

BIO311: Population Ecology  
*Mathematical Tools*

Timothée Bonnet &<sup>1</sup> Koen van Benthem

`timothee.bonnet@ieu.uzh.ch`  
`koen.vanbenthem@ieu.uzh.ch`

Spring 2016

<sup>1</sup>This document was co-authored by Tina Cornioley

# Contents

0.1	Why? . . . . .	2
0.2	The basic basics . . . . .	2
<b>1</b>	<b>Introduction to Calculus</b>	<b>4</b>
1.1	Derivatives . . . . .	4
1.2	Integrals . . . . .	13
1.3	Differential Equations . . . . .	20
<b>2</b>	<b>Introduction to Matrix Algebra</b>	<b>24</b>
2.1	Vectors . . . . .	24
2.2	Vector Transformations . . . . .	30
2.3	Eigenvalues and eigenvectors . . . . .	38

In this document, you will find in the main text the theory describing the essential mathematical tools for the course in population ecology. The theory is followed by exercises which difficulty is indicated by stars, \* . The more stars, the more difficult the exercise is. We expect you to solve the one star exercises whereas the several stars exercises are for the more mathematically inclined among you.

In white boxes such as this one, you will find additional information which will help you understand the main theory but is not essential. In blue boxes, you will find examples illustrating the theory.

## 0.1 Why?

Why more mathematics? Well, above all because it is important. If you don't believe us, please take a look at: <https://peerj.com/articles/285.pdf>. Besides, it is also fun! Although that last part is probably a bit subjective... Here we go...

## 0.2 The basic basics

Before we start, let's refresh the basics. First all, when we discuss multiplications we use  $\cdot$  (dot) rather than  $\times$ . When we want to say that 3 times 4 equals twelve, we write:

$$3 \cdot 4 = 12. \quad (1)$$

When a number ( $x$ ) is multiplied by itself, we use the power notation. So for  $x$  to the power 2:

$$x \cdot x = x^2. \quad (2)$$

And for higher powers:

$$x^5 = x \cdot x \cdot x \cdot x \cdot x. \quad (3)$$

We can also multiply different powers of  $x$  with each other:

$$x^2 \cdot x^3 = (x \cdot x \cdot x) \cdot (x \cdot x) = x \cdot x \cdot x \cdot x \cdot x = x^5. \quad (4)$$

Or more generally:

$$x^a \cdot x^b = x^{a+b}. \quad (5)$$

If for example  $a = 3$  and  $b = -1$ , the previous equation shows us that:

$$x^3 \cdot x^{-1} = x^{3-1} = x^2. \quad (6)$$

But we also know that  $x^3 = x \cdot x \cdot x$  and  $x^2 = x \cdot x$ . It thus has to be the case that:

$$x^{-1} = \frac{1}{x}. \quad (7)$$

And the more general case:

$$x^{-a} = \frac{1}{x^a}. \quad (8)$$

Finally there is the square root of  $x$  ( $\sqrt{x}$ ), this is the number that when squared equals  $x$ :

$$\sqrt{x} \cdot \sqrt{x} = x. \quad (9)$$

Because we know from equation 5 that:

$$x^{\frac{1}{2}} \cdot x^{\frac{1}{2}} = x^{(\frac{1}{2}+\frac{1}{2})} = x^1 = x, \quad (10)$$

we can see that:

$$\sqrt{x} = x^{\frac{1}{2}}. \quad (11)$$

In general the  $a$ -th root of  $x$  is that number that if it is raised to the power  $a$  it equals  $x$ :

$$(\sqrt[a]{x})^a = x. \quad (12)$$

Similar to the case of the square root, also higher order roots can be written as  $x$  to the power of a fraction:

$$\sqrt[a]{x} = x^{\frac{1}{a}}. \quad (13)$$

One last thing that can come in useful:

$$x^0 = 1. \quad (14)$$

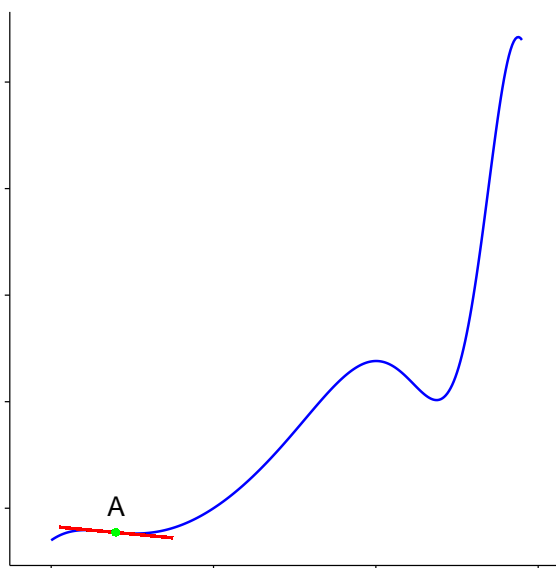
# Chapter 1

## Introduction to Calculus

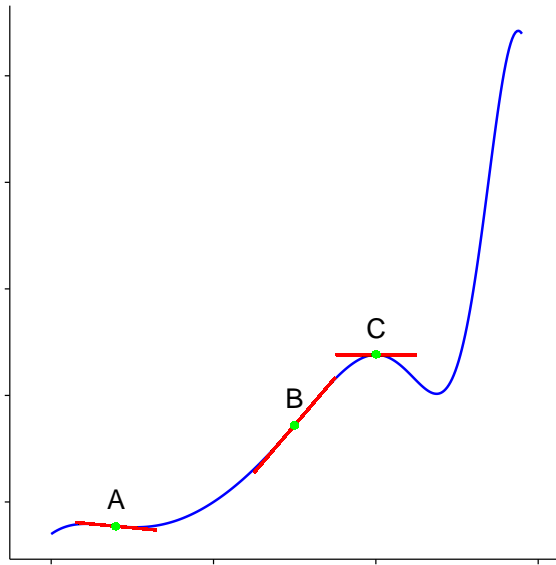
### 1.1 Derivatives

#### 1.1.1 Graphical definition

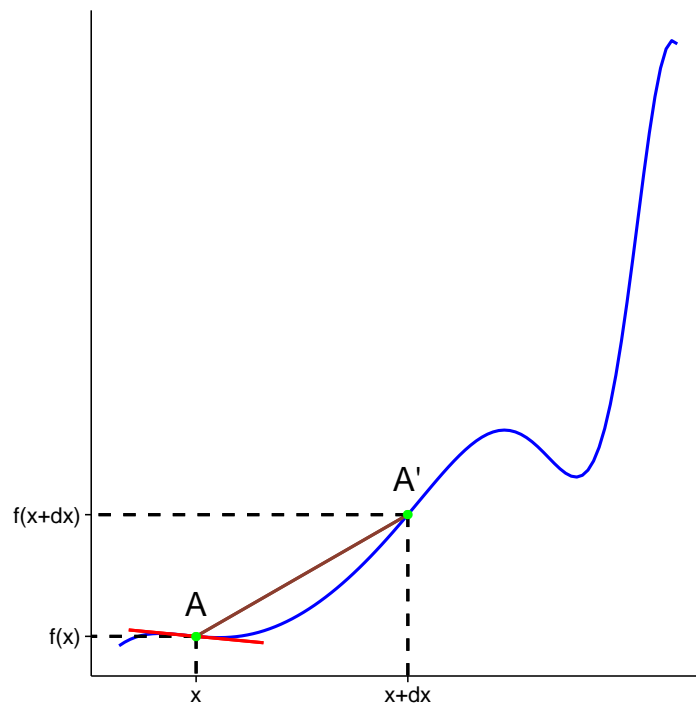
Graphically, a derivative is **the slope of a tangent line at a point on a curve**. In the figure below, the tangent in point A is the red line.



If the blue curve is a mountain, you can picture the tangent as the slope of your skis. For example by looking at the curve in the figure on the next page, you can see that the slope of the tangent is negative in point A (when you move from left to right you go down), positive in point B (when you move from left to right you go up) and zero in point C.



To calculate the slope of a line, we need to know two points the line goes through. We now take a second point, point  $A'$  that lies  $dx$  away from the first point. This is illustrated below.



The slope of the line between point A and point  $A'$  is given by the rise (height) over the run (length). Let the curve be described by  $y = f(x)$ . The coordinate

of point A is  $(x, f(x))$  and the coordinate of point A' is  $(x + dx, f(x + dx))$ . Thus the rise of the line between point A and point A' is given by

$$df = f(x + dx) - f(x) \quad (1.1)$$

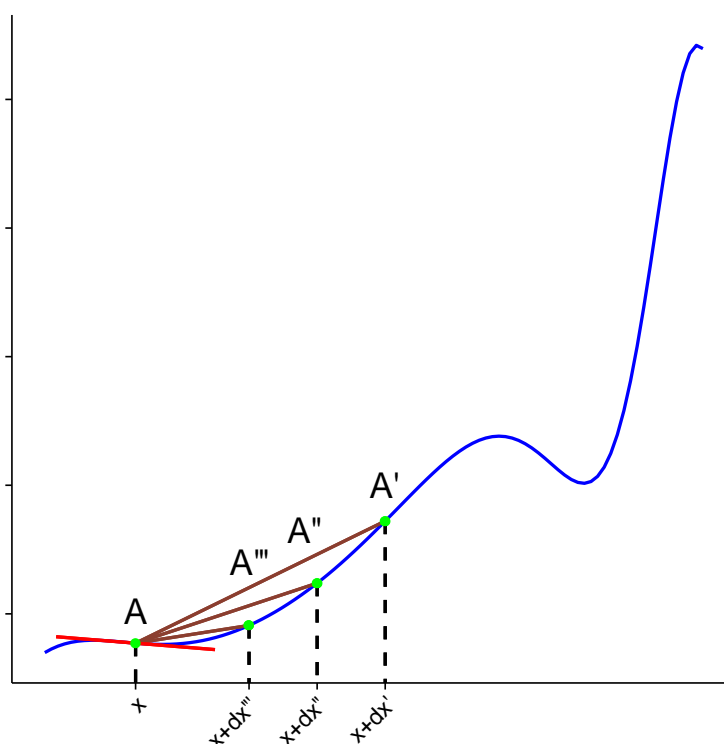
and the run is

$$dx = (x + dx) - x. \quad (1.2)$$

So the slope is given by

$$\frac{df}{dx} = \frac{f(x + dx) - f(x)}{dx}. \quad (1.3)$$

But if you look at the graph, the slope of the line between point A and point A' (brown line) is not identical to the slope of the tangent (red line). By reducing the distance between point A and A', which is the same as reducing  $dx$ , we can find lines that have slopes more similar to the tangent of point A. This is what we did in the next figure. We chose two more points, A'' and A''', that are both closer to A than A'. Note that the closer these points get to A, the more the line between them and A resembles the tangent of the curve in point A.



### 1.1.2 Formal definition

If we choose the points A and A' very close together, this line will have the same slope as the tangent in point A. We thus want  $dx$  to be reduced to such

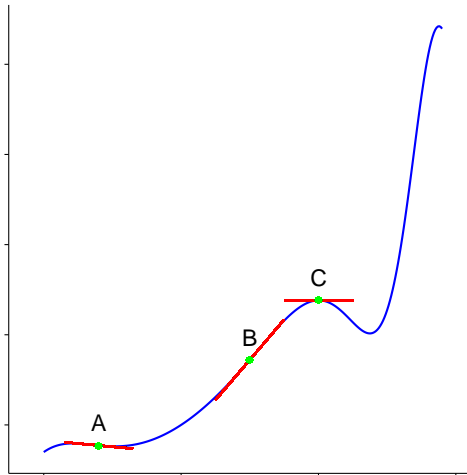
an extent that it tends towards zero. More precisely, we are interested in the limit when  $dx$  tends towards zero:

$$f'(x) = \lim_{dx \rightarrow 0} \frac{f(x + dx) - f(x)}{dx}. \quad (1.4)$$

Equation 1.4 is the formal definition of a derivative. In this equation we used the shorthand notation for  $\frac{df}{dx}$ :  $f'(x)$ .

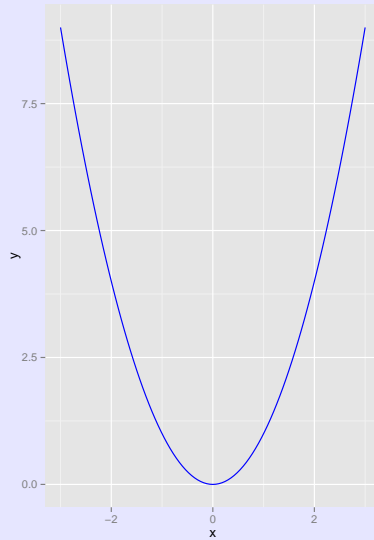
### 1.1.3 Interpretation of the derivative

The first interpretation of the derivative and the only one we consider in this course is **rate of change**. The derivative describes how a function changes with respect to the independent variable  $x$ . It tells us whether the function is decreasing, increasing or remaining constant at given points. For example, in the figure below, at point A the derivative is negative (we know this from the slope of the tangent). Thus at this point the function is decreasing. Similarly, in point B the derivative is positive so the function at this point is increasing. It is important that you understand this interpretation as it is frequently used in the rest of this course to describe population growth.



#### Example





As an example, let us work out the derivative of a simple quadratic function. The function of the curve in the figure is given by:

$$f(x) = x^2. \quad (1.5)$$

The derivative is given by the formal definition of the derivative (see previous section, equation 1.4) and by substituting  $f(x)$  by equation 1.5, we find:

$$f'(x) = \lim_{dx \rightarrow 0} \frac{(x + dx)^2 - x^2}{dx}. \quad (1.6)$$

When we multiply the brackets out, we are left with:

$$f'(x) = \lim_{dx \rightarrow 0} \frac{x^2 + 2xdx + dx^2 - x^2}{dx} \quad (1.7)$$

By simplification we get:

$$f'(x) = \lim_{dx \rightarrow 0} 2x + dx \quad (1.8)$$

But when  $dx$  tends towards zero we are only left with:

$$f'(x) = 2x \quad (1.9)$$

This means that the instantenous rate of change of any point on the curve is  $2x$ . Or in other words: the derivative of  $x^2$  with respect to  $x$  is  $2x$ . For instance, the slope of the tangent in  $x = 4$  is  $2 \cdot 4 = 8$ . Thus the instantenous rate of change in  $x = 4$  is 8.

#### 1.1.4 General rule

We do not demonstrate this here but in general the following holds:

$$f(x) = x^n \quad \rightarrow \quad f'(x) = nx^{n-1} \quad (1.10)$$

When a function is the sum of two functions (for example  $h(x)$  and  $u(x)$ ), its derivative is the sum of the derivatives of these two functions, thus:

$$f(x) = u(x) + h(x) \quad \rightarrow \quad f'(x) = u'(x) + h'(x) \quad (1.11)$$

When a function is multiplied by a constant, for example  $c$ , then the derivative is the derivative of the function multiplied by the constant.

$$f(x) = c \cdot u(x) \quad \rightarrow \quad f'(x) = c \cdot u'(x) \quad (1.12)$$

### 1.1.5 Chain rule

The general rule presented in the previous section is nice, but not very helpful when we move to more complicated formulas. What happens for example if functions become nested, what if for example a function  $h$  depends on  $u$  and  $u$  in turn depends on  $x$ . How does  $f$  now change if  $x$  changes? We are thus faced with the following function:

$$f(x) = h(u(x)) \quad (1.13)$$

Formally the chain rule states

$$f'(x) = h'(u)u'(x) = \frac{df}{du} \frac{du}{dx} \quad (1.14)$$

#### Example

So far it is all a bit vague, so let's move to an example:

$$f(x) = (x^2 - 4)^5 \quad (1.15)$$

Equation 1.15 can be decomposed in two functions; an “outer” function and an “inner” function. The outer function is raising to the power 5. The inner function is the content of the brackets. Let us call the inner function of  $f(x)$ ,  $u(x)$ .

$$u(x) = (x^2 - 4) \quad (1.16)$$

Let us call the outer function  $h(u)$

$$h(u) = u^5 \quad (1.17)$$

With the chain rule

$$f'(x) = h'(u) \cdot u'(x) \quad (1.18)$$

$$h'(u) = 5u^4 \quad (1.19)$$

$$u'(x) = 2x \quad (1.20)$$

We have now calculated all the separate terms in equation 1.18, all we need to do is fill in equations 1.19, 1.20 and 1.16 to obtain our final answer:

$$f'(x) = 5(x^2 - 4)^4 2x. \quad (1.21)$$

### 1.1.6 Product rule

Another useful rule to find the derivative of a function is the product rule. The product rule is used to find the derivative of the multiplication of two functions. Formally the product rule is

$$(h(x) \cdot g(x))' = h'(x) \cdot g(x) + h(x) \cdot g'(x) \quad (1.22)$$

#### Example

Let us work out as an example the derivative of the following function:

$$f(x) = 5x(x^2 - 3) \quad (1.23)$$

$f(x)$  is composed of two functions; Let

$$f(x) = h(x) \cdot g(x) \quad (1.24)$$

with

$$h(x) = 5x \quad (1.25)$$

and

$$g(x) = x^2 - 3 \quad (1.26)$$

The product rule states

$$f'(x) = h'(x)g(x) + h(x)g'(x) \quad (1.27)$$

The derivatives of the functions  $h(x)$  and  $g(x)$  are:

$$h'(x) = 5 \quad (1.28)$$

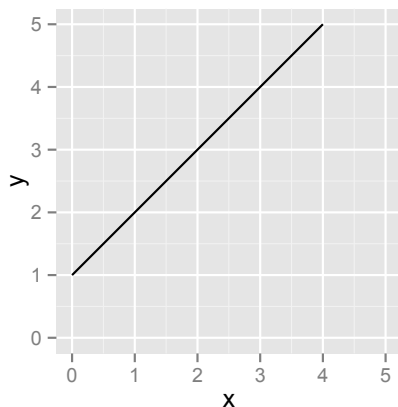
$$g'(x) = 2x \quad (1.29)$$

We now have all the information to fill in equation 1.27, by doing so we find:

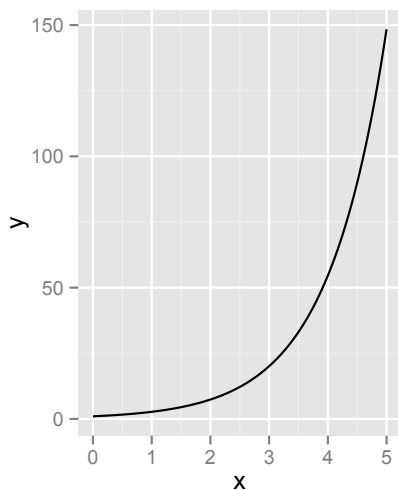
$$\begin{aligned} f'(x) &= 5(x^2 - 3) + (5x) \cdot (2x) \\ &= 5x^2 - 15 + 10x^2 \\ &= 15(x^2 - 1) \end{aligned} \quad (1.30)$$

### 1.1.7 Exercises

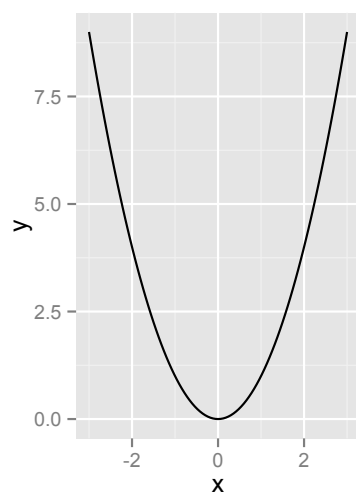
\* **Exercise 1.1**      Answer the following questions by looking at the graphs



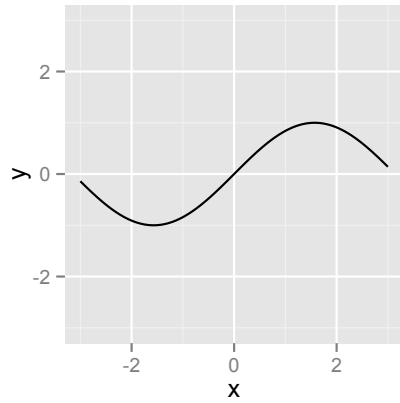
1. By looking at the graph, give the derivative of the function. Bonus question: Find the equation from the graph.



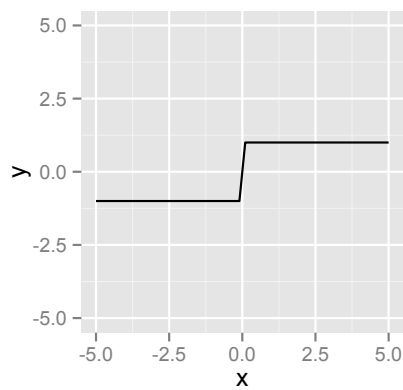
2. Show in the graph the part where the derivative is larger than 0. Bonus question: Find the equation from the graph.



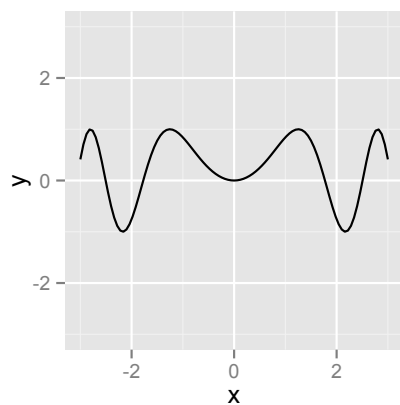
3. Show in the graph where the point where the derivative is 0. At point (2,4) is the derivative smaller, larger or equal to 0? Same question for point  $(-2, 4)$ . Bonus question: Find the equation from the graph.



4. Show in the graph where the derivative is 0. Give the derivative of the point  $(0,0)$ . Bonus question: Find the equation from the graph.



5. Give the derivatives at  $(-5, -1)$  and  $(3, 1)$ .



6. Show in the graph the points where the derivative is 0. How many such points are there? Also show in the graph the parts where the derivative is smaller than 0.

\* **Exercise 1.2** Find the derivatives of the following functions

$$a(x) = -12$$

$$b(x) = x^2$$

$$c(x) = 5x^3 + 2$$

$$d(x) = \frac{5}{2}x^2 - 3x$$

$$e(x) = (x - 3)^2$$

$$f(x) = (-2x - 1)^6$$

$$g(x) = \frac{x^2 - 1}{6x^4}$$

$$h(x) = \frac{(x^2 - 2)^2}{x^2}$$

\*\*\* **Exercise 1.3** (Optional) Find the derivatives of the following functions

$$a(x) = \sin(x^2 + x)$$

$$b(x) = \cos(x + 3)$$

*Hint:*

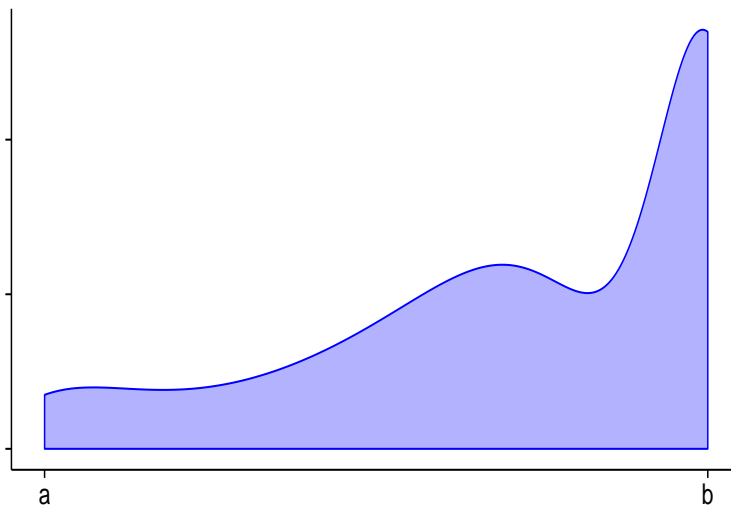
*if  $f(x) = \sin(x)$ , then  $f'(x) = \cos(x)$*

*if  $f(x) = \cos(x)$ , then  $f'(x) = -\sin(x)$ .*

## 1.2 Integrals

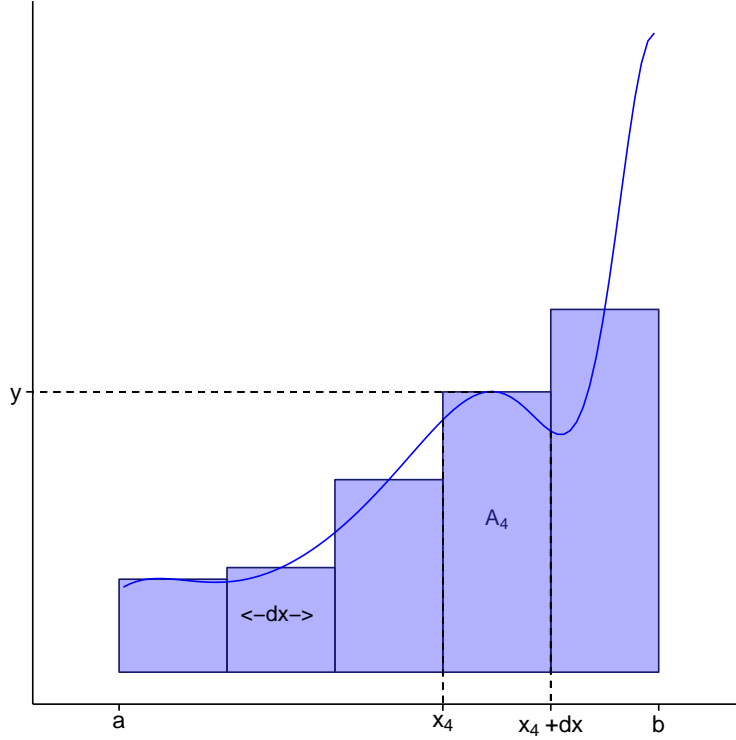
### 1.2.1 Graphical definition

Graphically, an integral is **the area below a curve**. The shaded area in the figure below is the integral of the blue curve between  $x = a$  and  $x = b$ .



An approximation of the shaded area can be obtained by dividing the area in small rectangles. This is interesting because calculating the area of a rectangle is simple: it is the width times the height. This is what is illustrated in the figure below. The  $x$ -axis has been divided in multiple segments of identical

width  $dx$ .



Let us look more closely at the fourth rectangle ( $A_4$ ). This rectangle lies between  $x_4$  and  $x_4 + dx$ . The height of the rectangle is equal to the height of the curve at the mid-point between  $x_4$  and  $x_4 + dx$ . We will call this point  $x_4^*$  (and thus  $x_4^* = x_4 + \frac{1}{2}dx$ ). Let  $f(x)$  be the function of the curve. The area of the rectangle,  $A_4$ , is  $f(x_4^*)$ , the height times  $dx$ , the width.

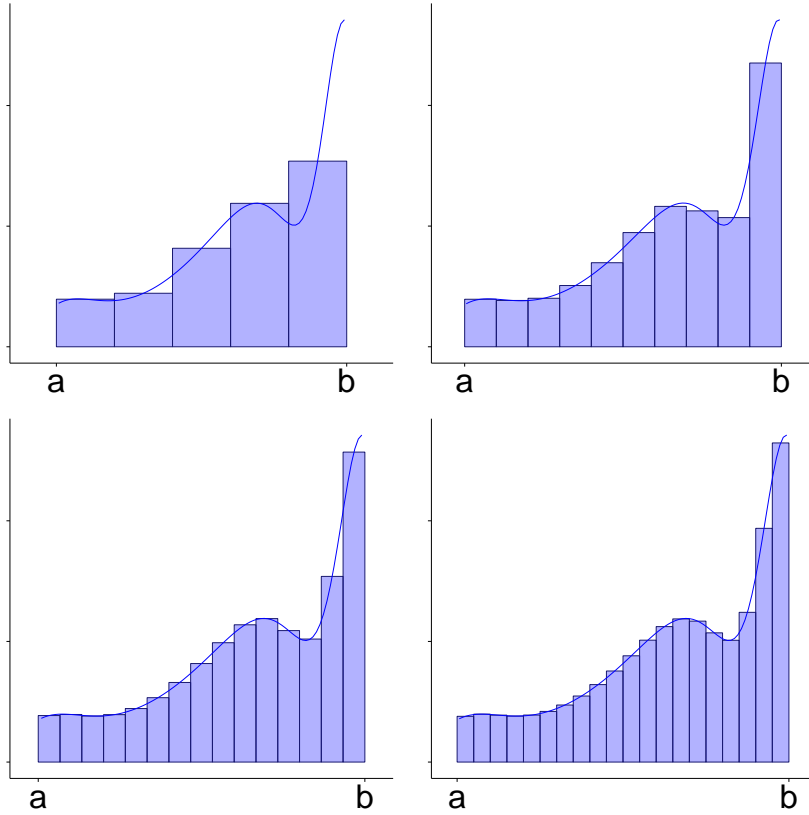
$$A_4 = f(x_4^*)dx \quad (1.31)$$

We can repeat this for all rectangles in the figure. The total area under the curve ( $A$ ) can then be approximated by the sum of all rectangles.

$$A \approx A_1 + A_2 + A_3 + A_4 + A_5 \quad (1.32)$$

Now we have divided the  $x$ -axis between  $a$  and  $b$  in five bins. If instead we divide it in  $n$  bins, we get the following expression:

$$\begin{aligned} A &\approx A_1 + A_2 + \dots A_n \\ &= f(x_1^*)dx + f(x_2^*)dx + \dots + f(x_n^*)dx \\ &= \sum_{i=1}^n f(x_i^*)dx \end{aligned} \quad (1.33)$$



The smaller the interval is, the closer the approximated area will be to the actual area. To state this differently: the larger the number of intervals ( $n$ ), the better the approximation.

There is of course a relationship between the number of intervals ( $n$ ) and the size of each interval ( $dx$ ). If we integrate from  $x = a$  to  $x = b$ , the length ( $L$ ) of this range is:

$$L = b - a$$

If we divide this length in  $n$  bins with equal size  $dx$ , this size will be:

$$dx = \frac{L}{n}$$

### 1.2.2 Formal definition

We want the number of intervals  $n$  to approach infinity. In other words we are interested in the case where the limit of  $n$  tends towards infinity: this is when the approximation is the true area under the curve.

$$A = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i^*) dx \quad (1.34)$$



Formally an integral is defined as:

$$\int f(x)dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i^*)dx \quad (1.35)$$

An integral is in a sense **the reverse of a derivative**. Given a function  $f(x)$ , the primitive of this function is any function  $F(x)$  such that

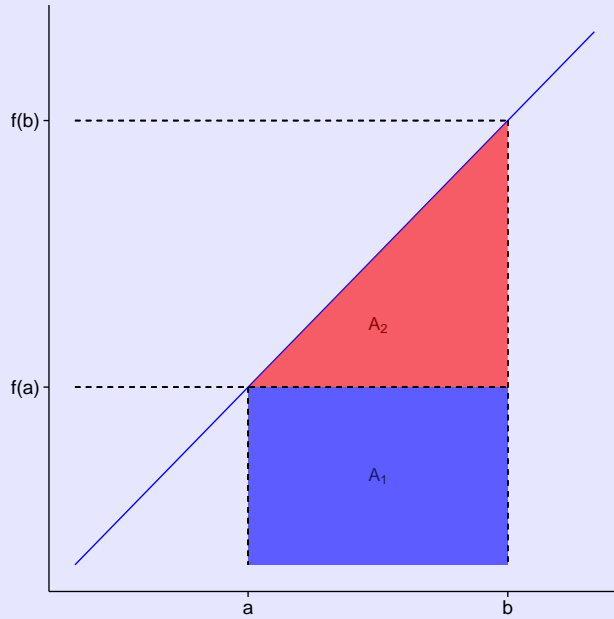
$$F'(x) = f(x) \quad (1.36)$$

The integral of  $f(x)$  over the interval  $[a, b]$ :

$$\int_{x=a}^b f(x)dx = F(b) - F(a) \quad (1.37)$$

### Example

We will now illustrate this by calculating the area of the curve  $f(x) = x$  between  $x = a$  and  $x = b$ .



Because  $f(x) = x$ ,  $f(a) = a$  and  $f(b) = b$ .

*Graphically*

We can divide the area that we are interested in into two smaller areas: a rectangle ( $A_1$ ) and a triangle ( $A_2$ ). For both of these it is easy to calculate the area.

First we calculate the area of the rectangle ( $A_1$ ), for this we multiply its width ( $b - a$ ) by its height ( $f(a) = a$ ):

$$A_1 = (b - a)a = ba - a^2$$

Next we calculate the area of the triangle ( $A_2$ ). This is equal to one half

times its width  $(b - a)$  times its height  $(f(b) - f(a) = b - a)$ :

$$A_2 = \frac{1}{2}(b - a)(b - a) = \frac{1}{2}(b^2 + a^2 - 2ba) = \frac{1}{2}b^2 + \frac{1}{2}a^2 - ba$$

We find the the total area under the curve by adding up the areas of the rectangle and the triangle:

$$A = A_1 + A_2 = \frac{1}{2}b^2 - \frac{1}{2}a^2.$$

### *Numerically*

We are looking for a function  $F(x)$ , such that  $F'(x) = f(x)$ . In words, we are looking for a function whose derivative is  $x$ , because a primitive is the reverse function of a derivative. Remember that the derivative of  $x^2$  is  $2x$ , so  $x^2$  is close to the function we are looking for. In fact, it is 2 times too high. So why not just divide it by 2?

$$F(x) = \frac{1}{2}x^2.$$

But is  $x$  really the derivative of  $\frac{1}{2}x^2$ ? Let us check by applying the general rule of derivatives:

$$\begin{aligned} F'(x) &= \frac{1}{2} \cdot 2x^{(2-1)} \\ &= x \end{aligned}$$

So yes,  $x$  is the derivative of  $\frac{1}{2}x^2$ ! Now we just need to fill in the boundaries  $[a, b]$  to obtain the area under the curve:

$$A = F(b) - F(a) = \frac{1}{2}b^2 - \frac{1}{2}a^2$$

### **Constants**

The derivative of a constant is zero (that is: a flat line has no slope). Therefore if we find a primitive  $F(x)$  such that:

$$F'(x) = f(x).$$

We know that if we add a constant ( $C$ ) to  $F(x)$ , this will also be a primitive of  $f(x)$ , because:

$$(F(x) + C)' = F'(x) + C' = F'(x) + 0 = F'(x) = f(x).$$

When we calculate the area under a curve, this extra constant has no influence. Say for example we define a new primitive  $G(x) = F(x) + C$ . If we calculate the area under the curve from  $x = a$  to  $x = b$ , we find that:

$$G(b) - G(a) = F(b) + C - (F(a) + C) = F(b) - F(a).$$

This the same result that we would have obtained if we had not added a constant to the primitive.

### 1.2.3 Rules of integration

Below we wrote down the primitives of the most common function and three lines that explain what happens when functions are combined (here  $a$  and  $C$  are constants):

$$\begin{array}{lll} f(x) = x^n & \rightarrow & F(x) = \frac{1}{n+1}x^{n+1} + C \\ f(x) = a \cdot g(x) & \rightarrow & F(x) = a \cdot G(x) \\ f(x) = g(x) + h(x) & \rightarrow & F(x) = G(x) + H(x) \\ f(x) = g(x) - h(x) & \rightarrow & F(x) = G(x) - H(x) \end{array}$$

#### Example

Let us work out a short example. what is the integral of  $f(x) = 2x^2 + x$  over the interval  $[a, b]$ ? From the rules we find that:

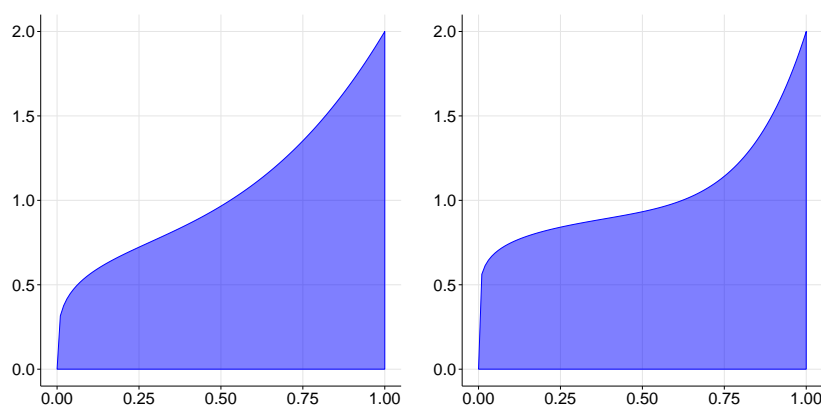
$$F(x) = \frac{2}{3}x^3 + \frac{1}{2}x^2 + C$$

and thus:

$$\int_{x=a}^b f(x)dx = F(b) - F(a) = \frac{2}{3}(b^3 - a^3) + \frac{1}{2}(b^2 - a^2)$$

### 1.2.4 Exercises

#### \* Exercise 1.4 Painting a wall



Your family just moved to a new apartment in Zurich and now it is time to paint the walls to make your house more of a home. Your creative uncle gave you two interesting designs and since you have two  $1 \times 2m$  walls in your house, why not

paint one wall with one design and the other with the other? One detail: how can you make sure that your sister has a larger surface to paint than you? Just to help you a little bit: the equation of the left design is:

$$f_{left}(x) = x^3 + \sqrt[4]{x}.$$

For the design on the right, the equation is:

$$f_{right}(x) = x^6 + \sqrt[8]{x}.$$

**\*\* Exercise 1.5 Find the primitives of the following functions**

$$a(x) = 2$$

$$e(x) = x^6 - 2x^3 + \frac{2}{3}$$

$$b(x) = x^3$$

$$f(x) = \sqrt{x}$$

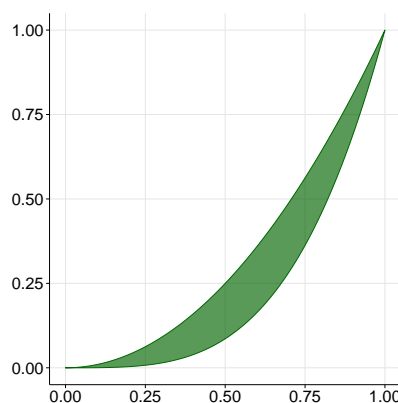
$$c(x) = 3x^2 - 6$$

$$g(x) = \sqrt{x} + \sqrt[3]{x}$$

$$d(x) = -\frac{3}{4}x^2 + x - 6$$

$$h(x) = \frac{5}{6}\sqrt[5]{x} + 2$$

**\*\*\* Exercise 1.6 A different shape**



Can you calculate the surface of the shape above? To help you, we will reveal that the curve that describes the upper boundary of the shape is given by:

$$f_{top}(x) = x^2$$

and the curve describing the lower boundary of the shape is given by:

$$f_{bottom}(x) = x^4 - \frac{1}{4}x^5 + \frac{1}{4}x^3.$$

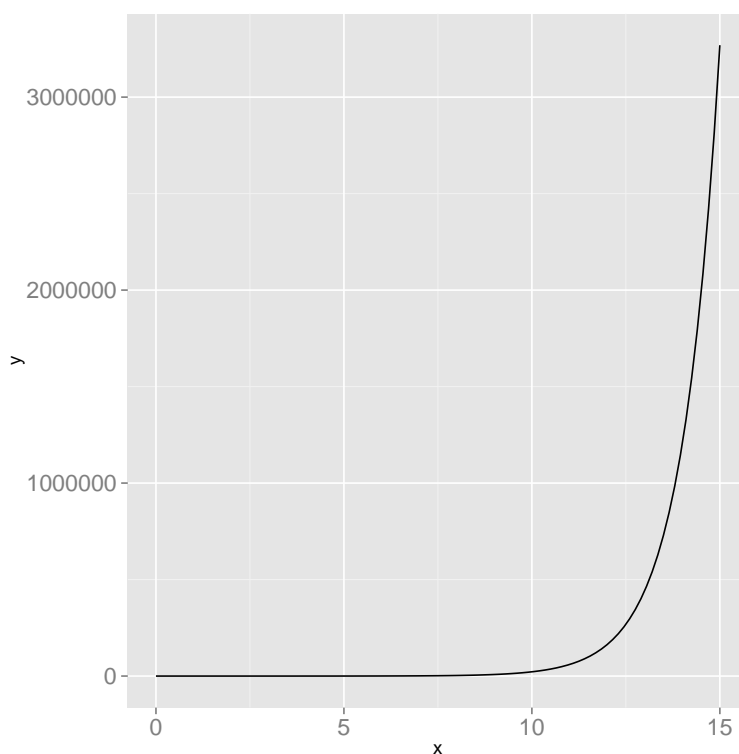
## 1.3 Differential Equations

### 1.3.1 The Exponential Function

An exponential functions look similar to a power function except that the variable of interest,  $x$ , is the exponent rather than the base. Exponential functions are functions of the form  $f(x) = a^x$ , here  $a$  is a positive constant. Graphically, an exponential function looks like the graph below. At low values of  $x$ , the curve is very close to one (because  $a^0 = 1$ ). When  $x$  increases by 1, the value of  $f(x)$  is multiplied by  $a$ , or mathematically put:

$$f(x + 1) = a \cdot f(x). \quad (1.38)$$

Therefore exponential functions are used to model system where  $a$  the factor of increase is constant.



**The function**  $f(x) = A \cdot e^x$

Most of the time, the exponential function refers to the natural exponential function, an exponential function in base  $e$  ( $e \approx 2.72$ ) multiplied by a constant ( $A$ );  $A \cdot e^x$ . When  $x$  tends towards  $\infty$ ,  $A \cdot e^x$  also tends towards  $\infty$ . This function is of specific interest to describe population growth and you will see it many times in the rest of this course. An important property of the exponential function is that its derivative is the same as the original function:

$$f(x) = A \cdot e^x \quad \rightarrow \quad f'(x) = A \cdot e^x \quad (1.39)$$

It means that **the slope of the graph at any point is the  $y$  value at that point**. It follows from this that the integral is also identical to the original function.

$$f(x) = A \cdot e^x \quad \rightarrow \quad F(x) = A \cdot e^x + c \quad (1.40)$$

Of course for this exponential function, the chain rule is still valid:

$$f(x) = A \cdot e^{u(x)} \quad \rightarrow \quad f'(x) = u'(x)f'(u) = u'(x)A \cdot e^{u(x)} \quad (1.41)$$

If  $u(x) = kx$ , with  $k$  a constant, this means that:

$$f(x) = A \cdot e^{kx} \quad \rightarrow \quad f'(x) = (kx)'f'(kx) = kA \cdot e^{kx} \quad (1.42)$$

This is of interest because if the increase of a variable (e.g. population size over time  $N(t)$ ) is proportional to the size of the variable, this means that:

$$N'(t) = k \cdot N(t). \quad (1.43)$$

This equation is called a differential equation because it contains both a function and the derivative of that function. For this differential equation we know that it describes exactly the property of the exponential function, we thus know that  $N(t)$  is an exponential function:

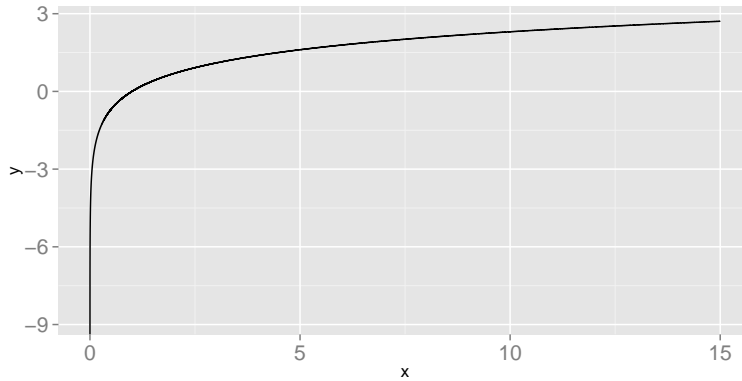
$$N(t) = A \cdot e^{k \cdot t}. \quad (1.44)$$

### 1.3.2 The Logarithmic Function

Just like division is the inverse operation of multiplication, the logarithm is the inverse operation of an exponential. If  $a^x = b$ , then  $\log_a(b) = x$ . The logarithm  $\log_a(b)$  gives the power at which we need to elevate the base ( $a$ ) to get  $b$ . Based on this definition one can find that:

$$a^{\log_a(b)} = \log_a(a^b) = b.$$

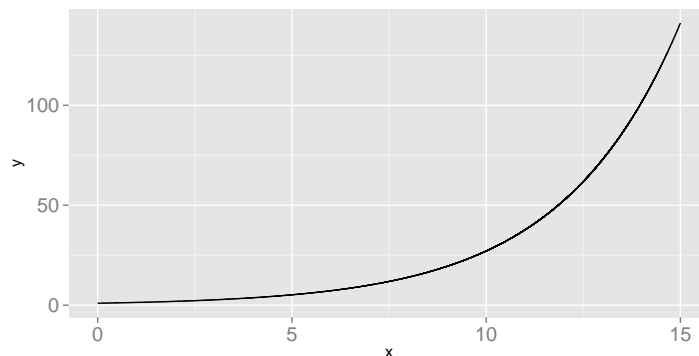
**The natural logarithm** The logarithm of special interest for us is the natural logarithm  $\ln(y)$  which is the inverse operation of the natural exponential function  $e^x$ . If  $\ln(y) = x$ , then  $e^x = y$ . The notation  $\ln$  is specific to the natural logarithm and is always in base  $e$ . However, the literature can be confusing as you may see  $\log(y)$  to mean  $\ln(y)$  or  $\log(y)$  to mean  $\log_{10}(y)$ . However, most of the time you can figure it out from the context.



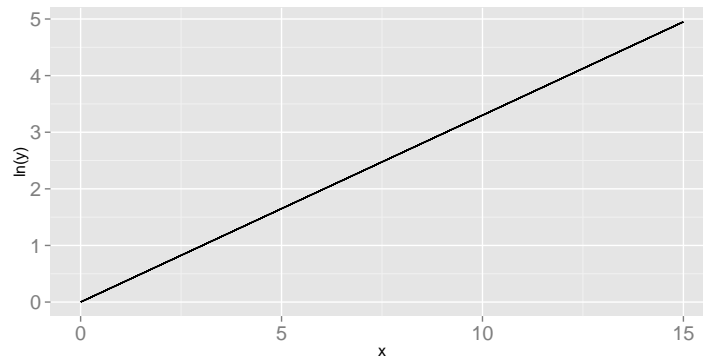
The figure above shows the curve of a natural logarithmic function. The closer  $x$  is to zero, the faster the value of  $\ln(x)$  gets smaller. Conversely, when  $x$  is far from zero, the value of  $\ln(x)$  keeps increasing but the larger  $x$ , the slower this increase. The limit of  $\ln(x)$  as  $x$  approaches infinity is equal to infinity.

#### Use of the natural logarithm

Let us take the function  $f(x) = e^{0.33 \cdot x}$ , which is a typical growth function in ecology. This function is plotted in the graph below.



This exponentially growing curve is nice but not very convenient to work with. A straight line would be much easier. The same graph can be made into a straight line with the use of the natural logarithmic function which is the inverse of the exponential. This is what you see in the graph below



So what is the function of this graph? It is  $\ln(f(x))$ , which is

$$\begin{aligned} \ln(f(x)) &= \ln(e^{0.33 \cdot x}) \\ &= 0.33 \cdot x \end{aligned} \tag{1.45}$$

Visually and computationally, it is much easier to find the function of a straight line and then find the initial function. This is why the natural logarithm is frequently used in combination with the exponential.

### 1.3.3 Exercises

**\* Exercise 1.7**      **Answer the following questions**

1. Around the exponential

a) Solve for  $x$  the following equations

$$2^x = 4$$

$$2 \cdot 3^x = 18$$

$$e^x = 1$$

$$e^x = 4$$

b) Find the derivatives of the following functions

$$a(x) = 5 \cdot e^x$$

$$b(x) = e^x + x^2$$

c) Draw the function  $e^x$ . Describe what happens when  $x$  is very small ( $x = -\infty$ ) and when  $x$  is very large ( $x = +\infty$ ).

2. Around the logarithm

a) Solve for  $x$  the following equations

$$\log_2 x = 4$$

$$\log_5 x = 3$$

$$\ln(x) = 1$$

$$\ln(x) = 3$$

b) Draw the function  $\ln(x)$ . Describe what happens when  $x < 0$  and  $x$  is very large ( $x = +\infty$ ).

3. Exponential and logarithm combined

a) Draw the function  $\ln(e^{0.4x})$ .

b) What is special about that function?



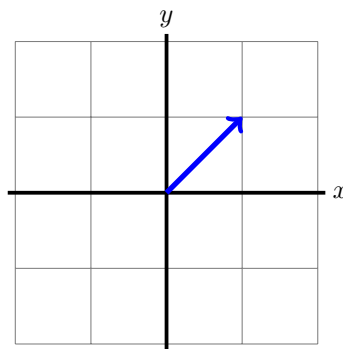
## Chapter 2

# Introduction to Matrix Algebra

### 2.1 Vectors

#### 2.1.1 Graphically

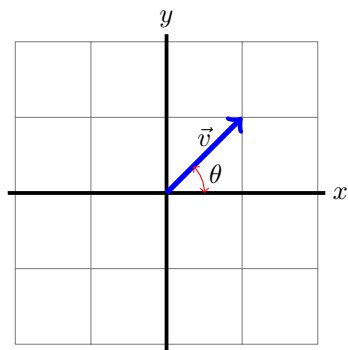
Vectors are probably the easiest to understand graphically. We drew an example of a vector below. Informally you could say that a vector is an arrow, for this reason usually when something is a vector, a small arrow is put on top of it to indicate that it is a vector:  $\vec{v}$ . It is different from a 'normal' number (a scalar) in the respect that a vector can not be described with one single number. We could give you the length of a vector, but you would still not know in which direction it points. On the other hand, we could specify what a vector points at, but you would still not know how long it is. **A vector is described by two components: a length and a direction.** The most common example of a vector is velocity: when you move around in a car there are two important components: your speed (the length of the vector) and your direction.



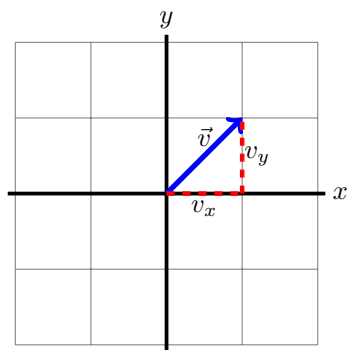
#### 2.1.2 With numbers

Pictures are easy to understand, but not always easy to use in communication. When a researcher here in Zürich wants to share one of his results with a collab-

orator on the other side of the ocean, it is a bit impractical to send a collection of drawings. Instead it would be nice to just send a few numbers per e-mail. In this case the researcher would have to send both bits of information: the direction and the length of the vector ( $\vec{v}$ ). Sending the length (we write this as  $|\vec{v}|$ ) is straightforward, the direction can be specified as the angle ( $\theta$ ) between the vector and the  $x$ -axis:



However, measuring angles can be annoying, so therefore a different notation is more frequently used. In this notation the part of the vector along the  $x$ -axis ( $v_x$ ) and the part along the  $y$ -axis ( $v_y$ ) are recorded:



This is usually written down as follows:

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix} \quad (2.1)$$

The two notations can be converted into one another. Using the Pythagorean theorem ( $a^2 + b^2 = c^2$ ), one can see that for the length of the vector:

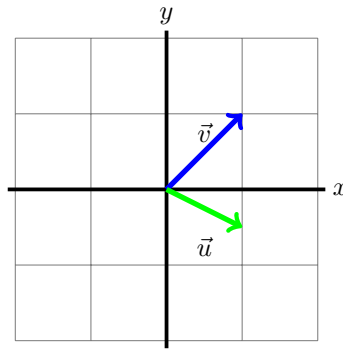
$$|\vec{v}| = \sqrt{v_x^2 + v_y^2}$$

Those who still remember the basics of trigonometry, will (hopefully) agree that for the angle  $\theta$  the following equation holds:

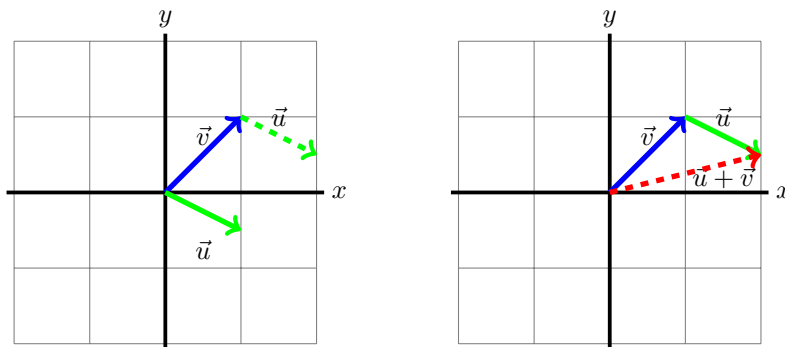
$$\tan(\theta) = \frac{v_y}{v_x}$$

### 2.1.3 Graphical addition

So far we have focused on one single vector. We will now slowly increase the excitement by doubling the number of vectors that we work with. Below two vectors are drawn:  $\vec{u}$  and  $\vec{v}$ :

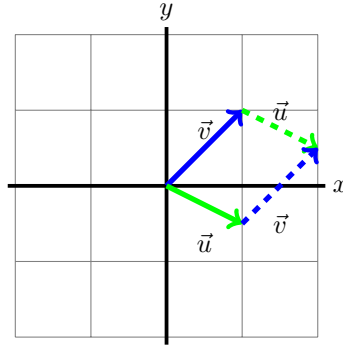


Like we can add up numbers ( $4 + 2 = 6$ ), we can also add up vectors. To do so in a drawing, we move the second vector such that it begins where the first one ends. The sum of the two vectors is the vector that starts at the origin  $((0,0))$  and ends where the second vector is pointing. Words are sometimes fuzzy so let's instead look at a picture.



In the left picture, first the second vector is moved such that it begins at the end of the first vector. Next, the sum of the two vectors is drawn from the origin to where the second vector is now pointing.

Please note that it is not important which vector is the first and which vector is the second vector. If we would have called  $\vec{v}$  the second vector and  $\vec{u}$  the first, the result would have been exactly the same:



Hopefully from this picture it has become clear that in general:

$$\vec{v} + \vec{u} = \vec{u} + \vec{v}$$

#### 2.1.4 Addition

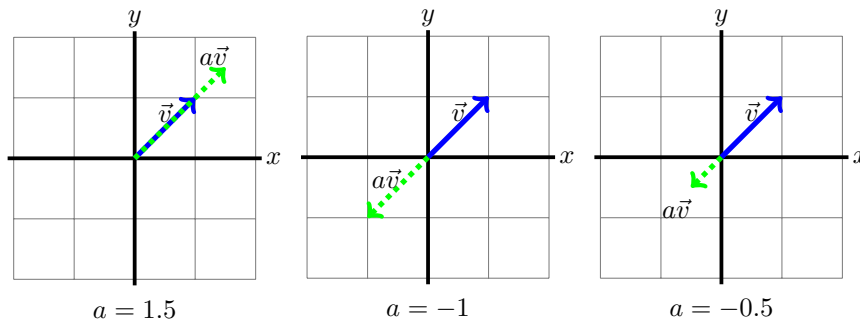
Fortunately addition of two vectors becomes very easy when we use the notation from equation 2.1:

$$\vec{u} + \vec{v} = \begin{pmatrix} u_x \\ u_y \end{pmatrix} + \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} u_x + v_x \\ u_y + v_y \end{pmatrix} \quad (2.2)$$

That is: we can simply add up the separate x contributions to get the total x contribution of the sum of the vectors and the y contributions to get the y contribution of the sum of the vectors.

#### 2.1.5 Multiplication of a vector by a scalar

Sometimes you want to express that the length of a vector changes while the direction of that vector does not change. An example is speeding up in a car: your speed increases but as long as you do not use the steering wheel your direction will (hopefully) remain the same. In the following images we illustrate the multiplication of the vector  $\vec{v}$  by a scalar  $a$ . If this scalar is negative, the multiplication will make the vector point in the opposite direction.



When facing numbers instead of figures, the multiplication of a vector with a scalar is performed by simply multiplying the separate elements of the vector by that scalar. That is:

$$a\vec{v} = a \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} a \cdot v_x \\ a \cdot v_y \end{pmatrix} \quad (2.3)$$

We will now show that equation 2.3 indeed leads to a multiplication of the length of a vector, while its direction remains the same.

First, we know that for a vector  $\vec{v}$  its length and direction are given as follows:

$$|\vec{v}| = \sqrt{v_x^2 + v_y^2}$$

$$\tan(\theta_v) = \frac{v_y}{v_x}$$

Let us now look at a new vector  $\vec{u}$ , where  $\vec{u} = a\vec{v}$ . That is:  $\vec{u}$  is the vector  $\vec{v}$  multiplied by a scalar. Now we know that:

$$\vec{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix} = a\vec{v} = \begin{pmatrix} a \cdot v_x \\ a \cdot v_y \end{pmatrix}$$

And thus:

$$|\vec{u}| = \sqrt{u_x^2 + u_y^2} = \sqrt{(av_x)^2 + (av_y)^2} = \sqrt{a^2(v_x^2 + v_y^2)} = a\sqrt{v_x^2 + v_y^2} = a|\vec{v}|$$

$$\tan(\theta_u) = \frac{u_y}{u_x} = \frac{av_y}{av_x} = \frac{v_y}{v_x} = \tan(\theta_v)$$

We see that indeed both  $\vec{v}$  and  $\vec{u}$  have the same direction ( $\theta$ ) but the length of  $\vec{u}$  is  $a$  times the length of  $\vec{v}$ .

To subtract one vector from another is the same as adding up one vector to the negative version of the other, that is:

$$\vec{u} - \vec{v} = \vec{u} + (-1 \cdot \vec{v})$$

### 2.1.6 Larger dimensions

So far we have been concerned only with two dimensional situation, not because these are the most relevant, but because these are the easiest to explain on paper. Fortunately it is relatively easy to extend everything we have discussed to 3 or even more dimensions. For every extra dimension that we add, we need to include one more number specifying the vector. That is, in three dimensions:

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \quad (2.4)$$

Adding up vectors and multiplying them with scalars is just the same as in two dimensions:

$$\vec{v} + \vec{u} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} + \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = \begin{pmatrix} v_x + u_x \\ v_y + u_y \\ v_z + u_z \end{pmatrix} \quad (2.5)$$

$$a\vec{v} = a \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} av_x \\ av_y \\ av_z \end{pmatrix} \quad (2.6)$$

We could go to even more dimensions. Instead of the labels  $x, y, z \dots$  we now number the different dimensions. If we have a total of  $N$  dimensions, the following relations hold:

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} \quad (2.7)$$

$$\vec{v} + \vec{u} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} v_1 + u_1 \\ v_2 + u_2 \\ \vdots \\ v_N + u_N \end{pmatrix} \quad (2.8)$$

$$a\vec{v} = a \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} = \begin{pmatrix} av_1 \\ av_2 \\ \vdots \\ av_N \end{pmatrix} \quad (2.9)$$

### 2.1.7 Basis Vectors

A special group of vectors are the basis vectors, these are vectors with length 1 that point along one of the axis. We denote the  $i^{th}$  basis vector as  $\vec{e}_i$ . In this notation we can use either numbers (1,2,3,...,N) or letters (x,y,z) for i. As an example we will show 3 basis vectors:

$$\vec{e}_x = \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_y = \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_N = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

This is just a definition that we will need in the next section. Note that there are as many basis vectors as there are dimensions: in 2 dimensional space there will only be two basis vectors ( $\vec{e}_x$  and  $\vec{e}_y$ ).

### 2.1.8 Exercises

**\* Exercise 2.1      Answer the following questions**

1. Vector manipulation

- a) Draw the vectors  $\vec{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$  and  $\vec{b} = \begin{pmatrix} -2 \\ -3 \end{pmatrix}$
- b) What is the angle between  $\vec{a}$  and  $\vec{b}$ ?
- c) Draw the vector  $\vec{a} + \vec{b}$  and then give the coordinates of that new vector. What is special about it?

2. Vector addition

- a) Draw the vectors  $\vec{c} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  and  $\vec{d} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$
- b) Draw the vector  $\vec{i} = \vec{c} - \vec{d}$  and then the vector  $\vec{f} = \vec{d} - \vec{c}$ . Give the coordinates for both vectors. Any comments?
- c) Find the magnitude of  $\vec{i}$ ,  $(|\vec{i}|)$ .

3. Vector multiplication

- a) Draw  $\vec{g} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$
- b) Draw  $2\vec{g}$  and give the coordinates of that new vector.

4. Multiple dimensional vectors

- a) What is the dimension of the following vector  $h = \begin{pmatrix} 2 \\ 3 \\ -4 \\ -6 \end{pmatrix}$ ?
- b) What are the coordinates of  $5\vec{h}$ ?

5. Basis vectors

- a) In a two dimensional space, give the coordinates of the basis vectors in the first dimension ( $\vec{e}_x$ ) and in the second dimension ( $\vec{e}_y$ )
- b) What is special about these vectors?

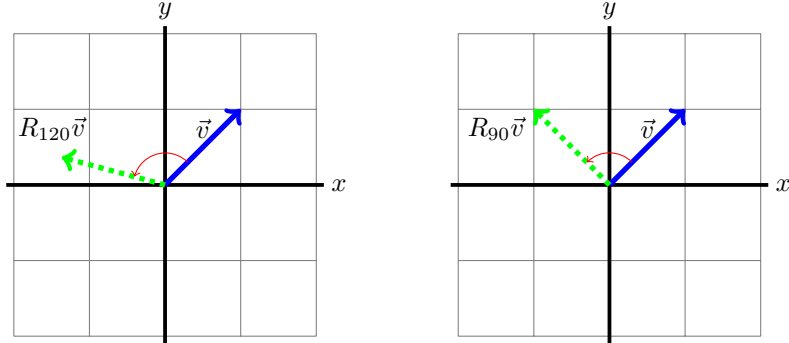
## 2.2 Vector Transformations

### 2.2.1 Graphically

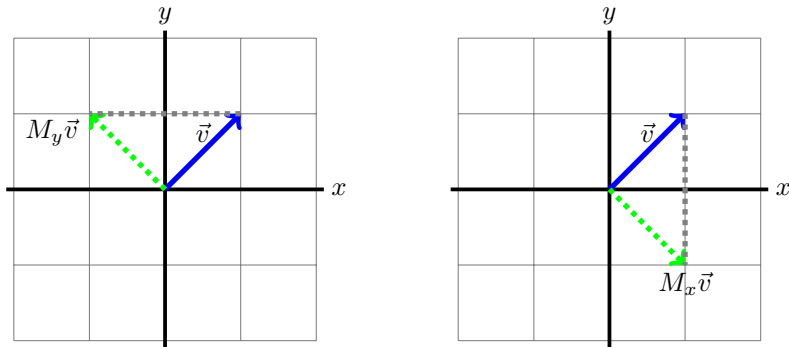
So far we have been concerned with vectors: what they are and what happens when they are combined. We will now deal with linear transformations of vectors. That is instead of vectors themselves we look at operations that can change vectors. In general we will denote such transformations with a capital letter. When we have a transformation called  $A$ , then  $A\vec{v}$  is the  $\vec{v}$  transformed by  $A$ . To give you an idea we will now discuss 5 examples of transformations of a 2 dimensional vector.

1. **Nothing** The simplest thing that can happen is nothing. The vector  $\vec{v}$  just remains  $\vec{v}$ . This is what we call **the identity transformation**.
2. **Rotation** A very common transformation is a rotation. Let us write  $R_\theta$  for a rotation of  $\theta$  degrees counter clockwise. Below two examples

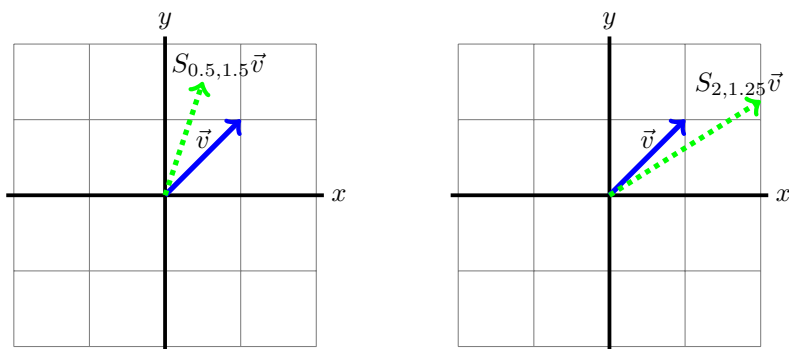
are shown: one for a rotation of  $120^\circ$  and one for a rotation of  $90^\circ$ .



3. **Mirror image** The next transformation is mirroring a vector with respect to a line. Below we illustrate this for (left) mirroring the vector with respect to the  $y$ -axis and (right) for mirroring the vector with respect to the  $x$ -axis. Mirroring can also be done with respect to other lines.



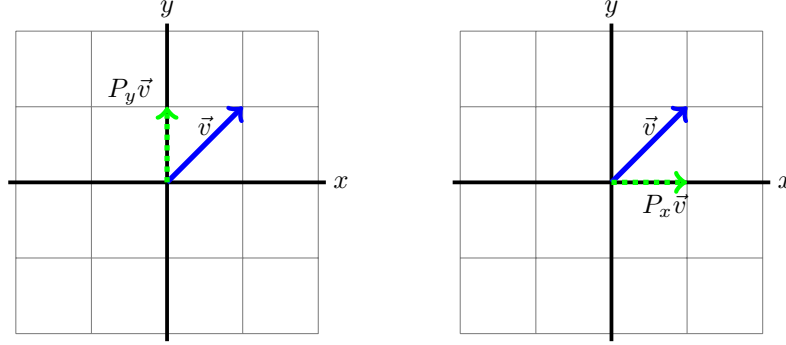
4. **Scaling** In the previous section we discussed multiplying a vector by a scalar. In this case both components of the vector would scale exactly the same. Instead it is also possible to scale the components separately, we could for example multiply the  $x$ -component by  $1.5$  and the  $y$ -component by  $0.5$ . Let us call that transformation  $S_{1.5,0.5}$ . Below are two graphical examples of such a transformation.



5. **Projection** The last transformation we will treat here is the projection. This transformation simply isolates a part of the vector, for example the part of the vector that is along the  $x$ -axis ( $P_x$ ) or the part that lies along



the y-axis ( $P_y$ ). Of course projections can also be made on other lines, but in the following figures for simplicity we will show projections on the x-axis and on the y-axis.



### 2.2.2 With numbers

The type of transformations that we are discussing are the so-called linear transformations. The nice property of these transformations is that when you add up two vectors and then perform the transformation on the sum of these vectors, it gives exactly the same result as first performing the transformation on the separate vectors and then adding up these transformed vectors. In an equation this is:

$$A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v} \quad (2.14)$$

Say that we want to know what happens to  $\vec{v}$  when it is transformed by  $A$ . What information do we need to make sure we are always able to find the result of the transformation? To answer this question we first write our vector in terms of the basis vectors:

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} v_x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ v_y \end{pmatrix} = v_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + v_y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.15)$$

We now use this way of writing  $\vec{v}$  together with equation 2.14 and we find:

$$A\vec{v} = Av_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + Av_y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.16)$$

In this equation  $v_x$  and  $v_y$  are not vectors but scalars. In general it is the case that the order in which a matrix  $A$  and a scalar  $a$  are multiplied does not matter:

$$Aa = aA \quad (2.17)$$

Using this we can write equation 2.16 as follows:

$$A\vec{v} = v_x A \begin{pmatrix} 1 \\ 0 \end{pmatrix} + v_y A \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.18)$$

What we have found is an important result: **in order to find out how any vector in 2 dimensional space is transformed by a linear transformation  $A$ , we only have to know how  $A$  transforms the basis vectors  $\vec{e}_x$  and  $\vec{e}_y$ .**

### 2.2.3 Matrices

Usually the information in equation 2.18 is written as a matrix, the columns of this matrix describe how each of the basis vectors look like after the transformation. That is, if we know that:

$$A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,1} \\ a_{1,2} \end{pmatrix} \quad (2.19)$$

$$A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{2,1} \\ a_{2,2} \end{pmatrix} \quad (2.20)$$

We summarise this information by putting all the relevant information together and we write:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \quad (2.21)$$

In this notation  $a_{i,j}$  is simply a scalar that tells what the value is of the  $i^{th}$  coordinate of the  $j^{th}$  basisvector after the transformation. For example  $a_{2,1}$  is the value of the second coordinate (the y-coordinate) of the transformed first basisvector.

In the matrix notation the element  $a_{i,j}$  always refers to the element in row  $i$  and column  $j$ . An easy way to remember that the first index refers to the row and the second to the column is by remembering that in Roman Catholic R also comes before C.

### 2.2.4 Examples

We will now discuss the matrices for the 2 dimensional examples in section 2.2.1:

1. **Nothing** If nothing happens  $\vec{e}_x$  remains  $\vec{e}_x$  and  $\vec{e}_y$  remains  $\vec{e}_y$ . This is what we call the identity matrix ( $\mathbb{1}$ ).

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

2. **Rotation** We will describe only a rotation of  $90^\circ$ . The more general matrix for rotations is left as an exercise. If the vector  $\vec{e}_x$  is rotated  $90^\circ$  counterclockwise it will be the vector  $\vec{e}_y$ . If the vector  $\vec{e}_y$  is rotated  $90^\circ$  counterclockwise it will be the vector  $-\vec{e}_x$ , or as a matrix:

$$R_{90} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

3. **Mirror image** Again we consider only one example of a mirror image, two other mirror images are left as an exercise. Here we will consider mirroring a vector in the x-axis. This operation will not change  $\vec{e}_x$ , but it will change  $\vec{e}_y$  into  $-\vec{e}_y$ , thus:

$$M_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

4. **Scaling** If we scale the x component of a vector with a factor 1.5 and the y component with a factor 0.75 the matrix describing this operation will look as follows:

$$S_{1.5,0.75} = \begin{pmatrix} 1.5 & 0 \\ 0 & 0.75 \end{pmatrix}$$

5. **Projection** Finally we consider a projection on the x-axis. This operation will not change  $\vec{e}_x$ . It will however completely remove the y-component of the vector, that is  $\vec{e}_y$  is transformed into 0.

$$P_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

### 2.2.5 Matrix vector Multiplication

In more dimensions everything is slightly more complicated (see next section), but for 2 dimensions it is not that bad:

$$A\vec{v} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} a_{1,1}v_x + a_{1,2}v_y \\ a_{2,1}v_x + a_{2,2}v_y \end{pmatrix} \quad (2.22)$$

This equation describes how  $\vec{v}$  will look like after it has been transformed by the matrix  $A$ .

### 2.2.6 Larger dimensions

So far we have just looked at transformations of 2 dimensional vectors. Of course 3 dimensional vectors can also be transformed by a matrix. In this case the matrix will have 3 rows and 3 columns. The first column describes what the first basis vector looks like after the transformation, the second column how the second will look like and the third what the third basis vector looks like. In general in 3 dimensional space:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \quad (2.23)$$

$$A\vec{e}_i = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ a_{3,i} \end{pmatrix} \quad (2.24)$$

In this last equation  $i$  can be 1, 2 or 3. Next we generalise matrices to even larger dimensions. If the dimension of the system is  $N$ , the previous equations look as follows:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,N} \\ a_{2,1} & a_{2,2} & \dots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \dots & a_{N,N} \end{pmatrix} \quad (2.25)$$

$$A\vec{e}_i = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{N,i} \end{pmatrix} \quad (2.26)$$

Now  $i$  can be any integer in the range from 1 to  $N$ .

### 2.2.7 Matrix vector multiplication in larger dimensions

In three dimensions a vector is transformed by a matrix as follows:

$$A\vec{v} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} a_{1,1}v_x + a_{1,2}v_y + a_{1,3}v_z \\ a_{2,1}v_x + a_{2,2}v_y + a_{2,3}v_z \\ a_{3,1}v_x + a_{3,2}v_y + a_{3,3}v_z \end{pmatrix} \quad (2.27)$$

This is maybe a bit difficult to remember, so instead let us just look at how the first entry of the transformed vector is calculated:

$$A\vec{v} = \begin{pmatrix} \boxed{a_{1,1} \quad a_{1,2} \quad a_{1,3}} \\ a_{2,1} \quad a_{2,2} \quad a_{2,3} \\ a_{3,1} \quad a_{3,2} \quad a_{3,3} \end{pmatrix} \begin{pmatrix} \boxed{v_x} \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} \boxed{a_{1,1}v_x + a_{1,2}v_y + a_{1,3}v_z} \\ a_{2,1}v_x + a_{2,2}v_y + a_{2,3}v_z \\ a_{3,1}v_x + a_{3,2}v_y + a_{3,3}v_z \end{pmatrix} \quad (2.28)$$

We take the first row of the matrix and multiply this number by the first entry of  $\vec{v}$ . To this we add the product of the second entry of the first row of the matrix and the second entry of  $\vec{v}$ . When we also add the product of the third entry of the first row of the matrix and the third entry of  $\vec{v}$ , this total sum gives us the first entry of the transformed vector. To get the second entry, we repeat the whole process, but instead of the first row of the matrix we now use the second row of the matrix.

$$A\vec{v} = \begin{pmatrix} \boxed{a_{1,1} \quad a_{1,2} \quad a_{1,3}} \\ \boxed{a_{2,1} \quad a_{2,2} \quad a_{2,3}} \\ a_{3,1} \quad a_{3,2} \quad a_{3,3} \end{pmatrix} \begin{pmatrix} \boxed{v_x} \\ \boxed{v_y} \\ v_z \end{pmatrix} = \begin{pmatrix} \boxed{a_{1,1}v_x + a_{1,2}v_y + a_{1,3}v_z} \\ \boxed{a_{2,1}v_x + a_{2,2}v_y + a_{2,3}v_z} \\ a_{3,1}v_x + a_{3,2}v_y + a_{3,3}v_z \end{pmatrix} \quad (2.29)$$

We will now write this procedure down as an equation. If  $\vec{u}$  is the transformed vector ( $\vec{u} = A\vec{v}$ ) then, for the  $i^{th}$  entry of  $\vec{u}$ :

$$u_i = \sum_{j=1}^3 a_{i,j}v_j \quad (2.30)$$

In this notation the  $\Sigma$  mean summation: it adds up all arguments that it has in it for which  $j$  is between 1 and 3 (these limits are given below and above the  $\Sigma$ -sign). For example:

$$\sum_{j=1}^3 a_{1,j} = a_{1,1} + a_{1,2} + a_{1,3} \quad (2.31)$$

In even larger dimensions, multiplying a vector by a matrix works in a similar manner, we just extend the number of values that  $j$  can have. If we have  $N$  dimensions and  $\vec{u} = A\vec{v}$ , then:

$$u_i = \sum_{j=1}^N a_{i,j}v_j \quad (2.32)$$

### 2.2.8 The order of matrix multiplications

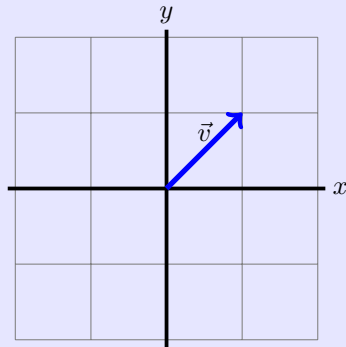
You probably remember that when we multiply two scalars, the order does not matter:

$$3 \cdot 4 = 4 \cdot 3 \quad (2.33)$$

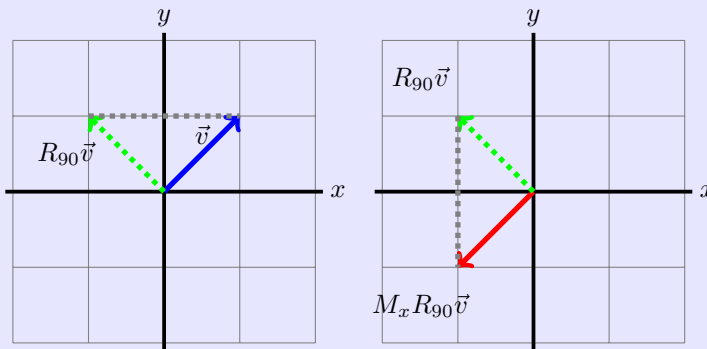
For matrices this is unfortunately not always true: in general the order of the multiplication does matter.

### Graphical example

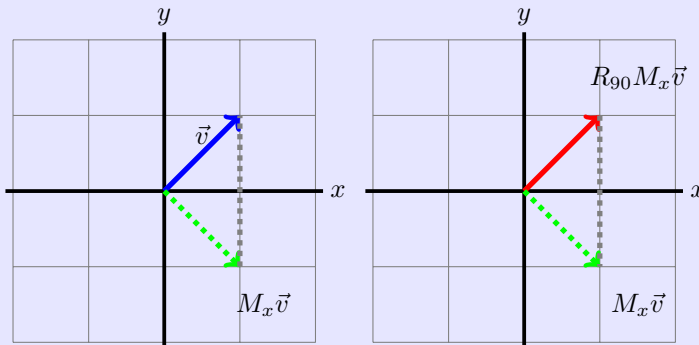
Let us consider the vector  $\vec{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ :



If we now first rotates this vector  $90^\circ$  counterclockwise ( $R_{90}$ ) and then mirror the resulting vector in the  $x$ -axis ( $M_x$ ), we find the following:



If instead we mirror the vector in the  $x$ -axis and then rotate it  $90^\circ$  counterclockwise, we find the following:



The final vectors (in red) point in opposite directions: the order in which we perform the transformation thus has an influence on the final outcome.

### Numerical Example

Let us now consider a three dimensional vector  $\vec{v}$ :

$$\vec{v} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

Next, we define two matrices  $A$  and  $B$ :

$$A = \begin{pmatrix} 0 & 2 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0.5 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 & 2 \\ 0.8 & 0 & 0 \\ 0 & 0.8 & 0.4 \end{pmatrix}$$

First we consider  $AB\vec{v}$ , this is the same as  $A(B\vec{v})$ . We calculate  $B\vec{v}$ :

$$B\vec{v} = \begin{pmatrix} 0 & 1 & 2 \\ 0.8 & 0 & 0 \\ 0 & 0.8 & 0.4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \cdot 1 + 1 \cdot 2 + 2 \cdot 1 \\ 0.8 \cdot 1 + 0 \cdot 2 + 0 \cdot 1 \\ 0 \cdot 1 + 0.8 \cdot 2 + 0.4 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 0.8 \\ 2 \end{pmatrix}$$

Now we apply matrix  $A$ :

$$\begin{aligned} AB\vec{v} &= A(B\vec{v}) = \begin{pmatrix} 0 & 2 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} 4 \\ 0.8 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \cdot 4 + 2 \cdot 0.8 + 3 \cdot 2 \\ 0.5 \cdot 4 + 0 \cdot 0.8 + 0 \cdot 2 \\ 0 \cdot 3 + 0.5 \cdot 0.8 + 0.5 \cdot 2 \end{pmatrix} \\ &= \begin{pmatrix} 7.6 \\ 2 \\ 1.4 \end{pmatrix} \end{aligned}$$

If we reverse the order, we find (this time we left out the middle step, if you have time, please do verify the equations below):

$$A\vec{v} = \begin{pmatrix} 0 & 2 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 0.5 \\ 1.5 \end{pmatrix}$$

And now:

$$BA\vec{v} = B(A\vec{v}) = \begin{pmatrix} 0 & 1 & 2 \\ 0.8 & 0 & 0 \\ 0 & 0.8 & 0.4 \end{pmatrix} \begin{pmatrix} 7 \\ 0.5 \\ 1.5 \end{pmatrix} = \begin{pmatrix} 3.5 \\ 5.6 \\ 1 \end{pmatrix}$$

So in this case we also note that  $BA\vec{v}$  is different from  $AB\vec{v}$ . The order in which we apply the matrices thus makes a difference in the final outcome.

### 2.2.9 Exercises

#### \* Exercise 2.2 A general matrix for scaling

1. Graphical exercise

- Draw  $A\vec{u}$ , where  $A\vec{u} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
- Draw  $B\vec{u}$ , where  $B\vec{u} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

2. Calculate:

$$\begin{aligned} \text{a) } A\vec{u} &= \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} 7 \\ 7 \end{pmatrix} \\ \text{b) } B\vec{v} &= \begin{pmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ \text{c) } C\vec{w} &= \begin{pmatrix} 1 & -4 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} -5 \\ 2 \end{pmatrix} \\ \text{d) } D\vec{z} &= \begin{pmatrix} 3 & 3 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & 7 \end{pmatrix} \begin{pmatrix} -5 \\ 0 \\ -6 \end{pmatrix} \end{aligned}$$

### \* Exercise 2.3 Population ecology and matrices

Consider a population of wandering albatrosses. This year the population contains 20 juveniles, 16 sub-adults and 12 adults. We have been observing the population for a few years now and from these observations we expect the following to happen in the coming year:

- Of the juveniles 40% will become sub-adults, 40% will remain juveniles and 20% die.
  - Of the sub-adults 50% remain a sub-adult, 25% become an adult and 25% dies.
  - Of the current adults, 25% give birth to 1 juvenile. Apart from that 25% of the adults die and thus 75% survive to next year
1. Calculate how many juveniles, sub-adults and adults there will be next year.
  2. Calculate the following expression:

$$\begin{pmatrix} 0.4 & 0 & 0.25 \\ 0.4 & 0.5 & 0 \\ 0 & 0.25 & 0.75 \end{pmatrix} \begin{pmatrix} 20 \\ 16 \\ 12 \end{pmatrix}$$

3. How can the matrix in question 2 help you solve question 1?
4. Good conditions It turns out that the wind was very favourable this year and instead of 25% of the adults, one third of the albatrosses gave birth to a juvenile. Can you:
  - a) Correct the matrix in question 2 based on this information
  - b) Use this matrix to calculate how the population actually looks like after one year

## 2.3 Eigenvalues and eigenvectors

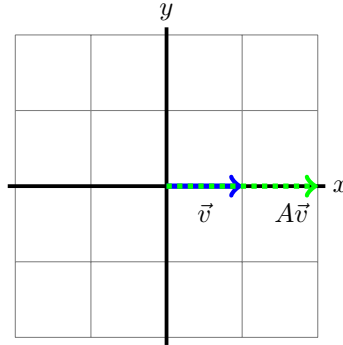
### 2.3.1 Graphically

We have seen that matrices can alter both the length and the direction of a vector. However, for each matrix there are vectors which direction is not changed

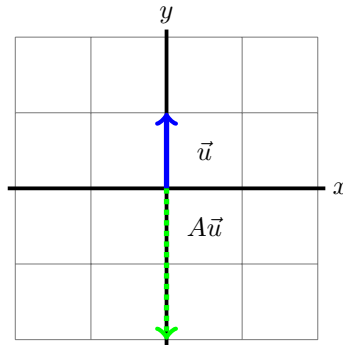
by the matrix. We call these eigenvectors of this matrix. Let us for example consider the matrix that multiplies the length of a vector by 2 and mirrors it in the x-axis:

$$A = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \quad (2.34)$$

Let us apply this matrix to the vector  $\vec{v}$ , which is such that the matrix does not change the direction of this vector:



Because  $A$  does not change the direction of the vector  $\vec{v}$ , we say that  $\vec{v}$  is an eigenvector of  $A$ .  $A$  does change the length of  $\vec{v}$  by a factor 2, therefore we say that the corresponding eigenvalue is 2. The matrix  $A$  has a second eigenvector:  $\vec{u}$ . This one is a bit more difficult to see.



Actually the direction of the vector did change: the new vector is pointing in the exact opposite direction as the original vector. Vectors for which this happens are also eigenvectors of  $A$ . The eigenvalue of corresponding to this eigenvector is  $-2$ . The minus sign indicates that the vector was inverted, the value of 2 indicates that the length of the vector was doubled.

### 2.3.2 With numbers

Numerically it is a bit easier to define the eigenvectors. The eigenvectors are those vectors of a matrix that when the matrix is applied on them, it is as if the vector is merely multiplied by a scalar instead of by a matrix. This scalar is the eigenvalue, which is usually denoted as  $\lambda$ .

Let us consider the example from the previous section: mirroring a vector in the x-axis. First we look at the effect of the matrix on the first eigenvector



that we identified in the previous section ( $\vec{v}$ ). From the graph we can read that  $\vec{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ . So

$$A\vec{v} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2\vec{v} \quad (2.35)$$

We see that when  $\vec{v}$  is transformed by  $A$ , this is the same effect as multiplying it with the scalar 2. We have thus confirmed that  $\vec{v}$  is an eigenvector. We can do the same for the second eigenvector. From the graph we can read that  $\vec{u} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ .

$$A\vec{u} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix} = -2\vec{u} \quad (2.36)$$

And so we see that also  $\vec{u}$  is an eigenvector of  $A$  and that this eigenvector has the eigenvalue  $\lambda = -2$ .

### 2.3.3 Exercises

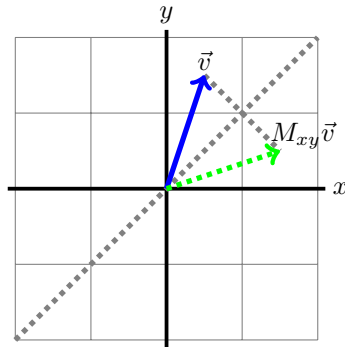
#### \* Exercise 2.4 Which of the two?

Below are four matrices, for each you can choose between two vectors. Choose the vector that is an eigenvector for that matrix and find the corresponding eigenvalue.

1. $\begin{pmatrix} 1 & 0.5 \\ 0.5 & 2 \end{pmatrix}$	a) $\begin{pmatrix} 0.38 \\ 0.92 \end{pmatrix}$	b) $\begin{pmatrix} 0.17 \\ 0.53 \end{pmatrix}$
2. $\begin{pmatrix} 3 & 1.5 \\ -0.5 & 1 \end{pmatrix}$	a) $\begin{pmatrix} 0.17 \\ 0.53 \end{pmatrix}$	b) $\begin{pmatrix} 0.95 \\ -0.32 \end{pmatrix}$
3. $\begin{pmatrix} 1 & 0 & 1.2 \\ 0.5 & 0.8 & 0 \\ 0 & 0.5 & 0.8 \end{pmatrix}$	a) $\begin{pmatrix} 0.63 \\ 0.27 \\ 0.31 \end{pmatrix}$	b) $\begin{pmatrix} 0.78 \\ 0.52 \\ 0.35 \end{pmatrix}$
4. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 0 \end{pmatrix}$	a) $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$	b) $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

#### \*\* Exercise 2.5 Finding eigenvectors

Consider the following transformation that mirrors every vector in the line  $x = y$  (we call this transformation  $M_{xy}$ ):



1. What are the two eigenvectors for this transformation? What are the corresponding eigenvalues?
2. Find the matrix that describes this transformation.
3. Show with this matrix that the eigenvectors and eigenvalues that you found in question 1 are indeed correct.

### \*\*\* Exercise 2.6 Re-arranging

Consider a 4-dimensional vector:

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}$$

#### Part I

It turns out that something went wrong in the notation and actually  $v_1$  should have been  $v_2$  and  $v_2$  should have been  $v_1$ :

$$P \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} v_2 \\ v_1 \\ v_3 \\ v_4 \end{pmatrix}$$

1. Find the matrix  $P$  that rearranges the vector.
2. Find the 4 eigenvectors and corresponding eigenvalues for  $P$ .
3. Give the matrix  $P^2$ .

*Hint: no need to do calculations, try to conceptually understand what happens when you apply  $P$  twice:  $PP\vec{v} = \dots$*

#### Part II

Consider now another re-arrangement in which all entries of  $\vec{v}$  have been shuffled. That is:

$$P_2 \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} v_2 \\ v_3 \\ v_4 \\ v_1 \end{pmatrix}$$

1. Write the matrix for this operation.
2. Find two of the four eigenvectors of this matrix and give the corresponding eigenvalues. (The remaining two eigenvectors contain complex numbers and are beyond the scope of this course)
3. Find  $n$  such that  $P_2^n = \mathbb{1}$ . Again, no need to do calculations.