CHAPTER 0

INTRODUCTION

In this chapter we introduce key concepts that will be used in later chapters. For this reason, unlike other chapters it contains many statements, sometimes given without thorough explanations or reasoning. While all of these statements are grounded in deep ideas and can be formulated in a rigorous manner, it is advised to first get an intuitive understanding of the ideas before diving into their more formal construction.

Note 0.1 In case you are already familiar with the topics

It is recommended for readers who are familiar with the topics to at least gloss over this chapter and make sure they know and understand all the concepts presented here.

0.1 EXERCISES

0.1. Write the following sets explicitly:

(i) $\{x \in \mathbb{N} \mid 1 < x \leq 7\}$

(ii) $\{x \in \mathbb{Z} \mid x < 5\}$

(iii) $\left\{x \in \mathbb{R} \mid x^2 = -1\right\}$

(iv) $\{x \in \mathbb{N} \wedge x \in \mathbb{Q}\}$

(v) $\left\{x \in \mathbb{R} \mid x^2 - 3x - 4 = 0\right\}$

(vi) $\{x \in \mathbb{R} \mid x < 5 \wedge x \geq 2\}$

0.2. Determine the relation between the sets:

(i) $A = \{1,\ 2,\ 3\},\ B = \{1,\ 2\}$

(ii) $A = \varnothing,\ B = \{2,\ -5,\ \pi\}$

(iii) $A = \mathbb{Z},\ B = \{\pm x \mid x \in \mathbb{N} \cup \{0\}\}$

(iv) $A = \left\{\pi, \mathrm{e}, \sqrt{2}\right\},\ B = \mathbb{Q}$

0.3. Write all elements in $S^2 \times W$, where $S = \{\alpha, \beta, \gamma\}$ and $W = \{x, y, z\}$. Find a condition that guarantees $S^2 \times W = W \times S^2$.

0.4. How many different injective functions $f : \{1,2\} \to \{1,2\}$ exist? How many injective functions $f : \{1,2,3\} \to \{1,2,3\}$ exist? How many inject functions $f : \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$ exist for a given $n \in \mathbb{N}$?

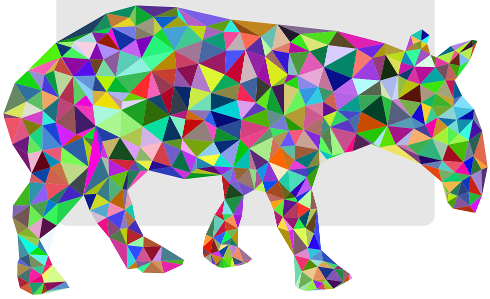
0.5. For each of the real functions below, find a set on which it is surjective (use a graphing calculator if you are not familiar with the shape of a function):

$$x^2,\ x^3 - 5,\ \mathrm{e}^{-x^2/2},\ \sin(x),\ \sin(x) + \cos(x),\ x\mathrm{e}^x.$$

0.6. Given two sets A, B such that $|B| = |A| - 1$, can a bijective function $f : A \to B$ exist? Explain your answer.

0.7. MORE EXERCISES TO BE WRITTEN...

CHAPTER 1



LINEAR ALGEBRA
(INTUITIVE APPROACH)

Linear algebra is one of the most important and often used fields, both in theoretical and applied mathematics. It brings together the analysis of systems of linear equations and the analysis of linear functions (in this context usually called linear transformations), and is employed extensively in almost any modern mathematical field, e.g. approximation theory, vector analysis, signal analysis, error correction, 3-dimensional computer graphics and many, many more.

In this book, we divide our discussion of linear algebra into to chapters: the first (this chapter) deals with a wider, birds-eye view of the topic: it aims to give an intuitive understanding of the major ideas of the topic. For this reason, in this chapter we limit ourselves almost exclusively to discussing linear algebra using 2- and 3-dimensional analysis (and higher dimensions when relevant) using real numbers only. This allows us to first create an intuitive picture of what is linear algebra all about, and how to use correctly the tools it provides us with.

The next chapter takes the opposite approach: it builds all concepts from the ground-up, defining precisely (almost) all basic concepts and proving them rigorously, and only then using them to build the next steps. This approach has two major advantages: it guarantees that what we build has firm foundations and does not fall apart at any future point, and it also allows us to generalize the ideas constructed during the process to such extent that they can be used as foundation to build ever newer tools we can apply in a wide range of cases.

1.1 VECTORS

1.1.1 Basics

Vectors are the fundamental objects of linear algebra: the entire field revolves around manipulation of vectors. In this chapter we deal with the so-called **real vectors**, which can be be defined in a geometric way:

Definition 1.1 Real vectors

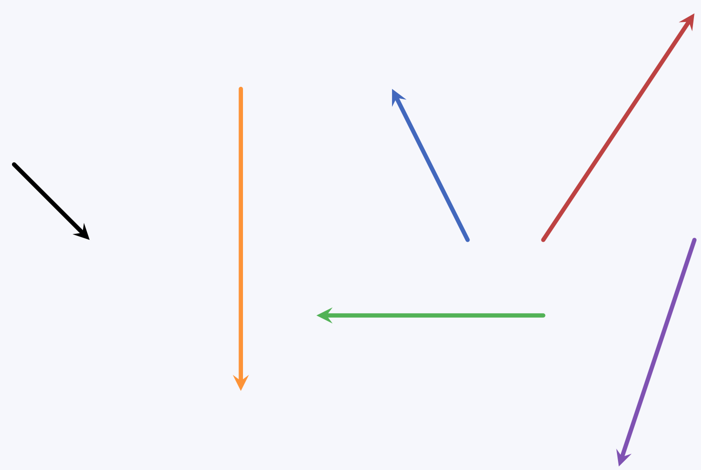
A *real vector* is an object with a **magnitude** (also called **norm**) and a **direction**.

π

In this chapter we refer to real vectors simply as *vectors*.

Example 1.1 Real vectors

The following are all vectors in 2-dimensional space depicted as arrows:



Vectors are usually denoted in one of the following ways:

- **Arrow above letter**: $\vec{u}$, $\vec{v}$, $\vec{x}$, $\vec{a}$, …

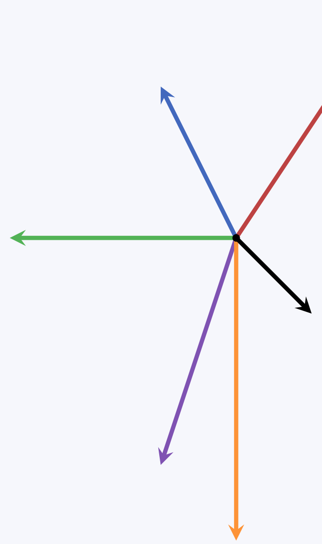
- **Bold letter**: $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{x}, \boldsymbol{a}, \ldots$
- **Bar below letter**: $\underline{u}, \underline{v}, \underline{x}, \underline{a}, \ldots$

In this book we use the first notation style, i.e. an arrow above the letter. In addition vectors will almost always be denoted using lowercase Lating script.

When discussing vectors in a single context, we always consider them starting at the same point, called the **origin**, and **translating** (moving) vectors around in space does not change their properties: only their norms and directions matter.

Example 1.2 Real vectors

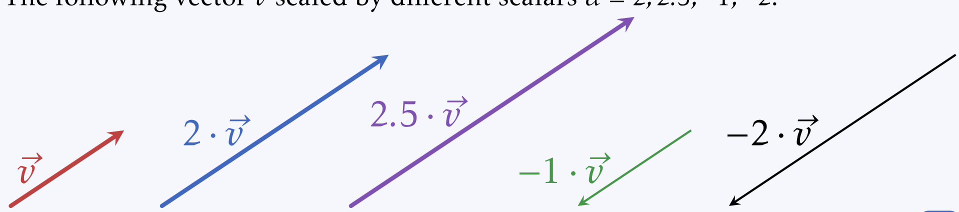
The vectors from the previous translated (moved) such that their origins all lie on the same point:



A vector can be scaled by a real number α: when this happens, its norm is multiplied by α while its direction stays the same. We call α a **scalar**.

Example 1.3 Scaling vectors

The following vector $\vec{v}$ scaled by different scalars $\alpha = 2, 2.5, -1, -2$:



Note 1.1 Negative scale

As can be seen in the example above, when scaling a vector by a negative amount its direction reverses. However, we consider two opposing direction (i.e. directions that are 180° apart) as being the same direction.

In this book we use the following notation for the norm of a vector $\vec{v}$: $\left\|\vec{v}\right\|$.

A vector $\vec{v}$ with norm $\left\|\vec{v}\right\| = 1$ is called a **unit vector**, and is usually denoted by replacing the arrow symbol by a hat symbol: $\hat{v}$. Any vector (except $\vec{0}$) can be scaled into a unit vector by scaling the vector by 1 over its own norm, i.e.

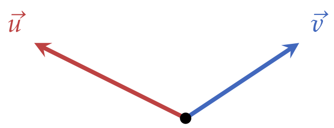
$$\hat{v} = \frac{1}{\left\|\vec{v}\right\|}\vec{v}. \tag{1.1.1}$$

The result of normalization is a vector of unit norm which points in the same direction of the original vector.

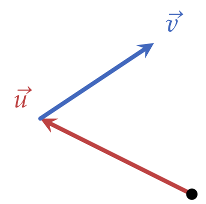
Two vectors can be added together to yield a third vector: $\vec{u} + \vec{v} = \vec{w}$. To find $\vec{w}$ we use the following procedure (depicted in Figure 1.1):

1.1. Move (translate) $\vec{v}$ such that its origin lies on the head of $\vec{u}$.

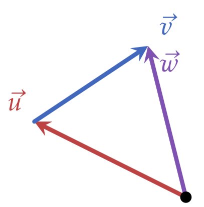
1.2. The vector $\vec{w}$ is the vector drawn from the origin of $\vec{u}$ to the head of $\vec{v}$.



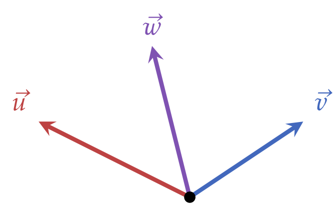
1 The vectors $\vec{u}$ and $\vec{v}$.



2 Translating $\vec{v}$ such that its origin lies at the head of $\vec{u}$.



3 Drawing the vector $\vec{w}$ from the origin to the head of $\vec{v}$.



4 Showing all three vectors.

Figure 1.1 Vector addition.

The addition of vectors as depicted here is commutative, i.e. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$. This can be seen by using the **parallogram law of vector addition** as depicted in Figure 1.2: drawing the two vectors $\vec{u}, \vec{v}$ and their translated copies (each such that its origin lies on the other vector's head) results in a parallelogram.

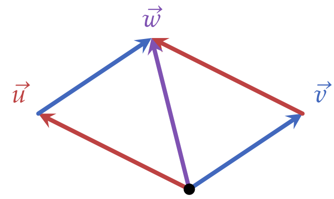


Figure 1.2 The parallogram law of vector addition.

An important vector is the **zero-vector**, denoted as $\vec{0}$. The zero-vector has a unique property: it is neutral in respect to vector addition, i.e. for any vector $\vec{v}$,

$$\vec{v} + \vec{0} = \vec{v}. \tag{1.1.2}$$

(we also say that $\vec{0}$ is the **additive identity** in respect to vectors.)

Any vector $\vec{v}$ always has an **opposite** vector, denoted $-\vec{v}$. The addition of a vector and its opposite always result in the zero-vector, i.e.

$$\vec{v} + (-\vec{v}) = \vec{0}. \tag{1.1.3}$$

1.1.2 Components

Vectors can be decomposed to their components, the number of which depends on the dimension of space we're using: 2-dimensional vectors can be decomposed into 2 components, 3-dimensional vectors can be decomposed into 3 components, etc. To decompose a vector, say $\vec{v}$, we first choose a coordinate system: the most commonly used system, and the one we will use for most of this chapter, is the Cartesian coordinate system. We place the vector in the coordinate system such that its origin lies at the origin of the system. We then draw a perpendicular line from its head to each of the axes in the system (see Figure 1.3), the point of interception on each axis is the component of the vector in that axis (we label these points v_x, v_y, v_z in the case of 2- or 3-dimensional spaces, and generally $v_1, v_2, v_3, \dots$). The vector can then be written as a column using these components:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}. \tag{1.1.4}$$

Note 1.2 Order of components

The order of the components of a vector is important, and should always be consistent. In the case of 2- and 3-dimensional the order is always v_x, v_y, v_z.

!

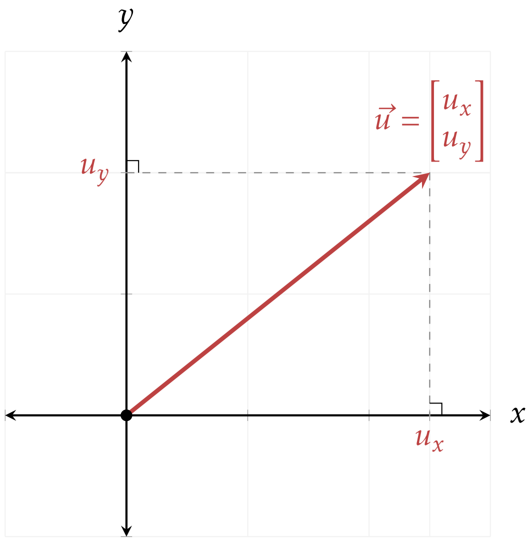
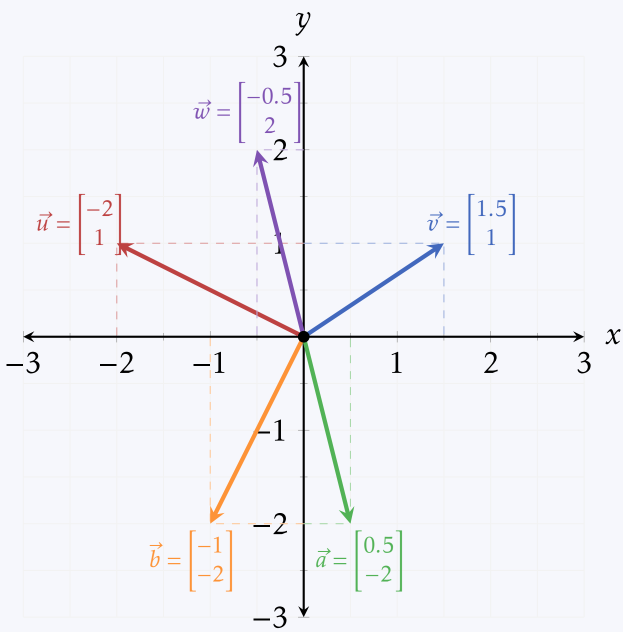


Figure 1.3 Placing a 2-dimensional vector $\vec{u}$ on the 2-dimensional Cartesian coordinate system, showing its x- and y-components.

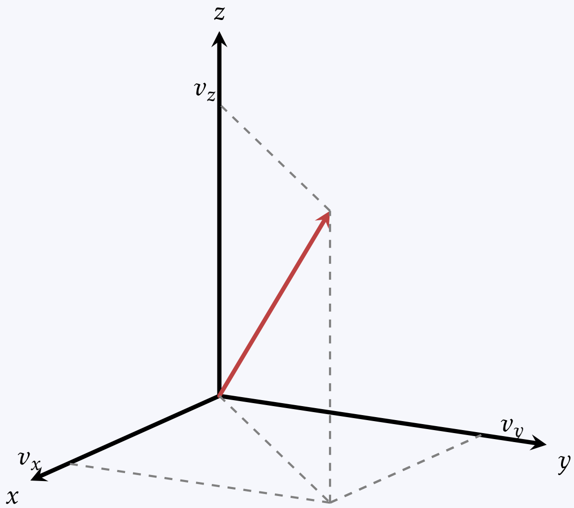
Example 1.4 Vector components in two dimensions

The following five 2-dimensional vectors are decomposed each into its x- and y-components:



Example 1.5 Vector components in three dimensions

The following 3-dimensional vector is decomposed into its x-, y- and z-components: (THIS NEEDS TO BE IMPROVED AND FINISHED)



The column form of a vector is essentially equivalent to an order list of n real numbers, i.e. $(v_1, v_2, \dots, v_n)$. Why then are we using the column form and not the list form (mostly known as **row vectors**)? In fact, we could use either form - and even using both interchangeably - and with only minor adjusments the entire chapter would stay the same as it is now. However, there are some advantages of using only a single form, and consider the other form as a different object altogether. This idea will become clear in future chapters, when discussing **covariant vectors**, **contravarient vectors**, and **tensors**. For now, we stick with the column form of vectors to stay consistent with common notation.

However, the row form of vectors highlights the space in which they exist: n-dimensional vectors live in a space we call $\mathbb{R}^n$. Recall from Chapter 0 that the set $\mathbb{R}^n$ is a Cartesian product made up of n times the set of real numbers, i.e.

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n}. \tag{1.1.5}$$

Each member of this set is a list of n real numbers, and their order inside the list matters - very similar to vectors, be they in row or column form. For this reason, we refer to $\mathbb{R}^n$ as the space of n-dimensional real vectors. As mentioned, in this chapter we use $\mathbb{R}^2$ (the 2-dimensional real space) and $\mathbb{R}^3$ (the 3-dimensional real space) for most ideas and examples.

Looking at vectors in $\mathbb{R}^2$, it is rather straight-forward to calculate their norm: since the origin, the head of the vector and the point v_x form a right triangle (see Figure 1.4), we can use the Pythagorean theorem to calculate the norm of the vector, which is equal to the hypotenous of said triangle:

$$\left\| \vec{v} \right\| = \sqrt{v_x^2 + v_y^2}. \tag{1.1.6}$$

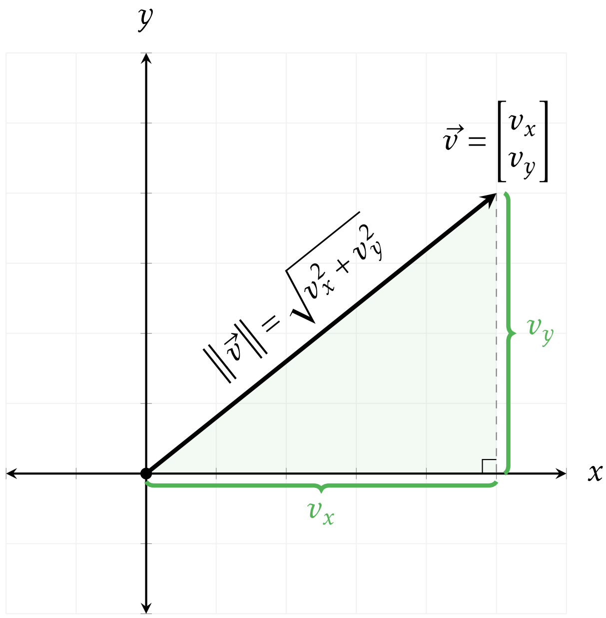


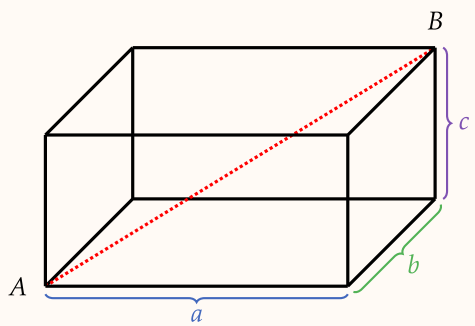
Figure 1.4 Calculating the norm of a 2-dimensional column vector.

In $\mathbb{R}^3$ the norm of a vector $\vec{v}$ is similarly

$$\|\vec{v}\| = \sqrt{v_x^2 + v_y^2 + v_z^2}. \tag{1.1.7}$$

Challange 1.1 Norm of a 3D vector

Show why Equation 1.1.7 is valid, by calculating the length AB in the following figure, depicting a box of sides $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$:



?

Generalizing the vector norms in $\mathbb{R}^2$ and $\mathbb{R}^3$ to $\mathbb{R}^n$ yields the following form:

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2 + \cdots + v_n^2} = \sqrt{\sum_{i=1}^{n} v_i^2}. \tag{1.1.8}$$

Note 1.3 Other norms

The norm shown here is called the 2-norm. There are other possible norm that can be defined, and are used in different situations, such as the 1-norm (also the called **taxicab norm**), general p-norm where $p \geq 1$ is a real number, the zero-norm, the max-norm, and many others. However, for the purpose of this chapter we use only the standard 2-norm, since it is the most useful for describing basic concepts of linear algebra and its uses.

!

Scaling a vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ by a real number α is done by multiplying each of its components by α, i.e.

$$\alpha\vec{v} = \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{bmatrix}. \tag{1.1.9}$$

We can prove Equation 1.1.9 by directly calculating the norm of a scaled vector $\vec{w} = \alpha\vec{v}$:

Proof 1.1 Scaling a column vector

Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{bmatrix}$, where $\alpha \in \mathbb{R}$. Then $\vec{w}$ has the following norm:

$$\begin{aligned}
\left\|\vec{w}\right\| &= \sqrt{\sum_{i=1}^{n} (\alpha v_i)^2} \\
&= \sqrt{(\alpha v_1)^2 + (\alpha v_2)^2 + \cdots + (\alpha v_1)^2} \\
&= \sqrt{\alpha^2 v_1^2 + \alpha^2 v_2^2 + \cdots + \alpha^2 v_n^2} \\
&= \sqrt{\alpha^2 \left(v_1^2 + v_2^2 + \cdots + v_n^2\right)} \\
&= \alpha\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \\
&= \alpha\left\|\vec{v}\right\|.
\end{aligned}$$

This shows that indeed $\vec{w} = \alpha\vec{v}$.

QED

Another idea we can prove in column form is vector normalization (Equation 1.1.1), by showing that dividing each component of a vector by its norm gives a vector of unit norm:

Proof 1.2 Norm of a vector

Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. Its norm is then $\left\|\vec{v}\right\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$. Scaling $\vec{v}$ by $\frac{1}{\|\vec{v}\|}$ yields

$$\hat{v} = \frac{1}{\left\|\vec{v}\right\|} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \frac{1}{\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

The norm of $\hat{v}$ is therefore

$$\begin{aligned} \|\hat{v}\| &= \sqrt{\frac{v_1^2}{v_1^2 + v_2^2 + \cdots + v_n^2} + \frac{v_2^2}{v_1^2 + v_2^2 + \cdots + v_n^2} + \cdots + \frac{v_n^2}{v_1^2 + v_2^2 + \cdots + v_n^2}} \\ &= \sqrt{\frac{1}{v_1^2 + v_2^2 + \cdots + v_n^2}\left(v_1^2 + v_2^2 + \cdots + v_n^2\right)} \\ &= \sqrt{1} = 1, \end{aligned}$$

i.e. $\hat{v}$ is indeed a unit vector.

QED

Example 1.6 Normalizing a vector

Let's normalize the vector $\vec{v} = \begin{bmatrix} 0 \\ 4 \\ -3 \end{bmatrix}$. Its norm is

$$\left\|\vec{v}\right\| = \sqrt{0^2 + 4^2 + (-3)^2} = \sqrt{0 + 16 + 9} = \sqrt{25} = 5.$$

Therefore $\hat{v}$ (the normalized $\vec{v}$) is

$$\hat{v} = \begin{bmatrix} 0 \\ \frac{4}{5} \\ -\frac{3}{5} \end{bmatrix}.$$

By calculating the norm of $\hat{v}$ directly, we can see that it is indeed a unit vector:

$$\|\hat{v}\| = \sqrt{0^2 + \frac{4^2}{5^2} + \frac{3^2}{5^2}} = \sqrt{\frac{0^2 + 4^2 + 3^2}{5^2}} = \sqrt{\frac{16 + 9}{25}} = \sqrt{\frac{25}{25}} = \sqrt{1} = 1.$$

The addition of two column vectors $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is done by adding their respective components together, i.e.

$$\vec{u} + \vec{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}. \tag{1.1.10}$$

TBW: how this addition is the same as the one shown in Figure 1.1.

Note 1.4 No addition of vectors of different number of components!

Two vectors can only be added together if they have the same number of components. The addition of vectors with different number of components is undefined.

1.1.3 Linear combinations, spans and linear dependency

As seen above, scaling a vector by a scalar results in a vector that has the same number of dimensions as the original vector. The same is true for adding two vectors: both of them must be of the same dimension, and the result is also a vector of the same dimension. Therefore, any combination of scaling and addition of vectors results in a vector of the same dimension as the original vector(s). This kind of combination is called a **linear combination**.

Let's define linear combinations a little more formaly:

Definition 1.2 Linear combinations

A linear combination of n vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ of the same dimension, using n scalars $\alpha_1, \alpha_2, \dots, \alpha_n$, is an expression of the form

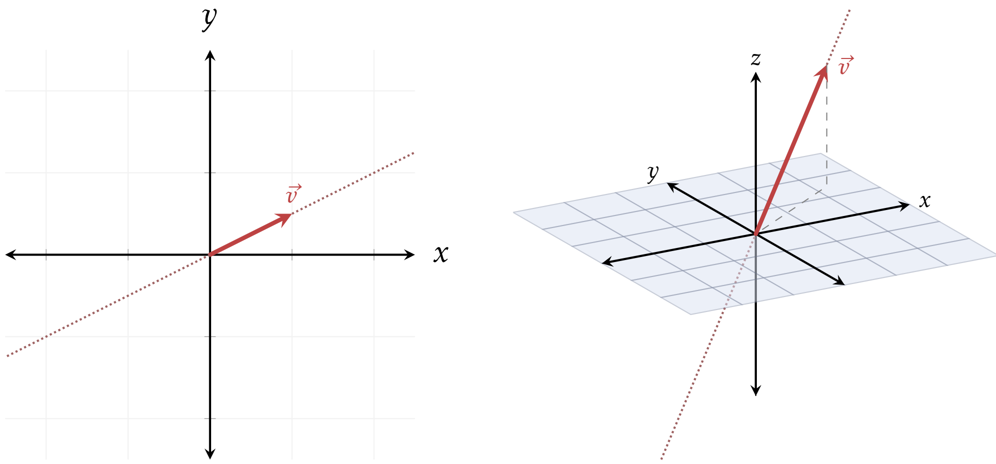
$$\vec{w} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n = \sum_{i=1}^{n} \alpha_i \vec{v}_i. \tag{1.1.11}$$

Linear combinations of real vectors have geometric meaningsc: we start with the set of all linear combinations of a single vector $\vec{v} \in \mathbb{R}^n$, i.e.

$$V = \{\alpha\vec{v} \mid \alpha \in \mathbb{R}\}. \tag{1.1.12}$$

The set V represents a line in the direction of $\vec{v}$ going through the origin (see Figure 1.5). The set V is itself a vector space of dimension 1, and as such a **subspace** of $\mathbb{R}^n$. We say that it is the **span** of the vector $\vec{v}$ (i.e. the vector $\vec{v}$ **spans** the subspace V).

Similarily, the set of all linear combinations of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ that are not scales



1 $\mathbb{R}^2$ **2** $\mathbb{R}^3$

Figure 1.5 The span of a single vector $\vec{v}$, shown as a dashed line: in $\mathbb{R}^2$ (left) and $\mathbb{R}^3$ (right).

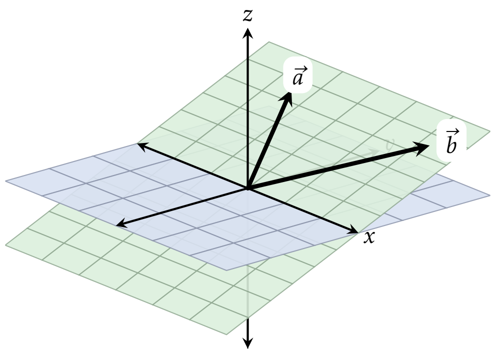


Figure 1.6 Two vectors $\vec{a}$ and $\vec{b}$ span a plane (colored green) in $\mathbb{R}^3$. The xy-plane (i.e. $z = 0$) is shown in blue for emphasis.

of each other (i.e. there is no such $\alpha \in \mathbb{R}$ for which $\vec{v} = \alpha\vec{u}$),

$$V = \{\alpha\vec{u} + \beta\vec{v} \mid \alpha, \beta \in \mathbb{R}\}, \tag{1.1.13}$$

is a plane that goes through the origin (see Figure 1.6). Such vectors are also said to be **non-collinear**.

Example 1.7 Spanning $\mathbb{R}^2$ using two non-collinear vectors

Since any two non-collinear vectors span a 2-dimensional subspace of $\mathbb{R}^n$, in $\mathbb{R}^2$ this means that any vector $\vec{w}$ can be written as a linear combination of any two vectors $\vec{u}, \vec{v}$ that are not a scale of each other. For example, we can take the vector

$$\vec{w} = \begin{bmatrix} 7 \\ -1 \end{bmatrix},$$

and write it as a linear combination of any two non-collinear vectors, say

$$\vec{u} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \ \vec{v} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}.$$

The equation which forces the relation is

$$\begin{bmatrix} 7 \\ -1 \end{bmatrix} = \alpha \begin{bmatrix} 2 \\ -3 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 5 \end{bmatrix},$$

and we should solve it for α and β. This is possible since the equation above is actually a system of two equations in two variables (namely α and β):

$$\begin{cases} 7 = 2\alpha, \\ -1 = -3\alpha + 5\beta. \end{cases}$$

The solution for the system is $\alpha = 3.5$ and $\beta = 1.9$, and therefore

$$\begin{bmatrix} 7 \\ -1 \end{bmatrix} = 3.5 \begin{bmatrix} 2 \\ -3 \end{bmatrix} + 1.9 \begin{bmatrix} 0 \\ 5 \end{bmatrix}.$$

As the reader, you should verify for yourself the above equation.

Generalizing the example above, any vector $\vec{w} = \begin{bmatrix} w_x \\ w_y \end{bmatrix}$ can be written as a linear combination of two vectors $\vec{u} = \begin{bmatrix} u_x \\ u_y \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$, as long as $\vec{u}$ and $\vec{v}$ are non-collinear. Let's prove this:

Proof 1.3 $\mathbb{R}^2$ is spanned by any two non-collinear vectors in $\mathbb{R}^2$

Let $\vec{u}, \vec{v} \in \mathbb{R}^2$ be two non-collinear vectors. Their non-collinearity means that the equation

$$\vec{u} = \alpha \vec{v} \tag{1.1.14}$$

has no solution, i.e. the system

$$\begin{cases} u_x = \alpha v_x \\ u_y = \alpha v_y \end{cases} \tag{1.1.15}$$

has no solution. The system has solution only when $u_x v_y = u_y v_x$, and so the restriction is translated to the simple equation

$$u_x v_y \neq u_y v_x. \tag{1.1.16}$$

The system which defines $\vec{w}$ as a linear combination of $\vec{u}$ and $\vec{v}$ is

$$\begin{cases} w_x = \alpha u_x + \beta v_x \\ w_y = \alpha u_y + \beta v_y \end{cases} \tag{1.1.17}$$

Isolating α using the first equation yields

$$\alpha = \frac{w_x - \beta v_x}{u_x}, \tag{1.1.18}$$

and subtituting it into the second equation yields

$$\beta = \frac{w_y - \alpha u_y}{v_y} = \frac{w_y - \frac{w_x - \beta v_x}{u_x}}{v_y}, \tag{1.1.19}$$

which rearranges into

$$\beta = \frac{u_x w_y - u_y w_x}{u_x v_y - u_y v_x}, \tag{1.1.20}$$

and thus

$$\alpha = \frac{-v_x w_y + v_y w_x}{u_x v_y - u_y v_x}. \tag{1.1.21}$$

We can see that α and β exist iff $u_x v_y \neq u_y v_x$, which is guaranteed by Equation 1.1.16. Therefore, α and β always exist when $\vec{u}$ and $\vec{v}$ are non-collinear, and thus any vector in $\mathbb{R}^2$ can be written as a linear combination of any two non-collinear vectors in $\mathbb{R}^2$, i.e. any two non-collinear vectors in $\mathbb{R}^2$ span $\mathbb{R}^2$.

QED

Going a step further, any three vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ that are not coplanar span a 3-dimensional subspace of $\mathbb{R}^n$ going through the origin. To generalize the notion of collinear and coplanar vectors to higher dimensions we introduct the concept of **linear dependency** of a set of vectors:

Definition 1.3 Linear dependent set of vectors

A set of n vectors

$$S = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} \tag{1.1.22}$$

is said to be linearly dependent if there exist a linear combination

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n = \vec{0}, \tag{1.1.23}$$

and **at least** one the coefficients $\alpha_i \neq 0$.

π

The following examples shows that the definition above reduces to colinarity and coplanary in the case of 2 and 3 vectors:

Example 1.8 Linear dependency of 2 vectors

Let $\vec{u}$ and $\vec{v}$ be two linearly dependent vectors in $\mathbb{R}^n$. Then there exist a linear combination

$$\alpha \vec{u} + \beta \vec{v} = \vec{0},$$

with either $\alpha \neq 0$ or $\beta \neq 0$ (or both). We can look at the different possible cases:

- $\alpha \neq 0,\ \beta = 0$: in this case $\alpha\vec{u} = \vec{0}$, i.e. $\vec{u} = 0$.
- $\alpha = 0,\ \beta \neq 0$: in this case $\beta\vec{v} = \vec{0}$, i.e. $\vec{v} = 0$.
- $\alpha \neq 0,\ \beta \neq 0$: in this case we can rearrange the equation and get

$$\vec{u} = -\frac{\beta}{\alpha}\vec{v},$$

i.e. $\vec{u}$ and $\vec{v}$ are scales of each other and thus are collinear.

What we learn from this is that two vectors form a linearly dependent set if at least one of the is the zero vector, or if they are collinear.

Example 1.9 Linear dependency of 3 vectors

Now, let $\vec{u}, \vec{v}$ and $\vec{w}$ be three linearly dependent vectors in $\mathbb{R}^n$. Then there exists a linear combination

$$\alpha\vec{u} + \beta\vec{v} + \gamma\vec{w} = \vec{0},$$

with either $\alpha \neq 0$ or $\beta \neq 0$ or $\gamma \neq 0$ or any combination where two of the coefficients are non-zero, or all of the coefficients are non-zero. Again, we look at all the possible cases:

- $\alpha \neq 0,\ \beta = \gamma = 0$: we get $\alpha\vec{u} = \vec{0}$, thus $\vec{u} = \vec{0}$.
- $\alpha = 0,\ \beta \neq 0,\ \gamma = 0$: we get $\beta\vec{v} = \vec{0}$, thus $\vec{v} = \vec{0}$.
- $\alpha = \beta = 0,\ \gamma \neq 0$: we get $\gamma\vec{w} = \vec{0}$, thus $\vec{w} = \vec{0}$.
- $\alpha \neq 0,\ \beta \neq 0, \gamma = 0$: we get that $\vec{u}$ and $\vec{v}$ are collinear, since this is exactly as the case for two linearly dependent vectors.
- $\alpha \neq 0,\ \beta = 0, \gamma \neq 0$: similar to the previous case, this time $\vec{u}$ and $\vec{w}$ are collinear.
- $\alpha = 0,\ \beta \neq 0, \gamma \neq 0$: similar to the previous case, this time $\vec{v}$ and $\vec{w}$ are collinear.
- $\alpha \neq 0,\ \beta \neq 0,\ \gamma \neq 0$: by rearranging we get

$$\vec{w} = -\frac{1}{\gamma}\left(\alpha\vec{u} + \beta\vec{v}\right),$$

i.e. $\vec{w}$ lies on the the plane spanned by $\vec{u}$ and $\vec{v}$. If we isolate $\vec{u}$ or $\vec{v}$ instead, we get the same result: the isolated vector is a lienar combination of the other two vectors, and thus lies on the plan spanned by these vectors.

From this example we learn that three vectors form a linearly dependent set if one or more of the vectors is the zero vector, or if any two vectors in the set are collinear, or if all three vectors are coplanar.

Just like the case of 2 and 3 vectors seen above, any set of $m \leq n$ vectors in $\mathbb{R}^n$ that are **not** linearly dependent span an m-dimensional subspace of $\mathbb{R}^n$ (which goes throught the origin) - i.e. any vector $\vec{v} \in \mathbb{R}^n$ can be written as a linear combination of these vectors. We call such a set a **basis set** of $\mathbb{R}^n$.

Example 1.10 Basis sets in n dimensions

The following three vectors are non coplanar (i.e. they are linearly independent), and thus form a basis set of $\mathbb{R}^3$:

$$B = \left\{ \begin{bmatrix} 0 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -5 \end{bmatrix} \right\}.$$

This means that any vector in $\mathbb{R}^3$ can be written as a linear combination of these vectors. We can show this by writing a generic vector $\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ as a linear combination of the vectors:

$$\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ 4 \\ 5 \end{bmatrix} + \beta \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 0 \\ -5 \end{bmatrix},$$

which can be expanded to the system of equations

$$\begin{cases} x = \cancel{0\alpha} + 4\beta + 1\gamma, \\ y = 4\alpha + 2\beta + \cancel{0\gamma}, \\ z = 5\alpha - 2\beta - 5\gamma. \end{cases}$$

The solution of the above system gives the coefficients of the linear combination to yield any vector in $\mathbb{R}^3$:

$$\begin{aligned} \alpha &= -\frac{5x}{31} + \frac{9y}{31} - \frac{z}{31}, \\ \beta &= \frac{10x}{31} - \frac{5y}{62} + \frac{2z}{31}, \\ \gamma &= -\frac{9x}{31} + \frac{10y}{31} - \frac{8z}{31}. \end{aligned}$$

For example, to yield the vector $\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ we sustitute $x = 1,\ y = -1,\ z = 0$ into the above solutions, and get that the following coefficients are needed:

$$\alpha = -\frac{28}{62},\ \beta = \frac{25}{62},\ \gamma = -\frac{38}{62},$$

i.e.

$$-\frac{28}{62} \begin{bmatrix} 0 \\ 4 \\ 5 \end{bmatrix} + \frac{25}{62} \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix} - \frac{38}{62} \begin{bmatrix} 1 \\ 0 \\ -5 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

(you, the reader, should verify this!)

Having described basis sets in somewhat general terms, we can now define them a bit more precisely:

Definition 1.4 Basis sets

Let B be a **linearly independent set** of vectors in $\mathbb{R}^n$. If any vector $\vec{v} \in \mathbb{R}^{\propto}$ can be written as a linear combination of the vectors in B, then B is called a basis set of $\mathbb{R}^n$. The **dimension** of B is the number of vectors in B.

π

The dimension of a basis set B of $\mathbb{R}^n$ is always n. In fact, in a later chapter we will see that the dimension of a vector space is defined by the dimension of its basis sets, i.e. given a vector space V and a basis set $B \subseteq V$, the dimension of V is equal to $|B|$, or mathematically

$$\dim(V) = |B|. \tag{1.1.24}$$

It can be easily shown that any set of vectors in $\mathbb{R}^n$ which has more than n vectors must be a linearly dependent set:

Proof 1.4 Sets with more than n vectors in $\mathbb{R}^n$

Let S be a set of $m \in \mathbb{N}$ vectors in $\mathbb{R}^n$, where $m > n$. Given a vector $\vec{v} \in S$ and the set of all vectors in S except $\vec{v}$ (call this set $\tilde{S}$), there are two possibilities:

- $\tilde{S}$ is a linearly dependent set in $\mathbb{R}^n$. In this case, the addition of $\vec{v}$ doesn't change this fact, i.e. the set S as a whole is linearly dependent.
- The set $\tilde{S}$ is linearly independent, and since it has n vectors it forms a basis set of $\mathbb{R}^n$. Therefore, $\vec{v}$ can be written as a linear combination of the vectors in $\tilde{S}$, and thus the inclusion of $\vec{v}$ in S makes S a linearly dependent set.

QED

Let us now take a vector, for example $\vec{v} = \begin{bmatrix} 1 \\ -3 \\ 7 \end{bmatrix}$, and span it by three different basis sets:

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \quad B_2 = \left\{ \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} \right\}, \quad B_3 = \left\{ \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \right\}.$$

As can be seen in Figure 1.7, for each basis set the coefficients (colored) are different. In this context we call the coefficients the **coordinates** of $\vec{v}$ in that basis set. In the basis set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ the coordinates of $\vec{v}$ are $(1, -3, 7)$ (as we will see next, it is not a coincidense that these are equal to its components as a column vector), and in the basis set $\left\{ \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} \right\}$ its coordinates are $(9, -23, -11)$.

Changing the coordinates of a vector between different basis sets is called **basis trans-**

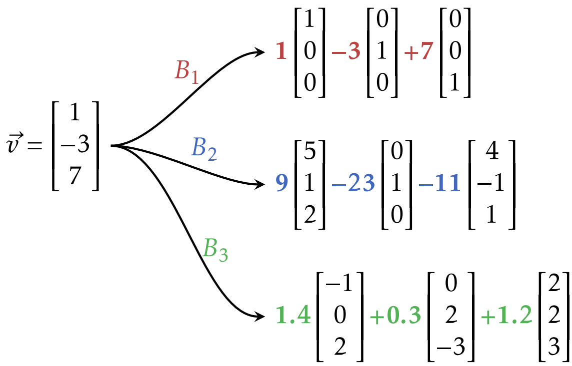


Figure 1.7 The vector $\vec{v} = \begin{bmatrix} 1 \\ -3 \\ 7 \end{bmatrix}$ spanned in three different basis sets.

formation, and is generally done using **matrices**. We will discuss this in more details in the next sections of this chapter. For now, let's look at a graphical representation of a vector being expressed in a different basis set (Figure 1.8): in the figure, we see that the vector $\vec{w} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ can be written in the basis set $B = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \end{bmatrix} \right\}$ using the coefficients 2 and $\frac{1}{2}$, i.e.

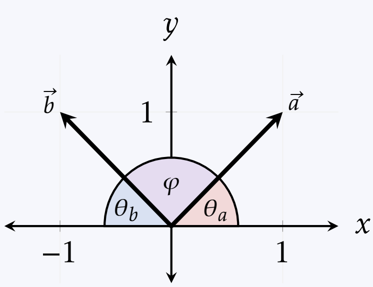
$$\vec{w} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -4 \\ 2 \end{bmatrix}.$$

Therefore, in the basis set B, the coordinates of $\vec{w}$ are $\left(2, \frac{1}{2}\right)$.

A bsis set B in which all vectors are **orthogonal** (i.e. are at 90°) to each other is called a **orthogonal basis set**. If all vectors are unit vectors as well, i.e. their norms all equal to 1, the basis set is then an **orthonormal basis set**.

Example 1.11 Orthogonal and orthonormal basis sets

The vectors $\vec{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ are linearly independent and thus form a basis set of $\mathbb{R}^2$. We can calculate their respective angles in relation to the x-axis (θ_a and θ_b) to find the angle between them (φ):



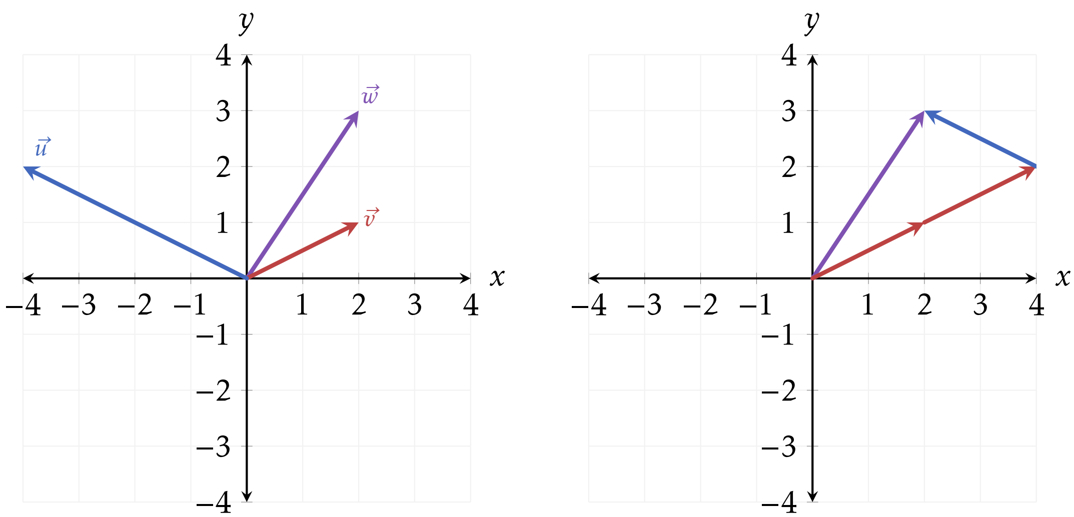


Figure 1.8 The vector $\vec{w} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is spanned using the vectors $\vec{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$, yielding the coordinates $\left(2, \frac{1}{2}\right)$ in the basis set B.

The angle of $\vec{a}$ is

$$\theta_a = \arctan\left(\frac{a_y}{a_x}\right) = \arctan(1) = \frac{\pi}{4}\ (= 45°).$$

Similarily, the angle α_b also equals $\frac{\pi}{4}$. Therefore, $\varphi = 2\frac{\pi}{4} = \frac{\pi}{2}$ $(= 90°)$ - i.e. $\vec{a}$ and $\vec{b}$ are orthogonal, and thus form an orthogonal basis set of $\mathbb{R}^2$.
To get a similar *orthonormal* basis set we can simply normalize the two vectors. We start with $\vec{a}$: its norm is

$$\left\|\vec{a}\right\| = \sqrt{1^2 + 1^2} = \sqrt{2}.$$

Thus, the vector $\hat{a} = \frac{1}{\sqrt{2}}\vec{a} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ is a unit vector. The same argument is valid for $\vec{b}$, i.e. $\hat{b} = \frac{1}{2}\vec{b} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$. We therefore get that

$$\left\{\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}\right\}$$

is an orthonormal basis set of $\mathbb{R}^2$.

Challange 1.2 Orthonormal basis sets of $\mathbb{R}^2$

Show that all orthonormal basis sets of $\mathbb{R}^2$ are rotations of the set

$$\left\{\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}, \begin{bmatrix}-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}\right\}$$

as a whole (i.e. each rotation angle is applied to both vectors).

?

See example below for such sets in $\mathbb{R}^2$ and $\mathbb{R}^3$.

One common orthonormal basis set in any $\mathbb{R}^n$ is the so-called **standard basis set**. We saw the standard basis set in $\mathbb{R}^3$ in Figure 1.7: it is the set $B_1 = \left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$. Note how in this set, each vector has a special structure: one of its components is 1 while the rest are 0. In the first basis vector the non-zero component is the first component of the vector, in the second basis vector it is the second component, and in the third basis vector it is the third component. In $\mathbb{R}^2$ the standard basis set is simply $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}$, and generally in $\mathbb{R}^n$ it is

$$B = \left\{\begin{bmatrix}1\\0\\0\\\vdots\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\\vdots\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\\vdots\\0\\0\end{bmatrix}, \dots, \begin{bmatrix}0\\0\\0\\\vdots\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\\vdots\\0\\1\end{bmatrix}\right\}, \tag{1.1.25}$$

i.e. in the n-th basis vector the n-th component is 1 while the rest are 0. The standard basis vectors are generally labeled as $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n$ - they get the “hat“ symbol since they are all unit length.

In $\mathbb{R}^2$ and $\mathbb{R}^3$ we give $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$ special notations: $\hat{x}$, $\hat{y}$ and $\hat{z}$, respectively (obviously $\hat{z}$ doesn't exists in $\mathbb{R}^2$). For historical reasons, these vectors are sometimes denoted in physics textbooks as $\hat{i}$, $\hat{j}$ and $\hat{k}$.

1.1.4 The scalar product

When given two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ it is often useful to know the angle between them: if the two vectors are linearly dependent then the angle is either $\theta = 0$ if they point in the same direction, or $\theta = \pi$ if the point in opposite directions (remember: we measure angles in radians). Otherwise, the angle θ can take any value in $(0, \pi)$. Angles are always measured on a plane, and in the case of two linearly independent vectors that plane is of course the one spanned by the two vectors (Figure 1.9).

If considering only the plane the vectors span, we can rotate it such that one of the vectors, say $\vec{u}$, lies horizotally (see Figure 1.10). We then drop a perpendicular line from

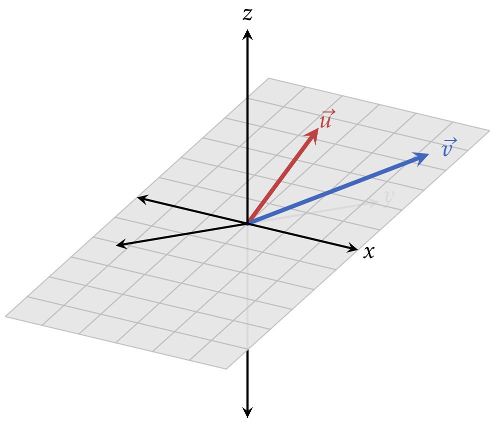


Figure 1.9 The angle between two linearly independent vectors lies on the plane spanned by the vectors.

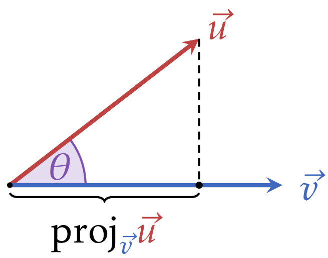


Figure 1.10 The projection of a vector $\vec{u}$ onto another vector $\vec{v}$ in the plane spanned by the two vectors.

the head of the $\vec{u}$ to the horizontal vector $\vec{v}$. We call the length from the origin to the intersection point of $\vec{v}$ and the perpendicular line the **projection** of $\vec{u}$ onto $\vec{v}$, and denote it as $\text{proj}_{\vec{v}}\vec{u}$.

Since the origin, the head of $\vec{u}$ and the intersection point of the perpendicular line with $\vec{v}$ form a right triangle, using basic trigonometry we find that the cosine of the angle θ is

$$\cos(\theta) = \frac{\text{proj}_{\vec{v}}\vec{u}}{\|\vec{u}\|}. \tag{1.1.26}$$

We can now use this construct to define a product between $\vec{u}$ and $\vec{v}$: their **scalar product**. We define it as following:

$$\vec{u}\cdot\vec{v} = \text{proj}_{\vec{v}}\vec{u}\cdot\|\vec{v}\|. \tag{1.1.27}$$

Subtituting Equation 1.1.26 into Equation 1.1.27 gives a very nice relation between the scalar product of two vectors and the angle between them:

$$\cos(\theta) = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|}. \tag{1.1.28}$$

The angle between the two vectors is then isolated by applying the arccos function on the right-hand side of Equation 1.1.28. A common form of this equation is the following:

$$\vec{u}\cdot\vec{v}=\left\|\vec{u}\right\|\left\|\vec{v}\right\|\cos\left(\theta\right). \tag{1.1.29}$$

Note that the scalar product returns a number, i.e. in the terms of linear algebra - a scalar, and hence its name. Since it is commonly denoted with a dot between the two vectors, it is sometimes refered to as the **dot product**. A common notation for the scalar product is the so-called **bracket notation**:

$$\langle\vec{a},\vec{b}\rangle.$$

Sometimes the comma in the notation is replaced by a vertical separator line:

$$\langle\vec{a}\,|\,\vec{b}\rangle.$$

This notation is very common in physics, and especially quantum physics where it is very useful and helps in simplifying many calculations. This will be discussed in more details in chapter/section TBD.

Later in the section we will examine some common properties of the scalar product, and see how we can calculate it directly from the vectors in their column form. Beofre we do that, let's use what we learned about the scalar product so far to solve some easy problems in the examples below.

Example 1.12 Angle between two vectors

Find the scalar product of the vectors

$$\vec{a}=\begin{bmatrix}1\\1\end{bmatrix},\ \vec{b}=\begin{bmatrix}-1\\1\end{bmatrix}.$$

Solution:
As seen in Example 1.11, the angle between $\vec{a}$ and $\vec{b}$ is $\frac{\pi}{2}$. Therefore, their scalar product is

$$\begin{aligned}\vec{a}\cdot\vec{b}&=\left\|\vec{a}\right\|\left\|\vec{b}\right\|\cos\left(\theta\right)\\&=\sqrt{2}\sqrt{2}\cos\left(\frac{\pi}{2}\right)\\&=2\cdot 0=0.\end{aligned}$$

Example 1.13 Scalar product of two vectors

Calculate the scalar product of the two vectors $\vec{u}=\begin{bmatrix}2\\3\\-1\end{bmatrix}$, $\vec{v}=\begin{bmatrix}-1\\0\\2\end{bmatrix}$, given that the angle between them is $\theta\approx 2.069\approx 118.561°$.

Solution:

The norms of the two vectors are

$$\left\|\vec{u}\right\| = \sqrt{2^2 + 3^2 + (-1)^2} = \sqrt{4+9+1} = \sqrt{14} \approx 3.742,$$
$$\left\|\vec{v}\right\| = \sqrt{(-1)^2 + 0^2 + 2^2} = \sqrt{1+4} = \sqrt{5} \approx 2.236.$$

Therefore, their scalar product is

$$\vec{u}\cdot\vec{v} \approx \sqrt{14}\sqrt{5}\cos(2.069) \approx -4.$$

The scalar product of any two vectors $\vec{u},\vec{v}$ has two important properties:

- It is commutative, i.e. $\vec{u}\cdot\vec{v} = \vec{v}\cdot\vec{u}$.
- Scalars can be taken out of the product, i.e. $(\alpha\vec{v})\cdot\vec{u} = \vec{v}\cdot(\alpha\vec{u}) = \alpha\left(\vec{u}\cdot\vec{v}\right)$.
- It equals zero in only one of two cases:
  1.1. One of the vectors (or both) is the zero vector, or
  1.2. The angle θ between the vectors is $\frac{\pi}{2}$, since then $\cos(\theta) = \cos\left(\frac{\pi}{2}\right) = 0$.

When the angle between two vectors is $\frac{\pi}{2}$ (remember: this is equivalent to 90°), we say that the two vectors are **orhtogonal** to eacth other. Note that in the special case of 2- and 3-dimensional we say that the vectors are **perpendicular** to each other.

This is such an important fact that we will put effort into framing it nicely, so you (the reader) could memorize it well. How well should you memorize this? Such that if someone wakes you up in the middle of the night and asked you, you could easily repeat it[1].

$$\vec{u}\cdot\vec{v} = 0$$
$$\Updownarrow$$
$$\vec{u} \text{ and } \vec{v} \text{ are orthogonal}$$

[1] For a humble fee, I'm willing to do this - just write me an email and we can discuss the terms ;)

Calculating the scalar product of two vectors in $\mathbb{R}^n$ using their column form is extremely straight-forward: it is nothing more than the sum of the component-wise product of the two vectors, i.e. given

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix},\ \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

the scalar product $\vec{u} \cdot \vec{v}$ is

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{k=1}^{n} u_i v_i. \tag{1.1.30}$$

Example 1.14 Angle between two vectors

Calculate the scalar product of the two vectors $\vec{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ using the above formula (Equation 1.1.30).

Solution:
We simply substitute $\vec{a}$ and $\vec{b}$ into the equation:

$$\vec{a} \cdot \vec{b} = 1 \cdot (-1) + 1 \cdot 1 = -1 + 1 = 0,$$

which is exactly the result we got using the previous method.

Example 1.15 Scalar product of two vectors - algebraicly

Calculate the scalar product $\vec{u} \cdot \vec{v}$ from Example 1.13 using Equation 1.1.30.

Solution:

$$\vec{u} \cdot \vec{v} = 2 \cdot (-1) + 3 \cdot 0 + (-1) \cdot 2 = -2 - 2 = -4,$$

exactly the result we got in Example 1.13.

The norm of a vector can be calculated using the scalar product: given a vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$,

$$\vec{v} \cdot \vec{v} = v_1 v_1 + v_2 v_2 + \cdots + v_n v_n = v_1^2 + v_2^2 + \cdots + v_n^2 = \left\| \vec{v} \right\|^2. \tag{1.1.31}$$

We therefore usually define the norm in terms of the scalar product:

$$\left\| \vec{v} \right\| = \sqrt{\vec{v} \cdot \vec{v}}. \tag{1.1.32}$$

This might seem unconsequential at the moment, but it will become very useful when we generalize linear algebra to more abstract vector spaces (Chapter 3).

Any vector can be **decomposed** into its projections on n orthogonal directions. In fact, this is exactly what we do when we write a vector as a linear combination of the vectors of an orthogonal basis: consider for example the vector

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

It can be written as the linear combination

$$\vec{v} = v_1\hat{e}_1 + v_2\hat{e}_2 + \cdots + v_n\hat{e}_n = \sum_{i=1}^{n} v_i\hat{e}_i,$$

where in turn any element v_i is the projection of $\vec{v}$ on the basis vector $\hat{e}_i$:

$$v_i = \operatorname{proj}_{\hat{e}_i}\vec{v}, \tag{1.1.33}$$

and thus the component $v_i\hat{e}_i = \left(\operatorname{proj}_{\hat{e}_i}\vec{v}\right)\hat{e}_i$ is itself a vector of norm v_i pointing at the direction $\hat{e}_i$. In general, given an orthogonal basis set $B = \left\{\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_n\right\}$, any vector in $\mathbb{R}^n$ can be decomposed as follows:

$$\vec{v} = \sum_{i=1}^{n} \left(\operatorname{proj}_{\hat{b}_i}\vec{v}\right)\hat{b}_i. \tag{1.1.34}$$

In the case where B is an orthonormal basis set, we know that each of its vector is a unit vector (i.e. $\left\|\vec{b}_i\right\| = 1$), and using Equation 1.1.27 we can re-write Equation 1.1.34 as

$$\vec{v} = \sum_{i=1}^{n} \left(\vec{v}\cdot\hat{b}_i\right)\hat{b}_i. \tag{1.1.35}$$

Example 1.16 Decomposing a vector

EXAMPLE TBD

1.1.5 The cross product

Another commonly used product of two vectors is the so-called **cross product**. Unlike the scalar product, it is only really valid in $\mathbb{R}^2$, $\mathbb{R}^3$ and $\mathbb{R}^7$, of which we will focus on $\mathbb{R}^3$ and touch a bit on its uses in $\mathbb{R}^2$. Also in contrast to the scalar product, the cross product in $\mathbb{R}^3$ results in a vector rather than a scalar - therefore the product is sometimes known as the **vector product**. The cross product uses the notation $\vec{a}\times\vec{b}$, from which it derives its name.

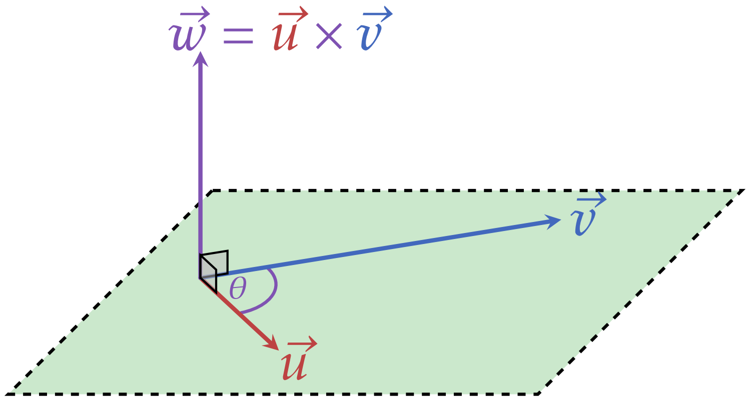


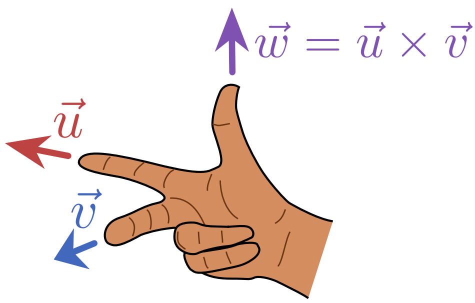
Figure 1.11 The cross product of the vectors $\vec{u}$ and $\vec{v}$ relative to the plane spanned by the two vectors.

Geometrically, the cross product of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^3$ is defined as a vector $\vec{w} \in \mathbb{R}^3$ which is **orthogonal to both** $\vec{u}$ and $\vec{v}$, and has a magnitude

$$\left\|\vec{w}\right\| = \left\|\vec{u}\right\| \left\|\vec{v}\right\| \sin(\theta), \tag{1.1.36}$$

where, like with the scalar product, θ is the angle between $\vec{u}$ and $\vec{v}$.

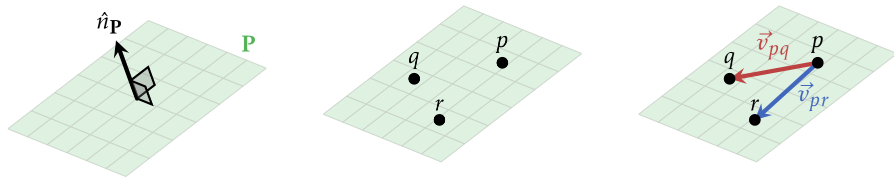
The direction of $\vec{u} \times \vec{v}$ is determined by the **right-hand rule**: using a person's right hand, when $\vec{u}$ points in the direction of their index finger and $\vec{v}$ points in the direction of their middle finger, then vector $\vec{w} = \vec{u} \times \vec{v}$ points in the direction of their thumb:



The cross product is **anti-commutative**, i.e. changing the order of the vectors results in inverting the product:

$$\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u}).$$

When the vectors are given as column vectors $\vec{u} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$, the resulting cross product is



1 The normal vector to **P**.

2 Finding three points on the plane.

3 Finding two vectors on the plane.

Figure 1.12 A normal vector $\hat{n}_{\mathbf{P}}$ to the plane **P**.

$$\vec{u}\times\vec{v}=\begin{pmatrix} u_y v_z - u_z v_y \\ u_z v_x - u_x v_z \\ u_x v_y - u_y v_x \end{pmatrix} \tag{1.1.37}$$

Note 1.5 The cross product of the standard basis vectors

The cross product of two of the standard basis vectors in $\mathbb{R}^3$ is the third basis vector. Its sign ($\pm$) is determined by a cyclic rule:

$$\operatorname{sign}\left(\hat{e}_i\times\hat{e}_j\right)=\begin{cases} 1 & \text{if } (i,j)\in\{(1,2),\ (2,3),\ (3,1)\}, \\ -1 & \text{if } (i,j)\in\{(3,2),\ (2,1),\ (1,3)\}, \\ 0 & \text{otherwise.} \end{cases}$$

!

Challange 1.3 Orthogonalily of the cross product

Using component calculation and utilizing the dot product, show that $\vec{a}\times\vec{v}$ is indeed orthogonal to both $\vec{a}$ and $\vec{b}$.

?

1.1.6 Normal vectors

A special kind of vector in $\mathbb{R}^3$ is the so-called **normal vector** to a plane **P**: this vector, usually denoted as $\hat{n}_{\mathbf{P}}$, is pointing at the orthogonal direction to any vector of the plane (see XXX). Given one knows three points on the plane, its normal vector can be calculated: say the following three points in **P** are given (for visualizing the following steps see YYY):

$$\begin{aligned} p&=(p_x,p_y,p_z)\\ q&=(q_x,q_y,q_z)\\ r&=(r_x,r_y,r_z), \end{aligned} \tag{1.1.38}$$

We can get two vectors lying on the plane by first considering the points as vectors, i.e.

$$\vec{p} = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}, \ \vec{q} = \begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix}, \ \vec{r} = \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix}. \tag{1.1.39}$$

Then, we calculate two vectors on the plane by subtraction, e.g.

$$\begin{aligned} \vec{v}_{pq} &= \vec{q} - \vec{p} = \begin{bmatrix} q_x - p_x \\ q_y - p_y \\ q_z - p_z \end{bmatrix}, \\ \vec{v}_{pr} &= \vec{r} - \vec{p} = \begin{bmatrix} r_x - p_x \\ r_y - p_y \\ r_z - p_z \end{bmatrix}. \end{aligned} \tag{1.1.40}$$

The normal vector $\hat{n}_p$ must be orthogonal to both $\vec{v}_{pq}$ and $\vec{v}_{pr}$ - and so we use the cross product to find its direction:

$$\vec{n}_{\mathbf{P}} = \vec{v}_{pq} \times \vec{v}_{pr} = \begin{bmatrix} (q_y - p_y)(r_z - p_z) - (r_y - p_y)(q_z - p_z) \\ (p_x - q_x)(r_z - p_z) - (r_x - p_x)(q_z - p_z) \\ (q_x - p_x)(r_y - p_y) - (r_x - p_x)(p_y - q_y) \end{bmatrix}. \tag{1.1.41}$$

Normalizing $\vec{n}_{\mathbf{P}}$ will then yield the normal vector $\hat{n}_{\mathbf{P}}$[2].

Note 1.6 Sign of normal vectors

The vector $\vec{m} = -\hat{n}_{\mathbf{P}}$ has all the properties of $\hat{n}_{\mathbf{P}}$, and is indeed a normal vector to **P**. The choice of which of the two vectors to use depends on the application. For now, we do not elaborate on this further.

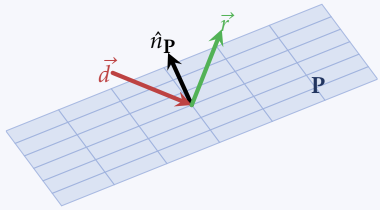
!

To wrap up the vectors section, we present and solve a single problem in the following example.

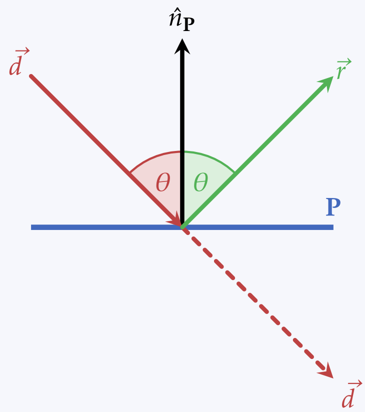
Example 1.17 Reflection of light rays

A ray light hits a mirror, modelled by the plane **P** which is defined by the normal vector $\hat{n}_{\mathbf{P}}$. The direction of the light ray is given by $\vec{d}$. What is the direction of the reflected light ray $\vec{r}$? Recall that both the incident and reflected rays are at the same angle in respect to the normal vector of $\hat{n}_{\mathbf{P}}$, and that the incident ray lie on the plane defined by $\vec{d}$ and $\hat{n}_{\mathbf{P}}$.

[2] I leave this as a challange to the reader, because I'm lazy.



We can rotate our viewpoint of the problem, looking at **P** from the side and in such a way that we look head-on at the plane spanned by $\hat{n}_{\mathbf{P}}$ and $\vec{d}$:



(the dashed red vector in the above figure represents the vector incident ray, $\vec{d}$, moved such that its origin lies at the origin of the other vectors)

As with any vector, we can decompose $\vec{d}$ to its projections on the vectors of an orthonormal basis set (Equation 1.1.35). Since we reduced the problem to two dimensions, we need a basis of two orthonormal directions: we choose one to be $\hat{n}_{\mathbf{P}}$, and the other orthogonal to it (in the figure above it is in the horizontal direction) which we call $\hat{p}$. The decomposition of $\vec{d}$ then reads:

$$\vec{d} = \left(\vec{d} \cdot \hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}} + \left(\vec{d} \cdot \hat{p}\right)\hat{p}.$$

Since there are only two vectors in the basis set $\{\hat{n}_{\mathbf{P}}, \hat{p}\}$, we can actually write the component $\left(\vec{d} \cdot \hat{p}\right)\hat{p}$ as $\vec{d} - \left(\vec{d} \cdot \hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}}$, yielding a rather silly looking expression for $\vec{d}$:

$$\vec{d} = \left(\vec{d} \cdot \hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}} + \left[\vec{d} - \left(\vec{d} \cdot \hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}}\right].$$

However, in closer inspection the above expression is not at all silly, and is actually very similar to the reflected vector $\vec{r}$: since they are both of same norm and oposing directions with respect to the direction $\hat{n}_{\mathbf{P}}$, we can write $\vec{r}$ as

$$\vec{r} = -\left(\vec{d}\cdot\hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}} + \left[\vec{d} - \left(\vec{d}\cdot\hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}}\right].$$

From the above expressions for $\vec{d}$ and $\vec{r}$ we can isolate an expression for $\vec{r}$ as a function of $\vec{d}$ and $\hat{n}_{\mathbf{P}}$:

$$\begin{aligned}\vec{r} &= d - \left(\vec{d}\cdot\hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}} - \left(\vec{d}\cdot\hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}}\\ &= d - 2\left(\vec{d}\cdot\hat{n}_{\mathbf{P}}\right)\hat{n}_{\mathbf{P}}.\end{aligned}$$

1.2 LINEAR TRANSFORMATIONS

1.2.1 Definition

In the previous section we introduced real vectors and their most important properties. In this section we explore a special set of operations that can act on vectors, namely **linear transformations**. As mentioned in Chapter 0, a “transformations“ is simply another name for a function. Thus in this context, linear transformations are some functions that act on vectors: a linear transformation T takes a vector as an input, and outputs another vector, possibly of a different dimension, i.e.

$$T : \mathbb{R}^{n} \to \mathbb{R}^{m}. \tag{1.2.1}$$

What makes linear transformations more “special” than other functions is their property of **linearity**, which entails the following two properties:

- Scalability: for any scalar α and vector $\vec{v}$,

$$T\left(\alpha\vec{v}\right) = \alpha T\left(\vec{v}\right).$$

- Additivity: for any two vectors $\vec{u}, \vec{v}$

$$T\left(\vec{u}+\vec{v}\right) = T\left(\vec{u}\right) + T\left(\vec{v}\right).$$

Example 1.18 A linear transformation

Claim: the following $\mathbb{R}^3 \to \mathbb{R}$ transformation is linear:

$$T\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right) = 2x + 3y - z.$$

Proof: We can show this using the properties of linear transformations.

- Scalability: given a scalar $\alpha \in \mathbb{R}$,

$$T\left(\begin{bmatrix}\alpha x\\ \alpha y\\ \alpha z\end{bmatrix}\right) = 2(\alpha x) + 3(\alpha y) - (\alpha z) = \alpha\,(2x + 3y - z) = \alpha T\left(\begin{bmatrix}x\\ y\\ z\end{bmatrix}\right).$$

- Additivity: given two vectors $\vec{u} = \begin{bmatrix}u_x\\ u_y\\ u_z\end{bmatrix}$ and $\vec{v} = \begin{bmatrix}v_x\\ v_y\\ v_z\end{bmatrix}$,

$$\begin{aligned}T\left(\begin{bmatrix}u_x\\ u_y\\ u_z\end{bmatrix} + \begin{bmatrix}v_x\\ v_y\\ v_z\end{bmatrix}\right) &= T\left(\begin{bmatrix}u_x + v_x\\ u_y + v_y\\ u_z + v_z\end{bmatrix}\right)\\ &= 2\,(u_x + v_x) + 3\left(u_y + v_y\right) - (u_z + v_z)\\ &= T\left(\begin{bmatrix}u_x\\ u_y\\ u_z\end{bmatrix}\right) + T\left(\begin{bmatrix}v_x\\ v_y\\ v_z\end{bmatrix}\right).\end{aligned}$$

Example 1.19 A non-linear transformation

Claim: the following $\mathbb{R}^3 \to \mathbb{R}$ transformation is **not** a linear transformation:

$$T\left(\begin{bmatrix}x\\ y\\ z\end{bmatrix}\right) = 2x^2 + 3y - z.$$

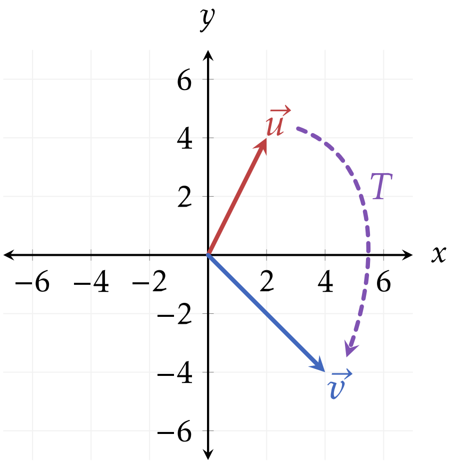
Proof: this time we only need to show a single case where linearity breaks - let's choose *scalability*. Given the vector $\vec{v} = \begin{bmatrix}v_x\\ v_y\\ v_z\end{bmatrix}$, on one hand

$$T\left(\alpha\begin{bmatrix}u_x\\ u_y\\ u_z\end{bmatrix}\right) = 2\,(\alpha u_x)^2 + 3\alpha u_y - \alpha u_z = 2\alpha^2 u_x^2 + 3\alpha u_y - \alpha u_z.$$

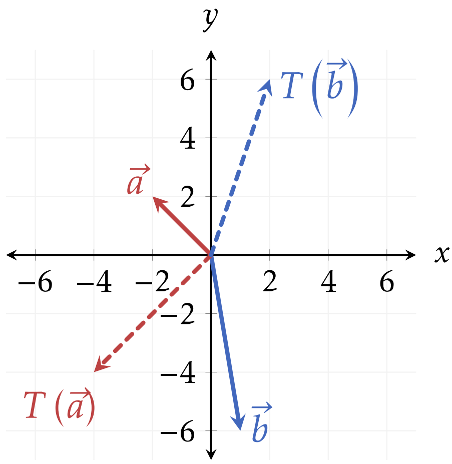
On the other hand

$$\alpha T\left(\begin{bmatrix}u_x\\ u_y\\ u_z\end{bmatrix}\right) = \alpha\left(2u_x^2 + 3u_y - u_z\right) = 2\alpha u_x^2 + 3\alpha u_y - \alpha u_z.$$

For any $a \notin \{0, 1\}$ we get that $T\,(\alpha\vec{v}) \neq \alpha T\,(\vec{v})$. Therefore, T is not linear.



1 The vector $\vec{u} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ is transformed by T yielding the vector $\vec{v} = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$.



2 The vectors $\vec{a} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ and $\vec{b}$ are transformed by the same T.

1.2.2 Developing intuition

Before moving on to explore the algebraic properties of linear transformations, we first shift our focus to gain some intuition about them. Much like in the last section, we do this using graphical representations of linear transformations in $\mathbb{R}^2$ and $\mathbb{R}^3$. We start with a single vector under transformation: let $\vec{u} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x \\ -y \end{bmatrix}. \tag{1.2.2}$$

(to the reader: verify that this transformation is indeed linear)

Applying T to $\vec{u}$ yields the vector $\vec{v} = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$ (see Figure 1.131), i.e. it scales the x-component of $\vec{u}$ by 2 and flippes over its y-component.

If we take other vectors, e.g. $\vec{a} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 1 \\ -6 \end{bmatrix}$ we see that T transforms them in the exact same manner: it scales their x-components by 2 and flipps over their y-components (Figure 1.132). This is a fundamental aspect of linear transformations: they always transform all vectors in the exact same manner. We can use this fact to help visualize transformations, by looking at how they transform the entire space. For example, we can draw all grid lines and observe how they are transformed.

In Figure 1.14 a schematic of $\mathbb{R}^2$ is shown before and after the application of a linear transformation T, by placing a transformed grid (blue) ontop of an untouched grid (gray). In this view, one can see how each point in space is transformed: assuming for example that each two adjacent grid points are 1 unit apart, the gray point at $(-2, 2)$ is

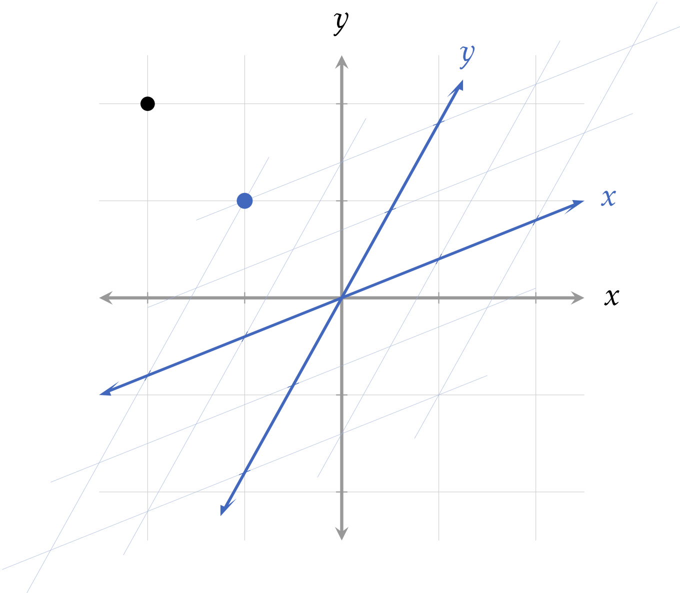


Figure 1.14 $\mathbb{R}^2$ after application of a linear transformation (blue), placed ontop of $\mathbb{R}^2$ before the transformation (gray). Note the black point at the top left at $(-2, 2)$ transforming into the blue point at $(-1, 1)$.

transformed to where the blue point is, i.e. $(-1, 1)$ when measured using the original axes.

For comparison, Figure 1.15 shows a non linear transformation applied to $\mathbb{R}^2$.

Figure 1.14 shows some important properties of linear transformation (cf. Figure 1.15):

1.1. The origin stays at the same place after application of the transformation, i.e. $T\left(\vec{0}\right) = \vec{0}$.

1.2. Parallel lines remain parallel after application of the transformation.

1.3. All areas are scaled by the same amount.

It is rather easy to prove the first two properties.

Proof 1.5 Two properties of linear transformations

1.1. Let T be a transformation that does not perserve the origin, i.e.

$$T\left(\vec{0}\right) = \vec{v} \neq \vec{0}.$$

We can scale $\vec{0}$ by a scalar $\alpha \neq 0$, which would yield

$$T\left(\alpha\vec{0}\right) = T\left(\vec{0}\right) = \vec{v}.$$

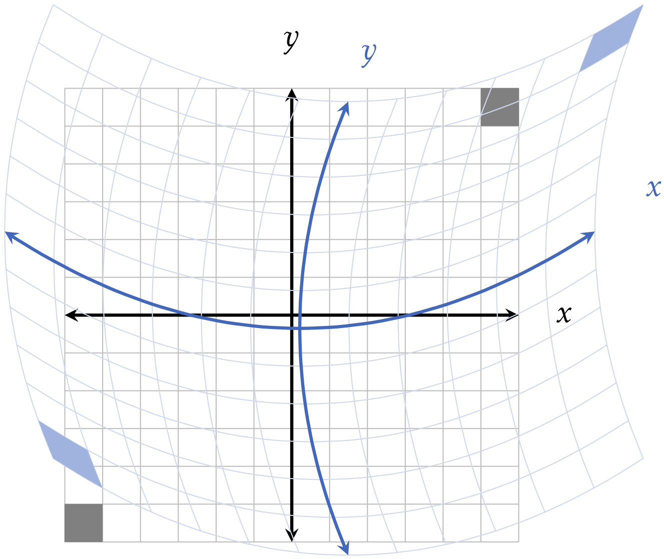


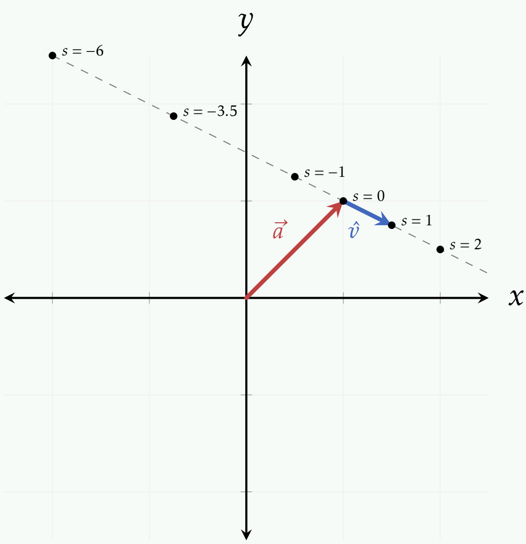
Figure 1.15 A non linear transformation applied to $\mathbb{R}^2$ for comparison.

However, for T to be linear we expect (due to scalability)

$$T\left(\alpha\vec{0}\right) = \alpha\vec{v},$$

but since $\alpha \neq 0$ and $\vec{v} \neq \vec{0}$ this does not happen. Therefore, T can not be linear - and in turn linear transformations must preserve the origin.

1.2. A line is defined using a point $\vec{a}$, and a direction $\hat{v}$ as the set of all the points $\left\{x = \vec{a} + s\hat{v}, s \in \mathbb{R}\right\}$:



Parallel lines have the same direction $\hat{s}$, i.e. $x_1 = \vec{a}_1 + s_1\hat{v}$ and $x_2 = \vec{a}_2 + s_2\hat{v}$ are parallel lines. Applying a linear transformation T to these lines yields (using the two defining properties of linear transformations)

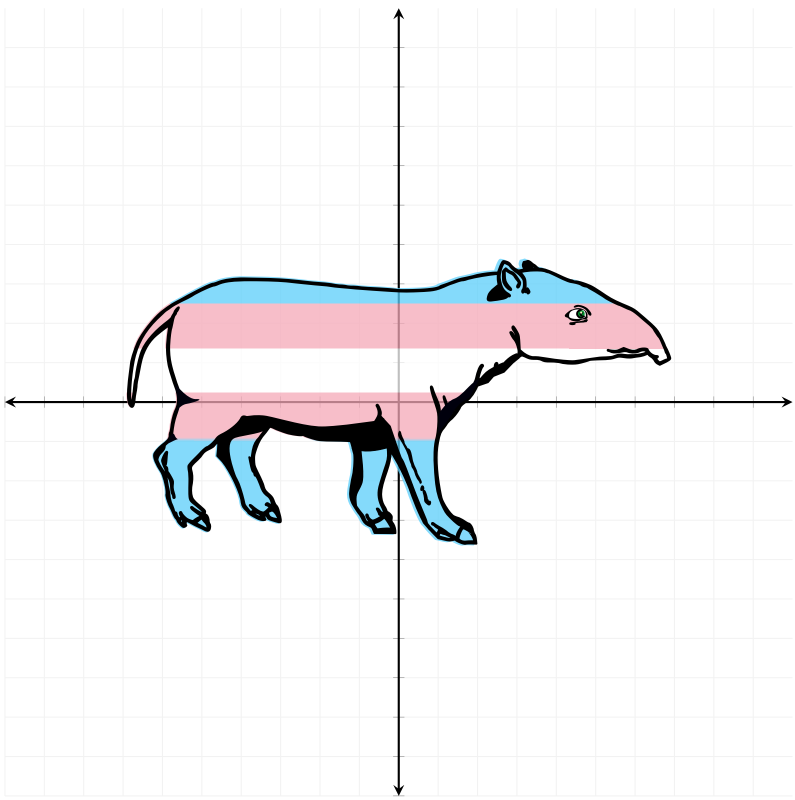
$$T(x_1) = T(\vec{a}_1) + s_1 T(\hat{v}),$$
$$T(x_2) = T(\vec{a}_2) + s_2 T(\hat{v}).$$

We can see that the two right-hand side equations represent two new lines with the same direction, i.e. $T(\hat{v})$. Therefore parallel lines remain parallel under a linear transformation.

QED

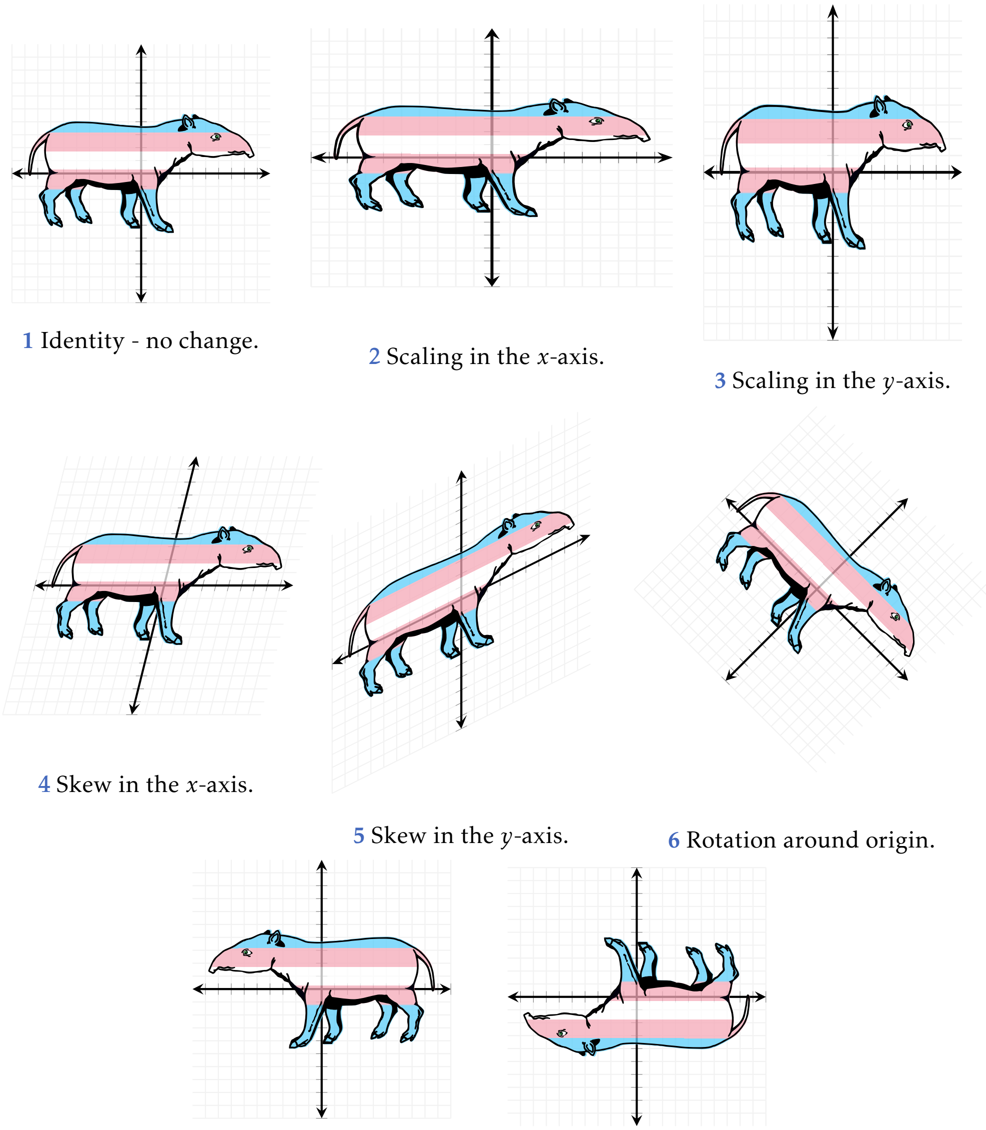
We will prove the the third property (all areas are scaled by the same amount) later in the chapter.

All linear transformations in $\mathbb{R}^2$ can be created by composing transformations from a set of linear transformation which we will refer to as the **fundamental linear transformations**[3]. To visualize these fundamental transformations we apply them on a figure of a tapir[4]:



[3] not an official name.
[4] They are here, they are a trans tapir. Get used to it.

Figure 1.16 shows the fundamental linear transformations applied to our happy tapir.



1 Identity - no change.

2 Scaling in the x-axis.

3 Scaling in the y-axis.

4 Skew in the x-axis.

5 Skew in the y-axis.

6 Rotation around origin.

7 Reflection about line.

8 Reflection through origin.

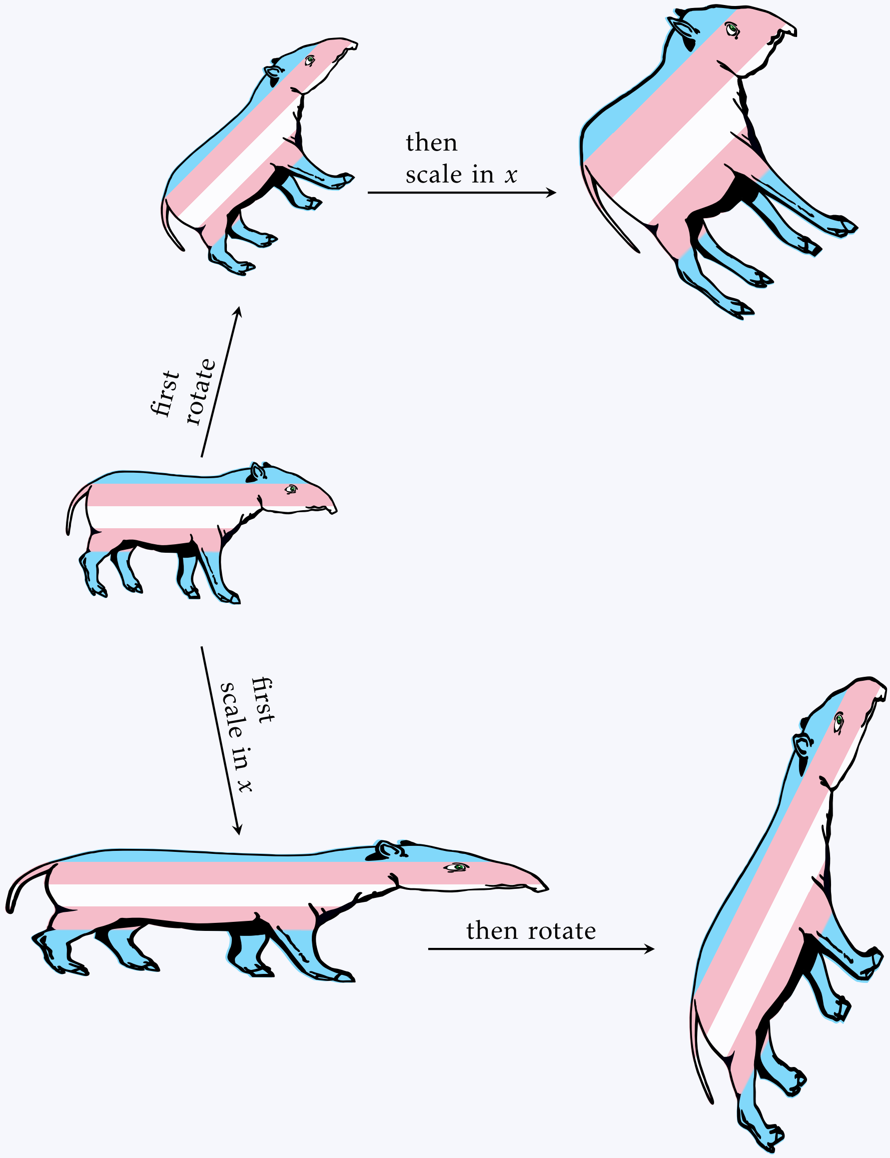
Figure 1.16 The "fundamental" linear transformations, exemplified using a very happy tapir.

Example 1.20 Composing fundamental linear transformations

Given the following two linear transformations:

1.1. Scale by 1.5 in the x-direction,

1.2. Rotate by $\frac{\pi}{4}$ anti-clockwise around the origin, two composite linear transformations can be created: first scale then rotate, and first rotate then scale. As can be seen in the figure bellow, changing the order of composition results in a different linear transformations all together:



This is not a suprising result: in Chapter 0 we learned that function composition is not a commutative operation.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

1.3 MATRICES

1.4 SYSTEMS OF LINEAR EQUATIONS

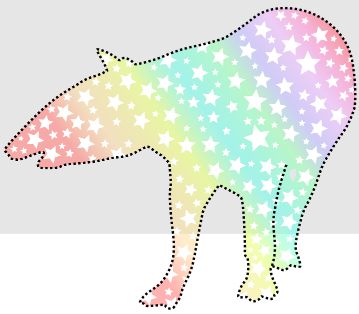
1.5 EIGENVECTORS AND EIGENVALUES

1.6 DECOMPOSITIONS

1.7 SOME REAL LIFE USES OF LINEAR ALGEBRA

1.8 EXERCISES

CHAPTER 2



CALCULUS IN 1D

2.1 SEQUENCES AND SERIES

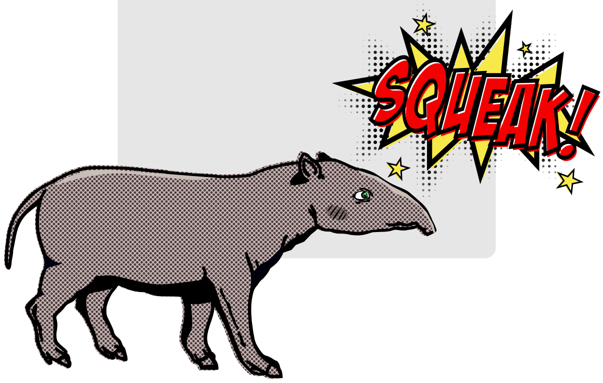
2.2 LIMITS OF REAL FUNCTIONS

2.3 DERIVATIVES

2.4 INTEGRALS

2.5 ANALYZING REAL FUNCTIONS

CHAPTER 3



LINEAR ALGEBRA
(RIGOROUS APPROACH)

Something about formalism, theorems, proofs, etc.

> **Note 3.1 Why present rigorous mathematics in this book?**
>
> Rigorous mathematics is rarely necessary for those who are interested in the tools mathematics provides us with, rather than the full and deep understanding of the concepts these tools are based on. However, it can be useful to students of scientific fields to experience rigorous mathematics at least once in their course of study. Usually, the choice for the topic to be analyzed rigorously is between linear algebra and calculus - for this book the latter was chosen.
>
> **!**

3.1 FIELDS

We begin our dive into the rigorous analysis of linear algebra by defining an algebraic construction call a **field**, which we need in order to properly define vector spaces later. In essence, a field has most of the important properties of the real numbers, namely the closure, commutativity, associativity, identity and inverse of addition and multiplication of any two elements in the field (except the product inverse of the field equivalent object for the number 0). In a later section we will use fields to construct the general notion of **vector spaces**.

Definition 3.1 Field

A field $\mathbb{F}$ is a set of objects together with two operations called **addition** and **multiplication** (denoted $+$ and $\cdot$, respectively), for which the following axioms hold:

- **Closure of under addition and multiplication**: for any $a, b \in \mathbb{F}$,
  1. $(a+b) \in \mathbb{F}$,
  2. $(a \cdot b) \in \mathbb{F}$.

- **Commutativity under addition multiplication**: for any $a, b \in \mathbb{F}$,
  1. $a+b=b+a$,
  2. $a \cdot b = b \cdot a$.

- **Associativity under addition and multiplication**: for any $a, b, c \in \mathbb{F}$,
  1. $a+(b+c)=(a+b)+c$,
  2. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

- **Additive and multiplicative identity**: there exist an element in $\mathbb{F}$ called the *additive identity* and denoted by 0, for which $a+0=a$ for any $a \in \mathbb{F}$.

  Similarity, there exists an element in $\mathbb{F}$ called the *multiplicative identity* and denoted by 1, for which $a \cdot 1 = a$ for any $a \in \mathbb{F}$.

- **Additive and multiplicative inverses**: for any element $a \in \mathbb{F}$ (except the additive identity) there exists:
  1. $b \in \mathbb{F}$ such that $a+b=0$, and
  2. $c \in \mathbb{F}$ such that $a \cdot c = 1$.

  (usually b is denoted as $-a$, while c is denoted as a^{-1})

- **Distributivity of multiplication over addition**: for any $a, b, c \in \mathbb{F}$,

$$a \cdot (b+c) = (a \cdot b) + (a \cdot c).$$

π

3.1.1 Infinite fields

We start with one of the most obvious examples of a field: the real numbers together with the standard addition and product.

Theorem 3.1 $\mathbb{R}$ as a field

The set of real numbers $\mathbb{R}$ forms a field together with the standard addition and product.

We leave the proof of 3.1 to the reader, as it is pretty straight forward using the known properties of the standard addition and product over $\mathbb{R}$ (and rather uninteresting). Instead, we jump forward to using 3.1 for proving the same idea about the complex numbers:

Theorem 3.2 $\mathbb{C}$ as a field and more more more

The set of complex numbers $\mathbb{C}$ forms a field together with the addition and product operations as defined in **??** (namely **??**, **??** and **??**).

Proof 3.1 $\mathbb{C}$ as a field

(note: in the following proof, equalities marked with ! use the respective property of the real numbers)

- **Closure under both operations**: for any two complex numbers $z_1 = a + ib$ and $z_2 = c + id$,
  - Addition: since addition in $\mathbb{R}$ is closed, $(a+c) \in \mathbb{R}$ and $(b+d) \in \mathbb{R}$. Therefore
  $$z = z_1 + z_2 = a + c + (b + d)\mathrm{i}$$
  is also a complex number with $\mathfrak{Re}(z) = a + c$ and $\mathfrak{Im}(z) = b + d$.
  - Multiplication: since multiplication in $\mathbb{R}$ is also closed, $(ac - bd) \in \mathbb{R}$ and $(ad + bc) \in \mathbb{R}$. Therefore
  $$z = z_1 \cdot z_2 = ac - bd + (ad + bc)\mathrm{i}$$
  is a complex number with $\mathfrak{Re}(z) = ac - bdc$ and $\mathfrak{Im}(z) = ad + bc$.
- **Commutativity of both operation**: for any two complex numbers $z_1 = a + ib$ and $z_2 = c + id$,
  - Addition: since addition in $\mathbb{R}$ is commutative, $a+c = c+a$ and $b+d = d+b$. Therefore
  $$z_1 + z_2 = a + c + (b + d)\mathrm{i} \stackrel{!}{=} c + a + (d + b)\mathrm{i} = z_2 + z_2.$$

• Multiplication: since multiplication in $\mathbb{R}$ is also commutative, $ac - bd = ca - db$ and $ad + bc = da + cb$. Therefore

$$z_1 \cdot z_2 = ac - bd + (ad + bc)\mathrm{i} \stackrel{!}{=} ca - db + (da + cb)\mathrm{i} = z_2 \cdot z_1.$$

• **Associativity of both operation**: for any three complex numbers $z_1 = a + \mathrm{i}b$, $z_2 = c + \mathrm{i}d$ and $z_3 = g + \mathrm{i}h$ (where $a, b, c, d, g, h \in \mathbb{R}$[a]),

• Addition: since addition in $\mathbb{R}$ is associative, $a + (c + g) = (a + c) + g$ and $b + (d + h) = (b + d) + h$. Therefore

$$z_1 + (z_2 + z_3) = a + (c + g) + [b + (d + h)]\mathrm{i} \stackrel{!}{=} (a + c) + g + [(b + d) + h]\mathrm{i} = (z_1 + z_2) + z_3.$$

• Multiplication: since multiplication in $\mathbb{R}$ is also associative, the following equalities apply:

$$\begin{aligned}
a \cdot (c \cdot g) &= (a \cdot c) \cdot g, \\
b \cdot (c \cdot h) &= (b \cdot c) \cdot h, \\
a \cdot (d \cdot h) &= (a \cdot d) \cdot h, \\
b \cdot (d \cdot g) &= (b \cdot d) \cdot g, \\
a \cdot (c \cdot h) &= (a \cdot c) \cdot h, \\
a \cdot (d \cdot g) &= (a \cdot d) \cdot g, \\
b \cdot (c \cdot g) &= (b \cdot c) \cdot g, \\
b \cdot (d \cdot h) &= (b \cdot d) \cdot h.
\end{aligned}$$

Therefore,

$$\begin{aligned}
z_1 \cdot (z_2 \cdot z_3) &= a \cdot (c \cdot g) - a \cdot (d \cdot h) - b \cdot (c \cdot h) - b \cdot (d \cdot g) \\
&\quad + [a \cdot (c \cdot h) + a \cdot (d \cdot g) + b \cdot (c \cdot g) - b \cdot (d \cdot h)]\mathrm{i} \\
&\stackrel{!}{=} (a \cdot c) \cdot g - (a \cdot d) \cdot h - (b \cdot c) \cdot h - (b \cdot d) \cdot g \\
&\quad + [(a \cdot c) \cdot h + (a \cdot d) \cdot g + (b \cdot c) \cdot g - (b \cdot d) \cdot h]\mathrm{i} \\
&= (z_1 \cdot z_2) \cdot z_3.
\end{aligned}$$

• **Identity for both operations**:

• Addition: the complex number $0 = 0 + 0i$ is the complex addition identity: for any real number $x \in \mathbb{R}$, $x + 0 = x$. Therefore, for any complex number $z = a + \mathrm{i}b$,

$$z + 0 = a + \mathrm{i}b + 0 + 0i = a + 0 + (b + 0)\mathrm{i} \stackrel{!}{=} a + \mathrm{i}b.$$

• Multiplication: the complex number $1 = 1 + 0\mathrm{i}$ is the complex multiplication identity: for any real number $x \in \mathbb{R}$, $x \cdot 1 = x$ and $x \cdot 0 = 0$. Therefore, for any complex number $z = a + \mathrm{i}b$,

$$z \cdot 1 = (a + \mathrm{i}b) \cdot (1 + 0\mathrm{i}) \stackrel{!}{=} a \cdot 1 - \cancel{b \cdot 0\mathrm{i}^2} + (\cancel{a \cdot 0\mathrm{i}} + b \cdot 1)\mathrm{i} = a + \mathrm{i}b.$$

- **Inverse for both operations**:

  - Addition: for any complex number $z_1 = a + \mathrm{i}b$, the number $z_2 = -a - \mathrm{i}b$ is also a complex number for which

$$z_1 + z_2 = a + \mathrm{i}b + -a - \mathrm{i}b \stackrel{!}{=} a - a + (b - b)\mathrm{i} = 0 + 0\mathrm{i} = 0.$$

  - Multiplication: for any complex number $z = re^{\theta i}$ where $r \neq 0$, the number $z^{-1} = \frac{1}{r}\mathrm{e}^{-\mathrm{i}\theta}$ is also a complex number for which

$$z \cdot z^{-1} = r\mathrm{e}^{\mathrm{i}\theta} \cdot \frac{1}{r}\mathrm{e}^{-\mathrm{i}\theta} \stackrel{!}{=} \frac{r}{r}\mathrm{e}^{\cancel{\mathrm{i}\theta - \mathrm{i}\theta}} = 1 \cdot 1 = 1.$$

  Note: for $z = a + \mathrm{i}b$,

$$z^{-1} = \frac{1}{r}\mathrm{e}^{-\mathrm{i}\theta} = \frac{1}{|z|} \cdot \frac{a - \mathrm{i}b}{|z|} = \frac{1}{|z|} \cdot \frac{\bar{z}}{|z|} = \frac{\bar{z}}{|z|^2}.$$

  Therefore, for any $z \neq 0$, $z^{-1} = \frac{\bar{z}}{|z|^2}$.

- **Distributivity of multiplication over addition**: for any $z_1 = a + \mathrm{i}b$, $z_2 = c + \mathrm{i}d$ and $z_3 = g + \mathrm{i}h$,

$$\begin{aligned}
z_1 \cdot (z_2 + z_3) &= (a + \mathrm{i}b) \cdot (c + \mathrm{i}d + g + \mathrm{i}h) = (a + \mathrm{i}b) \cdot (c + g + [d + h]\mathrm{i}) \\
&= ac + ag + (bd)\mathrm{i}^2 + (bh)\mathrm{i}^2 \\
&+ (ad)\mathrm{i} + (ah)\mathrm{i} + (bc)\mathrm{i} + (bg)\mathrm{i} \\
&= ac + ag - bd - bh + (ad + ah + bc + bg)\mathrm{i} \\
&= ac - bd + (ad + bc)\mathrm{i} + ag - bh + (ah + bg)\mathrm{i} \\
&= (z_1 \cdot z_2) + (z_1 \cdot z_3).
\end{aligned}$$

[a]The letters g and h are used instead of e and f to avoid confusion with Eurler's constant and the common notation for real functions, respectively.

QED

The sets $\mathbb{R}$ and $\mathbb{C}$ are examples of **infinite fields**, since they each have infinite number of elements. The set $\mathbb{Q}$ (rational numbers) can be shown to also be an infinite field, however unlike $\mathbb{R}$ and $\mathbb{C}$ it has **countable** number of elements, i.e. each number in $\mathbb{Q}$ can be assigned an index $1, 2, 3, \ldots$[1].

[1]For proof, see ...

Challange 3.1 $\mathbb{Q}$ as a field

Prove that $\mathbb{Q}$ (together with the usual addition and product operation) is indeed a field.

?

3.1.2 Finite fields

While all three examples of fields we encountered so far have each an infinite number of elements, some fields only have a finite number of elements (called their **order**). For example, consider the set $S = \{0, 1, a, b\}$ and the addition and product operations described using the following tables (left table describes addition, right table describes multiplication):

+	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

·	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

By examining the tables above, several points become clear:

- all the possible combinations of operands in both addition and multiplication give elements from S itself, meaning that the set is closed under both these operations.
- both tables are symmetric around their main diagonal, meaning that both addition and multiplication are commutative operations.
- in the addition table, the first row and first column both show that $x + 0 = x$ for any $x \in S$, meaning that 0 is the additive identity in S.
- in the product table, the second row and second column both show that $x \cdot 1 = x$ for any $x \in S$, meaning that 1 is the multiplicative identity in S.
- in the addition table, the element 0 appears in each row and each column exactly once. This means that every element x has a single additive inverse $y \in S$.
- in the product table, the element 1 appears in each row and each column exactly once, except for the first row and first column. This means that every element $x \neq 0$ has a single multiplicative inverse $z \in S$.

We therefore only need to prove two points to show that S is a field together with the operations described by the above tables: associativity of both operations and distributivity of multiplication over addition. We leave these proofs as a challenge to the reader. Such a field is sometime denoted as $\mathbb{F}_4$. There are, of course, infinitely many finite fields.

3.1.3 Modulo fields

Another example of finite fields are sets of integers of the form $\{0,1,2,3,\ldots,n\}$ where n is a prime, together with **modular addition** and **modular product**. To understand modular arithmetics, we recall the fact that on a circle, an angle can have a negative value but also greater than 360° values are possible (see **??**): 390° is equivalent to 30°, −30° is equivalent to 330°, etc. The set of integer values $0°,1°,2°,\ldots,359°$ on a circle is an example of a modular set: if for example we add together two angles of values deg 100 and 300° we get the equivalent angle deg 60. If we subtract deg 300 from deg 100 the result is an angle of deg 160.

We say that on a circle, the values $360°, 720°, -360°$ etc. are all **congruent** to 0 modulo 360. In mathematical notation we represent this fact as e.g.

$$720 \equiv 0 \pmod{360}. \tag{3.1.1}$$

Note that from this point forward we drop the degrees unit, and deal with pure integers. The notation for the set $\{0,1,2,\ldots,359\}$ is $\mathbb{Z}_{360}$. Generally speaking, the set $\{0,1,2,\ldots,n\}$ is denoted as $\mathbb{Z}_n$.

Note 3.2 About modulo set notation

It is not common to use the notation $\mathbb{Z}_n$ for the modulo-n set, since it is also used for a different algebraic construct, namely the n-adic ring. However, due to the simplicity of the notation, and the fact that we don't discuss rings in this chapter we are using it in this book. Common notations for the set are $\mathbb{Z}/n\mathbb{Z}$ and $\mathbb{Z}/n$.

!

Addition and multiplication on $\mathbb{Z}_n$ is done by the following rather straight forward definition:

Definition 3.2 Operations in $\mathbb{Z}_n$

In the set $\mathbb{Z}_n$ addition and multiplication are defined as the following:

- **Addition**: for any two elements $a,b \in \mathbb{Z}_n,\ a+b := (a+b) \pmod{n}$.
- **Multiplication**: for any two elements $a,b \in \mathbb{Z}_n,\ a\cdot b := (a\cdot b) \pmod{n}$.

π

Example 3.1 Operations in $\mathbb{Z}_n$

The tables below show addition and multiplication results of numbers in different modulo sets $\mathbb{Z}_n$ for some values of n:

n	$2+3$	$2\cdot 3$
4	1	2
5	0	1
6	5	0
7	5	6
8	5	6
9	5	6
10	5	6
11	5	6

n	$4+7$	$4\cdot 7$
8	3	4
9	2	1
10	1	8
15	11	13
20	11	8
27	11	1
28	11	0
30	11	28

Figure 3.1 shows the equivalency between integers and the elements of $\mathbb{Z}_5$.

$\mathbb{Z}$	$\cdots$	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	$\cdots$
$\downarrow$		$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	
$\mathbb{Z}_5$	$\cdots$	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	$\cdots$

Figure 3.1 An example of the periodicity of $\mathbb{Z}_5$: the top numbers are the ordinary integers, each showing their respective congruent modulo 5 below (blue dashed arrow).

Only the sets $\mathbb{Z}_n$ for which n is a prime number are also fields. Let's define this property precisely:

Theorem 3.3 $\mathbb{Z}_p$ is a field

Any modulo set $\mathbb{Z}_p$ where p is a prime number greater than 1 is also a field together with the operations as defined in 3.2.

In order to prove 3.3 we use two lemmas: the first is known as **Bézout's lemma**:

Lemma 3.1 Bézout's lemma

For any two positive integers a, b there exist two integers x, y such that

$$\gcd(a,b) = xa + yb.$$

Note 3.3 $\gcd(a,b)$

$\gcd(a,b)$ is the **greatest common deviser** of the two integers a and b. For example, $\gcd(36,24) = 12$ since the divisors of 36 are $1, 2, 3, 4, 6, 9, 12, 18, 36$, and the divisors of 24 are $1, 2, 3, 4, 6, 8, 12, 24$.

An example of Bézout's lemma is the following:

Example 3.2 Bézout's lemma in action

For the two positive integers $a = 60,\ y = 114$

$$\gcd(60, 114) = 6.$$

Therefore, Bézout's lemma says that there exist two integers x, y such that

$$6 = 60x + 114y.$$

Indeed, two such integers exist: $x = 2$ and $y = -1$.

(SHOULD WE PROVE THE LEMMA?..)

The second lemma we use is the following:

Lemma 3.2 $\gcd(n, p) = 1$

Given a positive prime number p, then for any positive integer $n < p$,

$$\gcd(p, n) = 1.$$

Proving the lemma:

Proof 3.2 $\gcd(n, p) = 1$

We assume that $\gcd(p, n) \neq 1$. Then there exist an integer $a \leq n < p$ which divides both n and p, meaning that p has a divider, contrary to the assumption that p is a prime number. Therefore $\gcd(n, p)$ must equal 1.

QED

Now we can proceed to the proof of 3.3:

Proof 3.3 $\mathbb{Z}_p$ is a field

- **Closure under both operations**: the definition of the modulo operator limit any $M \pmod{p}$ (where $M \in \mathbb{Z}$) to be in $[0, p-1]$. Therefore the result of using the operators given in 3.2 must be within the same range, and thus in $\mathbb{Z}_p$.

- **Commutativity and associativity of both operations**: for any two numbers $a, b \in \mathbb{Z}_p$ the result $a + b$ and $a \cdot b$ under $\mathbb{Z}$ is both commutative and associative. Therefore the result modulo n is the same no matter the order of operations.

- **Additive identity**: the number $0 \in \mathbb{Z}_p$ is the additive identity, since for each $a \in \mathbb{Z}_p,\ a + 0 = a$.

- **Multiplicative identity**: the number $1 \in \mathbb{Z}_p$ is the additive identity, since for each $a \in \mathbb{Z}_p$, $a \cdot 1 = a$.

- **Additive inverse**: for each $a \in \mathbb{Z}_p$ the element $n = p - a$ is in $\mathbb{Z}_p$ since $p > a$. Adding n to a results in 0:

$$a + n = a + (p - a) = p \equiv 0 \pmod{p}.$$

- **Multiplicative inverse**: let $a \in \mathbb{Z}_p$ and $a \neq 0$. Since p is a prime, $\gcd(a, p) = 1$ and from Bézout's theorem we know that there exist two integers x, y such that

$$xa + yp = 1.$$

Rearrangement gives $p = \frac{1-xa}{y}$ meaning that p divides $1 - xa$, and thus

$$xa \equiv 1 \pmod{p}.$$

Therefore x is the multiplicative inverse of a.

- **Distributivity of multiplication over addition**: ...

QED

The only part of the proof that uses the fact that p is a prime number is the multiplicative inverse. When n is not a prime, $\mathbb{Z}_n$ is not a field.

Challange 3.2 $\mathbb{Z}_n$ is not a field when n is not a prime number

Prove that the modulo set $\mathbb{Z}_n$ where n is **not** a prime number, is not a field. (**hint**: what property of prime numbers is used in the above proof to show that there is always a multiplicative inverse in $\mathbb{Z}_p$ where p is prime?)

?

3.2 VECTOR SPACES

As we've seen in Chapter 1 vectors are found at the heart of linear algebra. We first defined them in a geometric way as objects with magnitude and direction, and later as lists of real numbers, analyzing the connections between these two mostly parallel definitions. We also spoke about vector spaces of the type $\mathbb{R}^n$ as the structures vectors exist in. However, we haven't defined vectors nor vector spaces formally - which is exactly what we do in this section, by defining the concept of **vector spaces**.

Note 3.4 $\mathbb{R}^n$ as a guide to general vector spaces

While reading the definition below, it is worthwhile to reflect on each of the given axioms as it relates to the familiar vector space $\mathbb{R}^n$.

!

Definition 3.3 Vector space

A vector space over a field $\mathbb{F}$ is a set V which, together with two operations described below, fulfils a list of axioms. The two operations are

- Vector addition: an operation which takes two elements of V and returns a single element of V, i.e. $+ : V \times V \to V$.

- Scalar multiplication: an operation which takes a single element of $\mathbb{F}$ and a single element of V and returns a single element of V, i.e. $\cdot : \mathbb{F}, V \to V$.

The axioms to be fulfilled are:

- **Commutativity of vector addition**: for any $u, v \in V$,

$$u + v = v + u.$$

- **Associativity of vector addition**: for any $u, v, w \in V$,

$$u + (v + w) = (u + v) + w.$$

- **Additive identity**: there exist an element $0 \in V$ for which, for any $v \in V$,

$$v + 0 = v.$$

- **Scalar multiplicative identity**: for any $v \in V$

$$1 \cdot v = v,$$

  where 1 is the multiplicative identity in $\mathbb{F}$.

- **Additive inverse**: for any $v \in V$ there exist an element $u \in V$ for which

$$v + u = 0.$$

- **Associativity of scalar multiplication**: for any $\alpha, \beta \in \mathbb{F}$ and $v \in V$

$$\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v,$$

  where $\alpha\beta$ is the multiplication defined for $\mathbb{F}$.

- **Distributivity of vector addition**: for any $\alpha \in \mathbb{F}$ and $u, v \in V$,

$$\alpha \cdot (u + v) = (\alpha \cdot u) + (\alpha \cdot v).$$

- **Distributivity of scalar addition**: for any $\alpha, \beta \in \mathbb{F}$ and $v \in V$,

$$(\alpha + \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v).$$

The elements of V are then called **vectors**, and the elements of $\mathbb{F}$ are called **scalars**.

Since we discussed $\mathbb{R}^n$ thoroughly in Chapter 1, let's prove that it is indeed a vector space under the above definition. First, the claim:

Theorem 3.4 $\mathbb{R}^n$ is a vector space

The set of elements of the form

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

where $v_i \in \mathbb{R}$, forms a vector space over $\mathbb{R}$ together with the following two operations:

- **Vector addition**:

$$\vec{u} + \vec{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}.$$

- **Scalar multiplication**:

$$\alpha \cdot \vec{v} = \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{bmatrix}.$$

The proof itself is pretty easy, based on the fact that $\mathbb{R}$ is a field:

Proof 3.4 $\mathbb{R}^n$ is a vector space

Since the results of both operations defined for $\mathbb{R}^n$ only depend on the respective components of a vector $v \in \mathbb{R}^n$, all the axioms of a vector space apply, since they derive directly from the fact that $\mathbb{R}$ is a field. As an example, we will elaborate on two of the axioms:

- **Additive inverse**: Given a vector $\vec{v} \in \mathbb{R}^n$, each of its components v_i has an inverse under $\mathbb{R}$, namely $-v_i$. Therefore,

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{bmatrix} = \begin{bmatrix} v_1 - v_1 \\ v_2 - v_2 \\ \vdots \\ v_n - v_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{0},$$

which is the additive identity in $\mathbb{R}^n$.

- **Distributivity of vector addition**: for each component of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$, given the rules for vector addition and scalar multiplication, together with the distributivity of numbers in $\mathbb{R}$:

$$\begin{aligned} \alpha(\vec{u}+\vec{v}) &= \alpha \left(\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \right) = \alpha \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} \\ &= \begin{bmatrix} \alpha u_1 + \alpha v_1 \\ \alpha u_2 + \alpha v_2 \\ \vdots \\ \alpha u_n + \alpha v_n \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \vdots \\ \alpha u_n \end{bmatrix} + \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{bmatrix} = \alpha\vec{u} + \alpha\vec{v}. \end{aligned}$$

QED

(it is adviceable for the reader to go over the rest of the axioms and prove them for $\mathbb{R}^n$)

3.3 EXERCISES

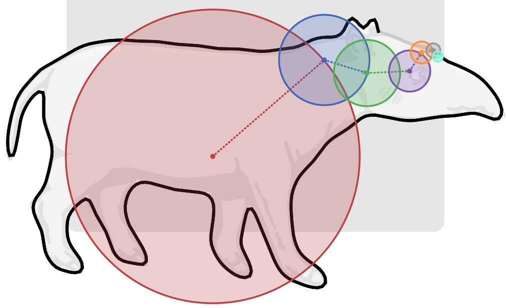
CHAPTER

4

DIFFERENTIAL EQUATIONS

4.1 EXERCISES

CHAPTER 5

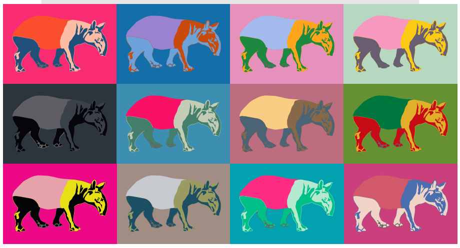


THE FOURIER TRANSFORM

5.1 EXERCISES

CHAPTER

6



SYMMETRY GOURPS

6.1 EXERCISES