MATH 2022 Groups and Vector Spaces

This outline is based on the previous lecturer's lectures notes, and is a useful resource, but not exactly the same as the course as now given (though I have tried to be consistent about notation). You should take your own notes in lectures, and may use these notes as back-up.

1 Definition and examples of groups; subgroups and order

$\mathbb{N} = \{0, 1, 2, 3, \ldots\}$, the set of natural numbers (note that some authors omit 0, but I do not). $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are the sets of integers, rational numbers, real numbers, and complex numbers respectively. All come with binary operations $+$ and $\times$.

1.1 Definition. Fix an integer $n \geq 1$. We let $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$, and we add and multiply members of $\mathbb{Z}_n$ 'modulo' n. That is, we add or multiply two given members of $\mathbb{Z}_n$ as usual, and then find the remainder of the answer on division by n. This is called the *ring of integers modulo n*.

Example. $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and $+$ and $\times$ are given by the tables

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\times$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

1.2 Definition. A *group* is a non-empty set G on which is defined an associative binary operation $\circ$ such that there is an identity e ($e \circ x = x$ and $x \circ e = x$ for all $x \in G$), and each $x \in G$ has an inverse in G (an element y such that $x \circ y = e$ and $y \circ x = e$).

The full notation is $(G, \circ)$, but if $\circ$ is understood, then we write G for short. Thus to check that G is a group, you have to check four things:

G is closed under the operation,
the operation is associative
there is an identity
every element has an inverse

(note that 'G is closed under the operation' means that if $x, y \in G$ then $x \circ y \in G$; this is not listed as an axiom, since it is part of what is meant by '$\circ$ is a binary operation on G'; also notice that you should avoid saying that G is 'closed', as that means something else - use the *whole phrase* 'G is closed under $\circ$').

1.3 Examples. (1) $(\mathbb{Z}_4, +)$ is a group. It is closed under the operation, since any remainder modulo 4 lies in the set $\{0, 1, 2, 3\}$. Associative. For example $(3 + 2) + 1 =$

$1 + 1 = 2$, while $3 + (2 + 1) = 3 + 3 = 2$. In fact it inherits associativity from $+$ for $\mathbb{Z}$. Identity element 0. The inverses of $0, 1, 2, 3$ are $0, 3, 2, 1$ respectively.

(2) More generally $(\mathbb{Z}_n, +)$ is a group for any n.

(3) $(\mathbb{Z}_4, \times)$ is not a group. It is closed under the operation as before, and the operation is associative (it inherits it from $\mathbb{Z}$). The identity element is 1. But 0 has no inverse, since there is no y with $0 \cdot y = 1$.

(4) The subset $\{1, 2, 3\}$ of $\mathbb{Z}_4$ is not a group under $\times$, since it is not closed under the operation.

(5) The subset $\{1, 3\}$ of $\mathbb{Z}_4$ *is* a group under $\times$. It is closed under the operation since $1 \times a = a$ and $3 \times 3 = 1$. It inherits associativity from $\times$ on $\mathbb{Z}_4$. The identity element is 1. The inverses of $1, 3$ are $1, 3$ respectively.

(6) $(\mathbb{R}, +)$ is a group. The identity element is 0, and the inverse of x is $-x$.

(7) $(\mathbb{R}, \times)$ is not a group. It is closed under $\times$, $\times$ is associative, the identity is 1, but 0 has no inverse.

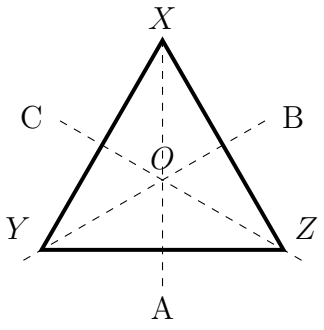
(8) We define $\mathbb{R}^* = \{x \in \mathbb{R} : x \neq 0\} = \mathbb{R} \setminus \{0\}$. Then $(\mathbb{R}^*, \times)$ is a group. The identity is 1, and the inverse of x is $1/x$.

(9) $(\mathbb{R}, -)$ is not a group. The operation is not associative, e.g. $(1 - 2) - 3 = -4$ but $1 - (2 - 3) = 2$.

1.4 Definition. We say that a group $(G, \circ)$ is *abelian* if the operation $\circ$ is commutative, that is, $x \circ y = y \circ x$ for all $x, y \in G$.

Examples. $(\mathbb{Z}_4, +)$, $(\mathbb{R}, +)$, $(\mathbb{R}^*, \times)$ are abelian groups. This next example is not.

1.5 Example. (The dihedral group D_3) Consider an equilateral triangle XYZ with centroid O.



The isometries of the plane preserving the triangle are

I	do nothing
R	rotate about O by angle $2\pi/3$ (clockwise)
S	rotate about O by angle $2\pi/3$ (anticlockwise)
A	reflect in line XO
B	reflect in line YO
C	reflect in line ZO.

The *dihedral group* D_3 is given by the set $D_3 = \{I, R, S, A, B, C\}$ with the operation $\circ$ given by composition. Specifically, $P \circ Q$ means 'do Q first, then P'. For example $A \circ R = C$, $R \circ A = B$, $A \circ B = S$. Then $\circ$ gives the following group table (where we write $P \circ Q$ in row P, column Q).

$\circ$	I	R	S	A	B	C
I	I	R	S	A	B	C
R	R	S	I	B	C	A
S	S	I	R	C	A	B
A	A	C	B	I	S	R
B	B	A	C	R	I	S
C	C	B	A	S	R	I

The identity element is I. The inverses of I, R, S, A, B, C are I, S, R, A, B, C respectively.

1.6 Notation. We only use the symbol $+$ for the operation if the group is abelian. We then say that the group is an *additive group*. In this case we denote the identity element by 0 and the inverse of x by $-x$.

But usually we write groups with *multiplicative notation*. We omit the symbol for the operation and just write xy. The identity element is often denoted by e or 1. The inverse of x is denoted by x^{-1}.

1.7 Proposition. *The identity element of a group is unique.*

Proof. Suppose that e and f are identity elements for the group G. Consider ef. Since e is an identity element $ef = f$. Since f is an identity element $ef = e$. Thus $e = f$. □

1.8 Proposition. *Given a group G and an element $x \in G$, there is only one inverse for x, that is, there is only one element y with $xy = 1$ and $yx = 1$.*

Proof. Say $xy = yx = 1$ and $xz = zx = 1$. Then $(yx)z = 1z = z$. But also $(yx)z = y(xz) = y1 = y$. Thus $y = z$. □

1.9 Proposition. *For elements x, y in a group G, the following properties hold.*
(1) If $xy = 1$, then $x = y^{-1}$ and $y = x^{-1}$.
(2) $(xy)^{-1} = y^{-1}x^{-1}$.
(3) $(x^{-1})^{-1} = x$.

1.10 Remark. By associativity a product xyz in a group is unambiguous without brackets. You can also leave out brackets in longer products, e.g. $xyzwu$.

1.11 Definition. Given an element x of a group G, and a positive integer n, we define the power $x^n \in G$ by

$$x^n = \underbrace{xx \dots x}_{n \text{ copies}} \in G.$$

We also define $x^0 = 1$ and negative powers by $x^{-n} = (x^n)^{-1}$. For an additive group we use the alternative notation $nx = x + x + \cdots + x$, $0x = 0$, $(-n)x = -(nx)$.

Example. For D_3, $R^2 = S$, $R^3 = I$, $R^4 = R$, $R^5 = S$, $R^6 = I$, $R^1 = R$, $R^0 = I$, $R^{-1} = S$, ...

Properties. For $n, m \in \mathbb{Z}$ we have $x^n x^m = x^{n+m}$ and $(x^n)^m = x^{nm}$. (In an additive group, $(n+m)x = nx + mx$ and $n(mx) = (nm)x$.)

1.12 Definition. We say that elements x, y in a group G *commute* if $xy = yx$.

Properties. x^n always commutes with x^m since $x^n x^m = x^{n+m} = x^m x^n$.

If x, y commute, then any power of x commutes with any power of y. For example $x^3y^2 = xxxyy = yyxxx = y^2x^3$.

If x, y commute then $(xy)^n = x^n y^n$ for all n. For example $(xy)^3 = xyxyxy = xxxyyy = x^3y^3$.

If x and y do not commute, then it is possible that $(xy)^n \neq x^n y^n$. For example in D_3, $AR = C$ but $A^2R^2 = S$, $C^2 = I$.

1.13 Proposition (Cancellation)**.** *For an elements x, y, z of a group G.*
(i) If $xy = xz$ then $y = z$.
(ii) If $yx = zx$ then $y = z$.

This says that the entries in a row or column of the multiplication table must be distinct. (i) is proved by multiplying the given equation on the left by x^{-1}, using associativity to rearrange, and the fact that $x^{-1}x$ is the identity. Similarly (ii) is proved my multiplying on the right by x^{-1}.

1.14 Proposition. *The multiplication table for a group is a 'Latin square', that is, each row and each column contains all group elements, once each.*

Proof. e.g. for rows. Cancellation shows that each element only occurs once, for if $xy = xy'$ then $y = y'$. Now the element z occurs in the row for x, at column y if $xy = z$, so we may take $y = x^{-1}z$. □

1.15 Example. The *Klein four group* V is a group with 4 elements such that every element x has $x^2 = 1$. This completely determines the multiplication table. Let $V = \{1, a, b, c\}$. Then

$\circ$	1	a	b	c
1	1	a	b	c
a	a	1	c	b
b	b	c	1	a
c	c	b	a	1

1.16 Example. We have additive groups $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_n$. On the other hand $\mathbb{N} = \{0, 1, 2, \dots\}$ is not a group under $+$ as 1 has no inverse.

1.17 Example. The following are groups under $\times$: $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$, $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

1.18 Example. For an integer $n > 1$ the set

$$\mathbb{Z}_n^* = \{x : x \text{ is coprime with } n\}$$

becomes a group under $\times$ modulo n. For example $\mathbb{Z}_2^* = \{1\}$, $\mathbb{Z}_3^* = \{1, 2\}$, $\mathbb{Z}_4^* = \{1, 3\}$, $\mathbb{Z}_5^* = \{1, 2, 3, 4\}$. In general, if n is a prime number, then

$$\mathbb{Z}_n^* = \{1, 2, \ldots, n-1\}.$$

1.19 Example. The set of invertible $n \times n$ matrices with real entries gives a group $\mathrm{GL}_n(\mathbb{R})$. The operation is matrix multiplication. The identity element is the identity matrix. The inverse is the usual inverse of a matrix. The closure of $\mathrm{GL}_n(\mathbb{R})$ under matrix multiplication follows from the fact that if A and B lie in $\mathrm{GL}_n(\mathbb{R})$, then A^{-1} and B^{-1} both exists, and hence so does $B^{-1}A^{-1}$, and then one can see that $B^{-1}A^{-1}$ is the inverse of AB, and it follows that AB is invertible, so lies in $\mathrm{GL}_n(\mathbb{R})$. The identity is certainly invertible, and if A is invertible, so is A^{-1} (and its inverse is A). Similarly one can allow rational or complex entries $\mathrm{GL}_n(\mathbb{Q})$, $\mathrm{GL}_n(\mathbb{C})$.

1.20 Example. The dihedral group D_n is the group of isometries of the plane (rotations and reflections) preserving a regular n-sided polygon.

$$D_n = \{I, R, R^2, \ldots, R^{n-1}, L_1, L_2, \ldots, L_n\}$$

where R is rotation about the centre by angle $2\pi/n$, and $L_1, \ldots, L_n$ are reflections.

1.21 Example. A *frieze* is a strip in the plane $\mathbb{R}^2$ with translational and possibly other symmetry. For example

$$\ldots \mathrm{AAAA} \ldots$$

Let G be the set of isometries of the plane preserving the frieze. It becomes a group under composition.

In the example, suppose that the copies of A are centred at the points with coordinates $(2n, 0)$. Let T_n be the translation $T_n(x, y) = (x + 2n, y)$. Let R_n be the reflection in the line $x = n$, so $R_n(x, y) = (2n - x, y)$. Then

$$G = \{T_n : n \in \mathbb{Z}\} \cup \{R_n : n \in \mathbb{Z}\}.$$

The multiplication table is

$\circ$	T_n	R_n
T_m	T_{m+n}	R_{m+n}
R_m	R_{m-n}	T_{m-n}

For example $(T_m \circ R_n)(x, y) = T_m(2n - x, y) = (2m + 2n - x, y) = R_{m+n}(x, y)$.
Of course, this is only *part* of the multiplication table, which is infinite, but it is sufficient to indicate how the whole table works.

More generally there are wallpaper groups and crystallographic groups.

1.22 Example. Let CUB be the group of rotations of $\mathbb{R}^3$ preserving a cube. A rotation can send a fixed face to any one of the six faces, and with any of 4 orientations. Thus there are $6 \times 4 = 24$ rotations in all, including the identity element. They can be classified as:

- Identity element (1 element)
- Rotation by angle $2\pi/3$ or $4\pi/3$ about an axis through opposite vertices (2 angles times 4 axes gives 8 elements)
- Rotation by angle $\pi/2$, π or $3\pi/2$ about an axis through two face centres (3 angles times 3 axes gives 9 elements).
- Rotation by angle π about an axis through two edge midpoints (1 angle times 6 axes gives 6 elements).

1.23 Example (Non-examinable)**.** Let G be the group of symmetries of the Rubik's cube. Pick it up, do some operations to it, and put it down again. Total number of elements is $24 \times 8! \times 3^7 \times (12!/2) \times 2^{11} = 1{,}038{,}048{,}078{,}587{,}756{,}544{,}000$.

1.24 Definition. Let $(G, \circ)$ be a group. A *subgroup* of $(G, \circ)$ is a subset H of G such that H becomes a group with the same operation $\circ$.

We write $H \leq G$ to mean that H is a subgroup of G.

Every group G has a *trivial subgroup*, the subset $\{1\}$ consisting of the identity element on its own. Also the set $H = G$, the whole group, is a subgroup. A *proper subgroup* is one which is different from G.

1.25 Examples. (1) The subgroups of $(\mathbb{Z}_4, +)$ are $\{0\}$, $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and $\{0, 2\}$. For example, suppose that H is a subgroup with $3 \in H$. Then also $3 + 3 = 2 \in H$, so $3 + 2 = 1 \in H$ and also $0 \in H$ so $H = \mathbb{Z}_4$.

(2) The subgroups of D_3 are $\{I\}$, $D_3 = \{I, R, S, A, B, C\}$, $\{I, R, S\}$, $\{I, A\}$, $\{I, B\}$, $\{I, C\}$. For example suppose that H is a subgroup with $A, B \in H$. Then $I \in H$. Also $AB = S$ and $BA = R$ are in H. Then $BS = C \in H$. Thus $H = D_3$.

(3) The set of rotations of a regular n-gon, $\{I, R, R^2, \ldots, R^{n-1}\}$, is a subgroup of D_n.

(4) $\mathbb{Z}$ is a subgroup of $\mathbb{Q}$ (additive groups), and these are both subgroups of $\mathbb{R}$, and these are all subgroups of $\mathbb{C}$.

(5) $\mathbb{Q}^*$ is a subgroup of $\mathbb{R}^*$, and these are both subgroups of $\mathbb{C}^*$.

1.26 Lemma. *If H is a subgroup of G, then*
(i) they have the same identity element (in particular H contains the identity of G), and
(ii) the inverse of any element of H is the same whether you use the group structure of H or that of G.

Proof. (i) Denote them by 1_H and 1_G. Then $1_G 1_H = 1_H$ and $1_H 1_H = 1_H$ so $1_G 1_H = 1_H 1_H$, and hence $1_G = 1_H$ by cancellation. (ii) Say $h \in H$ has inverse y in H. Then $hy = 1 = yh$. Then y is the inverse of h in G. □

1.27 Theorem (Subgroup criterion)**.** *Let $(G, \circ)$ be a group. A subset H of G is a subgroup if and only if it satisfies the following properties*
(i) $1 \in H$,
(ii) $xy \in H$ *for all* $x, y \in H$, *and*
(iii) $x^{-1} \in H$ *for all* $x \in H$.

Proof. First suppose that (i), (ii) and (iii) hold. Then (ii) says that H is closed under $\circ$, and it inherits associativity from G. Then $1 \in H$ by (i), and it is an identity for H. Also each element $x \in H$ has an inverse in $x^{-1} \in H$ by (iii). Thus H is a subgroup.

Conversely suppose that H is a subgroup. Then since H is closed under $\circ$, (ii) holds. Now (i) and (iii) follow from the lemma. □

1.28 Examples. (1) The set $\mathbb{R}^{>0} = \{x \in \mathbb{R} : x > 0\}$ is a subgroup of $\mathbb{R}^*$.

The identity element of $\mathbb{R}^*$ is 1, and this is in $\mathbb{R}^{>0}$. Also, if $x, y \in \mathbb{R}^{>0}$ then $xy \in \mathbb{R}^{>0}$. Also, if $x \in \mathbb{R}^{>0}$ then $x^{-1} = 1/x \in \mathbb{R}^{>0}$.

On the other hand $\mathbb{R}^{>1}$, $\mathbb{R}^{\geq 1}$ and $\mathbb{R}^{<0}$ are not subgroups.

(2) The set $H = \{2^n : n \in \mathbb{Z}\}$ is a subgroup of $\mathbb{R}^*$.

We have $1 = 2^0 \in H$. Also $2^n \times 2^m = 2^{n+m} \in H$ and $(2^n)^{-1} = 2^{-n} \in H$.

(3) For n a positive integer, let $H = \{z \in \mathbb{C} : z^n = 1\}$. For example for $n = 4$ this is $H = \{1, -1, i, -i\}$. This is a subgroup of $\mathbb{C}^*$.

We have $1^n = 1$, so $1 \in H$. If $z, w \in H$ then $(zw)^n = z^n w^n = 1$, so $zw \in H$. If $z \in H$ then $(1/z)^n = 1/(z^n) = 1$ so $1/z \in H$.

Similarly $\{z \in \mathbb{C} : |z| = 1\}$ is a subgroup of $\mathbb{C}^*$.

(4) $\mathbb{R}^*$ is a subset of $\mathbb{R}$, but it is not a subgroup. Recall that the operation for $\mathbb{R}$ is $+$. Now $1, -1 \in \mathbb{R}^*$ but $1 + (-1) \notin \mathbb{R}^*$, so it is not closed under the operation.

(5) For any integer k, the set $k\mathbb{Z}$ of multiples of k is a subgroup of $\mathbb{Z}$. For example $2\mathbb{Z}$ is the set of even integers.

The identity element for $\mathbb{Z}$ is $0 \in k\mathbb{Z}$. If $x, y \in k\mathbb{Z}$, then so is $x + y$. If $x \in k\mathbb{Z}$, then so is $-x$.

(6) The *special linear group* is $\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) : \det A = 1\}$. This is a subgroup of $\mathrm{GL}_n(\mathbb{R})$.

The identity matrix has determinant 1. If $A, B \in \mathrm{SL}_n(\mathbb{R})$ then $\det(AB) = \det(A)\det(B) = 1$, so $AB \in \mathrm{SL}_n(\mathbb{R})$. If $A \in \mathrm{SL}_n(\mathbb{R})$ then $\det(A^{-1}) = 1/\det(A) = 1$ so $A^{-1} \in \mathrm{SL}_n(\mathbb{R})$.

(7) The set

$$H = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a \in \mathbb{R}^*, b \in \mathbb{R} \right\}$$

is a subgroup of $\mathrm{GL}_2(\mathbb{R})$. The identity element for $\mathrm{GL}_2(\mathbb{R})$ is the identity matrix I, and it belongs to H. If $A, A' \in H$, say

$$A = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad A' = \begin{pmatrix} a' & b' \\ 0 & 1 \end{pmatrix},$$

then

$$AA' = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a' & b' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} aa' & ab'+b \\ 0 & 1 \end{pmatrix} \in H, \quad \text{and}$$

$$A^{-1} = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}^{-1} = \frac{1}{a} \begin{pmatrix} 1 & -b \\ 0 & a \end{pmatrix} = \begin{pmatrix} 1/a & -b/a \\ 0 & 1 \end{pmatrix} \in H.$$

Thus H is a subgroup.

1.29 Definition. The *order* of a group G, denoted by $|G|$, is the number of elements in the set G, either a positive integer or infinity.

For example the order of $\mathbb{Z}_4$ or $\mathbb{Z}_5^*$ is 4 and the group $\mathbb{R}^*$ is of infinite order.
Do not confuse with the following.

1.30 Definition. The *order* of an element x of a group G is the smallest integer $n > 0$ such that $x^n = 1$. If no such n exists we say that x has infinite order. (In an additive group the condition is $nx = 0$.)

The identity element has order 1. The order of every other element is > 1.

1.31 Examples. (1) In D_3 the orders of the elements I, R, S, A, B, C are 1, 3, 3, 2, 2, 2 respectively.

(2) In $\mathbb{Z}_7^*$ we have $2^2 = 4$, $2^3 = 1$, so 2 has order 3. Also $3^2 = 2$, $3^3 = 6$, $3^4 = 4$, $3^5 = 5$, $3^6 = 1$, so 3 has order 6. The orders of the elements are 1, 2, 3, 4, 5, 6 are 1 , 3, 6, 3, 6, 2 respectively.

(3) In $\mathbb{C}^*$ 1 has order 1, -1 has order 2, i has order 4, but 2 has order ∞, etc.

(4) In $\mathbb{Z}_4$ the orders of the elements 0, 1, 2, 3 are 1, 4, 2, 4 respectively.

1.32 Theorem. *Suppose $x \in G$.*
(i) If x has infinite order, then all powers x^k ($k \in \mathbb{Z}$) are distinct. In particular $x^k = 1$ if and only if $k = 0$.
(ii) If x has finite order n, then as k increases, the powers x^k repeat in cycles of length n. In particular $x^k = 1$ if and only if k is a multiple of n (even for negative k).

Proof. (i) Suppose that $x^j = x^k$ where $j < k$. Then $x^{k-j} = 1$, contrary to x having infinite order.
(ii) The first n powers $x^0, x^1, x^2, \ldots, x^{n-1}$ are all distinct, for if $x^j = x^k$ where $0 \le j < k \le n-1$ then $x^{k-j} = 1$ and $0 < k - j < n$, contrary to x having order n. Now for any integer N we can divide by n giving an integer quotient q and remainder r with $0 \le r \le n-1$. Then $N = nq + r$, and $x^N = x^{nq+r} = (x^n)^q x^r = 1^q x^r = 1x^r = x^r$. $\square$

2 Structure of groups

2.1 Definition. If x is an element of a group G we let

$$\langle x \rangle = \{x^n : n \in \mathbb{Z}\}.$$

(or in additive notation $\langle x \rangle = \{nx : n \in \mathbb{Z}\}$). It is a subgroup of G. We call it the *subgroup of G generated by x*. We say that G is *generated by x*, or that x is a *generator* for G if $G = \langle x \rangle$. We say that G is a *cyclic* group if it has a generator.

2.2 Lemma. *The order of the group $\langle x \rangle$ is equal to the order of the element x.*

Proof. Follows from Theorem 1.32. □

2.3 Theorem. *A finite group of order n is cyclic if and only if it contains an element of order n.*

Proof. If G is cyclic, then any generator is of order n. Conversely if G contains an element of order n, then $\langle x \rangle$ is a subgroup of G with the same number of elements, so we must have $\langle x \rangle = G$. □

2.4 Examples. (1) The group of rotations of an n-gon $\{I, R, R^2, \ldots, R^{n-1}\}$ is cyclic. R is a generator.

(2) The set of n-th roots of unity forms a group $\{z \in \mathbb{C} : z^n = 1\}$ under multiplication. It is cyclic. The element $e^{2\pi i/n}$ is a generator.

(3) $\mathbb{Z}_n$ is cyclic for every $n \geq 1$. The element 1 is a generator.

(4) For the group $\mathbb{Z}_5^* = \{1, 2, 3, 4\}$ we have $\langle 1 \rangle = \{1\}$, $\langle 2 \rangle = \{2, 4, 3, 1\}$, $\langle 3 \rangle = \{3, 4, 2, 1\}$. So it is cyclic. The possible generators are 2 and 3.

(5) The Klein four group $V = \{1, a, b, c\}$ is not cyclic, as 1 has order 1 and a, b, c all have order 2.

(6) $\mathbb{Z}$ is cyclic. The element 1 is a generator.

(7) $\mathbb{R}$ is not cyclic. There is no real number x with $\{nx : n \in \mathbb{Z}\} = \mathbb{R}$.

(8) $\mathbb{R}^{>0}$ is not cyclic. Suppose $x > 0$ has the property that $\{x^n : n \in \mathbb{Z}\} = \mathbb{R}^{>0}$. Then $x \neq 1$, so $x < 1$ or $x > 1$. By replacing x by $1/x$ if necessary, assume that $x > 1$. Then no member of $(1, x)$ is equal to an integral power of x, so $\{x^n : n \in \mathbb{Z}\} \neq \mathbb{R}^{>0}$ after all.

2.5 Proposition. *Any cyclic group is abelian.*

Proof. $(x^n)(x^m) = x^{n+m} = (x^m)(x^n)$. □

2.6 Theorem. *Any subgroup of $\mathbb{Z}$ is of the form $k\mathbb{Z}$ for some k (which is the same as $\langle k \rangle$ in this case), so is cyclic.*

Proof. Suppose H is a subgroup of $\mathbb{Z}$. If $H = \{0\}$ then $H = 0\mathbb{Z}$, so we may suppose that H contains a non-zero element. Then H contains a positive element. Let $k > 0$ be minimal with $k \in H$. Then $k\mathbb{Z} \subseteq H$. Suppose that $h \in H$ and $h \notin k\mathbb{Z}$. Dividing by k gives integer quotient q and remainder r with $0 < r < k$. Then by the choice of k we have $r \notin H$. But then $h = qk + r$, so $r = h - qk \in H$. Contradiction. □

2.7 Theorem. *Any subgroup of a cyclic group is cyclic.*

Proof. Suppose $G = \langle x \rangle$ and $H \leq G$. Define

$$K = \{k \in \mathbb{Z} : x^k \in H\}.$$

This is a subgroup of $\mathbb{Z}$, for $x^0 = 1 \in H$, so $0 \in K$. Also if $k, j \in K$ then $x^k, x^j \in H$, so $x^{k+j} = x^k x^j \in H$, so $k + j \in K$, and $x^{-k} = (x^k)^{-1} \in H$, so $-k \in K$. Thus by the previous theorem $K = n\mathbb{Z}$ for some n. Then $H = \{x^k : k \in K\} = \{x^{nj} : j \in \mathbb{Z}\} = \{((x^n)^j : j \in \mathbb{Z}\} = \langle x^n \rangle$, so H is cyclic. □

2.8 Definition. If G and H are groups, then we consider the cartesian product

$$G \times H = \{(g, h) : g \in G, h \in H\}$$

with the operation $\circ$ defined by

$$(g, h) \circ (g', h') = (gg', hh').$$

It is easy to see that it is a group. We call it the *direct product* of G and H. The identity element is $1 = (1_G, 1_H)$. The inverse of (g, h) is (g^{-1}, h^{-1}).

(If G and H are additive groups we use the notation $(g, h) + (g', h') = (g + g', h + h')$.)

We have $|G \times H| = |G| \cdot |H|$ (which applies even in the infinite case).

2.9 Example. Let $G = \{I, R, S\}$ be the group of rotations of an equilateral triangle, and $H = \mathbb{Z}_4^* = \{1, 3\}$. Then

$$G \times H = \{(I, 1), (I, 3), (R, 1), (R, 3), (S, 1), (S, 3)\}$$

For example $(I, 3)(R, 3) = (R, 1)$.

2.10 Lemma. *The order of $(g, h) \in G \times H$ is the least common multiple of the orders of g and h (or ∞ if g or h has infinite order).*

Proof. Suppose that the orders n and m of g and h are finite. The order of (g, h) is the least positive N with $(g, h)^N = 1 = (1_{G,H})$, so with $g^N = 1_G$ and $h^N = 1_H$. This holds if and only if N is a common multiple of n and m. □

2.11 Theorem. *If G and H are finite cyclic groups, then $G \times H$ is cyclic if and only the orders of G and H are coprime.*

Proof. Let $G = \langle x \rangle$ and $H = \langle y \rangle$ have order n and m.

Suppose n and m are coprime. Then (x, y) has order the least common multiple of n and m, but since they are coprime this is nm. This $G \times H$ is cyclic.

Conversely suppose that n and m are not coprime. Then their least common multiple ℓ is $< nm$. Then for any integers j, k we have $(x^j, y^k)^\ell = (x^{j\ell}, y^{k\ell}) = (1_G, 1_H)$ since $j\ell$ is a multiple of the order of n and $k\ell$ is a multiple of m. Thus every element of $G \times H$ has order $\leq \ell$. Thus no element has order equal to the order of $G \times H$, so $G \times H$ is not cyclic. □

2.12 Examples. (1) The group $G \times H$ as above is cyclic of order 6. $a = (R, 3)$ is a generator. Its powers are $a^0 = (I, 1)$, $a^1 = (R, 3)$, $a^2 = (S, 1)$, $a^3 = (I, 3)$, $a^4 = (R, 1)$, $a^5 = (S, 3)$, $a^6 = (I, 1)$.

(2) $\mathbb{Z}_7 \times \mathbb{Z}_8$ is cyclic of order 56. The element $(1, 1)$ is a generator.

(3) $\mathbb{Z}_6 \times \mathbb{Z}_9$ is not cyclic.

Recall that if X and Y are sets, a mapping $f : X \to Y$ gives an element $f(x) \in Y$ for each element $x \in X$.
It is *injective* (or *1-1*) if $f(x) = f(x')$ implies $x = x'$.
It is *surjective* (or *onto*) if for all $y \in Y$ there is some $x \in X$ with $\theta(x) = y$.
It is *bijective* if it is injective and surjective.

2.13 Definition. Let $(G, \circ)$ and $(H, \circ)$ be groups. A mapping $\theta : G \to H$ is a *homomorphism* if $\theta(g \circ g') = \theta(g) \circ \theta(g')$ for all $g, g' \in G$. It is an *isomorphism* if in addition it is a bijection.

We say that groups G and H are *isomorphic*, and write $G \cong H$, if there is an isomorphism $\theta : G \to H$.

2.14 Lemma. *If $\theta : G \to H$ is a homomorphism, then $\theta(1_G) = 1_H$ and $\theta(g^{-1}) = (\theta(g))^{-1}$ for all $g \in G$.*

Proof. $\theta(1_G)\theta(1_G) = \theta(1_G^2) = \theta(1_G) = \theta(1_G)1_H$, so $\theta(1_G) = 1_H$ by cancellation. Now $gg^{-1} = 1_G$, so $\theta(g)\theta(g^{-1}) = \theta(1_G) = 1_H$, giving $\theta(g^{-1}) = (\theta(g))^{-1}$. □

2.15 Examples. (1) The groups V and $\mathbb{Z}_{12}^* = \{1, 5, 7, 11\}$ are isomorphic.

The multiplication tables are

	1	a	b	c
1	1	a	b	c
a	a	1	c	b
b	b	c	1	a
c	c	b	a	1

	1	5	7	11
1	1	5	7	11
5	5	1	11	7
7	7	11	1	5
11	11	7	5	1

These become the same if we identify 1 with itself, $a = 5$, $b = 7$, $c = 11$. Thus the mapping $\theta : V \to \mathbb{Z}_{12}^*$, $\theta(1) = 1$, $\theta(a) = 5$, $\theta(b) = 7$, $\theta(c) = 11$ is an isomorphism.

(2) The group $\mathbb{Z}_4$ is isomorphic to the group $G = \{I, R, R^2, R^3\}$ of rotations of a square. The multiplication tables are

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\circ$	I	R	R^2	R^3
I	I	R	R^2	R^3
R	R	R^2	R^3	I
R^2	R^2	R^3	I	R
R^3	R^3	I	R	R^2

Thus the mapping $\theta : \mathbb{Z}_4 \to G$ defined by $\theta(j) = R^j$ is an isomorphism.

(3) The exponential map defines an isomorphism $\exp : \mathbb{R} \to \mathbb{R}^{>0}$, where $\mathbb{R}^{>0}$ is the group of positive real numbers under multiplication.

2.16 Proposition. *Suppose that $\theta : G \to H$ is an isomorphism. Then:*
(i) $|G| = |H|$.
(ii) $\theta(1_G) = 1_H$.
(iii) $\theta(g^{-1}) = (\theta(g))^{-1}$ for all $g \in G$.
(iv) For all $g \in G$ the elements g and $\theta(g)$ have the same order.
(v) For each n, the groups G and H have the same number of elements of order n.
(v) G is abelian if and only if H is abelian.
(vi) G is cyclic if and only if H is cyclic.

Proof. Straightforward, since G and H have the same multiplication table, and all of these properties can be read off from the multiplication table. □

2.17 Example. The groups V and $G = \{I, R, R^2, R^3\}$ are not isomorphic. In V all elements have order 1 or 2. In G there are elements of order 4, namely R and R^3. Alternatively, G is cyclic but V is not.

2.18 Theorem. *Two cyclic groups are isomorphic if and only if they have the same order.*

Proof. If they are isomorphic they must have the same order. For the converse, suppose the groups are $G = \langle x \rangle$ and $H = \langle y \rangle$ and that they have the same order.

If the order is infinite, then the elements x^k are all distinct, so we can define $\theta : G \to H$ by $\theta(x^k) = y^k$. It is a homomorphism since $\theta(x^j x^k) = \theta(x^{j+k}) = y^{j+k} = y^j y^k = \theta(x^j)\theta(x^k)$. Clearly it is bijective, so it is an isomorphism.

Thus suppose the order is finite, say n. The powers x^k repeat with period n, and similarly the powers y^k. Thus we can again define $\theta : G \to H$ with $\theta(x^k) = y^k$, and again get an isomorphism. □

2.19 Corollary (Chinese Remainder Theorem)**.** $\mathbb{Z}_n \times \mathbb{Z}_m \cong \mathbb{Z}_{nm}$ *if and only if n and m are coprime.*

Proof. If n and m are coprime, then $\mathbb{Z}_n \times \mathbb{Z}_m$ is cyclic. Also $\mathbb{Z}_{nm}$ is cyclic of the same order, so they must be isomorphic. If n and m are not coprime then $\mathbb{Z}_n \times \mathbb{Z}_m$ is not cyclic, so not isomorphic to $\mathbb{Z}_{nm}$. □

2.20 Example. $\mathbb{Z}_5 \times \mathbb{Z}_{12} \cong \mathbb{Z}_{20} \times \mathbb{Z}_3$, as both are isomorphic to $\mathbb{Z}_{60}$ or to $\mathbb{Z}_5 \times \mathbb{Z}_4 \times \mathbb{Z}_3$.

We now move on to the statement and proof of Lagrange's Theorem, which says that the order of a subgroup of a finite group G divides the order of G. The key idea of the proof is to show that G can be 'cut up' into pieces of equal size, one of which is H, from which it follows that the order of G is equal to the order of H times the number of pieces. The pieces are called 'right cosets' (there is an alternative proof using 'left cosets'). The most efficient way to get hold of these is as equivalence classes under a suitable equivalence relation.

2.21 Theorem. *(Lagrange): If H is a subgroup of the finite group G, then $|H|$ divides $|G|$.*

Proof. We define $\sim$ on G by letting $x \sim y$ if $xy^{-1} \in H$. We verify that this is an equivalence relation on G:
reflexivity: $x \sim x$ since $xx^{-1} = 1 \in H$ (as H contains the identity)
symmetry: if $x \sim y$ then $xy^{-1} \in H$. Hence $yx^{-1} = ((yx^{-1})^{-1})^{-1} = (xy^{-1})^{-1} \in H$ so $y \sim x$ (as H is closed under inverses)
transitivity: if $x \sim y$ and $y \sim z$ then $xy^{-1}, yz^{-1} \in H$, so $xz^{-1} = (xy^{-1})(yz^{-1}) \in H$ so $x \sim z$ (as H is closed under the operation).
Since $\sim$ is an equivalence relation, it partitions G into $\sim$-classes. The $\sim$-class containing x is $\{y : y \sim x\} = \{y : yx^{-1} \in H\} = \{y : yx^{-1} = h, \text{ some } h \in H\} = \{y : y = hx, \text{ some } h \in H\}$. This set is written Hx, and is called the *right coset* of H in G containing x.
To sum up, every member of G lies in some right coset of H, and the right cosets form a partition of G. Finally we see that each right coset Hx has $|H|$ members (noting that $H.1 = H$, so that H is itself a right coset). Map H to Hx by f where $f(h) = hx$. By definition of Hx this maps H onto Hx, and f is 1–1, since if $f(h_1) = f(h_2)$, then $h_1x = h_2x$, so by cancellation, $h_1 = h_2$. This f is a 1–1 map from H onto Hx, so Hx has $|H|$ members. □

2.22 Definition. Let H be a subgroup of a group G. A *(right) coset* of H in G is a subset of the form

$$Hx = \{hx : h \in H\}$$

for some $x \in G$. If G is an additive group we use the notation $H + x = \{h + x : h \in H\}$ instead.
Note that even if G is infinite, we still have the notion of 'right coset'. Finiteness is just used in the final part of the proof of Lagrange's Theorem.

2.23 Example. Recall the multiplication table for D_3.

$\circ$	I	R	S	A	B	C
I	I	R	S	A	B	C
R	R	S	I	B	C	A
S	S	I	R	C	A	B
A	A	C	B	I	S	R
B	B	A	C	R	I	S
C	C	B	A	S	R	I

(1) To find the cosets of the subgroup $H = \{I, R, S\}$ in D_3 we compute

$$HI = \{I, R, S\} = HR = HS, \quad HA = \{A, B, C\} = HB = HC.$$

Thus there are two right cosets $\{I, R, S\}$ and $\{A, B, C\}$.

(2) To find the cosets of the subgroup $K = \{I, A\}$ in G we compute

$$KI = \{I, A\} = KA, \quad KR = \{R, C\} = KC, \quad KS = \{S, B\} = KB$$

Thus there are three right cosets $\{I, A\}$, $\{R, C\}$ and $\{S, B\}$.

(3) The cosets of the subgroup $3\mathbb{Z} = \{\dots, -6, -3, 0, 3, 6, \dots\}$ in $\mathbb{Z}$ are

$$\begin{aligned}
&\mathbb{Z} + 0 = \{\dots, -6, -3, 0, 3, 6, \dots\} = \mathbb{Z} + 3 = \mathbb{Z} + 6 = \cdots = \mathbb{Z} + (-3) = \dots, \\
&\mathbb{Z} + 1 = \{\dots, -5, -2, 1, 4, 7, \dots\} = \mathbb{Z} + 4 = \mathbb{Z} + (-2) = \dots, \\
&\mathbb{Z} + 2 = \{\dots, -4, -1, 2, 5, 8, \dots\} = \mathbb{Z} + 5 = \mathbb{Z} + (-1) = \dots
\end{aligned}$$

From the above discussion, we read off the following facts about (right) cosets of a subgroup H of a group G:
(i) H itself is a coset.
(ii) Every element of G belongs to a coset.
(iii) If y is in a coset Hx, then $Hy = Hx$.
(iv) Two cosets Hx and Hy are either the same, or they are disjoint, $Hx \cap Hy = \emptyset$.
(v) $Hx = Hy$ if and only if $xy^{-1} \in H$.

2.24 Definition. If H is a subgroup of a finite group G, the *index* of H in G is the number of different cosets of H in G. We denote it by $|G : H|$.

2.25 Theorem. *If H is a subgroup of a finite group G, then $|G| = |H|.|G : H|$.*

This follows at once from Lagrange's Theorem.

2.26 Corollary. *The order of an element of a finite group divides the order of the group.*

Proof. The order of x is the same as the order of the subgroup $\langle x \rangle$ of G. □

2.27 Corollary. *Any group of prime order is cyclic.*

Proof. If G has order p, then any non-identity element has order $\neq 1$, so must have order p. Thus it generates the group. □

2.28 Corollary. *If G is a group of order n, then $x^n = 1$ for all $x \in G$.*

Proof. The order d of x divides n, so $n = dk$ for some k. Then $x^n = x^{dk} = (x^d)^k = 1^k = 1$. □

2.29 Corollary (Fermat's little theorem)**.** *If p is a prime number and a is coprime to p, then $a^{p-1} \equiv 1 \pmod{p}$ (which means that $a^{p-1} - 1$ is a multiple of p).*

Proof. Consider $a \in \mathbb{Z}_p^*$. We have $(a)^{p-1} = 1$, so $a^{p-1} = 1$. □

We are interested in classifying groups of small order up to isomorphism. For each positive integer n there is a cyclic group of order n. We denote it by C_n. It is unique up to isomorphism. Groups of prime order are cyclic. Thus, up to isomorphism, the only groups of orders 2, 3, 5 and 7 are C_2, C_3, C_5 and C_7.

2.30 Theorem. *Up to isomorphism the groups of order 4 are the cyclic group C_4 and the Klein four group $V \cong C_2 \times C_2$.*

Proof. If it is not cyclic, then every element has order 1 or 2, so every element has $x^2 = 1$. This determines the multiplication table to be that of the Klein four group. □

2.31 Theorem. *Up to isomorphism the groups of order 6 are the cyclic group C_6 and the dihedral group D_3.*

Proof. Suppose the group is not cyclic, so every element has order 1,2 or 3. Suppose first that there is no element of order 3. Then every element has order 1 or 2, from which it follows that the group is abelian, since $yx = y^{-1}x^{-1} = (xy)^{-1} = xy$. If a and b are distinct elements of order 2 then $\{1, a, b, ab\}$ is a subgroup if G. But this is impossible by Lagrange's Theorem.

Thus there is an element of order 3, say r. Then $H = \{1, r, r^2\}$ is a subgroup of G. Let x be an element not in this subgroup. Then $G = H \cup Hx = \{1, r, r^2, x, rx, r^2x\}$. If x has order 3 then we can't have $x^2 \in \{1, x, rx, r^2x\}$, so $x^2 \in \{r, r^2\}$, but then $x = (x^2)^2 \in \{r^2, r^4\} = \{r^2, r\}$, contrary to $x \notin \{1, r, r^2\}$. Thus x has order 2. Similarly rx and r^2x have order 2.

Then $rxrx = 1$, so, multiplying on the left by r^2 and on the right by x, one gets $xr = r^2x$. This is enough to fill in the multiplication table for $G = \{1, r, r^2, x, rx, r^2x\}$, giving the same table as D_3. □

2.32 Remark. Recall that we think of complex numbers as expressions of the form $a + bi$ and multiply them using the rule $i^2 = -1$. The *quaternions* are expressions of the form $a + bi + cj + dk$ with $a, b, c, d \in \mathbb{R}$, with an associative multiplication which can be derived from the rules $i^2 = j^2 = k^2 = -1$ and $ij = k$. For example $ik = i(ij) = (-1)j = -j$, $kj = (ij)j = -i$, ... The set $Q = \{\pm 1, \pm i, \pm j, \pm k\}$ becomes a group under multiplication, with table

	1	i	j	k	-1	$-i$	$-j$	$-k$
1	1	i	j	k	-1	$-i$	$-j$	$-k$
i	i	-1	k	$-j$	$-i$	1	$-k$	j
j	j	$-k$	-1	i	$-j$	k	1	$-i$
k	k	j	$-i$	-1	$-k$	$-j$	i	1
-1	-1	$-i$	$-j$	$-k$	1	i	j	k
$-i$	$-i$	1	$-k$	j	i	-1	k	$-j$
$-j$	$-j$	k	1	$-i$	j	$-k$	-1	i
$-k$	$-k$	$-j$	i	1	k	j	$-i$	-1

2.33 Theorem. *Up to isomorphism the groups of order 8 are $\mathbb{Z}_8$, $\mathbb{Z}_4 \times \mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$, the dihedral group D_4 and the quaternion group Q.*

Proof. Omitted. □

3 Permutations, homomorphisms, symmetric and alternating groups

3.1 Definition. A *permutation* of a set A is a bijective mapping from A to itself, $\pi : A \to A$. The set of all permutations of A forms a group under composition of mappings $\pi \circ \sigma$, where

$$(\pi \circ \sigma)(a) = \pi(\sigma(a))$$

for $a \in A$. The identity element is the identity map id. Since π is bijective, it has an inverse mapping π^{-1}, and that is the inverse to π in this group.

We shall only be interested in permutations of the set $A = \{1, 2, \ldots, n\}$ for n a positive integer. The set of all such permutations is called the *symmetric group of degree n* and denoted by S_n.

3.2 Example. We use *two row notation* for elements of the symmetric group. For example

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 4 & 1 \end{pmatrix}$$

denotes the permutation $\pi : \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4, 5\}$ with $\pi(1) = 3$, $\pi(2) = 5$, $\pi(3) = 2$, $\pi(4) = 4$ and $\pi(5) = 1$. Given another permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 5 & 2 & 3 \end{pmatrix}$$

the composition is

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 4 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 5 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 5 & 2 \end{pmatrix},$$

for example $(\pi \circ \sigma)(3) = \pi(\sigma(3)) = \pi(5) = 1$. Note that $\pi \circ \sigma$ means apply σ first, then π. The identity permutation in S_5 is

$$id = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

Inverses are given by swapping the rows and reordering the columns.

$$\pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 4 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & 5 & 2 & 4 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 4 & 2 \end{pmatrix}.$$

3.3 Proposition. *The group S_n has order $n!$.*

3.4 Definition. Let k, n be a positive integers with $k \leq n$ and let $a_1, a_2, \ldots, a_k$ be distinct elements in the set $\{1, 2, \ldots, n\}$. We denote by $(a_1\ a_2\ \ldots\ a_k)$ the permutation in S_n sending

$$a_1 \mapsto a_2 \mapsto a_3 \mapsto \ldots \mapsto a_k \mapsto a_1$$

and with $a \mapsto a$ for all a not in the list. It is called a *cycle of length k* or a *k-cycle*. A 2-cycle is also called a *transposition*.

For example for S_5, the 3-cycle (2 5 4) is the permutation $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 3 & 2 & 4 \end{smallmatrix}\right)$.

3.5 Remarks. (i) Cycle notation doesn't tell you which S_n you are working in. For example the cycle (2 5 4) could be a permutation in S_n for any $n \geq 5$.

(ii) A k-cycle can be written in k different ways. For example $(2\ 5\ 4) = (5\ 4\ 2) = (4\ 2\ 5)$. A 1-cycle is the identity.

(iii) A k-cycle has order k.

(iv) We say a collection if cycles is *disjoint* if there is no number a occurring in two of them. For example (2 5 4) and (1 3) are disjoint. Disjoint cycles commute, $(2\ 5\ 4)(1\ 3) = (1\ 3)(2\ 5\ 4)$.

3.6 Theorem. *Every permutation can be written as a product of disjoint cycles. The decomposition is essentially unique, apart from the order of the cycles and the different ways of writing a cycle.*

3.7 Example. To write the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 1 & 8 & 5 & 2 & 6 & 3 & 7 & 10 & 9 \end{pmatrix}$$

as a product of disjoint cycles, start with 1, and follow it under repeated applications of the permutation $1 \mapsto 4 \mapsto 5 \mapsto 2$, building up a cycle

$$(1\ 4\ 5\ 2)$$

The cycle closes since $2 \mapsto 1$, which is the number we started with. Then take the first number which hasn't yet appeared, and follow it, $3 \mapsto 8 \mapsto 7 \mapsto 3$, so

$$(1\ 4\ 5\ 2)(3\ 8\ 7)$$

Repeat with the next number which hasn't yet appeared, etc.

$$(1\ 4\ 5\ 2)(3\ 8\ 7)(6)(9\ 10)$$

Finally we can delete any 1-cycles, giving

$$\pi = (1\ 4\ 5\ 2)(3\ 8\ 7)(9\ 10).$$

3.8 Example. We have $S_3 = \{id, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$. The dihedral group D_3 is isomorphic to S_3. Number the vertices of the equilateral triangle $A = 1$, $B = 2$, $C = 3$. There is an isomorphism $\theta : D_3 \to S_3$ sending an isometry to the corresponding permutation of the vertices, $I \mapsto id$, $R \mapsto (1\ 2\ 3)$, $S \mapsto (1\ 3\ 2)$, $L \mapsto (2\ 3)$, $M \mapsto (1\ 3)$, $N \mapsto (1\ 2)$.

3.9 Corollary. *To find the order of a permutation, write it as a product of disjoint cycles and take the least common multiple of their lengths.*

Proof. Write π as a product of disjoint cycles, say $\pi = c_1 c_2 \dots c_k$. The order is the least $d > 0$ with $\pi^d = e$. Since disjoint cycles commute we get $\pi^d = c_1^d c_2^d \dots c_k^d = e$. Now the permutations $c_1^d, \dots, c_k^d$ act on disjoint subsets of $\{1, \dots, n\}$, so the only way that their product can be the identity is if each of them is the identity, so d must be a multiple of the orders of the cycles. □

3.10 Corollary. *Every permutation can be written as a product of transpositions.*

Proof. We have $(a_1\ a_2\ a_3\ \dots\ a_k) = (a_1\ a_k) \dots (a_1\ a_3)(a_1\ a_2)$. □

Let $\mathbf{e}_i$ $(1 \le i \le n)$ be the coordinate vectors in $\mathbb{R}^n$.

3.11 Definition. Given a permutation $\pi \in S_n$, the corresponding *permutation matrix* is the $n \times n$ matrix A_π whose jth column is $\mathbf{e}_{\pi(j)}$, for all j. Equivalently $A_\pi \mathbf{e}_j = \mathbf{e}_{\pi(j)}$. Explicitly $A_\pi = (a_{ij})$ where

$$a_{ij} = \begin{cases} 1 & (\text{if } i = \pi(j)) \\ 0 & (\text{otherwise}) \end{cases}$$

For example if $\pi = (1\ 2\ 4) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} \in S_4$ then $A_\pi = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$.

3.12 Definition. The *sign* or *signature* of a permutation π is $\epsilon(\pi) = \det(A_\pi)$.

3.13 Lemma. *For permutations $\pi, \sigma \in S_n$ we have $A_{\pi\sigma} = A_\pi A_\sigma$ and $\epsilon(\pi\sigma) = \epsilon(\pi)\epsilon(\sigma)$.*

Proof. $A_\pi A_\sigma \mathbf{e}_j = A_\pi \mathbf{e}_{\sigma(j)} = \mathbf{e}_{\pi(\sigma(j))} = A_{\pi\sigma} \mathbf{e}_j$. Then $\epsilon(\pi\sigma) = \det(A_{\pi\sigma}) = \det(A_\pi A_\sigma) = \det(A_\pi)\det(A_\sigma) = \epsilon(\pi)\epsilon(\sigma)$. □

3.14 Definition. A permutation which can be written as a product of an odd/even number of transpositions is called an *odd/even permutation.*

3.15 Theorem. *Every permutation is either odd or even, and not both. The sign of a permutation is 1 if it is even and -1 if it is odd. In particular the sign of a permutation is always in $\{\pm 1\}$.*

Proof. We know that any permutation can be written as a product of transpositions (although this expression is not unique). Also $\epsilon(\pi\sigma) = \epsilon(\pi)\epsilon(\sigma)$. It thus suffices to show that if τ is a transposition then $\epsilon(\tau) = -1$. But if $\tau = (a\ b)$ then A_τ is obtained from the identity matrix by exchanging rows a and b. Now the identity matrix has determinant 1, and exchanging any two rows changes the sign, so $\det A_\tau = -1$. □

3.16 Example. A k-cycle is of sign $(-1)^{k-1}$. Thus, for example the permutation $(1\ 5\ 9\ 2)(3\ 8)(7\ 4\ 9)$ is of sign $(-1)^3(-1)(-1)^2 = 1$, so it is even.

3.17 Definition. The set of even permutations in S_n (which forms a subgroup of S_n) is called the *alternating group A_n of degree n.*

3.18 Proposition. *For $n > 1$, we have $[S_n : A_n] = 2$, and so $|A_n| = n!/2$.*

Proof. Fix a transposition $\tau \in S_n$, for example $\tau = (1\ 2)$. For any odd permutation $\pi \in S_n$ we have $\pi\tau \in A_n$. Then $\tau^2 = e$, so $\pi = (\pi\tau)\tau \in A_n\tau$. Thus $S_n = A_n \cup A_n\tau$. Thus there are only two cosets of A_n in S_n. □

3.19 Theorem (Leibniz formula)**.** *If $A = (a_{ij})$ is an $n \times n$ matrix, then*

$$\det A = \sum_{\pi \in S_n} \epsilon(\pi) a_{\pi(1),1} a_{\pi(2),2} \dots a_{\pi(n),n}.$$

3.20 Examples. For a 2×2 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

the formula gives $\det A = a_{11}a_{22} - a_{21}a_{12}$ since $S_2 = \{id, (1\ 2)\}$. For a 3×3 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

expanding down the first column we get

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

The terms correspond to the permutations $id, (2\ 3), (1\ 2), (1\ 2\ 3), (1\ 3\ 2), (1\ 3)$.

Let G and H be groups. Recall that a mapping $\theta : G \to H$ is a *homomorphism* if $\theta(xy) = \theta(x)\theta(y)$ for all $x, y \in G$.

3.21 Examples. (i) The exponential map gives a homomorphism $\mathbb{R} \to \mathbb{R}^*$ or $\mathbb{C} \to \mathbb{C}^*$ since $\exp(x + y) = \exp(x)\exp(y)$.

(ii) For any group G and any element $g \in G$, there is a homomorphism $\mathbb{Z} \to G$, $n \mapsto g^n$.

(iii) The map $\mathbb{Z} \to \mathbb{Z}_n$, sending a to its remainder on dividing by n is a homomorphism.

(iv) $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$

(v) The mapping $S_n \to \mathrm{GL}_n(\mathbb{R})$ sending a permutation π to the permutation matrix A_π.

(vi) A composition of homomorphisms $G \xrightarrow{\theta} H \xrightarrow{\pi} K$ is a homomorphism. For example the composition of

$$S_n \xrightarrow{\pi \mapsto A_\pi} \mathrm{GL}_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^*$$

is the sign $\epsilon : S_n \to \mathbb{R}^*$, $\epsilon(\pi) = \det(A_\pi)$, so this is a homomorphism.

3.22 Proposition. *If $\theta : G \to H$ is a homomorphism, then*
(i) $\theta(1) = 1$, or, more precisely, $\theta(1_G) = 1_H$.
(ii) $\theta(g^{-1}) = \theta(g)^{-1}$ for $g \in G$.

Proof. (i) $\theta(1_G)\theta(1_G) = \theta(1_G 1_G) = \theta(1_G) = \theta(1_G)1_H$, so $\theta(1_G) = 1_H$ by cancellation.
(ii) $\theta(g^{-1})\theta(g) = \theta(g^{-1}g) = \theta(1_G) = 1_H$, so $\theta(g^{-1}) = \theta(g)^{-1}$. □

3.23 Definition. The *kernel* of a homomorphism $\theta : G \to H$ is the set $\ker\theta = \{g \in G : \theta(g) = 1\}$. It is a subset of G.

The *image* of a homomorphism $\theta : G \to H$ is the set $\mathrm{im}\,\theta = \{\theta(g) : g \in G\}$. It is a subset of H.

3.24 Proposition. *If $\theta : G \to H$ is a homomorphism between two groups, then* $\ker\theta$ *is a subgroup of G and* $\mathrm{im}\,\theta$ *is a subgroup of H.*

Proof. We have $1 \in \ker\theta$. If $g, g' \in \ker\theta$ then $\theta(gg') = \theta(g)\theta(g') = 1 \circ 1 = 1$, so $gg' \in \ker\theta$. If $g \in \ker\theta$ then $\theta(g^{-1}) = \theta(g)^{-1} = 1^{-1} = 1$, so $g^{-1} \in \ker\theta$. Thus $\ker\theta$ is a subgroup of G.
We have $\theta(1) = 1$, so $1 \in \mathrm{im}\,\theta$. If $h, h' \in \mathrm{im}\,\theta$, then $h = \theta(g)$ and $h' = \theta(g')$ for some $g, g' \in G$. Then $hh' = \theta(g)\theta(g') = \theta(gg') \in \mathrm{im}\,\theta$. Also $h^{-1} = \theta(g)^{-1} = \theta(g^{-1}) \in \mathrm{im}\,\theta$. Thus $\mathrm{im}\,\theta$ is a subgroup of H. □

3.25 Examples. (i) The exponential map $\exp : \mathbb{C} \to \mathbb{C}^*$ has kernel is $2\pi i\mathbb{Z}$ and the image is $\mathbb{C}^*$.

(ii) The kernel of the map $\mathbb{Z} \to G$, $k \mapsto g^k$ is $n\mathbb{Z}$, where n is the order of g, and the image is $\langle g \rangle$.

(iii) The kernel of the map $\mathbb{Z} \to \mathbb{Z}_n$ which send an element to its remainder under dividing by n, is $n\mathbb{Z}$. The image is all of $\mathbb{Z}_n$.

(iv) The kernel of the determinant map $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$ is $\mathrm{SL}_n(\mathbb{R})$. The image is $\mathbb{R}^*$.

(v) The kernel of the homomorphism $S_n \to \mathrm{GL}_n(\mathbb{R})$, $\pi \mapsto A_\pi$ is the trivial subgroup $\{id\}$. The image is the set of permutation matrices, that is, matrices of 1s and 0s, such that each row and each column contains exactly one 1.

(vi) The kernel of the sign map $S_n \to \mathbb{R}^*$ is A_n, and the image is the subgroup $\{1, -1\}$ of $\mathbb{R}^*$.

3.26 Proposition. *If $\theta : G \to H$ is a homomorphism, then θ is injective if and only if $\ker\theta = \{1\}$. In this case θ defines an isomorphism $G \cong \operatorname{im}\theta$.*

For example, in (v) above, the subgroup of $\mathrm{GL}_n(\mathbb{R})$ consisting of the permutation matrices is isomorphic to S_n.

Proof. If θ is injective and $x \in \ker\theta$ then $\theta(x) = 1 = \theta(1)$, so since θ is injective, $x = 1$. Thus $\ker\theta = \{1\}$.

Conversely suppose that $\ker\theta = \{1\}$. Suppose that $\theta(x) = \theta(y)$. Then $\theta(xy^{-1}) = \theta(x)\theta(y^{-1}) = \theta(x)\theta(y)^{-1} = \theta(x)\theta(x)^{-1} = 1$. Thus $xy^{-1} \in \ker\theta$, so $xy^{-1} = 1$. Thus $x = y$ and θ is injective.

Now θ gives an homomorphism $G \to \operatorname{im}\theta$ which is injective and surjective, so it is an isomorphism. □

3.27 Theorem. *(i) The group CUB of rotations preserving a cube is isomorphic to S_4.*
(ii) The group TET of rotations preserving a regular tetrahedron is isomorphic to A_4.

Proof. (i) There are four long diagonals (through opposite vertices). We number them 1,2,3,4. Consider the map $\theta : G \to S_4$ sending a rotation to the permutation it induces of the long diagonals. This is a homomorphism. The identity rotation is sent to id, and it is easy to see that none of the other rotations are sent to id. (Rotations about a long diagonal are sent to 3-cycles, rotations about an axis through face centres are sent to 4-cycles or products of two transpositions, and rotations about an axis through two edge midpoints are sent to transpositions.) Thus the homomorphism is injective. Therefore $\operatorname{im}\theta \cong G$, so it has 24 elements. Thus we must have $\operatorname{im}\theta = S_4$.

(ii) Number the vertices of the tetrahedron 1,2,3,4. Consider the map $\theta : TET \to S_4$ sending any rotation to the induced permutation of the vertices. This is clearly injective, and one can check it has image equal to A_4. □

3.28 Theorem (Cayley's Theorem)**.** *Any group of order n is isomorphic to a subgroup of S_n.*

Proof. Let $G = \{g_1, g_2, \ldots, g_n\}$. For each $g_i \in G$, the Latin square property gives a permutation $\pi(g_i) \in S_n$ with $g_i g_j = g_{\pi(g_i)(j)}$ for all j. Now

$$g_i(g_j g_k) = g_i g_{\pi(g_j)(k)} = g_{\pi(g_i)(\pi(g_j)(k))} \quad \text{and} \quad (g_i g_j)g_k = g_{\pi(g_i g_j)(k)}.$$

Thus $\pi(g_i)(\pi(g_j)(k)) = \pi(g_i g_j)(k)$ for all i, j, k, so $\pi(g_i) \circ \pi(g_j) = \pi(g_i g_j)$ for all i, j, so π defines a homomorphism $G \to S_n$. It is injective since if $\pi(g_i) = id$ then $g_i = id$. □

3.29 Definition. Elements x, y of a group G are said to be *conjugate* in G if there is $g \in G$ with $y = g^{-1}xg$. The set of all elements conjugate to a given element x is called a *conjugacy class*. The conjugacy class containing x is

$$\mathrm{conj}_G(x) = \{g^{-1}xg : g \in G\}.$$

3.30 Theorem. *G is the disjoint union of its conjugacy classes.*

Proof. It suffices to prove that conjugacy is an equivalence relation. So define a relation $\sim$ on G by $x \sim y$ if $y = g^{-1}xg$ for some $g \in G$. Then

(reflexive) For any $x \in G$ the condition $x \sim x$ holds since $x = 1^{-1}x1$.

(symmetric) If $x \sim y$ then $y = g^{-1}xg$. Then $x = (g^{-1})^{-1}y(g^{-1})$, so $y \sim x$.

(transitive) If $x \sim y$ and $y \sim z$ then $y = g^{-1}xg$ and $z = h^{-1}yh$. Then $z = (gh)^{-1}x(gh)$, so $x \sim z$. □

3.31 Proposition. *Conjugate elements have the same order.*

Proof. If $y = g^{-1}xg$ and $n > 0$, then

$$y^n = 1 \Leftrightarrow \underbrace{(g^{-1}xg)(g^{-1}xg)\dots(g^{-1}xg)}_{n \text{ copies}} = 1 \Leftrightarrow g^{-1}x^ng = 1 \Leftrightarrow x^n = 1.$$

□

3.32 Examples. (i) The identity element always forms a conjugacy class $\{1\}$ on its own, since $g^{-1}1g = g^{-1}g = 1$.

(ii) If G is abelian, then every conjugacy class consists of one element, since $g^{-1}xg = g^{-1}gx = x$.

(iii) In D_3 the conjugacy classes are $\{I\}$, $\{R, S\}$, $\{A, B, C\}$. For example $ARA^{-1} = CA = S$, $RAR^{-1} = BS = C$.

(iv) In the symmetric group S_n the conjugates of a permutation have the same *cycle type*. For example in S_6,

$$\text{if } g^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ a & b & c & d & e & f \end{pmatrix}, \quad \text{then } g^{-1}(1\ 2\ 6)(3\ 4)(5)g = (a\ b\ f)(c\ d)(e).$$

Thus the conjugacy classes are given by all permutations of the same cycle type. For example for S_4 they are:
– $\{id\}$
– $\{(1\ 2), (1\ 3), (1\ 4), (2\ 3), (2\ 4), (34)\}$
– $\{(1\ 2\ 3), (1\ 3\ 2), (1\ 2\ 4).(1\ 4\ 2), (1\ 3\ 4), (1\ 4\ 3), (2\ 3\ 4), (2\ 4\ 3)\}$
– $\{(1\ 2\ 3\ 4), (1\ 2\ 4\ 3), (1\ 3\ 2\ 4), (1\ 3\ 4\ 2), (1\ 4\ 2\ 3), (1\ 4\ 3\ 2)\}$
– $\{(1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$.

(v) Consider the group CUB of rotations preserving a cube. Let x and y be rotations about axes A and B, by angles θ and ϕ respectively. Then the rotations x and y in G are conjugate if and only if there is a rotation g with $y = g^{-1}xg$, and this is if and only if there is a rotation g sending B to A in such a way that the angles θ and ϕ correspond. Thus the conjugacy classes for G are:
– $\{id\}$
– The 6 rotations by angle π about an axis through edge mid-points.

– The 8 rotations about a long diagonal by angle $2\pi/3$ or $4\pi/3$.
– The 6 rotations about an axis through two face centres by angle $\pi/2$ or $3\pi/2$.
– The 3 rotations about an axis through two face centres by angle π.
Under the isomorphism $G \cong S_4$, the conjugacy classes correspond.

3.33 Definition. If x is an element of a group G, the *centralizer* of x in G is the set $C_G(x) = \{g \in G : gx = xg\}$. It is easy to see that it is a subgroup of G.

3.34 Theorem. *The conjugacy class of $x \in G$ has size $|\mathrm{conj}_G(x)| = [G : C_G(x)]$.*

Proof. $g^{-1}xg = (g')^{-1}xg' \Leftrightarrow xg(g')^{-1} = g(g')^{-1}x \Leftrightarrow g(g')^{-1} \in C_G(x) \Leftrightarrow$ the cosets $C_G(x)g$ and $C_G(x)g'$ are equal. Thus the number of different conjugates of x is equal to the number of different cosets of $C_G(x)$ in G. □

3.35 Definition. A subgroup H of a group G is said to be *normal* if $g^{-1}hg \in H$ for all $h \in H$ and $g \in G$. It is equivalent that H is a union of conjugacy classes. We denote this by $H \lhd G$.

3.36 Theorem. *If $\theta : G \to G'$ is a homomorphism then $\ker\theta$ is a normal subgroup of G.*

Proof. If $x \in \ker\theta$ and $g \in G$ then $\theta(g^{-1}xg) = \theta(g)^{-1}\theta(x)\theta(g) = \theta(g)^{-1}\theta(g) = 1$, so $g^{-1}xg \in \ker\theta$. □

3.37 Examples. (i) If G is abelian, then any subgroup is normal.

(ii) In D_3 the subgroup $\{I, R, S\}$ is normal, but $\{I, A\}$, $\{I, B\}$ and $\{I, C\}$ are not.

(iii) The normal subgroups of S_4 are $\{id\}$, S_4, A_4 and $\{id, (1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3)\}$.

(iv) $\mathrm{SL}_n(\mathbb{R})$ is a normal subgroup of $\mathrm{GL}_n(\mathbb{R})$ as it is the kernel of the determinant homomorphism.

Recall that the cosets given by a subgroup H are the subsets $Hg = \{hg : h \in H\}$. These are really the *right* cosets. There are also *left cosets* $gH = \{gh : h \in H\}$.

3.38 Theorem. *A subgroup H of G is normal if and only if $Hg = gH$ for all $g \in G$, so that the right cosets are the same as the left cosets.*

Proof. If H is normal and $g \in G$ we need to show $Hg \subseteq gH$ and $gH \subseteq Hg$. If $h \in H$ then $g^{-1}hg \in H$ and $hg = g(g^{-1}hg) \in gH$, giving the first inclusion. Also $ghg^{-1} = (g^{-1})^{-1}h(g^{-1}) \in H$, and $gh = (ghg^{-1})g \in Hg$ giving the second inclusion.

Conversely, if $Hg = gH$ and $h \in H$, then $hg = gh'$ for some $h' \in H$, so $g^{-1}hg = h' \in H$, so H is normal. □

3.39 Proposition. *Any subgroup of index 2 in a group is normal.*

Proof. Since there are only two right cosets, they must be H and $G \setminus H$. Similarly the left cosets must be H and $G \setminus H$. Thus the right cosets of H are the same as the left cosets. □

3.40 Example. Consider the group $G = D_3$ and the subgroup $H = \{I, R, S\}$. The cosets are $HI = \{I, R, S\}$ and $HA = \{A, B, C\}$. The multiplication table for G has block form.

	I	R	S	A	B	C
I	I	R	S	A	B	C
R	R	S	I	C	A	B
S	S	I	R	B	C	A
A	A	B	C	I	R	S
B	B	C	A	S	I	R
C	C	A	B	R	S	I

We spot that this defines a natural multiplication of cosets

	HI	HA
HI	HI	HA
HA	HA	HI

This turns the set of cosets $\{HI, HA\}$ into a group. We denote it by G/H. This is only possible since H is a normal subgroup of D_3.

3.41 Lemma. *Let H be a subgroup of a group G. The following are equivalent*
(i) H is a normal subgroup of G.
(ii) For all $g, g' \in G$ we have: if $x \in Hg$ and $y \in Hg'$ then $xy \in H(gg')$.

Proof. Suppose (i) holds. Let $x = hg$ and $y = h'g'$. Then $xy = h(gh'g^{-1})(gg') \in H(gg')$ since $gh'g^{-1} \in H$.
Conversely if (ii) holds and $h \in H$, then $g^{-1} \in Hg^{-1}$ and $hg \in Hg$ so $g^{-1}hg \in H(g^{-1}g) = H1 = H$. □

3.42 Definition. If H is a normal subgroup of G, then we denote by G/H the set of cosets of H in G, and we equip it with the multiplication defined by $(Hg)(Hg') = H(gg')$. The lemma shows that this is well-defined. It turns G/H into a group, called the *quotient group* of G by H. The map $\theta : G \to G/H$, $\theta(g) = Hg$ is a homomorphism.

3.43 Example. The group $\mathbb{Z}_4$ is really the quotient group $\mathbb{Z}/4\mathbb{Z}$. Similarly for $\mathbb{Z}_n$.

The subgroup $4\mathbb{Z}$ of $\mathbb{Z}$ is normal since $\mathbb{Z}$ is abelian. The cosets are $4\mathbb{Z} = 4\mathbb{Z}+0$, $4\mathbb{Z}+1$, $4\mathbb{Z}+2$ and $4\mathbb{Z}+3$. We denote them by 0, 1, 2 and 3. Then $\mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\}$ with the usual operation of addition modulo 4.

3.44 Theorem (First isomorphism theorem)**.** *If $\theta : G \to G'$ is a homomorphism, then there is an isomorphism $\overline{\theta} : G/\ker\theta \to \operatorname{im}\theta$ defined by $\overline{\theta}(Hg) = \theta(g)$, where $H = \ker\theta$.*

Proof. The map $\overline{\theta}$ is well-defined and injective since $Hx = Hy \Leftrightarrow xy^{-1} \in H = \ker\theta \Leftrightarrow \theta(xy^{-1}) = 1 \Leftrightarrow \theta(x)\theta(y)^{-1} = 1 \Leftrightarrow \theta(x) = \theta(y)$. It is clearly surjective, and it is a homomorphism by the definition of the product in G/H. □

3.45 Examples. (i) The exponential map gives an isomorphism $\mathbb{C}/2\pi i\mathbb{Z} \cong \mathbb{C}^*$.

(ii) The map $\mathbb{R} \to \mathbb{C}^*$, $x \mapsto \exp(2\pi i x)$, is a homomorphism with kernel $\mathbb{Z}$ and image the subgroup $T = \{z \in \mathbb{C} : |z| = 1\}$ of $\mathbb{C}^*$. Thus $\mathbb{R}/\mathbb{Z} \cong T$.

(iii) The sign map $\epsilon : S_n \to \mathbb{R}^*$ has kernel A_n and image the subgroup $K = \{1, -1\}$ of $\mathbb{R}^*$. Thus we get an isomorphism $S_n/A_n \cong K$.

3.46 Definition. A group G is *simple* if it has no non-trivial proper normal subgroups. That is, if the only normal subgroups are $\{1\}$ and G.

There is very long and difficult classification of all finite simple groups.

4 Vector spaces

You have already studied vectors

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n}$$

To save space we also write them as $\mathbf{v} = (v_1, \ldots, v_n)^T$ (T for 'transpose'). We write vectors using bold or underlined letters ($\mathbf{v}$ or $\underline{v}$). Now $\mathbb{R}^n$ becomes an additive group under addition

$$(v_1, \ldots, v_n)^T + (w_1, \ldots, w_n)^T = (v_1 + w_1, \ldots, v_n + w_n)^T.$$

The identity element is the zero vector is $\mathbf{0} = (0, 0, \ldots, 0)$. There is also *scalar multiplication* $a\mathbf{v}$ is defined by

$$a(v_1, \ldots, v_n)^T = (av_1, \ldots, av_n)^T.$$

for a vector $\mathbf{v}$ and a scalar $a \in \mathbb{R}$. We shall make two generalizations:
(1) We will allow $\mathbb{R}$ to be replaced by any field F, for example $\mathbb{C}$, $\mathbb{Q}$, $\mathbb{Z}_p$ (p prime).
(2) We want to allow changing coordinate systems, or having no coordinates at all.

4.1 Definition. A *field* consists of a set F with binary operations $+$ and $\cdot$ satisfying
(i) The operation $+$ turns F into an additive group. The identity element is denoted by 0.
(ii) The product $a \cdot b$ is defined and in F for all $a, b \in F$, it is associative and commutative, and it turns $F^* = \{x \in F : x \neq 0\}$ into an abelian group.
(iii) The product $\cdot$ is distributive over $+$, that is, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

4.2 Definition. Let F be a field. A *vector space over* F, or *an* F*-vector space* consists of a set V, whose elements are called *vectors*, together with operations of addition of vectors, $+$, and scalar multiplication satisfying the following axioms.

(addition) The set V of vectors is an additive group under $+$.

(closure) Scalar multiplication $a\mathbf{v}$ is defined and in V for all scalars $a \in F$ and $\mathbf{v} \in V$.

(compatibility of multiplication) $(ab)\mathbf{v} = a(b\mathbf{v})$ for all $a, b \in F$ and $\mathbf{v} \in V$.

(identity) $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$.

(distributivity)
$a(\mathbf{v} + \mathbf{w}) = (a\mathbf{v}) + (a\mathbf{w})$ for all $a \in F$ and $\mathbf{v}, \mathbf{w} \in V$.
$(a + b)\mathbf{v} = (a\mathbf{v}) + (b\mathbf{v})$ for all $a + b \in F$ and $\mathbf{v} \in V$.

We denote by $\mathbf{0}$ the identity element for V under $+$. The zero vector. We can define subtraction for vectors by defining $\mathbf{u} - \mathbf{v}$ to be equal to $\mathbf{u} + (-\mathbf{v})$.

4.3 Examples. (a) The set F^n of column vectors with entries in F, with the usual operations of addition and scalar multiplication.

(b) F is a vector space over itself. Think of it as F^1.

(c) The set F^∞ of vectors with infinitely many components

$$\mathbf{v} = (v_0, v_1, v_2, \dots)^T, \quad v_i \in F.$$

with addition and scalar multiplication defined 'componentwise', as for F^n.

(d) The set $M_{m\times n}(F)$ of $m \times n$ matrices with entries in F is a vector space. The operations are addition of matrices and scalar multiplication.

(e) The field $\mathbb{C}$ of complex numbers can be considered as a vector space over $\mathbb{R}$ with the usual addition of complex numbers, and scalar multiplication is given by the usual product az, but limited to the case where a is a real number and z is a complex number.

(f) The set $\mathbb{R}[x]$ of polynomials $p(x) = c_n x^n + \cdots + c_1 x + c_0$ in an indeterminate x with real coefficients is a vector space over $\mathbb{R}$. The operations are addition of polynomials $p(x) + q(x)$ and multiplication by a constant $ap(x)$.

4.4 Proposition. *Suppose that V is a vector space over F.*
(i) If $a \in F$ is any scalar and $\mathbf{0} \in V$ is the zero vector, then $a\mathbf{0} = \mathbf{0}$.
(ii) If 0 is the zero element of the field F and $\mathbf{v} \in V$ is any vector, then $0\mathbf{v} = \mathbf{0}$.
(iii) If $a \in F$ and $\mathbf{v} \in V$ and $a\mathbf{v} = \mathbf{0}$, then either $a = 0$ or $\mathbf{v} = 0$.
(iv) If $\mathbf{v} \in V$ then $(-1)\mathbf{v} = -\mathbf{v}$, and in general $(-a)\mathbf{v} = -(a\mathbf{v})$, for any $a \in F$.

Proof. These are straightforward consequences from the axioms. For example for (i), observe that $\mathbf{0} + \mathbf{0} = \mathbf{0}$ since $\mathbf{0}$ is the additive identity. Thus $a(\mathbf{0} + \mathbf{0}) = a\mathbf{0}$, so $a\mathbf{0} + a\mathbf{0} = a\mathbf{0}$ by distributivity. Subtracting $a\mathbf{0}$ from both sides gives $a\mathbf{0} = \mathbf{0}$. □

The analogue of a subgroup of a group is the notion of a subspace.

4.5 Definition. Let V be a vector space over a field F. By a *subspace* of V we mean a subset U of V such that U becomes a vector space with the same operations of addition of vectors and scalar multiplication in V.

In particular U is a subgroup of V considered just as an additive group.

4.6 Theorem (Subspace criterion)**.** *Let V be a vector space over a field F. A subset U of V is a subspace if and only if it satisfies the following properties*
(i) $\mathbf{0} \in U$.
(ii) For all $\mathbf{u}, \mathbf{u}' \in U$ *we have* $\mathbf{u} + \mathbf{u}' \in U$*, and*
(iii) For all scalars $a \in F$ *and elements* $\mathbf{u} \in U$ *we have* $a\mathbf{u} \in U$.

Note that (ii) and (iii) can be replaced by the one condition:
(ii+iii) For all $a, b \in F$ and all $\mathbf{u}, \mathbf{u}' \in U$ we have $a\mathbf{u} + b\mathbf{u}' \in U$.

Proof. Similar to subgroup criterion. □

4.7 Examples. (1) For any vector space V, the subset $\{\mathbf{0}\}$ is a subspace of V called the *trivial* subspace. (Note that $\{\mathbf{0}\}$ is a set with one element, the zero vector for V, so it is a subset of V.)

(2) The subset consisting of the whole of V is a subspace of V. A *proper* subspace of V is one which is not equal to V.

(3) A subset given by a homogeneous linear equation, or a system of simultaneous homogeneous linear equations, is normally a subspace. For example $U = \{(x, y, z) \in \mathbb{R}^3 : 2x + y - 3z = 0\}$ is a subspace of $\mathbb{R}^3$.

Proof. (i) The equation $2x + y - 3z = 0$ holds for $(x, y, z)^T = (0, 0, 0)^T$, so $\mathbf{0} = (0, 0, 0)^T \in U$.

(ii) Suppose $\mathbf{u}, \mathbf{u}' \in U$. Say $\mathbf{u} = (x, y, z)^T$ and $\mathbf{u}' = (x', y', z')^T$. Then $2x + y - 3z = 0$ and $2x'+y'-3z' = 0$. Now $\mathbf{u}+\mathbf{u}' = (x'', y'', z'')^T$ where $x'' = x+x'$, $y'' = y+y'$, $z'' = z+z'$ and $2x''-y''+3z'' = 2(x+x')-(y+y')+3(z+z') = (2x-y+3z)+(2x'-y'+3z') = 0+0 = 0$ so $\mathbf{u} + \mathbf{u}' = (x'', y'', z'')^T \in U$.

(iii) Suppose $\mathbf{u} \in U$ and $a \in \mathbb{R}$. Say $\mathbf{u} = (x, y, z)^T$. Then $2x + y - 3z = 0$. Now $a\mathbf{u} = (ax, ay, az)^T$, and this is in U since $2(ax) - ay + 3(az) = a(2x - y + 3z) = a0 = 0$.

(4) Other equations typically don't give subspaces. Examples:

$W = \{(x, y, z)^T \in \mathbb{R}^3 : x + y + 2z = 3\}$ is not a subspace of $\mathbb{R}^3$, since $\mathbf{0} \notin W$.

$U = \{(x, y, z)^T : y = x^2\}$ is not a subspace of $\mathbb{R}^3$ since $(1, 1, 0)^T, (2, 4, 0)^T \in U$ but $(1, 1, 0)^T + (2, 4, 0)^T = (3, 5, 0)^T \notin U$ (or $2(1, 1, 0)^T = (2, 2, 0)^T \notin U$).

(5) The subspaces of $\mathbb{R}^2$ are the trivial subspace, the whole space, and lines through the origin. For example the line of slope 2 through the origin is $\{(x, y)^T : y - 2x = 0\}$. The subspaces of $\mathbb{R}^3$ are the trivial subspace, the whole space, lines through the origin and planes containing the origin.

(6) We denote by P_n the subset of $\mathbb{R}[x]$ of all polynomials of degree $\leq n$. Thus

$$P_n = \{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 : a_n, \ldots, a_1, a_0 \in \mathbb{R}\}.$$

This is a subspace of $\mathbb{R}[x]$. Thus we have subspaces $P_0 \subseteq P_1 \subseteq P_2 \subseteq \cdots \subseteq \mathbb{R}[x]$.

The analogue of a homomorphism of groups is the notion of a linear map.

4.8 Definition. Let V, W be vector spaces over a field F. A mapping $\theta : V \to W$ is called a *linear mapping* (or *linear transformation*, *linear operator*, or *homomorphism of vector spaces*) if
(i) $\theta(\mathbf{v} + \mathbf{v}') = \theta(\mathbf{v}) + \theta(\mathbf{v}')$ for all $\mathbf{v}, \mathbf{v}' \in V$, and
(ii) $\theta(a\mathbf{v}) = a\theta(\mathbf{v})$ for all $a \in F$ and $\mathbf{v} \in V$.

(It follows that $\theta(a\mathbf{v} + b\mathbf{v}') = a\theta(\mathbf{v}) + b\theta(\mathbf{v}')$ for all $a, b \in F$ and $\mathbf{v}, \mathbf{v}' \in V$. In fact this can be used as a characterization of linear mappings.)

An *isomorphism of vector spaces* is a linear map which is a bijection. If so, we write $V \cong W$.

Example. The mapping $\theta : \mathbb{R}^3 \to \mathbb{R}^2$ given by $\theta(x, y, z)^T = (2x + y, y - z)^T$ is a linear mapping.

Proof (i) Let $\mathbf{v} = (x, y, z)^T$ and $\mathbf{v}' = (x', y', z')^T$. Then $\mathbf{v} + \mathbf{v}' = (x + x', y + y', z + z')^T$, so $\theta(\mathbf{v}+\mathbf{v}') = (2(x+x')+(y+y'), (y+y')-(z+z'))^T$. On the other hand $\theta(\mathbf{v})+\theta(\mathbf{v}') = (2x + y, y - z)^T + (2x' + y', y' - z')^T$. These are equal.

(ii) $\theta(a\mathbf{v}) = \theta(ax, ay, az)^T = (2ax + ay, ay - az)^T = a(2x + y, y - z)^T = a\theta(\mathbf{v})$.

4.9 Proposition. *Given an $m \times n$ matrix A with entries in F, one gets a linear map $\theta_A : F^n \to F^m$ given by $\theta_A(\mathbf{v}) = A\mathbf{v}$. Conversely any linear map $\theta : F^n \to F^m$ is of the form θ_A, where A is the matrix whose columns are $\theta(\mathbf{e}_1), \theta(\mathbf{e}_2), \ldots, \theta(\mathbf{e}_n)$.*

Example. The linear map $\theta : \mathbb{R}^3 \to \mathbb{R}^2$ above has $\theta(1, 0, 0)^T = (2, 0)^T$, $\theta(0, 1, 0)^T = (1, 1)^T$, $\theta(0, 0, 1)^T = (0, -1)^T$, so it is given by the matrix $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$. We check

$$\theta_A(x, y, z) = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y \\ y - z \end{pmatrix} = \theta(x, y, z).$$

Recall that a homomorphism of groups has a kernel and image. Thus a linear map $\theta : V \to W$ has a kernel $\ker \theta = \{\mathbf{v} \in V : \theta(\mathbf{v}) = \mathbf{0}\}$ and image $\text{im } \theta = \{\theta(\mathbf{v}) : \mathbf{v} \in V\}$.

4.10 Proposition. *If $\theta : V \to W$ is a linear transformation, then* $\ker \theta$ *is a subspace of V and* $\text{im } \theta$ *is a subspace of W.*

4.11 Examples. (1) Differentiation $p(x) \mapsto \frac{dp}{dx}$ defines a linear mapping $\mathbb{R}[x] \to \mathbb{R}[x]$ or $P_n \to P_n$,

$$\frac{d}{dx}(a_n x^n + \cdots + a_1 x + a_0) = n a_n x^{n-1} + \cdots + a_1.$$

More generally a linear differential operator such as $D(p(x)) = \frac{d^2p}{dx^2} + 2x\frac{dp}{dx} + (x^2+1)p(x)$ gives a linear map $D : \mathbb{R}[x] \to \mathbb{R}[x]$. The kernel is the set of polynomial solutions of the differential equation.

$$\ker D = \{p(x) \in \mathbb{R}[x] : \frac{d^2p}{dx^2} + 2x\frac{dp}{dx} + (x^2 + 1)p(x) = 0\}$$

One could look for more general solutions by using more general vector spaces of functions.
(2) Evaluation at $\alpha \in \mathbb{R}$ defines a linear map $P_n \to \mathbb{R}$, $p(x) \mapsto p(\alpha)$. The kernel is $\{p(x) \in P_n : p(\alpha) = 0\}$ so this is a subspace of P_n.

(3) The mapping $M_{n\times n}(F) \to M_{n\times n}(F)$ sending a matrix A to its transpose A^T is a linear mapping. So is the mapping sending A to $A - A^T$. The kernel is the set of symmetric matrices

$$\{A \in M_{n\times n}(F) : A = A^T\}.$$

4.12 Lemma. *Given a finite set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in a vector space V over F, the mapping $\phi_S : F^n \to V$ given by*

$$\phi_S(a_1, \ldots, a_n)^T = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

is a linear map.

(Note that we really need to think of S as an *ordered set*, since ϕ_S depends on the ordering of the $\mathbf{v}_j$. We shall ignore this subtlety.)

Proof. Let $\mathbf{a} = (a_1, \ldots, a_n)^T, \mathbf{a}' = (a_1', \ldots, a_n')^T \in F^n$. Then $\phi_S(\mathbf{a} + \mathbf{a}') = \phi_S(a_1 + a_1', \ldots, a_n + a_n') = (a_1 + a_1')\mathbf{v}_1 + \cdots + (a_n + a_n')\mathbf{v}_n = (a_1\mathbf{v}_1 + \ldots a_n\mathbf{v}_n) + (a_1'\mathbf{v}_1 + \cdots + a_n'\mathbf{v}_n) = \phi_S(\mathbf{a}) + \phi_S(\mathbf{a}')$. Also $\phi_S(\lambda\mathbf{a}) = \phi_S(\lambda a_1, \ldots, \lambda a_n)^T = (\lambda a_1)\mathbf{v}_1 + \cdots + (\lambda a_n)\mathbf{v}_n = \lambda(a_1\mathbf{v}_1 + \ldots a_n\mathbf{v}_n) = \lambda\phi_S(\mathbf{a})$. □

4.13 Definition. The *span* of a finite set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in a vector space V is the set of all linear combinations of them,

$$\operatorname{span} S = \{a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n : a_1, \ldots, a_n \in F\}.$$

By convention $\operatorname{span} \varnothing = \{\mathbf{0}\}$. Clearly $\operatorname{span} S = \operatorname{im} \phi_S$, so it is a subspace of V. It is the unique smallest subspace of V which contains S, since any subspace which contains S must contain $\operatorname{span} S$,

4.14 Examples. (1) If $\mathbf{v}$ is a nonzero vector, then $\operatorname{span}\{\mathbf{v}\} = \{a\mathbf{v} : a \in \mathbb{R}\}$. Geometrically, if $\mathbf{v} \in \mathbb{R}^n$, then $\operatorname{span}\{\mathbf{v}\}$ is the line through the origin containing $\mathbf{v}$.

(2) A plane containing the origin in $\mathbb{R}^3$ is the span of two vectors. For example consider $U = \{(x, y, z)^T \in \mathbb{R}^3 : 2x - y + 3z = 0\}$. Using the equation to determine y, we get

$$\begin{aligned} U &= \{(x, 2x + 3z, z)^T : x, z \in \mathbb{R}\} \\ &= \{x(1, 2, 0)^T + z(0, 3, 1)^T : x, z \in \mathbb{R}\} \\ &= \operatorname{span}\{(1, 2, 0)^T, (0, 3, 1)^T\}. \end{aligned}$$

(3) Find a spanning set for the subspace $U = \{(x, y, z, t)^T \in \mathbb{R}^4 : x + y + z + t = 0, x + 2y + 3z + 4t = 0\}$ of $\mathbb{R}^4$. We write the equations in matrix form, row reduce, and can read off the solutions

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{matrix} \\ R_2' = R_2 - R_1 \\ \\ \sim \end{matrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix}$$

Thus $y = -2z - 3t$ and $x = -y - z - t = z + 2t$. Thus

$$\begin{aligned} U &= \{(z+2t, -2z-3t, z, t)^T : z, t \in \mathbb{R}\} \\ &= \{z(1,-2,1,0)^T + t(2,-3,0,1)^T : z, t \in \mathbb{R}\} \\ &= \operatorname{span}\{(1,-2,1,0)^T, (2,-3,0,1)^T\}. \end{aligned}$$

4.15 Definition. Let V be a vector space and let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a finite subset of V. We say that S is *linearly independent* if there is no linear relation between the elements of S of the form

$$a_1\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n = \mathbf{0}$$

with $a_1, \ldots, a_n \in F$, other than the trivial one with $a_1 = \ldots = a_n = 0$. Otherwise S is said to be *linearly dependent*.

Notes. (i) By convention the empty set is linearly independent.
(ii) Any subset of a linearly independent set is linearly independent.
(iii) If some $\mathbf{v}_i = \mathbf{0}$, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent.
(iv) $\{\mathbf{v}\}$ is linearly independent if and only if $\mathbf{v} \neq \mathbf{0}$.
(v) $\{\mathbf{v}, \mathbf{w}\}$ is linearly independent if and only if neither vector is a multiple of the other.
(vi) A linear relation corresponds to an element $(a_1, \ldots, a_n) \in \ker \phi_S$. Thus S is linearly independent if and only if $\ker \phi_S = \{\mathbf{0}\}$.

4.16 Proposition. *$\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent if and only if some $\mathbf{v}_i$ is a linear combination of its predecessors, that is, $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}$.*

Proof. Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent, so there is a relation $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$ with some coefficient nonzero. Let i be maximal with $a_i \neq 0$. Then $a_1\mathbf{v}_1 + \cdots + a_i\mathbf{v}_i = \mathbf{0}$, so

$$\mathbf{v}_i = (-\frac{a_1}{a_i})\mathbf{v}_1 + \cdots + (-\frac{a_{i-1}}{a_i})\mathbf{v}_{i-1} \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}.$$

Conversely, if $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}$, then $\mathbf{v}_i = b_1\mathbf{v}_1 + \cdots + b_{i-1}\mathbf{v}_{i-1}$ for some scalars $b_1, \ldots, b_{i-1}$, giving a relation $b_1\mathbf{v}_1 + \cdots + b_{i-1}\mathbf{v}_{i-1} - \mathbf{v}_i = \mathbf{0}$. □

4.17 Theorem. *Given vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ in F^n, write them as the rows of a matrix A, and row reduce to echelon form giving a matrix B. Then*
(i) $\operatorname{span} S$ *= span of the rows of B. This is equal to $F^n \Leftrightarrow B$ has non-zero leading elements in every column.*
(ii) the rows of A are linearly independent $\Leftrightarrow$ the rows of B are linearly independent $\Leftrightarrow$ B has no rows which are entirely zero.

4.18 Example. Consider $S = \{(1,1,2)^T, (1,3,0)^T, (3,7,2)^T\}$ in $\mathbb{R}^3$. We form the matrix with these rows, and row reduce

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 0 \\ 3 & 7 & 2 \end{pmatrix} \begin{matrix} R_2' = R_2 - R_1 \\ R_3' = R_3 - 3R_1 \\ \sim \end{matrix} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & -2 \\ 0 & 4 & -4 \end{pmatrix} \begin{matrix} R_2' = \frac{1}{2}R_2 \\ \sim \end{matrix} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 4 & -4 \end{pmatrix} \begin{matrix} R_3' = R_3 - 4R_2 \\ \sim \end{matrix} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Then $\operatorname{span} S = \operatorname{span}\{(1,1,2)^T,(0,1,-1)^T,(0,0,0)^T\} = \operatorname{span}\{(1,1,2)^T,(0,1,-1)^T\}$. The final matrix has no non-zero leading element in column 3, so S does not span $\mathbb{R}^3$. The final matrix has a zero row, so S is not linearly independent.

4.19 Theorem. *In any vector space, if I is a linearly independent set and S is a spanning set, then $|I| \leq |S|$.*

Proof. Write each element of the linearly independent set $I = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ as as a linear combination of the vectors in the spanning set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, say

$$\begin{array}{lcccccc} \mathbf{w}_1 & = & a_{11}\mathbf{v}_1 & + & \ldots & + & a_{1n}\mathbf{v}_n \\ \mathbf{w}_2 & = & a_{21}\mathbf{v}_1 & + & \ldots & + & a_{2n}\mathbf{v}_n \\ \ldots & & & & & & \\ \mathbf{w}_m & = & a_{m1}\mathbf{v}_1 & + & \ldots & + & a_{mn}\mathbf{v}_n \end{array}$$

The coefficients give an $m \times n$ matrix A. The rows of A are linearly independent because any relation between them would give a relation between the $\mathbf{w}_j$. Thus, after row reducing, the matrix has no zero rows. But this is only possible if the number of rows is $\leq$ the number of columns, that is, $m \leq n$. □

4.20 Definition. Let V be a vector space. We say that a finite set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a *basis* of V if it is linearly independent and it spans V (i.e. $\operatorname{span} S = V$).

4.21 Examples. (a) $\{(1,0,0)^T,(0,1,0)^T,(0,0,1)^T\}$ is a basis of $\mathbb{R}^3$. But there are many more bases, for example $\{(1,2,4)^T,(0,1,2)^T,(0,0,3)^T\}$.

(b) The vector space $\{(x,y,z)^T \in \mathbb{R}^3 : x+y+z=0\}$ has basis $\{(1,0,-1)^T,(0,1,-1)^T\}$. Again, there are many other bases, for example $\{(2,-1,-1)^T,(3,-1,-2)^T\}$.

4.22 Theorem. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a finite subset of V. Then the following are equivalent*
(i) S is a basis of V
(ii) $\phi_S : F^n \to V$ is an isomorphism of vector spaces
(iii) every $\mathbf{v} \in V$ can be written in a unique way as a linear combination $\mathbf{v} = a_1\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n$.

Proof. (i)⇔(ii). Since $\operatorname{span} S = \operatorname{im} \phi_S$, S spans V if and only if ϕ_S is surjective. Also S is linearly independent if and only if $\ker\phi_S = \{\mathbf{0}\}$, which is if and only if ϕ_S is injective.

(ii)⇔(iii). Clear. □

4.23 Theorem. *Any two bases of a vector space have the same number of elements.*

Proof. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ and $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ be bases of V. Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is independent and $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ spans, so $k \leq n$ by Theorem 4.19. Also $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is independent and $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ spans, so $n \leq k$. Thus $n = k$. □

4.24 Definition. If a vector space V has a basis with n elements, then we say that V has *dimension n*. We call V *finite-dimensional* in this case, and write $\dim V = n$. If V does not have a (finite) basis, then it is said to be *infinite-dimensional*.

Note that the empty set is a basis for the vector space $V = \{\mathbf{0}\}$, so it has dimension 0. It is the only vector space of dimension 0.

4.25 Examples. The *standard basis* of F^n is $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ where $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ is the coordinate vector with a 1 in the ith place, so $\dim F^n = n$.

(ii) The vector space P_n of real polynomials of degree $\leq n$ has basis $\{1, x, x^2, \ldots, x^n\}$, so $\dim P_n = n + 1$.

(iii) The vector space $\mathbb{R}[x]$ of all real polynomials is infinite dimensional. If it had dimension d, then any linearly independent set would have at most d elements. But for any n, the set $\{1, x, x^2, \ldots, x^n\}$ is linearly independent.

(iv) Every $z \in \mathbb{C}$ can be written uniquely in the form $a1 + bi$ with $a, b \in \mathbb{R}$. Thus $\mathbb{C}$, considered as a vector space over $\mathbb{R}$, has basis $\{1, i\}$, so it has dimension 2.

4.26 Theorem. *(i) In a vector space, any spanning set contains a basis. Thus any spanning set in a vector space of dimension n has $\geq n$ elements, and if it has exactly n, then it is a basis.*
(ii) In a finite-dimensional vector space, any linearly independent set can be extended to a basis. Thus any linearly independent set in a vector space of dimension n has $\leq n$ elements, and if it has exactly n, then it is a basis.

Proof. (i) Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a spanning set. If S is linearly independent, then it is already a basis. Thus assume that S is linearly dependent. Then some $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}$. It follows that $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$ is spanning. Now either S' is linearly independent, so a basis of V, or we can continue in the same way, eliminating further elements. Eventually we obtain a basis of V.

(ii) Let I be the linearly independent set and let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a basis. If $\mathbf{u}_1 \notin \operatorname{span} I$, replace I by $I \cup \{\mathbf{u}_1\}$. It is still linearly independent. Now if $\mathbf{u}_2 \notin \operatorname{span} I$, replace I by $I \cup \{\mathbf{u}_2\}$, and so on. At the end, we have enlarged I to a linearly independent set whose span contains all elements in the basis, so it is a basis. □

4.27 Examples. (a) The set $\{(1,1,1)^T, (1,2,3)^T, (1,4,9)^T, (1,8,27)^T\}$ spans $\mathbb{R}^3$, but can't be a basis as there are too many elements. Use trial and error to discard elements, and row reduction to check it still spans. In fact $\{(1,1,1)^T, (1,2,3)^T, (1,4,9)^T\}$ is a basis.

(b) The set $\{(1,2,3,0)^T, (0,2,3,0)^T\}$ is linearly independent, but can't be a basis of $\mathbb{R}^4$. We can get a basis by adjoining vectors in the standard basis. Use trial and error, and row reduction to check. In fact $\{(1,2,3,0)^T, (0,2,3,0)^T, (0,1,0,0)^T, (0,0,0,1)^T\}$ is a basis.

4.28 Theorem. *If W is a subspace of a finite-dimensional vector space V, then W is finite-dimensional and $\dim W \leq \dim V$. Moreover, if $\dim W = \dim V$ then $W = V$.*

Proof. Any linearly independent subset S of W is linearly independent in V so has at most $\dim V$ elements. Thus we can choose one with as many elements as possible. Every element $\mathbf{w} \in W$ is in $\operatorname{span} S$, for otherwise $S \cup \{\mathbf{w}\}$ is linearly independent by Proposition 4.16. Thus S is a basis for W. □

4.29 Theorem. *Two finite-dimensional vector spaces over F are isomorphic if and only if they have the same dimension.*

Proof. If they both have dimension n, then they are both isomorphic to F^n, so they are isomorphic to each other.

Conversely, if $\theta : V \to W$ is an isomorphism, and V is finite-dimensional, with basis $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then it is easy to see that $\{\theta(\mathbf{v}_1), \ldots, \theta(\mathbf{v}_n)\}$ is a basis for W, so W also has dimension n. □

4.30 Definition. Let V be a vector space over F. If U is a subspace of V, then the *quotient vector space* V/U is the quotient group under addition, with scalar multiplication defined by $a(U + \mathbf{v}) = U + a\mathbf{v}$. It is easy to see that the natural map $V \to V/U$, $\mathbf{v} \mapsto U + \mathbf{v}$ is a linear map.

4.31 Proposition. *If V is a finite-dimensional vector space and U is a subspace of V, then* $\dim V/U = \dim V - \dim U$.

Proof. Take a basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of U. It can be extended to give a basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ of V. We check that $\{U + \mathbf{v}_1, \ldots, U + \mathbf{v}_\ell\}$ is a basis of V/U.

Span. Any element of V/U is of the form $U + \mathbf{v}$. We can write $\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k + b_1\mathbf{v}_1 + \cdots + b_\ell\mathbf{v}_\ell$ for some a_i, b_i. Then

$$U+\mathbf{v} = U+a_1\mathbf{u}_1+\cdots+a_k\mathbf{u}_k+b_1\mathbf{v}_1+\cdots+b_\ell\mathbf{v}_\ell = U+b_1\mathbf{v}_1+\cdots+b_\ell\mathbf{v}_\ell = b_1(U+\mathbf{v}_1)+\cdots+b_\ell(U+\mathbf{v}_\ell).$$

Linear independence. Say $b_1(U + \mathbf{v}_1) + \cdots + b_\ell(U + \mathbf{v}_\ell) = U + \mathbf{0}$. Then $U + b_1\mathbf{v}_1 + \cdots + b_\ell\mathbf{v}_\ell = \mathbf{0}$. Then $b_1\mathbf{v}_1 + \cdots + b_\ell\mathbf{v}_\ell \in U$. Thus there are a_i with $a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k + b_1\mathbf{v}_1 + \cdots + b_\ell\mathbf{v}_\ell = \mathbf{0}$. But then all $a_i = 0$ and $b_i = 0$. □

4.32 Theorem (First isomorphism theorem for vector spaces)**.** *If $\theta : V \to W$ is a linear map, then it induces an isomorphism of vector spaces $\overline{\theta} : V/\ker\theta \to \operatorname{im}\theta$.*

Proof. Same as the first isomorphism theorem for groups. □

4.33 Definition. If $\theta : V \to W$ is a linear map, then the *rank* of θ is $r(\theta) = \dim \operatorname{im}\theta$ and the *nullity* of θ is $n(\theta) = \dim \ker\theta$.

4.34 Corollary (Rank-nullity formula)**.** *If $\theta : V \to W$ is a linear map with V finite-dimensional, then $r(\theta) + n(\theta) = \dim V$.*

Proof. $r(\theta) = \dim \operatorname{im}\theta = \dim V/\ker\theta = \dim V - \dim\ker\theta = \dim V - n(\theta)$. □

4.35 Example. If $\theta : \mathbb{R}^4 \to \mathbb{R}^2$ is defined by $\theta(x, y, z, w)^T = (x + y, 3x + 3y)^T$, then $\ker\theta$ is all solutions to $x + y = 3x + 3y = 0$, i.e., parametrized by $(a, -a, b, c)^T$ and $n(\theta) = 3$. Likewise, $\operatorname{im}\theta$ is parametrized as $(d, 3d)^T$, and $r(\theta) = 1$. Then $r(\theta) + n(\theta) = 4 = \dim\mathbb{R}^4$.

4.36 Corollary. *If $\theta : V \to W$ is a linear map with $\dim V = \dim W$, then θ is injective if and only if it is surjective.*

Proof. Surjective $\Leftrightarrow r(\theta) = \dim W \Leftrightarrow r(\theta) = \dim V \Leftrightarrow n(\theta) = 0 \Leftrightarrow$ injective. □

5 Diagonalizability and the orthogonal group

We can use a basis to introduce coordinates on a vector space.

5.1 Definition. Suppose that $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of a vector space V over F. In this case the map $\phi_S : F^n \to V$ is an isomorphism. Thus for each $\mathbf{v} \in V$ there is a unique vector $\mathbf{x} = (x_1, \ldots, x_n)^T \in F^n$ such that $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$. We call it the *coordinates of* $\mathbf{v}$ *with respect to* S, and denote it by $[\mathbf{v}]_S$.

Special case: if $V = F^n$ and S is the standard basis of F^n, then $[\mathbf{v}]_S = \mathbf{v}$ for $\mathbf{v} \in F^n$.

5.2 Definition. Let $\theta : V \to W$ be a linear map, let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of V and let $R = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ be a basis of W. The *matrix of* θ *with respect to the basis* S *of* V *and the basis* T *of* W is the matrix $A = (a_{ij})$ whose jth column is the coordinates of $\theta(\mathbf{v}_j)$ with respect to R.

Thus

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

where

$$\begin{aligned} \theta(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m \\ \theta(\mathbf{v}_2) &= a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m \\ &\cdots \\ \theta(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m. \end{aligned}$$

or $\theta(\mathbf{v}_j) = \sum_{i=1}^n a_{ij}\mathbf{w}_i$.

Special case. If $\theta : V \to V$ is a linear map from a vector space to itself, and we use the same basis for both the source and target copies of V, then we speak of the *matrix of* θ *with respect to* S.

5.3 Proposition. *If* $\theta : V \to W$, S *is a basis of* V *and* R *is a basis of* W *then* $[\theta(\mathbf{v})]_R = A[\mathbf{v}]_S$ *for* $\mathbf{v} \in V$.

Proof. If $\mathbf{x} = [\mathbf{v}]_S$, then $\mathbf{v} = \sum_{j=1}^n x_j\mathbf{v}_j$, so

$$\theta(\mathbf{v}) = \theta(\sum_{j=1}^n x_j\mathbf{v}_j) = \sum_{j=1}^n x_j \sum_{i=1}^n a_{ij}\mathbf{w}_i = \sum_{i=1}^n (\sum_{j=1}^n a_{ij}x_j)\mathbf{w}_i = \sum_{i=1}^n (A\mathbf{x})_i\mathbf{w}_i.$$

Thus $[\theta(\mathbf{v})]_R = A\mathbf{x} = A[\mathbf{v}]_S$. □

5.4 Examples. (1) Recall that an $m \times n$ matrix A gives a linear map $\theta_A : F^n \to F^m$, $\theta_A(\mathbf{v}) = A\mathbf{v}$. One recovers A as the matrix of θ_A with respect to the standard bases of F^n and F^m.

(2) Let $\theta : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear map $\theta(x,y)^T = (x+y, 2y, x-y)^T$. We want to find the matrix A of θ with respect to the basis $S = \{(0,1)^T, (1,1)^T\}$ of $\mathbb{R}^2$ and the basis $\{(1,0,0)^T, (1,1,0)^T, (1,1,1)^T\}$ of $\mathbb{R}^3$. We compute

$$\begin{array}{rcccl} \theta(0,1)^T & = & (1,2,-1)^T & = & -1(1,0,0)^T + 3(1,1,0)^T - 1(1,1,1)^T \\ \theta(1,1)^T & = & (2,2,0)^T & = & 0(1,0,0)^T + 2(1,1,0)^T + 0(1,1,1)^T \end{array} \quad \text{so} \quad A = \begin{pmatrix} -1 & 0 \\ 3 & 2 \\ -1 & 0 \end{pmatrix}.$$

(3) The vector space P_3 of polynomials over $\mathbb{R}$ of degree ≤ 3 has basis $\{1, x, x^2, x^3\}$. Consider the linear map $\frac{d}{dx} : P_3 \to P_3$. We have

$$\begin{array}{rcccl} \frac{d}{dx}(1) & = & 0 & = & 0.1 + 0.x + 0.x^2 + 0.x^3 \\ \frac{d}{dx}(x) & = & 1 & = & 1.1 + 0.x + 0.x^2 + 0.x^3 \\ \frac{d}{dx}(x^2) & = & 2x & = & 0.1 + 2.x + 0.x^2 + 0.x^3 \\ \frac{d}{dx}(x^3) & = & 3x^2 & = & 0.1 + 0.x + 3.x^2 + 0.x^3. \end{array}$$

Thus the matrix of $\frac{d}{dx}$ with respect to the basis $\{1, x, x^2, x^3\}$ of P_3, is:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix of θ depends on the basis we use.

5.5 Definition. If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $S' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ are bases of V then the *transition matrix from S to S'* is the matrix $P = (p_{ij})$ whose jth column is the coordinates of $\mathbf{v}'_j$ with respect to S. Thus $\mathbf{v}'_j = \sum_{i=1}^n p_{ij}\mathbf{v}_i$.

We have $[\mathbf{v}]_S = P[\mathbf{v}]_{S'}$ for $\mathbf{v} \in V$ since if $\mathbf{x} = [\mathbf{v}]_{S'}$, then

$$\mathbf{v} = \sum_{j=1}^n x_j \mathbf{v}'_j = \sum_{j=1}^n x_j \sum_{i=1}^n p_{ij}\mathbf{v}_i = \sum_{i=1}^n (\sum_{j=1}^n p_{ij}x_j)\mathbf{v}_i = \sum_{i=1}^n (P\mathbf{x})_i \mathbf{v}_i.$$

Note that P is invertible; its inverse is the transition matrix in the opposite direction.

Special case: for F^n, the transition matrix from the standard basis to a basis S' is the matrix whose columns are given by the vectors in S'.

5.6 Theorem (Change of basis)**.** *Let $\theta : V \to V$ be a linear map from a vector space to itself.*
Let A be the matrix of θ with respect to a basis S of V.
Let A' be the matrix of θ with respect to a basis S' of V.
Then $A' = P^{-1}AP$ where P is the transition matrix from S to S'.

Proof. For $\mathbf{v} \in V$, we have $AP[\mathbf{v}]_{S'} = A[\mathbf{v}]_S = [\theta(\mathbf{v})]_S = P[\theta(\mathbf{v})]_{S'} = PA'[\mathbf{v}]_{S'}$. Since this holds for all $\mathbf{v}$, we must have $AP = PA'$. □

5.7 Definition. Two $n \times n$ matrices A, A' are *similar* if there is an invertible matrix P with $A' = P^{-1}AP$.

It is easy to see that similarity is an equivalence relation for matrices. Also, similar matrices have the same determinant since $\det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P)$ and $\det(P^{-1}) = 1/\det(P)$.

Observe that similarity is like conjugacy in the group $\mathrm{GL}_n(F)$, except that we don't require that A, A' be invertible matrices.
Let A be an $n \times n$ matrix over F. Recall that an *eigenvector* for A with *eigenvalue* $\lambda \in F$ is a non-zero vector $\mathbf{v} \in F^n$ with $A\mathbf{v} = \lambda\mathbf{v}$. Then $\lambda \in F$ is an eigenvalue
$\Leftrightarrow$ there is a non-zero vector $\mathbf{v} \in F^n$ with $(A - \lambda I)\mathbf{v} = \mathbf{0}$
$\Leftrightarrow$ the matrix $A - \lambda I$ is not invertible
$\Leftrightarrow \det(A - \lambda I) = 0$
$\Leftrightarrow \lambda$ is a root of the *characteristic polynomial* $\chi_A(t) = \det(tI - A)$.

For $\lambda \in F$, the corresponding *eigenspace* is

$$Esp(\lambda) = \{\mathbf{v} \in F^n : A\mathbf{v} = \lambda\mathbf{v}\} = \{\mathbf{v} \in F^n : (A - \lambda I)\mathbf{v} = 0\} = \ker \theta_{A-\lambda I}$$

The eigenvectors are thus the non-zero elements of the eigenspace, and λ is an eigenvalue if $Esp(\lambda) \neq \{\mathbf{0}\}$.

5.8 Example. The matrix

$$A = \begin{pmatrix} 1 & -4 & 2 & 0 \\ 1 & -5 & 4 & 0 \\ 1 & -8 & 7 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

has characteristic polynomial

$$\chi_A(t) = \det(tI - A) = \begin{vmatrix} t-1 & 4 & -2 & 0 \\ -1 & t+5 & -4 & 0 \\ -1 & 8 & t-7 & 0 \\ 0 & 0 & 0 & t-3 \end{vmatrix} = (t-3)\begin{vmatrix} t-1 & 4 & -2 \\ -1 & t+5 & -4 \\ -1 & 8 & t-7 \end{vmatrix}.$$

Subtracting row 2 from row 3 this becomes

$$\chi_A(t) = (t-3)\begin{vmatrix} t-1 & 4 & -2 \\ -1 & t+5 & -4 \\ 0 & 3-t & t-3 \end{vmatrix} = (t-3)^2\begin{vmatrix} t-1 & 4 & -2 \\ -1 & t+5 & -4 \\ 0 & -1 & 1 \end{vmatrix} = (t-3)^2(t^2+1),$$

where we have taken out another factor $t - 3$, expanded along row 3, and calculated the resulting 2×2 determinants. Thus the eigenvalues are 3, i and $-i$.

The $\lambda = 3$ eigenspace $Esp(3)$ is the set of $\mathbf{v} = (x, y, z, w) \in \mathbb{C}^4$ satisfying $(A-3I)\mathbf{v} = \mathbf{0}$, so

$$\begin{pmatrix} -2 & -4 & 2 & 0 \\ 1 & -8 & 4 & 0 \\ 1 & -8 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Row reducing this matrix we find that $Esp(3) = \{(x,y,z,w)^T \in \mathbb{C}^4 : x = 0, z = 2y\}$, so it has basis $\{(0,1,2,0)^T, (0,0,0,1)^T\}$.

The $\lambda = i$ eigenspace $Esp(i)$ is the set of $\mathbf{v} = (x,y,z,w)^T \in \mathbb{C}^4$ satisfying $(A - iI)\mathbf{v} = \mathbf{0}$, so

$$\begin{pmatrix} 1-i & -4 & 2 & 0 \\ 1 & -5-i & 4 & 0 \\ 1 & -8 & 7-i & 0 \\ 0 & 0 & 0 & 3-i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Row reducing this matrix we find that $V_i = \{(x,y,z,w)^T \in \mathbb{C}^4 : x = (1+i)y, y = z, w = 0\}$, so $Esp(i)$ has basis $\{(1+i,1,1,0)^T\}$. Similarly the eigenspace $Esp(-i)$ has basis $\{(1-i,1,1,0)^T\}$.

5.9 Definition. Suppose A is an $n \times n$ matrix and $\lambda \in F$.
Geometric multiplicity of λ = dimension of the λ-eigenspace $Esp(\lambda)$ for A.
Algebraic multiplicity of λ = multiplicity of λ as a root of the characteristic poly $\chi_A(t)$.

One can show that the geometric multiplicity of $\lambda \leq$ algebraic multiplicity of λ.

5.10 Theorem. *Let A be an $n \times n$ matrix over F. The following are equivalent:*
(i) A is diagonalizable, meaning that it is similar to a diagonal matrix, so

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

for some invertible matrix $P \in \mathrm{GL}_n(F)$.
(ii) A has n linearly independent eigenvectors (so they form a basis of F^n).
(iii) The characteristic polynomial of A has n roots in F, counted with multiplicity (which always holds if $F = \mathbb{C}$), and for each eigenvalue λ, the geometric multiplicity of λ is equal to the algebraic multiplicity of λ.

Proof. Sketch. (i)⇒(iii) Similar matrices have the same characteristic polynomial, for

$$\chi_{P^{-1}AP}(t) = \det(tI - P^{-1}AP) = \det(P^{-1}(tI - A)P) = \det(tI - A) = \chi_A(t).$$

Thus if (i) holds then $\chi_A(t) = (t-\lambda_1)(t-\lambda_2)\dots(t-\lambda_n)$, so it has n roots in F. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be the columns of P. Since P is invertible, the set $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of F^n. Also, letting D be the diagonal matrix of λs, we have $D = P^{-1}AP$, so $AP = PD$. The ith column in this equation gives $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$, so the $\mathbf{v}_i$ are eigenvectors for the λ_i.

(iii)⇒(ii) Combining bases of each of the eigenspaces, one gets n eigenvectors. One can show that this set is linearly independent.

(ii)⇒(i) The matrix of θ_A with respect to this basis of eigenvectors is diagonal. □

In the example, working over $\mathbb{C}$, the basis of eigenvectors is

$$\{(0,1,2,0)^T, (0,0,0,1)^T, (1+i,1,1,0)^T, (1-i,1,1,0)^T\}.$$

Then $D = P^{-1}AP$ is diagonal, where P is the transition matrix from the standard basis to this one, so with columns given by the vectors in the basis, and D is the diagonal matrix given by the eigenvalues:

$$P = \begin{pmatrix} 0 & 0 & 1+i & 1-i \\ 1 & 0 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad D = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}.$$

Note that we can diagonalize this matrix over $\mathbb{C}$, but not over $\mathbb{R}$.

5.11 Corollary. *If A is an $n \times n$ matrix, and its characteristic polynomial has n distinct roots in F, then A is diagonalizable.*

Proof. Each eigenvalue has algebraic multiplicity 1, which must therefore also be its geometric multiplicity. □

5.12 Theorem. *Let A be an $n \times n$ matrix. If the characteristic polynomial of A has n roots in F, counted with multiplicity (which always holds if $F = \mathbb{C}$), then A is similar to an upper triangular matrix.*

The proof is omitted, but it is based on the following lemma.

5.13 Lemma. *Let A be an $n \times n$ matrix and $\mathbf{v}$ an eigenvector with eigenvalue λ. Extend to a basis $\{\mathbf{v}_1 = \mathbf{v}, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of F^n, and let P be the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Then $P^{-1}AP$ has block form*

$$\begin{pmatrix} \lambda & * \\ 0 & B \end{pmatrix}$$

where B is an $(n-1) \times (n-1)$ matrix and $*$ is a $1 \times (n-1)$ matrix.

We work over $\mathbb{R}$. The vector space $\mathbb{R}^n$ has dot product

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \ldots v_n w_n$$

where $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$. It can also be computed as $\mathbf{v}^T \mathbf{w}$. The *length* of a vector $\mathbf{v} \in \mathbb{R}^n$ is $|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$, and two vectors $\mathbf{v}$ and $\mathbf{w}$ are *orthogonal* if $\mathbf{v} \cdot \mathbf{w} = 0$.

5.14 Definition. A set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is *orthogonal* if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$. It is *orthonormal* if also $|\mathbf{v}_i| = 1$ for all i, so

$$\mathbf{v}_i \cdot \mathbf{v}_i = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j). \end{cases}$$

5.15 Definition. A real $n \times n$ matrix P is said to be *orthogonal* if it is invertible and $P^{-1} = P^T$, that is, $P^T P = I = P P^T$.

(In fact you only need to check that $P^T P = I$. It follows that $\det P \neq 0$, so P is invertible, so $P^{-1} = P^T$.)

The set of orthogonal matrices forms a subgroup $O_n(\mathbb{R})$ of $\mathrm{GL}_n(\mathbb{R})$, the *orthogonal group*. The set of orthogonal matrices of determinant 1 forms a subgroup $SO_n(\mathbb{R})$, the *special orthogonal group*.

5.16 Proposition. *The determinant of an orthogonal matrix is* ± 1.

Proof. $\det P^T = \det P$, so $P^T P = I$ gives $(\det P)^2 = 1$. □

5.17 Proposition. *A matrix P is orthogonal if and only if its columns are an orthonormal set of vectors.*

Proof. If $\mathbf{v}_i$ is the ith column of P, then $\mathbf{v}_i^T$ is the ith row of P^T, and the (i, j) entry of $P^T P$ is $\mathbf{v}_i^T \mathbf{v}_j$. Thus the set of columns $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is orthonormal if and only if $P^T P = I$. □

5.18 Proposition. *Any orthonormal set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent.*

Proof. If $a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k = \mathbf{0}$ is a linear relation, then for all i we have

$$0 = \mathbf{v}_i \cdot \mathbf{0} = \mathbf{v}_i \cdot (a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k) = a_1 \mathbf{v}_i \cdot \mathbf{v}_1 + \cdots + a_k \mathbf{v}_i \cdot \mathbf{v}_k = a_i.$$

□

5.19 Theorem (Gram-Schmidt process)**.** *Given any linearly independent set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ in $\mathbb{R}^n$, we can find an orthonormal set $S' = \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ with the same span. In particular, any subspace of $\mathbb{R}^n$ has an orthonormal basis.*

Proof. We construct

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 \\
&\cdots \\
\mathbf{u}_k &= \mathbf{v}_k - \sum_{i=1}^{k-1} \frac{\mathbf{v}_k \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \mathbf{u}_i
\end{aligned}$$

By construction $\mathbf{u}_j \cdot \mathbf{u}_i = 0$ for all $i < j$. Thus $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is orthogonal. It is easy to see that it is linearly independent and has the same span as S. Now all $\mathbf{u}_i \neq \mathbf{0}$, so we can normalize them, letting $\mathbf{w}_i = \mathbf{u}_i / |\mathbf{u}_i|$, to get an orthonormal set S'. □

5.20 Example. Find an orthonormal basis for the subspace $U = \{(x, y, z, t)^T : x+y+z+t = 0\}$ of $\mathbb{R}^4$. We have $U = \{(x, y, z, -x-y-z)^T : x, y, z \in \mathbb{R}\} = \{x(1, 0, 0, -1)^T + y(0, 1, 0, -1)^T + z(0, 0, 1, -1)^T : x, y, z \in \mathbb{R}\} = \text{span}\{(1, 0, 0, -1)^T, (0, 1, 0, -1)^T, (0, 0, 1, -1)^T\}$. Apply the Gram-Schmidt process to the vectors $\mathbf{v}_1 = (1, 0, 0, -1)^T$, $\mathbf{v}_2 = (0, 1, 0, -1)^T$, $\mathbf{v}_3 = (0, 0, 1, -1)^T$.

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 = (1, 0, 0, -1)^T. \\
\mathbf{u}_2 &= \mathbf{v}_2 - \tfrac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\
&= (0, 1, 0, -1)^T - \tfrac{1}{2}(1, 0, 0, -1)^T \\
&= (-\tfrac{1}{2}, 1, 0, -\tfrac{1}{2})^T. \\
\mathbf{u}_3 &= \mathbf{v}_3 - \tfrac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \tfrac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \\
&= (0, 0, 1, -1)^T - \tfrac{1}{2}(1, 0, 0, -1)^T - \tfrac{1/2}{3/2}(-\tfrac{1}{2}, 1, 0, -\tfrac{1}{2})^T \\
&= (-\tfrac{1}{3}, -\tfrac{1}{3}, 1, -\tfrac{1}{3})^T.
\end{aligned}$$

Then our orthonormal basis for U is $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ where $\mathbf{w}_1 = \frac{1}{\sqrt{2}}(1, 0, 0, -1)^T$, $\mathbf{w}_2 = \frac{1}{\sqrt{6}}(-1, 2, 0, -1)^T$, $\mathbf{w}_3 = \frac{1}{2\sqrt{3}}(-1, -1, 3, -1)^T$.

5.21 Theorem. *A real symmetric matrix has real eigenvalues, and eigenvectors for distinct eigenvalues are orthogonal.*

Proof. Suppose that $\lambda \in \mathbb{C}$ is an eigenvalue with associated eigenvector $\mathbf{v} \in \mathbb{C}^n$. We compute $\overline{\mathbf{v}}^T A\mathbf{v}$ in two ways. We have $A\mathbf{v} = \lambda\mathbf{v}$, so $\overline{\mathbf{v}}^T A\mathbf{v} = \overline{\mathbf{v}}^T \lambda\mathbf{v} = \lambda \sum_{i=1}^n |v_i|^2$. On the other hand, starting with $A\mathbf{v} = \lambda\mathbf{v}$, taking the conjugate, and using A real, gives $A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$. Now taking the transpose and using the fact that A is symmetric, we get $\overline{\mathbf{v}}^T A = \overline{\lambda}\overline{\mathbf{v}}^T$. Thus $\overline{\mathbf{v}}^T A\mathbf{v} = \overline{\lambda}\overline{\mathbf{v}}^T\mathbf{v} = \overline{\lambda} \sum_{i=1}^n |v_i|^2$. Thus $\lambda \sum_{i=1}^n |v_i|^2 = \overline{\lambda} \sum_{i=1}^n |v_i|^2$, so λ is real.

If $A\mathbf{v} = \lambda\mathbf{v}$ and $A\mathbf{w} = \mu\mathbf{w}$, then $\mathbf{v}^T A\mathbf{w} = \mathbf{v}^T \mu\mathbf{w} = \mu\mathbf{v} \cdot \mathbf{w}$. But also $\mathbf{v}^T A$ is the transpose of $A\mathbf{v}$, so it is $\lambda\mathbf{v}^T$, so $\mathbf{v}^T A\mathbf{w} = \lambda\mathbf{v}^T\mathbf{w} = \lambda\mathbf{v} \cdot \mathbf{w}$. Thus $\mu\mathbf{v} \cdot \mathbf{w} = \lambda\mathbf{v} \cdot \mathbf{w}$, so $\mathbf{v} \cdot \mathbf{w} = 0$. □

5.22 Lemma. *If a matrix A is a symmetric, then so is $P^{-1}AP$ for P orthogonal.*

Proof. $(P^{-1}AP)^T = P^T A^T (P^{-1})^T = P^{-1}AP$. □

5.23 Theorem. *Any real symmetric matrix A can be diagonalized by an orthogonal matrix, that is, there is an orthogonal matrix P with $D = P^{-1}AP$ diagonal. Equivalently, $\mathbb{R}^n$ has an orthonormal basis consisting of eigenvectors of A.*

Proof. Take an eigenvalue λ of A. It is real, so has an eigenvector $\mathbf{v} \in \mathbb{R}^n$. We may assume that $|\mathbf{v}| = 1$. We can extend this to a basis $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{R}^n$ where $\mathbf{v}_1 = \mathbf{v}$, and, by the Gram-Schmidt process, we may assume that S is an orthonormal basis of $\mathbb{R}^n$. Let P_0 be the matrix whose columns are the vectors $\mathbf{v}_i$. It is an orthogonal

matrix. If $\theta_A : \mathbb{R}^n \to \mathbb{R}^n$ is the linear map $\mathbf{v} \mapsto A\mathbf{v}$, then $P_0^{-1}AP_0$ is the matrix of θ_A with respect the basis S. Since $\theta_A(\mathbf{v}) = \lambda\mathbf{v}$, it takes upper block shape, so

$$P_0^{-1}AP_0 = \begin{pmatrix} \lambda & * \\ 0 & B \end{pmatrix}$$

Since P_0 is an orthogonal matrix, $P_0^{-1}AP_0$ is symmetric. Thus the $*$ term is zero, so the matrix is block diagonal. Now B is symmetric and smaller so by induction there is an orthogonal matrix Q with $C = Q^{-1}BQ$ diagonal. Then

$$Q' = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$$

is orthogonal, hence so is $P = P_0Q'$, and

$$P^{-1}AP = (Q')^{-1}(P_0^{-1}AP_0)Q' = \begin{pmatrix} 1 & 0 \\ 0 & Q^{-1} \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & C \end{pmatrix},$$

which is diagonal. For the last part, an orthogonal matrix diagonalizes A if and only if its columns are an orthonormal basis of eigenvectors of A. □

5.24 Example. Find an orthogonal matrix P diagonalizing the matrix

$$A = \begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix}.$$

We compute the eigenspaces of A and find bases of them. Then we use the Gram-Schmidt process to find orthonormal bases of the eigenspaces. Then we combine them to get an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors. Then P is the matrix which has these vectors as its columns, and D is the diagonal matrix of the eigenvectors.

The characteristic polynomial is

$$\chi_A(t) = \det(tI - A) = \begin{vmatrix} t-4 & -4 & -2 \\ -4 & t-4 & -2 \\ -2 & -2 & t-1 \end{vmatrix} = \begin{vmatrix} t-4 & -4 & -2 \\ 0 & t & -2t \\ -2 & -2 & t-1 \end{vmatrix}$$

$$= \begin{vmatrix} t-4 & -4 & -10 \\ 0 & t & 0 \\ -2 & -2 & t-5 \end{vmatrix} = t \begin{vmatrix} t-4 & -10 \\ -2 & t-5 \end{vmatrix} = t(t^2 - 9t) = t^2(t-9).$$

The $\lambda = 9$ eigenspace $Esp(9)$ is the set of $\mathbf{v} = (x, y, z)^T \in \mathbb{R}^3$ satisfying $(A - 9I)\mathbf{v} = \mathbf{0}$, so

$$\begin{pmatrix} -5 & 4 & 2 \\ 4 & -5 & 2 \\ 2 & 2 & -8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Row reducing this matrix we find that

$$Esp(9) = \{(x, y, z)^T \in \mathbb{R}^3 : x + y - 4z = 0, y - 2z = 0\} = \{(2z, 2z, z)^T : z \in \mathbb{R}\}.$$

This has basis $\{(2,2,1)^T\}$. It has orthonormal basis $\{\frac{1}{3}(2,2,1)^T\}$.

The $\lambda = 0$ eigenspace $Esp(0)$ is the set of $\mathbf{v} = (x,y,z)^T \in \mathbb{R}^3$ satisfying $A\mathbf{v} = \mathbf{0}$, so $Esp(0) = \{(x,y,z)^T \in \mathbb{R}^3 : 2x+2y+z=0\} = \{(x,y,-2x-2y)^T : x,y \in \mathbb{R}\}$. Thus $Esp(0)$ has basis $\{(1,0,-2)^T, (0,1,-2)^T\}$.

Now apply the Gram-Schmidt process to the vectors $\mathbf{v}_1 = (1,0,-2)^T$, $\mathbf{v}_2 = (0,1,-2)^T$, giving

$$\mathbf{u}_1 = \mathbf{v}_1 = (1,0,-2)^T, \quad \mathbf{u}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\,\mathbf{u}_1 = (0,1,-2)^T - \frac{4}{5}(1,0,-2)^T = \frac{1}{5}(-4,5,-2)^T.$$

Then we normalize this to get an orthonormal basis $\{\frac{1}{\sqrt{5}}(1,0,-2)^T, \frac{1}{\sqrt{45}}(-4,5,-2)^T\}$ for $Esp(0)$.

Combining, we get an orthonormal basis $\{\frac{1}{3}(2,2,1)^T, \frac{1}{\sqrt{5}}(1,0,-2)^T, \frac{1}{\sqrt{45}}(-4,5,-2)^T\}$ for $\mathbb{R}^3$. Then

$$P = \begin{pmatrix} \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{-4}{\sqrt{45}} \\ \frac{2}{3} & 0 & \frac{5}{\sqrt{45}} \\ \frac{1}{3} & \frac{-2}{\sqrt{5}} & \frac{-2}{\sqrt{45}} \end{pmatrix}, \qquad D = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

An *isometry* of $\mathbb{R}^n$ is a mapping $\mathbb{R}^n \to \mathbb{R}^n$ which preserves distances.

5.25 Theorem. *If P is an orthogonal matrix, then the map $\theta_P : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{v} \mapsto P\mathbf{v}$ is an isometry. Conversely, any isometry of $\mathbb{R}^n$ that fixes the origin is of this form.*

Proof. For any $\mathbf{v} \in \mathbb{R}^n$ we have $|P\mathbf{v}| = |\mathbf{v}|$ since

$$|P\mathbf{v}|^2 = (P\mathbf{v}) \cdot (P\mathbf{v}) = (P\mathbf{v})^T (P\mathbf{v}) = \mathbf{v}^T P^T P \mathbf{v} = \mathbf{v}^T \mathbf{v} = \mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|^2.$$

Now the distance between $\theta_P(\mathbf{v})$ and $\theta_P(\mathbf{w})$ is

$$|\theta_P(\mathbf{v}) - \theta_P(\mathbf{w})| = |P\mathbf{v} - P\mathbf{w}| = |P(\mathbf{v}-\mathbf{w})| = |\mathbf{v}-\mathbf{w}|.$$

Conversely suppose that θ is an isometry fixing the origin. Then for any $\mathbf{v}, \mathbf{w}$ we have $|\theta(\mathbf{v}) - \theta(\mathbf{w})| = |\mathbf{v}-\mathbf{w}|$, so

$$(\theta(\mathbf{v}) - \theta(\mathbf{w})) \cdot (\theta(\mathbf{v}) - \theta(\mathbf{w})) = (\mathbf{v}-\mathbf{w}) \cdot (\mathbf{v}-\mathbf{w}).$$

In particular, taking $\mathbf{w} = \mathbf{0}$ and using $\theta(\mathbf{0}) = \mathbf{0}$, we get $\theta(\mathbf{v}) \cdot \theta(\mathbf{v}) = \mathbf{v} \cdot \mathbf{v}$. Expanding the displayed formula above, and substituting in this formula and the corresponding one for $\mathbf{w}$, gives

$$\theta(\mathbf{v}) \cdot \theta(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}.$$

Thus $\{\theta(\mathbf{e}_1), \ldots, \theta(\mathbf{e}_n)\}$ is an orthonormal basis of $\mathbb{R}^n$, so the matrix P whose columns are the $\theta(\mathbf{e}_j)$ is orthogonal. Now for any vector $\mathbf{v} = (v_1, \ldots, v_n)$ we can write $\theta(\mathbf{v}) = \lambda_1 \theta(\mathbf{e}_1) + \cdots + \lambda_n \theta(\mathbf{e}_n)$ for some scalars $\lambda_1, \ldots, \lambda_n$. Then $\lambda_i = \theta(\mathbf{v}) \cdot \theta(\mathbf{e}_i) = \mathbf{v} \cdot \mathbf{e}_i = v_i$, so $\theta(\mathbf{v}) = v_1 \theta(\mathbf{e}_1) + \cdots + v_n \theta(\mathbf{e}_n) = P\mathbf{v}$. □

5.26 Example. Any orthogonal 2×2 matrix is of the form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

since the first column is a vector of length 1, say at angle α, and then the second column is perpendicular to this, so at angle $\alpha+\pi/2$ or $\alpha-\pi/2$. The first type has determinant 1, and corresponds to a rotation by angle α. The second type has determinant -1, and corresponds to a reflection in the line with angle $\alpha/2$.

[To check the last statement, a reflection in this line is the same as a rotation by angle $-\alpha/2$, followed by a reflection in the x-axis, followed by a rotation by angle $\alpha/2$. This is given by the matrix

$$\begin{pmatrix} \cos\frac{\alpha}{2} & -\sin\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\frac{\alpha}{2} & \sin\frac{\alpha}{2} \\ -\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} \end{pmatrix}.]$$

5.27 Lemma. *Every matrix $P \in SO_3(\mathbb{R})$ has 1 as an eigenvalue.*

Proof. It suffices to show that $\det(P - I) = 0$. Recall that $\det P = \det(P^T)$, so $\det(P^T) = 1$. Since P is orthogonal $P^T(P - I) = (I - P)^T$. Then

$$\det(P - I) = \det P^T(P - I) = \det(I - P)^T = \det(I - P).$$

But for any 3×3 matrix, B, $\det(-B) = -\det B$. Thus $\det(P - I) = -\det(P - I)$, so $\det(P - I) = 0$. □

5.28 Theorem. *The rotations of $\mathbb{R}^2$ and $\mathbb{R}^3$ fixing the origin are the linear maps θ_P given by matrices P in $SO_2(\mathbb{R})$ and $SO_3(\mathbb{R})$ respectively.*

Proof. For $\mathbb{R}^2$ this is clear. Sketch for $\mathbb{R}^3$. Say P is in $SO_3(\mathbb{R})$. It has 1 as an eigenvalue. Take an eigenvector of length 1 and extend to an orthonormal basis of $\mathbb{R}^3$. This gives an an orthogonal matrix Q with $Q^{-1}PQ$ having upper triangular block form

$$Q^{-1}PQ = \begin{pmatrix} 1 & * \\ 0 & B \end{pmatrix}.$$

But this matrix is orthogonal (since P and Q are), which implies that the $*$ block must be zero. Now the block B must be in $SO_2(\mathbb{R})$, so a 2×2 rotation matrix. Then the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & B \end{smallmatrix}\right)$ is a rotation about the axis $(1, 0, 0)^T$, and P is the corresponding rotation in the coordinate system given by the columns of Q^{-1}. □

5.29 Corollary (Euler's Theorem)**.** *The composition of two rotations of $\mathbb{R}^3$ about axes through the origin is also a rotation.*

5.30 Theorem. *Every finite subgroup of $SO_3(\mathbb{R})$ is one of the following*
- the group of planar rotations of a regular n-gon ($\cong C_n$)
- the full group of symmetries of a regular n-gon ($\cong D_n$)
- the group of rotations preserving a regular tetrahedron ($\cong A_4$)
- the group of rotations preserving a cube (or regular octahedron) ($\cong S_4$)
- the group of rotations preserving a regular icosahedron (or dodecahedron) ($\cong A_5$)

This question paper consists of
4 printed pages, each of which
is identified by the reference MATH202201.

All calculators must carry
an approval sticker issued
by the School of Mathematics.

School of Mathematics

January 2018

MATH202201

Groups and Vector Spaces

Time allowed: 2 hours 30 minutes

You must attempt to answer 4 questions.
If you answer more than 4 questions, only your best 4 answers will be counted towards your final mark for this exam.
All questions carry equal marks.

1. (i) Define the terms *group* and *subgroup* of a group.

Draw up the group table for the quaternion group Q of order 8 having elements $\pm 1, \pm \mathrm{i}, \pm \mathrm{j}, \pm \mathrm{k}$ where $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1$, $\mathrm{ij} = -\mathrm{ji} = \mathrm{k}$, $\mathrm{jk} = -\mathrm{kj} = \mathrm{i}$, $\mathrm{ki} = -\mathrm{ik} = \mathrm{j}$.

Determine which of the following are subgroups of Q, giving reasons:

(a) $\{\mathrm{i}, -\mathrm{i}, \mathrm{j}, -\mathrm{j}\}$, (b) $\{1, -1, \mathrm{k}, -\mathrm{k}\}$, (c) $\{1, \mathrm{i}, \mathrm{j}, -1\}$, (d) $\{1, -1\}$.

(ii) Let $\mathbb{Z}_n^*$ stand for the group of integers in $\{1, 2, \ldots, n-1\}$ coprime with n under multiplication modulo n. Which of (a) $\mathbb{Z}_{13}^*$, (b) $\mathbb{Z}_{14}^*$, (c) $\mathbb{Z}_{15}^*$, (d) $\mathbb{Z}_{16}^*$ are cyclic? Given that two of the groups (a), (b), (c), (d) are isomorphic, find which they are, explaining your reasons.

2. (i) State Lagrange's Theorem, and deduce from it that the order of any element of a finite group divides the order of the group.

(ii) Prove that $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ are non-isomorphic groups of order 4, but that any group of order 4 is isomorphic to either $\mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$.

(iii) Describe a non-abelian group G of order 12, and state the orders of all its elements. Does G have a subgroup of order 6?

(iv) Determine the conjugacy classes of the dihedral group D_4 whose group table is shown, explaining your reasoning.

	I	R	R^2	R^3	H	V	D	D'
I	I	R	R^2	R^3	H	V	D	D'
R	R	R^2	R^3	I	D'	D	H	V
R^2	R^2	R^3	I	R	V	H	D'	D
R^3	R^3	I	R	R^2	D	D'	V	H
H	H	D	V	D'	I	R^2	R	R^3
V	V	D'	H	D	R^2	I	R^3	R
D	D	V	D'	H	R^3	R	I	R^2
D'	D'	H	D	V	R	R^3	R^2	I

3. (i) Define *odd* and *even* permutations of a finite set X. Show that any permutation is either even or odd, and that the family of even permutations forms a normal subgroup of the group of all permutations of X.

(ii) Express the permutations f and g of $\{1,2,3,4,5,6,7,8,9\}$ given by

$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 9 & 7 & 3 & 5 & 8 & 1 & 6 \end{pmatrix}$ and $g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 2 & 9 & 6 & 7 & 4 & 3 & 1 & 8 \end{pmatrix}$

as (a) products of disjoint cycles, (b) products of transpositions.

Which of f, g, fg, gf, $f^{-1}gf$ are even?

(iii) Define *normal subgroup* of a group, and state the first isomorphism theorem for groups.

Prove that the map $\theta : (\mathbb{R}, +) \to \mathbb{C}^*$ given by $\theta(x) = \mathrm{e}^{2\pi \mathrm{i} x}$ is a homomorphism. Find the kernel and image of θ. Prove that $\mathbb{R}/\mathbb{Z} \cong \{z \in \mathbb{C} : |z| = 1\}$ (where the operation on this set is multiplication).

4. (i) Let V be a vector space over a field F. Define the terms *linearly independent* subset of V, and *spanning subset* of V.

(ii) Determine with reasons which of the following sets are linearly independent or spanning:

(a) $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\}$, in $\mathbb{R}^3$

(b) $\{1 + x, x - x^3, 3 + x + 2x^3\}$, in the space of real polynomials of degree ≤ 3.

(iii) If f is a linear transformation from a vector space V to a vector space W, show that for any $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, if $\{f\mathbf{v}_1, \ldots, f\mathbf{v}_n\}$ is a linearly independent subset of W, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a linearly independent subset of V.

(iv) If U and W are subspaces of the vector space V, define the *sum* $U + W$ of U and W, and state the circumstances under which this is a *direct sum* (written $U \oplus W$). For which of the following is the sum $U + W$ direct?

(a) $U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : 2x + y - z = 0 \right\}$, $W = \left\{ \begin{pmatrix} a + b \\ 2a \\ a - 3b \end{pmatrix} : a, b \in \mathbb{R} \right\}$ in $\mathbb{R}^3$,

(b) $U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : 2x + y = z - 3t = 0 \right\}$, $W = \left\{ \begin{pmatrix} a + 2b \\ a - b \\ 2a - b \\ 3a + b \end{pmatrix} : a, b \in \mathbb{R} \right\}$ in $\mathbb{R}^4$.

5. (i) Consider the linear transformation $\theta : \mathbb{C}^3 \to \mathbb{C}^3$ given by $\theta \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+2y \\ y+\mathrm{i}z \\ z-\mathrm{i}x \end{pmatrix}$.

Determine the matrix A of θ with respect to the standard basis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Find the transition matrix from the standard basis to the basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ where $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Hence or otherwise find the matrix B of θ with respect to $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

(ii) Define the *algebraic multiplicity* and *geometric multiplicity* of an eigenvalue of a square matrix A. Determine whether or not $A = \begin{pmatrix} 3 & 4 & -4 \\ 4 & 3 & -4 \\ 4 & 4 & -5 \end{pmatrix}$ is diagonalizable, and if it is, find a non-singular matrix P such that $P^{-1}AP$ is diagonal.

(iii) Define *orthogonal vectors* and *orthonormal basis*. Use the Gram–Schmidt process to find an orthonormal basis of the subspace of $\mathbb{R}^4$ spanned by $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \end{pmatrix}$, and $\begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$.

End.

This question paper consists of
3 printed pages, each of which
is identified by the reference MATH202201.

All calculators must carry
an approval sticker issued
by the School of Mathematics.

School of Mathematics

January 2019

MATH202201

Groups and Vector Spaces

Time allowed: 2 hours 30 minutes

Answer all questions.

1. (i) Determine which of the following are groups under the stated operations. For those which are groups, state the identity and inverses, and for those which are not, give one axiom that fails.

(a) The set of 2×2 non-singular matrices, under addition,

(b) The set of positive real numbers, under division,

(c) $\{0, 2, 4, 6, 8\}$, under addition modulo 10,

(d) $\{2, 4, 6, 8, 10, 12\}$ under multiplication modulo 14.

(ii) Define a *subgroup* of a group G, and prove that the intersection of any two subgroups of G is a subgroup. (You may assume without proof that a subgroup has the same identity as G, and the inverse of an element of the subgroup is the same evaluated in G or the subgroup.) Give examples of subgroups H and K of $(\mathbb{Z}, +)$ whose union is not a subgroup, and calculate what $H \cap K$ is.

2. (i) Define the *order* of a group, and the *order of an element* of a group.

(ii) Prove Lagrange's Theorem, that the order of a subgroup of a finite group G divides the order of G.

(iii) The group table of the quaternion group Q of order 8 is given. Find the orders of all elements of Q, and find all the right cosets of the subgroups $H = \{1, -1\}$ and $J = \{1, -1, j, -j\}$.

	1	-1	i	$-i$	j	$-j$	k	$-k$
1	1	-1	i	$-i$	j	$-j$	k	$-k$
-1	-1	1	$-i$	i	$-j$	j	$-k$	k
i	i	$-i$	-1	1	k	$-k$	$-j$	j
$-i$	$-i$	i	1	-1	$-k$	k	j	$-j$
j	j	$-j$	$-k$	k	-1	1	i	$-i$
$-j$	$-j$	j	k	$-k$	1	-1	$-i$	i
k	k	$-k$	j	$-j$	$-i$	i	-1	1
$-k$	$-k$	k	$-j$	j	i	$-i$	1	-1

3. (i) Define a *permutation* of a set X. Let f and g be the permutations of $X = \{0, 1, 2, 3, 4, 5, 6\}$ given by $f(x) = x + 3$ and $g(x) = 3x$, where each is evaluated modulo 7. Write f and g as products of disjoint cycles. Also express f^2, g^2, fg, and gf as products of disjoint cycles, and for each of f, g, f^2, g^2, fg, and gf, determine whether it is even or odd.

(ii) Prove that the mapping θ from a group G to itself given by $\theta(g) = g^2$ is a homomorphism if and only if G is abelian.

(iii) State the first isomorphism theorem for groups.

By calculating the kernel and image of the map given in part (ii), deduce that $\mathbb{R}^*/\{1, -1\} \cong \mathbb{R}^+$, where $\mathbb{R}^*$ and $\mathbb{R}^+$ are the sets of non-zero and positive reals respectively (and the operation is multiplication).

4. (i) Define *linear transformation* θ from a vector space V to a vector space W over the same field F. Define the *null space* and *image* of θ. Prove that the null space is a subspace of V and the image is a subspace of W.

(ii) Determine which of the following are linear transformations from $\mathbb{R}^3$ to itself, and for those which are, find the null space and image.

(a) $\theta_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+y+1 \\ y-z-2 \\ z+x+5 \end{pmatrix}$

(b) $\theta_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix}$

(c) $\theta_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y+z \\ y-z+x \\ z+2x+3y \end{pmatrix}$

(d) $\theta_4 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+3y+4z \\ x-y+2z \\ x+9y-2z \end{pmatrix}$

5. (i) What is the *matrix* A of a linear transformation θ from $\mathbb{C}^4$ to itself with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ of $\mathbb{C}^4$?

Find the matrix of the linear transformation θ from $\mathbb{C}^4$ to itself given by

$$\theta \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x+\mathrm{i}z \\ y-\mathrm{i}t \\ x-\mathrm{i}z \\ y+\mathrm{i}t \end{pmatrix}$$

where $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \mathbf{v}_4 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$.

(ii) Define the terms *symmetric* matrix, and *orthogonal matrix* (over $\mathbb{R}$). Find the eigenvalues and orthogonal eigenvectors of the matrix $B = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 5 & 2 \\ 0 & 2 & 2 \end{pmatrix}$ and hence find an orthogonal matrix P such that $P^T BP$ is diagonal.

End.

MATH2022 Groups and vector spaces, January 2018 Answers

1. (i) A *group* is a non-empty set G on which is defined an associative binary operation so that there is an *identity* (an element e such that $eg = ge = g$ for all $g \in G$) and such that every element g has an inverse (i.e. h such that $gh = hg = e$). A *subgroup* of G is a subset H which is itself a group and such that for $h_1, h_2 \in H$, $h_1 h_2$ is the same whether evaluated in H or G.

	1	-1	i	$-i$	j	$-j$	k	$-k$
1	1	-1	i	$-i$	j	$-j$	k	$-k$
-1	-1	1	$-i$	i	$-j$	j	$-k$	k
i	i	$-i$	-1	1	k	$-k$	$-j$	j
$-i$	$-i$	i	1	-1	$-k$	k	j	$-j$
j	j	$-j$	$-k$	k	-1	1	i	$-i$
$-j$	$-j$	j	k	$-k$	1	-1	$-i$	i
k	k	$-k$	j	$-j$	$-i$	i	-1	1
$-k$	$-k$	k	$-j$	j	i	$-i$	1	-1

(a) $\{\pm\mathrm{i}, \pm\mathrm{j}\}$ is not a subgroup as it doesn't contain the identity.

(b) $\{\pm 1, \pm\mathrm{k}\}$ is a subgroup. It is closed under the operation since $k^2 = -1$, it contains the identity, 1, and contains inverses of all its elements as $1^{-1} = 1$, $(-1)^{-1} = -1$, $k^{-1} = -k$, $(-k)^{-1} = k$.

(c) This is not a subgroup as it is not closed under the operation, $ij = k$ which does not lie in the set.

(d) This is a subgroup. It is closed under the operation, contains the identity, and is closed under inverses, $1^{-1} = 1$, $(-1)^{-1} = -1$.

(ii) We find the orders of the elements:

(a) In $\mathbb{Z}_{13}^* = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, the powers of 2 are $2^0 = 1$, $2^1 = 2$, $2^2 = 4$, $2^3 = 8$, $2^4 = 3$, $2^5 = 6$, $2^6 = 12$, $2^7 = 11$, $2^8 = 9$, $2^9 = 5$, $2^{10} = 10$, $2^{11} = 7$, $2^{12} = 1$. So 2 has order $12 = |\mathbb{Z}_{13}^*|$. This is cyclic.

(b) In $\mathbb{Z}_{14}^* = \{1, 3, 5, 9, 11, 13\}$, the powers of 3 are $3^0 = 1$, $3^1 = 3$, $3^2 = 9$, $3^3 = 13$, $3^4 = 11$, $3^5 = 5$, $3^6 = 1$. So 3 has order $6 = |\mathbb{Z}_{14}^*|$. This is also cyclic.

(c) In $\mathbb{Z}_{15}^* = \{1, 2, 4, 7, 8, 11, 13, 14\}$, the elements have these orders: 1 has order 1; 2, 7, 8, 13 have order 4; 4, 11, 14 have order 2. This is not cyclic.

(d) In $\mathbb{Z}_{16}^* = \{1, 3, 5, 7, 9, 11, 13, 15\}$, the elements have these orders: 1 has order 1; 3, 5, 13 have order 4; 7, 9, 11, 15 have order 2. This is also not cyclic.

The only two of these groups of equal order are $\mathbb{Z}_{15}^*$ and $\mathbb{Z}_{16}^*$, so we deduce that they are isomorphic.

2. (i) If H is a subgroup of a finite group G, then the order of H divides the order of G.

Let $g \in G$, and consider $\langle g \rangle = \{g^m : m \in \mathbb{Z}\}$, the cyclic subgroup generated by g. Since G is finite, not all g^m are distinct, and we deduce from this and cancellation that for some least $n > 0$, $g^n = 1$, the 'order' of g. Since $\langle g \rangle = \{1, g, g^2, \ldots, g^{n-1}\}$ has order n, it follows by Lagrange's Theorem that n divides $|G|$.

(ii) $\mathbb{Z}_4$ is cyclic—having an element of order 4, but in $\mathbb{Z}_2 \times \mathbb{Z}_2$, all elements have order 1 or 2, so $\mathbb{Z}_4 \ncong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Let G be any group of order 4.

Case 1: G has an element of order 4. Then G is cyclic, hence isomorphic to $\mathbb{Z}_4$ (an explicit isomorphism could be given, but is not required).

Case 2: G has no elements of order 4. Then all elements have order 1 or 2. Pick $a \neq 1$. Then a has order 2, and $\{1, a\}$ is a subgroup and $a^2 = 1$. Pick $b \in G \setminus \{1, a\}$. Then $1, a, b, ab$ are all distinct, since $ab = 1 \Rightarrow b = a^{-1} = a$, $ab = a \Rightarrow b = 1$, and $ab = b \Rightarrow a = 1$. Hence $G = \{1, a, b, ab\}$. Also b and ab have order 2. Hence $abab = 1$, so $ba = a^{-1}b^{-1} = ab$ and G is abelian. We see by inspection that the group table for G is the same as that for $\mathbb{Z}_2 \times \mathbb{Z}_2$ (again, an explicit isomorphism may be given, but is not required).

(iii) There are two possible answers.

Answer 1: $G = D_6$, the symmetries of a regular hexagon $= \{1, r, r^2, r^3, r^4, r^5, s, rs, r^2s, r^3s, r^4s, r^5s\}$ where r is a $\pi/3$ rotation and s is a reflection. Since $sr = r^5s$ this is not abelian. Orders of elements are: 1 has order 1; $r^3, s, rs, r^2s, r^3s, r^4s, r^5s$ have order 2; r, r^5 have order 6; r^2, r^4 have order 3. This *has* a subgroup of order 6, namely $\langle r \rangle$.

Answer 2: $G = \mathcal{A}_4$, the alternating group on $\{1, 2, 3, 4\}$. Orders of the elements are: *id* has order 1; $(1\,2\,3), (1\,3\,2), (1\,2\,4), (1\,4\,2), (1\,3\,4), (1\,4\,3), (2\,3\,4), (2\,4\,3)$ have order 3; $(1\,2)(3\,4)$, $(1\,3)(2\,4), (1\,4)(2\,3)$ have order 2. This has no subgroup of order 6 (proof not expected).

(iv) We find centralizers and their indices in D_4 to determine the sizes of the conjugacy classes. I and R^2 commute with all elements, and so they form conjugacy classes on their own. The centralizer of R is $\{I, R, R^2, R^3\}$, so it has 2 conjugates, R and $H^{-1}RH = R^3$. The centralizer of H is $\{I, H, R^2, V\}$ so its conjugacy class is $\{H, R^{-1}HR\} = \{H, V\}$. Similarly, the conjugacy class of D is $\{D, D'\}$.

3. (i) A *transposition* is a permutation which interchanges two points and fixes all others. A permutation is *odd* or *even* if it can be written as the product of an odd or even number of transpositions, respectively. Given any permutation f of X, write f as a product of disjoint cycles, and then note that $(a_1\, a_2\, \ldots\, a_k) = (a_1\, a_2)(a_2\, a_3)\ldots(a_{k-1}\, a_k)$, so f may be written as a (finite) product of transpositions. This number is either odd or even, so f itself is either odd or even.

Let $\mathrm{Alt}(X)$ denote the set of even permutations of X where $|X| = n$. Then $\mathrm{Alt}(X)$ is closed under the operation since even + even is even. It contains the identity since 0 is an even number, and it is closed under forming inverses since $[(a_1\, b_1)(a_2\, b_2)\ldots(a_k\, b_k)]^{-1} = (a_k\, b_k)(a_{k-1}\, b_{k-1})\ldots(a_1\, b_1)$. Thus $\mathrm{Alt}(X)$ is a subgroup. To see that $\mathrm{Alt}(X)$ is normal note that if $f \in \mathrm{Alt}(X)$ and $g \in \mathrm{Sym}(X)$ then f, g can be written as the product of m, n transpositions where m is even, and so $g^{-1}fg$ can be written as the product of $m + 2n$ transpositions, which is also even.

(ii) (a) $f = (1\,2\,4\,7\,8)(3\,9\,6\,5)$, $g = (1\,5\,7\,3\,9\,8)(4\,6)$.

(b) $f = (1\,2)(2\,4)(4\,7)(7\,8)(3\,9)(9\,6)(6\,5)$, $g = (1\,5)(5\,7)(7\,3)(3\,9)(9\,8)(4\,6)$.

Thus f is odd and g is even.

Hence just g and $f^{-1}gf$ in the list are even, and the others are odd.

(iii) If $N \subseteq G$, then N is a *normal subgroup* if it is a subgroup, i.e. N is closed under the operation and formation of inverses, and contains the identity, *and* $n \in N, g \in G \Rightarrow g^{-1}ng \in N$.

The first isomorphism theorem: If θ is a homomorphism from a group G to a group H, then $N = ker\theta = \{g \in G : \theta(g) = 1\}$ is a normal subgroup of G, and the natural map φ given by $\varphi(Ng) = \theta(g)$ is an isomorphism from the quotient group G/N to the image of θ. Conversely, if N is a normal subgroup of G, then the natural map θ from G to G/N given by $\theta(g) = Ng$ is a homomorphism having kernel equal to N.

The given map is a homomorphism since $\theta(x + y) = \mathrm{e}^{2\pi\mathrm{i}(x+y)} = \mathrm{e}^{2\pi\mathrm{i}x}\mathrm{e}^{2\pi\mathrm{i}y} = \theta(x)\theta(y)$.

$x \in ker\theta \Leftrightarrow \mathrm{e}^{2\pi\mathrm{i}x} = 1 \Leftrightarrow x \in \mathbb{Z}$, so $ker\theta = \mathbb{Z}$. The image of θ is $\{z \in \mathbb{C} : |z| = 1\}$, so that fact that $\mathbb{R}/\mathbb{Z} \cong \{z \in \mathbb{C} : |z| = 1\}$ follows at once from the first isomorphism theorem.

4. (i) A subset I of V is *linearly independent* if for any finitely many distinct members $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of I, and scalars λ_i, if $\sum_{i=1}^n \lambda_i \mathbf{v}_i = \mathbf{0}$, then all λ_i are 0. The set S *spans* V if every member of V may be written in the form $\sum_{i=1}^n \lambda_i \mathbf{v}_i$ for (finitely many) members $\mathbf{v}_i$ of S, and scalars λ_i.

(ii) (a) We test for linear independence. Let $\alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + \gamma \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} = \mathbf{0}$. Hence $\alpha + 3\gamma = \alpha + \beta = 2\beta + \gamma = 0$. From these we readily deduce that $\alpha = \beta = \gamma = 0$, so these vectors *are* linearly independent. Since the space is 3-dimensional, it is also spanning.

(b) Again testing for linear independence, we try

$\alpha(1+x)+\beta(x-x^3)+\gamma(3+x+2x^3) = 0$, and we find that $\alpha+3\gamma = \alpha+\beta+\gamma = -\beta+2\gamma = 0$. Hence $\alpha = -3\gamma$, $\beta = 2\gamma$, so a possible solution is given by $\alpha = -3$, $\beta = 2$, $\gamma = 1$. Hence these are *not* linearly independent.

(iii) Suppose that $\lambda_1 \mathbf{v}_1 + \ldots \lambda_n \mathbf{v}_n = \mathbf{0}$. Applying f, $f(\lambda_1 \mathbf{v}_1 + \ldots \lambda_n \mathbf{v}_n) = \mathbf{0}$, and since f is linear, $\lambda_1 f\mathbf{v}_1 + \ldots \lambda_n f\mathbf{v}_n = \mathbf{0}$. As $f\mathbf{v}_1, \ldots, f\mathbf{v}_n$ are linearly independent, $\lambda_i = 0$ for all i, as desired.

(iv) The sum $U + W = \{\mathbf{u} + \mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}$. It is *direct* if $U \cap W = \{\mathbf{0}\}$.

(a) Both U and W are 2-dimensional in a 3-dimensional space, so this is not direct.

(b) Suppose $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in U \cap W$. Then for some a and b, $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} a+2b \\ a-b \\ 2a-b \\ 3a+b \end{pmatrix}$. Since it lies in U, $2(a+2b) + (a-b) = (2a-b) - 3(3a+b) = 0$. Hence $3a + 3b = 0$ and $-7a - 4b = 0$, from which it follows that $a = b = 0$. Hence $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \mathbf{0}$, so the sum is direct.

5. (i) $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & \mathrm{i} \\ -\mathrm{i} & 0 & 1 \end{pmatrix}$. The transition matrix is $P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

We calculate the inverse, $P^{-1} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, and it follows that $B = P^{-1}AP = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & \mathrm{i} \\ -\mathrm{i} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+\mathrm{i} & \mathrm{i} & \mathrm{i} \\ 2-\mathrm{i} & 1-\mathrm{i} & 2-\mathrm{i} \\ -\mathrm{i} & -\mathrm{i} & 1 \end{pmatrix}$.

(ii) The *algebraic multiplicity* of the eigenvalue λ is the largest power of $t - \lambda$ that divides the characteristic polynomial. Its *geometric multiplicity* is the dimension of the eigenspace $\{\mathbf{v} \in V : A\mathbf{v} = \lambda\mathbf{v}\}$.

$$det(tI-A) = \begin{vmatrix} t-3 & -4 & 4 \\ -4 & t-3 & 4 \\ -4 & -4 & t+5 \end{vmatrix} = \begin{vmatrix} t-3 & -4 & 4 \\ -t-1 & t+1 & 0 \\ -4 & -4 & t+5 \end{vmatrix} = (t+1)\begin{vmatrix} t-3 & -4 & 4 \\ -1 & 1 & 0 \\ -4 & -4 & t+5 \end{vmatrix} =$$
$$(t+1)\begin{vmatrix} t-7 & -4 & 4 \\ 0 & 1 & 0 \\ -8 & -4 & t+5 \end{vmatrix} = (t+1)^2(t-3).$$

Eigenvectors for -1: $\begin{pmatrix} -4 & -4 & 4 \\ -4 & -4 & 4 \\ -4 & -4 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$, so $z = x + y$. There are linearly independent solutions $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ (so algebraic multiplicity = geometric multiplicity = 2).

Eigenvectors for 3: $\begin{pmatrix} 0 & -4 & 4 \\ -4 & 0 & 4 \\ -4 & -4 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$, so $x = y = z$, solution $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ (so algebraic multiplicity = geometric multiplicity = 1).

The matrix is diagonalizable, and we may let $P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

(iii) (Column) vectors $\mathbf{v}$ and $\mathbf{w}$ are *orthogonal* if $\mathbf{v}^T\mathbf{w} = 0$. An *orthonormal basis* is a basis consisting of vectors of length 1, such that any two are orthogonal.

In the notation given in the lectures, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$, and

following the method given, $\mathbf{u}_1 = \mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{u}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \end{pmatrix} - \frac{2}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$, $\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{(-2)}{4}\begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

Finally, dividing by the lengths to get unit vectors, we have $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$.

END

MATH2022 Groups and vector spaces, January 2019 Answers

1. (i) (a) This is not a group, as it is not closed under the operation. For instance $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ are both non-singular, but their sum is the zero matrix, which is certainly not non-singular.

(b) This is not a group, since the operation is not associative, e.g. $8 \div (4 \div 2) = 4$ whereas $(8 \div 4) \div 2 = 1$. (The answer that says that there is no identity would also be acceptable, but not that there are no inverses, since that statement doesn't make any sense unless one already has an identity.)

(c) This is a group. The identity is 0, and the inverse of 0 is itself, and for the other elements x is $10 - x$.

(d) This is a group. The identity is 8, and the inverses are as follows: $2^{-1} = 4, 4^{-1} = 2, 6^{-1} = 6, 8^{-1} = 8, 10^{-1} = 12, 12^{-1} = 10$.

(ii) A *subgroup* H of a group G is a subset of G which is itself a group under the same operation inherited from G. (From this one can deduce that G and H have the same identities, and that the inverse of an element of H is the same evaluated in H or G, but these are *not* part of the definition, which is why the question allows us to assume this.)

Suppose that H and K are subgroups of G. We see that $H \cap K$ is a group under the operation inherited from G. Clearly $H \cap K \subseteq G$. First, $H \cap K$ is closed under the operation. For suppose that $a, b \in H \cap K$. Since $a, b \in H$ and H is a subgroup, it follows that $ab \in H$. Similarly, $ab \in K$. Hence $ab \in H \cap K$. By what we are allowed to assume, the identity (of G) lies in both H and K, where in each case it is the identity. Hence it also lies in $H \cap K$. Let $a \in H \cap K$. Then as $a \in H$, $a^{-1} \in H$, and similarly, $a^{-1} \in K$. Hence $a^{-1} \in H \cap K$. This shows that $H \cap K$ is a group under the same operation, hence is a subgroup.

The subgroups of $(\mathbb{Z}, +)$ are of the form $k\mathbb{Z}$ for integers $k \geq 0$ (proof not asked for, or expected), so the easiest example of what is requested is $2\mathbb{Z} =$ the set of even integers, and $3\mathbb{Z}$, the set of multiples of 3. Here $2, 3 \in H \cup K$, but $2 + 3 = 5 \notin H \cup K$.

$H \cap K = 6\mathbb{Z}$, the set of multiples of 6.

2. (i) The *order* of a group G is its number of elements (finite or infinite). If $g \in G$, the *order of* g is equal to the least positive integer n such that $g^n = 1$, if any; if none exists, the g is said to have *infinite* order.

(ii) Let H be a subgroup of G. Define $\sim$ on G by $x \sim y$ if $xy^{-1} \in H$. We verify that $\sim$ is an equivalence relation:

$\sim$ is reflexive since $xx^{-1} = 1 \in H$, which gives $x \sim x$.

$\sim$ is reflexive, since if $x \sim y$, then $xy^{-1} \in H$, hence $yx^{-1} = (xy^{-1})^{-1} \in H$, as H has the inverses of its elements.

$\sim$ is transitive, since if $x \sim y$ and $y \sim z$, then xy^{-1}, $yz^{-1} \in H$. Hence $xz^{-1} = (xy^{-1})(yz^{-1}) \in H$ as H is closed under the operation, so $x \sim z$.

Since $\sim$ is an equivalence relation, it partitions G into disjoint sets. The $\sim$-class containing x is $\{y : y \sim x\} = \{y : yx^{-1} \in H\} = \{y : yx^{-1} = h, \text{ some } h \in H\}$
$= \{y : y = hx, \text{ some } h \in H\}$. This set is written Hx, and is called a *right coset* of H in G. Thus G is partitioned into right cosets.

Finally, the size of any right coset is equal to $|H|$, since we can map H 1–1 onto Hx by the mapping $f(h) = hx$. This is 1–1 since $h_1x = h_2x \Rightarrow h_1 = h_2$.

We deduce that $|G|$ is equal to $|H|$ times the number of cosets, and hence $|H|$ divides $|G|$.

(iii) The orders are as follows: 1 has order 1; -1 has order 2; all other elements have order 4.

Right cosets of H are $\{1, -1\}$, $\{i, -i\}$ $\{j, -j\}$ $\{k, -k\}$.

Right cosets of J are $\{1, -1, j, -j\}$ and $\{i, -i, k, -k\}$.

3. (i) A *permutation* of X is a 1–1 and onto mapping from X to X.

$f = (0\,3\,6\,2\,5\,1\,4)$, $g = (1\,3\,2\,6\,4\,5)$.

$f^2 = (0\,6\,5\,4\,3\,2\,1)$, $g^2 = (1\,2\,4)(3\,6\,5)$.

$fg = (0\,3\,5\,4\,1\,6)$, $gf = (0\,2\,1\,5\,3\,4)$.

f is a cycle of odd length, so is an even permutation, and similarly, g is odd. We deduce that f^2, g^2 are even, and fg, gf are odd.

(ii) The condition for θ to be a homomorphism is that $(gh)^2 = g^2h^2$ for all $g, h \in G$. This is equivalent to saying that $ghgh = gghh$, and on cancelling on left and right, that $hg = gh$ for all g and h, which is precisely the condition for G to be abelian.

(iii) **The first isomorphism theorem**: If θ is a homomorphism from a group G to a group H, then $N = ker\,\theta = \{g \in G : \theta(g) = 1\}$ is a normal subgroup of G, and the natural map φ given by $\varphi(Ng) = \theta(g)$ is an isomorphism from the quotient group G/N to the image of θ. Conversely, if N is a normal subgroup of G, then the natural map θ from G to G/N given by $\theta(g) = Ng$ is a homomorphism having kernel equal to N.

Note that $\mathbb{R}^*$ is abelian, so we can use part (ii). The kernel of the squaring map is the set of elements whose square is the identity, which in $\mathbb{R}^*$ is precisely $\{\pm 1\}$. Since all squares are positive, and every positive real number is the square of some real number, the squaring map is onto $\mathbb{R}^+$, so the stated isomorphism follows at once from the First Isomorphism Theorem.

4. (i) A *linear transformation* θ from V to W is a mapping such that for all $\mathbf{v}_1, \mathbf{v}_2 \in V$, and $\lambda \in F$, $\theta(\mathbf{v}_1 + \mathbf{v}_2) = \theta(\mathbf{v}_1) + \theta(\mathbf{v}_2)$ and $\theta(\lambda \mathbf{v}_1) = \lambda\theta(\mathbf{v}_1)$.

The *null space* N of θ is $\{\mathbf{v} \in V : \theta(\mathbf{v}) = \mathbf{0}\}$ and the *image* of θ is $\{\theta(\mathbf{v}) : \mathbf{v} \in V\}$.

N is a subspace of V: N is closed under $+$ since if $\mathbf{v}_1, \mathbf{v}_2 \in N$ then $\theta(\mathbf{v}_1 + \mathbf{v}_2) = \theta(\mathbf{v}_1) + \theta(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$. It is closed under multiplication by scalars, since if $\mathbf{v} \in N$ and $\lambda \in F$, then $\theta(\lambda\mathbf{v}) = \lambda\theta(\mathbf{v}) = \lambda\mathbf{0} = \mathbf{0}$. It contains $\mathbf{0}$ since $\theta(\mathbf{0}) = \mathbf{0}$.

$im(\theta)$ is a subspace of W: $im(\theta)$ is closed under $+$ since $\theta(\mathbf{v}_1) + \theta(\mathbf{v}_2) = \theta(\mathbf{v}_1 + \mathbf{v}_2)$. It is closed under multiplication by scalars since $\lambda\theta(\mathbf{v}) = \theta(\lambda\mathbf{v})$, and it contains $\mathbf{0}$ since $\mathbf{0} = \theta(\mathbf{0})$.

(ii) (a) Not a linear transformation since $\theta_1(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}) \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

(b) Not a linear transformation since $\theta_2(\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}) = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix} \neq 2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 2\theta_2(\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix})$.

(c) This is a linear transformation. For the null space N we solve $\left.\begin{array}{c} x+y+z=0 \\ y-z+x=0 \\ z+2x+3y=0 \end{array}\right\}$, so $z = x + y = -z$ giving $z = 0$. Also $x + y = 2x + 3y = 0$ giving $x = y = z = 0$. The null space is therefore $\left\{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right\}$, and the image is the whole of $\mathbb{R}^3$.

(d) This is a linear transformation. For the null space N we solve $\left.\begin{array}{c} 2x+3y+4z=0 \\ x-y+2z=0 \\ x+9y-2z=0 \end{array}\right\}$. The first two equations show that $y = 0$, so $x + 2z = 0$ and $x - 2z = 0$. Hence the null space is again $\left\{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right\}$, and the image is the whole of $\mathbb{R}^3$.

5. (i) The matrix A *representing* a linear transformation $\theta : V \to V$ with respect to $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is given by $A = (a_{ij})$ where $\theta(\mathbf{v}_j) = \sum_{i=1}^{4} a_{ij}\mathbf{v}_i$.

$$\theta(\mathbf{v}_1) = \begin{pmatrix} 1+\mathrm{i} \\ 0 \\ 1-\mathrm{i} \\ 0 \end{pmatrix} = 1.\mathbf{v}_1 + 0.\mathbf{v}_2 + \mathrm{i}.\mathbf{v}_3 + 0.\mathbf{v}_4,$$

$$\theta(\mathbf{v}_2) = \begin{pmatrix} 0 \\ 1-\mathrm{i} \\ 0 \\ 1+\mathrm{i} \end{pmatrix} = 0.\mathbf{v}_1 + 1.\mathbf{v}_2 + 0.\mathbf{v}_3 + \mathrm{i}.\mathbf{v}_4,$$

$$\theta(\mathbf{v}_3) = \begin{pmatrix} 1-\mathrm{i} \\ 0 \\ 1+\mathrm{i} \\ 0 \end{pmatrix} = 1.\mathbf{v}_1 + 0.\mathbf{v}_2 + -\mathrm{i}.\mathbf{v}_3 + 0.\mathbf{v}_4, \qquad A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ \mathrm{i} & 0 & -\mathrm{i} & 0 \\ 0 & \mathrm{i} & 0 & \mathrm{i} \end{pmatrix}$$

$$\theta(\mathbf{v}_4) = \begin{pmatrix} 0 \\ -1-\mathrm{i} \\ 0 \\ -1+\mathrm{i} \end{pmatrix} = 0.\mathbf{v}_1 + (-1).\mathbf{v}_2 + 0.\mathbf{v}_3 + \mathrm{i}.\mathbf{v}_4.$$

(ii) A is *symmetric* if $A = A^T$; it is *orthogonal* if $AA^T = I$.

$$det(tI - B) = \begin{vmatrix} t-6 & 0 & 0 \\ 0 & t-5 & -2 \\ 0 & -2 & t-2 \end{vmatrix} = (t-6)(t^2 - 7t + 6) = (t-6)^2(t-1).$$

Eigenvectors for 6: $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$, so orthogonal solutions are $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$.

Eigenvectors for 1: $\begin{pmatrix} -5 & 0 & 0 \\ 0 & -4 & -2 \\ 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$, solution $\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$.

Normalizing, we get $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$, and $P^TBP = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

END