]_{B_W} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} = A$$

and

$$[T]_B = \left(\begin{array}{c|c} A & X \\ \hline 0 & Y \end{array}\right).$$

The characteristic polynomial of T_W is $p_W(x) = \det(xI_k - A)$, and characteristic polynomial of T is

$$\begin{aligned} p(x) &= \det \left(\begin{array}{c|c} xI_k - A & -X \\ \hline 0 & xI_l - Y \end{array} \right) \\ &= \det(xI_k - A) \cdot \det(xI_l - Y) \\ &= p_W(x) \cdot q(x). \end{aligned}$$

So $p_W(x)$ divides $p(x)$. $\square$

Example $V =$ polynomials of deg ≤ 2, $T(p(x)) = p(x+1)$, $W = x^T =$ Span $(x, x+1)$. Take basis $B_W = \{1, x\}$, $B = \{1, x, x^2\}$. Then

$$\begin{aligned} [T]_{B_W} &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \\ [T]_B &= \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Characteristic polynomial of T_W is $(x-1)^2$, characteristic polynomial of T is $(x-1)^3$.

Proposition 13.5 *Let $T : V \to V$ be a linear transformation. Let $v \in V$, $v \neq 0$, and*

$$W = v^T = \text{Span}\left(T^i(v) \mid i \geq 0\right).$$

Let $k = \dim W$. Then

$$\left\{v, T(v), T^2(v), \ldots, T^{k-1}(v)\right\}$$

is a basis of W.

Proof We show that $\{v, T(v), \ldots, T^{k-1}(v)\}$ is linearly independent, hence a basis of W. Let j be the largest integer such that the set $\{v, T(v), \ldots, T^{j-1}(v)\}$ is linearly independent. So $1 \leq j \leq k$. Aim to show that $j = k$. Let

$$S = \{v, T(v), \ldots, T^{j-1}(v)\}$$

and

$$X = \text{Span}(S).$$

Then $X \subseteq W$ and $\dim X = j$. By the choice of j, the set

$$\{v, T(v), \ldots, T^{j-1}(v), T^j(v)\}$$

is linearly dependent. This implies that $T^j(v) \in \text{Span}(S) = X$. Say

$$T^j(v) = b_0 v + b_1 T(v) + \cdots + b_{j-1} T^{j-1}(v).$$

So

$$T^{j+1}(v) = b_0 T(v) + b_1 T^2(v) + \cdots + b_{j-1} T^j(v) \in X.$$

Similarly $T^{j+2}(v) \in X$, $T^{j+3}(v) \in X$ and so on. Hence $T^i(v) \in X$ for all $i \geq 0$. This implies

$$W = \text{Span}(T^i(v) \mid i \geq 0) \subseteq X.$$

As $X \subseteq W$ this means $X = W$, so $j = \dim X = \dim W = k$. Hence $\{v, T(v), \ldots, T^{k-1}(v)\}$ is linearly independent, as required. □

Proposition 13.6 *Let $T : V \to V$, let $v \in V$ and $W = v^T = \text{Span}\left(T^i(v) \mid i \geq 0\right)$, with basis $B_W = \{v, T(v), \ldots, T^{k-1}(v)\}$ as in 13.5. Then*

(1) there exist scalars a_i such that

$$a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0,$$

(2) the characteristic polynomial of T_W is

$$p_W(x) = x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0,$$

(3)] $p_W(T)(v) = 0$.

Proof

(1) is clear, since $T^k(v) \in W$ and B_W is a basis of W.

(2) Clearly

$$[T_W]_{B_W} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix}$$

(for the last column $T(T^{k-1}(v)) = T^k(v) = -a_0 v - a_1 T(v) - \cdots - a_{k-1} T^{k-1}(v)$).
By Sheet 8 Q4, the characteristic polynomial of this matrix is

$$p_W(x) = x^k + a_{k-1} x^{k-1} + \cdots + a_0.$$

(3) This is clear from (1) and (2). □

Completion of the proof of Cayley-Hamilton 13.1

We have $T : V \to V$ with characteristic polynomial $p(x)$. Let $v \in V$, let $W = v^T$ with basis $\{v, T(v), \ldots, T^{k-1}(v)\}$. Let $p_W(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_0$ to be the characteristic polynomial of T_W. By 13.6(3),

$$p_W(T)(v) = 0.$$

By 13.4, $p_W(x)$ divides $p(x)$, say $p(x) = q(x)p_W(x)$, so $p(T) = q(T)p_W(T)$. Then

$$\begin{aligned} p(T)(v) &= (q(T)p_W(T))(v) \\ &= q(T)(p_W(T)(v)) \\ &= q(T)(0) = 0. \end{aligned}$$

Thus $p(T)(v) = 0$ for all $v \in V$, which means that $p(T) = 0$. This completes the proof.

14 Invariants of matrices

Recall that two $n \times n$ matrices A, B are *similar* if there is an invertible matrix P such that $B = P^{-1}AP$. Similar matrices share many common properties:

Proposition 14.1 *If A, B are similar $n \times n$ matrices, they have*

(i) *the same characteristic polynomial*

(ii) *the same eigenvalues and algebraic multiplicities*

(iii) *the same geometric multiplicities*

(iv) *the same determinant*

(v) *the same rank and nullity*

(vi) *the same trace, where* trace$(A) = \sum a_{ii}$*, the sum of the diagonal entries.*

Proof (i) is Sheet 8 Q2, and (ii) follows from (i).

(iii) Let $V = F^n$ (where $F = \mathbb{R}$ or $\mathbb{C}$), and define $T : V \to V$ by $T(v) = Av$. Choose bases E and F of V such that $[T]_E = A$ and $[T]_F = B$ (i.e. take E to be the standard basis, and F the basis with P as its change of basis matrix from E). Then for any evalue λ, the dimension of the λ-eigenspace of A or B is equal to $\dim\ker(T - \lambda I)$. Hence (iii).

(iv) is 10.14.

(v) The nullity of A is the dimension of the 0-eigenspace, so (v) follows from (iii).

(vi) The char poly of A is

$$\det(xI - A) = x^n - x^{n-1}(a_{11} + \cdots + a_{nn}) + \cdots$$

so the coefficient of x^{n-1} is $-\text{trace}(A)$. Hence $\text{trace}(A) = \text{trace}(B)$ by (i) □

We summarise 14.1 by saying that the char poly, eigenvalues, geometric mults, trace. etc. of a matrix A are quantities which are *invariant under similarity.*

Note however that there properties do not determine A: there are many pairs of non-similar matrices which have the same char poly, determinant, trace, etc. Here's an example:

Example Let

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then A, B have the same char poly $(x-1)^4$, the same geom mult $g(1) = 2$, the same determinant 1, the same rank 4, the same trace 4. Yet A and B are not similar (see the next section to justify this).

Aim: to find invariants of a matrix A which are sufficient to determine A up to similarity. Will do this in the next section.

15 The Jordan Canonical Form

Definition Let $\lambda \in \mathbb{C}$ and define the $n \times n$ matrix

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ & & & \dots & & \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

Such a matrix is called a *Jordan block.*

For example

$$J_2(5) = \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix}, \; J_3(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \; J_1(\lambda) = (\lambda).$$

Proposition 15.1 *Let $J = J_n(\lambda)$.*

(1) The char poly of J is $(x - \lambda)^n$.

(2) λ is the only eigenvalue of J: its algebraic mult is n and its geometric mult is 1.

(3) $J - \lambda I = J_n(0)$, and multiplication by $J - \lambda I$ sends the standard basis vectors

$$e_n \to e_{n-1} \to \cdots \to e_2 \to e_1 \to 0.$$

(4) $(J - \lambda I)^n = 0$, and for $i < n$, $(J - \lambda I)^i$ sends $e_n \to e_{n-i}$, $e_{n-1} \to e_{n-i-1}$ and so on.

The proof is routine.

Block diagonal matrices

If $A_1, \ldots, A_k$ are square matrices, where A_i is $n_i \times n_i$, we define the *block diagonal* matrix

$$A_1 \oplus A_2 \oplus \cdots \oplus A_k = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & A_k \end{pmatrix}$$

This is $n \times n$, where $n = \sum n_i$.

For example, if $A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}$ and $B = (3)$, then

$$A \oplus B = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Proposition 15.2 *Let $A = A_1 \oplus \cdots \oplus A_k$ and let $p_i(x)$ be the char poly of A_i.*

(1) The char poly of A is $\prod_1^k p_i(x)$.

(2) The set of eigenvalues of A is the union of the set of eigenvalues of the A_i's.

(3) For any polynomial $q(x)$,

$$q(A) = q(A_1) \oplus \cdots \oplus q(A_k).$$

(4) For any eigenvalue λ *of* A*, its geometric mult for* A *is the sum of its geometric mults for the* A_i*, i.e.* $\dim E_\lambda(A) = \sum \dim E_\lambda(A_i)$.

Proof Parts (1)-(3) are clear, and (4) is Sheet 9, Q3.

Here is the main theorem of this section, indeed one of the main theorems in the whole of linear algebra.

Theorem 15.3 (Jordan Canonical Form) *Let* A *be an* $n \times n$ *matrix over* $\mathbb{C}$*. Then* A *is similar to a matrix of the form*

$$J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_k}(\lambda_k)$$

where $\sum n_i = n$ *(note that the evalues* λ_i *are not necessarily distinct). This is called the* **Jordan canonical form (JCF)** *of* A*, and is unique, apart from changing the order of the Jordan blocks.*

Proof later.

Here are a few examples of JCFs:

$$J_2(1) \oplus J_2(1) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \; J_3(1) \oplus J_1(1) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

(the theorem says these are not similar – see the end of the last section),

$$J_1(0) \oplus J_2(-i) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -i & 1 \\ 0 & 0 & -i \end{pmatrix}.$$

Notice that the only diagonal JCF matrices are of the form $J_1(\lambda_1) \oplus \cdots \oplus J_1(\lambda_k)$ – so in some sense "most" matrices are not diagonalisable.

Notice also that a JCF matrix is upper triangular, so one consequence of the theorem is that every $n \times n$ matrix over $\mathbb{C}$ can be "triangularised", i.e. is similar to a triangular matrix.

At this point I have become somewhat cheesed off with typing all these notes, so I am going to stop here and tell you to rely on the excellent notes you wrote in the lectures. I have put some notes on the proof of the JCF theorem on the website, so you can't complain too much.