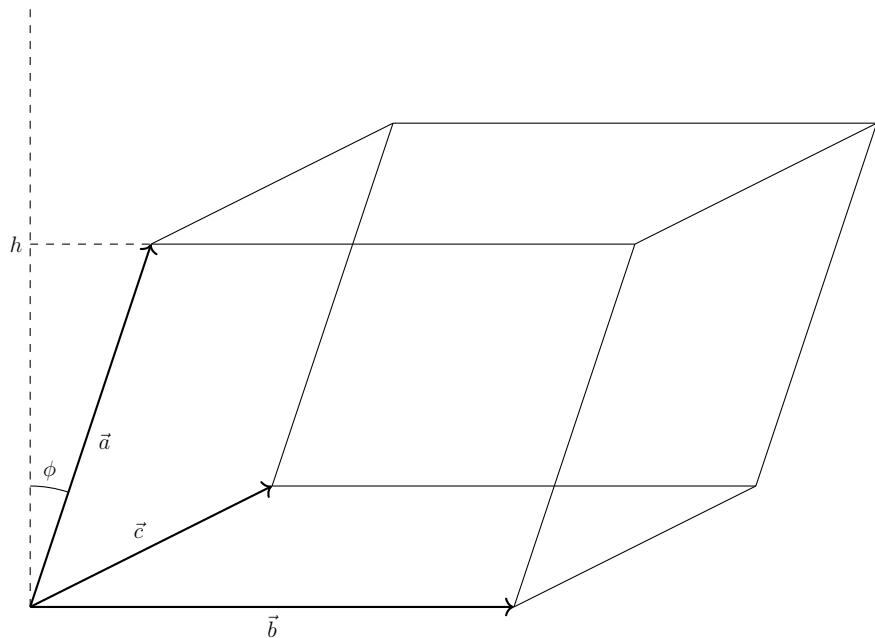


F17ZD

**Mathematics for Data Scientists 4**

**Course Notes**

Semester 2, 2025/26



Name: \_\_\_\_\_

- Copyright Heriot Watt University 2026. All rights reserved. For use of students of Xidian and Heriot Watt Universities enrolled in the course F17ZD.
- Based on notes by Jack Carr, Des Johnston, Alex Martin and Thomas Wong.

## Contents

<b>1 Complex Numbers</b>	<b>3</b>
1.1 Definitions and Arithmetic . . . . .	4
1.2 Solving Quadratic Equations . . . . .	9
1.3 The Argand Diagram . . . . .	11
1.4 Polar Form . . . . .	14
1.5 Exponential Form . . . . .	23
1.6 De Moivre's Theorem . . . . .	29
1.7 Complex Arguments in Trigonometric/Hyperbolic Functions . . . . .	34
1.8 Complex Mappings . . . . .	36
1.9 Iterating complex maps . . . . .	39
<b>2 An aside: Quaternions</b>	<b>42</b>
2.1 Definitions and Arithmetic . . . . .	42
2.2 Conjugate, Norm and Inverse . . . . .	43
2.3 Unit Quaternions and 3D Rotations . . . . .	44
<b>3 Vectors, Spans and Bases in <math>\mathbb{R}^2</math>, <math>\mathbb{R}^3</math> and <math>\mathbb{R}^n</math>.</b>	<b>47</b>
3.1 Vectors in $\mathbb{R}^2$ . . . . .	47
3.2 Vectors in $\mathbb{R}^3$ . . . . .	48
3.3 Vectors in $\mathbb{R}^n$ . . . . .	48
3.4 Systems of linear equations in vector form. . . . .	51
3.5 Spans of vectors and spanning families in $\mathbb{R}^n$ . . . . .	52
3.6 Spanning families. . . . .	55
3.7 Linearly independent families in $\mathbb{R}^n$ . . . . .	56
3.8 Bases of $\mathbb{R}^n$ . . . . .	59
3.9 Characterisation of bases of $\mathbb{R}^n$ . . . . .	59
<b>4 Vector Spaces in General</b>	<b>63</b>
4.1 Examples and definition . . . . .	63
4.2 Linear combinations in abstract vector spaces . . . . .	64
4.3 Basis and dimension . . . . .	66
4.4 Dimension of a vector space. . . . .	66
4.5 Vector subspaces . . . . .	68
<b>5 The solution space of a homogeneous system of linear equations.</b>	<b>72</b>
<b>6 Matrices</b>	<b>76</b>
6.1 Image and rank of a matrix. . . . .	76
6.2 Rank of a matrix. . . . .	78
6.3 Rank and systems of linear equations. . . . .	79
6.4 The Rank-Nullity Theorem for matrices. . . . .	80
<b>7 Eigenvalues and diagonalisability</b>	<b>85</b>
7.1 A motivating problem: predicting the dynamics of a population . . . . .	85
7.2 Eigenvectors and Eigenvalues . . . . .	86
7.3 Diagonalisability. . . . .	87
7.4 Eigenspaces and diagonalisability. . . . .	88
7.5 Characteristic polynomial . . . . .	89

7.6	Trace, determinants, and consistency check.	91
7.7	Diagonalisability of a matrix	92
7.8	Application: Dynamical systems	96
7.9	Application: Dynamics of populations.	97
7.10	Powers of a matrix.	100
<b>8</b>	<b>Stochastic Matrices</b>	<b>101</b>
8.1	A motivating problem: Optimising a distribution of bikes	101
8.2	Markov diagrams and transition matrices	102
8.3	The Perron-Frobenius Theorem for stochastic matrices	106
8.4	Application: stationary distributions in dynamical systems.	108
8.5	Application: The PageRank algorithm	109

# 1 Complex Numbers

To introduce complex numbers, we'll take a quick (and somewhat abridged) tour through the history of mathematics. The first numbers that appeared historically (evidence in cave art etc.), and the first numbers that you learned about, are the 'counting numbers' or natural numbers. Aside from counting objects,

- 5 cookies
- 2 sheep,

basic operations can also be performed,

- add 2 sheep and 3 sheep to get 5 sheep
- 7 cookies take away 3 cookies leaves 4 cookies.

With these operations comes the ability to solve equations:  $? + 3 = 8$  can be solved with  $? = 5$ . What about trying to solve the following?  $? + 5 = 3$ .

We are now all quite happy with the idea of negative numbers and the solution  $-2$ . However this hasn't always been the case; negative numbers initially appeared to some to be 'unnatural' (what does  $-2$  cookies mean?), and even a recently as the Renaissance in Europe they were still widely mistrusted. But now their use has been recognised. For example, a bank account with a negative balance makes perfect sense!

We now write  $x + 3 = 8$  and solve to get  $x = 5$ . How about solving the equation

$$x^2 - 2x - 3 = 0$$

We have a nice formula for solving such equations (courtesy of thousands of years of work dating from the Babylonians, the Indians, the Greeks, the Chinese, the Persians and the Egyptians):

$$x = \frac{2 \pm \sqrt{4 + 12}}{2} = 1 \pm 2 = -1, 3.$$

Now let us try to solve the equation

$$\frac{x^2}{2} - x + 1 = 0.$$

Proceeding as before gives

$$x = \frac{1 \pm \sqrt{1 - 2}}{1} = 1 \pm \sqrt{-1}.$$

Up until now, we would stop here and say that the equation has no (real) roots. This is consistent with most of human history. It took many years from initial work by Bombelli (1572 AD) for them to be accepted in the work of Euler (1707-1783) and Gauss (1777-1855).

## 1.1 Definitions and Arithmetic

We first define some notation of complex numbers:

### Definition 1.1 Complex Number

We define  $i = +\sqrt{-1}$  to be the positive square root of  $-1$ . So  $i^2 = -1$ . A *complex number*,  $z$ , is any expression of the form  $z = a + ib$  for real numbers  $a$  and  $b$ . For a complex number  $z = a + ib$ , we say  $a$  is the *real part* and  $b$  is the *imaginary part*. This is denoted

$$\operatorname{Re}(z) = a \quad \operatorname{Im}(z) = b$$

 The definition allows for  $b = 0$ . Hence all real numbers are included in the definition of complex numbers.

 In Electrical Engineering the notation  $j = \sqrt{-1}$  often used instead of  $i$ .

### Example 1.2

Examples of complex numbers include:

$$5 + i2, \quad 7 - i, \quad -4 + 3i, \quad e, \quad 3, \quad i\pi.$$

The real and imaginary parts of the above numbers are as follows:

$z$	$5 + i2$	$7 - i$	$-4 + 3i$	$e$	$3$	$i\pi$
$\operatorname{Re}(z)$	5	7	-4	$e$	3	0
$\operatorname{Im}(z)$	2	-1	3	0	0	$\pi$

Notes:

- $7 - i$  is the same as  $7 + (-1)i$ .
- $i3 = 3i$ .
- $e = e + i0$  and  $i\pi = 0 + i\pi$

As for real numbers, we can apply some basic arithmetic operations to complex numbers.

### Definition 1.3 Equality

Two complex numbers  $z = a + ib$  and  $w = c + id$  are equal, denoted  $z = w$  if  $a = c$  and  $b = d$ .

For example: We have  $1 + 2i = 2i + 1$  and  $1 + 2i \neq 2 + 2i$ .

Basic operations on complex numbers behave the way we expect as per real numbers

### Definition 1.4 Basic operations

Given complex numbers  $z = a + ib$  and  $w = c + id$ , we have the following operations:

- Addition:

$$\begin{aligned} z + w &= (a + ib) + (c + id) \\ &= (a + c) + i(b + d) \end{aligned}$$

- Subtraction:

$$\begin{aligned} z - w &= (a + ib) - (c + id) \\ &= (a - c) + i(b - d) \end{aligned}$$

- Multiplication:

$$\begin{aligned} z \cdot w &= (a + ib) \cdot (c + id) \\ &= (ac - bd) + i(ad + bc) \end{aligned}$$

We can verify these operations by expanding the brackets and then regrouping by real and imaginary components.

### Example 1.5

Take complex multiplication for example:

$$\begin{aligned} (a + ib) \cdot (c + id) &= ac + iad + ibc + i^2bd \\ &= ac + iad + ibc - bd \\ &= (ac - bd) + i(ad + bc) \end{aligned}$$

where we used the identity  $i^2 = -1$ .

### Example 1.6

Addition and subtraction:

$$\begin{aligned} (5 + 4i) + (2 - 3i) &= (5 + 2) + (4 - 3)i = 7 + i \\ (\pi + 2i) - (\pi + i) &= (\pi - \pi) + (2 - 1)i = i \end{aligned}$$

Multiplication:

$$\begin{aligned} (1 - i)(2 + 3i) &= 2 + 3 + 3i - 2i = 5 + i. \\ (1 - i)^2 &= (1 - i)(1 - i) = (1 - 1) + i(-1 - 1) = -2i \\ (1 + 2i)(1 - 2i) &= (1 + 4) + i(2 - 2) = 5. \end{aligned}$$

We can also combine operations together using our standard order of operations

$$(1+i)((2-3i)+(12+8i)) = (1+i)(14+5i) = 9+19i.$$

You'll notice that division is not on this list yet. That's because division is a bit more complicated (as for real numbers)<sup>1</sup>. Recall from real numbers, we can convert division into multiplication by using the reciprocal. That is  $\frac{a}{b} = a \times \frac{1}{b}$ . For complex numbers we will use a similar approach. Given  $z = a + ib$ , we will attempt to find a number  $w = p + iq$  such that  $w = \frac{1}{z}$ .

Before we get to that, we will introduce a new concept that will be useful

### Definition 1.7 Complex conjugate

Given a complex number  $z = a + ib$ , the complex conjugate,  $\bar{z}$ , is defined to be

$$\bar{z} = a - ib$$

In particular, we have  $\operatorname{Re}(\bar{z}) = \operatorname{Re}(z)$  and  $\operatorname{Im}(\bar{z}) = -\operatorname{Im}(z)$ .

 Some resources will use  $z^*$  to denote complex conjugate

### Example 1.8

- If  $z = 1 + i$  then  $\bar{z} = 1 - i$ .
- If  $z = 2i$ , then  $\bar{z} = -2i$ .
- If  $z = 4$ , then  $\bar{z} = 4$ .

Note that if  $r$  is a real number (i.e.  $\operatorname{Im}(r) = 0$ ), then  $r = \bar{r}$ . This is why the conjugate does not make an appearance prior to our study of complex numbers.

Equipped with the complex conjugate, we can now rewrite the reciprocal of a complex number.

### Example 1.9

Let  $z = a + ib$ . We can rewrite  $\frac{1}{z}$  by multiplying by 1 in a specific way.

$$\begin{aligned}\frac{1}{z} &= \frac{1}{a+ib} \\ &= \frac{1}{a+ib} \cdot \frac{a-ib}{a-ib} && \text{Complex Conjugate} \\ &= \frac{a-ib}{(a+ib)(a-ib)} \\ &= \frac{a-ib}{a^2+b^2} \\ &= \frac{a}{a^2+b^2} + i \frac{-b}{a^2+b^2}\end{aligned}$$

<sup>1</sup>Unlike real numbers, it is not clear what the expression  $\frac{4+3i}{1+2i}$  means pictorially (yet)

In particular, if  $w = \frac{1}{z}$ , we can express  $w = p + iq$  with

$$p = \frac{a}{a^2 + b^2} \quad q = \frac{-b}{a^2 + b^2}$$

The essence of this process is to transfer the imaginary number  $i$  from the denominator to the numerator. The denominator  $a^2 + b^2$  is derived from the difference of squares and is guaranteed to be a real number. This process is often referred to as *realising the denominator*.

Further, the denominator tells us that  $z\bar{z} = a^2 + b^2$  which will be a real number.

### Exercise 1.10

Let  $z = 1 + 2i$ . Express  $\frac{1}{z}$  as a complex number  $w = c + id$ .

**Solution:** Using the derivation we had above:

$$\begin{aligned}\frac{1}{z} &= \frac{1}{1+2i} \\ &= \frac{1}{1^2 + 2^2} + i \frac{-2}{1^2 + 2^2} \\ &= \frac{1}{5} + i \frac{-2}{5}\end{aligned}$$

We can also fill in the details as we go.

### Exercise 1.11

Express  $\frac{1}{3+4i}$  in the form  $a + ib$ .

**Solution:** The complex conjugate of  $3 + 4i$  is  $3 - 4i$ . Hence, we have

$$\begin{aligned}\frac{1}{3+4i} &= \frac{1}{3+4i} \cdot \frac{3-4i}{3-4i} \\ &= \frac{3-4i}{3^2 + 4^2} \\ &= \frac{3}{25} - \frac{4}{25}i\end{aligned}$$

### Exercise 1.12

Write  $\frac{1}{i}$  in the form  $a + ib$ .

**Solution:**

$$\frac{1}{i} = \frac{1}{i} \frac{-i}{(-i)^2} = \frac{-i}{1} = -i.$$

This property  $\frac{1}{i} = -i$  is a particularly useful one to remember.

With this in mind, we can now look at complex division

**Example 1.13**

To evaluate the following:

$$\begin{aligned}\frac{4+3i}{1+2i} &= (4+3i) \cdot \frac{1}{1+2i} \\ &= (4+3i) \cdot \left(\frac{1}{5} + i\frac{-2}{5}\right) \\ &= \left(4 \cdot \frac{1}{5} - 3 \cdot \frac{-2}{5}\right) + i\left(4 \cdot \frac{-2}{5} + 3 \cdot \frac{1}{5}\right) \\ &= \left(\frac{4}{5} + \frac{6}{5}\right) + i\left(\frac{-8}{5} + \frac{3}{5}\right) \\ &= \frac{10}{5} + i\frac{-5}{5} \\ &= 2 - i.\end{aligned}$$

**Example 1.14**

The algebraic operations for complex numbers satisfy the *same properties* as the corresponding operations for real numbers. For example

$$z_1(z_2 + z_3) = z_1z_2 + z_1z_3$$

$$z_1z_2 = z_2z_1$$

$$z^{-3} = \frac{1}{z^3}$$

## 1.2 Solving Quadratic Equations

We can extend the definition of  $i$  to take the square root of any negative number.

### Example 1.15

If we have  $d > 0$ , then  $-d < 0$ , so we have

$$\sqrt{-d} = \sqrt{(-1) \cdot d} = \sqrt{-1} \cdot \sqrt{d} = i\sqrt{d}$$

### Exercise 1.16

Solve the equation  $z^2 + 16 = 0$

**Solution:** We isolate  $z$  by rewriting

$$\begin{aligned} z^2 + 16 &= 0 \\ z^2 &= -16 \\ z &= \pm\sqrt{-16} \\ z &= \pm\sqrt{16}\sqrt{-1} \\ z &= \pm 4\sqrt{-1} \\ z &= \pm 4i. \end{aligned}$$

With complex numbers, we can revisit our knowledge of solving quadratics using the quadratic formula.

### Theorem 1.17 Quadratic formula

Given the quadratic equation  $az^2 + bz + c = 0$ , with real constants  $a, b, c$ . The solutions  $z$  are given by

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

The quantity under the square root  $\Delta = b^2 - 4ac$  is called the *discriminant*. In particular:

- If  $\Delta > 0$ , the two solutions are *real and distinct*.
- If  $\Delta = 0$ , the two solutions are *real and overlap*.
- If  $\Delta < 0$ , the two solutions are *complex and distinct*.

The first two cases for the discriminant should be familiar. In the case for complex solutions  $z_1, z_2$ , we have the additional property that  $\overline{z_1} = z_2$ . That is the two complex solutions form a *complex conjugate pair*.

**Exercise 1.18**

Solve the equation  $z^2 + 2z + 2 = 0$

**Solution:** The quadratic formula gives

$$\begin{aligned} z &= \frac{-2 \pm \sqrt{2^2 - 4 \times 1 \times 2}}{2}, \\ &= \frac{-2 \pm \sqrt{-4}}{2} \\ &= \frac{-2 \pm 2\sqrt{-1}}{2} \\ &= -1 \pm \sqrt{-1} = -1 \pm i. \end{aligned}$$

Thus the solutions are  $z = -1 + i$  and  $z = -1 - i$ .

**Exercise 1.19**

Find the 2 solutions  $z_1, z_2$  of the equation  $z^2 + 14z + 58 = 0$ , and check that  $\overline{z_1} = z_2$ .

**Solution:** From the quadratic formula

$$\begin{aligned} z &= \frac{-14 \pm \sqrt{(14)^2 - 4 \times 58}}{2} \\ &= \frac{-14 \pm \sqrt{-36}}{2} \\ &= \frac{-14 \pm 6\sqrt{-1}}{2} = -7 \pm 3i \end{aligned}$$

The two solutions are  $z_1 = -7 + 3i$  and  $z_2 = -7 - 3i$ . We also have  $\overline{z_1} = -7 - 3i = z_2$ .

### 1.3 The Argand Diagram

Operations on complex numbers lend themselves very nicely to visual interpretations. To construct a visual representation of complex numbers, we start with the real number line.

#### Example 1.20

We can think of real numbers as points along the real number line.

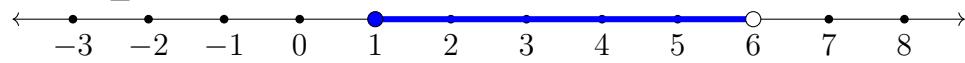


Inequalities in real numbers translates to intervals along the real number line. For example:

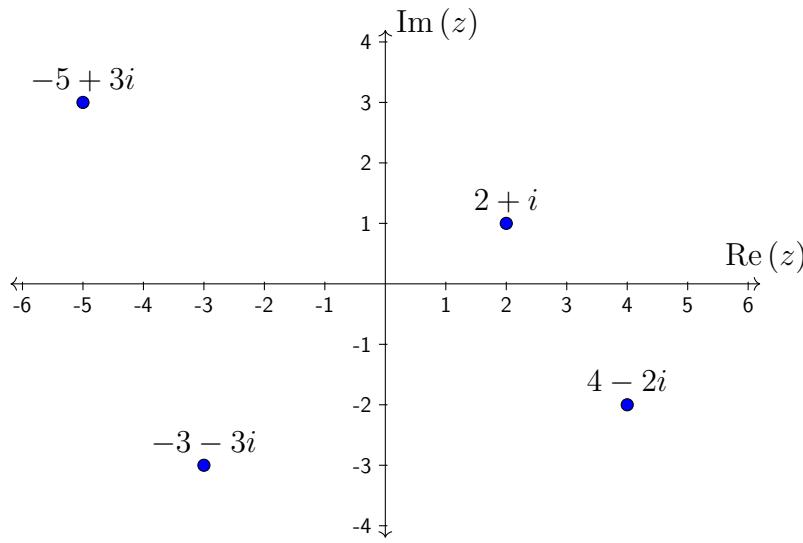
- The interval  $-1 < x \leq 3$  translates to



- The interval  $1 \leq x < 6$  translates to



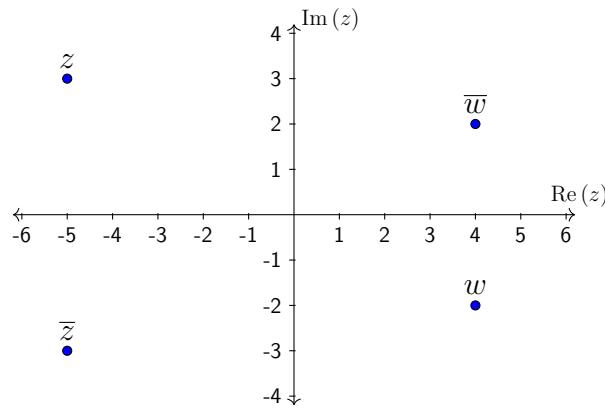
We can extend this idea to complex numbers. Since two pieces of data is required to describe a complex number (a real part and an imaginary part). We can use those two data as co-ordinates on the plane. The complex number  $z = a + bi$  is represented by the point with co-ordinates  $(a, b)$  in the plane. Complex numbers written in  $z = a + bi$  is often called *Cartesian form* for its connection with coordinates in the complex plane.



This is known as an *Argand Diagram* or the *complex plane*. The Argand diagram provides a simple visual way of representing many of the key properties of complex numbers.

**Example 1.21**

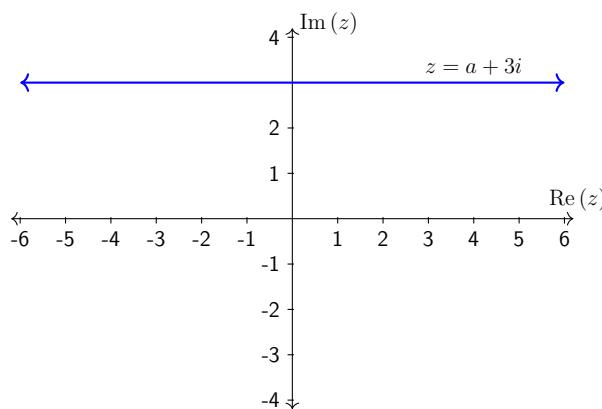
Let  $z$  be a complex number. The Argand diagram of  $\bar{z}$ , the complex conjugate of  $z$ , is the point obtained by reflection in the real axis.

**Exercise 1.22**

Plot all complex numbers of the form  $z = a + 3i$  for real numbers  $a$

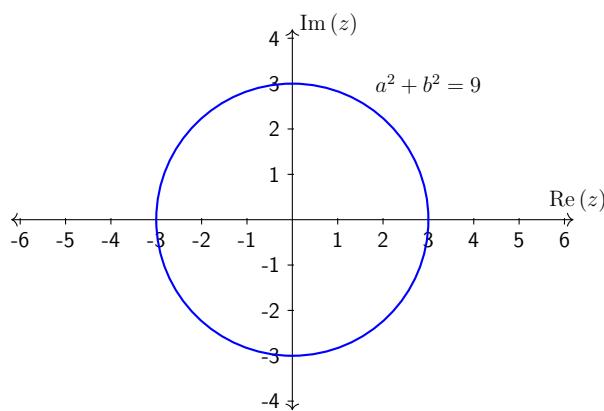
**Solution:**

The imaginary part of  $z$  is always 3 while the real part can vary. This leads to the following picture:

**Exercise 1.23**

Plot all complex numbers  $z = a + ib$  with the property that  $a^2 + b^2 = 9$ .

**Solution:** Recall that  $a^2 + b^2 = 9$  is the equation of a circle with radius 3. Thus a complex number  $a + ib$  satisfies  $a^2 + b^2 = 9$  if it lies on this circle.

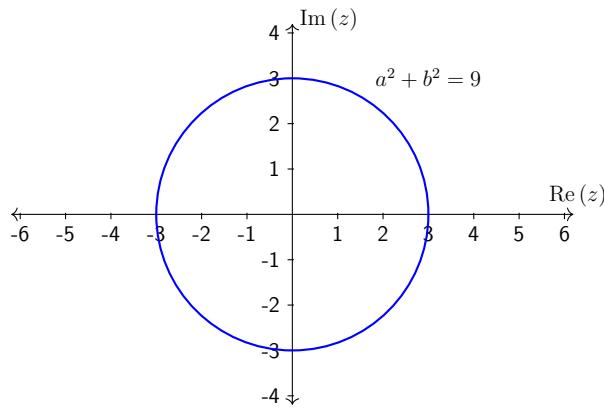


Each point on the circle is a complex number satisfying the requirement.

### Exercise 1.24

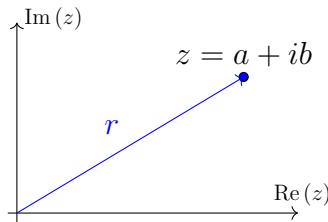
Plot all complex numbers  $z = a + ib$  with the property that  $z\bar{z} = 9$ .

**Solution:** This is a retread of the previous question since  $z\bar{z} = a^2 + b^2 = 9$  is the equation of a circle with radius 3.



## 1.4 Polar Form

Now that we have a visual representation of complex numbers, we can ask new questions about them. For example, what is the *size* of a complex number. One way to define size is to take the distance between the number and the origin of the Argand diagram<sup>2</sup>.



Using Pythagoras' Theorem we get

$$r^2 = a^2 + b^2 \quad r = \sqrt{a^2 + b^2}$$

### Definition 1.25 Modulus

The *modulus* of the complex number  $z = a + ib$  is denoted  $|z| = |a + ib|$  and is defined by

$$|z| = |a + ib| = \sqrt{a^2 + b^2}$$

### Exercise 1.26

If  $z = 3 - i$  calculate  $|z|$ .

**Solution:** Since  $z = a + ib$  with  $a = 3$  and  $b = -1$ , we have

$$|z| = \sqrt{3^2 + (-1)^2} = \sqrt{10}$$

### Exercise 1.27

Find the modulus of the following complex numbers

1.  $z = i$
2.  $z = -9$ ,
3.  $z = \left| \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i \right|$

**Solution:** Applying the theorem, we have

1.  $|i| = \sqrt{0^2 + 1^2} = 1$
2.  $|-9| = \sqrt{9^2 + 0^2} = 9$

<sup>2</sup>This idea of using distance to origin as size is consistent with our knowledge of vectors

$$3. \left| \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i \right| = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1.$$

**Exercise 1.28**

In the Argand diagram, draw the set of all complex numbers  $z$  with modulus  $|z| = \sqrt{2}$ .

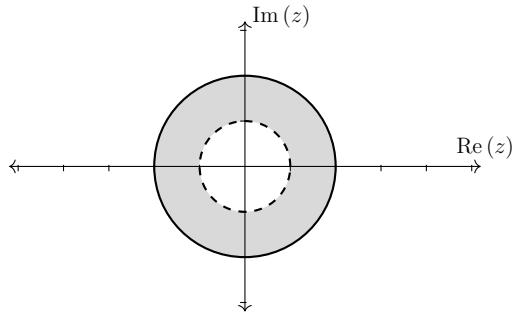
**Solution:** A complex number  $z = a + ib$  has modulus  $|z| = \sqrt{2}$  if  $\sqrt{a^2 + b^2} = \sqrt{2}$  or  $a^2 + b^2 = 2$  (See Exercise 1.23).

**Solution:** (Alternate) We know that the modulus of a complex number gives the distance to the origin. Thus we want all the points at distance  $\sqrt{2}$  from the origin. This again leads to the circle of radius  $\sqrt{2}$  (See Exercise 1.23).

**Exercise 1.29**

Plot the set of complex numbers  $z$  satisfying  $1 < |z| \leq 2$  in the complex plane.

**Solution:** Since the modulus is the distance to the origin the set consists of all those complex numbers which are at a distance between 1 and 2 from the origin. This gives the area shown below.



⚠️ Notice the difference between the inner and outer boundaries. A solid line indicates that the boundary is included and a dashed line indicates that the boundary is not included<sup>3</sup>.

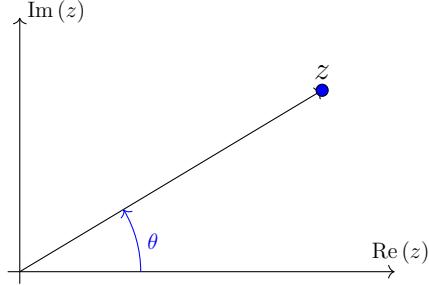
To uniquely define a complex number with the modulus, we would need another piece of information. This second piece of information is called the *argument*.

<sup>3</sup>This is analogous to using filled and hollowed circles to indicate endpoints of a real interval.

### Definition 1.30 (Principal) Argument

Let  $z$  be a complex number. The principal argument of  $z$ , denoted as  $\text{Arg}(z)$ , is the angle  $\theta$  in radians that  $z$  makes with the positive real axis in a counter-clockwise direction, chosen so that  $-\pi < \theta \leq \pi$ .

A point  $z$  in the Argand diagram determines an angle  $\theta$  with the positive real axis.

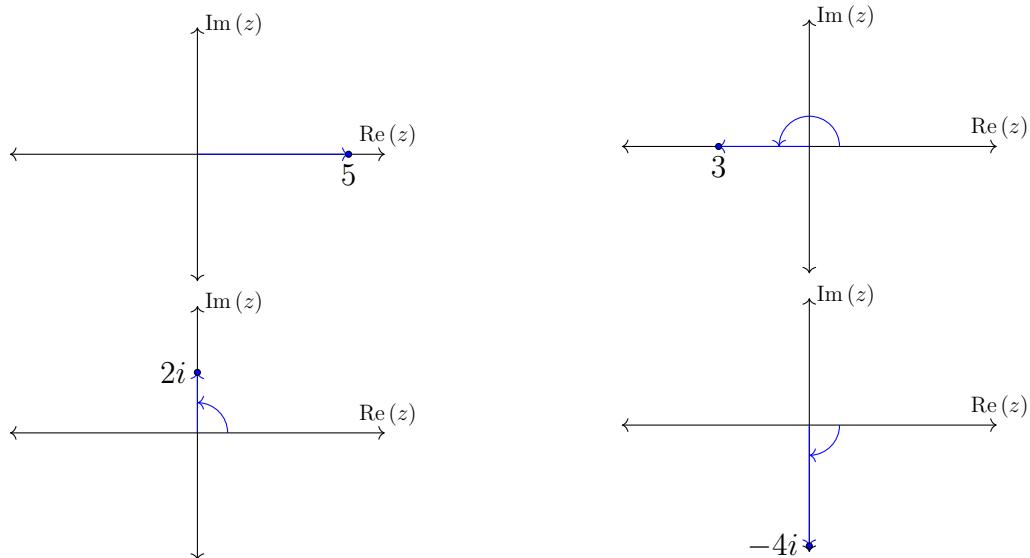


The choice of  $\theta$  such that  $-\pi < \theta \leq \pi$  means:

- $0 < \theta \leq \pi$  if  $z$  is above the real axis
- $-\pi < \theta < 0$  if  $z$  is below the real axis

### Example 1.31

The most straight forward method of determining the argument is to draw the diagram.



Based on the diagrams, we have:

$$\text{Arg}(5) = 0$$

$$\text{Arg}(2i) = \frac{\pi}{2}$$

$$\text{Arg}(-3) = \pi$$

$$\text{Arg}(-4i) = -\frac{\pi}{2}$$

A similar approach would show that

$$\operatorname{Arg}(1+i) = \frac{\pi}{4}$$

$$\operatorname{Arg}(1-i) = -\frac{\pi}{4}$$

$$\operatorname{Arg}(-1+i) = \frac{3\pi}{4}$$

$$\operatorname{Arg}(-1-i) = -\frac{3\pi}{4}$$

### Example 1.32

We can use our knowledge of coordinate axes to determine:

$$\operatorname{Arg}(1) = 0$$

$$\operatorname{Arg}(i) = \frac{\pi}{2}$$

$$\operatorname{Arg}(-1) = \pi$$

$$\operatorname{Arg}(-i) = -\frac{\pi}{2}$$

Further, we can use the special triangles to get

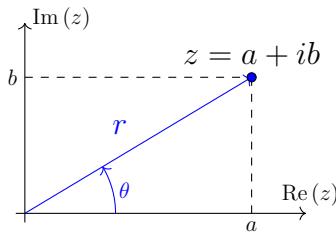
$$\operatorname{Arg}(\sqrt{3} + i) = \frac{\pi}{6}$$

$$\operatorname{Arg}(1 + i\sqrt{3}) = \frac{\pi}{3}$$

These are all cases we can obtain the angles by drawing and using our experience with simple triangles. We can extend this and develop a more systematic method to use for more general complex numbers

### Example 1.33

Suppose  $z = a + ib$  is a complex number which determines an angle  $\theta$  with the real axis.



We can use our knowledge of trigonometry to get that

$$\tan(\theta) = \frac{b}{a}$$

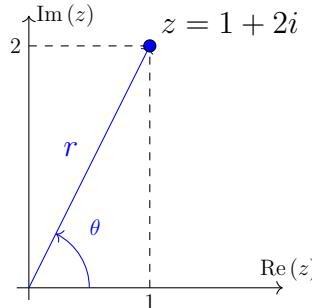
In other words, we can determine the argument  $\theta$  from the Cartesian form.

! In the case where  $a = 0$ , we are on the imaginary axis and can resolve the angle using more simple methods.

### Exercise 1.34

Find  $\operatorname{Arg}(1 + 2i)$ .

**Solution:** We start by plotting a diagram



Using the above formula we know that  $\tan(\theta) = \frac{b}{a} = 2$ . Using a calculator we get

$$\theta = \arctan(2) = 1.107$$

to 3 decimal places (in radians!).

⚠ Even though we have  $\tan(\theta) = \frac{b}{a}$ , we have cases where  $\theta \neq \arctan\left(\frac{b}{a}\right)$ .

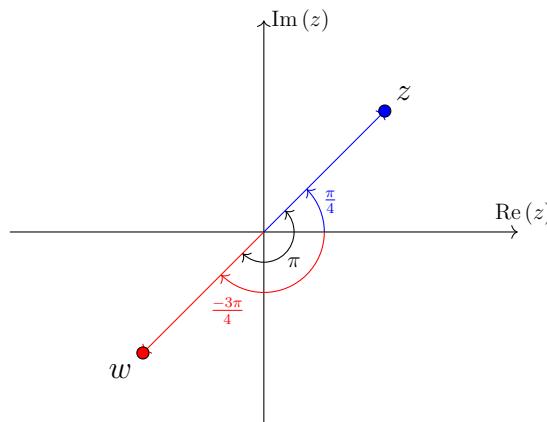
### Example 1.35

Suppose we want to find  $\operatorname{Arg}(-1 - 2i)$  using the same method. We will that  $\frac{b}{a} = \frac{-2}{-1} = 2$ , which will give  $\theta = \arctan(2) = 1.107$ . However this is clearly not correct based on our previous exercise. That is  $w = -1 - 2i$  and  $z = 1 + 2i$  are not in the same direction.

The reason for this is because the inverse tangent function produces values  $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ . In other words, it assumes  $a > 0$ . If we want to obtain the argument for the complex number in the exact opposite direction, we have to account for it by adding or subtracting  $\pi$ .

### Example 1.36

Consider the numbers  $z = 1 + i$  and  $w = -1 - i$ . We know  $\operatorname{Arg}(z) = \frac{\pi}{4}$  and  $\operatorname{Arg}(w) = -\frac{3\pi}{4}$  and  $\tan\left(\frac{\pi}{4}\right) = \tan\left(-\frac{3\pi}{4}\right) = 1$ . If we plot a diagram, we see why this is the case and how we can resolve it.



Be sure to have your calculator set to **radians** mode

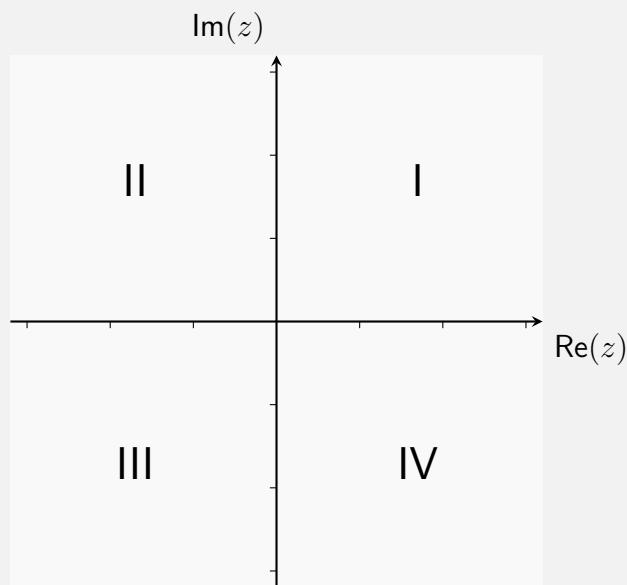
### Summary 1.37

To find  $\theta = \operatorname{Arg}(a + ib)$ ,

1. Draw a picture showing the angle. Stop if the angle is from a special triangle or on the axes.
2. Determine which *quadrant* the point is in and use the following:

$\theta = \arctan\left(\frac{b}{a}\right) + \pi$	$\theta = \arctan\left(\frac{b}{a}\right)$
$a < 0, b > 0$	$a > 0, b > 0$
$\theta = \arctan\left(\frac{b}{a}\right) - \pi$	$\theta = \arctan\left(\frac{b}{a}\right)$
$a < 0, b < 0$	$a > 0, b < 0$

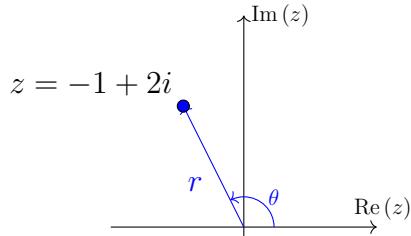
The four *quadrants* of the complex plane are usually numbered anti-clockwise.



**Exercise 1.38**

Find  $\operatorname{Arg}(-1 + 2i)$ .

**Solution:** Starting with the picture:



We see that  $0 < \operatorname{Arg}(-1 + 2i) < \pi$ . Using the above formula we know that  $\tan(\theta) = \frac{2}{-1} = -2$ . Using a calculator we get

$$\arctan(-2) = -1.107$$

to 3 decimal places. Since this is not in the correct range for  $\theta$ . Adding  $\pi$  gives

$$\theta = -1.107 + \pi = 2.034$$

**Exercise 1.39**

Find  $\operatorname{Arg}(3 - 4i)$ .

**Solution:** We see that  $-\frac{\pi}{2} < \operatorname{Arg}(3 - 4i) < 0$ , so the inverse tangent value should give us the right argument. This gives:

$$\operatorname{Arg}(3 - 4i) = \arctan\left(\frac{-4}{3}\right) = -0.927.$$

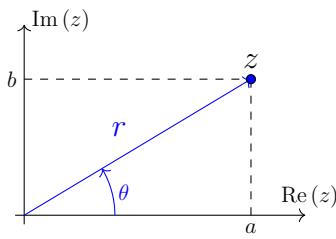
Now that we have we can determine the modulus and argument of a complex number, we can try to reverse the process.

**Definition 1.40 Polar form**

Let  $z$  be a complex number with modulus  $r$  and principal argument  $\theta$ . The polar form of  $z$  is given by

$$\begin{aligned} z &= r \cos(\theta) + ir \sin(\theta) \\ &= r(\cos(\theta) + i \sin(\theta)) \end{aligned}$$

We can verify this by looking at the picture and applying a bit of trigonometry.



From the picture, we see that the base  $a = r \cos(\theta)$  and height  $b = r \sin(\theta)$ . Hence, we have

$$z = r \cos(\theta) + i r \sin(\theta)$$

### Exercise 1.41

Suppose  $|z| = 2$  and  $\text{Arg}(z) = \frac{\pi}{3}$ . Write  $z$  in polar form.

**Solution:** Set  $r = |z| = 2$  and  $\theta = \text{Arg}(z) = \frac{\pi}{3}$ . Then

$$\begin{aligned} z &= r (\cos(\theta) + i \sin(\theta)) \\ &= 2 \left( \cos\left(\frac{\pi}{3}\right) + i \sin\left(\frac{\pi}{3}\right) \right) \end{aligned}$$

If we wished to convert it to Cartesian form, we can further simplify

$$\begin{aligned} z &= 2 \left( \cos\left(\frac{\pi}{3}\right) + i \sin\left(\frac{\pi}{3}\right) \right) \\ &= 2 \left( \frac{1}{2} + i \frac{\sqrt{3}}{2} \right) \\ &= 1 + i\sqrt{3} \end{aligned}$$

In general, if we want to write a complex number in polar form, we:

1. Compute the modulus,  $r = |z|$ .
2. Computer the principal argument,  $\theta = \text{Arg}(z)$ .
3. Write  $z$  as

$$z = r (\cos(\theta) + i \sin(\theta))$$

We will see in the next subsection that we can also write this as  $z = r e^{i\theta}$ .

### Exercise 1.42

Write  $z = 1 + i$  in polar form.

**Solution:** First we calculate the modulus and principal argument.

- Modulus:  $|z| = \sqrt{1^2 + 1^2} = \sqrt{2}$ .
- Principal Argument:  $\text{Arg}(z) = \arctan\left(\frac{1}{1}\right) = \frac{\pi}{4}$ . Since  $z$  is in the first quadrant, the argument does not need altering.

Thus the polar form of  $z = 1 + i$  is

$$z = \sqrt{2} \left( \cos\left(\frac{\pi}{4}\right) + i \sin\left(\frac{\pi}{4}\right) \right)$$

### Exercise 1.43

Write  $z = -1 + 2i$  in polar form.

**Solution:** First we calculate the modulus and principal argument.

- Modulus:  $|z| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$ .
- Principal Argument:  $\text{Arg}(z) = \arctan\left(\frac{2}{-1}\right) = -1.107$ . Since  $z$  is in the upper-left quadrant, we need to alter the argument (see Exercise 1.38) to get  $\text{Arg}(z) = 2.034$

Thus the polar form of  $z = -1 + 2i$  is

$$z = \sqrt{5} (\cos(2.034) + i \sin(2.034))$$

## 1.5 Exponential Form

The introduction of complex numbers connects two key areas of mathematics: Trigonometry and the exponential. It is possible to write a Maclaurin series for the exponential

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

As it turns out, the exponential can be extended to complex numbers in the analogous way.

### Definition 1.44 Complex exponential

For a complex number  $z$ , the complex exponential is defined to be:

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

It is not at all obvious, but it is a fact that the infinite summation results in something finite for all complex  $z$  so long as  $|z| < \infty$ .

If we consider the complex number  $z = i\theta$  for an angle  $\theta$ , we get:

$$e^{i\theta} = 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!}$$

By resolving the powers of  $i$ , we get

$$\begin{aligned} e^{i\theta} &= 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \cdots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} + \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \cdots\right) + i \left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \cdots\right) \end{aligned}$$

Recalling the Maclaurin series for  $\sin(x)$  and  $\cos(x)$ , we see that the two summations give:

$$e^{i\theta} = \cos(\theta) + i \sin(\theta)$$

This gives a fundamental result in complex numbers

### Theorem 1.45 Euler's Formula

For any angle  $\theta$ , we have

$$e^{i\theta} = \cos(\theta) + i \sin(\theta)$$

In particular, if  $\theta = \pi$ , we have the famous equation<sup>4</sup>

$$e^{i\pi} + 1 = 0$$

### Example 1.46

With Euler's formula, we have

$$\begin{aligned} e^{i1} &= \cos(1) + i \sin(1). \\ e^{i\frac{\pi}{2}} &= \cos\left(\frac{\pi}{2}\right) + i \sin\left(\frac{\pi}{2}\right) = 0 + 1i = i. \end{aligned}$$

Using Euler's formula we can get a new and very useful representation for complex numbers in polar form.

### Definition 1.47 Exponential form

The exponential form of the complex number  $z$  with modulus  $r$  and argument  $\theta$  is written as

$$z = re^{i\theta}$$

We derive this form by applying Euler's Formula to the polar form

$$z = r(\cos(\theta) + i \sin(\theta)) = re^{i\theta}$$

To write a complex number  $z$  in exponential form:

1. Calculate the modulus  $r = |z|$ ,
2. Calculate the principle argument  $\theta = \text{Arg}(z)$ ,
3. Write  $z = re^{i\theta}$ .

### Exercise 1.48

Write  $z = 1 + i$  in exponential form.

**Solution:** Building on Exercise 1.42, we know  $|z| = \sqrt{2}$  and  $\text{Arg}(z) = \frac{\pi}{4}$ . Hence, the exponential form is

$$z = 1 + i = \sqrt{2}e^{i\frac{\pi}{4}}.$$

### Exercise 1.49

◆ Write  $-i$  and  $-1$  in exponential form.

---

<sup>4</sup>This remarkable formula was obtained by Leonhard Euler in around 1768. The special case (one of the Top 10 equations ever!) links 5 fundamentally important numbers: 1,0,  $e$ ,  $\pi$  and  $i$  that have their origin in different branches of mathematics.

Now that we have shown how to convert a complex number from Cartesian form  $a + ib$  to exponential form  $re^{i\theta}$ , we can go in the reverse direction.

**Exercise 1.50**

Write the complex number  $z = 2e^{i\frac{\pi}{4}}$  in Cartesian form  $a + ib$ .

**Solution:** Using Euler's formula, we have

$$\begin{aligned} z &= 2e^{i\frac{\pi}{4}} \\ &= 2 \left( \cos\left(\frac{\pi}{4}\right) + i \sin\left(\frac{\pi}{4}\right) \right) \\ &= 2 \left( \frac{1}{\sqrt{2}} + i \frac{1}{\sqrt{2}} \right) \\ &= \sqrt{2} + i\sqrt{2} \end{aligned}$$

**Exercise 1.51**

Let  $z = 5 + 4i$ . Write  $e^z$  in the form  $a + ib$ .

**Solution:** We have (to 2 decimal point)

$$\begin{aligned} e^{5+4i} &= e^5 e^{4i} \\ &= e^5 (\cos(4) + i \sin(4)) \\ &= e^5 \cos(4) + ie^5 \sin(4) \\ &\approx -97.01 - 112.32i \end{aligned}$$

Recall that multiplication and division of complex numbers can be quite an involved process (See Example 1.13). However, we can utilise properties of the exponential function (and hence the exponential form) help simplify this process.

Let  $z$  and  $w$  be complex numbers. Then we have:

$$e^z \cdot e^w = e^{z+w} \quad \frac{e^z}{e^w} = e^{z-w}$$

We can express this as:

### Theorem 1.52 Multiplication and division in exponential form

Let  $z = re^{i\theta}$  and  $w = se^{i\phi}$  be complex numbers in exponential form

$$z \cdot w = (r \cdot s)e^{i(\theta+\phi)} \quad \frac{z}{w} = \frac{r}{s}e^{i(\theta-\phi)}.$$

Converting the right-hand side to *polar form* gives:

$$z \cdot w = (r \cdot s)(\cos(\theta + \phi) + i \sin(\theta + \phi))$$

$$\frac{z}{w} = \frac{r}{s}(\cos(\theta - \phi) + i \sin(\theta - \phi)).$$

### Exercise 1.53

Let  $z = 2e^{\frac{i\pi}{4}}$ , and  $w = e^{-\frac{i\pi}{4}}$ . Find  $z \cdot w$  and  $\frac{z}{w}$ .

**Solution:** Reading off the complex numbers, we get  $r = 2, s = 1, \theta = \frac{\pi}{4}, \phi = -\frac{\pi}{4}$ . So we get:

$$\begin{aligned} z \cdot w &= (2 \cdot 1) \left[ \cos\left(\frac{\pi}{4} + \left(-\frac{\pi}{4}\right)\right) + i \sin\left(\frac{\pi}{4} + \left(-\frac{\pi}{4}\right)\right) \right] \\ &= 2 [\cos(0) + i \sin(0)] = 2 \\ \frac{z}{w} &= \frac{2}{1} \left[ \cos\left(\frac{\pi}{4} - \left(-\frac{\pi}{4}\right)\right) + i \sin\left(\frac{\pi}{4} - \left(-\frac{\pi}{4}\right)\right) \right] \\ &= \frac{2}{1} [\cos(\pi/2) + i \sin(\pi/2)] = 2i. \end{aligned}$$

Complex exponentials are extremely useful when dealing with multiplication and division<sup>5</sup>. However, its usefulness extends beyond simple arithmetic. For the rest of this section, we will take a (very brief) look into areas where complex exponentials might be useful.

### Complex exponentials and trigonometric functions

Euler's formula gives us an alternative way to write the sine and cosine functions which is used very widely in science and engineering. In particular, we have

$$\begin{aligned} e^{i\theta} + e^{-i\theta} &= (\cos(\theta) + i \sin(\theta)) + (\cos(\theta) - i \sin(\theta)) = 2 \cos(\theta) \\ e^{i\theta} - e^{-i\theta} &= \cos(\theta) + i \sin(\theta) - (\cos(\theta) - i \sin(\theta)) = 2i \sin(\theta) \end{aligned}$$

We can rearrange to get the following:

<sup>5</sup>but not so effective when dealing with addition and subtraction.

**Theorem 1.54 Trigonometric functions in exponential forms**

We can express the trigonometric functions:

$$\cos(\theta) = \frac{1}{2} (e^{i\theta} + e^{-i\theta}),$$

$$\sin(\theta) = \frac{1}{2i} (e^{i\theta} - e^{-i\theta}).$$

Notice that these relations are very similar to the ones for the hyperbolic trigonometric functions ( $\cosh$  and  $\sinh$ ). The introduction of complex numbers illustrates why they share so many similar identities.

**Exercise 1.55**

◆ Use complex exponentials to show that  $2 \sin(\theta) \cos(\theta) = \sin(2\theta)$ .

**Calculus with complex exponentials**

The complex exponential is also commonly involved in the solution of differential equations, particularly in cases where solutions oscillate (like in AC electronics, quantum mechanics, etc.). It satisfies the same rules of differentiation and integration as any other exponential function<sup>6</sup>.

**Example 1.56**

For derivatives involving  $i$ , we have

$$\frac{d}{dx} e^{i\omega x} = i\omega e^{i\omega x}$$

$$\frac{d^2}{dx^2} e^{i\omega x} = (i\omega)^2 e^{i\omega x} = -\omega^2 e^{i\omega x}$$

Integration works in a similar way

$$\int e^{i\omega x} dx = \frac{e^{i\omega x}}{i\omega} + C$$

**Exercise 1.57**

◆ Verify that  $y(t) = Ce^{-i4t}$  is a solution of  $y'' = -16y$ .

**Logarithms of complex numbers**

Now that we have seen the exponential in complex numbers, it is natural to question how logarithm behave in complex numbers? The short answer is that logarithms are very complicated beasts,

<sup>6</sup>In essence, we can treat the complex number  $i$  like a constant.

with properties far beyond this introductory course. ⚠ Use with caution!

Suppose we “naively” apply logarithms on complex numbers using the rules of real numbers, we have the following.

### Example 1.58

Let  $z = re^{i\theta}$  be a complex number in exponential form. We can apply the logarithm as

$$\begin{aligned}\text{Log}(z) &= \text{Log}(re^{i\theta}) \\ &= \text{Log}(r) + \text{Log}(e^{i\theta}) \\ &= \text{Log}(r) + i\theta \\ &= \text{Log}|z| + i\text{Arg}(z)\end{aligned}$$

### Exercise 1.59

Let  $z = 42e^{i\frac{\pi}{7}}$ . Find  $\text{Log}(z)$ .

**Solution:** Since  $|z| = 42$  and  $\text{Arg}(z) = \frac{\pi}{7}$ , we have  $\text{Log}(z) = \text{Log}(42) + i\frac{\pi}{7}$ .

⚠ Whilst this approach is not wrong, it also does not capture the entire picture<sup>7</sup>.

<sup>7</sup>Recall from Definition 1.30 that we have  $-\pi < \text{Arg}(z) \leq \pi$ . However, this is a choice we made and there are other (non-principal) arguments we can take to be the angle of a complex number. For example, we could have chosen  $0 < \text{Arg}(z) \leq 2\pi$ , which would affect the resulting logarithm.

## 1.6 De Moivre's Theorem

Through repeated application of Theorem 1.52, we can extend exponential multiplication to powers of exponentials. We start with

$$e^{i\theta} \cdot e^{i\theta} = e^{i(\theta+\theta)} = e^{i2\theta}$$

which can be extended to any integer  $n$  giving

$$(e^{i\theta})^n = e^{in\theta}$$

If we convert both sides of the above equation into polar form, we obtain a very important result.

### Theorem 1.60 De Moivre's Theorem

Let  $n$  be an integer. For an angle  $\theta$ , we have

$$(\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta).$$

For a complex number  $z = re^{i\theta}$  in exponential form, the following are equivalent

$$\begin{aligned} z^n &= r^n e^{in\theta} = r^n (\cos(n\theta) + i \sin(n\theta)) \\ &= r^n (e^{i\theta})^n = r^n (\cos(\theta) + i \sin(\theta))^n \end{aligned}$$

De Moivre's Theorem is particularly useful when dealing with powers of complex numbers.

### Exercise 1.61

Let  $z = (1 + i)$  and compute  $z^8$

**Solution:** We first write  $z$  in exponential form as

$$z = 1 + i = \sqrt{2} \left( \cos\left(\frac{\pi}{4}\right) + i \sin\left(\frac{\pi}{4}\right) \right) = \sqrt{2} e^{i\frac{\pi}{4}}$$

This gives

$$\begin{aligned} z^8 &= (1 + i)^8 \\ &= \left(\sqrt{2} e^{i\frac{\pi}{4}}\right)^8 \\ &= \sqrt{2}^8 e^{i\frac{8\pi}{4}} \\ &= 16 e^{i2\pi} = 16 \end{aligned}$$

We can reach the same solution by multiplying 8 times. But this would be rather tedious and not recommended.

**Exercise 1.62**

Compute  $(-1 + 2i)^{10}$

**Solution:** Putting  $z = (-1 + 2i)$  in exponential form, we find

$$z = (-1 + 2i) = \sqrt{5}(\cos(2.034) + i\sin(2.034)) = \sqrt{5}e^{2.034i}$$

Thus we have

$$z^{10} = (-1 + 2i)^{10} = (\sqrt{5}e^{2.034i})^{10} = 3125e^{20.34i} = 250.8 + 3115i$$

De Moivre's theorem is also useful for computing roots of complex numbers.

**Theorem 1.63 Roots of complex numbers**

Let  $z = r e^{i\theta}$  be a complex number in exponential form and  $n$  be a positive integer. There are precisely  $n$  different  $n$ -th roots of  $z$  and they are given by

$$z^{\frac{1}{n}} = r^{\frac{1}{n}} e^{i\frac{\theta+2\pi k}{n}} \quad \text{for } k = 0, 1, \dots, n-1$$

where  $r^{\frac{1}{n}}$  is the positive  $n$ -th root of  $r$ .

**Example 1.64**

For *square roots* of  $z = re^{i\theta}$  we have  $n = 2$ . So the *two roots*  $w_1, w_2$  are given by

$$\begin{aligned} w_1 &= \sqrt{r} e^{i\frac{\theta+2\pi 0}{2}} = \sqrt{r} e^{i\frac{\theta}{2}} \\ w_2 &= \sqrt{r} e^{i\frac{\theta+2\pi 1}{2}} = \sqrt{r} e^{i\frac{\theta+2\pi}{2}} \end{aligned}$$

and  $\sqrt{r}$  is the positive square root of  $r$ .

For *cube roots* of  $z = re^{i\theta}$  we have  $n = 3$ . So the *three roots*  $w_1, w_2, w_3$  are given by

$$\begin{aligned} w_1 &= \sqrt[3]{r} e^{i\frac{\theta+2\pi 0}{3}} = \sqrt[3]{r} e^{i\frac{\theta}{3}} \\ w_2 &= \sqrt[3]{r} e^{i\frac{\theta+2\pi 1}{3}} = \sqrt[3]{r} e^{i\frac{\theta+2\pi}{3}} w_3 \\ &\qquad\qquad\qquad = \sqrt[3]{r} e^{i\frac{\theta+2\pi 2}{3}} = \sqrt[3]{r} e^{i\frac{\theta+4\pi}{3}} \end{aligned}$$

and  $\sqrt[3]{r}$  is the positive cubed root of  $r$ .

**Exercise 1.65**

Let  $z = 9e^{i\frac{\pi}{3}}$ . Find the two values of  $\sqrt{z}$  and verify your answer.

**Solution:** Apply Theorem 1.63 with  $r = 9$ ,  $\theta = \frac{\pi}{3}$  and  $n = 2$  gives

$$\begin{aligned} w_1 &= \sqrt{9}e^{i\frac{\frac{\pi}{3}+2\pi}{2}} \\ &= 3e^{i\frac{\pi}{6}} \\ w_2 &= \sqrt{9}e^{i\frac{\frac{\pi}{3}+2\pi}{2}} \\ &= 3e^{i\frac{7\pi}{6}} \end{aligned}$$

We can verify the solutions by squaring the answer

$$\begin{aligned} (w_1)^2 &= \left(3e^{i\frac{\pi}{6}}\right)^2 \\ &= 9e^{2i\frac{\pi}{6}} \\ &= 9e^{i\frac{\pi}{3}} \\ &= z \\ (w_2)^2 &= \left(3e^{i\frac{7\pi}{6}}\right)^2 \\ &= 9e^{2i\frac{7\pi}{6}} \\ &= 9e^{i\frac{14\pi}{6}} \\ &= 9e^{i\frac{2\pi}{6}}e^{i2\pi} \\ &= z \cdot 1 = z \end{aligned}$$

### Exercise 1.66

- ◆ Find the 5-th roots of  $z = -1 + 2i$ . Confirm your answers by plotting it on the complex plane.

**Example 1.67**

Plotting the  $n$ -th roots of a complex number  $z$  gives some interesting patterns.

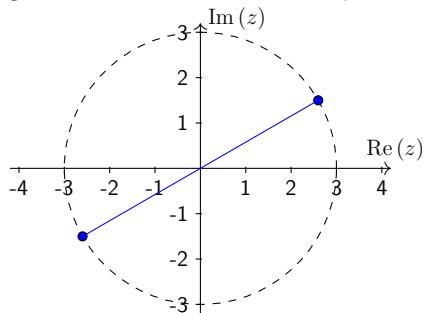


Figure 1.1: Square roots of  $z = 9e^{i\frac{\pi}{3}}$

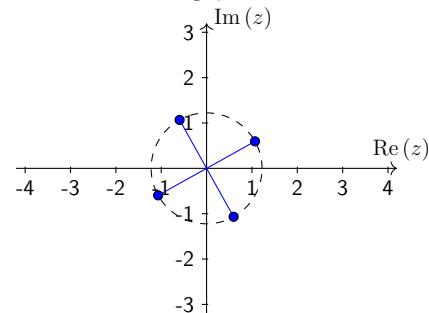


Figure 1.3: 4-th roots of  $z = -1 + 2i$

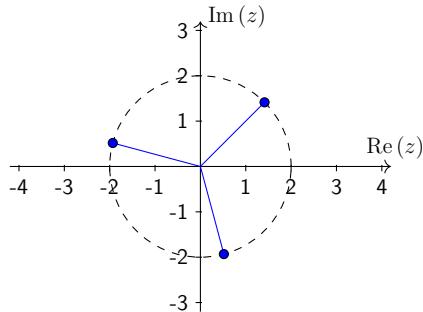


Figure 1.2: Third roots of  $z = \frac{8(-1+i)}{\sqrt{2}}$

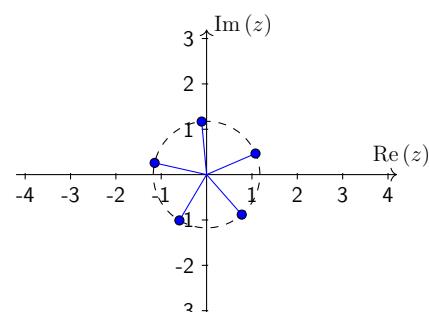


Figure 1.4: 5-th roots of  $z = -1 + 2i$

The  $n$ -th roots of  $z = re^{i\theta}$  are spaced equally in wedges with angle  $\frac{2\pi}{n}$  around a circle of radius  $r^{\frac{1}{n}}$ .

As a final application of De Moivre's theorem, we can use it to derive some useful trigonometric identities.

**Example 1.68**

Consider the case of  $n = 2$  in De Moivre's Theorem. We get

$$\begin{aligned} \cos(2\theta) + i \sin(2\theta) &= (\cos(\theta) + i \sin(\theta))^2 \\ &= \cos^2(\theta) - \sin^2(\theta) + i2 \sin(\theta) \cos(\theta) \end{aligned} \quad \text{Expanding}$$

In particular, if we equate the real and imaginary parts of the equation, we recover a couple of well-known identities

$$\begin{aligned} \cos(2\theta) &= \cos^2(\theta) - \sin^2(\theta) \\ \sin(2\theta) &= 2 \sin(\theta) \cos(\theta) \end{aligned}$$

The theorem gives us two trigonometric identities at the same time.

**Exercise 1.69**

◆ Prove the identity  $\sin(3\theta) = 3 \cos^2(\theta) \sin(\theta) - \sin^3(\theta)$  using De Moivre's Theorem.

## 1.7 Complex Arguments in Trigonometric/Hyperbolic Functions

We have seen that for any real angle  $\theta$ , *Euler's formula* gives

$$e^{i\theta} = \cos(\theta) + i \sin(\theta)$$

which can be used to write  $\sin(\theta), \cos(\theta)$  as

$$\begin{aligned}\cos(\theta) &= \frac{1}{2} (e^{i\theta} + e^{-i\theta}), \\ \sin(\theta) &= \frac{1}{2i} (e^{i\theta} - e^{-i\theta}).\end{aligned}$$

This can, in turn be used to write

$$\begin{aligned}\cos(z) &= \frac{1}{2} (e^{iz} + e^{-iz}), \\ \sin(z) &= \frac{1}{2i} (e^{iz} - e^{-iz}).\end{aligned}$$

for *complex arguments*  $z$ .

If we expand the RHS using  $z = x + iy$ , we find

$$\cos(z) = \cos(x) \cosh(y) - i \sin(x) \sinh(y)$$

and

$$\sin(z) = \sin(x) \cosh(y) + i \cos(x) \sinh(y).$$

This is also true for  $\cosh(z), \sinh(z)$  with complex arguments  $z$ , i.e.

$$\begin{aligned}\cosh(z) &= \frac{1}{2} (e^z + e^{-z}), \\ \sinh(z) &= \frac{1}{2} (e^z - e^{-z}).\end{aligned}$$

with *complex arguments*  $z = x + iy$ .

### Example 1.70

Expand  $\cosh(z)$  for  $z = x + iy$ .

Since

$$\begin{aligned}\cosh(z) &= \frac{1}{2} (e^z + e^{-z}) = \frac{1}{2} (e^{x+iy} + e^{-x+iy}) \\ &= \frac{1}{2} (e^x (\cos(y) + i \sin(y)) + e^{-x} (\cos(y) - i \sin(y))) \\ &= \frac{1}{2} ((e^x + e^{-x}) \cos(y) + i (e^x - e^{-x}) \sin(y)) \\ &= \cosh(x) \cos(y) + i \sinh(x) \sin(y)\end{aligned}$$

The exponential function can be defined for complex arguments

$$\exp(z) = \exp(x + iy) = \exp(x) \exp(iy) = \exp(x) (\cos(y) + i \sin(y))$$

which shows that

$$\operatorname{Re} \exp(z) = \exp(x) \cos(y), \quad \operatorname{Im} \exp(z) = \exp(x) \sin(y).$$

The Logarithm can also be defined for complex arguments

$$\operatorname{Log}(z) = \operatorname{Log}|z| + i\operatorname{Arg}(z)$$

which can be written as

$$\operatorname{Log}(z) = \frac{1}{2}\operatorname{Log}(x^2 + y^2) + i \arctan\left(\frac{y}{x}\right).$$

You might worry that  $\operatorname{Arg}()$ , (i.e.  $\arctan$ ), might lead to complications in the definition of  $\operatorname{Log}$  for complex arguments - it does! Disentangling these is for another course.

## 1.8 Complex Mappings

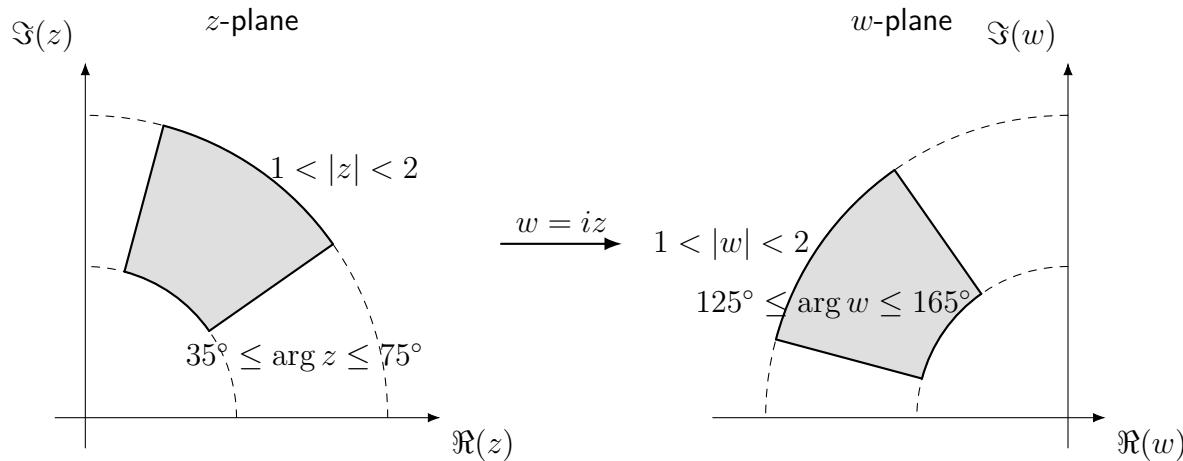
A nice geometrical way to picture complex functions is to think of them as *Complex Mappings*. A complex function  $w = u + iv = f(z)$  takes the point  $z = x + iy$  in the complex plane and maps it to  $u + iv$ .

### Example 1.71

Show the effect of considering  $w = iz$  as a complex map that acts on the segment  $1 \leq |z| \leq 2$ ,  $35^\circ \leq \arg z \leq 75^\circ$  in the complex  $z$  plane.

If we use polar form,  $i = e^{i\frac{\pi}{2}}$ , so it represents a (positive, anticlockwise) rotation by  $\pi/2$  radians.

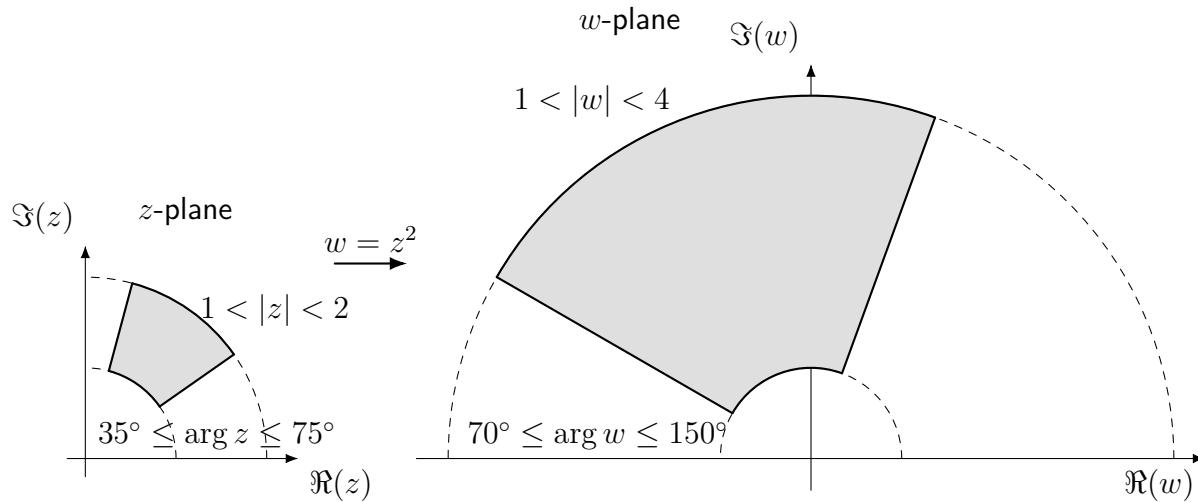
I have sneakily used an alternative notation in the figures for  $\operatorname{Re}(z)$ , namely  $\Re(z)$ , and  $\operatorname{Im}(z)$ , namely  $\Im(z)$ .



**Example 1.72**

Show the effect of considering  $w = z^2$  as a complex map that acts on the segment  $1 \leq |z| \leq 2$ ,  $35^\circ \leq \arg z \leq 75^\circ$  in the complex  $z$  plane.

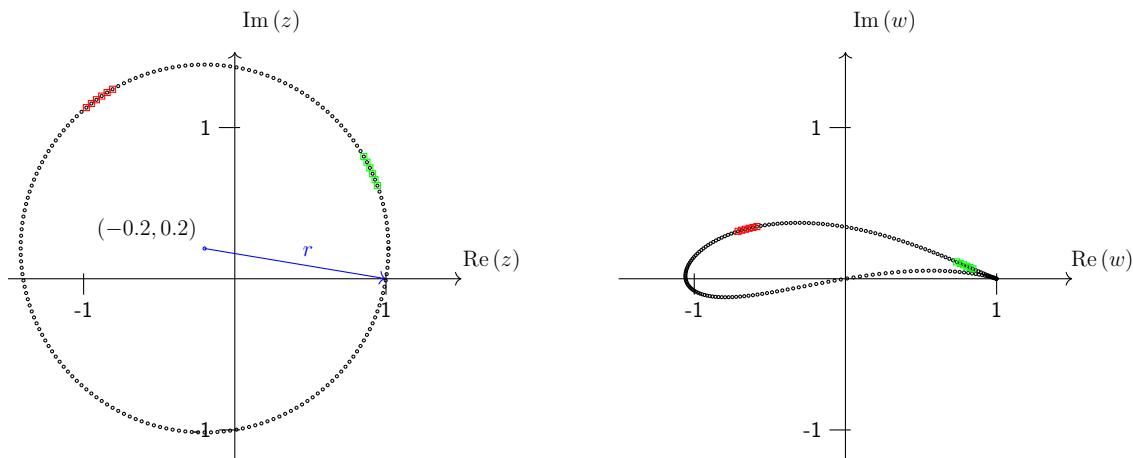
Again, polar form is useful. If  $z = re^{i\theta}$ , then  $w = r^2 e^{2i\theta}$ .

**Example 1.73**

We can use a complex function to convert a circle (with certain properties) to an aerofoil shape. Each dot on the left plot (the  $z$  values) is connected to one on the right (the  $w$  values) by the formula

$$w(z) = \frac{1}{2} \left( z + \frac{1}{z} \right).$$

The highlighted circles and diamonds show where their  $z$  values end up on the  $w$  value plot. Note that the circle in the  $z$  plane is slightly off-centre.



$w(z)$  in this example is a *complex* function. The argument  $z$  is a complex number and the function returns a complex number  $w(z)$ .

Let us pick apart  $w(z)$  in Cartesian form.

$$\begin{aligned} w(z) &= \frac{1}{2} \left( z + \frac{1}{z} \right) = \frac{1}{2} \left( x + iy + \frac{1}{x+iy} \right) \\ &= \frac{1}{2} \left( x + iy + \frac{x-iy}{x^2+y^2} \right) \\ &= \frac{1}{2} \left( x + \frac{x}{x^2+y^2} + i \left( y - \frac{y}{x^2+y^2} \right) \right) \end{aligned}$$

We can see that the *complex* function  $w(z)$  may be written using two *real*-valued functions of  $x, y$

$$w(z) = u(x, y) + iv(x, y)$$

where

$$u(x, y) = x + \frac{x}{x^2+y^2}, \quad v(x, y) = y - \frac{y}{x^2+y^2}$$

This is true in general.

### Example 1.74

Let the initial circle now be centred on the origin with radius  $R$ , i.e.  $|z| = R$  with  $R > 0$ . What is  $w(z)$ ?

Parametrize it by

$$z = Re^{i\theta}, \quad 0 \leq \theta < 2\pi.$$

Then

$$\begin{aligned} w(z) &= \frac{1}{2} \left( z + \frac{1}{z} \right) = \frac{1}{2} \left( Re^{i\theta} + \frac{1}{R} e^{-i\theta} \right) \\ &= \frac{1}{2} \left( (R + \frac{1}{R}) \cos \theta + i(R - \frac{1}{R}) \sin \theta \right). \end{aligned}$$

Writing  $w = u + iv$ , we have

$$u = \frac{R + \frac{1}{R}}{2} \cos \theta, \quad v = \frac{R - \frac{1}{R}}{2} \sin \theta.$$

Eliminating  $\theta$  gives

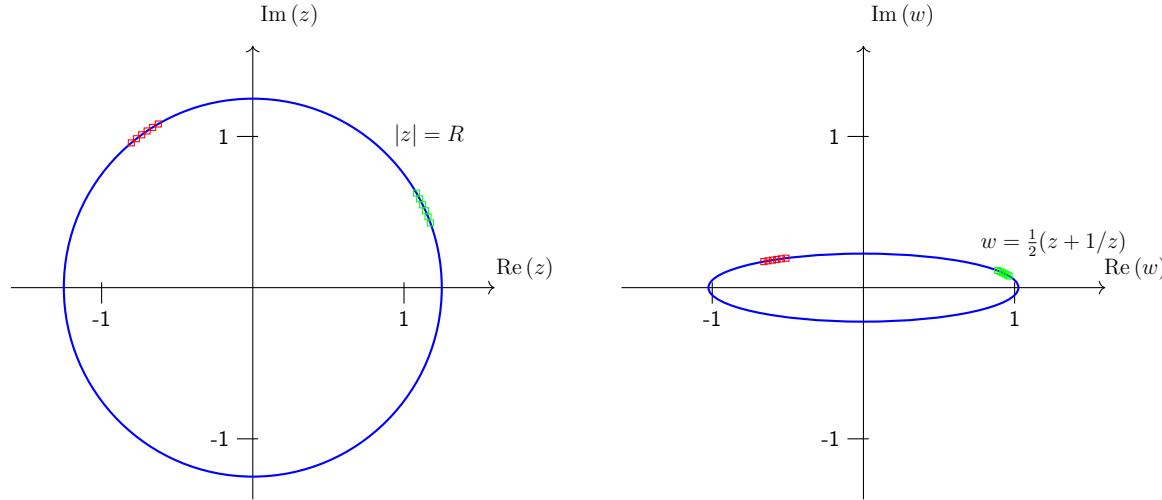
$$\frac{u^2}{\left(\frac{R+\frac{1}{R}}{2}\right)^2} + \frac{v^2}{\left(\frac{|R-\frac{1}{R}|}{2}\right)^2} = 1.$$

Hence the image of the circle  $|z| = R$  under  $w(z) = \frac{1}{2}(z + 1/z)$  is an ellipse centered at the origin, with semiaxes

$$a = \frac{R + \frac{1}{R}}{2} \quad (\text{along the real axis}), \quad b = \frac{|R - \frac{1}{R}|}{2} \quad (\text{along the imaginary axis}).$$

In the special case  $R = 1$ , the ellipse degenerates to the line segment.

$$w = \cos \theta \in [-1, 1].$$



### Exercise 1.75

Consider the *quadratic map*,  $f(z) = z^2 + c$  where  $c = a + ib$  (with  $a, b$  real) is a complex constant.

What is  $f(z)$  when written in terms of real functions  $g, h$ . i.e.  $f(z) = f(x + iy) = g(x, y) + ih(x, y)$ ?

**Solution:** Substitute  $z = x + iy$  into  $f(z) = z^2 + c$  and gather the real and imaginary parts:

$$\begin{aligned} f(z) &= f(z + iy) = (x + iy)^2 + a + ib \\ &= (x^2 - y^2 + a) + i(2xy + b) \end{aligned}$$

hence  $g(x, y) = x^2 - y^2 + a$  and  $h(x, y) = 2xy + b$

## 1.9 Iterating complex maps

One of the most striking uses of complex numbers is in *complex dynamics*, where we repeatedly apply a map

$$z_{n+1} = f(z_n), \quad n = 0, 1, 2, \dots$$

starting from some initial value  $z_0 \in \mathbb{C}$ . The resulting sequence  $\{z_n\}$  is called the *orbit* of  $z_0$ . Even very simple choices of  $f$  can produce complicated behaviour: orbits may converge to a fixed point, fall into a periodic cycle, or escape to infinity.

We have just encountered the *quadratic map*

$$f_c(z) = z^2 + c,$$

where  $c \in \mathbf{C}$  is a complex parameter. A useful practical fact is the *escape radius*: if at some step  $|z_n| > 2$  (when  $|c| < 2$ ), then the orbit will diverge to infinity (so we do not need to keep iterating).

### Example 1.76

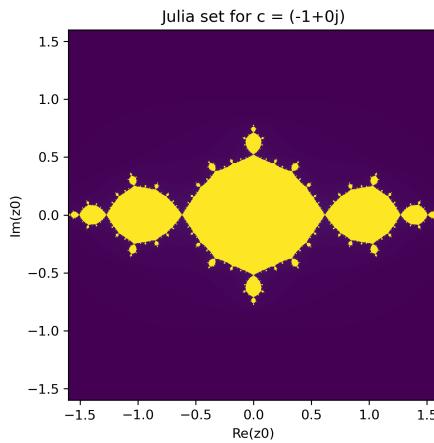
If we consider a few iterates in the complex plane: Take  $c = -0.123 + 0.745i$  and start at  $z_0 = 0$ . The first few iterates are

$$z_1 = c, \quad z_2 \approx -0.6629 + 0.5617i, \quad z_3 \approx 8.9 \times 10^{-4} + 2.6 \times 10^{-4}i,$$

and then the orbit returns close to  $z_1$  and  $z_2$  again.

### Example 1.77

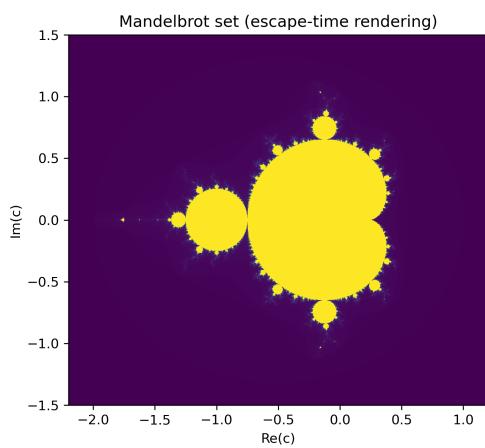
**Julia sets (bounded starting points):** Fix  $c \in \mathbf{C}$ . The *filled Julia set*  $K_c$  is the set of starting values  $z_0$  whose orbits under  $z_{n+1} = z_n^2 + c$  stay bounded. (The *Julia set*  $J_c$  is the boundary of  $K_c$ .) The picture below is a coarse “escape-time” plot for  $c = -1$ : points  $z_0$  that do *not* escape after a fixed number of iterations (200) are drawn as yellow dots.



### Example 1.78

The *Mandelbrot set* (bounded parameters):

Instead of fixing  $c$  and varying  $z_0$ , we can fix the starting point  $z_0 = 0$  and vary the parameter  $c$ . The *Mandelbrot set*  $M$  is the set of parameters  $c$  for which the orbit of 0 under  $z_{n+1} = z_n^2 + c$  remains bounded. The plot below is again an escape-time picture (very low resolution), now in the  $c$ -plane.



## 2 An aside: Quaternions

Complex numbers extend the real numbers by adjoining a new unit  $i$  with  $i^2 = -1$ . Quaternions extend the complex numbers further by adjoining *three* imaginary units.

Quaternions were introduced by the Irish mathematician William Rowan Hamilton in 1843 as a way to extend complex numbers to describe rotations in three dimensions. After years of trying (and failing) to build a consistent “three-dimensional complex arithmetic”, Hamilton realised that the key was to move to four components and to accept that multiplication need not commute. On 16th October 1843, while walking in Dublin, he famously carved the fundamental relations

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1.$$

into the stone of Brougham Bridge.



Quaternions quickly attracted attention in 19th-century mathematical physics and geometry, but they were later overshadowed in many applications by the vector calculus of Gibbs and Heaviside. In the late 20th century they saw a major revival in engineering and computing—especially in robotics, aerospace, and computer graphics—because unit quaternions provide a numerically stable and efficient way to represent 3D rotations without the singularities that can occur with Euler angles, which appear when using vectors.

### 2.1 Definitions and Arithmetic

#### Definition 2.1 Quaternion

A *quaternion* is an expression of the form

$$q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

where  $a, b, c, d \in \mathbb{R}$  and the symbols  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  satisfy

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1.$$

The set of all quaternions is denoted  $\mathbb{H}$ .

From the defining relations one can derive the multiplication rules

.	i	j	k
i	-1	k	-j
j	-k	-1	i
k	j	-i	-1

⚠ Quaternion multiplication is *not commutative*. For example,

$$ij = k \quad \text{but} \quad ji = -k.$$

### Definition 2.2 Scalar and vector parts

Write a quaternion as

$$q = a + \vec{v}, \quad \text{where} \quad \vec{v} = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}.$$

We call  $a$  the *scalar part* of  $q$  and  $\vec{v}$  the *vector part*.

Addition and subtraction are componentwise (exactly as for complex numbers):

$$(a, b, c, d) \pm (e, f, g, h) = (a \pm e, b \pm f, c \pm g, d \pm h).$$

Multiplication is defined using distributivity together with the table above.

### Example 2.3

Compute  $(1 + \mathbf{i})(1 + \mathbf{j})$ .

$$\begin{aligned} (1 + \mathbf{i})(1 + \mathbf{j}) &= 1 + \mathbf{j} + \mathbf{i} + \mathbf{ij} \\ &= 1 + \mathbf{i} + \mathbf{j} + \mathbf{k}. \end{aligned}$$

If we swap the factors,

$$\begin{aligned} (1 + \mathbf{j})(1 + \mathbf{i}) &= 1 + \mathbf{i} + \mathbf{j} + \mathbf{ji} \\ &= 1 + \mathbf{i} + \mathbf{j} - \mathbf{k}, \end{aligned}$$

so the order matters.

## 2.2 Conjugate, Norm and Inverse

Quaternions have an analogue of complex conjugation.

**Definition 2.4 Quaternion conjugate**

For  $q = a + bi + cj + dk$ , the *conjugate* is

$$\bar{q} = a - bi - cj - dk.$$

**Definition 2.5 Norm**

The *norm* of  $q$  is defined by

$$\|q\| = \sqrt{q\bar{q}} = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

**Example 2.6**

Let  $q = 2 - i + 2j + k$ . Then

$$\begin{aligned}\bar{q} &= 2 + i - 2j - k, \\ \|q\| &= \sqrt{2^2 + (-1)^2 + 2^2 + 1^2} = \sqrt{10}.\end{aligned}$$

Just like complex numbers, the norm is useful for division.

**Theorem 2.7 Inverse of a nonzero quaternion**

If  $q \neq 0$ , then  $q$  has a multiplicative inverse

$$q^{-1} = \frac{\bar{q}}{\|q\|^2}.$$

**Example 2.8**

Find the inverse of  $q = 1 + i + j + k$ .

$$\begin{aligned}\bar{q} &= 1 - i - j - k, \\ \|q\|^2 &= 1^2 + 1^2 + 1^2 + 1^2 = 4, \\ q^{-1} &= \frac{1}{4}(1 - i - j - k).\end{aligned}$$

⚠ Because multiplication is not commutative, one must distinguish between left and right division in more advanced settings. In this course we will only divide by placing the inverse on the *right*:  $p/q := p q^{-1}$ .

## 2.3 Unit Quaternions and 3D Rotations

A key application of quaternions is the description of rotations in  $\mathbb{R}^3$ .

**Definition 2.9 Pure imaginary quaternion**

A quaternion with zero scalar part,

$$\vec{v} = b\mathbf{i} + c\mathbf{j} + d\mathbf{k},$$

is called *pure imaginary*. We identify it with the vector  $(b, c, d) \in \mathbb{R}^3$ .

**Definition 2.10 Unit quaternion**

A quaternion  $q$  with  $\|q\| = 1$  is called a *unit quaternion*.

Any rotation by angle  $\theta$  about a unit axis  $\hat{n} \in \mathbb{R}^3$  can be encoded by the unit quaternion

$$q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} (n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}), \quad \hat{n} = (n_1, n_2, n_3), \quad \|\hat{n}\| = 1.$$

Given a vector  $\vec{v} \in \mathbb{R}^3$  (viewed as a pure imaginary quaternion), the rotated vector is

$$\vec{v}' = q \vec{v} q^{-1}.$$

**Example 2.11**

Rotate  $(1, 0, 0)$  by  $90^\circ$  about the  $z$ -axis.

The axis is  $\hat{n} = (0, 0, 1)$  and  $\theta = \pi/2$ , so

$$q = \cos \frac{\pi}{4} + \sin \frac{\pi}{4} \mathbf{k} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \mathbf{k}, \quad q^{-1} = \bar{q} = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} \mathbf{k}.$$

Let  $\vec{v} = \mathbf{i}$  (since  $(1, 0, 0)$  corresponds to  $\mathbf{i}$ ). Then

$$\begin{aligned} \vec{v}' &= q \mathbf{i} q^{-1} = \left( \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \mathbf{k} \right) \mathbf{i} \left( \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} \mathbf{k} \right) \\ &= \left( \frac{\sqrt{2}}{2} \mathbf{i} + \frac{\sqrt{2}}{2} \mathbf{k} \mathbf{i} \right) \left( \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} \mathbf{k} \right) = \left( \frac{\sqrt{2}}{2} \mathbf{i} + \frac{\sqrt{2}}{2} \mathbf{j} \right) \left( \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} \mathbf{k} \right) \\ &= \mathbf{j}. \end{aligned}$$

Thus  $(1, 0, 0)$  rotates to  $(0, 1, 0)$ , as expected.

**Exercise 2.12**

Use the multiplication table to compute  $\mathbf{i}\mathbf{k}$  and  $\mathbf{k}\mathbf{i}$ . What do you notice?

**Solution:** From the table,  $\mathbf{i}\mathbf{k} = -\mathbf{j}$  and  $\mathbf{k}\mathbf{i} = \mathbf{j}$ . They differ by a minus sign, illustrating non-commutativity.

**Exercise 2.13**

Let  $q = 3 - 2\mathbf{i} + \mathbf{j}$ . Compute  $\bar{q}$ ,  $\|q\|$ , and  $q^{-1}$ .

**Solution:** We have  $\bar{q} = 3 + 2\mathbf{i} - \mathbf{j}$  and

$$\|q\| = \sqrt{3^2 + (-2)^2 + 1^2 + 0^2} = \sqrt{14}.$$

Hence

$$q^{-1} = \frac{\bar{q}}{\|q\|^2} = \frac{1}{14} (3 + 2\mathbf{i} - \mathbf{j}).$$

**Exercise 2.14**

Let  $\hat{n} = (1, 0, 0)$  and  $\theta = \pi$ . Write down the unit quaternion  $q$  describing rotation by  $\pi$  about the  $x$ -axis, and compute  $q\mathbf{j}q^{-1}$ .

**Solution:** Here  $q = \cos(\pi/2) + \sin(\pi/2)\mathbf{i} = \mathbf{i}$  and  $q^{-1} = \bar{q} = -\mathbf{i}$ . Then

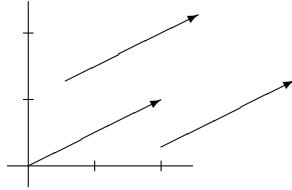
$$q\mathbf{j}q^{-1} = \mathbf{i}\mathbf{j}(-\mathbf{i}) = -(\mathbf{i}\mathbf{j})\mathbf{i} = -\mathbf{k}\mathbf{i} = -\mathbf{j}.$$

So the  $y$ -axis is sent to  $-y$ , as expected for a  $180^\circ$  rotation about the  $x$ -axis.

### 3 Vectors, Spans and Bases in $\mathbb{R}^2$ , $\mathbb{R}^3$ and $\mathbb{R}^n$ .

#### 3.1 Vectors in $\mathbb{R}^2$

We briefly recall some facts about vectors in  $\mathbb{R}^2$ . Intuitively, a *vector* of  $\mathbb{R}^2$  can be thought as an arrow encoding a length and a direction. A vector however does not have an origin, so one may draw many arrows in the plane corresponding to the same vector:

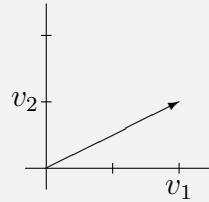


One may identify the set of all vectors in  $\mathbb{R}^2$  with the set of points of the plane  $\mathbb{R}^2$ , by associating to each point of the plane the vector from the origin to that point. Under this identification, the origin corresponds to a vector called the *null vector* and denoted  $\vec{0}$ .

##### Definition 3.1

Each point and thus each vector  $\vec{v} \in \mathbb{R}^2$  can be denoted by a pair of real numbers (thus the notation  $\mathbb{R}^2$ ): the horizontal displacement and the vertical displacement. Such a vector is denoted by the  $2 \times 1$  matrix

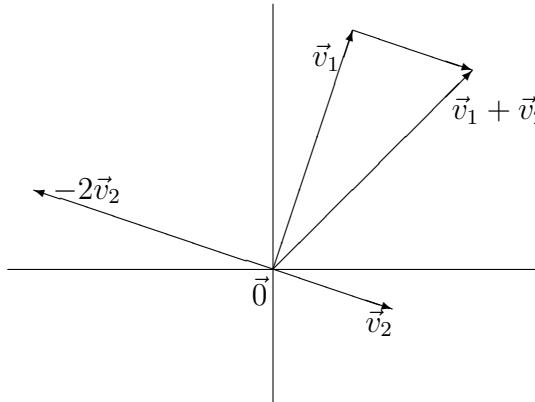
$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$



It is a very useful convention to write vectors vertically, i.e. as  $2 \times 1$  matrices. For typesetting reasons however, we often write that vector as  $\vec{v} = (v_1, v_2)^T$ , i.e. as a  $1 \times 2$  matrix instead.

We have the following rules for adding two vectors  $\vec{v} = (v_1, v_2)^T$  and  $\vec{w} = (w_1, w_2)^T$  and for multiplying a vector by a number  $\lambda \in \mathbb{R}$  (called a *scalar*):

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} := \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} \quad \text{and} \quad \lambda \vec{v} := \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \end{pmatrix}.$$



With our convention, the null vector is the vector with coordinates  $(0, 0)^T$ . Note that if we multiply a vector by 0, we obtain the null vector:  $0\vec{v} = \vec{0}$  for all  $\vec{v} \in \mathbb{R}^2$ .

## 3.2 Vectors in $\mathbb{R}^3$

### Definition 3.2

Most of the notions introduced on  $\mathbb{R}^2$  straightforwardly extend to  $\mathbb{R}^3$ . A point in the Euclidean dimensions encodes a vector  $\vec{v} \in \mathbb{R}^3$ , i.e. an arrow from the origin  $\vec{0}$  to that point. We describe the vector  $\vec{v}$  by a  $3 \times 1$  matrix

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix},$$

or sometimes as a  $1 \times 3$  matrix  $\vec{v} = (v_1, v_2, v_3)^T$  for typesetting reasons.

We can add vectors and multiply them by scalars as before:

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix} \quad \text{and} \quad \lambda\vec{v} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \lambda v_3 \end{pmatrix}.$$

## 3.3 Vectors in $\mathbb{R}^n$

The generalisation to  $\mathbb{R}^n$  should now be clear.

### Definition 3.3

We define:

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in \mathbb{R}, \quad i = 1, \dots, n \right\}.$$

An element of  $\mathbb{R}^n$  is called a *vector* (of  $\mathbb{R}^n$ ), and is identified with an  $n \times 1$  matrix.

**Definition 3.4**

For vectors  $\vec{x}, \vec{y} \in \mathbb{R}^n$  and a scalar  $\lambda \in \mathbb{R}$ , we define the following operations on vectors:

- addition:  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix},$
- multiplication by a scalar:  $\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$

From a given family of vectors, one can construct new vectors using additions and multiplication by scalars.

**Definition 3.5**

Let  $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$ . A *linear combination* of  $\vec{v}_1, \dots, \vec{v}_k$  is a vector of the form

$$\lambda_1 \vec{v}_1 + \dots + \lambda_k \vec{v}_k \quad \text{for some } \lambda_1, \dots, \lambda_k \in \mathbb{R}.$$

**Example 3.6**

The vector  $\begin{pmatrix} 1 \\ 3 \end{pmatrix} \in \mathbb{R}^2$  is a linear combination of  $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$  and  $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ , since

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

**Exercise 3.7**

Let

$$\vec{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

- Compute  $\vec{v} + \vec{w}$ ,  $\vec{v} - \vec{w}$  and  $3\vec{v} - 2\vec{w}$ .
- Find all  $\lambda \in \mathbb{R}$  such that  $\vec{v} + \lambda \vec{w} = \vec{0}$  (if any).
- Are  $\vec{v}$  and  $\vec{w}$  scalar multiples of one another?

**Solution:**

a)

$$\vec{v} + \vec{w} = \begin{pmatrix} 1 - 1 \\ -1 + 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \vec{v} - \vec{w} = \begin{pmatrix} 1 + 1 \\ -1 - 3 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \end{pmatrix}.$$

Also

$$3\vec{v} - 2\vec{w} = 3 \begin{pmatrix} 1 \\ -1 \end{pmatrix} - 2 \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \end{pmatrix} - \begin{pmatrix} -2 \\ 6 \end{pmatrix} = \begin{pmatrix} 5 \\ -9 \end{pmatrix}.$$

b) Solve  $\vec{v} + \lambda\vec{w} = \vec{0}$ :

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{cases} 1 - \lambda = 0, \\ -1 + 3\lambda = 0. \end{cases}$$

The first gives  $\lambda = 1$  while the second gives  $\lambda = \frac{1}{3}$ , which is impossible. Hence there is no  $\lambda$  with  $\vec{v} + \lambda\vec{w} = \vec{0}$ .

c) If  $\vec{v} = \mu\vec{w}$ , then  $1 = -\mu$  and  $-1 = 3\mu$ , giving  $\mu = -1$  and  $\mu = -\frac{1}{3}$ , a contradiction. So they are not scalar multiples. This is just rephrasing (b) with  $\mu = -\lambda$ .

Writing a vector as a linear combination of other vectors can be thought of as ‘decomposing’ that vector. Given a family of vectors  $\vec{v}_1, \dots, \vec{v}_k$  of  $\mathbb{R}^n$ , we will study the following questions:

- Can every vector of  $\mathbb{R}^n$  be written as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ ?
- If not, which vectors of  $\mathbb{R}^n$  can be written as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ ?
- In how many ways can a vector of  $\mathbb{R}^n$  be written as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ ?

**The standard basis of  $\mathbb{R}^n$ .** There is already a standard family of vectors for which these questions have a simple answer.

### Definition 3.8

We introduce the following vectors of  $\mathbb{R}^n$ :

$$\vec{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \dots, \vec{e}_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

This family of vectors is generally called the *standard basis* of  $\mathbb{R}^n$ .

### Theorem 3.9

Every vector  $\vec{x} = (x_1, \dots, x_n)^T$  of  $\mathbb{R}^n$  can be written in a unique way as a linear combination of  $\vec{e}_1, \dots, \vec{e}_n$ , namely:

$$\vec{x} = x_1\vec{e}_1 + \dots + x_n\vec{e}_n.$$

### Example 3.10

The vector  $\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \in \mathbb{R}^3$  can be written as a linear combination of the standard basis vectors

for  $\mathbb{R}^3$ ,  $\vec{e}_1, \vec{e}_2, \vec{e}_3$  since

$$\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = \vec{e}_1 + 3\vec{e}_2 + 5\vec{e}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 5 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Note that it is also possible to use a *non-standard* basis to describe the vectors, as in the following example.

### Example 3.11

The vector  $\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \in \mathbb{R}^3$  from the previous example can also be written as a linear combination of the non-standard basis vectors for  $\mathbb{R}^3$ ,  $\vec{e}_1, \vec{f}_2, \vec{f}_3$  where

$$\vec{f}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{f}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

since

$$\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = \vec{e}_1 + 4\vec{f}_2 - 1\vec{f}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

### Exercise 3.12

◆ Can I use  $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$  along with  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  as basis for  $\mathbb{R}^3$ ?

## 3.4 Systems of linear equations in vector form.

Consider a system of linear equations whose associated matrix is an  $m \times n$  matrix of the form:

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \left( \text{col}_1(A) \mid \dots \mid \text{col}_n(A) \right).$$

For  $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$  a vector of  $\mathbb{R}^n$ , we define :

$$A\vec{x} := x_1\text{col}_1(A) + \dots + x_n\text{col}_n(A) \in \mathbb{R}^m.$$

In other words, for  $1 \leq i \leq m$ , the  $i$ -th component of  $A\vec{x}$  is  $(A\vec{x})_i = \sum_{1 \leq k \leq n} a_{ik}x_k$ .

Consider the following system of linear equations:

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & = & b_m \end{array} \right.$$

Let  $A$  be the matrix associated to this system, and let  $\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ . A vector  $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  is solution of that system if and only we have:

$$A\vec{x} = \vec{y}.$$

Solving that system is thus equivalent to the following problem:

*Is  $\vec{y} \in \mathbb{R}^n$  a linear combination of  $\text{col}_1(A), \dots, \text{col}_n(A)$ ?*

Thus, a  $m \times n$  matrix is not ‘just a bunch of numbers’, but we can use it to associate to a vector  $\vec{x} \in \mathbb{R}^n$  a new vector  $A\vec{x} \in \mathbb{R}^m$ . In other words, we can associate to any  $m \times n$  matrix a map from  $\mathbb{R}^n$  to  $\mathbb{R}^m$ .

### 3.5 Spans of vectors and spanning families in $\mathbb{R}^n$

Given a family  $\vec{v}_1, \dots, \vec{v}_k$  of vectors of  $\mathbb{R}^n$ , we start by considering the questions: What vectors of  $\mathbb{R}^n$  can be obtained as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ ? We first introduce some definition:

#### Definition 3.13

The *span* of a family  $\vec{v}_1, \dots, \vec{v}_k$  of vectors of  $\mathbb{R}^n$  is the set  $\text{span}(\vec{v}_1, \dots, \vec{v}_k)$  of vectors of  $\mathbb{R}^n$  that can be written as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ . In other words,

$$\text{span}(\vec{v}_1, \dots, \vec{v}_k) = \left\{ \sum_{i=1}^k \lambda_i \vec{v}_i \mid \lambda_i \in \mathbb{R} \text{ for every } i \right\}.$$

Checking whether a given vector is a linear combination of a family of vectors is checked by solving a system of linear equations. Here is an example:

#### Example 3.14

Let us determine whether the vector  $(3, -4, 2)^T$  is a linear combination of  $(1, 0, 2)^T$  and

$(1, 1, 3)^T$ . We have to solve the equation

$$\lambda \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \kappa \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -4 \\ 2 \end{pmatrix}$$

with variables  $\lambda, \kappa \in \mathbb{R}$ . Using coordinates, we express this as a system of linear equations, and perform Gaussian elimination:

$$\left( \begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 1 & -4 \\ 2 & 3 & 2 \end{array} \right) \xrightarrow[R_3 \leftrightarrow R_3 - 2R_1]{\sim} \left( \begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 1 & -4 \\ 0 & 1 & -4 \end{array} \right) \xrightarrow[R_3 \rightarrow R_3 - R_2]{\sim} \left( \begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{array} \right)$$

From this echelon form, we see that the system has exactly one solution (all variables are pivot variables and no inconsistent line). We solve the resulting system by substitution, which yields  $\kappa = -4$  and  $\lambda = 3 - \kappa = 7$ . Thus,

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

### Exercise 3.15 Linear combination test in $\mathbb{R}^3$

Determine whether the vector  $\begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$  is a linear combination of  $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$ . If it is, find scalars  $\lambda, \kappa \in \mathbb{R}$  such that

$$\lambda \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \kappa \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}.$$

**Solution:** We must solve

$$\lambda \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \kappa \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}.$$

In coordinates this becomes the linear system

$$\begin{cases} \lambda = 2, \\ \lambda + 2\kappa = 5, \\ \kappa = 1. \end{cases}$$

which is inconsistent.

Equivalently, we row-reduce the augmented matrix:

$$\left( \begin{array}{cc|c} 1 & 0 & 2 \\ 1 & 2 & 5 \\ 0 & 1 & 1 \end{array} \right) \xrightarrow[R_2 \leftrightarrow R_3]{\sim} \left( \begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 2 & 3 \\ 0 & 1 & 1 \end{array} \right) \xrightarrow[R_3 - 2R_2]{\sim} \left( \begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 3 \end{array} \right) \xrightarrow[R_3 - 2R_2]{\sim} \left( \begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right).$$

The last row represents  $0\lambda + 0\kappa = 1$ , which is impossible. Hence the system has no solution, so  $\begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$  is *not* a linear combination of  $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$ .

Since checking whether a given vector  $\vec{v}$  is a linear combination of vectors  $\vec{v}_1, \dots, \vec{v}_k$  is checked by solving a system of linear equations, it follows that there is either no way to write  $\vec{v}$  as a linear combination  $\vec{v}_1, \dots, \vec{v}_k$  (precisely when  $\vec{v}$  is not in the span of  $\vec{v}_1, \dots, \vec{v}_k$ ), exactly one way or infinitely many ways.

### Theorem 3.16

Let  $\vec{u}_1, \dots, \vec{u}_k$  be a family of vectors of  $\mathbb{R}^n$  and consider  $V := \text{span}(\vec{u}_1, \dots, \vec{u}_k)$ . We have the following:

- $\vec{0} \in V$ ,
- For every  $\vec{x}, \vec{y} \in V$ , we also have  $\vec{x} + \vec{y} \in V$ .
- For every  $\vec{x} \in V$  and  $\lambda \in \mathbb{R}$ , we also have  $\lambda\vec{x} \in V$ .

More generally, any linear combination of vectors of  $V$  is again in  $V$ .

### Definition 3.17

A subset of  $\mathbb{R}^n$  containing the null vector and stable under linear combinations is called a *subspace* of  $\mathbb{R}^n$ .

The algebraic properties of spans of vectors mentioned above have a geometric counterpart: Spans of vectors, seen as subsets of  $\mathbb{R}^n$ , have a very simple shape: line or plane through the origin in  $\mathbb{R}^3$  for instance. More complicated shapes, such as spheres, hyperboloids, etc. can never be spans.

As an illustration, we now list the various possibilities for the span of two vectors of  $\mathbb{R}^2$ . In particular, we see that such spans are geometrically very simple.

### Theorem 3.18

The span of two vectors of  $\mathbb{R}^2$  is either:

- (i)  $\{\vec{0}\}$ ,
- (ii) a straight line through the origin,
- (iii) all of  $\mathbb{R}^2$ .

*Proof.* Consider two arbitrary vectors  $\vec{v} = (v_1, v_2)^T, \vec{w} = (w_1, w_2)^T$ . If  $\vec{v} = \vec{w} = \vec{0}$ , the span of  $\vec{v}$  and  $\vec{w}$  is just  $\{\vec{0}\}$ . If they are collinear and for instance  $\vec{v}$  is non-vanishing, then the span of  $\vec{v}$  and  $\vec{w}$  is the straight line parallel to  $\vec{v}$  going through the origin.

If they are not collinear, then we now show that  $\vec{v}$  and  $\vec{w}$  span  $\mathbb{R}^2$ . Let  $\vec{u} = (a, b)^T$  be an arbitrary vector of  $\mathbb{R}^2$ . We want to write  $\vec{u}$  as a linear combination of  $\vec{v}$  and  $\vec{w}$ . In other words, we want to find numbers  $\lambda, \mu \in \mathbb{R}$  such that  $\lambda\vec{v} + \mu\vec{w} = \vec{u}$ . By taking coordinates, this yields the following equations:

$$\begin{cases} \lambda v_1 + \mu w_1 = u_1 \\ \lambda v_2 + \mu w_2 = u_2 \end{cases}$$

Here, the variables are  $\lambda$  and  $\mu$ , and the coefficients  $u_1, u_2, v_1, v_2, w_1, w_2$  are constants. We thus have a system of linear equations. We perform the following row operations:

$$\xrightarrow[R_2 \rightarrow v_1 R_2 - v_2 R_1]{R_1 \rightarrow w_2 R_1 - w_1 R_2} \begin{cases} (v_1 w_2 - v_2 w_1) \lambda = u_1 w_2 - u_2 w_1 \\ (v_2 w_1 - v_1 w_2) \mu = u_2 v_1 - u_1 v_2 \end{cases}$$

Since  $\vec{v}$  and  $\vec{w}$  are not collinear, we have  $v_1 w_2 - v_2 w_1 \neq 0$ , so we can find solutions  $\lambda$  and  $\mu$  in terms of the other constants. Thus, there is a solution to our system of equations, so  $\vec{u}$  is a linear combination of  $\vec{v}$  and  $\vec{w}$ .  $\square$

## 3.6 Spanning families.

### Definition 3.19

We say that a family  $\vec{v}_1, \dots, \vec{v}_k$  of vectors of  $\mathbb{R}^n$  spans  $\mathbb{R}^n$  (or is a *spanning family* of  $\mathbb{R}^n$ , or that  $\mathbb{R}^n$  is spanned by  $\vec{v}_1, \dots, \vec{v}_k$ ), if  $\text{span}(\vec{v}_1, \dots, \vec{v}_k) = \mathbb{R}^n$ , that is, if every vector of  $\mathbb{R}^n$  can be written as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ .

### Theorem 3.20

Let  $\vec{v}_1, \dots, \vec{v}_k$  be a spanning family of  $\mathbb{R}^n$ . Then  $k \geq n$ . In other words, a spanning family of  $\mathbb{R}^n$  contains at least  $n$  vectors.

*Proof.* Given a vector  $\vec{b} = (b_1, \dots, b_n)^T$  of  $\mathbb{R}^n$ , we have to consider the equation

$$\lambda_1 \vec{v}_1 + \dots + \lambda_k \vec{v}_k = \vec{b},$$

where  $\lambda_1, \dots, \lambda_k$  are variables. Taking coordinates gives us a system of  $n$  equations (one for each coordinate of  $\mathbb{R}^n$ ) with  $k$  variables. We can now perform Gaussian elimination to get a system in echelon form. We will now show that each row contains a pivot, which will be enough to conclude that  $k \geq n$ : Since each pivot must be strictly to the right of the pivots of the previous rows, this means that there must be at least as many columns as rows in the associated matrix, hence  $k \geq n$ .

If there was a row without a pivot, then the last row of the augmented matrix the echelon form would be of the form

$$\left( \begin{array}{ccc|c} 0 & \cdots & 0 & \beta \end{array} \right),$$

where  $\beta$  is a non-trivial linear combination of  $b_1, \dots, b_n$ . Note that if  $\beta \neq 0$ , then we have a forbidden row, and the system has no solution. We can now choose specific values of  $b_1, \dots, b_n$  such that  $\beta \neq 0$ . This corresponds to an equation of the form

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

which also has no solution. Thus, the vector  $\vec{b} = (b_1, \dots, b_n)^T$  is not a linear combination of  $\vec{v}_1, \dots, \vec{v}_n$ , hence these vectors do not form a spanning family.  $\square$

### 3.7 Linearly independent families in $\mathbb{R}^n$

We now study family of vectors  $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$  for which there is a unique way to write a given in the span as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ .

#### Definition 3.21

A family of vectors of  $\mathbb{R}^n$  is *linearly dependent* if one of the vectors is a linear combination of the other vectors. Equivalently, a family of vectors  $\vec{v}_1, \dots, \vec{v}_k$  is linearly dependent if there exist scalars  $\lambda_1, \lambda_k \in \mathbb{R}$  not all zero such that

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k = \vec{0}.$$

A family of vectors  $\vec{v}_1, \dots, \vec{v}_k$  is *linearly independent* if for every  $\lambda_1, \lambda_k \in \mathbb{R}$ , we have

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k = \vec{0} \implies \lambda_1 = \cdots = \lambda_k = 0.$$

A family of vectors  $\vec{v}_1, \dots, \vec{v}_k$  is linearly independent if and only if there is a unique way to write any vector of the span as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ .

Determining whether vectors are linearly independent amounts to solving a system of linear equations, which we do using Gaussian elimination.

#### Theorem 3.22

To determine whether vectors  $\vec{v}_1, \dots, \vec{v}_k$  are linearly independent, one uses Gaussian elimination to find all solutions to the system

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots + c_k \vec{v}_k = \vec{0}.$$

If all variables  $c_1, \dots, c_k$  are pivot variables, this homogeneous system of linear equations has only the trivial solution  $c_1 = c_2 = \cdots = c_k = 0$ . In this case,  $\vec{v}_1, \dots, \vec{v}_k$  are linearly independent. Otherwise, nontrivial solutions exist and the vectors are linearly dependent.

Let us consider some examples:

### Example 3.23

We want to determine whether the vectors  $(2, 1, 1)^T$ ,  $(1, 2, 1)^T$ , and  $(1, 1, 2)^T$  are linearly independent. We need to study the equation

$$\lambda \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \kappa \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

with variables  $\lambda, \kappa, \mu \in \mathbb{R}$ . Taking coordinates, this leads to the system of linear equations:

$$\begin{cases} 2\lambda + \kappa + \mu = 0 \\ \lambda + 2\kappa + \mu = 0 \\ \lambda + \kappa + 2\mu = 0 \end{cases}$$

We use Gaussian elimination to solve this system. We get

$$\begin{array}{ccc|c} 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 0 \end{array} \xrightarrow[R_1 \leftrightarrow R_2]{\sim} \begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 2 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 \end{array} \xrightarrow[R_2 \rightarrow R_2 - 2R_1]{R_3 \rightarrow -R_3 + R_1} \begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -3 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{array} \xrightarrow[R_3 \leftrightarrow 3R_3 + R_1]{\sim} \begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -3 & -1 & 0 \\ 0 & 0 & -4 & 0 \end{array}$$

Thus, all the variables are pivot variables and the system is consistent as it is homogeneous, so the system has exactly one solution. Since the system has at least one solution, namely  $\lambda = \kappa = \mu = 0$ , this must be the only solution. As a consequence, the vectors  $(2, 1, 1)^T$ ,  $(1, 2, 1)^T$ , and  $(1, 1, 2)^T$  are linearly independent.

### Exercise 3.24 Linearly dependent vectors

Determine whether the vectors  $(1, 2, 3, -1)^T$ ,  $(2, 1, 3, 1)^T$  and  $(4, 5, 9, -1)^T$  are linearly dependent.

**Solution:** We must find all solutions  $(c_1, c_2, c_3)$  of the homogeneous system:

$$c_1 \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \\ 3 \\ 1 \end{pmatrix} + c_3 \begin{pmatrix} 4 \\ 5 \\ 9 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We reduce the associated system using Gaussian elimination:

$$\left\{ \begin{array}{l} c_1 + 2c_2 + 4c_3 = 0 \\ 2c_1 + c_2 + 5c_3 = 0 \\ 3c_1 + 3c_2 + 9c_3 = 0 \\ -c_1 + c_2 - c_3 = 0 \end{array} \right. \rightsquigarrow \left( \begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 2 & 1 & 5 & 0 \\ 3 & 3 & 9 & 0 \\ -1 & 1 & -1 & 0 \end{array} \right)$$

$$\xrightarrow{\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - 3R_1 \\ R_4 \rightarrow R_4 + R_1}} \left( \begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 3 & 3 & 0 \end{array} \right) \xrightarrow{\substack{R_3 \rightarrow R_3 - R_2 \\ R_4 \rightarrow R_4 + \frac{1}{3}R_2}} \left( \begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

Not all the variables are pivot variables ( $c_3$  is a free variable) and therefore the system has infinitely many solutions, which implies that the vectors are linearly dependent. More precisely, if we rewrite it as a system of linear equations, we get

$$\left\{ \begin{array}{l} c_1 + 2c_2 + 4c_3 = 0 \\ c_2 + c_3 = 0 \end{array} \right. \rightarrow c_1 = -2c_2 - 4c_3 = -2\alpha \quad .$$

For instance, by setting  $\alpha = 1$  we obtain

$$-2 \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 9 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

### 3.8 Bases of $\mathbb{R}^n$ .

We now study families of vectors such that every vector of  $\mathbb{R}^n$  can be written *in a unique way* as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ . We have already seen one example of such a family, namely the standard basis of  $\mathbb{R}^n$ :

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

This family of vectors is generally referred to as the *standard basis* of  $\mathbb{R}^n$ . This notion can be generalised as follows:

#### Definition 3.25

A *basis* of  $\mathbb{R}^n$  is a family of vectors  $\vec{v}_1, \dots, \vec{v}_k$  such that every vector of  $\mathbb{R}^n$  can be written in a unique way as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ . In other words, a family of vectors is a basis if and only if it is both spanning (existence of a linear combination) and linearly independent (uniqueness of a linear combination).

#### Theorem 3.26 A

basis of  $\mathbb{R}^n$  contains exactly  $n$  vectors. □

### 3.9 Characterisation of bases of $\mathbb{R}^n$

#### Theorem 3.27

Let  $\vec{u}_1, \dots, \vec{u}_n$  be a family of **exactly**  $n$  vectors of  $\mathbb{R}^n$ . The following are equivalent:

- (i)  $\vec{u}_1, \dots, \vec{u}_n$  is a basis of  $\mathbb{R}^n$ .
- (ii)  $\vec{u}_1, \dots, \vec{u}_n$  spans  $\mathbb{R}^n$ .
- (iii)  $\vec{u}_1, \dots, \vec{u}_n$  is linearly independent.

Note that (i)  $\Rightarrow$  (ii), (iii). Let us show that (ii)  $\Rightarrow$  (i). We want to show that  $\vec{u}_1, \dots, \vec{u}_n$  is linearly independent. To that end, we try to solve the homogeneous system of equations  $\lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n = \vec{0}$  by Gaussian elimination. Since we have  $n$  equations with  $n$  variables, either every variable is a pivot variable, or there exists a free variable. In the latter case, we can find a vector  $\vec{u}$  of  $\mathbb{R}^n$  such that reducing the system of equations  $\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n = \vec{u}$  would give a bottom row of the form

$$( 0 \ \dots \ 0 \mid 1 ).$$

This would imply that the system associated to the equation  $\lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n = \vec{u}$  is inconsistent, contradicting the fact that the family spans  $\mathbb{R}^n$ . Thus, every variable is a pivot variable, and we know the system has exactly one solution:  $(\lambda_1, \dots, \lambda_n) = (0, \dots, 0)$ .

Let us now show that  $(iii) \Rightarrow (i)$ , using a similar strategy. We want to show that  $\vec{u}_1, \dots, \vec{u}_n$  spans  $\mathbb{R}^n$ . To that end, we try to solve the system of equations  $\lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n = \vec{u}$ , for some vector  $\vec{u}$  of  $\mathbb{R}^n$ , by Gaussian elimination. Since we have  $n$  equations with  $n$  variables, either every variable is a pivot variable, or there exists a free variable. In the latter case, the system of equations  $\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n = \vec{0}$  would have infinitely many solutions, contradicting the fact that the family is linearly independent. Thus, every variable is a pivot variable, which implies that the system of equations  $\lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n = \vec{u}$  admits at least one solution.

The implication  $(iii) \Rightarrow (i)$  gives us a possible method to show that a given set of vectors forms a basis of  $\mathbb{R}^n$ . However, there is often a faster way to show that a given family of vectors forms a basis of  $\mathbb{R}^n$ :

### Theorem 3.28

Let  $\vec{v}_1, \dots, \vec{v}_n$  be a family of  $n$  vectors of  $\mathbb{R}^n$ , and let

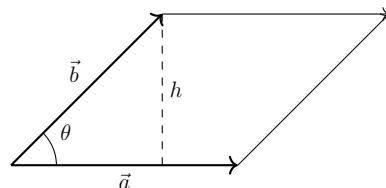
$$A := \left( \begin{array}{c|c|c|c} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{array} \right)$$

be the associated matrix. Then:

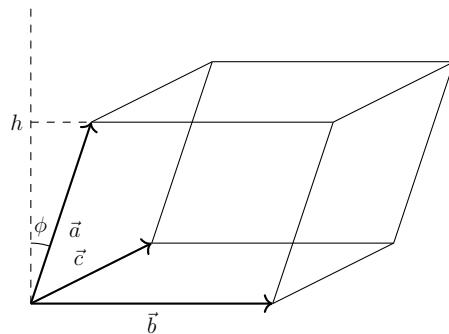
$$\vec{v}_1, \dots, \vec{v}_n \text{ is a basis of } \mathbb{R}^n \iff \det(A) \neq 0.$$

This works because the determinant is calculating a *volume* (or area in 2 dimensions) defined by the vectors.

In  $\mathbb{R}^2$  the determinant gives the area of the parallelogram defined by  $\vec{a}, \vec{b}$ . If  $\vec{b}$  is collinear with  $\vec{a}$  the area is zero.



In  $\mathbb{R}^3$ , on the other hand, three vectors  $\vec{a}, \vec{b}, \vec{c}$  generically define a parallelepiped, as shown below. If one of the vectors is linearly dependent on the other two it will be coplanar with them and the parallelepiped will collapse to have zero volume. It is harder to picture in higher dimensions, but the idea is the same.



**Example 3.29**

Using a determinant to test for a basis of  $\mathbb{R}^3$ :

Let

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Form the associated matrix with these vectors as columns:

$$A := (\vec{v}_1 \mid \vec{v}_2 \mid \vec{v}_3) = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 2 & -1 & 1 \end{pmatrix}.$$

Compute its determinant (expand along the first row, for example):

$$\det(A) = 1 \cdot \det \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} - 0 \cdot (\dots) + 3 \cdot \det \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}.$$

Hence

$$\det(A) = (1 \cdot 1 - 2 \cdot (-1)) + 3(0 \cdot (-1) - 1 \cdot 2) = (1 + 2) + 3(-2) = 3 - 6 = -3 \neq 0.$$

Therefore, by the theorem,

$\vec{v}_1, \vec{v}_2, \vec{v}_3$  is a basis of  $\mathbb{R}^3$ .

We can also use the determinant test when the vectors do *not* form a basis:

**Example 3.30**

Using a determinant to test for a basis of  $\mathbb{R}^3$  (an example where the vectors don't form a basis).

Let

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

(Notice that  $\vec{v}_3 = \vec{v}_1 + \vec{v}_2$ .)

Form the associated matrix with these vectors as columns:

$$A := (\vec{v}_1 \mid \vec{v}_2 \mid \vec{v}_3) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & -1 & 1 \end{pmatrix}.$$

Compute its determinant (expand along the first row, for example):

$$\det(A) = 1 \cdot \det \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} - 0 \cdot (\dots) + 1 \cdot \det \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}.$$

Hence

$$\det(A) = (1 \cdot 1 - 1 \cdot (-1)) + (0 \cdot (-1) - 1 \cdot 2) = (1 + 1) + (-2) = 0.$$

Therefore, by the theorem,

$\vec{v}_1, \vec{v}_2, \vec{v}_3$  do not form a basis of  $\mathbb{R}^3$ .

(In fact they are linearly dependent since  $\vec{v}_3 = \vec{v}_1 + \vec{v}_2$ .)

## 4 Vector Spaces in General

In the previous chapter, we introduced the space  $\mathbb{R}^n$  of vectors. The key operations for vectors in  $\mathbb{R}^n$  are *adding* two vectors and *multiplying* a vector by a scalar. In this section, we will develop the general abstract framework allowing us to treat various examples of vector spaces in a uniform way.

### 4.1 Examples and definition

To be able to treat various cases with a single mathematical framework, we now introduce the notion of an *abstract vector space*. This is the notion that makes rigorous this idea of 'spaces of objects where one can add vectors and multiply them by a scalar'.

#### Definition 4.1

A *real vector space* is a set  $V$  (whose elements are called *vectors*) endowed with two operations:

- (i) an *addition* denoted  $+$  that associates to two vectors  $\vec{u}, \vec{v} \in V$  a vector  $\vec{u} + \vec{v} \in V$ ,
- (ii) a *multiplication by a scalar* denoted  $\cdot$  that associates to a vector  $\vec{v} \in V$  and a scalar  $\lambda \in \mathbb{R}$  a vector  $\lambda \cdot \vec{v} \in V$  (or simply  $\lambda\vec{v}$ ).

We further require that the following *vector space axioms* are satisfied:

- (1) *null vector*: There is an element  $\vec{0} = \vec{0}_V \in V$ , the zero or *null vector*, such that  $\vec{v} + \vec{0} = \vec{v}$  for all  $\vec{v} \in V$ .
- (2) *opposite*: For all  $\vec{v} \in V$ , there is an element  $-\vec{v} \in V$  such that  $\vec{v} + (-\vec{v}) = \vec{0}$ .
- (3) *commutativity of the addition*:  $\vec{v} + \vec{w} = \vec{w} + \vec{v}$  for all  $\vec{v}, \vec{w} \in V$ ,
- (4) *associativity of the addition*:  $(\vec{v} + \vec{w}) + \vec{u} = \vec{v} + (\vec{w} + \vec{u})$  for all  $\vec{v}, \vec{w}, \vec{u} \in V$ ,
- (5) *distributivity of the scalar multiplication*:  $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$  and  $(a+b)\vec{v} = a\vec{v} + b\vec{v}$ ,  $a, b \in \mathbb{R}$ ,  $\vec{v}, \vec{w} \in V$ ,
- (6) *associativity of the scalar multiplication*:  $a(b\vec{v}) = (ab)\vec{v}$ ,  $a, b \in \mathbb{R}$ ,  $\vec{v} \in V$ ,
- (7) *compatibility*:  $1 \cdot \vec{v} = \vec{v}$ .

This list of axioms (1)-(7) may seem long and technical. However, you should convince yourself that they encode the usual properties that one expects from addition and multiplication by a number for vectors, functions, etc. Moreover, the strength of this abstract framework is that, once we know that something is a vector space, we can treat objects of that space (functions, sequences, or more exotic objects) as if they were vectors, and use our geometric intuition to solve problems.

- (i) The space  $\mathbb{R}^n$  is a vector space for the addition and multiplication by a scalar introduced in the previous chapter. The null vector  $(0, \dots, 0)^T$  satisfies axiom (1). The opposite of a vector  $(v_1, \dots, v_n)^T$  is defined as  $(-v_1, \dots, -v_n)^T$  to satisfy axiom (2). One then checks that axioms (3)-(7) hold. Indeed, one has to check the equations coordinate by coordinate, and each of these equations then boils down to a standard rule about addition and multiplication of real numbers (commutativity of addition, distributivity of the multiplication, etc.)

- (iii) The space  $\mathbb{R}^{\mathbb{N}}$  is a vector space for the addition and multiplication by a scalar defined above. Here, a vector is a sequence

$$\begin{pmatrix} u_0 \\ u_1 \\ \vdots \end{pmatrix},$$

the zero vector is the *zero sequence*

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \end{pmatrix},$$

and the opposite of a vector is given by

$$-\begin{pmatrix} u_0 \\ u_1 \\ \vdots \end{pmatrix} := \begin{pmatrix} -u_0 \\ -u_1 \\ \vdots \end{pmatrix}.$$

Here again, one then checks that axioms (3)-(7) hold, by checking each equation pointwise.

**Complex vector spaces.** We have just defined *real* vector spaces, that is, vector spaces where the set of scalars is  $\mathbb{R}$ . These will be the almost sole focus of this course. However, one can analogously define *complex vectors spaces* by setting the set of scalars to be  $\mathbb{C}$  and having the same list of axioms. Examples of complex vectors spaces are:

- The space  $\mathbb{C}^n$  of vectors of the form  $(z_1, \dots, z_n)^T$  with  $z_i \in \mathbb{C}$  for every  $i$ .
- The space  $\mathcal{F}(\mathbb{C})$  of functions from  $\mathbb{C}$  to  $\mathbb{C}$ .
- The space  $\mathcal{P}(\mathbb{C})$  of polynomial functions with complex coefficients.
- The space  $\mathbb{C}^{\mathbb{N}}$  of complex sequences  $(z_0, z_1, \dots)$  with  $z_i \in \mathbb{C}$  for every  $i \geq 0$ .

## 4.2 Linear combinations in abstract vector spaces

We now generalise to abstract vector spaces the notions we introduced in  $\mathbb{R}^n$ .

### Definition 4.2

A vector  $\vec{v} \in V$  is called a *linear combination* of the vectors  $\vec{u}_1, \dots, \vec{u}_k \in V$ , if it can be written as

$$\vec{v} = a_1 \vec{u}_1 + a_2 \vec{u}_2 + \dots + a_k \vec{u}_k, \quad a_1, \dots, a_k \in \mathbb{R}.$$

This is an extension of our definition of linear combination of vectors in  $\mathbb{R}^n$ .

### Example 4.3

Consider the vector space  $\mathcal{P}_2(\mathbb{R})$  of polynomials of degree at most 2. Linear combinations of the two functions (=vectors) defined  $P_1(x) = x$  and  $P_2(x) = x^2$  are polynomial functions of

the form

$$a_1x + a_2x^2, \quad \text{for } a_1, a_2 \in \mathbb{R}.$$

For instance, the polynomial function defined  $P(x) = 5x^2 - 2x$  is a linear combination of  $x$  and  $x^2$ , since we have  $P = 5P_2 - 2P_1$ .

#### Definition 4.4

The *span* of a family of vectors  $\vec{u}_1, \dots, \vec{u}_k$ , usually denoted by  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$ , is the set of all possible linear combinations of  $\vec{u}_1, \dots, \vec{u}_k$ :

$$\text{span}(\vec{u}_1, \dots, \vec{u}_k) := \{a_1\vec{u}_1 + \dots + a_k\vec{u}_k \mid a_1, \dots, a_k \in \mathbb{R}\}.$$

We also say that  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$  is spanned by  $\vec{u}_1, \dots, \vec{u}_k$  or that these vectors span  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$ .

#### Example 4.5

- (i) The vectors  $(1, 0, 0)^T, (0, 1, 0)^T, (0, 0, 1)^T$  span  $\mathbb{R}^3$ .
- (ii) The vector space of polynomials up to degree  $n$ ,  $\mathcal{P}_n(\mathbb{R})$ , is spanned by the polynomials  $1, x, x^2, \dots, x^n$ , as every polynomial function in  $\mathcal{P}_n(\mathbb{R})$  can be written as  $a_0 \times 1 + \dots + a_n x^n$ .
- (iii) The polynomial function  $2x^2 + x - 1$  belongs to  $\text{span}(x^2 + x, x + 1)$  since  $2x^2 + x - 1 = 2(x^2 + x) - (x + 1)$ .

#### Definition 4.6

We say that vectors  $\vec{v}_1, \dots, \vec{v}_k$  are *linearly independent* if

$$c_1\vec{v}_1 + \dots + c_k\vec{v}_k = \vec{0} \implies c_1 = \dots = c_k = 0.$$

Otherwise, we say that the vectors are *linearly dependent*. This is an extension of our definition of linear independence for  $\mathbb{R}^n$ .

As in the case of  $\mathbb{R}^n$ , linear dependence as a simple interpretation: a family of vectors  $\vec{u}_1, \dots, \vec{u}_k$  is linearly dependent, if and only if one (i.e. at least one) of the  $\vec{u}_i$ 's can be expressed as a linear combination of the others.

#### Example 4.7

Let us show that the functions  $f_1(x) := \cos(x), f_2(x) := \cos(2x), f_3(x) := \cos(3x)$  are linearly independent vectors of  $\mathcal{F}(\mathbb{R})$ .

Let  $a, b, c \in \mathbb{R}$  be scalars such that  $af_1 + bf_2 + cf_3$  is the zero vector of  $\mathcal{F}(\mathbb{R})$ , that is, the zero function. This is equivalent to:

$$a \cos(x) + b \cos(2x) + c \cos(3x) = 0 \quad \text{for every } x \in \mathbb{R}.$$

By evaluating at  $x = \pi/2$ , we get  $b = 0$ , hence  $a \cos(x) + c \cos(3x) = 0$  for every  $x \in \mathbb{R}$ . By evaluating this equation at  $x = \pi/6$ , we get  $a \times \sqrt{3}/2 = 0$ , hence  $a = 0$ . Finally, we have  $c \cos(3x) = 0$  for every  $x \in \mathbb{R}$ , and evaluating at  $x = 0$  yields  $c = 0$ .

We thus have  $a = b = c = 0$ , and it follows that  $f_1, f_2, f_3$  are linearly independent.

### 4.3 Basis and dimension

#### Definition 4.8

A family of vectors  $\vec{u}_1, \dots, \vec{u}_n$  is called a *basis* of  $V$ , if it is a linearly independent set of vectors that span  $V$ .

#### Example 4.9

We have already seen the standard basis  $\vec{e}_1 = (1, 0, \dots, 0)^T, \dots, \vec{e}_n = (0, \dots, 0, 1)^T$  of  $\mathbb{R}^n$ .

Note that a real vector space has infinitely many bases. For instance,  $(1, 1)^T$  and  $(1, -1)^T$  is also a basis of  $\mathbb{R}^2$  Why?

Here is an important example:

#### Theorem 4.10

A basis for the vector space of polynomials of degree at most  $n$  is  $1, x, x^2, \dots, x^n$ .

*Proof.* The family clearly spans  $\mathcal{P}_n(\mathbb{R})$ , as every polynomial function of  $\mathcal{P}_n(\mathbb{R})$  can be written as a linear combination  $a_0 + a_1x + \dots + a_nx^n$  for some scalars  $a_0, \dots, a_n \in \mathbb{R}$ . Let us show that this family is free. Suppose that we have a linear combination  $a_0 + a_1x + \dots + a_nx^n$  that is the zero vector, that is,  $a_0 + a_1x + \dots + a_nx^n = 0$  for every  $x \in \mathbb{R}$ . By evaluating at  $x = 0$ , we get  $a_0 = 0$ . Now since the polynomial function  $a_0 + a_1x + \dots + a_nx^n$  is the zero function, so its derivative, so we get  $a_1 + 2a_2x + \dots + na_nx^{n-1} = 0$  for every  $x \in \mathbb{R}$ . Evaluating again at  $x = 0$ , we get  $a_1 = 0$ . By repeating the same procedure (differentiating and evaluating at  $x = 0$ ), we prove successively that  $a_0 = a_1 = \dots = a_n = 0$ . Thus, the family of polynomial functions  $1, x, \dots, x^n$  is linearly independent, so it is a basis of  $\mathcal{P}_n(\mathbb{R})$ .  $\square$

### 4.4 Dimension of a vector space.

In order to define the dimension of a vector space, we need the following important result:

#### Theorem 4.11

Any two bases for a vector space  $V$  contain the same number of vectors.

Before proving it, we need the following result:

**Theorem 4.12**

If a family of vectors  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$  is a basis of a vector space  $V$ , then every family of vectors of  $V$  containing more than  $n$  vectors is linearly dependent.

*Proof.* Let  $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_m$  be a family of  $m > n$  vectors of  $V$ . We will show that there exist  $c_1, c_2, \dots, c_m$  not all zero such that

$$c_1\vec{w}_1 + c_2\vec{w}_2 + \dots + c_m\vec{w}_m = \vec{0}. \quad (4.1)$$

Since  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$  spans  $V$ , each  $\vec{w}_i$  can be expressed as a linear combination of the  $\vec{v}_i$ 's:

$$\begin{aligned} \vec{w}_1 &= a_{11}\vec{v}_1 + a_{12}\vec{v}_2 + \dots + a_{1n}\vec{v}_n, \\ \vec{w}_2 &= a_{21}\vec{v}_1 + a_{22}\vec{v}_2 + \dots + a_{2n}\vec{v}_n, \\ &\vdots && \vdots \\ \vec{w}_m &= a_{m1}\vec{v}_1 + a_{m2}\vec{v}_2 + \dots + a_{mn}\vec{v}_n. \end{aligned}$$

Plugging this into (4.1), we have

$$c_1(a_{11}\vec{v}_1 + \dots + a_{1n}\vec{v}_n) + \dots + c_m(a_{m1}\vec{v}_1 + \dots + a_{mn}\vec{v}_n) = \vec{0}.$$

To have all the coefficients of the  $\vec{v}_1, \dots, \vec{v}_n$  vanish, note that it is sufficient to find  $c_1, \dots, c_m$  such that

$$\begin{aligned} a_{11}c_1 + a_{21}c_2 + \dots + a_{m1}c_m &= 0, \\ a_{12}c_1 + a_{22}c_2 + \dots + a_{m2}c_m &= 0, \\ &\vdots && \vdots \\ a_{1n}c_1 + a_{2n}c_2 + \dots + a_{mn}c_m &= 0. \end{aligned}$$

This is a homogeneous system of linear equations with more unknowns ( $m$ ) than equations ( $n$ ) and thus there is a solution with not all of the  $c_i$  being zero. It follows that (4.1) has a solution besides the trivial solution and so  $S'$  is linearly dependent.  $\square$

*Proof of Proposition 4.4.* Let  $B = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$  and  $B' = (\vec{v}'_1, \vec{v}'_2, \dots, \vec{v}'_m)$  be bases for  $V$ . From the above lemma, we conclude that since  $B$  is a basis and  $B'$  is linearly independent,  $m \leq n$ . Equally, since  $B'$  is a basis and  $B$  is linearly independent,  $n \leq m$ . Altogether, we have  $m = n$ .  $\square$

We are now able to define properly the dimension of a vector space:

**Theorem 4.13**

Let  $V$  be a vector space. We say that  $V$  is *finite-dimensional* if it has a finite basis, and *infinite-dimensional* otherwise. We define the *dimension* of  $V$  to be the number of vectors in any basis of  $V$ .

**Example 4.14**

We have constructed bases of several vector spaces. We have the following dimensions:

Space	$\mathbb{R}^2$	$\mathbb{R}^3$	$\mathbb{R}^n$	$\mathcal{P}_n(\mathbb{R})$
Dimension	2	3	$n$	$n + 1$

The following result is useful in finding a lower bound for the dimension of a vector space:

**Theorem 4.15**

Let  $V$  be a vector space, and let  $\vec{v}_1, \dots, \vec{v}_k$  be a linearly independent family. Then we have:

$$\dim V \geq k.$$

## 4.5 Vector subspaces

The previous examples are the main sources of vectors spaces for this course. However, we are often not interested in the space of all vectors, or of all functions. Instead, we are often interested in particular subsets of elements that satisfy some equation: system of linear equations, differential equations, etc.

In this section, we introduce the notion of *vector subspace* as the natural notion of subset of a vector space that is compatible with the operations of addition and multiplication by a scalar. We will see that the set of solutions of various equations naturally form a vector subspace of the associated vector space.

**Definition 4.16**

Let  $V$  be a vector space and let  $W$  be a subset of  $V$ . We say that  $W$  is a *vector subspace* (or simply a subspace) of  $V$  if the following holds:

- (i)  $\vec{0} \in W$ ,
- (ii) for all  $\vec{w}_1, \vec{w}_2 \in W$ , we have  $\vec{w}_1 + \vec{w}_2 \in W$  and
- (iii) for all  $\lambda \in \mathbb{R}$  and  $\vec{w} \in W$ , we have  $\lambda\vec{w} \in W$ .

A vector subspace  $W$  of  $V$  is itself a vector space, when endowed with the addition and scalar multiplication coming from  $V$ . (This is actually an equivalence: a subset  $W$  is a subspace if and only if it  $W$  is a vector space when endowed with the addition and scalar multiplication from  $V$ .)

As a consequence, we can talk of the dimension or of bases of a given subspace.

**Example 4.17**

- (i) In a vector space  $V$ , the whole space  $V$  and the trivial subset  $\{\vec{0}\}$  are always vector subspaces.
- (ii) Consider the subset  $W = \{(x, y)^T \in \mathbb{R}^2 : x + y = 0\}$  of  $\mathbb{R}^2$ . That is,  $W$  consists of all

the vectors of the form  $(x, -x)^T$  for  $x \in \mathbb{R}$ . The null vector  $(0, 0)^T$  is clearly in  $W$ . If we add two vectors or multiply a vector by a scalar in  $W$ , we end up back in  $W$ :

$$\begin{pmatrix} x_1 \\ -x_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ -x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ -(x_1 + x_2) \end{pmatrix} \quad \text{and} \quad \lambda \begin{pmatrix} x_1 \\ -x_1 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ -\lambda x_1 \end{pmatrix}.$$

Thus,  $W$  is a subspace of  $\mathbb{R}^2$ . Note that every vector of  $W$  is of the form

$$\begin{pmatrix} x \\ -x \end{pmatrix} = x \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{for some } x \in \mathbb{R},$$

so  $W$  has dimension 1 and a basis of  $W$  is given by the vector  $(1, -1)^T$ .

**Remark.** If a subspace  $W$  contains a vector  $\vec{w}$ , it also contains its opposite  $-\vec{w}$ , as  $W$  is stable under multiplication by a scalar and  $-\vec{w} = (-1) \cdot \vec{w}$ .

### Example 4.18

The subset  $W' = \{(x, y)^T | x + y = 1\}$  is not a vector subspace of  $\mathbb{R}^2$ , as  $\vec{0} = (0, 0)^T$  is not an element of  $W'$ . Another reason is that the sum of two elements  $w_1, w_2 \in W'$  is not always in  $W'$ . For instance,  $(1, 0)^T$  and  $(0, 1)^T$  belong to  $W'$ , but their sum  $(1, 1)^T$  does not.

### Example 4.19

- (i) Let  $V = \mathbb{R}^2$ . Then  $W = \{(x, y)^T | x = 3y\}$  is a vector subspace, since vectors of  $W$  are of the form

$$\begin{pmatrix} 3y \\ y \end{pmatrix} = y \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \text{for some } y \in \mathbb{R},$$

and we have:

$$\begin{pmatrix} 3y_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} 3y_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3(y_1 + y_2) \\ y_1 + y_2 \end{pmatrix} \quad \text{and} \quad \lambda \begin{pmatrix} 3y \\ y \end{pmatrix} = \begin{pmatrix} 3\lambda y \\ \lambda y \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Thus,  $W$  is a subspace of  $\mathbb{R}^2$  of dimension 1, and a basis of it is given by the vector  $(3, 1)^T$ .

In general, lines through the origin of  $\mathbb{R}^2$  form vector subspaces of  $\mathbb{R}^2$ .

- (ii) Lines in  $\mathbb{R}^2$  that do not pass through the origin do not contain  $\vec{0}$  and thus are not vector subspaces of  $\mathbb{R}^2$ , cf. example in 4.5.

**Subspaces of  $\mathbb{R}^2$ .** We have seen already that we can completely describe the subspaces of  $\mathbb{R}^2$ :

- (i) the trivial subspace  $\{\vec{0}\}$ ,

- (ii) lines through the origin,
- (iii) the whole space  $\mathbb{R}^2$ .

There is a similar picture in  $\mathbb{R}^3$ , where subspaces can be lines through the origin, planes through the origin, etc. Geometrically, being stable under addition and scalar multiplication makes vector subspaces ‘look flat’.

**Solutions of systems of linear differential equations as subspaces.** We saw in a previous example how the set of solutions of certain differential equations may be described as a span, and hence is a subspace of  $\mathcal{F}(\mathbb{R})$ . Even without an explicit description of the solutions, it is possible to show that the set of solutions forms a subspace.

Consider the following differential equation (linearised simple pendulum):

$$y'' + \omega^2 y = 0.$$

Then the set  $W$  of solutions of this equation is a subspace of  $\mathcal{F}(\mathbb{R})$ .

*Proof.* Let  $y_1, y_2 \in W$  and  $\lambda \in \mathbb{R}$ . Then:

$$\begin{aligned} (y_1 + y_2)'' + \omega^2(y_1 + y_2) &= (y_1'' + \omega^2 y_1) + (y_2'' + \omega^2 y_2) = 0 + 0 = 0, \text{ hence } y_1 + y_2 \in W \\ (\lambda y_1)'' + \omega^2(\lambda y_1) &= \lambda(y_1'' + \omega^2 y_1) = \lambda 0 = 0, \text{ hence } \lambda y_1 \in W. \end{aligned}$$

We conclude that  $W$  is a vector subspace of  $\mathcal{F}(\mathbb{R})$ . □

**Dimension of subspaces.** Since a subspace of vector space is itself a vector space, it also has a dimension. It is natural to wonder whether the dimension behaves well with respect to subspaces: For instance, is the dimension of a subspace at most the dimension of the original vector space? While this intuitively obvious, it requires a proof. And indeed, things go extremely well:

### Theorem 4.20

Let  $V$  be a finite-dimensional vector space, and let  $W$  be a subspace. Then

$$\dim W \leq \dim V.$$

Moreover, we have

$$\dim W = \dim V \iff W = V.$$

This result can provide a simple way to show the equality between two subspaces: Instead of showing both inclusions, it is only necessary to show one inclusion and the equality of dimensions, something that is generally easier to handle. We will see applications in the next chapter.

The previous theorem relies on the following results:

Let  $V$  be an  $n$ -dimensional vector space and let  $\vec{v}_1, \dots, \vec{v}_k$  be a linearly independent family of vectors. Then  $\vec{v}_1, \dots, \vec{v}_k$  can be extended to a basis  $\vec{v}_1, \dots, \vec{v}_n$  of  $V$ . □

Suppose that a family  $\vec{v}_1, \dots, \vec{v}_k$  of vectors spans a (finite-dimensional) vector space  $V$ . Then there exists a subfamily that is a basis for  $V$ .  $\square$

## 5 The solution space of a homogeneous system of linear equations.

In F17ZB we looked at solving  $A\vec{x} = \vec{b}$ , now we are going to look in more detail at solving the *homogeneous* system of equations

$$A\vec{x} = \vec{0}$$

This leads on naturally to a discussion of *rank* for matrices (we have already seen one definition of this in F17ZB).

### Theorem 5.1

The set of solutions of a *homogeneous* system of linear equations in  $n$  unknowns is a vector subspace of  $\mathbb{R}^n$  whose dimension is the number of free variables, after reducing the system to its echelon form.

*Proof.* Let  $W$  be the set of solutions. By definition, we have  $W = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$  for some matrix  $A$ . In particular, the null vector belongs to  $W$ . Let  $\vec{x}_1, \vec{x}_2 \in W$  and  $\lambda \in \mathbb{R}$ . We have:

$$\begin{aligned} A(\vec{x}_1 + \vec{x}_2) &= A\vec{x}_1 + A\vec{x}_2 = \vec{0} + \vec{0} = \vec{0}, \text{ hence } \vec{x}_1 + \vec{x}_2 \in W \\ A(\lambda\vec{x}_1) &= \lambda A\vec{x}_1 = \lambda\vec{0} = \vec{0}, \text{ hence } \lambda\vec{x}_1 \in W. \end{aligned}$$

Thus,  $W$  is a vector subspace of  $\mathbb{R}^n$ . We solve the system  $A\vec{x} = \vec{0}$  using Gaussian elimination. As usual, we can express solutions in terms of the  $k \leq n$  free variables  $\alpha_1, \dots, \alpha_k$  of the system. In particular, we write the general solution as a linear combination of the form  $\alpha_1\vec{v}_1 + \dots + \alpha_k\vec{v}_k$  for some vectors  $\vec{v}_1, \dots, \vec{v}_k$  of  $\mathbb{R}^n$ . In particular, we see that the family  $\vec{v}_1, \dots, \vec{v}_k$  spans the subspace of solutions. To show that this family is also free, notice that if a free variable  $\alpha_i$  corresponds to the variable  $x_j$ , then the  $j$ -th component of the vector  $\alpha_1\vec{v}_1 + \dots + \alpha_k\vec{v}_k$  is exactly  $\alpha_i$ . In particular, if  $\alpha_1\vec{v}_1 + \dots + \alpha_k\vec{v}_k = \vec{0}$ , then all the components must be zero, and it follows that  $\alpha_1 = \dots = \alpha_k = 0$ .  $\square$

**Remark.** Note that  $\{x \in \mathbb{R}^n : A\vec{x} = \vec{b}\}$  with  $\vec{b} \neq \vec{0}$  is *not* a vector subspace. Indeed, this subset does not contain the null vector as  $A\vec{0} = \vec{0} \neq \vec{b}$ . This is analogous to lines in  $\mathbb{R}^2$  not running through the origin.

To find a basis of the space of solutions of a homogeneous system of linear equations, we can apply the following algorithm:

- Solve the system using Gaussian elimination.
- Express solutions in terms of the free variables  $\alpha_1, \dots, \alpha_k$ .
- Decompose the general solution as a linear combination of the form  $\alpha_1\vec{v}_1 + \dots + \alpha_k\vec{v}_k$ .
- A basis of the space of solutions is  $\vec{v}_1, \dots, \vec{v}_k$ .

**Example 5.2**

Consider the following system of linear equations:

$$\begin{aligned}x_1 - x_2 + 2x_3 + x_4 &= 0, \\2x_1 + x_2 - x_3 + x_4 &= 0, \\4x_1 - x_2 + 3x_3 + 3x_4 &= 0, \\x_1 + 2x_2 - 3x_3 &= 0.\end{aligned}$$

We would like to determine a basis of the subspace  $W$  of solutions. We first perform Gaussian elimination to reduce the system:

$$\left( \begin{array}{cccc|c} 1 & -1 & 2 & 1 & 0 \\ 2 & 1 & -1 & 1 & 0 \\ 4 & -1 & 3 & 3 & 0 \\ 1 & 2 & -3 & 0 & 0 \end{array} \right) \xrightarrow{\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - 4R_1 \\ \sim \\ R_4 \rightarrow R_4 - R_1}} \left( \begin{array}{cccc|c} 1 & -1 & 2 & 1 & 0 \\ 0 & 3 & -5 & -1 & 0 \\ 0 & 3 & -5 & -1 & 0 \\ 0 & 3 & -5 & -1 & 0 \end{array} \right) \xrightarrow{\substack{R_3 \rightarrow R_3 - R_2 \\ R_4 \rightarrow R_4 - R_2 \\ \sim}} \left( \begin{array}{cccc|c} 1 & -1 & 2 & 1 & 0 \\ 0 & 3 & -5 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

There are two free variables,  $x_3$  and  $x_4$ , and the general solution is thus of the form

$$x_3 = \alpha, \quad x_4 = \beta, \quad x_2 = \frac{1}{3}(5\alpha + \beta), \quad x_1 = x_2 - 2x_3 - x_4 = -\frac{1}{3}\alpha - \frac{2}{3}\beta.$$

Any solution can be rewritten as

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -\alpha - 2\beta \\ 5\alpha + \beta \\ 3\alpha \\ 3\beta \end{pmatrix} = \frac{\alpha}{3} \begin{pmatrix} -1 \\ 5 \\ 3 \\ 0 \end{pmatrix} + \frac{\beta}{3} \begin{pmatrix} -2 \\ 1 \\ 0 \\ 3 \end{pmatrix},$$

and we find that  $\frac{1}{3}(-1, 5, 3, 0)^T, \frac{1}{3}(-2, 1, 0, 3)^T$  is a basis of the solution space of the system.

**The span of a family of vectors.** Another important type of subspace is given by spans of families of vectors. The following generalises a result we have seen in  $\mathbb{R}^n$ :

**Theorem 5.3**

Let  $V$  be a vector space and let  $\vec{u}_1, \dots, \vec{u}_k$  be vectors of  $V$ . Then  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$  is a vector subspace of  $V$ .

*Proof.* The span contains the null vector since  $\vec{0} = 0\vec{u}_1 + \dots + 0\vec{u}_k$ . Let  $a_1\vec{u}_1 + \dots + a_k\vec{u}_k, b_1\vec{u}_1 +$

$\dots + b_k \vec{u}_k$  be vectors in  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$ , and let  $\lambda \in \mathbb{R}$ . Then we have

$$(a_1 \vec{u}_1 + \dots + a_k \vec{u}_k) + (b_1 \vec{u}_1 + \dots + b_k \vec{u}_k) = (a_1 + b_1) \vec{u}_1 + \dots + (a_k + b_k) \vec{u}_k \in \text{span}(\vec{u}_1, \dots, \vec{u}_k),$$

$$\lambda(a_1 \vec{u}_1 + \dots + a_k \vec{u}_k) = (\lambda a_1) \vec{u}_1 + \dots + (\lambda a_k) \vec{u}_k \in \text{span}(\vec{u}_1, \dots, \vec{u}_k).$$

Thus,  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$  is a subspace of  $V$ .  $\square$

### Example 5.4

- A solution of the differential equation  $y'' + \omega^2 y = 0$  is of the form  $a \cos(\omega x) + b \sin(\omega x)$  with  $a, b \in \mathbb{R}$ . Thus, the set of solutions of the differential equation  $y'' + \omega^2 y = 0$  is the span of the functions  $x \mapsto \cos(\omega x)$  and  $x \mapsto \sin(\omega x)$ . In particular, we recover the fact that it is a vector subspace of  $\mathcal{F}(\mathbb{R})$ .

**Remark.** A vector subspace can sometimes be spanned by many different sets of vectors: for instance, both the pairs of vectors  $((1, 0)^T, (0, 1)^T)$  and  $((1, 0)^T, (1, 1)^T)$  span  $\mathbb{R}^2$ . Indeed, given a vector  $(x, y)^T \in \mathbb{R}^2$ , we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and also} \quad \begin{pmatrix} x \\ y \end{pmatrix} = (x - y) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

### Definition 5.5

The *rank* of a family of vectors  $\vec{v}_1, \dots, \vec{v}_k$  is the dimension of the subspace they span.

For this type of subspaces, there is also a simple way to determine their dimension and find a basis. However, the proof of this result requires tools that will be introduced in the next chapter, so we postpone its proof for now.

Finding the dimension and a basis of the span of a family of vectors of  $\mathbb{R}^n$ . We can find a basis for  $\text{span}(\vec{v}_1, \dots, \vec{v}_k)$  as follows:

- (1) Write down a matrix whose  $i$ th column is  $\vec{v}_i$ .
- (2) Perform elementary row operations to bring the matrix into row echelon form.
- (3) The rank of  $(\vec{v}_1, \dots, \vec{v}_k)$  is the number of pivot variables in the row echelon form. If we denote by  $j_1, \dots, j_k$  the columns of the row echelon form that contain a pivot, then a basis for  $\text{span}(\vec{v}_1, \dots, \vec{v}_k)$  is given by  $\vec{v}_{j_1}, \dots, \vec{v}_{j_k}$ .

### Example 5.6

We wish to find a basis for the subspace of  $\mathbb{R}^3$  spanned by

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}.$$

We apply the previous algorithm (place the vectors as columns and row-reduce):

$$\left( \begin{array}{ccc} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{array} \right) \xrightarrow[R_3 \rightarrow R_3 - R_1]{\sim} \left( \begin{array}{ccc} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{array} \right) \xrightarrow[R_3 \rightarrow R_3 - R_2]{\sim} \left( \begin{array}{ccc} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array} \right).$$

There are two pivot columns (columns 1 and 2), so the rank of this family of vectors is 2 and the vectors span a two-dimensional subspace of  $\mathbb{R}^3$  (a plane through the origin). Hence a basis for the span is given by

$$\left( \begin{array}{c} 1 \\ 0 \\ 1 \end{array} \right), \left( \begin{array}{c} 2 \\ 1 \\ 3 \end{array} \right).$$

Moreover, since the third column is non-pivot,  $v_3$  is a linear combination of  $v_1$  and  $v_2$ ; indeed the row-reduction shows  $v_3 = v_1 + v_2$ .

Since finding a span is a useful skill, let's consider a further example

### Example 5.7

We wish to find a basis for the subspace of  $\mathbb{R}^4$  spanned by  $(1, 2, 3, 0)^T, (2, 1, 2, 1)^T$  and  $(1, 5, 7, -1)^T, (0, 0, 1, 2)^T$ . We apply the previous algorithm:

$$\left( \begin{array}{cccc} 1 & 2 & 1 & 0 \\ 2 & 1 & 5 & 0 \\ 3 & 2 & 7 & 1 \\ 0 & 1 & -1 & 2 \end{array} \right) \xrightarrow[R_2 \rightarrow R_2 - 2R_1]{\sim} \left( \begin{array}{cccc} 1 & 2 & 1 & 0 \\ 0 & -3 & 3 & 0 \\ 3 & 2 & 7 & 1 \\ 0 & 1 & -1 & 2 \end{array} \right) \xrightarrow[R_3 \rightarrow 3R_3 - 4R_2]{\sim} \left( \begin{array}{cccc} 1 & 2 & 1 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & -4 & 4 & 1 \\ 0 & 1 & -1 & 2 \end{array} \right) \xrightarrow[R_4 \rightarrow 3R_4 + R_2]{\sim} \left( \begin{array}{cccc} 1 & 2 & 1 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 2 \end{array} \right) \xrightarrow[R_4 \rightarrow -3R_4 + 2R_3]{\sim} \left( \begin{array}{cccc} 1 & 2 & 1 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

There are three pivot variables, corresponding to columns 1, 2, and 4, so the rank of this family of vectors is 3 and a basis for its span is given by  $(1, 2, 3, 0)^T, (2, 1, 2, 1)^T, (0, 0, 1, 2)^T$ .

## 6 Matrices

### 6.1 Image and rank of a matrix.

#### Definition 6.1

Given an  $m \times n$  matrix written in column form as

$$A = \left( \begin{array}{c|c|c} \vec{c}_1 & \cdots & \vec{c}_n \end{array} \right),$$

the *image* of  $A$ , denoted  $\text{Im}(A)$ , is the subspace of  $\mathbb{R}^m$  spanned by  $\vec{c}_1, \dots, \vec{c}_n$ . Its dimension is called the *rank* of  $A$  and denoted  $\text{rk}(A)$  or  $\text{rank}(A)$ .

#### Example 6.2

Consider the following matrix:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \begin{array}{ll} \text{row space: } \text{span}\{(1, 2, 3), (4, 5, 6)\}, & 2 = \text{rk}_r(A) \leq 2, \\ \text{column space: } \text{span}\{(1, 4)^T, (2, 5)^T, (3, 6)^T\}, & 2 = \text{rk}_c(A) \leq 3. \end{array}$$

The column space of  $A$  is  $\text{span}((1, 4)^T, (2, 5)^T, (3, 6)^T)$  and a basis of it (obtained by Gaussian elimination) is  $(1, 4)^T, (2, 5)^T$ .

In this example, we see that  $\text{rk}_r(A) = \text{rk}_c(A) = 2$ .

#### Theorem 6.3

The inhomogeneous system of linear equations  $A\vec{x} = \vec{b}$  is consistent if and only if  $\vec{b}$  is in the column space of  $A$ .

*Proof.*

$$\begin{aligned} \{A\vec{x} : \vec{x} \in \mathbb{R}^n\} &= \left\{ \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\} \\ &= \left\{ x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \right\} \end{aligned}$$

The last expression is the column space of  $A$  and thus consistency of the system of linear equations  $A\vec{x} = \vec{b}$  requires that  $\vec{b}$  is in the column space of  $A$ .  $\square$

**Theorem 6.4**

Elementary row operations do not change the row rank of a matrix.

Determining the row rank and a basis of the row space. We can determine a basis of the row space as follows:

- (i) Bring the matrix to row echelon form.
- (ii) A basis for the row space consists of the family of non-zero rows of the matrix in echelon form, and the row rank is the number of non-zero row in the row echelon form.

**Example 6.5**

Consider the following matrix:

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

After performing Gaussian elimination, we find the following row echelon form:

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

As the row vectors in the echelon form are linearly independent, we have that a basis of the row space of  $A$  is  $(1, 2, -1, 3)$ ,  $(0, -5, 4, -5)$ ,  $(0, 0, 1, 2)$ , and in particular  $\text{rk}_r(A) = 3$ .

Now consider determining the *column* rank and a basis of the column space. We can determine the column space of a matrix  $A$  and its column rank as follows:

- (i) Compute the transpose  $A^T$ .
- (ii) Bring  $A^T$  to row echelon form.
- (iii) The column rank of  $A^T$  is the number of non-zero rows of the echelon form of  $A^T$ , and a basis for the column space is given by the transpose of the non-zero rows of the echelon form.

**Example 6.6**

Let us consider the same matrix as in the previous example. We bring its transpose to row

echelon form:

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & -1 & 2 & 1 \\ 4 & 3 & 0 & 7 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 2 & 4 & 0 \\ 2 & -1 & 3 & 0 \\ -1 & 2 & 0 & 1 \\ 3 & 1 & 7 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 4 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus,  $((1, 2, 4, 0)^T, (0, 1, 1, 0)^T, (0, 0, 0, 1)^T)$  is a basis for the column space of  $A$  and the column rank is  $\text{rk}_c(A) = 3$ .

## 6.2 Rank of a matrix.

Again, we saw in the previous two examples that  $\text{rk}_r(A) = \text{rk}_c(A) = 3$ . This is a consequence of the following general result:

### Theorem 6.7

For an  $n \times m$  matrix  $A$ , we have

$$\text{rk}_c(A) = \text{rk}_r(A).$$

This quantity, simply denoted  $\text{rk}(A)$ , is called the *rank* of the matrix  $A$ .

*Proof.* Let  $(\vec{e}_1, \dots, \vec{e}_k)$  be a basis of the row space:  $\text{span}\{\vec{r}_1, \dots, \vec{r}_m\} = \text{span}\{\vec{e}_1, \dots, \vec{e}_k\}$ . We then have:

$$\begin{aligned} A &= \begin{pmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vdots \\ \vec{r}_m \end{pmatrix} = \begin{pmatrix} a_{11}\vec{e}_1 + \dots + a_{1k}\vec{e}_k \\ a_{21}\vec{e}_1 + \dots + a_{2k}\vec{e}_k \\ \vdots \\ a_{m1}\vec{e}_1 + \dots + a_{mk}\vec{e}_k \end{pmatrix} \\ &= \begin{pmatrix} a_{11}e_{11} + \dots + a_{1k}e_{k1} & a_{11}e_{12} + \dots + a_{1k}e_{k2} & \dots & a_{11}e_{1n} + \dots + a_{1k}e_{kn} \\ \vdots & \vdots & & \vdots \\ a_{m1}e_{11} + \dots + a_{mk}e_{k1} & a_{m1}e_{12} + \dots + a_{mk}e_{k2} & \dots & a_{m1}e_{1n} + \dots + a_{mk}e_{kn} \end{pmatrix}. \end{aligned}$$

The column space is thus spanned by  $\{(a_{11}, \dots, a_{m1})^T, \dots, (a_{1k}, \dots, a_{mk})^T\}$ . It follows that  $\text{rk}_r(A) \geq \text{rk}_c(A)$ . Interchanging rows and columns in this argument leads to  $\text{rk}_c(A) \geq \text{rk}_r(A)$ , and altogether, we have  $\text{rk}_c(A) = \text{rk}_r(A)$ .  $\square$

**Example 6.8**

$$\begin{array}{ll} \left( \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right) \text{ has rank 0} & \left( \begin{array}{ccc} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{array} \right) \text{ has rank 2} \\ \left( \begin{array}{ccc} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{array} \right) \text{ has rank 1} & \left( \begin{array}{ccc} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \text{ has rank 3} \end{array}$$

**6.3 Rank and systems of linear equations.**

Let us reinterpret some old results in term of the rank. There is nothing new in this paragraph (and you should convince yourself of it):

**Theorem 6.9**

Let  $A$  be an  $n \times n$  matrix, Then:

$$A \text{ is invertible} \iff \text{rk}(A) = n.$$

**Theorem 6.10**

Let  $A$  be an  $m \times n$  matrix. Then the system of linear equations  $A\vec{x} = \vec{b}$  is consistent for all  $\vec{b} \in \mathbb{R}^m$  if and only if  $\text{rk}(A) = m$ .

*Proof.* The statement that  $A\vec{x} = \vec{b}$  is consistent for all  $\vec{b} \in \mathbb{R}^m$  is equivalent to the fact that the image of  $A$  is  $\mathbb{R}^m$  and thus that the (column) rank of  $A$  is maximal:  $\text{rk}(A) = m$ .  $\square$

**Theorem 6.11**

Let  $A$  be an  $m \times n$ -matrix. We have:

- (i)  $\text{rk}(A) < n \Rightarrow A\vec{x} = \vec{0}$  has infinitely many solutions.
- (ii)  $\text{rk}(A) = n \Rightarrow A\vec{x} = \vec{0}$  has only one solution  $\vec{x} = \vec{0}$ .
- (iii)  $\text{rk}(A) > n$  is not possible, as the row space of  $A$  is a subspace of  $\mathbb{R}^n$ .

## 6.4 The Rank-Nullity Theorem for matrices.

### Definition 6.12

Let  $A$  be an  $m \times n$  matrix. The *kernel* (also called the *nullspace*) of  $A$  is the subspace of  $\mathbb{R}^n$  consisting of the solutions of the system of linear equations

$$A\vec{x} = \vec{0}.$$

Its dimension is called the *nullity* of  $A$ .

### Theorem 6.13

Let  $A$  be an  $m \times n$  matrix. Then we have:

$$\text{rk}(A) + \dim \ker(A) = n.$$

*Proof.* Performing Gaussian elimination, we obtain a matrix with exactly  $\text{rk}(A)$  pivots. In particular,  $\text{rk}(A)$  of the  $n$  variables are pivot variables, while the other  $n - \text{rk}(A)$  are free variables. But we know from previous results that the rank of a matrix is the number of pivot variables in any row echelon form, while the nullity of a matrix is the number of free variables in any row echelon form.  $\square$

We have already seen a rather different looking definition of the rank of a matrix in F17ZB.

### Definition 6.14 Rank of a matrix

The rank of a matrix  $A$  of size  $m \times n$  is the smallest  $k \leq \min(m, n)$  such that there exist vectors  $\vec{u}_\ell \in \mathbb{R}^m$  and  $\vec{v}_\ell \in \mathbb{R}^n$  with

$$A = \sum_{\ell=1}^k \vec{u}_\ell \vec{v}_\ell^T.$$

Matrix  $A$  is said to be if full rank of  $k = \min(m, n)$ . Since

$$A^T = \sum_{\ell=1}^k \vec{v}_\ell \vec{u}_\ell^T$$

the rank of  $A$  and  $A^T$  are the same.

Let's take the same example used there and see how Gaussian Elimination gives the same value for the rank.

**Example 6.15**

In F17ZB we considered the matrix

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

and stated that this was a rank-2 matrix (not obvious at all from just looking at it!) since

$$A = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} (1 \ -1 \ 1 \ 1) + \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} (1 \ 2 \ -1 \ 1.)$$

**Example 6.16**

To determine the *column rank* of

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

We row-reduce (row operations do not change linear relations among the columns):

$$\left( \begin{array}{cccc} 2 & 1 & 0 & 2 \\ 3 & 3 & -1 & 3 \\ 4 & 5 & -2 & 4 \\ 5 & 7 & -3 & 5 \end{array} \right) \xrightarrow{\substack{R_2 \leftarrow 2R_2 - 3R_1 \\ R_3 \leftarrow R_3 - 2R_1 \\ R_4 \leftarrow 2R_4 - 5R_1}} \left( \begin{array}{cccc} 2 & 1 & 0 & 2 \\ 0 & 3 & -2 & 0 \\ 0 & 3 & -2 & 0 \\ 0 & 9 & -6 & 0 \end{array} \right) \xrightarrow{\substack{R_3 \leftarrow R_3 - R_2 \\ R_4 \leftarrow R_4 - 3R_2}} \left( \begin{array}{cccc} 2 & 1 & 0 & 2 \\ 0 & 3 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

There are exactly two nonzero rows in echelon form, hence

$$\text{rank}(A) = 2.$$

The pivot columns are columns 1 and 2, so a basis for the column space is given by the corresponding original columns:

$$\mathcal{B}_{\text{col}(A)} = \left\{ \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \end{pmatrix} \right\}.$$

Moreover, the remaining columns are dependent:

$$\mathbf{c}_4 = \mathbf{c}_1, \quad \mathbf{c}_3 = \frac{1}{3} \mathbf{c}_1 - \frac{2}{3} \mathbf{c}_2,$$

so all columns lie in the span of  $\mathbf{c}_1, \mathbf{c}_2$ .

**Example 6.17**

Let

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

From row-reduction we know that  $\text{rank}(A) = 2$  and that the pivot columns are the first two columns. Set

$$\mathbf{c}_1 = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \quad \mathbf{c}_2 = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \end{pmatrix}, \quad C = (\mathbf{c}_1 \ \mathbf{c}_2) \in \mathbb{R}^{4 \times 2}.$$

Moreover, the remaining columns are linear combinations of  $\mathbf{c}_1, \mathbf{c}_2$ :

$$\mathbf{c}_3 = \frac{1}{3}\mathbf{c}_1 - \frac{2}{3}\mathbf{c}_2, \quad \mathbf{c}_4 = \mathbf{c}_1.$$

Therefore, if we record these coefficients column-by-column, we obtain the  $2 \times 4$  coefficient matrix

$$R = \begin{pmatrix} 1 & 0 & \frac{1}{3} & 1 \\ 0 & 1 & -\frac{2}{3} & 0 \end{pmatrix}.$$

Then  $A$  factors as

$$A = CR = (\mathbf{c}_1 \ \mathbf{c}_2) \begin{pmatrix} 1 & 0 & \frac{1}{3} & 1 \\ 0 & 1 & -\frac{2}{3} & 0 \end{pmatrix}.$$

Equivalently, writing  $R$  in terms of its rows

$$\mathbf{r}_1^T = \left( 1 \ 0 \ \frac{1}{3} \ 1 \right), \quad \mathbf{r}_2^T = \left( 0 \ 1 \ -\frac{2}{3} \ 0 \right),$$

we obtain a decomposition of  $A$  as a sum of two outer products (a sum of rank-1 matrices):

$$A = \mathbf{c}_1 \mathbf{r}_1^T + \mathbf{c}_2 \mathbf{r}_2^T.$$

Since this is a sum of two rank-1 matrices, it also makes it transparent that  $\text{rank}(A) \leq 2$  (and in fact  $\text{rank}(A) = 2$  because  $\mathbf{c}_1, \mathbf{c}_2$  are independent).

The decomposition

$$A = \sum_{k=1}^2 \mathbf{u}_k \mathbf{v}_k^T$$

is highly non-unique in general.

However, if you fix the left vectors to be a specific basis of the column space (e.g., the first two columns), then the corresponding right vectors are unique (given that choice).

**Exercise 6.18**

Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 2 & 4 & 6 \end{pmatrix}.$$

- (a) Use *row operations* to determine a set of *independent columns* of  $A$  and hence compute  $\text{rank}(A)$ .
- (b) Compute  $\text{rank}(A)$  by writing  $A$  as a sum of outer products (equivalently, a rank factorisation  $A = CR$ ).

**Solution:** (a) **Row operations to find independent columns.** Row-reduce  $A$  to echelon form:

$$\left( \begin{array}{ccc} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 2 & 4 & 6 \end{array} \right) \xrightarrow{\substack{R_2 \leftarrow R_2 - R_1 \\ R_3 \leftarrow R_3 - 2R_1}} \left( \begin{array}{ccc} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{array} \right).$$

The leading (pivot) entries occur in columns 1 and 2. Therefore the *corresponding original columns* of  $A$  form a basis for the column space:

$$\mathbf{c}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{c}_2 = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}.$$

Hence the column rank is the number of pivot columns:

$$\text{rank}(A) = 2.$$

(b) **Outer-product / rank factorisation.** Let the columns of  $A$  be  $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ . From the matrix we have

$$\mathbf{c}_3 = \begin{pmatrix} 3 \\ 1 \\ 6 \end{pmatrix} = -\mathbf{c}_1 + 2\mathbf{c}_2,$$

so every column of  $A$  lies in  $\text{span}\{\mathbf{c}_1, \mathbf{c}_2\}$ .

Define

$$C = (\mathbf{c}_1 \ \mathbf{c}_2) = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 4 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Then the columns of  $CR$  are

$$CR_{(:,1)} = \mathbf{c}_1, \quad CR_{(:,2)} = \mathbf{c}_2, \quad CR_{(:,3)} = -\mathbf{c}_1 + 2\mathbf{c}_2 = \mathbf{c}_3,$$

so  $A = CR$ . Writing  $R$  by rows,

$$\mathbf{r}_1^T = (1 \ 0 \ -1), \quad \mathbf{r}_2^T = (0 \ 1 \ 2),$$

we obtain the outer-product decomposition

$$A = \mathbf{c}_1 \mathbf{r}_1^T + \mathbf{c}_2 \mathbf{r}_2^T.$$

This expresses  $A$  as a sum of two rank-1 matrices, so  $\text{rank}(A) \leq 2$ . Since  $\mathbf{c}_1$  and  $\mathbf{c}_2$  are not multiples of each other, they are independent, so  $\text{rank}(A) \geq 2$ . Therefore,

$$\text{rank}(A) = 2.$$

## 7 Eigenvalues and diagonalisability

### 7.1 A motivating problem: predicting the dynamics of a population

A simple model to study a population of animals is the following: the population is divided into two age groups: adults and juveniles. At the start of the observation, the population consists of  $a_0$  adults and  $j_0$  juveniles. From one year to the next, the adults will each produce on average  $\gamma$  juveniles. Adults will survive from one year to the next with probability  $\alpha$ , while juveniles will survive into the next year and become adults with the probability  $\beta$ .

We wish to understand the evolution of the population over time: does the population collapse, does it converge to a stable state? And how does this behaviour depend on the fertility rate  $\gamma$ ? Notice that we have the following equations between the population in two consecutive years:

$$\begin{aligned} a_{n+1} &= \alpha a_n + \beta j_n \\ j_{n+1} &= \gamma a_n \end{aligned}$$

This can be rewritten as

$$\begin{pmatrix} a_{n+1} \\ j_{n+1} \end{pmatrix} = A \begin{pmatrix} a_n \\ j_n \end{pmatrix} \quad \text{with } A := \begin{pmatrix} \alpha & \beta \\ \gamma & 0 \end{pmatrix}.$$

In particular, we see by induction that

$$\begin{pmatrix} a_n \\ j_n \end{pmatrix} = A^n \begin{pmatrix} a_0 \\ j_0 \end{pmatrix}.$$

To understand the population at a given time, we thus need to compute powers of the matrix  $A$ . This is a priori a non-trivial problem. There is however one case where computing powers poses no problem: the case of diagonal matrices.

A natural strategy would be to try to find a ‘simplest possible basis’ where the matrix  $A$  becomes diagonal, compute the powers of the matrix in that basis, and go back to the original basis. If we restate this problem in terms of matrices, we want to find an invertible matrix  $P$  and a diagonal matrix  $D$  such that  $A = PDP^{-1}$ . We have:

$$A^n = (PDP^{-1})^n = PD \underbrace{P^{-1}P}_{=I_2} D P^{-1} \cdots \underbrace{P^{-1}P}_{=I_2} D P^{-1} = PD^n P^{-1}.$$

Since  $D^n$  is very easy to compute, it follows that computing  $A^n$  itself becomes much easier to compute, provided we know how to compute  $P$ . This leads to the following questions:

- Given a square matrix  $A$ , does there always exist an invertible matrix  $P$  such that  $P^{-1}AP$  is diagonal?
- If so, how to compute such a matrix  $P$ ?

The goal of this chapter is to answer these questions and see applications to various problems.

## 7.2 Eigenvectors and Eigenvalues

### Definition 7.1

Let  $V$  be an  $n$ -dimensional vector space and let  $T : V \rightarrow V$  be a linear map. Then  $\lambda \in \mathbb{R}$  is called a (real) *eigenvalue* of  $T$ , if there is a vector  $\vec{x} \in V$ ,  $\vec{x} \neq \vec{0}$ , such that

$$T(\vec{x}) = \lambda \vec{x}.$$

The vector  $\vec{x}$  is called an *eigenvector* of  $T$  for the eigenvalue  $\lambda$ .

Recall that an  $n \times n$  matrix  $A$  can be seen as a linear map  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\vec{x} \mapsto A\vec{x}$ . In particular, the notions of eigenvalues and eigenvectors are well defined for matrices:  $\lambda \in \mathbb{R}$  is an *eigenvalue* of  $A$ , if there is a vector  $\vec{x} \in \mathbb{R}^n$ ,  $\vec{x} \neq \vec{0}$ , such that  $A\vec{x} = \lambda \vec{x}$ , and the vector  $\vec{x}$  is called an *eigenvector* of  $A$  for the eigenvalue  $\lambda$ .

### Example 7.2

- (i) For  $A = I_n$  the identity matrix, every vector  $\vec{x} \neq \vec{0}$  is an eigenvector with eigenvalue 1.
- (ii) The rotation of angle  $\pi/2$ ,  $r_{\pi/2} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  does not have any eigenvalue, since  $r_{\pi/2}(\vec{x})$  is never collinear to  $\vec{x}$  (but is orthogonal to it and of the same norm).

### Example 7.3

Consider the following diagonal matrix:

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}, \quad d_1, \dots, d_n \in \mathbb{R}.$$

Then the eigenvalues of  $D$  are  $d_1, \dots, d_n$ . Indeed, for a vector  $\vec{x} = (x_1, \dots, x_n)^T$ , we have  $D\vec{x} = (d_1 x_1, \dots, d_n x_n)^T$ . If  $\vec{x}$  is an eigenvector of  $D$  for an eigenvalue  $\lambda$ , then we also have  $D\vec{x} = (\lambda x_1, \dots, \lambda x_n)^T$ . As one of the  $x_i$  is non-zero (since  $\vec{x} \neq \vec{0}$ ), it follows that  $\lambda = d_i$  for some  $i$ .

Note that if the  $d_1, \dots, d_n$  are pairwise distinct, then the eigenvectors associated to  $d_1, \dots, d_n$  are of the form  $(c_1, 0, \dots, 0)^T, \dots, (0, \dots, 0, c_n)^T$  respectively, where  $c_1, \dots, c_n \in \mathbb{R}$ .

### 7.3 Diagonalisability.

#### Definition 7.4

A linear map  $T : V \rightarrow V$  is called *diagonalisable* if there exists a basis of eigenvectors. Equivalently,  $T$  is diagonalisable if there exists a basis  $\mathcal{B}$  of  $V$  such that the associated matrix  $[T]_{\mathcal{B}}$  is of the form:

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

#### Example 7.5

- (i) The rotation  $r_{\pi/4} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is not diagonalisable as it has no (real) eigenvalue.
- (ii) The orthogonal projection  $s_{\pi/4} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is diagonalisable as the matrix of  $s_{\pi/4}$  in the basis  $\mathcal{B}_{\pi/4}$  is

$$[s_{\pi/4}]_{\mathcal{B}_{\pi/4}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

- (iii) The map

$$T : \mathcal{P}_n(\mathbb{R}) \rightarrow \mathcal{P}_n(\mathbb{R}), \quad P(x) \mapsto xP'(x)$$

is diagonalisable. Indeed, in the basis  $\mathcal{B} = (1, x, \dots, x^n)$ , we have

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 2 & \\ & & & \ddots \\ & & & n \end{pmatrix}.$$

By seeing an  $n \times n$  matrix as a linear map  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , we are led to the following definition:

#### Definition 7.6

An  $n \times n$  matrix  $A$  is *diagonalisable* (over  $\mathbb{R}$ ) if there exists a basis of  $\mathbb{R}^n$  made of eigenvectors of  $A$ .

Equivalently,  $A$  is diagonalisable if there exists an invertible  $n \times n$  matrix  $P$  such that  $P^{-1}AP$  is of the form:

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

**Remark.** The two definitions of diagonalisability for maps and matrices are compatible: a linear map  $T$  is diagonalisable if and only if the associated matrix  $[T]_{\mathcal{B}}$  is diagonalisable for some (hence every) basis  $\mathcal{B}$ .

**Distinct eigenvalues and linear independence.** Here is an useful criterion:

**Theorem 7.7**

An  $n \times n$  matrix has at most  $n$  eigenvalues.

This theorem is a direct consequence of the following result, together with the fact that a linearly independent family in an  $n$ -dimensional vector space contains at most  $n$  vectors.

**Theorem 7.8**

Let  $\vec{x}_1, \dots, \vec{x}_k$  be eigenvectors of a matrix  $A$  corresponding to distinct eigenvalues  $\lambda_1, \dots, \lambda_k$ . Then  $\vec{x}_1, \dots, \vec{x}_k$  are linearly independent.

*Proof.* Assume that we have constants  $c_1, \dots, c_k$  such that  $c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_k\vec{u}_k = \vec{0}$ . Let us denote this equation  $(E_1)$ . By applying  $A$  on both sides of the equation, we get  $c_1\lambda_1\vec{u}_1 + c_2\lambda_2\vec{u}_2 + \dots + c_k\lambda_k\vec{u}_k = \vec{0}$ . Let us denote this equation  $(E_2)$ . Now by considering  $(E_2) - \lambda_1(E_1)$ , we get the equation  $c_2(\lambda_2 - \lambda_1)\vec{u}_2 + \dots + c_k(\lambda_k - \lambda_1)\vec{u}_k = \vec{0}$ . We are thus back to an other linear combination of fewer eigenvectors, and we can now prove the result by induction.  $\square$

Here is a useful corollary:

**Theorem 7.9**

Let  $V$  be an  $n$ -dimensional vector space, and let  $T : V \rightarrow V$  be a linear map that has exactly  $n$  distinct real eigenvalues  $\lambda_1, \dots, \lambda_n$ , then  $T$  is diagonalisable.

**Remark.** The previous result is not an equivalence, there exist diagonalisable maps that have fewer than  $n$  eigenvalues. For instance, the identity map is diagonalisable, but its only eigenvalue is 1.

## 7.4 Eigenspaces and diagonalisability.

**Definition 7.10**

Given an eigenvalue  $\lambda \in \mathbb{R}$ , the associated *eigenspace* is the following subset:

$$E_\lambda(T) := \{\vec{x} \in U \mid T(\vec{x}) = \lambda\vec{x}\} = \{\vec{0}\} \cup \{\vec{x} \in U \mid \vec{x} \text{ is an eigenvector for } \lambda\}.$$

**Theorem 7.11**

An eigenspace is a subspace of  $V$ .

*Proof.* By definition, we have that  $E_\lambda(T) = \ker(T - \lambda Id)$ , and we know that kernels of linear maps are vector subspaces.  $\square$

The proof of the following result is a slight variation on the proof of Proposition 7.3:

**Theorem 7.12**

Let  $\lambda_1, \dots, \lambda_k$  be distinct eigenvalues of a linear map  $T$ . Let  $\mathcal{B}_1, \dots, \mathcal{B}_k$  be bases of the eigenspaces  $E_{\lambda_1}(T), \dots, E_{\lambda_k}(T)$  respectively. Then  $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$  is a linearly independent family of vectors.

**Theorem 7.13**

Let  $T : V \rightarrow V$  be a linear map and let  $\lambda_1, \dots, \lambda_k$  be the distinct eigenvalues of  $T$ . Then  $T$  is diagonalisable if and only if:

$$\dim E_{\lambda_1}(T) + \dots + \dim E_{\lambda_k}(T) = n.$$

The next natural question to answer is: how do we find the eigenvalues of a given linear map or matrix?

## 7.5 Characteristic polynomial

Finding an eigenvalue  $\lambda$  and an associated eigenvector  $\vec{x}$  amounts to finding solutions to the equation  $T(\vec{x}) = \lambda \vec{x}$ , or equivalently  $(T - \lambda I_n)\vec{x} = \vec{0}$ . Up to choosing coordinates, we can assume that  $T$  is represented by an  $n \times n$  matrix  $A$ . The problem thus amounts to solving a linear system of equations  $A\vec{x} = \lambda \vec{x}$ , or equivalently  $(A - \lambda I_n)\vec{x} = \vec{0}$ . This is a homogeneous system of  $n$  equations in  $n$  unknowns. It has a non-trivial solution if and only if  $(A - \lambda I_n)$  is not invertible, and this is equivalent to  $\det(A - \lambda I_n) = 0$ . This motivates the following definition:

**Definition 7.14**

Given an  $n \times n$ -matrix  $A$ , the expression

$$\chi_A(\lambda) := \det(A - \lambda I_n) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}$$

is a polynomial of degree  $n$  in  $\lambda$ , called the *characteristic polynomial* of the matrix  $A$ .

**Remark.** In this course, we will mostly compute characteristic polynomials for  $2 \times 2$  and  $3 \times 3$  matrices. When computing characteristic polynomials in general, one computes this determinant using the expansion technique.

The definition of the characteristic polynomial is motivated by the following result:

### Theorem 7.15

Given an  $n \times n$ -matrix  $A$ , we have:

$$\lambda \text{ is an eigenvalue of } A \Leftrightarrow \chi_A(\lambda) = 0.$$

*Proof.* We have:

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A &\Leftrightarrow \exists \vec{x} \neq \vec{0} \text{ such that } A\vec{x} = \lambda\vec{x}. \\ &\Leftrightarrow \exists \vec{x} \neq \vec{0} \text{ such that } (A - \lambda I_n)\vec{x} = \vec{0}. \\ &\Leftrightarrow \ker(A - \lambda I_n) \neq \{\vec{0}\}. \\ &\Leftrightarrow A - \lambda I_n \text{ is not invertible.} \\ &\Leftrightarrow \det(A - \lambda I_n) = 0 \quad \square \end{aligned}$$

### Example 7.16

Let us find the eigenvalues of the matrix

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

The characteristic polynomial of  $A$  reads as (rule of Sarrus)

$$\det \begin{pmatrix} 3 - \lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3 - \lambda \end{pmatrix} = -(3 - \lambda)^2\lambda + 16 + 16 + 16\lambda - 4(3 - \lambda) - 4(3 - \lambda),$$

or  $\chi_A(\lambda) = -\lambda^3 + 6\lambda^2 + 15\lambda + 8$ . We see that  $\lambda = -1$  is a root of this polynomial and factorise it as  $-(\lambda + 1)(\lambda + 1)(\lambda - 8)$ . That is, the eigenvalues of  $A$  are  $-1$  and  $8$ .

**The characteristic polynomial seen as a polynomial over  $\mathbb{C}$ .** By the fundamental theorem of algebra, the characteristic polynomial of an  $n \times n$  matrix  $A$  can be factorised into  $n$  linear factors:

$$\chi_A(\lambda) = \pm(\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n),$$

where  $\lambda_i$  are (not necessarily distinct) *complex*. These  $\lambda_i$  are called the *complex eigenvalues* of  $A$ , and the eigenvalues we defined so far are the complex eigenvalues that are real. Since a

polynomial of degree  $n \geq 1$  has at most  $n$  roots, we recover the fact that  $A$  has at most  $n$  complex eigenvalues, hence at most  $n$  real eigenvalues.

**Remark.** A matrix may have no real eigenvalues, e.g.

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = (\lambda + i)(\lambda - i) = \chi_A(\lambda)$$

has a characteristic polynomial  $\chi_A(\lambda)$  without real roots. However, since a polynomial of degree  $\geq 1$  always has at least one root by the fundamental theorem of algebra,  $A$  always has at least one complex eigenvalue.

There is one simple case where the eigenvalues can be read directly from a matrix:

### Theorem 7.17 L

et

$$A = \begin{pmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & d_n \end{pmatrix} \in M_n(\mathbb{R})$$

be an upper-triangular  $n \times n$  matrix. Then the eigenvalues of  $A$  are  $d_1, \dots, d_n$ .

The same result holds for lower-triangular matrices.

*Proof.*  $\chi_A(x) = (d_1 - x) \cdots (d_n - x)$ , and the result follows from Theorem 7.5.  $\square$

When computing the row echelon form of a matrix via Gaussian elimination, the diagonal entries in the echelon form are **not** the eigenvalues of  $A$  in general.

## 7.6 Trace, determinants, and consistency check.

### Definition 7.18

The *trace* of a square matrix  $A$ , denoted  $\text{tr}(A)$ , is the sum of its diagonal coefficients.

We have the following properties:

- The trace defines a linear map  $\text{tr} : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ ,  $A \mapsto \text{tr}(A)$ .
- We have  $\text{tr}(AB) = \text{tr}(BA)$  for every pair of matrices  $A, B \in M_n(\mathbb{R})$ .

In particular, we get that for every invertible matrix  $P$ ,  $\text{tr}(PAP^{-1}) = \text{tr}(AP^{-1}P) = \text{tr}(A)$ . Similarly, we have  $\det(PAP^{-1}) = \det(P)\det(A)\det(P)^{-1} = \det(A)$ . In other words, the trace and the determinant remain the same when changing basis. If  $A$  is diagonalisable, then these

quantities can be computed in terms of the eigenvalues of  $A$  in an appropriate basis. This leads to the following useful result.

### Theorem 7.19 L

Let  $A$  be a diagonalisable  $n \times n$  matrix, and let  $\lambda_1, \dots, \lambda_n$  be its eigenvalues counted with multiplicity, i.e. the roots of  $\chi_A$  counted with multiplicity. Then we have

$$\det(A) = \lambda_1 \times \dots \times \lambda_n.$$

$$\text{tr}(A) = \lambda_1 + \dots + \lambda_n.$$

This result is useful to check your computations, as the trace in particular is very easy to compute. Try to always use this consistency check!

### Example 7.20

Let us go back to the matrix

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$

whose eigenvalues we computed previously:  $-1, -1, 8$ . In particular, we get that the sum of the eigenvalues is 6, which does indeed coincide with  $\text{tr}(A) = 3 + 0 + 3$ . Similarly, the product of the eigenvalues is 8, which we recover from Sarrus' rule. We can thus be reasonably confident that we did not make mistakes in computing the eigenvalues of  $A$ .

## 7.7 Diagonalisability of a matrix

### Definition 7.21

Let  $A \in M_n(\mathbb{R})$  and let  $\lambda$  be an eigenvalue of  $A$ .

- the *geometric multiplicity* of  $\lambda$  is the dimension of the corresponding eigenspace  $E_\lambda(A)$ .
- the *algebraic multiplicity* of  $\lambda$  is number of factors  $(x - \lambda)$  that appear in  $\chi_A$ .

### Example 7.22

Let us consider again the matrix

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

We have  $\chi_A(x) = (8 - x)(x + 1)^2$ , so the eigenvalues of  $A$  are  $-1$  and  $8$ , and their algebraic multiplicity are 2 and 1 respectively.

Let us compute the dimension of the corresponding eigenspaces using Gaussian elimination.

For the eigenvalue 8, we obtain the system:

$$\left\{ \begin{array}{lcl} 3x + 2y + 4z & = & 8x \\ 2x + 2z & = & 8y \\ 4x + 2y + 3z & = & 8z \end{array} \right. \rightsquigarrow \left\{ \begin{array}{lcl} 2x - 8y + 2z & = & 0 \\ -5x + 2y + 4z & = & 0 \\ 4x + 2y - 5z & = & 0 \end{array} \right.$$

We use Gaussian elimination:

$$\left( \begin{array}{ccc|c} 1 & -4 & 1 & 0 \\ -5 & 2 & 4 & 0 \\ 4 & 2 & -5 & 0 \end{array} \right) \rightsquigarrow \left( \begin{array}{ccc|c} 1 & -4 & 1 & 0 \\ 0 & -18 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

The row echelon form has exactly one free variable, so the dimension of the corresponding eigenspace is 1, so the geometric dimension of the eigenvalue  $\lambda = 8$  is 1.

For the eigenvalue  $\lambda = -1$ , we get the system of equations

$$\left\{ \begin{array}{lcl} 3x + 2y + 4z & = & -x \\ 2x + 2z & = & -y \\ 4x + 2y + 3z & = & -z \end{array} \right. \rightsquigarrow \left\{ \begin{array}{lcl} 4x + 2y + 4z & = & 0 \\ 2x + y + 2z & = & 0 \\ 4x + 2y + 4z & = & 0 \end{array} \right.$$

All of these equations are multiples of  $2x + y + 2z = 0$ , so the system boils down to the equation  $2x + y + 2z = 0$ . This system has exactly two free variables ( $y$  and  $z$ ), and so the dimension of the corresponding eigenspace is 2. In other words, the geometric multiplicity of the eigenvalue  $\lambda = -1$  is 2.

In particular, we see in this particular example that geometric and algebraic multiplicity coincide.

### Example 7.23

Consider the following matrix:

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

Since  $A$  is upper-triangular, we have  $\chi_A(x) = (1-x)^3$ . Thus,  $\lambda = 1$  is the unique eigenvalue of  $A$  and its algebraic multiplicity is 3.

We find the geometric multiplicity by solving the system  $(A - 1 \times I_3)\vec{x} = \vec{0}$ , which yields the system:

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

This system is already in row echelon form and has exactly one free variable. So the geometric multiplicity of  $\lambda = 1$  is 1.

In this particular example, the algebraic and geometric multiplicity differ.

For every eigenvalue of  $A$ , the geometric multiplicity is less than or equal to the algebraic multiplicity.

### Theorem 7.24

An  $n \times n$  matrix  $A$  is diagonalisable (over  $\mathbb{R}$ ) if and only if:

- all the complex roots of  $\chi_A$  are real numbers.
- for every eigenvalue  $\lambda$  of  $A$ , the algebraic multiplicity of  $\lambda$  equals the geometric multiplicity of  $\lambda$ .

### Example 7.25

Here are examples of matrices which are not diagonalisable:

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

We have  $\chi_A(x) = x^2 + 1$ , so the matrix  $A$  has no real eigenvalue (but two complex eigenvalues:  $i$  and  $-i$ ), and therefore cannot be diagonalised as a real matrix.

We have  $\chi_B(x) = (x - 1)^2$ , so the only eigenvalue of the matrix  $B$  is 1 (note that  $B$  is upper triangular), which has algebraic multiplicity 2, but the associated eigenspace has dimension 1, so in particular the algebraic and geometric multiplicity do not coincide. Thus,  $B$  is not diagonalisable over  $\mathbb{R}$ .

### Example 7.26

Let us continue the study of the matrix  $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$ . We already know that  $\chi_A(x) = (8 - x)(x + 1)^2$ ,

so the eigenvalues of  $A$  are  $-1$  and  $8$ , and we showed in previous examples that the algebraic and geometric multiplicity coincide for each eigenvalue. By the previous theorem,  $A$  is diagonalisable.

We have the following method to determine whether a given matrix is diagonalisable (over  $\mathbb{R}$ ):

- (1) Compute the characteristic polynomial  $\chi_A$  of  $A$  and compute its roots. For  $n > 3$ , you will generally be given some eigenvalues to help you with the computations.
- (2) If  $\chi_A$  possesses a non-real complex root, then  $A$  is not diagonalisable over  $\mathbb{R}$ .
- (3) Otherwise, let  $\lambda_1, \dots, \lambda_k$  be the (real) roots of  $\chi_A$ . Compute the algebraic multiplicity of each eigenvalue by factorising  $\chi_A$ .
- (4) For each eigenvalue  $\lambda_i$ , compute the geometric multiplicity of  $\lambda_i$  by computing the dimension of the solution space of the system  $A\vec{x} - \lambda_i\vec{x} = \vec{0}$  using Gaussian elimination.

- (5) If for some eigenvalue  $\lambda$ , the geometric and algebraic multiplicities of  $\lambda$  differ, then  $A$  is not diagonalisable.
- (6) To find a basis of eigenvectors, compute for each eigenvalue  $\lambda_i$  a basis  $\mathcal{B}_i$  of the corresponding eigenspace, using Gaussian elimination. A basis of eigenvectors is then given by  $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ .

In the previous theorem, it is enough to check that algebraic and geometric multiplicity coincide for eigenvalues of algebraic multiplicity at least 2, as the algebraic and geometric multiplicity necessarily coincide for eigenvalues of algebraic multiplicity 1

### Example 7.27

Let us finish the study of the matrix  $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$ . We already know that  $\chi_A(x) = (8-x)(x+1)^2$ , so the eigenvalues of  $A$  are  $-1$  and  $8$ , and we showed in previous examples that the algebraic and geometric multiplicity coincide for each eigenvalue. By the previous theorem,  $A$  is diagonalisable.

Let us compute a basis of eigenvectors. For the eigenvalue  $8$ , we solve the system:

$$\left\{ \begin{array}{lcl} 3x + 2y + 4z & = & 8x \\ 2x + 2z & = & 8y \\ 4x + 2y + 3z & = & 8z \end{array} \right. \rightsquigarrow \left\{ \begin{array}{lcl} 2x - 8y + 2z & = & 0 \\ -5x + 2y + 4z & = & 0 \\ 4x + 2y - 5z & = & 0 \end{array} \right.$$

We use Gaussian elimination:

$$\left( \begin{array}{ccc|c} 1 & -4 & 1 & 0 \\ -5 & 2 & 4 & 0 \\ 4 & 2 & -5 & 0 \end{array} \right) \xrightarrow{\substack{R_2 \rightarrow R_2 + 5R_1 \\ R_3 \rightarrow R_3 - 4R_1}} \left( \begin{array}{ccc|c} 1 & -4 & 1 & 0 \\ 0 & -18 & 9 & 0 \\ 0 & 18 & -9 & 0 \end{array} \right) \xrightarrow{\substack{R_3 \rightarrow R_2 + R_3 \\ R_2 \rightarrow R_2 / 9}} \left( \begin{array}{ccc|c} 1 & -4 & 1 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

By substitution, the solutions of this homogeneous system are  $\vec{x} = (\alpha, \frac{1}{2}\alpha, \alpha)^T$ , so a basis for  $E_8(A)$  is  $\mathcal{B}_8 = \{(1, \frac{1}{2}, 1)^T\}$ .

For the eigenvalue  $\lambda = -1$ , we get the system of equations

$$\left\{ \begin{array}{lcl} 3x + 2y + 4z & = & -x \\ 2x + 2z & = & -y \\ 4x + 2y + 3z & = & -z \end{array} \right. \rightsquigarrow \left\{ \begin{array}{lcl} 2x + y + 2z & = & 0 \\ 0 & = & 0 \\ 0 & = & 0 \end{array} \right.$$

By substitution, we find that the solution to this system of linear equations is of the form  $(-\frac{1}{2}(2\alpha + \beta), \beta, \alpha)^T$  for  $\alpha, \beta \in \mathbb{R}$ . Thus, a basis of  $E_{-1}(A)$  is given by  $\mathcal{B}_{-1} = \{(-1, 0, 1)^T, (-\frac{1}{2}, 1, 0)^T\}$ . Finally, a basis of eigenvectors is  $\mathcal{B} = \{(1, \frac{1}{2}, 1)^T, (-1, 0, 1)^T, (-\frac{1}{2}, 1, 0)^T\}$ .

Seeing  $A$  as a linear map  $T_A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , let us perform a change of basis. Let

$$P := \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

be the matrix of change of basis from the standard basis to our basis of eigenvectors. In particular, the matrix of  $T_A$  in the basis of eigenvectors  $\mathcal{B}$  is:

$$P^{-1}AP = \begin{pmatrix} 8 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

## 7.8 Application: Dynamical systems

### Definition 7.28

The *dynamical system* associated to an  $n \times n$  matrix  $A$  and a choice of initial condition  $\vec{x}_0 \in \mathbb{R}^n$  is the sequence  $(\vec{x}_k)_{k \geq 0}$  of vectors of  $\mathbb{R}^n$  defined by:

$$\vec{x}_{k+1} = A\vec{x}_k.$$

In particular, we get by induction that:

$$\vec{x}_k = A^k \vec{x}_0.$$

### Theorem 7.29

Let  $A$  be a diagonalisable  $n \times n$  matrix with a chosen basis of eigenvectors  $\vec{e}_1, \dots, \vec{e}_n$  and corresponding eigenvalues  $\lambda_1, \dots, \lambda_n$ . Let  $(\vec{x}_k)$  be the dynamical system associated to a given initial condition  $\vec{x}_0 \in \mathbb{R}^n$ . Let  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$  such that

$$\vec{x}_0 = \alpha_1 \vec{e}_1 + \dots + \alpha_n \vec{e}_n.$$

Then we have for every  $k \geq 0$ ,

$$\vec{x}_k = A^k \vec{x}_0 = \alpha_1 \lambda_1^k \vec{e}_1 + \dots + \alpha_n \lambda_n^k \vec{e}_n.$$

To determine the asymptotic behaviour (convergence, divergence, possible limit) of a dynamical system associated to a  $n \times n$  diagonalisable matrix  $A$ :

- (1) Compute a basis of eigenvectors  $\vec{e}_1, \dots, \vec{e}_n$  of  $A$ , with corresponding eigenvalues  $\lambda_1, \dots, \lambda_n$ , as in the previous method.
- (2) Using Gaussian elimination, find constants  $\alpha_1, \dots, \alpha_n$  such that

$$\vec{x}_0 = \alpha_1 \vec{e}_1 + \dots + \alpha_n \vec{e}_n.$$

- (3) We then have

$$\vec{x}_k = A^k \vec{x}_0 = \alpha_1 \lambda_1^k \vec{e}_1 + \dots + \alpha_n \lambda_n^k \vec{e}_n.$$

This formula can be used to determine the asymptotic behaviour of the dynamical system.

## 7.9 Application: Dynamics of populations.

Let us go back to the problem of understanding the dynamics of some animal population. We have:

$$\begin{pmatrix} a_{n+1} \\ j_{n+1} \end{pmatrix} = A \begin{pmatrix} a_n \\ j_n \end{pmatrix} \quad \text{with } A := \begin{pmatrix} \alpha & \beta \\ \gamma & 0 \end{pmatrix}.$$

For simplicity and concreteness, assume that  $\alpha = \frac{1}{2}$ ,  $\beta = \frac{3}{10}\gamma$ . The characteristic polynomial of  $A$  is:

$$\chi_A(x) = \left(\frac{1}{2} - x\right)(-x) - \frac{3}{10}\gamma = x^2 - \frac{1}{2}x - \frac{3}{10}\gamma.$$

The eigenvalues of  $A$  are  $\lambda_1 = \frac{1}{20}(5 - \sqrt{5}\sqrt{5 + 24\gamma})$  and  $\lambda_2 = \frac{1}{20}(5 + \sqrt{5}\sqrt{5 + 24\gamma})$ . In particular,  $A$  has two distinct eigenvalues  $\lambda_1 < \lambda_2$ , hence it is diagonalisable by Theorem 7.3. We can thus find a basis  $\vec{e}_1, \vec{e}_2$  of eigenvectors for  $\mathbb{R}^2$ . If we write the initial vector in this basis as:

$$\begin{pmatrix} a_0 \\ j_0 \end{pmatrix} = \alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2,$$

we get

$$\begin{pmatrix} a_n \\ j_n \end{pmatrix} = A^n \begin{pmatrix} a_0 \\ j_0 \end{pmatrix} = \alpha_1 \lambda_1^n \vec{e}_1 + \alpha_2 \lambda_2^n \vec{e}_2.$$

We thus see that the dynamics of this population depend on the eigenvalues, and hence on the fecundity rate  $\gamma$ :

- (i) If  $\lambda_1, \lambda_2 < 1$ , that is, if  $\gamma < \frac{5}{3}$ , then  $a_n, j_n \rightarrow 0$  and the population collapses.
- (ii) If  $\lambda_2 > 1$ , that is, if  $\gamma > \frac{5}{3}$ , then  $a_n, j_n \rightarrow \infty$  and the population explodes.
- (iii) Finally, if  $\lambda_2 = 1$  and  $\lambda_1 < 1$ , that is, if  $\gamma = \frac{5}{3}$ , then

$$\begin{pmatrix} a_n \\ j_n \end{pmatrix} \rightarrow \alpha_2 \vec{e}_2 = \alpha_2 \begin{pmatrix} \frac{3}{5} \\ 1 \end{pmatrix}.$$

In particular, the population converges to a stable population where the number  $a, j$  of adults and juveniles satisfy:

$$a = \frac{3}{5}j.$$

This mathematical model to understand the evolution of a population by dividing it into several age groups is known as the *Leslie model*.

**Application: The Fibonacci sequence.** Consider the sequence satisfying the following recurrence relation:

$$a_{n+1} = a_n + a_{n-1},$$

and with the initial values  $a_0 = 0$  and  $a_1 = 1$ . The numbers  $a_i$  are known as the *Fibonacci numbers*: 0, 1, 1, 2, 3, 5, 8, 13, .... Let us introduce the following vector of  $\mathbb{R}^2$ :

$$\vec{x}_n := \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} \quad \text{for } n \geq 0.$$

The recurrence relation gets encoded into the following matrix equation:

$$\vec{x}_{n+1} = A\vec{x}_n$$

with

$$A := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

To compute  $a_{n+1}$  from given  $a_1$  and  $a_0$ , we just have to apply  $A^n$  to the initial vector. The characteristic polynomial  $\chi_A$  is

$$\chi_A(x) = (1-x)(-x) - 1 = x^2 - x - 1,$$

and its eigenvalues are  $\varphi := \frac{1+\sqrt{5}}{2}$  and  $\varphi' := \frac{1-\sqrt{5}}{2}$ . In particular, the  $2 \times 2$  matrix  $A$  has two distinct eigenvalues, hence it is diagonalisable by Theorem 7.3. We compute the corresponding eigenvectors  $\vec{e}_\varphi = (\frac{1}{2}(1+\sqrt{5}), 1)^T$  and  $\vec{e}_{\varphi'} = (\frac{1}{2}(1-\sqrt{5}), 1)^T$ . One then writes the initial vector  $\vec{x}_0$  in this basis of eigenvectors by solving the system of linear equations

$$\vec{x}_0 = a\vec{e}_\varphi + b\vec{e}_{\varphi'},$$

for some  $a, b \in \mathbb{R}$ , using Gaussian elimination. One finds:

$$\vec{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}}\vec{e}_\varphi - \frac{1}{\sqrt{5}}\vec{e}_{\varphi'}.$$

If we apply  $A^n$  to our initial vector, we obtain

$$\vec{x}_n = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}}\varphi^n \vec{e}_\varphi - \frac{1}{\sqrt{5}}(\varphi')^n \vec{e}_{\varphi'}.$$

and by taking the second component of  $\vec{x}_n$ , one gets

$$a_n = \frac{\varphi^n - (\varphi')^n}{\sqrt{5}}.$$

**Sequences satisfying a linear recurrence relation.** More generally, for a sequence  $(a_n)$  satisfying a linear recurrence relation of the form

$$a_{n+m} = c_{m-1}a_{n+m-1} + \dots + c_0a_n$$

for some constants  $c_0, \dots, c_m \in \mathbb{R}$ , we can compute  $a_n$  using the same procedure. First introduce the following vector of  $\mathbb{R}^m$ :

$$\vec{x}_n = \begin{pmatrix} a_n \\ \vdots \\ a_{n+m-1} \end{pmatrix} \quad \text{for } n \geq 0,$$

And notice that we have

$$\vec{x}_{n+1} = A\vec{x}_n$$

with

$$A = \begin{pmatrix} c_{m-1} & c_{m-2} & \dots & c_1 & c_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

One then compute the characteristic polynomial  $\chi_A$ . One can show that

$$\chi_A(x) = x^m - c_{m-1}x^{m-1} - \dots - c_1x - c_0.$$

One can then compute the eigenvalues and eigenvectors. If  $A$  turns out to be diagonalisable, with a basis  $\mathcal{B}$  of eigenvectors  $\vec{e}_1, \dots, \vec{e}_m$  and associated eigenvalues  $\lambda_1, \dots, \lambda_m$  respectively, one then express the initial vector  $\vec{x}_0$  in the basis  $\mathcal{B}$  by solving the associated system of linear equations. If

$$\vec{x}_0 = \alpha_1 \vec{e}_1 + \dots + \alpha_m \vec{e}_m,$$

then it follows that

$$\vec{x}_n = A^n \vec{x}_0 = \alpha_1 \lambda_1^n \vec{e}_1 + \dots + \alpha_m \lambda_m^n \vec{e}_m,$$

and one finally obtains  $a_n$  by taking the first component of  $\vec{x}_n$ .

**Diagonalisability over  $\mathbb{C}$  and dynamical systems.** In this course, we have mostly been dealing with real vector spaces and diagonalisability over  $\mathbb{R}$ . When it comes to diagonalisability however, using complex numbers instead can be very useful: A matrix may be diagonalisable over  $\mathbb{C}$  (in which case the previous methods apply) even though the same matrix is not diagonalisable over  $\mathbb{R}$ .

Here is a concrete example. Consider a sequence satisfying the following linear recurrence relation:

$$u_{n+2} = -u_{n+1} - u_n.$$

By applying the previous method, we end up studying a dynamical system associated to a  $2 \times 2$  matrix whose characteristic polynomial is  $x^2 + x + 1$ . This polynomial does not admit any real root, so the associated matrix is not diagonalisable over  $\mathbb{R}$ . However, over  $\mathbb{C}$ , the characteristic polynomial has two roots,  $\lambda_1 = e^{2i\pi/3}$  and  $\lambda_2 = e^{-2i\pi/3}$ . One can show that the associated matrix is diagonalisable over  $\mathbb{C}$ , and as a result, there exists constants  $\alpha_1, \alpha_2 \in \mathbb{C}$  such that for every  $n \geq 0$ , we have

$$u_n = \alpha_1 \lambda_1^n + \alpha_2 \lambda_2^n.$$

We can then compute the values of  $\alpha_1, \alpha_2$  using the initial conditions  $u_0, u_1$ . For instance, for  $u_0 = 2$  and  $u_1 = 1$ , we find  $\alpha_1 = \alpha_2 = 1$ , and finally

$$u_n = e^{2in\pi/3} + e^{-2in\pi/3} = 2 \cos(2n\pi/3).$$

## 7.10 Powers of a matrix.

Diagonalising a matrix helps to compute its powers. Assume that we have diagonalised an  $n \times n$  matrix  $A$ :  $A = PDP^{-1}$ , where  $D$  is a diagonal matrix with entries  $d_1, \dots, d_n$ . We then have

$$A^k = (PDP^{-1})^k = PDP^{-1}PDP^{-1} \dots PDP^{-1} = PD^kP^{-1}.$$

Here,  $D^k$  is simply the diagonal matrix with entries  $d_1^k, \dots, d_n^k$ .

We have the following algorithm to compute the powers of a diagonalisable matrix  $A$ :

- 1) Using the previous method, compute a diagonal matrix  $D$  and an invertible matrix  $P$  such that  $A = PDP^{-1}$ .
- 2) We then have  $A^k = PD^kP^{-1}$  for every  $k \in \mathbb{Z}$ .

### Example 7.30

Let

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

Let us compute  $A^{20}$ .

We compute the characteristic polynomial  $\chi_A(\lambda) = \lambda^2 - 5\lambda + 4$ , and so the eigenvalues of  $A$  are  $\lambda = 1$  and  $\lambda = 4$ . In particular,  $A$  is a  $2 \times 2$  matrix with two distinct eigenvalues so it is diagonalisable. Using Gaussian elimination, we find that a basis of  $E_1(A)$  is the vector  $(2, -1)^T$  and a basis of  $E_4(A)$  is the vector  $(1, 1)^T$ . We thus define the matrix of change of basis

$$P = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$$

whose inverse we compute by Gaussian elimination:

$$P^{-1} = \frac{1}{3} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}.$$

We thus have

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix} = PDP^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}.$$

and it follows that

$$\begin{aligned} A^{20} &= PD^{20}P^{-1} = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4^{20} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 2 + 4^{20} & -2 + 2 \times 4^{20} \\ -1 + 4^{20} & 1 + 2 \times 4^{20} \end{pmatrix} \end{aligned}$$

## 8 Stochastic Matrices

In the previous chapter, we introduced new tools to study (linear) dynamical systems such as dynamics of animal population. In such a model, animals die and are born, and the total number of animals can vary (and as we have seen, the asymptotic behaviour is closely related to the eigenvalues of the associated matrix.)

In this chapter, we focus on a class of dynamical systems where the total number of ‘elements’ under study stays the same, modelling instead a phenomenon of *redistribution*. As we will see in the following motivating problem, many such dynamical systems reach some equilibrium, and the goal of this chapter will be to understand why.

### 8.1 A motivating problem: Optimising a distribution of bikes

A company is planning to introduce to-rent bikes in Edinburgh. The plan is to have various spots scattered in four main zones: Old Town (zone 1), Haymarket (zone 2), Fountainbridge (zone 3), and Newington (zone 4). During a trial period, the company monitored the displacements of 200 bikes to understand the behaviour of their possible customers. They noticed the following:

- (i) On average, 50% of bikes in zone 1 are parked back in zone 1 at the beginning of the next day, 20% are parked in zone 2, 20% in zone 3, and 10% in zone 4.
- (ii) On average, 20% of bikes in zone 2 are parked back at zone 2 at the beginning of the next day, 40% are parked in zone 1, 20% in zone 3, and 20% in zone 4.
- (iii) On average, 40% of bikes in zone 3 are parked back in zone 3 at the beginning of the next day, 30% are parked in zone 1, 20% in zone 2, and 10% in zone 4.
- (iv) On average, 50% of bikes in zone 4 are parked back in zone 4 at the beginning of the next day, 10% are parked in zone 1, 20% in zone 2, and 20% in zone 3.

A natural question (in order to decide the amount of parking slots needed) is to decide whether there exists a stable distribution of bikes, that is, so that the number of bikes parked at the beginning of every day remains constant over time. If we denote by  $(x_1, \dots, x_4)^T$  the vector whose components encode the distribution of bikes in zone 1, …, 4 respectively, this amounts to solving the following equation:

$$\begin{pmatrix} 0.5 & 0.4 & 0.3 & 0.1 \\ 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0.1 & 0.2 & 0.1 & 0.5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

This computation was done in Tutorial 1, and we found that the unique solution to this system was the vector  $(70, 40, 50, 40)^T$ .

However, the company might just have decided to start with an equal distribution of 50 bicycles at each location. In that situation, one can study the evolution over time of the distribution of

bicycles. Indeed, if we denote by  $\vec{x}_k$  the vector whose components encode the average distribution of bicycles in zones 1, to 4 after  $k$  days, then

$$\vec{x}_k = \begin{pmatrix} 0.5 & 0.4 & 0.3 & 0.1 \\ 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0.1 & 0.2 & 0.1 & 0.5 \end{pmatrix}^k \begin{pmatrix} 50 \\ 50 \\ 50 \\ 50 \end{pmatrix}.$$

We find:

$$\vec{x}_1 = \begin{pmatrix} 65 \\ 40 \\ 50 \\ 45 \end{pmatrix}, \quad \vec{x}_2 = \begin{pmatrix} 68 \\ 40 \\ 50 \\ 42 \end{pmatrix}, \quad \vec{x}_3 = \begin{pmatrix} 69.2 \\ 40 \\ 50 \\ 40.8 \end{pmatrix}, \quad \vec{x}_4 \simeq \begin{pmatrix} 69.7 \\ 40 \\ 50 \\ 40.3 \end{pmatrix}, \dots$$

In particular, we see that, over time, the distribution seems to converge (rather quickly) to the stable distribution we computed before. There are thus two natural questions to ask oneself:

- Does that sort of system (and what sort of system exactly are we looking at?) always have a stable distribution?
- Does the distribution always evolve over time towards the same stable distribution, regardless of the initial conditions?

The goal of this final chapter will be to answer in the affirmative these two questions, and to see further applications.

## 8.2 Markov diagrams and transition matrices

The dynamical systems we wish to encode correspond to the evolution over time of a set of ‘elements’ (bikes, people, animals, etc.) that can be in a certain number of ‘states’. At every time, each element can either stay in the same state or change to a different state with a given probability. The only condition is that this process is *without memory*, that is, the probability from moving to a state  $s_j$  to a state  $s_i$  depends only on the states  $s_i, s_j$ , and not of the states the process was at at previous times. Such processes, heavily studied in probability theory and statistics, are called *finite Markov chains*.

### Example 8.1

We give here some examples to illustrate the type of processes we are trying to model:

- (i) The study of bikes mentioned in the introduction can be modelled in this way: the set of elements correspond to the 200 bikes considered, the states correspond to the four possible zones, and the averaged probabilities measured in the introduction give the probability of a bike to move from one state to another after one day.
- (ii) We flip a fair coin and record the occurrences of heads and tails. Here, the set of elements considered consists of the coin itself, and the two possible states are ‘heads’ and ‘tails’.

Since the coin is fair, the probability from moving from one state to the other is always  $\frac{1}{2}$ .

### Example 8.2

Here is an example of a process that cannot be modelled in this way. We play a game by flipping a fair coin as follows: We start with 10 pounds. Every time we get heads we win 1 pound, and every time we get tails we lose 1 pound. However, being reasonable, we decide to stop playing if we lose three times in a row. One way we could try to model this game is by taking the set of elements to consist of a single player, and the three states can be: ‘win’, ‘lose’, and ‘stop playing’, and we go from one state to the other depending on the outcome of the coin flips. This process however does not fit the previous description. Indeed, knowing that we just lost 1 pound (i.e. we are currently on the ‘lose’ state), we cannot tell the probability to move to the other states without knowledge of the prior events: if we had already lost twice before, then we will stop playing and move to state ‘stop playing’ with probability 1. Otherwise, we will keep playing and move to state ‘win’ or ‘lose’ with probability  $\frac{1}{2}$ .

Now that we have some intuition about the type of processes we are trying to model, we will present an algebraic framework to study them. The key thing is to encode the various probabilities of moving from one state to another in a workable way. This is done as follows, using a *Markov diagram*.

### Definition 8.3

*Markov diagram* is a finite labelled oriented graph:

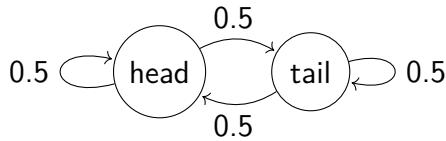
- for each pair of vertices  $v_i, v_j$ , there is at most one oriented edge from  $v_j$  to  $v_i$ , and the label of that edge is an element  $a_{ij} \in (0, 1]$ .
- for each vertex  $v_j$ , we have:

$$\sum_{i=1}^n a_{ij} = 1.$$

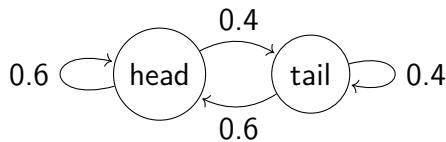
Intuitively, a Markov diagram encodes (in a diagrammatic way) the types of processes we are interested: the vertices correspond to the states of our process, and an oriented edge with label  $a_{ij}$  from the vertex  $v_j$  to the vertex  $v_i$  corresponds to a probability  $a_{ij}$  of moving from state  $v_j$  to state  $v_i$ . We only indicate non-zero probabilities in the diagram, hence the requirement that  $a_{ij} \in (0, 1]$ . Moreover, the second condition simply means that the diagram encodes all the possible transitions from one state to the other: if we are at a state  $v_j$  at a given time, with probability 1 ( $= \sum_{i=1}^n a_{ij}$ ) we will move to one of the other states indicated in the diagram by following an oriented edge starting at  $v_j$ .

### Example 8.4

Let us consider again the example of the coin flip, and the coin is assumed to be fair: heads and tails happen with probability  $\frac{1}{2}$ . The Markov diagram for this process is the following:



If instead we considered an unfair coin giving a head with a probability of 60%, then the Markov diagram would be:



Another way to encode this data is by means of a matrix, called a *transition matrix*, which is particularly useful to understand the evolution over time of the process.

### Definition 8.5

The *transition matrix* associated to a Markov diagram with  $n$  states is the  $n \times n$  matrix defined by:

$$a_{ij} = \text{probability of going from state } s_j \text{ at time } n \text{ to state } s_i \text{ at time } n + 1.$$

### Example 8.6

- (i) The transition matrix associated to the fair flip coin above is:

$$A = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix},$$

while the transition matrix associated to the unfair flip coin above is:

$$A = \begin{pmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{pmatrix},$$

- (ii) The introductory example involving bikes can be modelled using a Markov diagram with four states  $s_1, \dots, s_4$  corresponding to the four zones  $1, \dots, 4$  respectively. The associated transition matrix is:

$$A = \begin{pmatrix} 0.5 & 0.4 & 0.3 & 0.1 \\ 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0.1 & 0.2 & 0.1 & 0.5 \end{pmatrix}.$$

Note that a transition matrix is a particular example of a stochastic matrix, as defined below:

**Definition 8.7**

An  $n \times n$  matrix is called a *stochastic matrix* if the following holds:

- (S1)  $a_{ij} \geq 0$  for every  $1 \leq i, j \leq n$ ,
- (S2) for every column  $C_j$  of  $A$ , the sum of its entries is 1.

The transition matrix can be used to model the evolution over time of the distribution of a given number of elements taking at each time one of the possible states. Suppose that we start originally with  $x_1$  elements in the state  $s_1$ ,  $x_2$  elements in the state  $s_2$ , etc. We encode this by a vector  $\vec{x}_0 = (x_1, \dots, x_n)^T$ , whose coordinates record the initial distribution of objects under study. At each iteration of the process, each object moves from one state to another with following the probabilities encoded in the Markov diagram (or equivalently, in the transition matrix). At a given time, we have the following:

**Theorem 8.8**

Let  $\vec{x}_0 = (x_1, \dots, x_n)^T$  be the vector whose coordinates record the initial distribution of elements in states  $s_1, \dots, s_n$ . Let us denote by  $\vec{x}_k$  the vector whose components record the expected distribution after  $k$  iterations of the process of elements in states  $s_1, \dots, s_n$  respectively. We have for every  $k \geq 0$ :

$$\vec{x}_{k+1} = A\vec{x}_k \text{ for every } k \geq 0.$$

In particular, we get by induction:

$$\vec{x}_k = A^k \vec{x}_0.$$

**Definition 8.9**

A vector  $\vec{x}$  such that  $A\vec{x} = \vec{x}$  is called a *stationary distribution*. In other words, it is an eigenvector of  $A$  for the eigenvalue 1.

A stationary distribution represents the distribution over the possible states when the system is at equilibrium, i.e. does not change over time.

**Example 8.10**

In the Markov chain modelling the distribution of bikes mentioned in the introduction, the vector  $(70, 40, 50, 40)^T$  is a stationary distribution.

### 8.3 The Perron-Frobenius Theorem for stochastic matrices

#### Definition 8.11

Let  $A$  be an  $n \times n$  matrix.

- We say that  $A$  is *non-negative* if  $a_{ij} \geq 0$  for every  $i, j$ , and *positive* if  $a_{ij} > 0$  for every  $i, j$ .
- We say that a non-negative matrix  $A$  is *regular* if  $A^k$  is positive for some integer  $k \geq 1$ .

The transition matrix associated to a Markov diagram is non-negative, but not necessarily positive as there may be a probability 0 of going from some state  $s_i$  to some other state  $s_j$ .

The integer  $k$  in the definition above depends on the matrix. In particular, for every  $k$  it is possible to construct a regular matrix  $A$  such that  $A^k$  is positive but  $A, A^2, \dots, A^{k-1}$  are not positive.

#### Theorem 8.12

The transition matrix associated to a Markov diagram is regular if and only if there exists an integer  $n \geq 1$  such that one can go from any vertex of its Markov diagram to any other by a sequence of exactly  $n$  oriented edges.

*Proof.* The  $(i, j)$  term of  $A^k$  is given by:

$$(A^k)_{i,j} = \sum_{1 \leq i_1, \dots, i_{k-1} \leq n} a_{i,i_1} a_{i_1,i_2} \cdots a_{i_{k-2},i_{k-1}} a_{i_{k-1},j}.$$

Since all the coefficients  $a_{i,j}$  are non-negative,  $A^k$  is positive if and only if for every  $i, j$ , there exists  $1 \leq i_1, \dots, i_{k-1} \leq n$  such that  $a_{i,i_1} a_{i_1,i_2} \cdots a_{i_{k-2},i_{k-1}} a_{i_{k-1},j}$  are all non-zero. Since a directed edge in the Markov diagram corresponds to a positive probability to move from one state to the other, this is equivalent to saying that for every  $i, j$ , there exists a sequence of  $n$  directed edges from the state  $i$  to the state  $j$ .  $\square$

#### Example 8.13

The following matrix is a regular stochastic matrix:

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

Indeed, it is clearly stochastic. Moreover, since  $A = A^1$  is positive, it is also regular.

#### Example 8.14

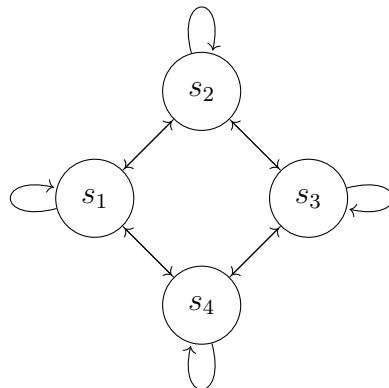
The following matrix is a regular stochastic matrix:

$$A := \begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix}.$$

Indeed, the fact that  $A$  is stochastic is clear, and we have:

$$A^2 = \begin{pmatrix} 1/3 & 2/9 & 2/9 & 2/9 \\ 2/9 & 1/3 & 2/9 & 2/9 \\ 2/9 & 2/9 & 1/3 & 2/9 \\ 2/9 & 2/9 & 2/9 & 1/3 \end{pmatrix},$$

which is positive. The corresponding Markov diagram is the following (where for simplicity we omit the probabilities, and we represent arrows from  $s_i$  to  $s_j$  and from  $s_j$  to  $s_i$  simply by a double arrow):



and we see that it is indeed possible to find a directed path of length 2 between any pair of states.

The key theorem of this chapter is the following:

### Theorem 8.15

Perron-Frobenius Theorem for stochastic matrices Let  $A$  be a regular stochastic matrix. Then the following holds:

- (i) There exists stationary distributions for  $A$ . In other words, 1 is an eigenvalue of  $A$ . Moreover, there exists a stationary distribution all of whose components are strictly positive.
- (ii) For every vector  $\vec{x} = (x_1, \dots, x_n)^T$  such that  $\sum_i x_i \neq 0$ , we have that the sequence of vectors  $A^k \vec{x}$  converges to a stationary distribution of  $A$ .
- (iii)  $\dim E_1(A) = 1$ .
- (iv) For every other eigenvalue  $\lambda$  of  $A$ , we have  $|\lambda| < 1$ .

The proof of the Perron-Frobenius is rather long and technical (you can find the proof on the

internet, or come and ask me after class/tutorials.). I will just present the proofs of two of the fours results, that illustrate nicely some of the concepts we have encountered in the course so far.

For any square matrix, we have

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

A square matrix and its transpose have the same eigenvalues.

*Proof.* Indeed, since  $\det(A) = \det(A^T)$  for any square matrix, it follows that

$$\chi_{A^T}(\lambda) = \det(A^T - \lambda I_n) = \det((A - \lambda I_n)^T) = \det(A - \lambda I_n) = \chi_A(\lambda).$$

In particular,

$$\lambda \text{ is an eigenvalue of } A \iff \chi_A(\lambda) = 0 \iff \chi_{A^T}(\lambda) = 0 \iff \lambda \text{ is an eigenvalue of } A^T.$$

□

*Proof.* (i) Notice that, since  $A$  is stochastic, the entries in each column of  $A$  add up to 1. This can be reformulated as follows:

$$A^T \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

In particular, we see that 1 is an eigenvalue of  $A^T$ . But  $A$  and  $A^T$  have the same eigenvalues by the above theorem. Since 1 is an eigenvalue of  $A^T$ , it is also an eigenvalue of  $A$ .

(ii) Let us prove this statement in the case where  $A$  is diagonalisable. We choose a basis  $\vec{e}_1, \dots, \vec{e}_n$  of eigenvectors for the eigenvalues  $1 = \lambda_1, \lambda_2, \dots, \lambda_n$  respectively. Let  $\vec{x} \in \mathbb{R}^n$  and let us write  $\vec{x} = \alpha_1 \vec{e}_1 + \dots + \alpha_n \vec{e}_n$ , for some  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ . We then have for  $k \geq 1$ :

$$\begin{aligned} A^k \vec{x} &= \alpha_1 A^k \vec{e}_1 + \alpha_2 A^k \vec{e}_2 + \dots + \alpha_n A^k \vec{e}_n \\ &= \alpha_1 \vec{e}_1 + \alpha_2 \lambda_2^k \vec{e}_2 + \dots + \alpha_n \lambda_n^k \vec{e}_n. \end{aligned}$$

Since  $|\lambda_2|, \dots, |\lambda_n| < 1$  by (iv), it follows that  $A^k \vec{x}$  converges to  $\alpha_1 \vec{e}_1$ . To conclude that the limit is an eigenvector (for the eigenvalue  $\lambda_1 = 1$ ), it remains to show that  $\alpha_1 \vec{e}_1 \neq \vec{0}$ . (an eigenvector cannot be the null vector.) But since  $A$  is stochastic, the sum of the components of  $A^k \vec{x}$  remains the same at every stage. Since we started from a vector  $\vec{x}$  such that  $\sum_i x_i \neq 0$ , it follows that the sum of the components of  $\alpha_1 \vec{e}_1$  is non-zero, and in particular  $\alpha_1 \vec{e}_1 \neq \vec{0}$ . □

## 8.4 Application: stationary distributions in dynamical systems.

Let us go back to the problem stated in the introduction. The problem of modelling the changes in the distribution of bicycles can be modelled using a Markov diagram, whose transition matrix  $A$  was given in the introduction. This associated transition matrix is regular, as it is positive. Let us denote by  $\vec{x}_0 = (50, 50, 50, 50)^T$  the vector corresponding to the initial distribution of bikes. After  $k$  days, the expected distribution of bikes is given by  $A^k \vec{x}_0$ . By (iv), this sequence

of vectors converges towards a stationary distribution  $\vec{x}$ . This stationary distribution is a vector whose components add up to 200, since each vector  $A^k \vec{x}_0$  satisfies the same property (intuitively, the bikes are just redistributed, there is no gain nor loss of bike in the process). By (ii), since the eigenspace  $E_1(A)$  is a line (i.e. has dimension 1), there is a unique vector of  $E_1(A)$  whose components add up to 200, and this is exactly our stationary distribution  $\vec{x}$ . Note that this characterisation of  $\vec{x}$  does not depend on the original distribution  $\vec{x}_0$ . This explains in hindsight why the stationary distribution is unique, and does not depend on the original vector  $\vec{x}_0$ .

One natural question would be to understand the speed of convergence. When the transition  $A$  matrix is diagonalisable, then there exists a basis  $\vec{e}_1, \dots, \vec{e}_n$  of  $\mathbb{R}^n$  for the associated eigenvalues  $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_n > -1$ . In particular, if we write our initial vector  $\vec{x}$  as a linear combination of the  $\vec{e}_i$ , as  $\vec{x} = \alpha_1 \vec{e}_1 + \dots + \alpha_n \vec{e}_n$ , we have for  $k \geq 1$ :

$$A^k \vec{x} = \alpha_1 \vec{e}_1 + \alpha_2 \lambda_2^k \vec{e}_2 + \dots + \alpha_n \lambda_n^k \vec{e}_n.$$

In particular, we see that  $A^k \vec{x} \rightarrow_{\infty} \alpha_1 \vec{e}_1$  and that  $\|A^k \vec{x} - \alpha_1 \vec{e}_1\|$  converges to zero exponentially, and as fast as  $\mu^k$  (up to a multiplicative constant), where

$$\mu = \max_{2 \leq i \leq n} |\lambda_i|.$$

Rephrased using Landau's notations, we have:

$$\|A^k \vec{x} - \vec{e}_1\| = O(\mu^k)$$

If  $A$  is not diagonalisable, a slightly weaker inequality holds, namely:

$$\|A^k \vec{x} - \vec{e}_1\| = o(\mu^k)$$

for every  $\max_{2 \leq i \leq n} |\lambda_i| < \mu < 1$ . We will not prove this stronger statement in this course.

## 8.5 Application: The PageRank algorithm

One can imagine the internet as a (very large) oriented graph. Vertices correspond to pages and oriented edges correspond to links from a page to another.

To this oriented graph, we can associate a finite Markov chain whose states correspond to the pages, by assuming that a person on a given webpage will click with equal probability on one of the links of the page. To construct the associated transition matrix, we denote by  $n_{ij}$  the number of links from page  $j$  to page  $i$ , and by  $d_j$  the total number of links on page  $j$ , then the probability  $a_{ij}$  that someone on page  $j$  clicks on a link leading to page  $i$  is:

$$a_{ij} = \frac{n_{ij}}{d_j}.$$

Let us consider a small example of a very small network consisting of four webpages  $p_1, \dots, p_4$ . The table representing the links between pages is given below, where the coefficient in row  $i$  and column  $j$  represents the number of links from page  $j$  to page  $i$ :

	$p_1$	$p_2$	$p_3$	$p_4$
$p_1$	0	0	2	2
$p_2$	2	0	2	1
$p_3$	1	1	0	2
$p_4$	1	1	1	0

The transition matrix associated to this Markov chain is:

$$A = \begin{pmatrix} 0 & 0 & 2/5 & 2/5 \\ 1/2 & 0 & 2/5 & 1/5 \\ 1/4 & 1/2 & 0 & 2/5 \\ 1/4 & 1/2 & 1/5 & 0 \end{pmatrix}.$$

However, this model fails to take into account that quite often people will go to a webpage not because it was linked from a previous page, but because they started thinking of something completely different, independently of the content of the current page. This can be modelled by another Markov process where at each time, the probability of going from any page to any other page is exactly  $1/n$ , where  $n$  is the number of pages (in our example,  $n = 4$ ). This finite Markov chain has a transition matrix a  $n \times n$  matrix of the form

$$U = \begin{pmatrix} 1/n & \cdots & 1/n \\ \vdots & \ddots & \vdots \\ 1/n & \cdots & 1/n \end{pmatrix}.$$

What Google does to take both processes into consideration is to consider that what is going on is a combination of the previous cases. There exists a constant  $\alpha$ , called *damping factor*, such that with probability  $\alpha$ , a person will choose the next page by clicking on one of the links, and with probability  $(1 - \alpha)$  that person will choose the next page completely at random. A standard value for  $\alpha$  is  $\alpha = 0.85$ . This corresponds to a Markov process whose transition matrix is given by the following barycenter of  $A$  and  $U$ :

$$G := \alpha A + (1 - \alpha)U.$$

Note that this transition matrix is now strictly positive, and in particular regular. We can thus apply the Perron-Frobenius Theorem, which tells us that there exists a stationary distribution with positive components and with sum 1, called the *ranking vector*. By the convergence property, this corresponds to the expected distribution of people on pages when this process has been repeated infinitely many times. Pages with a comparatively high distribution correspond to ‘popular’ pages, and are thus ranked highly by Google. This ranking vector is (a simplified version of) what Google uses to rank pages.

**Definition 8.16**

A *ranking vector* used in the Google PageRank algorithm is a stationary distribution of the 'Google matrix'  $G = \alpha A + (1 - \alpha)U$  with positive components. (Such a vector is unique up to multiplication by a positive constant, by the Perron-Frobenius Theorem)  
The ranking of the webpages is obtained by ranking the components of that vector.

Let us go back to our simple example. Taking the damping factor to be 0.85, the new transition matrix for the process is

$$\begin{aligned} G &= 0.85 \begin{pmatrix} 0 & 0 & 2/5 & 2/5 \\ 1/2 & 0 & 2/5 & 1/5 \\ 1/4 & 1/2 & 0 & 2/5 \\ 1/4 & 1/2 & 1/5 & 0 \end{pmatrix} + 0.15 \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} \\ &= \begin{pmatrix} 0.0375 & 0.0375 & 0.3775 & 0.3775 \\ 0.4625 & 0.0375 & 0.3775 & 0.2075 \\ 0.25 & 0.4625 & 0.0375 & 0.3775 \\ 0.25 & 0.4625 & 0.2075 & 0.0375 \end{pmatrix}. \end{aligned}$$

Finding the stationary distribution corresponds to solving the system  $G\vec{x} = \vec{x}$ , which we do as usual using Gaussian elimination, with a computer as the computations are too cumbersome. The unique stationary distribution  $\vec{r}$  whose components are positive and add up to 1 is (with approximations to the second digit)

$$\vec{r} \simeq \begin{pmatrix} 0.21 \\ 0.26 \\ 0.28 \\ 0.24 \end{pmatrix}.$$

So the ranking of these pages, from most popular to least popular, is: page 3, page 2, page 4, page 1.

**Speed of convergence (extra-curricular).** We know from the Perron-Frobenius Theorem that a way to obtain a ranking vector is to look at the asymptotic behaviour of  $G^k\vec{x}$  for an arbitrary vector  $\vec{x}$  with  $\sum x_i = 1$ . We wish to find an upper bound on the norm  $\|G^k\vec{x} - \vec{r}\|$ . However, instead of looking at this norm, we will be looking at a slightly different one that is easier to use in this situation:

$$\|\vec{x}\|_1 = \sum_i |x_i|.$$

Note that we still have the triangle identity  $\|\vec{x} + \vec{y}\|_1 \leq \|\vec{x}\|_1 + \|\vec{y}\|_1$ . Moreover, for every vector  $\vec{x} = (x_1, \dots, x_n)^T$ , we have:

$$\|A\vec{x}\|_1 = \sum_i |\sum_j a_{ij}x_j| \leq \sum_i \sum_j a_{ij}|x_j| \leq \sum_j |x_j|(\sum_i a_{ij}) \leq \sum_i |x_i| = \|\vec{x}\|_1.$$

**Theorem 8.17 F**

or every  $k \geq 1$  and every positive vector  $\vec{x}$  whose components add up to 1, we have:

$$\|G^k \vec{x} - \vec{r}\|_1 \leq 2\alpha^k.$$

*Proof.* We have:

$$G\vec{x} = \alpha A\vec{x} + (1 - \alpha)U\vec{x} = \alpha A\vec{x} + (1 - \alpha) \begin{pmatrix} 1/n \\ \vdots \\ 1/n \end{pmatrix}.$$

In particular, for non-negative vectors  $\vec{x}, \vec{y}$ , we get

$$G\vec{x} - G\vec{y} = \alpha A(\vec{x} - \vec{y}),$$

and

$$\|G\vec{x} - G\vec{y}\|_1 = \alpha \|A(\vec{x} - \vec{y})\|_1 \leq \alpha \|\vec{x} - \vec{y}\|_1.$$

By induction, we get

$$\|G^k \vec{x} - G^k \vec{y}\|_1 \leq \alpha^k \|\vec{x} - \vec{y}\|_1$$

and thus, for every positive vector  $\vec{x}$  whose components add up to 1, we get:

$$\|G^k \vec{x} - \vec{r}\|_1 = \|G^k \vec{x} - G^k \vec{r}\|_1 \leq \alpha^k \|\vec{x} - \vec{r}\|_1 \leq \alpha^k (\|\vec{x}\|_1 + \|\vec{r}\|_1) \leq 2\alpha^k.$$

□

Note in particular that this speed of convergence is independent on the size of the matrix. For instance, for  $\alpha = 0.85$ , then the power method gives an approximation of the ranking vector with error less than  $10^{-6}$  after 90 iterations.