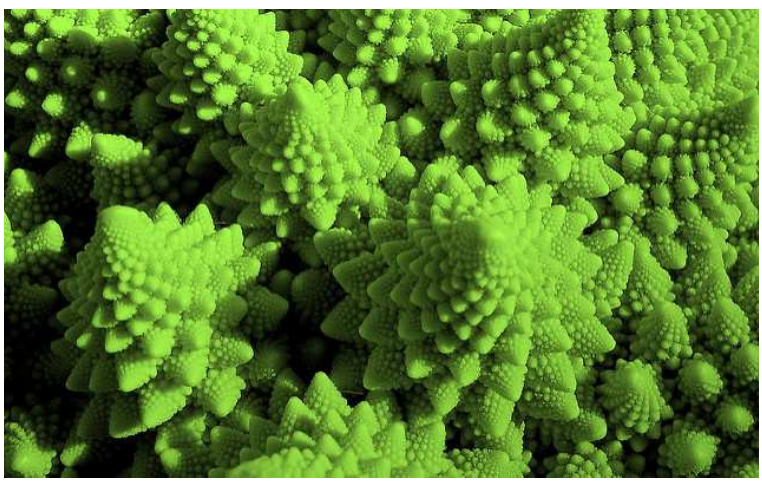
Maths 4Biology

Alberto Pascual-García

This project contains supporting notes and scripts for the course "Maths for Biology" of the Master for Computational Methods in Ecology and Evolution at Imperial College London. It is a fork of the the project "Mathematics Boot Camp" proposed by Viktor Jirsa for graduate courses at the Center for Complex Systems & Brain Sciences at Florida Atlantic University. Similar in spirit, by no means it is intended to compete with the traditional introductory courses in mathematics taught at universities. Rather it should be viewed as a synopsis of the mathematical tools most probably encountered in scientific applications. With respect to Jistra's notes, the present project expands some of the topics, and it further includes First-Order Differential Equations, of oustanding importance in Ecology. On the other hand, Fourier analysis is not considered in the course, although it is kept in the repository for completeness.



Zürich, December 2018
Institute of Integrative Biology
ETH

Contents

List of Figures

1 Function Theory

1.1 Foundations

Definition: Consider two sets of elements A and B. A *function* f is a rule which unambiguously assigns $y \in B$ to each $x \in A$. We call the set A the *domain* of the function and the set B its *codomain*.

Note: "$x \in A$" means that "x is an element of A".

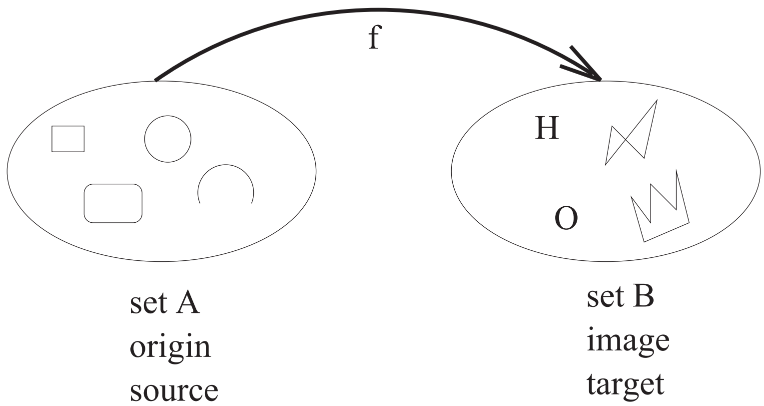


Figure 1: Illustration of a function

Notation:

$$\begin{aligned} f: &\quad A \to B \\ f: &\quad A \to B \qquad x \in A,\ y \in B \\ & y = f(x) \qquad x \in A,\ y \in B \\ & y = y(x) \qquad x \in A,\ y \in B \end{aligned}$$

Examples:

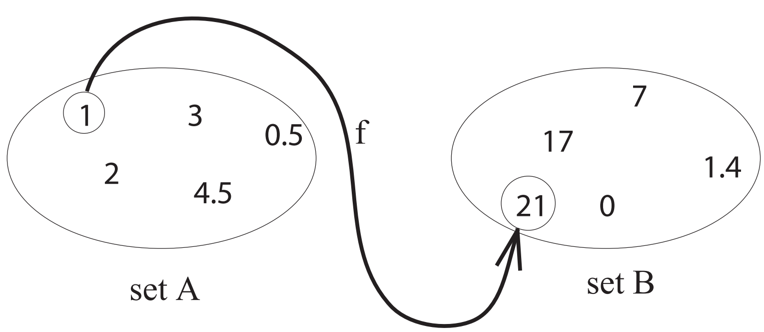
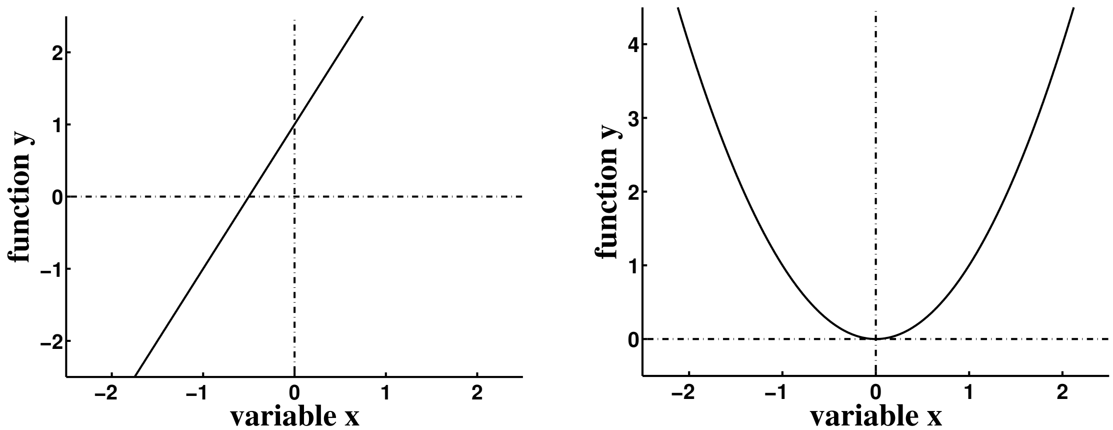


Figure 2: Assigning an element $\{21\} \in B$ to $\{1\} \in A$

Most commonly, functions are defined by equations: a) $y = f(x) = 2x + 1$, b) $y = f(x) = x^2$

Graphical representation:



(a) Function $y = 2x + 1$

(b) Function $y = x^2$

Figure 3: A linear and a quadratic function

Types: Some important types of functions are:

1. Injective functions: Those functions that preserve distinctiveness, i.e. they never map distinct elements of its domain to the same elements of its codomain.
2. Surjective functions: Every element y of the codomain has a correspondent element x of the domain that is mapping it.
3. Bijective functions: Those functions that are both injective and surjective.

1.2 Domain, parity and periodicity of functions

Domain of a function is the set of input values in which the function can be defined. Therefore, given a map between two sets, to define a function we should determine a valid domain identifying possible *pathological* input values.

Example: The map $f(x) = 1/x$ has a pathological value at $x = 0$ meaning that the map diverges to infinity. In order for this map to be a function, we should *restrict the domain* and define the function as:

$$f(x) = \begin{cases} 1/x & \text{if } x \neq 0. \\ 0 & \text{if } x = 0. \end{cases}$$

Parity of a function. Let $f(x)$ be a real-valued function, then:

1. f is *even* if the following equation holds for all x in the domain of f:

$$f(x) = f(-x) \quad \rightarrow \quad f(x) - f(-x) = 0$$

2. f is *odd* if

$$-f(x) = f(-x) \quad \rightarrow \quad f(x) + f(-x) = 0.$$

Note: These properties will be very useful to guess some properties of derivatives and integrals of functions.

Periodicity of a function. We say that a function $f(x)$ is periodic if for all x in the domain of f there is a value $\lambda \in \mathbb{R}$ such that for a every $n \in \mathbb{N}$ it holds that $f(x) = f(x + n\lambda)$, where λ is the *period* of the function.

1.3 Inverse Functions

f^{-1} denotes the inverse of the function f.

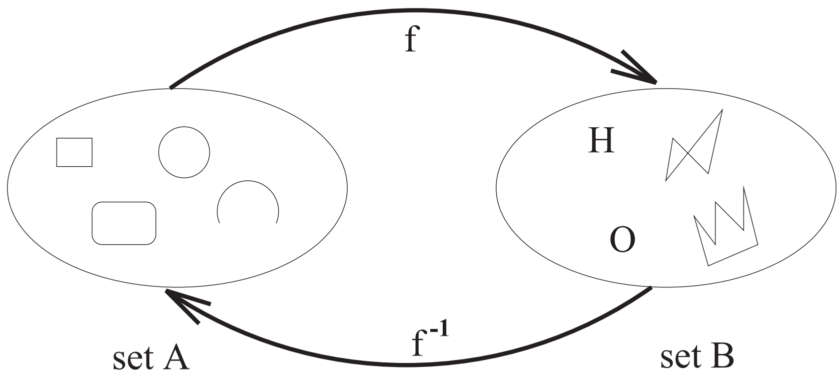


Figure 4: Inverse function

Notation:

$$f^{-1}: \; B \rightarrow A$$

$$x = f^{-1}(y) \qquad \text{where} \qquad y = f(x)$$

Graphically the inverse can be constructed as the mirror image of the function at the first bisector. This method always works, but caution is asked for, because the inverse may not be unique and require more detailed discussion.

Example:

$$y = f(x) = 2x + 1 \qquad x = f^{-1}(y) = \frac{1}{2}(y - 1)$$

Note: There is not always an inverse function!

Example:

$$y = f(x) = x^2$$

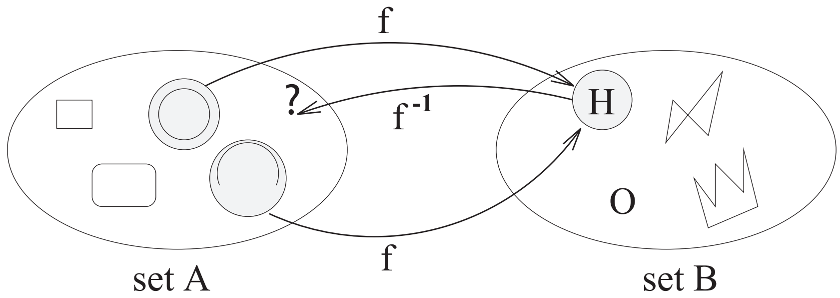
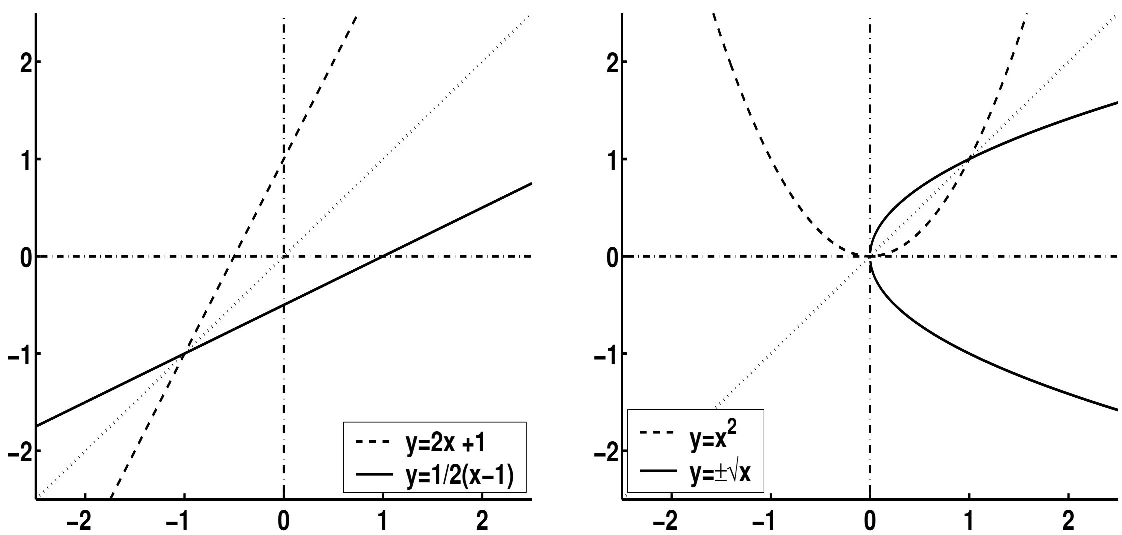


Figure 5: The inverse f^{-1} is not unique, thus not a function.

$$x = \sqrt{y} \qquad \text{or} \qquad x = -\sqrt{y}$$



(a) $f(x) = 2\,x + 1 \;\rightarrow\; f^{-1}(x) = \frac{1}{2}(x-1)$ (b) $f(x) = x^2 \;\rightarrow\; f^{-1}(x) = \pm\sqrt{x}$

Figure 6: Graphical construction of the inverse function

1.4 Implicit Functions

A function is not given explicitly as in $y = f(x)$, but implicitly by $F(x, y) = 0$. Note that an implicit equation may not necessarily fulfill the conditions of a function when written explicitly, unless some additional conditions are imposed.

Example: Unit circle: $F(x, y) = x^2 + y^2 - 1 = 0$. In the plot, we observe that each x point corresponds to two y points and hence it is not a function.

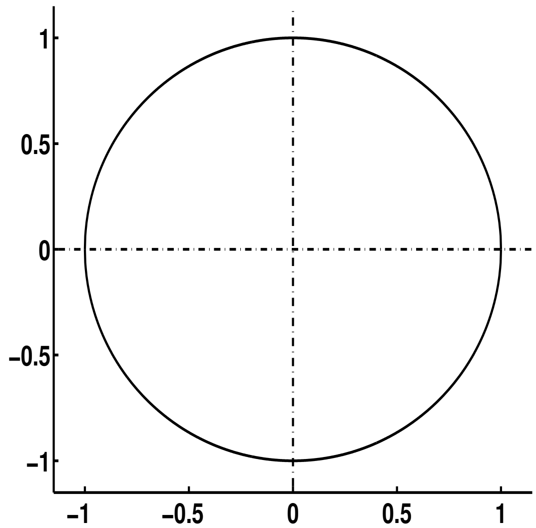


Figure 7: Unit circle

The implicit representation of the unit circle needs additional conditions to become unique and thus a function: a local neighborhood has to be defined, e.g. $y = \sqrt{1 - x^2}$ for $x \in (-1; 1)$, $y > 0$ and $y = -\sqrt{1 - x^2}$ for $x \in (-1; 1)$, $y < 0$.

Note: The implicit representation is particularly important for algebraic equations of the form:

$$a_n(x)y^n + a_{n-1}y^{n-1} + ... + a_0(x) = 0$$

since those of order $n > 5$ may not have an explicit representation.

1.5 Polynomials

Polynomials are defined as a class of functions of the form

$$y = a_0 + a_1\,x + a_2\,x^2 + \cdots + a_N\,x^N = \sum_{n=0}^{N} a_n\,x^n$$

where the function y is said to be a polynomial of order N.

Example:

$$y = \underbrace{2}_{a_2}x^2 + \underbrace{8}_{a_1}x + \underbrace{4}_{a_0}$$

Goal: To achieve a qualitative understanding of a given function without computing each value.

Approach:

$$y = \sum_{n=0}^{N} a_n\, x^n \qquad \text{where we assume } a_N > 0$$

Step 1: If N is even, then $x \to \pm\infty : y \to ?$; if N is odd, then $x \to \pm\infty : y \to ?$. **Note:** If $a_N < 0$, the behavior is the opposite. **Note:** I'm sorry about the question marks, but it is an exercise you should solve!

Step 2: (Fundamental theorem of algebra) A polynomial of order N has N roots which are the solutions of $f(x) = 0$.

Set $y = 0 : f(x) = x^N + a_{N-1}\, x^{N-1} + \cdots + a_1\, x + a_0 = 0$

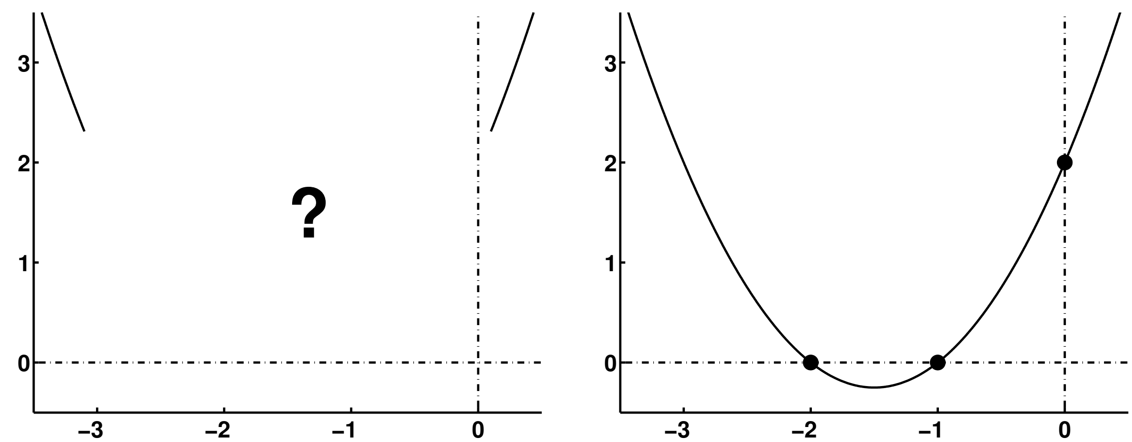
$$y = 17x + 4 \qquad y = 0 : x = -\tfrac{4}{17} \qquad \to \quad 1 \text{ root}$$

$$y = 2x^2 + 8x + 4 \qquad y = 0 : x = \tfrac{1}{4}(-8 \pm \sqrt{8^2 - 4 \cdot 2 \cdot 4}) = \tfrac{1}{4}(-8 \pm \sqrt{32)} = -2 \pm \sqrt{2} \qquad \to \quad 2 \text{ roots}$$

$\Rightarrow$ the roots are the locations where $f(x)$ crosses the x-axis.

Note: Analytic formulas to find the roots of polynomials are known until 4th degree.

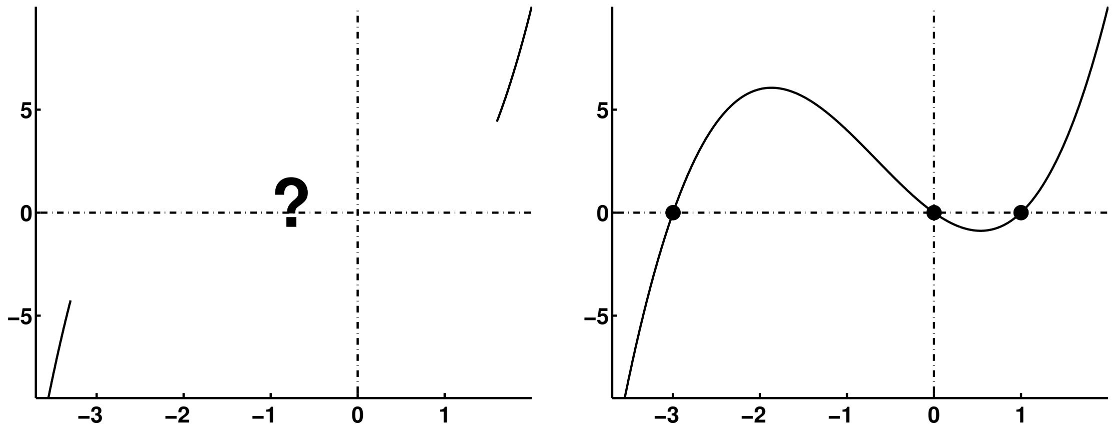
Examples: Construct graph of $y = f(x)$



(a) Step 1: $y = x^2 + 3x + 2$

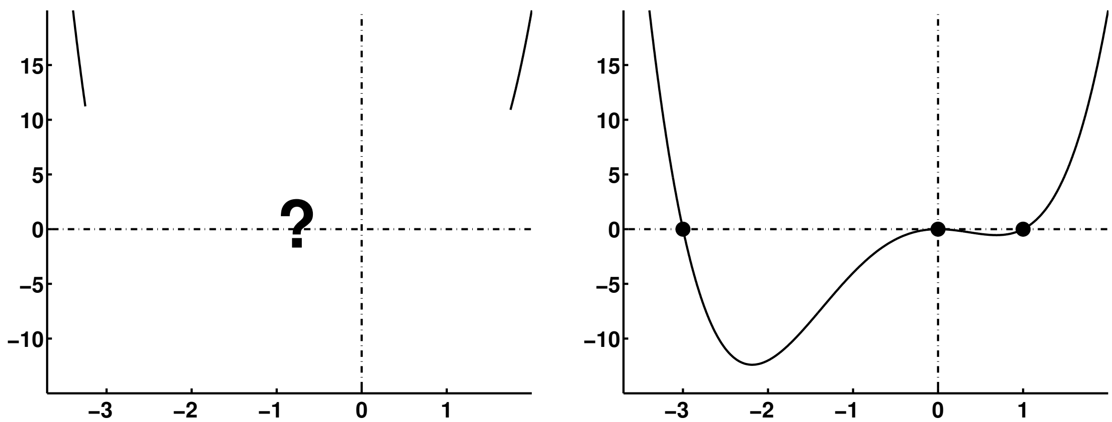
(b) Step 2: $y = x^2 + 3x + 2 = 0$

Figure 8: Graphical construction (the roots are $x = -1$ and $x = -2$)



(a) Step 1: $y = x^3 + 2x^2 - 3x$ (b) Step 2: $y = x^3 + 2x^2 - 3x = 0$

Figure 9: Graphical construction (the roots are $x = -3, x = 0$ and $x = 1$)



(a) Step 1: $y = x^4 + 2x^3 - 3x^2$ (b) Step 2: $y = x^4 + 2x^3 - 3x^2 = 0$

Figure 10: Graphical construction (the roots are $x = -3, x = 0$ and $x = 1$)

Horizontal translation is a horizontal shift of a function by x_0

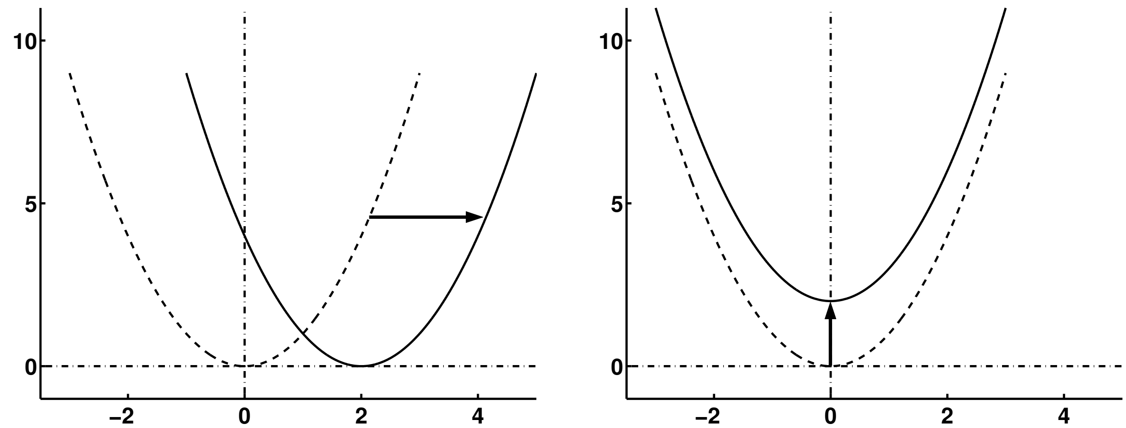
$$y = f(x) \ \rightarrow \ y = f(x - x_0)$$

Example: $y = x^2$ shift by $x_0 = 2$: $y = (x-2)^2 = x^2 - 4x + 4$

Vertical translation is a vertical shift of a function by y_0

$$y = f(x) \ \rightarrow \ y = f(x) + y_0$$

Example: $y = x^2$ shift by $y_0 = 2$: $y = x^2 + 2$



(a) Horizontal shift by $x_0 = 2 : y = (x-2)^2$ (b) Vertical shift by $y_0 = 2 : y = x^2 + 2$

Figure 11: Vertical and horizontal shift of $y = x^2$

1.6 Rational functions

Given two polynomial functions $D(x)$ and $d(x)$ a rational function has the form:

$$f(x) = \frac{D(x)}{d(x)}$$

Working with these functions can be cumbersome if the polynomials are complicated. It is useful to learn how to factorize the polynomials, because it will allow us to find equivalent (simplified) functions.

1. Binomial expressions. Second order polynomials can be factorized in the product of two first order polynomials, i.e. $ax^2 + bx + c = (x+u)(x+v)$. Some immediate factorizations are:

$$a^2x^2 + b^2 + 2abx = (ax+b)(ax+b)$$
$$a^2x^2 + b^2 - 2abx = (ax-b)(ax-b)$$
$$a^2x^2 - b^2 = (ax+b)(ax-b)$$

Nevertheless, other expressions are not so easy. In general, we know that we are looking for something like $(x+u)(x+v)$ and hence, if we have for instance

$$x^2 + bx + c = (x+u)(x+v) = x^2 + (u+v)x + uv,$$

we can see that $u + v = b$ and $uv = c$, which are two equations with two unknowns that we can solve. A little bit more general expression would be:

$$ax^2 + bx + c = (ax+u)(x+v) = ax^2 + (u+av)x + uv,$$

leads to $u + av = b$ and $uv = c$.

Example:

$$\frac{x^4+8x^2+7}{3x^5-3x}=\frac{(x^2+7)(x^2+1)}{3x(x^2+1)(x^2-1)}=\frac{x^2+7}{3x(x^2-1)}$$

2. Degree of $D(x)$ is larger than $d(x)$. In this situation we can simply divide the numerator by the denominator! Calling $q(x)$ to the quotient and $r(x)$ to the remainder of the division, we get:

$$\frac{D(x)}{d(x)}=q(x)+\frac{r(x)}{d(x)}.$$

Example:

$$\frac{3x^3-2x^2+4x-3}{x^2+3x+3}=(3x-11)+\frac{28x+30}{x^2+3x+3},$$

3. Degree of $D(x)$ is smaller than $d(x)$. To solve this case we proceed in three steps that we illustrate with one example:

Step 1 – Factorize the denominator.

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

Step 2 – Rewrite the function as a sum of rational functions whose numerators are polynomials with one degree lower than those of the denominator, and whose coefficients are unknown. Then sum up the functions and group the terms by their degree on x.

$$\frac{x}{(x-1)(x-2)}=\frac{A}{(x-1)}+\frac{B}{(x-2)}=$$
$$=\frac{A(x-2)+B(x-1)}{(x-1)(x-2)}=\frac{(A+B)x-(2A+B)}{(x-1)(x-2)}$$

Step 3 – Identify the terms in the numerator of the last expression with those of the original function, and find the coefficients A, B, etc.

$$(A+B)x-(2A+B)=x.$$

Since the coefficient of x in the numerator of the original function is 1 we get $A+B=1$ and, since the independent term is zero, we get $2A+B=0$. Considering both equations for the variables A and B we conclude that $A=-1$ and $B=2$, and the final factorization is:

$$\frac{x}{(x-1)(x-2)}=\frac{-1}{(x-1)}+\frac{2}{(x-2)}$$

Example: It may happen that the multiplicity of any of the factors in the denominator is larger than one, in which case the decomposition at Step 2 becomes:

$$\frac{x}{(x-1)^3} = \frac{A}{(x-1)} + \frac{B}{(x-1)^2} + \frac{C}{(x-1)^3}$$

and we proceed similarly

$$\frac{x}{(x-1)^3} = \frac{A(x-1)^2 + B(x-1) + C}{(x-1)^3} = \frac{Ax^2 + (B-2A)x + (A+C-B)}{(x-1)^3}$$

finding that $A = 0$, $B = 1$ and $C = 1$.

1.7 Trigonometric Functions

Trigonometric functions are a class of periodic functions such as

$$y = f(x) = A\sin(kx + \phi) \quad \text{and} \quad y = f(x) = A\cos(kx + \phi)$$

The constant parameters are the amplitude A, the frequency k and the phase (angle) ϕ. The period is defined as $\lambda = 2\pi/k$ and the trigonometric functions fulfill the relation $f(x+\lambda) = f(x)$.

Special values of trigonometric functions (midnight stuff)								
x	0	$\pi/6 = 30^\circ$	$\pi/4 = 45^\circ$	$\pi/3 = 60^\circ$	$\pi/2 = 90^\circ$	$\pi = 180^\circ$	$3\pi/2 = 270^\circ$	$2\pi = 360^\circ$
$\sin x$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1	0	-1	0
$\cos x$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0	-1	0	1

Horizontal translation (shift) by means of ϕ:

$$\sin(x + \frac{\pi}{2}) = \cos x \qquad \cos(x + \frac{\pi}{2}) = -\sin x \qquad \sin(x - \frac{\pi}{2}) = -\cos x \qquad \cos(x - \frac{\pi}{2}) = \sin x$$

Useful relations between $\cos x$ and $\sin x$:

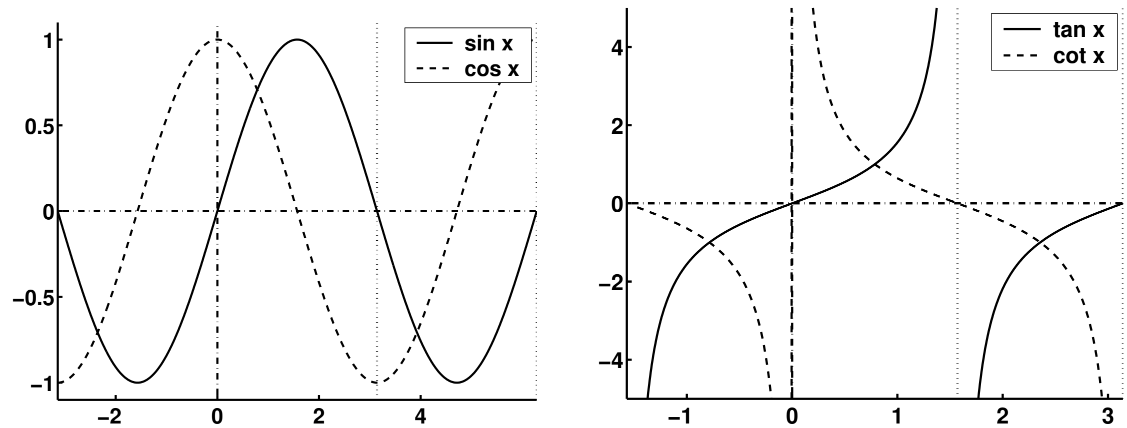
$$\cos^2 x + \sin^2 x = 1$$

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y \qquad \cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$$

$$\sin 2x = 2\sin x \cos x \qquad \cos 2x = \cos^2 x - \sin^2 x$$

Other trigonometric functions:

$$\tan x = \frac{\sin x}{\cos x} \qquad \cot x = \frac{\cos x}{\sin x}$$



(a) The functions $y = \sin x$ and $y = \cos x$

(b) The functions $y = \tan x$ and $y = \cot x$. Note that these are π-periodic

Figure 12: The most common trigonometric functions

1.8 Exponential Functions

Exponential functions are functions most commonly used in the form

$$y = Ae^{kx} = A \exp kx$$

with the constant parameters: A amplitude, k growth rate if $k > 0$ and the damping or fall off, if $k < 0$, and e Euler number: 2.718....

Note that $e^0 = 1$ and $e^{-x} = \frac{1}{e^x}$.

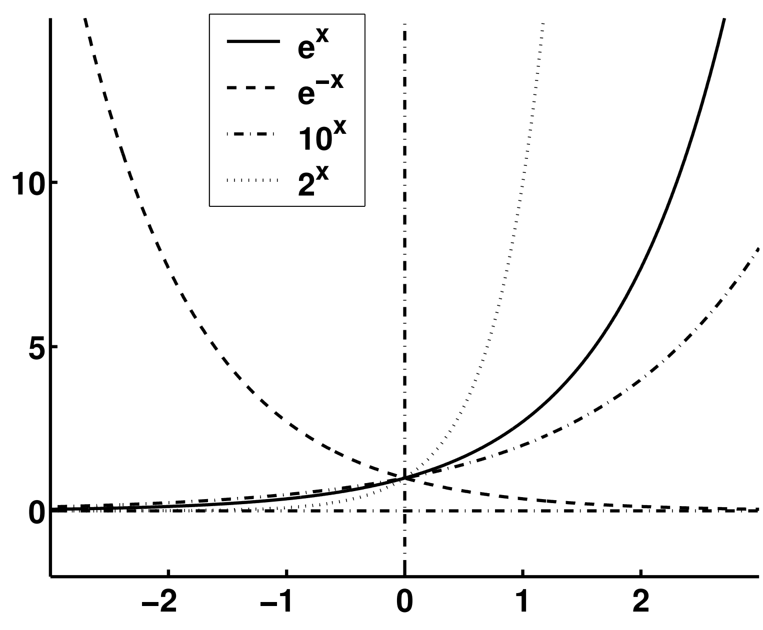


Figure 13: Exponential functions

1.9 Hyperbolic Functions

Hyperbolic functions are of the form

$$y = f(x) = \cosh x = \frac{1}{2}(e^x + e^{-x}) \qquad \text{hyperbolic cosine}$$

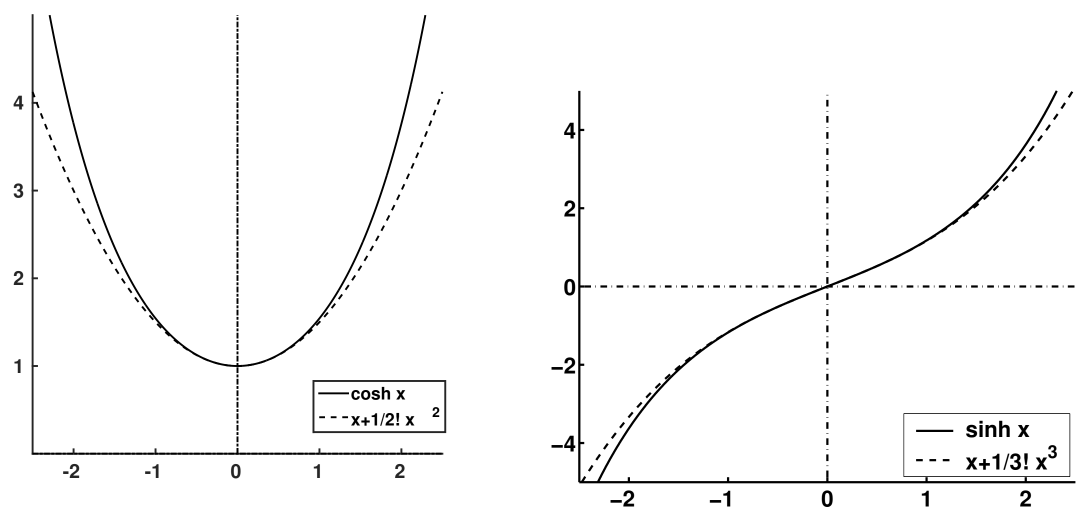
and

$$y = f(x) = \sinh x = \frac{1}{2}(e^x - e^{-x}) \qquad \text{hyperbolic sine}$$

They have similar properties as the trigonometric functions such as a representation by exponentials (as we shall see later), and their derivatives convert into each other. But the hyperbolic functions are *not* periodic.

Other hyperbolic functions:

$$y = \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad y = \coth x = \frac{\cosh x}{\sinh x} = \frac{e^x + e^{-x}}{e^x - e^{-x}}$$



(a) $y = \cosh x$ and $y = 1 + \frac{1}{2!}\,x^2$ (b) $y = \sinh x$ and $y = x + \frac{1}{3!}\,x^3$

Figure 14: Hyperbolic functions

1.10 Basic Inverse Functions

1.10.1 Logarithms

The logarithms are the inverse of the exponential functions:

$$y = a^x \quad \leftrightarrow \quad x = \log_a y \qquad \text{where} \quad 0 < y < \infty$$

Special cases:

$$\begin{array}{llllll} a=e: & y & = & \log_e x & = & \ln x & \text{natural logarithm} \\ a=10: & y & = & \log_{10} x & = & \mathrm{lg}x & \text{decimal logarithm} \\ a=2: & y & = & \log_2 x & = & \mathrm{ld}x & \text{dual logarithm} \end{array}$$

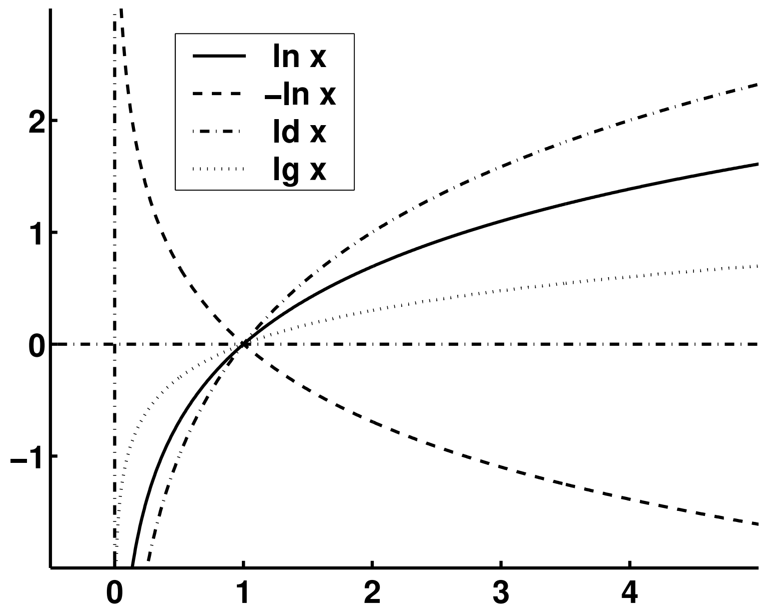


Figure 15: The logarithmic functions $y = \ln(x)$, $y = \mathrm{lg}(x)$, $y = \mathrm{ld}(x)$, and $y = -\ln(x)$.

Remark: The most commonly used logarithm is $\ln x$, but there are certain applications for other logarithms as well. For instance, the decimal logarithm can be used to find the number of digits in a decimal number ($\mathrm{lg}4821 = 3.683 \rightarrow$ taking the whole number in front of the decimal point and adding 1 gives the number of digits, 4). Similarly, the dual logarithm can be used to find the number of bits or binary digits that are necessary to represent a number n in binary format, i.e. as zeros and ones.

Rules and tricks for dealing with logarithms:

$$\log_a x^n = n \log_a x$$

$$\log_a x_1 x_2 = \log_a x_1 + \log_a x_2 \qquad \log_a \frac{x_1}{x_2} = \log_a x_1 - \log_a x_2$$

$$\log_a x = \frac{\log_b x}{\log_b a}$$

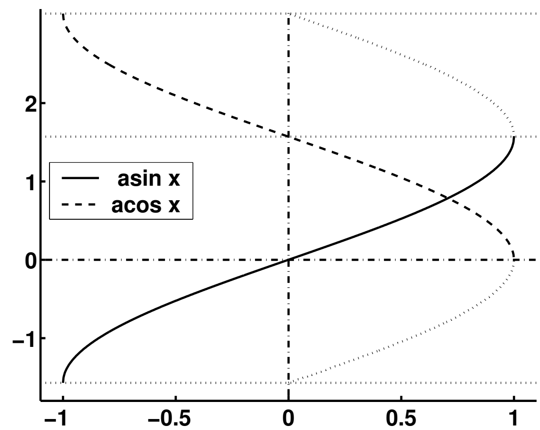
Note: The last expression tells us that changing the base simply changes the values of the function by a constant. Every logarithm can be expressed in terms of the natural logarithm, and every exponential function can be written in terms of the basis e

A particularly useful relation is the following:

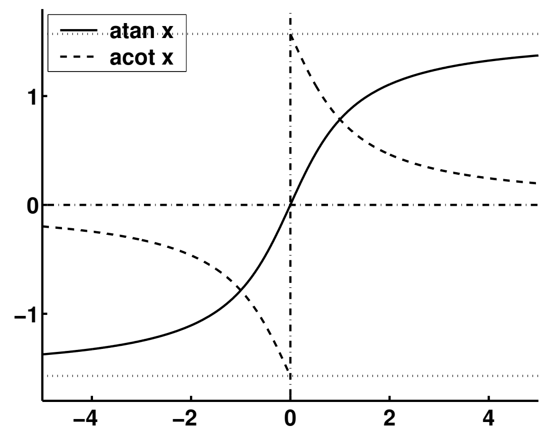
$$a^x = e^{\ln a^x} = e^{x \ln a} \quad \text{with} \quad a > 0$$

1.10.2 Other inverse functions

$$
\begin{array}{llll}
y = \sin x & \rightarrow & x = \arcsin y & \text{arc sine} \\
y = \cos x & \rightarrow & x = \arccos y & \text{arc cosine} \\
y = \tan x & \rightarrow & x = \arctan y & \text{arc tangent} \\
y = \cot x & \rightarrow & x = \text{arccot }\, y & \text{arc cotangent} \\
y = \sinh x & \rightarrow & x = \text{arcsinh }\, y & \\
 & & x = \ln(y + \sqrt{y^2+1}) & \text{area sine hyperbolic} \\
y = \cosh x & \rightarrow & x = \text{arccosh }\, y & \\
 & & x = \ln(y + \sqrt{y^2-1}) & \text{area cosine hyperbolic}
\end{array}
$$



(a) Arc sine, and arc cosine



(b) Arc tangent, and arc cotangent

Figure 16: The inverse of the trigonometric functions

1.11 Elementary Combinations of Functions

1.11.1 Superposition

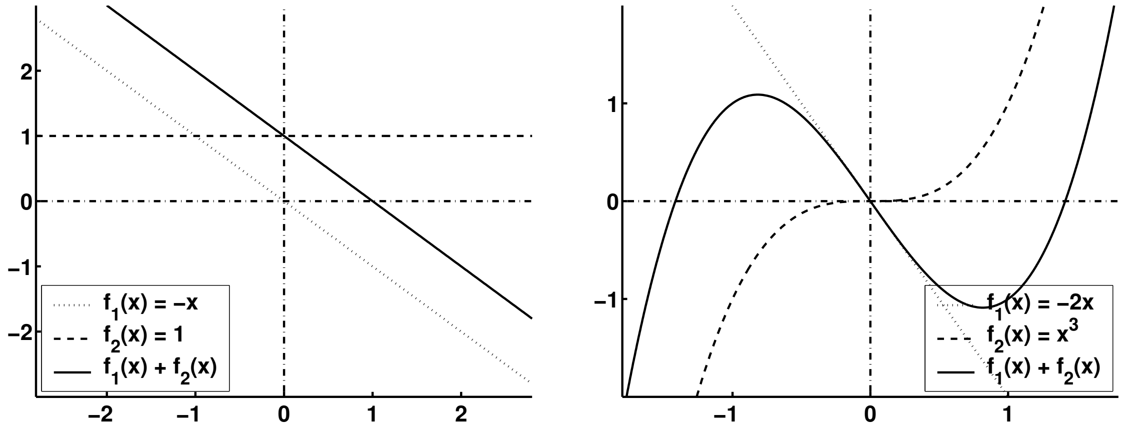
Two functions are superimposed on each other by adding their values for the same x.

$$y = f_1(x) + f_2(x)$$

1.11.2 Modulation

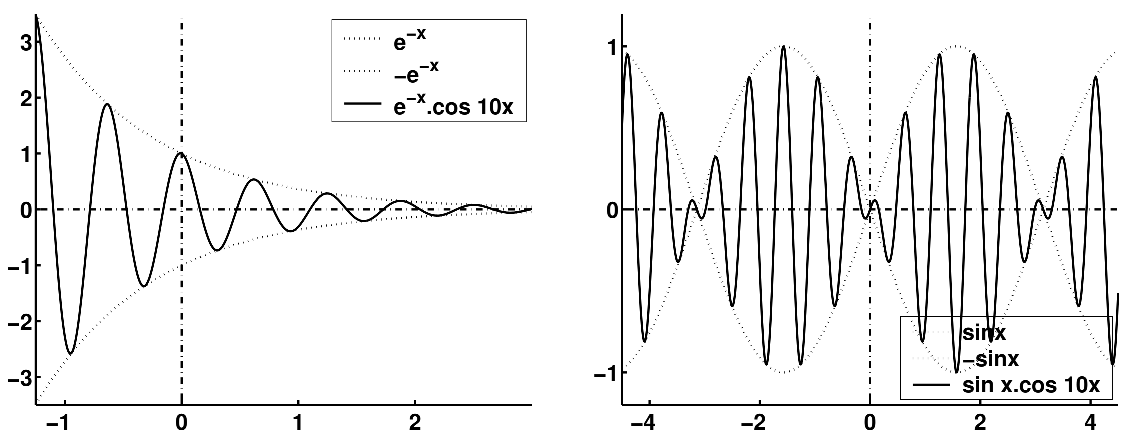
A function is modulated by another function by multiplying their values for the same x.

$$y = f_1(x)\, f_2(x)$$



(a) Superposition of $f_1(x)\!=\!-x$ and $f_2(x)\!=\!1$ (b) Superposition of $f_1(x)\!=\!-2x$ and $f_2(x)\!=\!x^3$

Figure 17: Superposition of lines and functions



(a) e^{-x} is the envelope function of $\cos 10x$ (b) $\cos 10x$ is modulated with $\sin x$

Figure 18: Modulation of functions

2 Limits and Derivatives

2.1 Limits

Many times we are interested in studying the behaviour of a function in the proximity of a certain value. This value can be, in principle, any value, but the use of limits typically concerns those values that are singular, perhaps because do not belong to the domain of the function or because we are interested in its behaviour at the infinity. A typical question of interest in biology, may be which is the expected behaviour of a population if we assume that its size is so large that we consider it infinite. In that case, we will take the limit of the function describing the population at the infinite, and this is many times useful because the function may be simplified, and will allow us to perform further analytical development.

As we said, the limit of a function at an arbitrary point may be easy to compute. To compute the limit of a function $f(x)$ as x approaches a, that we write $\lim_{x\to a} f(x)$, we start evaluating $f(a)$.

Example:

$$\lim_{x\to 1} f(x) = \lim_{x\to 1} \frac{1}{2}(x+3) \underbrace{=}_{f(1)} \frac{1}{2}(1+3) = 2.$$

In this case, it was very easy and not very interesting. Let's see a precise definition of limit which is, perhaps, the most abstract definition we will learn in this course, and how we can solve more complicated examples. Learning the definition will be a good training for our mind to get into the abstraction of concepts such as the infinity, and to open the door towards the important world of the derivatives.

Definition: The limit of a function $f(x)$ as x approaches a, that we write $\lim_{x\to a} f(x)$, is a number l such that, given any target distance ε between $f(x)$ and l, it is always possible to find a value x such that its distance δ with respect to a is such that the distance between $f(x)$ and l remain lower than ε, i.e.

$$\lim_{x\to a} f(x) = l \Leftrightarrow (\forall\varepsilon > 0\ \exists\delta > 0\ /\ 0 < |x-a| < \delta \Rightarrow |f(x) - l| < \varepsilon)$$

Wow, the definition is really ugly. Let's try to solve the above example with this definition.

Example: Demonstrate that the $\lim_{x\to 1} \frac{1}{2}(x+3) = 2$ using the definition of limit.

What we look for is a positive distance δ such that if we fix an arbitrary ε, if $|x-1| < \delta$ then $|\frac{1}{2}(x+3) - 2| < \varepsilon$. We can rewrite the latter condition:

$$|\frac{x+3-4}{2}| < \varepsilon \Rightarrow |\frac{x-1}{2}| < \varepsilon \Rightarrow |x-1| < 2\varepsilon.$$

And it looks like it is easy to find δ, because if we make $\delta = 2\varepsilon$, it actually happens that:

$$0 < |x-1| < 2\varepsilon \Rightarrow |\frac{1}{2}(x+3) - 2| < \varepsilon,$$

which fulfills the definition of limit. Let's now look for a more interesting example, because we look for a limit at a pathological value.

Example: Demonstrate that the $\lim_{x\to 0} x\sin(1/x) = 0$ using the definition of limit.

Again, we look for a positive distance δ such that if $|x-0| < \delta$ then $x\sin(1/x) - 0 < \varepsilon$. Given that the image of the sin is bounded between zero and one, we note that $0 \leq \sin(1/x) \leq 1$ and, hence, $|x\sin(1/x)| \leq |x| \leq \varepsilon$. Therefore, it is true that this is the limit, because it is enough to say that $\delta = \varepsilon$ to see that if $|x| < \varepsilon$ also $|x\sin(1/x)| < \varepsilon$, which is what we were willing to demonstrate.

2.2 Lateral limits, continuity of functions and assymptotes

The lateral limit is the limit of a function when we approximate a value a either from x values smaller than a (it is said *from the left* and we write $\lim_{x\to a^-} f(x)$) or from x values larger than a (*from the right*, $\lim_{x\to a^+} f(x)$). The formal definitions are:

$$\lim_{x\to a^-} f(x) = l \Leftrightarrow (\forall\varepsilon > 0\ \exists\delta > 0\ /\ 0 < x - a < \delta \Rightarrow |f(x) - l| < \varepsilon)$$
$$\lim_{x\to a^+} f(x) = l \Leftrightarrow (\forall\varepsilon > 0\ \exists\delta > 0\ /\ 0 < a - x < \delta \Rightarrow |f(x) - l| < \varepsilon).$$

The definition of lateral limits leads to two important theorems:

Theorem: The limit of a function exists if and only if its lateral limits exist and are equal.

Theorem: A function is *continuous* in x_0 if the limit exist and it is equal to the value of the function at x_0.

With these theorems we can determine if a function is continuous (we will investigate its pathological values) and, if it is not continuous, which kind of discontinuity it has.

Example: Continuous function.

$$f(x) = \begin{cases} x+3 & \text{if } x \leq 1. \\ 4 & \text{if } x > 1. \end{cases}$$

In this case, we observe a possible pathological value at $x = 1$. Nevertheless, the function is continuous because $\lim_{x\to 1^+} f(x) = \lim_{x\to 1^-} f(x) = f(1)$.

Example: Function with a "removable" discontinuity.

$$g(x) = \begin{cases} 3 & \text{if } x \neq 0. \\ 2 & \text{if } x = 0. \end{cases}$$

With this function it happens that, at $x = 0$, the lateral limits exists and are equal, but the function takes a different value, i.e. $\lim_{x\to 0^+} g(x) = \lim_{x\to 0^-} g(x) \neq g(0)$. We say that the function is discontinuous but, since the limit exists and it is finite, we call it a "removable" discontinuity (in some sense we think of the limit as the "true" value).

Example: Function with a "finite jump" discontinuity.

$$h(x) = \begin{cases} 3 & \text{if } x < 0. \\ 4 & \text{if } x > 0. \\ 2 & \text{if } x = 0. \end{cases}$$

This function it happens that, at $x = 0$, the lateral limits exists but are not equal, but there is a finite difference between both values so we say that it is a finite discontinuity.

Example: Function with an infinite discontinuity.

$$i(x) = \frac{x+1}{x-2} \Rightarrow \begin{cases} \lim_{x\to 2^-} i(x) = -\infty \\ \lim_{x\to 2^+} i(x) = \infty \end{cases}$$

Definition: We will say that a function $fa(x)$ approaches assymptotically to a line (e.g. $y = a$ or $x = a$), or that the line is an assymptote of the function, if the distance between the function and the curve approaches to zero when one or both x and $f(x)$ tend to infinity.

Examples:

1. Vertical assymptote: $\lim_{x\to a} f(x) = \infty$
2. Horizontal assymptote: $\lim_{x\to\infty} f(x) = a$

2.3 Limits with an indeterminate form

In many situations, when we evaluate a limit we do not have enough information to determine if the limit exists, in which case we face an *indeterminate form.* These forms are functions that, after being evaluated, lead to expressions of this kind:

$$\frac{0}{0}, \frac{\infty}{\infty}, \infty - \infty, 1^{\infty}, 0^0, 0\infty, \infty^0,$$

and require special techniques to solve them. In the following subsections we summarize some of the most common techniques.

2.3.1 Rational limits to infinity

For rational functions, we should consider how fast the functions in the numerator and in the denominator tend to infinity. There is an order on how fast functions tend to infinity:

$$x^{kx} > b^x > x^m > \log x$$

Where the symbol $>$ means that one function grows faster than the other one, and $k > 0$.

Examples:

1. $\lim_{x\to\infty} \frac{e^{3x}}{4x^2} = \infty$
2. $\lim_{x\to\infty} \frac{\log(x)}{x} = 0$
3. $\lim_{x\to\infty} \frac{18x^2+1}{32x^2+3} = \frac{9}{16}$

2.3.2 Infinitesimal equivalents

On the other hand, when two functions become infinitesimally small when they converge towards the same point, we will say that they are *infinitesimal equivalents*, and we can use this fact to find limits of rational functions. We say that $f(x)$ and $g(x)$ are infinitesimal equivalents around a if

$$\lim_{x\to a} \frac{f(x)}{g(x)} = 1.$$

Some common infinitesimal equivalents are:

1. When $x \to 0$:
$$x \simeq \sin(x);\ x \simeq \tan(x);\ x \simeq \log(1+x);\ x \simeq e^x - 1$$
$$x \simeq \arcsin(x);\ x \simeq \arctan(x);\ 1 - \cos(x) \simeq \frac{x^2}{2}$$
2. When $x \to 1$:
$$x - 1 \simeq \log(x);$$

Example:

$$\lim_{x\to 0} \frac{\sin(x)}{\ln(1+x)} = \lim_{x\to 0} \frac{x}{x} = 1$$

Example:

$$\lim_{x\to 0} \frac{\tan(x)(1-\cos(x))}{x^3} = \lim_{x\to 0} \frac{x(1-\cos(x))}{x^3} = \lim_{x\to 0} \frac{x^2/2}{x^2} = \frac{1}{2}.$$

2.3.3 Algebraic operations

Many times, we can simplify the expression or find an appropriate change of variables before we compute the limit, that will solve the indeterminacy.

Example: Rational factorization

$$\lim_{x\to 3}\frac{x^2-9}{x^2-5x+6}=\lim_{x\to 3}\frac{(x+3)(x-3)}{(x-2)(x-3)}=\lim_{x\to 3}\frac{x+3}{x-2}=6.$$

Example: Rational factorization

$$\lim_{x\to 3}\frac{\sqrt{x+1}-2}{x-3}=\lim_{x\to 3}\frac{(\sqrt{x+1}-2)(\sqrt{x+1}+2)}{(x-3)(\sqrt{x+1}+2)}=$$
$$=\lim_{x\to 3}\frac{(x+1)-4}{(x-3)(\sqrt{x+1}+2)}=\lim_{x\to 3}\frac{1}{\sqrt{x+1}+2}=\frac{1}{4}.$$

Example: Change of variables

$$\lim_{x\to 2}\frac{\sin(x^3-8)}{x-2}\quad \underbrace{=}_{\uparrow}\lim_{t\to 0}\frac{\sin((t+2)^3-8)}{t}$$
$$\text{change vars: }\begin{cases} t=x-2\Rightarrow x=t+2\\ x\to 2\Rightarrow t\to 0\end{cases}$$

The change of variables does not seem to help much, but if we remind the formula for the cube of a binomial, which we recall here:

$$(a+b)^3=a^3+b^3+3a^2b+3ab^2$$
$$(a-b)^3=a^3-b^3-3a^2b+3b^2a,$$

and we apply it to $(t+2)^3=t^3+6t^2+12t+8$, the numerator simplifies:

$$\lim_{t\to 0}\frac{\sin(t^3+6t^2+12t)}{t}\quad \underbrace{=}_{\uparrow}\lim_{t\to 0}\frac{t(t^2+6t+12)}{t}=12.$$
$$\text{Infinitesimal equivalents:}\qquad \sin(f(x))\simeq f(x)$$

2.3.4 L'Hôpital rule

For indeterminate forms of the type 0/0 or ∞/∞ there is a rule that may allow us to find a limit. But we need to learn first derivatives! We will come back to this question in the section of Applications of Derivatives.

2.4 Derivatives: the Difference Quotient

First derivatives of simple functions were studied by Galileo Galilei (1564-1642) and Johannes Kepler (1571-1630). A systematic theory of differential calculus was developed by Isaac Newton (1643-1727) and Gottfried Wilhelm Leibniz (1646-1710).

The difference quotient becomes the differential in the limit $h \to 0$ and describes the slope of a function $y = f(x)$ at a given point x.

$$y'(x) = \frac{dy}{dx} = \lim_{h\to 0} \frac{y(x+h) - y(x)}{h}$$

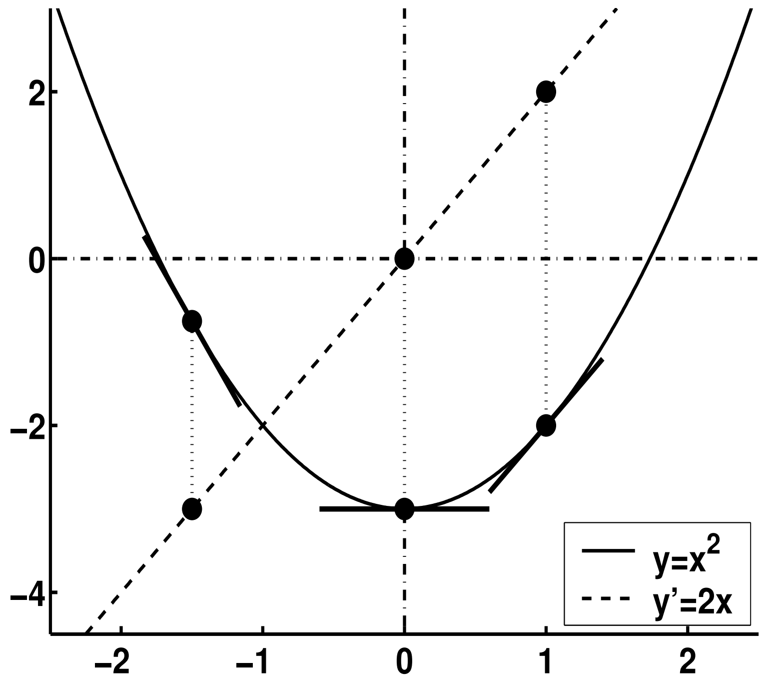


Figure 19: The slope of a curve is found from its derivative.

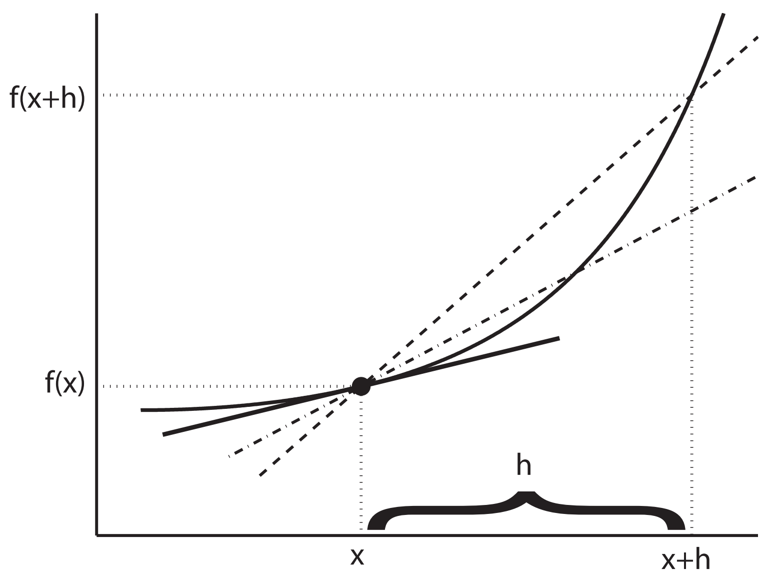


Figure 20: Slope as h→ 0.

Notation: The limit value of the difference quotient is called the derivative of a function $f(x)$.

Derivatives are denoted by

$$y'(x)\ ,\ \frac{dy}{dx}\ ,\ \frac{df}{dx}\ ,\ \frac{d}{dx}f(x) \qquad \text{or sometimes in physics: } \dot{y}(t)$$

Note: Here we consider first-order derivatives only.

Example: $y = f(x) = x^2$

$$y' \;=\; \frac{dy}{dx} = \lim_{h\to 0}\frac{(x+h)^2 - x^2}{h} \;=\; \lim_{h\to 0}\frac{x^2 + 2hx + h^2 - x^2}{h} \;=\; \lim_{h\to 0}\frac{2hx + h^2}{h} = 2x$$

2.5 Derivatives of Elementary Functions

2.5.1 Polynomials

$$y = x^2 \quad \to \quad \frac{dy}{dx} = 2x \qquad\qquad \text{more general:} \quad y = x^n \quad \to \quad \frac{dy}{dx} = nx^{n-1}$$

2.5.2 Trigonometric functions

$$y = \sin x \quad \to \quad \frac{dy}{dx} = \cos x \qquad\qquad y = \cos x \quad \to \quad \frac{dy}{dx} = -\sin x$$

2.5.3 Exponential functions

$$y = e^x \quad \to \quad \frac{dy}{dx} = e^x$$

2.5.4 Hyperbolic functions

$$y = \sinh x \quad \to \quad \frac{dy}{dx} = \cosh x \qquad\qquad y = \cosh x \quad \to \quad \frac{dy}{dx} = \sinh x$$

2.5.5 Logarithms

$$y = \ln x \quad \to \quad \frac{dy}{dx} = \frac{1}{x}$$

2.6 The Basic Rules for Calculating Derivatives

If the derivatives of two functions $u(x)$ and $v(x)$ exist on an interval $a < x < b$, then the derivatives of their combinations exist as well, i.e.

$$u + v, \quad \alpha\, u \;\text{ with }\; \alpha \in \mathbb{R}, \quad u\, v, \quad \frac{u}{v} \;\text{ if }\; v(x) \neq 0 \;\text{ for }\; a < x < b$$

Rules:

$$(u+v)' = \tfrac{d}{dx}\{u+v\} = u'+v' \qquad \textbf{derivatives are additive}$$

$$(\alpha\, u)' = \tfrac{d}{dx}\{\alpha\, u\} = \alpha\, u' \qquad \textbf{multiplication with a scalar}$$

$$(u\, v)' = \tfrac{d}{dx}\{u\, v\} = u'\, v + u\, v' \qquad \textbf{product rule}$$

$$(\tfrac{u}{v})' = \tfrac{d}{dx}\{\tfrac{u}{v}\} = \tfrac{u'\, v - u\, v'}{v^2} \qquad \textbf{quotient rule}$$

Examples:

$$\frac{d}{dx}\{x^{17} + \cos x\} = 17x^{16} - \sin x$$

$$\frac{d}{dx}\{35\cosh x\} = 35\frac{d}{dx}\,\cosh x = 35\sinh x$$

$$\frac{d}{dx}\{\cos x e^x\} = -\sin x e^x + \cos x e^x = e^x(\cos x - \sin x)$$

$$\frac{d}{dx}\{\frac{\cos x}{e^x}\} = \frac{-\sin x\, e^x - \cos x\, e^x}{e^{2x}} = \frac{-e^x\,(\sin x + \cos x)}{e^{2x}} = -\frac{\sin x + \cos x}{e^x}$$

2.7 The Chain Rule

If $u(x)$ and $v(x)$ have derivatives and the image of $v(x)$ is part of the source set of $u(x)$, then $u(v(x))$ has a derivative.

To understand what this complicated sentence means, consider $\ln(\cos x)$. Here $u(x) = \ln x$ and $v(x) = \cos x$. The source set of $\cos x$ are all real numbers $[-\infty, \infty]$, the image set of the cosine are the numbers in the interval [-1, 1], and the source set of the logarithm are all positive real numbers $]0, \infty]$. Therefore the image set of the cosine and the source set of the logarithm overlap in the interval $]0, 1]$. The source set of $\cos x$ that corresponds to the image set $]0, 1]$ is given by all numbers where $\cos x$ is positive, i.e. $]-\frac{\pi}{2}, -\frac{\pi}{2}[$, $]\frac{3\pi}{2}, -\frac{5\pi}{2}[$, etc., and the function $\ln(\cos x)$ exists and has a derivative for these values of x.

$$[u(v(x))]' = \frac{d}{dx}\{u(v(x))\} = \frac{d\,u(v)}{dv}\,\frac{d\,v(x)}{dx} \qquad \textbf{chain rule}$$

Examples:

$$f(x) = \cos(\alpha x) \quad \rightarrow \quad u(v) = \cos v \quad \text{and} \quad v(x) = \alpha x$$

$$\frac{d}{dx}\,\cos\alpha x = \frac{d\,\cos\alpha x}{d\,\alpha x}\,\frac{d\,\alpha x}{dx} = (-\sin\alpha x)\,\alpha = -\alpha\sin\alpha x$$

$$f(x) = (2x+5)^3 \quad \rightarrow \quad u(v) = v^3 \quad \text{and} \quad v(x) = 2x+5$$

$$\frac{d}{dx}\,(2x+5)^3 = \frac{d\,(2x+5)^3}{d\,(2x+5)}\,\frac{d\,(2x+5)}{dx} = 3\,(2x+5)^2\;2 = 6\,(2x+5)^2$$

2.8 Selected problems (the page from hell):

Important note: Now we can take the derivative of ANY analytic function !!!

$$f(x) = e^{\ln x} \quad \rightarrow \quad u(v) = e^{v} \quad \text{and} \quad v(x) = \ln x$$

$$f'(x) = e^{\ln x}\,\frac{1}{x} = x\,\frac{1}{x} = 1 \qquad \text{of course we started with} \quad f(x) = x \ \rightarrow \ f'(x) = 1$$

$$f(x) = \sqrt{\sin(3\,\alpha^2\,x^5)} = [\sin(3\,\alpha^2\,x^5)]^{\frac{1}{2}} = u(v(w(x)))$$

$$\rightarrow \quad u(v) = v^{\frac{1}{2}} \quad v(w) = \sin(3\,\alpha^2\,x^5) \quad w(x) = 3\,\alpha^2\,x^5$$

$$f'(x) = \frac{d\,u(v(w(x)))}{dv}\,\frac{d\,v(w(x))}{dw}\,\frac{d\,w(x)}{dx} = \frac{1}{2}\,[\sin(3\,\alpha^2\,x^5)]^{\frac{1}{2}-1}\ \cos(3\,\alpha^2\,x^5)\ 3\,\alpha\,5\,x^{5-1}$$

$$= \frac{15\,\alpha\,x^4\ \cos(3\alpha^2x^5)}{2\sqrt{\sin(3\alpha^2x^5)}} \qquad \text{who guessed this result ???}$$

$$f(x) = \frac{3x^2 + \cos kx}{\cosh x} \quad \rightarrow \quad f'(x) = \frac{(6x - k\sin kx)\cosh x + (3x^2 + \cos kx)\sinh x}{\cosh^2 x}$$

Also quite ugly, but technically correct !!!

$$f(x) = \cos^2 kx = \cos kx\ \cos kx \quad \rightarrow \quad f'(x) = 2\ \cos kx(-\sin kx)\,k = -2k\ \cos kx \sin kx$$

$$\text{or} \quad \rightarrow \quad (-\sin kx)\,k\ \cos kx + \cos kx(-\sin kx)\,k = -2k\ \cos kx \sin kx$$

$$f(x) = y = (x^5 + e^{\cos kx})^{1/2} \quad \rightarrow \quad y' = \frac{1}{2}(x^5 + e^{\cos kx})^{-1/2}\,(5\,x^4 + e^{\cos kx}(-k\ \sin kx))$$

$$= \frac{5\,x^4 - k\ \sin 2kx\,e^{\cos kx}}{2\,(x^5 + e^{\cos kx})^{1/2}}$$

$$y = x^x = e^{x\ln x} \quad (\text{remember: } a^x = e^{x\ln a}) \quad \rightarrow \quad y' = e^{x\ln x}\,(\ln x + 1) = x^x(\ln x + 1)$$

2.9 Applications of derivatives

2.9.1 Maxima, minima, inflection points and convexity

One of the most useful applications of derivatives is that they allow us to find properties of functions such as maxima, minima and inflection points. Given a differentiable function f, the procedure to find these points would be:

1. Find the first derivative $f'(x)$.
2. Find the second derivative $f''(x)$
3. Solve $f'(x) = 0$ and find $X = \{x_1, x_2, \ldots, x_N / f'(x) = 0,\ \forall i = 1, \ldots, N\}$
4. Substitute the values of X into $f''(x)$ and:
   * if $f''(x_i) < 0$ we have a maxima at x_i.
   * if $f''(x_i) < 0$ we have a minima at x_i.
   * if $f''(x_i) = 0$ continue computing derivatives unitil $f^{n)}(x_i) \neq 0$.
     i. If n is even and $f^{n)}(x_i) < 0$ we have a *local* maxima, and if $f^{n)}(x_i) > 0$ we have a *local* minima .
     ii. If n is odd we get an inflection point.
5. Solve $f''(x) = 0$ and find $X = \{x_1, x_2, \ldots, x_N / f''(x) = 0,\ \forall i = 1, \ldots, N\}$. This is a necessary condition for being an inflection point. Sufficient condition will be that $f^{3)}(x) \neq 0$.

Note: These are local minima and maxima at x_0, to be global we further require that:

$$f(x) \leq f(x_0), \quad \forall x \text{ (maxima)}$$

$$f(x) \geq f(x_0), \quad \forall x \text{ (minima)}$$

Similarly, we can find conditions for the concavity of functions. Given a function f and an interval (a, b), we say that the function is concave (up or down) in the interval if:

$$f'(a) \leq f'(b) \Rightarrow f''(x) > 0, \text{ (concave up)}$$

$$f'(a) \geq f'(b) \Rightarrow f''(x) < 0, \text{ (concave down)}$$

Example: Find and analyze the critical points of the function$f(x) = x^3 - 12x^2 + 45x - 30$.

First, we solve $f'(x) = 0$ and $f''(x) = 0$:

$$f'(x) = 3x^2 - 24x + 45 = 3(x^2 - 8x + 15) = 3(x-3)(x-5) = 0 \quad \rightarrow x_1 = 3 \text{ and } x_2 = 5.$$

$$f''(x) = 6x - 24 \quad \rightarrow x_3 = 4.$$

We evaluate the three points we obtained. From the first derivatives, we found two points that we evaluate in the second derivative: $f''(x_1) = f''(3) = -6$ (local maxima), and $f''(x_2) = f''(5) = 6$ (local minima). Then, we evaluate the point found in the second derivative at the third derivative: $f^{3)}(x_3) = f^{3)}(4) = 6$, which means that we have an inflection point. Try now to plot the result!

2.9.2 L'Hôpital theorem

This theorem will allow us to solve some limits that led to an indeterminate form. **Theorem:** If f and g are two functions derivable in a neighborhood of $a \in \mathbb{R}$ (meaning that it works even if there is a hole) and if $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\infty}{\infty}$ or $\frac{0}{0}$ (and a could be ∞):

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = \lim_{x\to a} \frac{f(x)}{g(x)},$$

provided that $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists.

Example: To compute the following limit, we applied L'Hôpital twice:

$$\lim_{x\to 1} \left(\frac{x}{x-1} - \frac{1}{\log x} \right) = \lim_{x\to 1} \frac{x\log x - x + 1}{(x-1)\log x} =$$
$$\underbrace{=}_{\text{L'Hôpital}} \lim_{x\to 1} \frac{\log x + 1 - 1}{\log x + \frac{x-1}{x}} \underbrace{=}_{\text{L'Hôpital}} \lim_{x\to 1} \frac{1/x}{1/x + \frac{x-x+1}{x^2}} = \frac{1}{2}.$$

...smooth!

2.9.3 A myriad of theorems

There are many theorems involving derivatives more or less evident, and whose application is more or less powerful as well. Let's see just one example, the Mean Value Theorem due to Lagrange.

Theorem: Given a function f which is continuous in $[a, b]$ and differentiable in (a, b), then $\exists x_0 \in (a, b)$ such that:

$$f'(x_0) = \frac{f(b) - f(a)}{b - a} = \tan\alpha.$$

Plot it and convince yourself that this is true!

2.9.4 Taylor Series

A Taylor series is an approximation of a function $f(x)$ in the neighborhood of a given point in terms of its derivatives. This technique has been developed by Brook Taylor$(1685 - 1731)$ and first published in 1715.

In a first step we can approximate the function $f(x)$ in the neighborhood around x_0 by the tangent through x_0

$$f(x) = f(x_0) + f'(x_0)\,(x - x_0) + \text{error}$$

Can we do better than this? Yes if higher order derivatives are considered!

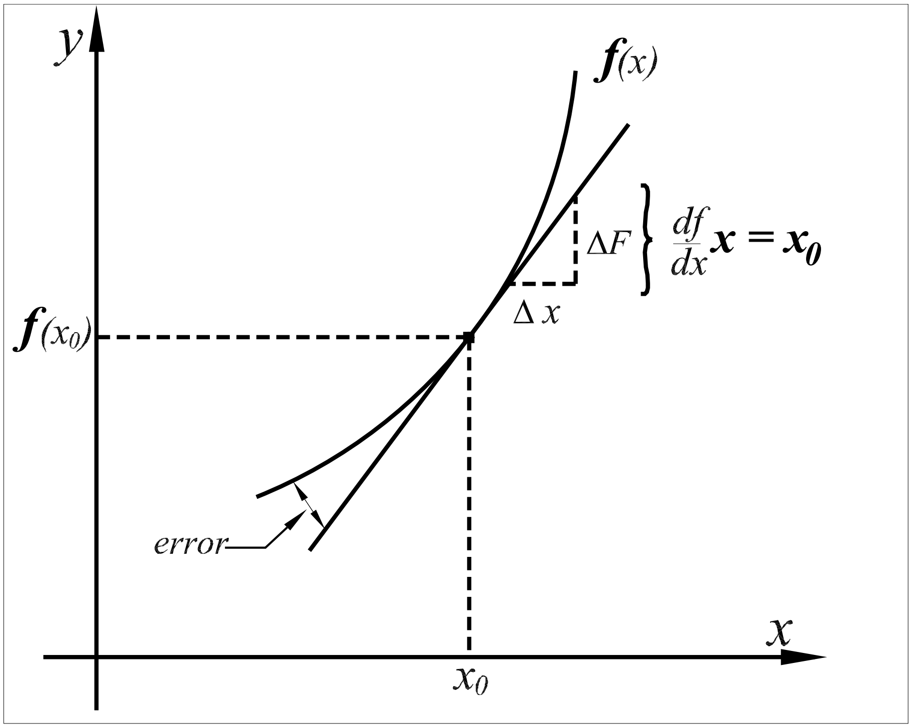


Figure 21: Approximation of a curve at a point x_0 with the tangent at x_0

$$f(x) = f(x_0) + f'(x_0)\,(x - x_0) + \frac{1}{2!}\,f''(x_0)(x - x_0)^2 + \frac{1}{3!}\,f'''(x_0)(x - x_0)^3 + ...$$

Taylor series: $$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!}\,\frac{d^n f(x)}{dx^n}\bigg|_{x=x_0} (x - x_0)^n \qquad \text{with} \quad n! = 1 \cdot 2 \cdot 3 \cdot ... \quad \text{n-factorial}$$

A function $f(x)$ may be approximated by truncating a Taylor Expansion around x_0 at the m-th order.

⇒ Polynomial representations of function work well if the error approaches 0 as the order n increases.

Error estimate:

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!}\,\frac{d^n f(x)}{dx^n}\bigg|_{x=x_0} (x - x_0)^k + \underbrace{R_n(x)}_{\text{error}}$$

Lagrange formulation of the error R_n in a Taylor expansion that is truncated at order n

$$R_n(x) = \frac{(x - x_0)^{n+1}}{(n+1)!}\,\frac{d^n f(x)}{dx^n}\bigg|_{x=\xi} \qquad \xi = x_0 + \delta\,(x - x_0) \quad 0 < \delta < 1$$

Examples:

Approximate the function $\sin x$ up to the 7-th order for x around $x_0 = 0$ (Maclaurin series) using a Taylor expansion

$$\sin x \approx \underbrace{\sin 0}_{0} + \underbrace{\cos 0 \cdot x}_{x} + \underbrace{-\sin 0 \frac{1}{2!}\, x^2}_{0} - \frac{1}{3!}\, x^3 + \frac{1}{5!}\, x^5 - \frac{1}{7!}\, x^7 \qquad \text{symmetries!!}$$

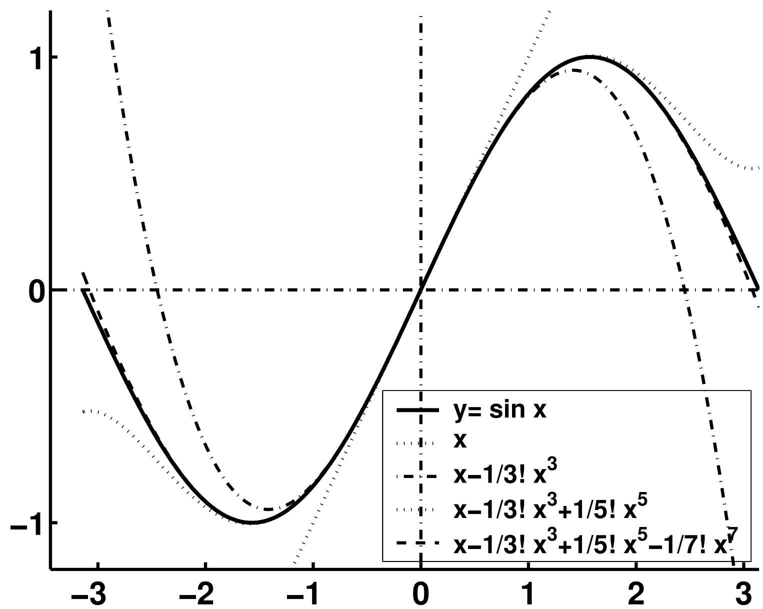


Figure 22: Steps of a Taylor expansion of $\sin x$ around $x_0 = 0$

The further away you move from the expansion point, the more significant the higher order terms!!

Specific expansion of important functions:

$$\sin x = x - \frac{1}{3!}\, x^3 + \frac{1}{5!}\, x^5 - \;....\; \frac{(-1)^{n+1}}{(2\,n+1)!}\, x^{2\,n+1} \qquad \text{only odd terms}$$

$$\cos x = 1 - \frac{1}{2!}\, x^2 + \frac{1}{4!}\, x^4 - \;....\; \frac{(-1)^n}{2\,n!}\, x^{2\,n} \qquad \text{only even terms}$$

$$e^x = 1 + \frac{1}{2!}\, x^2 + \frac{1}{3!}\, x^3 + \;....\; \frac{1}{n!}\, x^n$$

$$\ln(1+x) = x - \frac{x^2}{2!} + \frac{x^3}{3!} + \;....\; \frac{(-1)^{n+1}}{n!}\, x^{n+1}$$

Euler's Formula: $e^{i\theta} = \cos\theta + i\,\sin\theta$

$$e^{i\,\theta} = 1 + i\theta + \frac{1}{2!}\underbrace{(i\,\theta)^2}_{-\theta^2} + \frac{1}{3!}\underbrace{(i\,\theta)^3}_{-i\,\theta^3} + \frac{1}{4!}\underbrace{(i\,\theta)^4}_{\theta^4} + \frac{1}{5!}\underbrace{(i\theta)^5}_{i\,\theta^5} + ... \qquad \text{Taylor expansion around } \theta = 0$$

$$\Rightarrow \qquad e^{i\,\theta} = \underbrace{1 - \frac{1}{2!}\,\theta^2 + \frac{1}{4!}\,\theta^4 + ...}_{\cos\theta} \quad + i\,(\underbrace{\theta - \frac{1}{3!}\,\theta^3 + \frac{1}{5!}\,\theta^5 + ...}_{\sin\theta}) \qquad \text{q.e.d.}$$

3 Integrals

3.1 Integral Calculus: Definite Integrals

How do you determine the area A enclosed by a function $f(x)$ and the horizontal axis. It is simple if $f(x) = f_0 = const.$

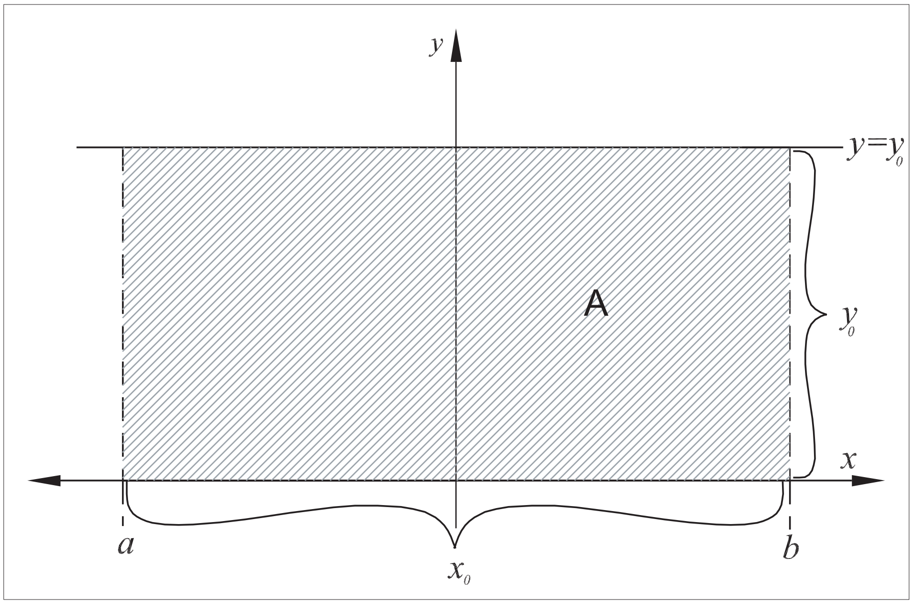


Figure 23: Area A enclosed by the horizontal axis and a horizontal line.

For the general case, divide the area A into subareas A_ν between $x_{\nu-1}$ and x_ν.

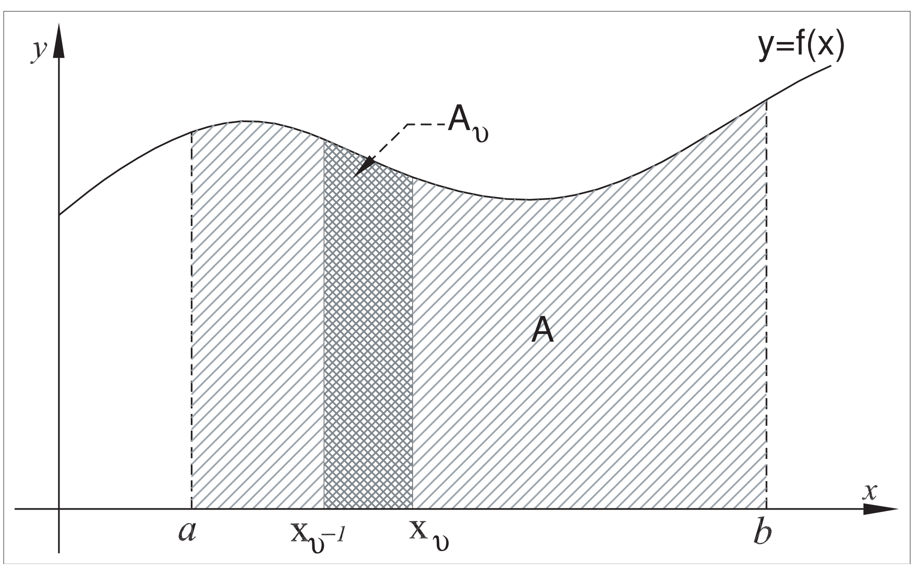


Figure 24: Area enclosed by the function $f(x)$ and the horizontal axis.

Then the subarea A_ν may be approximated by $A_\nu = f(\xi_\nu)(x_\nu - x_{\nu-1})$ for $x_{\nu-1} < \xi_\nu < x_\nu$. There exists always a ξ_ν such that this is true.

Reconstruct the area A as follows:

$$A = \sum_{\nu=1}^{N} = A_\nu = \sum_{\nu=1}^{N} f(\xi_\nu) \underbrace{(x_\nu - x_{\nu-1})}_{\Delta x}$$

This sum is called the *Riemann sum*. For instance, for $N = 3$ the sum becomes

$$A = f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + f(\xi_3)(x_3 - x_2),$$

but, if we take few terms, our approximation to the area will be poor. However, we know already how to take limits, so we can take the limit of the Riemann sum, and hence the area A will be computed precisely, and it will allow us to define the integral.

$$A = \lim_{N\to\infty} \sum_{\nu=1}^{N} f(\xi_\nu)\Delta x = \int_{x=a}^{x=b} f(x)\, dx$$

The area enclosed by $f(x)$ and the horizontal x-axis over an interval $x \in [a, b]$ is given by definite integral

$$\int_a^b f(x)\, dx = F(x) \Big|_{x=a}^{x=b} = F(x) \Big|_a^b = F(b) - F(a)$$

where $F(x)$ is called the *anti-derivative* of $f(x)$ and

$$f(x) = \frac{dF(x)}{dx} = F'(x) \qquad \text{or, which is equivalent,} \qquad F(x) = \int f(x)\, dx + \text{const}$$

Integration is to some extend the inverse operation of differentiation.

Example: Try to guess the area under $f(x) = x^2$ within the interval $[-1, 1]$.

a) Looking at Fig. 25, the shaded area is given by

$$A = \int_{-1}^{1} f(x)\, dx = F(1) - F(-1)$$

In order to guess a solution, remember that we know that $f(x) = x^2 = F'(x)$. Therefore, we guess $F(x) = 1/3x^3 + c$, and the are becomes:

$$A = F(1) - F(-1) = \frac{1}{3}1^3 + c - \{\frac{1}{3}(-1)^3 + c\} = \frac{2}{3}$$

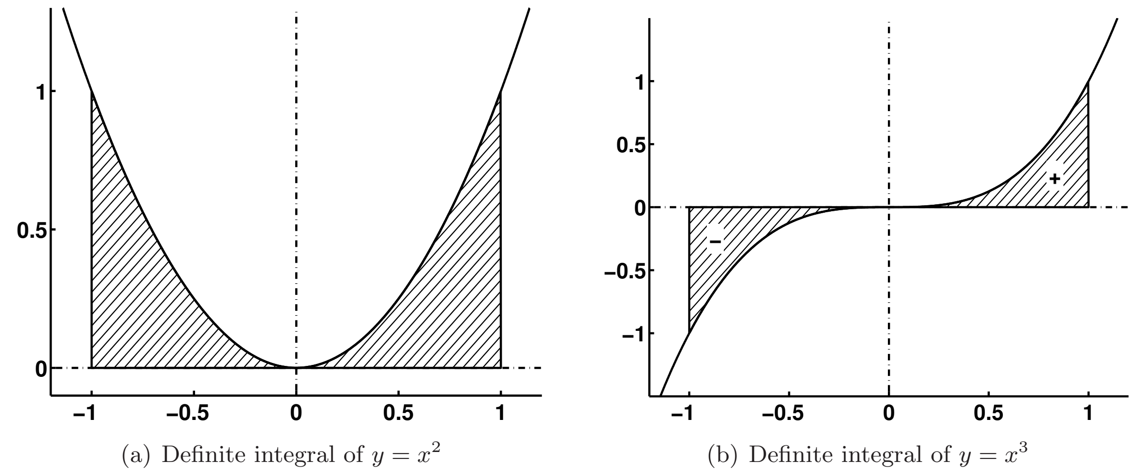


Figure 25: Definite integrals as areas under curves

Example: Try to guess the area under $f(x) = x^3$ within the interval $[-1, 1]$.

b) Again, the shaded area shown in Fig. 25 is given by

$$A = \int_{-1}^{1} f(x)\,dx = F(1) - F(-1)$$

We guess $F(x) = \frac{1}{4}x^4 + c$ and find

$$A = \int_{-1}^{1} f(x)\,dx = F(1) - F(-1) = \frac{1}{4}1^4 + c - \{\frac{1}{4}(-1)^4 + c\} = 0$$

Why does the area A vanish? It actually consists of two areas, A_1 and A_2, which both have the same size, but opposite sign $A_1 = -A_2$.

$$A_1 = F(0) - F(-1) = \frac{1}{4}0^4 + c - \frac{1}{4}(-1)^4 - c = -\frac{1}{4} = -A_2$$

Note: In an integral the area *below* the x-axis is counted negative. In order to calculate the shaded area we have to evaluate all pieces between intersections of the curve with the horizontal axis separately and add up their magnitudes. Here: $A = \mid A_1 \mid + \mid A_2 \mid = \mid -\frac{1}{4} \mid + \mid \frac{1}{4} \mid = \frac{1}{2}$.

3.2 Indefinite integral

We say that $F(x)$ is the primitive function or antiderivative if $F'(x) = f(x)$ for all the domain of f.

3.2.1 Polynomials

$$\int x^p dx = \frac{x^{p+1}}{p+1} + C \quad \text{for } p \neq -1$$

3.2.2 Logarithms

$$\int \frac{1}{x} dx = \log|x| + C$$

3.2.3 Exponential functions

$$\int e^x dx = e^x + C; \quad \int f'(x) e^f(x) dx = e^f(x) + C$$

3.2.4 Trigonometric functions

$$\int \sin x dx = -\cos x + C$$

$$\int \cos x dx = \sin x + C$$

$$\int \sec^2 x dx = \int \frac{dx}{\cos^2 x} = \int (1 + \tan^2 x) dx = \tan x + C$$

$$\int \csc^2 x dx = \int \frac{dx}{\sin^2 x} = \int (1 + \cot^2 x) dx = -\cot x + C$$

The following three are particularly important:

$$\int \frac{dx}{1+x^2} = \arctan x + C$$

$$\int -\frac{dx}{\sqrt{1-x^2}} = \arccos x + C$$

$$\int \frac{dx}{\sqrt{1+x^2}} = \arcsin x + C$$

Note: How did you end up with these last expressions? Let's see one example:

$$y = \arcsin x \quad \rightarrow \sin y = x \quad \underbrace{\rightarrow}_{\text{Differentiate}} \quad y' \cos y = 1$$

$$y' = \frac{1}{\cos y} \underbrace{=}_{\sin^2 y + \cos^2 y = 1} \frac{1}{\sqrt{1 - \sin^2 y}} = \frac{1}{\sqrt{1-x^2}}.$$

3.2.5 Hyperbolic functions

$$\int \sinh x dx = \cosh x + C$$

$$\int \cosh x dx = \sinh x + C$$

3.3 Properties of integrals

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

$$\int_a^b f(x)\,dx = F(b) - F(a) = -(F(a) - F(b)) = -\int_b^a f(x)\,dx$$

$$\int_a^b (f_1(x) + f_2(x))\,dx = \int_a^b f_1(x)\,dx + \int_a^b f(x)_2\,dx$$

$$\int_a^b c\,f(x)\,dx = c\int_a^b f(x)\,dx \qquad \text{where } c \text{ is a constant}$$

3.4 Methods of Integration

3.4.1 Guess and tables

Find $F(x)$ such that $\frac{dF(x)}{dx} = f(x)$. For polynomials: $F(a\,x^n) = \frac{a}{n+1}\,x^{n+1}$.

Note: Here n can be negative or any rational number except -1.
$F(x)$ may be also looked up in mathematical tables of anti-derivatives and/or definite integrals.

3.4.2 Partial integration

Corresponds to the product rule but only works for special cases.

$$\int \underbrace{g(x)}_{u}\ \underbrace{f(x)dx}_{dv} = \underbrace{g(x)}_{u}\ \underbrace{F(x)}_{v} - \int \underbrace{F(x)}_{v}\ \underbrace{g'(x)\,dx}_{du}$$

$$\int udv = uv - \int vdu$$

Memo: Susan **u**ses a **d**e**v**ice **v**ery **u**nique **s**ince **v**isualizes the **d**eep **u**niverse.

Typically, we will have the products of two functions, one of which is easy to differentiate (logarithms, polynomials, arccos-like functions) that will be u, and another one that is easy to integrate such as polynomials or sin-like functions, will be v.

Example: Solve the definite integral.

$$I = \int_a^b \underbrace{x}_{u} \underbrace{\cosh x \, dx}_{dv}$$

First of all, we find v and du, respectively:

$$x = u \quad \rightarrow dx/dx = 1 = du$$
$$dv = \cosh x \, dx \quad \rightarrow v = \int \cosh x \, dx = \sinh x.$$

Now we substitute using the memo, be aware that the integral is definite!

$$I = x \sinh x|_a^b - \int_a^b \sinh x dx =$$
$$= b \sinh b - a \sinh a - (\cosh b - \cosh a).$$

Example: Solve the indefinite integral

$$I = \int \operatorname{arccot} x dx$$

Substitution: Corresponds to the chain rule but again only works for special cases.

$$\int_{x=a}^{x=b} f(\phi(x)) \, \phi'(x) \, dx = \int_{u=\phi(a)}^{u=\phi(b)} f(u) \, du \qquad \text{where} \qquad u = \phi(x)$$

$$\int_0^\pi \cos^2 x \; \sin x \, dx \quad \text{substitute:} \quad u = \cos x = \phi(x)$$

$$u' = \frac{du}{dx} = -\sin x = \phi'(x) \quad \rightarrow \quad du = -\sin x dx = \phi'(x) dx \quad \rightarrow \quad dx = -\frac{du}{\sin x}$$

Substitute the integral:

$$\int_{x=0}^{x=\pi} \cos^2 x \; \sin x \frac{-du}{\sin x} = -\int_{x=0}^{x=\pi} \cos^2 x \, du = -\int_{x=0}^{x=\pi} u^2 \, du$$

Express the boundaries in terms of u:

$$x = 0 \quad \rightarrow \quad u = \cos 0 = 1 \qquad\qquad x = \pi \quad \rightarrow \quad u = \cos \pi = -1$$

Insert them and perform the integration:

$$\int_0^{\pi} \cos^2 x \ \sin x \, dx = -\int_{u=1}^{u=-1} u^2 \, du = -\frac{1}{3} u^3 \Big|_1^{-1} = -\frac{1}{3}(-1)^3 + \frac{1}{3} 1^3 = \frac{2}{3}$$

3.5 Symmetries

A function $f(x)$ is called an *even function* if $f(-x) = f(x)$; a function $g(x)$ is called an *odd function* if $g(-x) = -g(x)$. The product of two even functions or the product of two odd functions is an even function; the product of an odd and an even function is an odd function.

The integral over a symmetric interval around $x = 0$ of an odd function vanishes.

$$\int_{-a}^{b=a} g(x) \, dx = \int_{-a}^{b=a} f(x) \, g(x) \, dx = 0 \qquad \text{if } \ f(-x) = f(x) \quad \text{and} \quad g(-x) = -g(x)$$

Example:

$$\int_{-1}^{1} \underbrace{x^2}_{f(x)} \underbrace{\sin 3x}_{g(x)} \, dx = 0$$

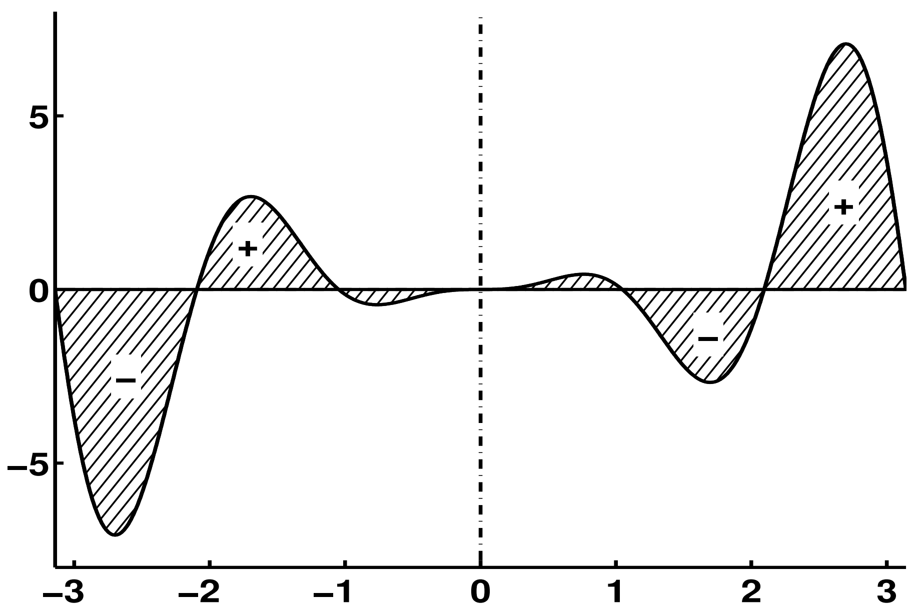


Figure 26: Due to symmetry the integral $\int_{-1}^{1} x^2 \, \sin 3x \, dx$ vanishes

3.6 Orthogonality of trigonometric functions

The cosine is an even function $\cos(-x) = \cos x$, and the sine is an odd function $\sin(-x) = -\sin x$. Moreover, these trigonometric functions are 2π-periodic, hence it is sufficient to consider integration over windows of 2π only.

$$\int_0^{2\pi} \underbrace{\sin x\,\cos x}_{\frac{1}{2}\sin 2x}\,dx = \frac{1}{2}\int_0^{2\pi} \sin 2x\,dx = -1/4\cos 2x\,\Big|_0^{2\pi} = 0 \quad \text{or equivalent:} \quad \int_{-\pi}^{\pi} \sin x\,\cos x\,dx = 0$$

$$\int_{-\pi}^{\pi} \sin 2x\,\sin x\,dx = 2\int_{-\pi}^{\pi} \underbrace{\sin^2 x}_{u^2}\,\underbrace{\cos x\,dx}_{du} = 2\int_{u=0}^{u=0} u^2\,du = 0$$

Here we used the substitution $u = \sin x$ and $du = \cos x\,dx$ with the boundaries $x = \pi \to u = 0$ and $x = -\pi \to u = 0$.

More general cases:

$$\int_{-\pi}^{\pi} \cos mx\,\cos nx\,dx = \int_{-\pi}^{\pi} \sin mx\,\sin nx\,dx = \pi\,\delta_{mn}$$

$$\int_{-\pi}^{\pi} \cos mx \sin nx\,dx = 0\;\;\forall\;\; m,\,n \qquad \text{where "}\forall\text{" means "for all"}$$

δ_{mn} is called the *Kronecker delta* which is defined as $\delta_{mn} = 1$ if $m = n$ and $\delta_{mn} = 0$ else.

3.7 Integrals to Infinity

If one or both boundaries of an integral are infinite this does not mean that the area under this curve cannot be finite. A trivial example is given by the integral from $-\infty$ to $+\infty$ over an odd function. This integral vanishes, as seen above, independent of the function as long as it is odd.

A nontrivial example is

$$\int_1^{\infty} \frac{1}{x^2}\,dx = -\,\frac{-1}{x}\Big|_1^{\infty} = -(\underbrace{\frac{1}{\infty}}_{=0} - \frac{1}{1}) = 1$$

however

$$\int_1^{\infty} \frac{1}{x}\,dx = \ln x\,\Big|_1^{\infty} = \ln\infty - \underbrace{\ln 1}_{=0} = \infty$$

In the same way even if a function has a singularity like $\frac{1}{\sqrt{x}}$ for $x \to 0$, the area can still be finite

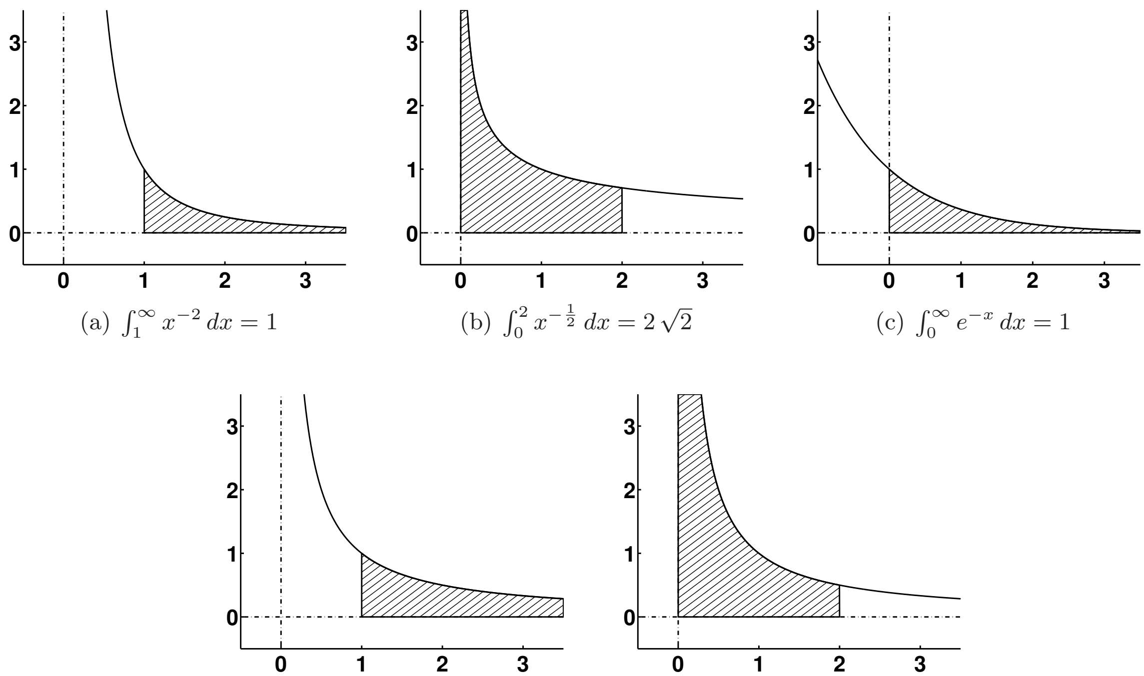
$$\int_0^2 x^{-\frac{1}{2}}\,dx = 2\,x^{\frac{1}{2}}\,\Big|_0^2 = 2\sqrt{x}\,\Big|_0^2 = 2\,(\sqrt{2} - \sqrt{0}) = 2\sqrt{2}$$

but again

$$\int_0^2 x^{-1}\,dx = \int_0^2 \frac{1}{x}\,dx = \ln x\,\Big|_0^2 = \ln 2 - \underbrace{\ln 0}_{=-\infty} = \infty$$

And finally an exponential function

$$\int_0^{\infty} e^{-x}\,dx = -e^{-x}\,\Big|_0^{\infty} = -(\underbrace{e^{-\infty}}_{=0} - \underbrace{e^{0}}_{=1}) = 1$$



(a) $\int_1^\infty x^{-2}\,dx = 1$ (b) $\int_0^2 x^{-\frac{1}{2}}\,dx = 2\sqrt{2}$ (c) $\int_0^\infty e^{-x}\,dx = 1$

(d) $\int_1^\infty x^{-1}\,dx = \infty$ (e) $\int_0^2 x^{-1}\,dx = \infty$

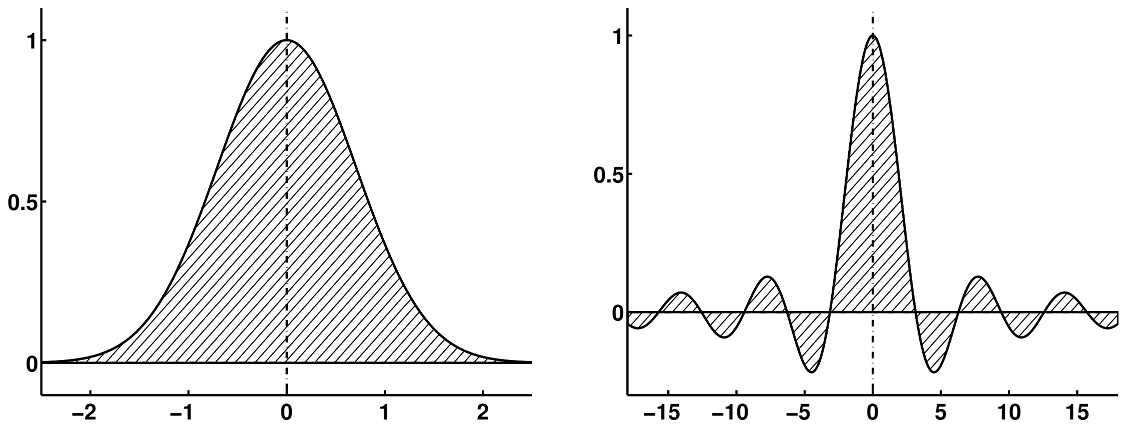
Figure 27: Definite integrals that involve infinities

3.8 Functions with no Antiderivative

As we have seen, it quite straightforward to calculate the derivatives of quite complicated "monsters" of functions. On the other hand, it is much more difficult to find antiderivatives. To make things worse there are certain functions with very important applications for which an antiderivative does not exist, i.e. it cannot be expressed in terms of elementary functions.

One of these simple functions which do not have an antiderivative is $f(x) = e^{-x^2}$. This is very inconvenient because this function is the famous bell-shaped Gaussian which rules the entire field of statistics, because the probability that an event occurs within a certain interval of a parameter is given by area under this curve. This area unfortunately cannot be calculated using a pocket calculator that has only elementary functions. The definite integral can be found numerically or looked up in tables, and it also has a name: the "error function" erf(x)

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\,du \qquad \text{note:} \qquad \int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$$



(a) $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$ (b) $\int_{-\infty}^{\infty} \frac{\sin x}{x}\,dx = \pi$

Figure 28: Definite integrals over functions with no antiderivatives

A second example of such a function with no antiderivative is the so-called "integral sine" $\mathrm{Si}(x)$

$$\mathrm{Si}(x) = \int_0^x \frac{\sin u}{u}\,du \qquad \text{note:} \qquad \int_{-\infty}^{\infty} \frac{\sin x}{x}\,dx = \pi$$

4 Differential equations

In many relevant problems, we aim to determine a function satisfying an equation containing one or more derivatives of the unknown function.

Example: Consider a function $y = y(t)$ such that verifies the differential equation:

$$\frac{dy}{dt} = -Ky \qquad \text{where } K \text{ is a constant.}$$

We aim to find the explicit expression of the function $y(t)$. We can rearrange this equation separating terms containing the variable y from those containing the variable t:

$$\frac{dy}{y} = -Kdt,$$

and now we integrate both sides

$$\int \frac{dy}{y} = \int -Kdt \quad \rightarrow \quad \ln|y| + C_1 = -Kt + C_2$$

where C_i are constants that we can merge into a single one, i.e. $C = C_1 + C_2$. Solving for y we obtain the explicit expression $y(t)$ that we look for:

$$y = e^{-Kt+C} = Ce^{-Kt}$$

4.1 General classes of differential equations – and which ones we will NOT study

To solve differential equations we should first recognize different properties in the equations, that allow us to classify them into types. This is important because there are different methods for each particular type of equation, so many times the difficulty relies in the identification of the equation and, eventually, in the transformation of the equation into one type we know how to solve. In this course, we will focus on *first-order ordinary differential equations.* To see what does this mean we will see here the type of equations that we will NOT study.

a) Ordinary differential equation (ODE) vs. partial differential equation (PDE). If the unknown function only depends on one independent variable we are dealing with an ODE, otherwise it is a PDE. We will study ODEs only.

Example: The difussion equation is a PDE. If the diffusion constant D is independent of time (t) and space ($\overrightarrow{r}$), the one-dimensional ($\overrightarrow{r} = x$) equation is:

$$D\frac{\partial^2 y(x,t)}{\partial x^2} = \frac{\partial y(x,t)}{\partial t}$$

b) Order of a differential equation. The order of a differential equation is given by the order of the highest derivative found in the equation, i.e. the highest number of times that the function y is differentiated in any term of the equation. In this course we will focus on first-order ODEs. Many times in natural systems we only need to consider first and second order differential equations to describe processes.

Example: The Newton's law is a second-order differential equation.

$$m\frac{d\overrightarrow{r}(t)}{dt^2} = F(t, \overrightarrow{r}(t), d\overrightarrow{r}(t)/dt)$$

c) Systems of differential equations In the same way that we studied systems of linear equations, we may have systems of differential equations, where there is one equation for each unknown function.

Example: Consider the population dynamics of some plant, $P = P(t)$, and herbivore species, $H = H(t)$:

$$\frac{dP}{dt} = aP - \beta HP$$
$$\frac{dH}{dt} = -bH + \gamma PH.$$

Where the parameters a (b) describe the intrinsic growth (death) rates of plants (herbivores), and the β (γ) parameters the rate of decrease (increase) in the populations due to the prey-predator interaction. These equations are also known as Lotka-Volterra equations. In this course, we will not study systems of differential equations that will be studied in the module of Ecological Modelling.

4.2 Differential equations that we WILL study

In this course, we would like to solve 1st-order ODEs that can be written in general in this form:

$$N(x, y)y' + M(x, y) = Q(x).$$

There are four main types of analytically solvable 1st-order ODEs, that we will define from the above equation. The four types are: separable, homogeneous, exact and linear ODEs. We will learn how to solve these types of ODEs and we will see some examples of equations that can be transformed into these ones. It is important to know that homogeneity and linearity are properties that can be defined for higher-order ODEs and hence ODEs can be classified as linear vs. non-linear and as homogeneous vs. non-homogeneous. For simplicity, we will present here the definition for first-order ODEs but we should keep that these two types are particularly important.

a) Separable ODEs We will say that a first-order ODE is separable if $N(x, y) = N(y)$ and $M(x, y) = M(x)$:

$$N(y)y' + M(x) = 0.$$

Example: The first example we discussed in this chapter was a separable ODE, and it was very easy to solve it!

b) Homogeneous ODEs We say that the equation is *homogeneous* if $Q(x) = 0$ and if $M(x, y)$ and $N(x, y)$ are *homogeneous functions* of the same order. What is a homogeneous function?

Definition: We say that a function $M(x, y)$ is homogeneous if, when we multiply each variable by a factor λ, we can factor it out:

$$M(\lambda x, \lambda y) = \lambda^k M(x, y),$$

calling the constant k the *degree* of the (homogeneous) function.

Example: The following function is homogeneous of degree two :

$$N(x, y) = 10x^2 + y^2 \sin(x/y).$$

Multiplying both variables times λ we can factorize out a term λ^2,

$$N(\lambda x, \lambda y) = 10(\lambda x)^2 + (\lambda y)^2 \sin(\lambda x/(\lambda y)) = \lambda^2(10x^2 + y^2 \sin(x/y)) = \lambda^2 N(x, y).$$

Example: If we look now for another homogeneous function of degree two, for instance $M(x, y) = xy$ (prove that this is true), we can propose a first-order homogeneous equation:

$$(10x^2 + y^2 \sin(x/y))y' + xy = 0$$

This equation is, however, non-linear. Would be possible to convert it into a linear one with the approximation we did for $\sin(x)$ in the example of the pendulum?

Note: If you compare the definition of separable and homogeneous ODEs, you will see that both have $Q(x) = 0$. Then, the conditions over the functions $M(x, y)$ and $N(x, y)$ are a little bit uglier for homogeneous ODEs, but we will see that it is possible to convert homogeneous ODEs into separable ODEs.

c) Exact ODEs We say that an ODE is *exact* if $Q(x) = 0$ and if there exist a function $\Psi = \Psi(x, y)$ such that:

$$\frac{\partial \Psi}{\partial x} = M(x, y)$$
$$\frac{\partial \Psi}{\partial y} = N(x, y).$$

Why such a weird condition? Well, because if these conditions hold, we will be able to recognize our beloved chain rule:

$$N(x,y)y' + M(x,y) = 0 \to \frac{\partial \Psi}{\partial x} + \frac{\partial \Psi}{\partial y}\frac{dy}{dx} = 0 \to \frac{d}{dx}\Psi(x, y(x)) = 0$$

And we know that the last equation has the solution $\Psi(x,y) = C$.

Note: We have seen so far three types of ODEs having $Q(x) = 0$ plus some properties on the functions M and N that will allow us to solve them analytically. What if $Q(x) \neq 0$?

d) Linear ODEs We say that the differential equation is *linear* if $N(x,y) = N(x)$ and $M(x,y) = M(x)y$ (otherwise, it is non-linear):

$$N(x)y' + M(x)y = Q(x).$$

What we should note is that we do not have non-linear functions on y or y'. For instance, the following ODE is non-linear: $yy' = Q(x)$ due to the product yy'.

Example: The ODE describing the oscillations of a pendulum is non-linear. Consider a pendulum of lenght L, that oscillates with an angle $\theta = \theta(t)$ driven by the acceleration of gravity g. Its motion is described by the equation:

$$\frac{d^2\theta}{dt^2} + \frac{g}{L}\sin(\theta) = 0,$$

which is non-linear, due to the term $sin(\theta)$. Nevertheless, for small oscillations we can approximate $\sin(\theta) \approx \theta$ and it becomes linear. What would be the order of the error? (hint: think in MacLaurin series).

Note: Now the term $Q(x)$ is not equal to zero in general, as it was in the three previous types, but note that there are ODEs that are both homogeneous and linear!

4.3 Existence and uniqueness of solutions

It is apparent that, in order to solve things analytically, we need very particular conditions over the functions M, N, and Q. Why is this the case? Well, because, in general, it is difficult to solve differential equations. This is why there is intense research to understand if, given a differential equation, we are able to say (at least) that there exist a solution for some initial conditions (the problem of existence) and, if it exists, if the solution is unique (the problem of uniqueness).

For the sake of illustration and before we get into the world of easy-to-solve problems, we show here one theorem on the existence and uniqueness of solutions. Given the following initial value problem:

$$\frac{dy}{dx} = f(x,y) \quad \text{with } y(x_0) = y_0,$$

when does a solution exist? If it exist, is it unique? The following conditions:

1. f is continuous in some rectangle $R = \{(x, y)/|x - x_0| \leq a, |y - y_0| \leq b\}$, and
2. $\frac{\partial f}{\partial y}$ is continuous in R,

are *sufficient* to guarantee a solution in an interval $[x_0 - h, x_0 + h]$ with $h \leq a$.

Notes:

1. The above expression is kind of general, at least three out of four examples discussed above belong to this class.
2. This is only true for a neighbourhood around (x_0, y_0) and we don't know how far we can go beyond that.
3. These are *sufficient* conditions but are not *necessary*, so there may be cases in which the solution exists and it is unique even if these conditions do not hold.

4.4 Separable ODEs

We consider ODEs that have the form

$$N(y)y' + M(x) = 0.$$

In this case, we can separate the equation in terms on y and x. Since $y' = dy/dx$ we get that

$$N(y)dy = -M(x)dx,$$

that we can integrate independently and then solve for y.

Example. Solve the following ODE the initial value $y(0) = 1$:

$$\frac{dy}{dx} = \frac{3x^2 + 4x + 2}{2(y - 1)},$$

The problem asks for a particular solution, the one corresponding to the initial value $x = 0$ that it says should lead to $y = 1$. The ODE is separable, rearranging a little bit we see it easily:

$$2(y - 1)dy = (3x^2 + 4x + 2)dx$$

integrating both sides we get,

$$y^2 - 2y = x^3 + 2x^2 + 2x + C,$$

where we merged the constants arising from both integrals into C. We should determine the constant C considering the initial value given by the problem. Since for $x = 0$ we know that $y = 1$ we get that $C = 3$.

Let's make the solution explicit:

$$y = 1 \pm \sqrt{x^3 + 2x^2 + 2x + 4}.$$

Do we have two solutions for the starting conditions given by the problem? Let's double check the initial conditions, because if we make $x = 0$ we see that only the negative solution works, which would be the solution required.

Example Solve the following ODE for the initial value $y(0) = 1$:

$$\frac{dy}{dx} = \frac{y \cos x}{1 + 2y^2}.$$

Again, rearranging the equation we see that it is separable,

$$(1/y)dy + 2ydy = \cos x dx$$

which, after integration, leads to the general solution

$$\ln |y| + y^2 = \sin x + C.$$

Looking for the particular solution requires to substitute $x = 0$, and to find the value of the constant C such that $y = 1$, which is $C = 1$. Therefore, the final solution is

$$\ln |y| + y^2 = \sin x + 1,$$

that we can just keep implicit.

4.5 Homogeneous ODEs

Homogeneous ODEs have the form

$$N(x, y)y' + M(x, y) = 0,$$

where $N(x, y)$ and $M(x, y)$ are homogeneous functions of the same degree. This fact allows to rewrite the equation as follows:

$$y' = \frac{-M(x, y)}{N(x, y)} = f(y/x),$$

i.e. we will deal with a function $f(y/x)$ which makes the change of variables $u = y/x$ appropriate to transform the homogeneous ODE into a separable one. Indeed, since $y = ux \;\rightarrow\; y' = u'x + u$, we get for the above general expression

$$x\frac{du}{dx} + u = f(u) \quad \rightarrow \frac{du}{f(u) - u} = \frac{dx}{x}.$$

Integrating the right hand side will always lead to $\ln|x| + c$, and the left side will depend on $f(u) - u$ which will be the function that will change through the different problems.

Example: Solve the following ODE:

$$\frac{dy}{dx} = \frac{x^2 + xy + y^2}{x^2}.$$

Although the equation is not presented in the general form, it is easy to see that the functions in the numerator and denominator are both homogeneous of degree two. Since we know that the change $u = y/x$ works well for these equations, let's transform it first multiplying up and down times x^2, which leads to

$$\frac{dy}{dx} = \frac{1 + y/x + y^2/x^2}{1}.$$

Making the suggested change $y = ux \;\rightarrow\; y' = u'x + u$ we get

$$x\frac{du}{dx} + u = u^2 + u + 1 \quad \rightarrow \int \frac{du}{u^2 + 1} = \int \frac{dx}{x},$$

which leads to the general solution,

$$\arctan(u) = \ln|u| + C \quad \rightarrow \arctan(y/x) - \ln|x| = C,$$

where, in the last step, we changed the variables back to the original ones.

Example: Solve the following ODE:

$$\left(x\sin\left(\frac{y}{x}\right) - y\cos\left(\frac{y}{x}\right)\right)dx + x\cos\left(\frac{y}{x}\right)dy = 0$$

Both functions are homogeneous of degree one and hence we can apply the usual change, getting

$$\frac{dy}{dx} = -\left(\frac{x\sin\left(\frac{y}{x}\right) - y\cos\left(\frac{y}{x}\right)}{x\cos\left(\frac{y}{x}\right)}\right) \Rightarrow \frac{du}{dx}x + u = -\left(\frac{x\sin(u) - xu\cos(u)}{x\cos(u)}\right),$$

where it is possible to remove x in the right-hand of the equation:

$$\frac{du}{dx}x = -\left(\frac{\sin(u) - u\cos(u)}{\cos(u)}\right) - u = -\frac{\sin u}{\cos u}$$

separate variables...

$$\frac{\cos u}{\sin u}du = -\frac{dx}{x}$$

and integrating we get

$$\ln|\sin u| = -\ln|x| + \ln|C|$$

and we can just leave the solution in implicit form, for instance as

$$\left|x \sin\left(\frac{y}{x}\right)\right| = C.$$

4.6 Exact ODEs

We start with the same general expression than for homogeneous equations,

$$N(x,y)y' + M(x,y) = 0,$$

but now the functions $N(x,y)$ and $M(x,y)$ fulfill the relation

$$\frac{\partial\Psi}{\partial x} = M(x,y)$$
$$\frac{\partial\Psi}{\partial y} = N(x,y).$$

Above we showed that if these conditions hold, a solution of the ODE is $\Psi(x,y) = C$. The question is, how can we obtain such a function? Well, since $\partial\Psi/\partial x = M(x,y)$ we could simply integrate $M(x,y)$ with respect to x (or we could integrate $N(x,y)$ with respect to y, since the problem is symmetric). But this may not be enough, let's see why. Let's assume that $\Psi(x,y) = f(x,y)+h(y)$, if we do $\partial\Psi/\partial x = \partial f(x)/\partial x = M(x,y)$ and, if we integrate back:

$$\int M(x,y)dx = \int \frac{\partial f(x)}{\partial x}dx = f(x) + C \neq \Psi(x,y) \quad !!$$

The problem is that, if $\Psi(x,y)$ has any term which depends only on the variable y, differentiating with respect to x will make this term to dissapear like tears in the rain. Let's say then that Ψ is the ouctome of integrating $M(x,y)$ with respect to x plus possibly some function $h(y)$:

$$\Psi(x,y) = \int M(x,y)dx + h(y) = f(x,y) + h(y),$$

and now let's use the second condition, namely that $\partial\Psi/\partial y = N(x,y)$:

$$\frac{\partial \Psi(x,y)}{\partial y} = \frac{\partial f(x,y)}{\partial y} + h'(y) = N(x,y).$$

And we can get $h(y)$ just solving the above expression for $h'(y)$ and integrating.

Example: Solve the exact ODE:

$$\underbrace{(y\cos(x) + 2xe^y)}_{M(x,y)} + \underbrace{(\sin(x) + x^2e^y - 1)}_{N(x,y)}\, y' = 0.$$

Let's first check if the ODE is indeed exact:

$$\frac{\partial M(x,y)}{\partial x} = \cos(x) + 2xe^y; \quad \frac{\partial N(x,y)}{\partial y} = \cos(x) + 2xe^y.$$

It is! therefore there exist a function Ψ, let's look for the function:

$$\begin{aligned} \Psi(x,y) &= \int M(x,y)dx + h(y) = \int y\cos(x)dx + \int 2xe^y dx + h(y) \\ &= y\sin(x) + x^2e^y + h(y) + C. \end{aligned}$$

Now we should determine $h(y)$, using the relation $\partial\Psi/\partial y = N(x,y)$:

$$\frac{\partial \Psi(x,y)}{\partial y} = \sin(x) + x^2e^y + h'(y) = \underbrace{\sin(x) + x^2e^y - 1}_{N(x,y)},$$

which means that $h'(y) = -1 \quad \to h(y) = -y + C$. Therefore, the final solution is

$$\Psi(x,y) = y\sin(x) + x^2e^y - y + C.$$

4.7 Linear ODEs

Before we get into linear equations, we need to learn first an important concept: the *integrating factor*. Basically, it is a term that will multiply our ODE and will allow us to solve it, and it is a tool very frequently used to solve differential equations. In particular, for linear equations there is a general method to find such a factor. It is not difficult but, to introduce it properly, we will first get some intuition with an example of a separable ODE, then we will present the method and demonstrate why it works, and finally we will do some exercises.

Intuition behind the integrating factor. Let's start considering the ODE:

$$\frac{dy}{dx} + \frac{1}{2}y = \frac{3}{2}.$$

This equation is separable:

$$\int \frac{dy}{y-3} = \int \frac{-1}{2} dx \quad \rightarrow \ln|y-3| = e^{(-x/2+C)},$$

and leads to the solution

$$y = 3 + Ce^{-x/2}.$$

Now, let's multiply both sides of the solution by $e^{x/2}$:

$$ye^{x/2} = 3e^{x/2} + C$$

and take derivatives in both sides

$$\underbrace{y'e^{x/2} + \frac{y}{2}e^{x/2}}_{\frac{d(ye^{x/2})}{dx}} = \frac{3}{2}e^{x/2} \quad \rightarrow y' + \frac{1}{2}y = \frac{3}{2},$$

which is the starting ODE! Since from the solution we can get the original ODE back simply multiplying by $e^{x/2}$ and taking derivatives, it should be also easy to find the solution of the original ODE if we multiply by the same term, and we say that $e^{x/2}$ is an integrating factor. Let's see:

$$e^{x/2}y' + \frac{1}{2}ye^{x/2} = \frac{3}{2}e^{x/2} \quad \rightarrow \int \frac{d(ye^{x/2})}{dx} = \int \frac{3}{2}e^{x/2}$$

And we get

$$ye^{x/2} = \frac{3}{2}2e^{x/2} + C \quad \rightarrow y = 3 + Ce^{-x/2}.$$

So far so good, but we found the integral factor after finding the solution! Is it any general method to find this factor in advance? For linear ODEs there is.

Method to find the integrating factor in linear equations. Let's start considering the general expression for linear equations presented above and divide both sides by $N(x)$:

$$N(x)y' + M(x)y = Q(x) \quad \rightarrow y' + \underbrace{\frac{M(x)}{N(x)}}_{p(x)} y = \underbrace{\frac{Q(x)}{N(x)}}_{g(x)},$$

where we are just renaming the functions for simplicity. With this notation the general expression for linear equations is now:

$$y' + p(x)y = g(x),$$

and the integrating factor $\mu(x)$ can be easily defined as

$$\mu(x) = e^{\int p(x)dx},$$

from which we can derive a general expression for the solution of linear ODEs:

$$y = \frac{\int \mu(x)g(x)dx + C}{\mu(x)}.$$

Proof: One can memorize the above formula for the solution but it is not really needed. Let's multiply the general expression of linear ODEs by the integrating factor:

$$\underbrace{e^{\int p(x)dx}y' + p(x)e^{\int p(x)dx}y}_{(uv)'=uv'+u'v} = g(x)e^{\int p(x)dx},$$

we see that it happens as in the example we discussed above, that the left hand side of the equation is the result of the derivative of the product of two functions $u = e^{\int p(x)dx}$ and $v = y$, and hence

$$\underbrace{\frac{d}{dx}\left(e^{\int p(x)dx}y\right)}_{(uv)'} = g(x)e^{\int p(x)dx}$$

which, integrating both sides leads to

$$\underbrace{e^{\int p(x)dx}}_{\mu(x)}y + C = \int g(x)e^{\int p(x)dx}dx$$

and this is how we found the above solution

$$y = \frac{\int \mu(x)g(x)dx + C}{\mu(x)}.$$

Example: Solve the ODE with starting conditions $y(0) = 0$

$$y' - 2xy = x.$$

Since $p(x) = -2x$ the integrating factor is $\mu(x) = e^{\int -2xdx} = e^{-x^2}$. Multiplying both sides of the ODE by $\mu(x)$ we get

$$\underbrace{e^{-x^2}y' - 2xye^{-x^2}}_{\text{is this } \frac{d}{dx}(e^{-x^2}y)?} = xe^{-x^2},$$

and the answer to the above question is yes, so we can write

$$\int d\left(e^{-x^2}y\right) \quad = \int xe^{-x^2} \quad \underbrace{=}_{\substack{\uparrow \\ \text{change vars: } \begin{cases} u = -x^2 \to du = -2xdx \\ -\int \frac{1}{2}e^u du = -\frac{1}{2}e^u \end{cases}}} \frac{1}{2}e^{-x^2} + .C$$

The problem asks for the particular solution $y(0) = 0$, from which we get $C = 1/2$, and the final solution is:

$$y = \frac{1}{2}\left(e^{x^2} - 1\right).$$

Exercise: Solve the following ODE:

$$y' + \frac{1}{x}y = x^2.$$

The ODE is clearly linear so we start computing the integrating factor: $\mu(x) = e^{\int \frac{1}{x}dx} = e^{\ln|x|} = x$. We multiply both sides of the equation by $\mu(x)$ getting

$$\underbrace{xy' + y}_{\text{Is } xy'+y=(xy)'?} \quad = x^3$$

Again, the answer is yes, so we can easily integrate:

$$\int d(xy) = \int x^3 dx \quad \to xy = \frac{x^4}{4} + C,$$

and, solving for y we get

$$y = \frac{x^3}{4} + \frac{C}{x}.$$

5 Algebra

5.1 Vectors

Until now we have dealt only with *scalars* which are one-dimensional entities consisting of a magnitude and a sign. Higher-dimensional entities are composed of several scalars each of which is related to a particular direction. These objects are called vectors and are represented in print by either using bold symbols like $\boldsymbol{x}$ or with an arrow on top as in $\overrightarrow{x}$. An n-dimensional vector has n components x_i with $i = 1, ..., n$. Its magnitude is given by $|\overrightarrow{x}| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$.

Notation: $\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$ is a column vector and $\overrightarrow{x} = (x_1, x_2, ..., x_n)$ is row vector.

Sometimes a row vector is specifically denoted as $\overrightarrow{x}^T$ (T for transposed).

A vector is graphically represented by an arrow. The vector's magnitude $|\overrightarrow{x}|$ is denoted by the arrow's length. If the starting point of the vector coincides with the origin of the coordinate system, then its end point corresponds to the coordinates of the vector components. Such a vector is called a coordinate vector.

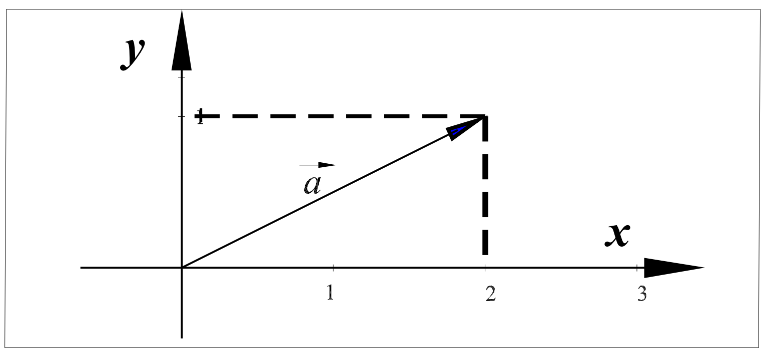
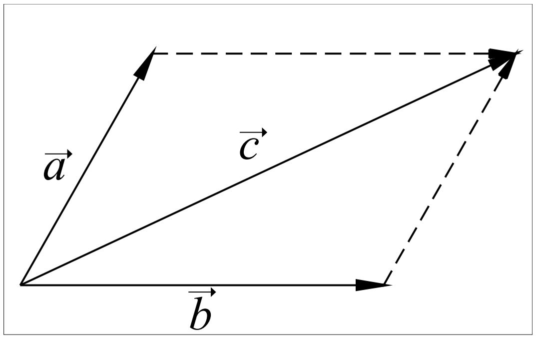


Figure 29: The vector $(1, 2)$ is an arrow from the origin to the point $x = 1$ and $y = 2$.

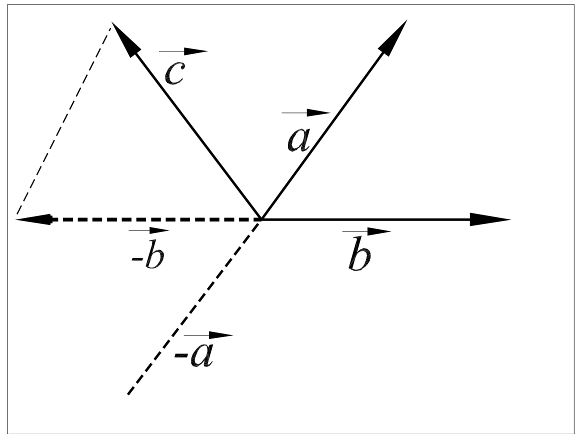
5.2 Elementary Vector Operations

5.2.1 Addition and Subtraction

The sum two vectors can be obtained graphically by either shifting the tail of the second arrow to the head of the first, or by constructing the parallelogram that is defined by the two arrows. The difference between two vectors can be found by adding the vector that has the same length but points into the opposite direction.



(a) The sum of two vectors $\vec{c} = \vec{a} + \vec{b}$



(b) The difference between two vectors $\vec{c} = \vec{a} - \vec{b}$

Figure 30: Addition and subtraction of vectors

In components: $\quad \vec{a} + \vec{b} = (a_1 + b_1, ..., a_n + b_n) = (c_1, ..., c_n) = \vec{c}$

Properties:

$$\begin{array}{rcll} \vec{a} + \vec{b} & = & \vec{b} + \vec{a} & \qquad \textbf{commutative} \\ (\vec{a} + \vec{b}) + \vec{c} & = & \vec{a} + (\vec{b} + \vec{c}) & \qquad \textbf{associative} \end{array}$$

A closed polygon corresponds to the vector sum equal $\vec{0}$.

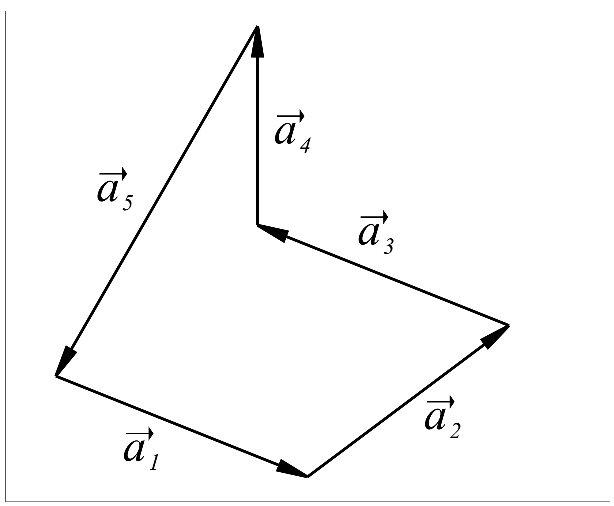


Figure 31: $\vec{a_1} + \vec{a_2} + \ldots + \vec{a_5} = \vec{0}$

Important note: Make sure you understand that $\vec{0} \neq 0$!!!!

5.2.2 Multiplication of a Vector with a Scalar

A vector can be multiplied with a scalar by multiplying each of the components which results in either stretching or squeezing of the vector and may change its orientation.

$$\vec{a} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \vec{b} = -2\,\vec{a} = -2\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

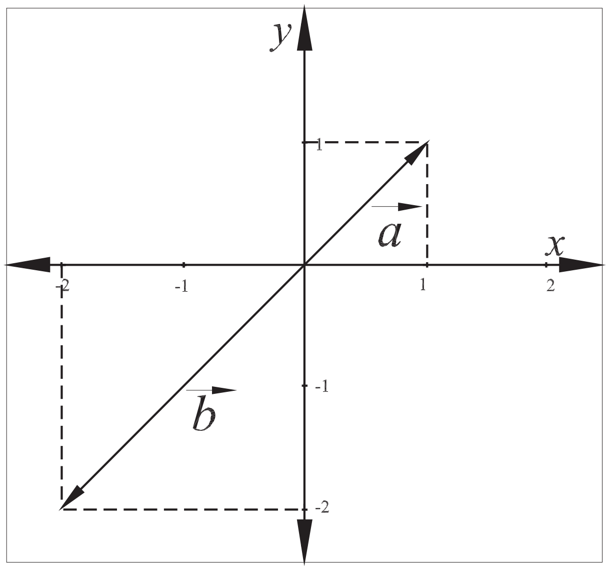


Figure 32: The multiplication of a vector with a scalar $\vec{b} = -2\,\vec{a}$

Linear dependence of vectors:

n vectors $\vec{a_1}, \cdots, \vec{a_n}$ are called *linearly independent*, if the only way to fulfill

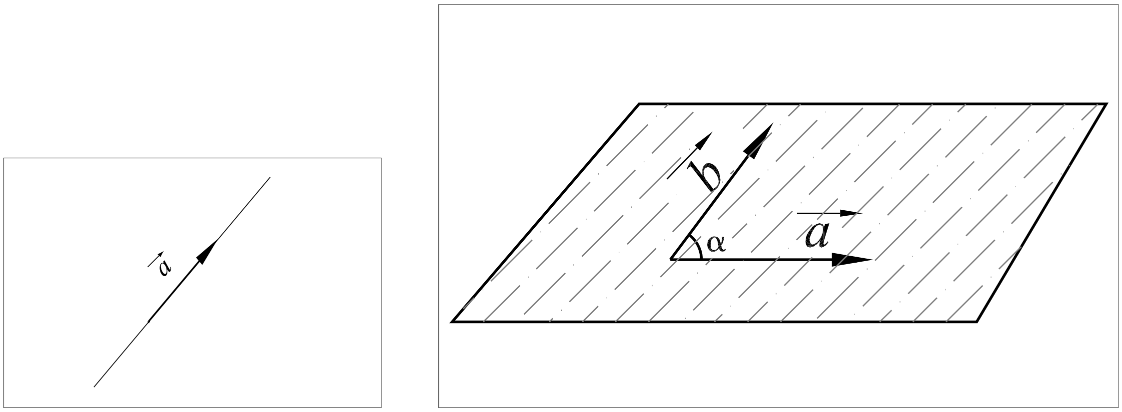
$$\alpha_1\,\vec{a_1} + \alpha_2\,\vec{a_2} + \cdots + \alpha_n\,\vec{a_n} = 0 \quad \text{is} \quad \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$$

If this relation can be fulfilled with certain $\alpha_i \neq 0$, then the vectors are said to be *linearly dependent*. For instance, imagine $\alpha_1 \neq 0$, all others are free. Then $\vec{a_1}$ may be expressed by the other vectors and is redundant.

$$\vec{a_1} = -\frac{1}{\alpha_1}\sum_{i=2}^{n} \alpha_i\,\vec{a_i} \tag{1}$$

One-dimensional: $\alpha\,\vec{a}$ represents all vectors on a straight line. Such vectors are called collinear.

Two-dimensional: $\alpha_1\,\vec{a_1} + \alpha_2\,\vec{a_2}$ represents all vectors in the plane. These vectors are coplanar.



(a) Collinear vectors define a line (b) Two non-collinear vectors span a plane

Figure 33: Collinear and coplanar vectors

5.2.3 Scalar Product

Two vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ can be multiplied such that the result is a scalar c. This operation is called the *scalar, dot* or *inner* product.

$$\overrightarrow{a} \cdot \overrightarrow{b} = \mid \overrightarrow{a} \mid \mid \overrightarrow{b} \mid \cos\alpha \qquad \text{where } \alpha \text{ is the angle between } \overrightarrow{a} \text{ and } \overrightarrow{b}$$

$$\overrightarrow{a} \cdot \overrightarrow{b} = a_1\, b_1 + a_2\, b_2 + ... + a_n\, b_n \qquad \text{scalar product in components}$$

The scalar product measures the contribution of vector $\overrightarrow{a}$ to vector $\overrightarrow{b}$ If the angle between $\overrightarrow{a}$ and $\overrightarrow{b}$ is 90° the two vectors are orthogonal, there are no contributions at all.

Properties:

$$\overrightarrow{a} \cdot \overrightarrow{b} = \overrightarrow{b} \cdot \overrightarrow{a} \qquad \textbf{commutative}$$

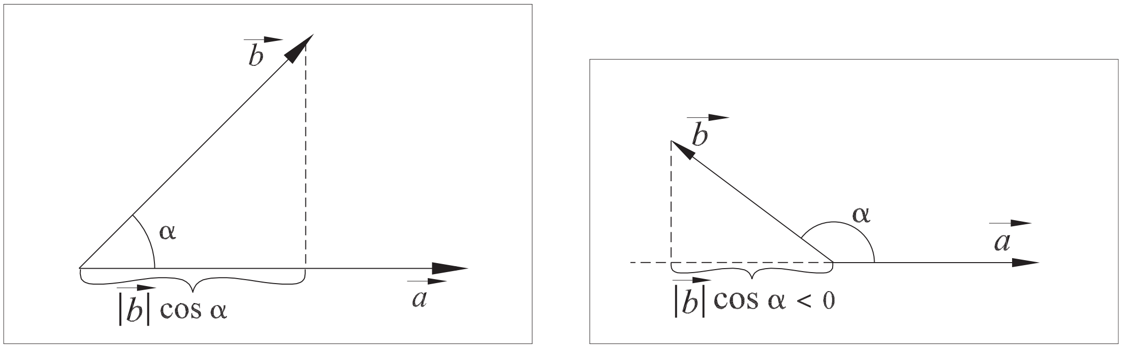
$$(c\,\overrightarrow{a}) \cdot \overrightarrow{b} = c\,(\overrightarrow{a} \cdot \overrightarrow{b}) = \overrightarrow{a} \cdot (c\,\overrightarrow{b}) \qquad \textbf{associative}$$

$$(\overrightarrow{a_1} + \overrightarrow{a_2}) \cdot \overrightarrow{b} = \overrightarrow{a_1} \cdot \overrightarrow{b} + \overrightarrow{a_2} \cdot \overrightarrow{b} \qquad \textbf{distributive}$$

Examples:

$$\overrightarrow{a} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad \overrightarrow{b} = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \qquad \overrightarrow{a} \cdot \overrightarrow{b} = 2 \cdot 1 + 0 \cdot (-2) = 2$$

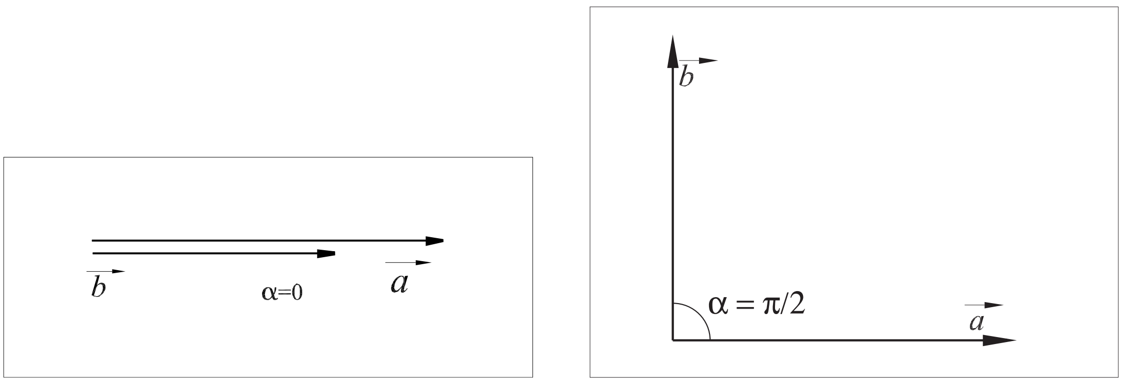
$$\mid \overrightarrow{a} \mid = 2 \qquad \mid \overrightarrow{b} \mid = \sqrt{5} \quad \rightarrow \quad \cos\alpha = \frac{\overrightarrow{a} \cdot \overrightarrow{b}}{\mid \overrightarrow{a} \mid \mid \overrightarrow{b} \mid} = \frac{2}{2\sqrt{5}} \quad \rightarrow \quad \alpha = \arccos\frac{1}{\sqrt{5}} = 1.107 \approx 63°$$



(a) Projection of $\overrightarrow{b}$ on $\overrightarrow{a}$

(b) If $\alpha > 90^0$ then $\cos\alpha < 0$ and the scalar product is negative

Figure 34: Scalar Product



(a) The dot product has is maximal value in the case $\alpha = 0 \ \rightarrow \cos\alpha = 1$

(b) For $\alpha = 90^\circ$ the scalar product vanishes because $\cos\alpha = 0$

Figure 35: Scalar product for parallel and orthogonal vectors

5.2.4 Vector Product

Two vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ can be multiplied such that the result is a vector $\overrightarrow{c}$. This operation is called the *vector, cross* or *outer* product.

The vector product exists only in three dimensions !!!

$$\overrightarrow{a} \times \overrightarrow{b} = \overrightarrow{c} \qquad\qquad | \overrightarrow{c} | = | \overrightarrow{a} \times \overrightarrow{b} | = | \overrightarrow{a} | | \overrightarrow{b} | \sin\alpha$$

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix} \qquad \text{vector product in components}$$

The result of a vector product between two non-collinear vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ is a vector $\overrightarrow{c}$ which has a magnitude of $| \overrightarrow{a} | | \overrightarrow{b} | \sin\alpha$ and points into the direction perpendicular to the plane defined by $\overrightarrow{a}$ and $\overrightarrow{b}$ such that $\overrightarrow{a}$, $\overrightarrow{b}$ and $\overrightarrow{c}$ form a right-handed system. To find this direction have your right thumb point into the direction of $\vec{a}$, the right index into the direction of $\vec{b}$, and the right middle finger perpendicular to the plane defined by $\vec{a}$ and $\vec{b}$. There is only one way to do that without hurting yourself seriously. Now the middle finger points into the direction of $\vec{c}$.

Hint: It is imperative that you use the **right** hand for this.

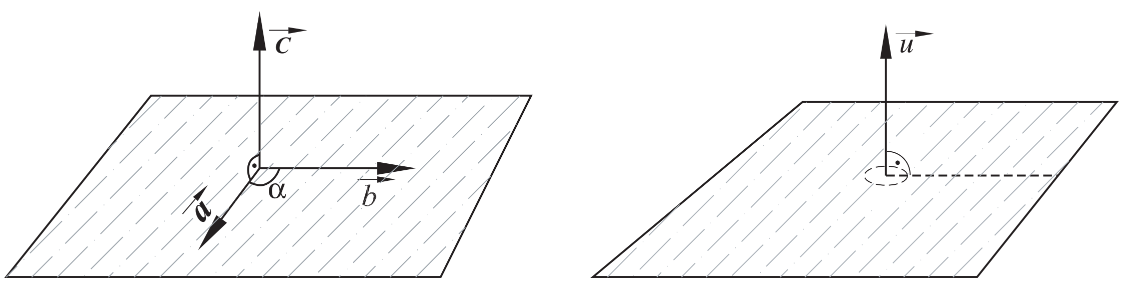
Properties:

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a} \qquad \text{anti-commutative}$$

$$(c\,\vec{a}) \times \vec{b} = c\,(\vec{a} \times \vec{b}) = \vec{a} \times (c\,\vec{b}) \qquad \text{associative}$$

$$(\vec{a_1} + \vec{a_2}) \times \vec{b} = \vec{a_1} \times \vec{b} + \vec{a_2} \times \vec{b} \qquad \text{distributive}$$

Note: In 3 dimensions a plane can be defined by a point in the plane and its *normal vector* $\vec{n}$.



(a) $\vec{a}$ and $\vec{b}$ span a plane

(b) A plane defined by a point and its normal vector

Figure 36: Vectors in 3-dimensional space

5.3 Matrices

A matrix **A** operates on a vector $\vec{x}$ and transforms, i.e. stretches, squeezes or rotates it.

$$\vec{y} = \boldsymbol{A}\,\vec{x} \qquad \text{where} \quad \boldsymbol{A} = \begin{pmatrix} A_{11} & A_{12} & A_{1n} \\ A_{21} & \ddots & \dots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} = A_{ij} \quad \text{is a } \; n \times n \text{ matrix}$$

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} A_{11}\,x_1 + A_{12}\,x_2 + \dots A_{1n}\,x_n \\ A_{21}\,x_1 + A_{22}\,x_2 + \dots A_{2n}\,x_n \\ \vdots \\ A_{n1}\,x_1 + A_{n2}\,x_2 + \dots A_{nn}\,x_n \end{pmatrix}$$

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}}_{\boldsymbol{A}} \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} 2 \\ -1 \end{pmatrix}}_{\vec{y}}$$

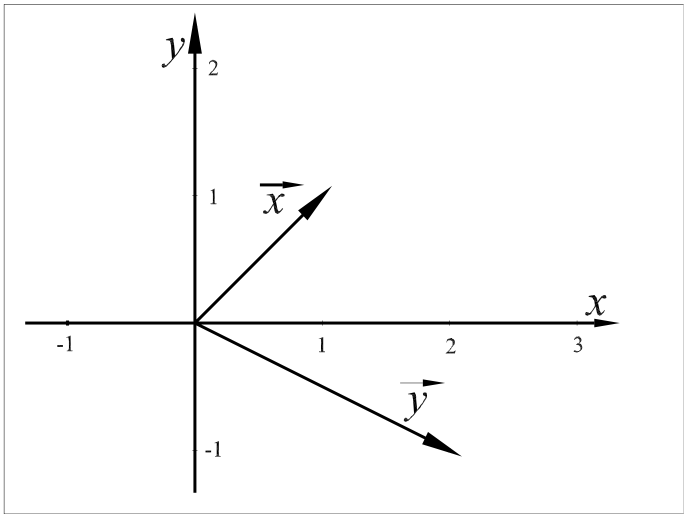


Figure 37: Rotation and scaling of a vector

Properties:

$$\begin{aligned} \boldsymbol{A} = \boldsymbol{B} \quad &\rightarrow \quad a_{ij} = b_{ij} \\ \boldsymbol{A} + \boldsymbol{B} \quad &\rightarrow \quad a_{ij} + b_{ij} \\ c\,\boldsymbol{A} \quad &\rightarrow \quad c\,a_{ij} \\ \boldsymbol{A} + \boldsymbol{B} &= \boldsymbol{B} + \boldsymbol{A} \end{aligned}$$

5.4 Multiplication of Matrices

The product of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is found by calculating the scalar products between the *rows* of matrix $\boldsymbol{A}$ and the *columns* of matrix $\boldsymbol{B}$. This implies that the number of columns of matrix $\boldsymbol{A}$ must be the same as the number of rows of matrix $\boldsymbol{B}$.

Examples:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}\,b_{11} + a_{12}\,b_{21} & a_{11}\,b_{12} + a_{12}\,b_{22} \\ a_{21}\,b_{11} + a_{22}\,b_{21} & a_{21}\,b_{12} + a_{22}\,b_{22} \end{pmatrix}$$

$$\boldsymbol{A}\,\boldsymbol{B} = \begin{pmatrix} 5 & 3 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} 2 & -3 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 5 \cdot 2 + 3 \cdot (-1) & 5 \cdot (-3) + 3 \cdot 4 \\ 2 \cdot 2 + 7 \cdot (-1) & 2 \cdot (-3) + 7 \cdot 4 \end{pmatrix} = \begin{pmatrix} 7 & -3 \\ -3 & 22 \end{pmatrix}$$

$$\boldsymbol{B}\,\boldsymbol{A} = \begin{pmatrix} 2 & -3 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 2 & 7 \end{pmatrix} = \begin{pmatrix} 4 & -15 \\ 3 & -25 \end{pmatrix} \neq \boldsymbol{A}\,\boldsymbol{B}$$

The multiplication of matrices is NOT commutative!!!!!

In general a $n \times m$ matrix can be multiplied with a $m \times n$ matrix and the result is a $n \times n$ matrix.

$$\boldsymbol{C} = \boldsymbol{A}\,\boldsymbol{B} = \begin{pmatrix} \Sigma_i\, a_{1i}\, b_{i1} & \Sigma_i\, a_{1i}\, b_{i2} & \dots & \Sigma_i\, a_{1i}\, b_{in} \\ \Sigma_i\, a_{2i}\, b_{i1} & \Sigma_i\, a_{2i}\, b_{i2} & \dots & \Sigma_i\, a_{2i}\, b_{in} \\ \vdots & \vdots & & \vdots \\ \Sigma_i\, a_{ni}\, b_{i1} & \Sigma_i\, a_{ni}\, b_{i2} & \dots & \Sigma_i\, a_{ni}\, b_{in} \end{pmatrix} \qquad \text{with } \Sigma_i = \sum_{i=1}^{m}$$

5.5 Transposed Matrix

The transposed of a matrix is found by exchanging the row and column vectors.

$$\boldsymbol{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & \cdot & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \dots & A_{mn} \end{pmatrix} \quad \rightarrow \quad \boldsymbol{A^T} = \begin{pmatrix} A_{11} & A_{21} & A_{31} & \dots \\ A_{12} & A_{22} & \cdot & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \dots & A_{nm} \end{pmatrix}$$

Examples:

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 4 & 0 \end{pmatrix} \quad \rightarrow \quad \boldsymbol{A^T} = \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 5 & 0 \end{pmatrix}$$

$$\boldsymbol{A}\,\boldsymbol{A^T} = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 4 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 5 & 0 \end{pmatrix} = \begin{pmatrix} 38 & 14 \\ 14 & 17 \end{pmatrix}$$

$$\boldsymbol{A^T}\,\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 5 & 0 \end{pmatrix} \begin{pmatrix} 2 & 3 & 5 \\ 1 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 10 & 10 \\ 10 & 25 & 15 \\ 10 & 15 & 25 \end{pmatrix}$$

This should eliminate any remaining doubts that matrix multiplication could be commutative. These matrices are not even the same size.

5.6 Basis vectors

Basis vectors span a coordinate system and can be represented in various ways

$$\vec{i}\,,\ \vec{j}\,,\ \vec{k} \qquad \vec{e_1}, \vec{e_2}, \vec{e_3} \qquad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\vec{s} = \begin{pmatrix} x \\ y \end{pmatrix} = x \overbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}^{\vec{e_1}} + y \overbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}^{\vec{e_2}} = x\,\vec{e_1} + y\,\vec{e_2}$$

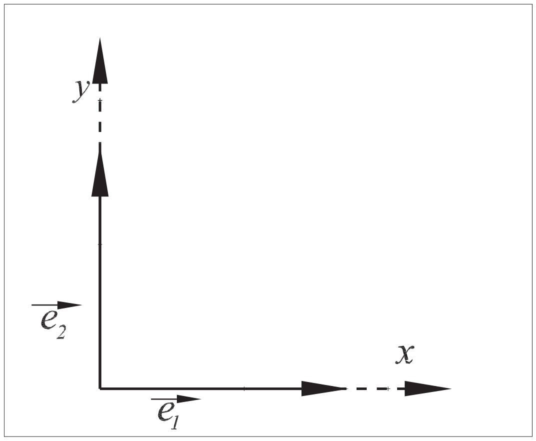


Figure 38: Basis vectors.

5.7 Transformation of coordinate systems

In general the component of a vector depend on the coordinate system used. Coordinate systems can be transformed, which changes the vector components in a certain way and a vector $\vec{s}$ in the old coordinate system becomes the vector $\widetilde{\vec{s}}$ in the new coordinates. The two easiest transformations of a coordinate system are a translation or shift and a rotation around the origin.

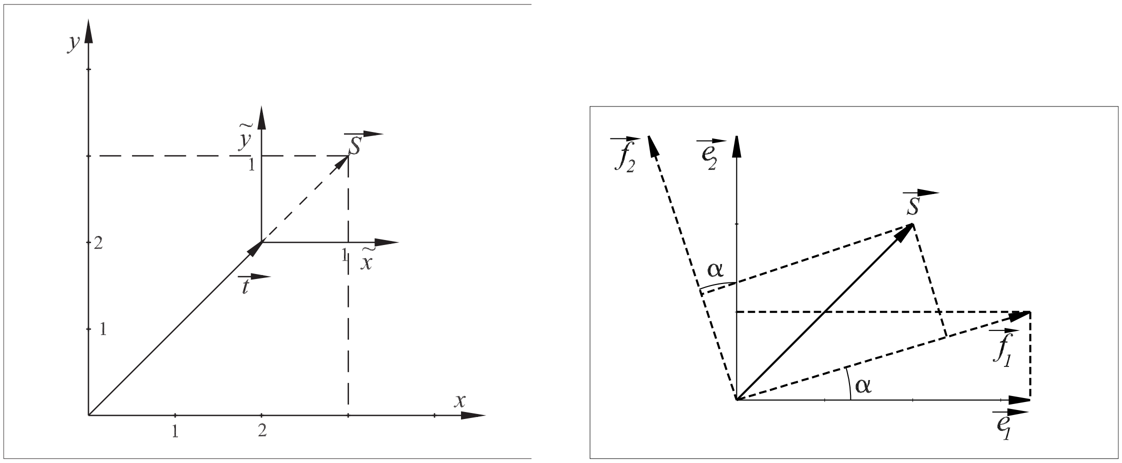
5.7.1 Translation

A translation of the coordinate system is performed by adding a constant vector

$$\vec{s} \to \widetilde{\vec{s}} = \vec{s} + \vec{t} \qquad \text{shifts the coordinate system by } \vec{t} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$$

Components of $\vec{s}$ in old system: $\vec{s} = \begin{pmatrix} x + t_1 \\ y + t_2 \end{pmatrix} = \begin{pmatrix} 1 + t_1 \\ 2 + t_2 \end{pmatrix}$

Components of $\widetilde{\vec{s}}$ in the new system: $\widetilde{\vec{s}} = \begin{pmatrix} \widetilde{x} \\ \widetilde{y} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$



(a) Translation of the coordinate system (b) Rotation of the coordinate system

Figure 39: Coordinate transformations: Translation and Rotation

5.7.2 Rotation

A rotation of the coordinate system by an angle α around the origin is performed by applying the rotation matrix $\boldsymbol{R}$ to the vector $\widetilde{\vec{s}}$

$$\boldsymbol{R} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \qquad \widetilde{\vec{s}} = \boldsymbol{R}\,\vec{s} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\,\cos\alpha - y\,\sin\alpha \\ x\,\sin\alpha + y\,\cos\alpha \end{pmatrix}$$

The rotation matrix $\boldsymbol{R}$ can be found by calculating the basis vectors $\widetilde{\vec{e_1}}$ and $\widetilde{\vec{e_2}}$ for the new coordinate system

$$\widetilde{\vec{e_1}} = \cos\alpha\,\vec{e_1} + \sin\alpha\,\vec{e_2} \qquad \text{and} \qquad \widetilde{\vec{e_2}} = -\sin\alpha\,\vec{e_1} + \cos\alpha\,\vec{e_2}$$

Representation of a point $\vec{s}$:

$$\vec{s} = \begin{pmatrix} x \\ y \end{pmatrix} = \underbrace{x\,\vec{e_1} + y\,\vec{e_2}}_{\text{old system}} = \underbrace{\tilde{x}\,\widetilde{\vec{e_1}} + \tilde{y}\,\widetilde{\vec{e_2}}}_{\text{new system}} = \begin{pmatrix} \tilde{x}\,\cos\alpha + \tilde{y}\,\sin\alpha \\ -\tilde{x}\,\sin\alpha + \tilde{y}\,\cos\alpha \end{pmatrix}$$

Relation between old and new coordinates:

$$\vec{x} = \boldsymbol{A}\,\widetilde{\vec{x}} \qquad \text{where} \qquad \boldsymbol{A} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

$$\widetilde{\vec{x}} = \boldsymbol{A^T}\,\vec{x} \qquad \text{with} \qquad \boldsymbol{A^T} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \qquad \text{and} \quad \boldsymbol{A}\,\boldsymbol{A^T} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{aligned} \vec{s} = x\,\vec{e_1} + y\,\vec{e_2} \quad &= \quad \tilde{x}\,\underbrace{(\cos\alpha\,\vec{e_1} + \sin\alpha\,\vec{e_2})}_{\widetilde{\vec{e_1}}} + \tilde{y}\,\underbrace{(-\sin\alpha\,\vec{e_1} + \cos\alpha\,\vec{e_2})}_{\widetilde{\vec{e_2}}} \\ &= \quad \underbrace{(\tilde{x}\,\cos\alpha - \tilde{y}\,\sin\alpha)}_{x}\,\vec{e_1} + \underbrace{(\tilde{x}\,\sin\alpha + \tilde{y}\,\cos\alpha)}_{y}\,\vec{e_2} \end{aligned}$$

5.7.3 Polar coordinates

Polar coordinates are often used if the problem under consideration has a certain symmetry. They are represented by a vector $\overrightarrow{e_r}$ from the origin to a point in the plane and a vector $\overrightarrow{e_\varphi}$ from that point with the direction tangentially to the unit circle.

$$\overrightarrow{S} = \begin{pmatrix} x \\ y \end{pmatrix} = x\,\overrightarrow{e_1} + y\,\overrightarrow{e_2} = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix} = r\,\overrightarrow{e_r} + \varphi\,\overrightarrow{e_\varphi} \quad \rightarrow \quad \begin{pmatrix} r \\ \varphi \end{pmatrix}_{pol}$$

$$\text{with} \quad r = \sqrt{x^2 + y^2} \quad \varphi = \arctan\frac{y}{x}$$

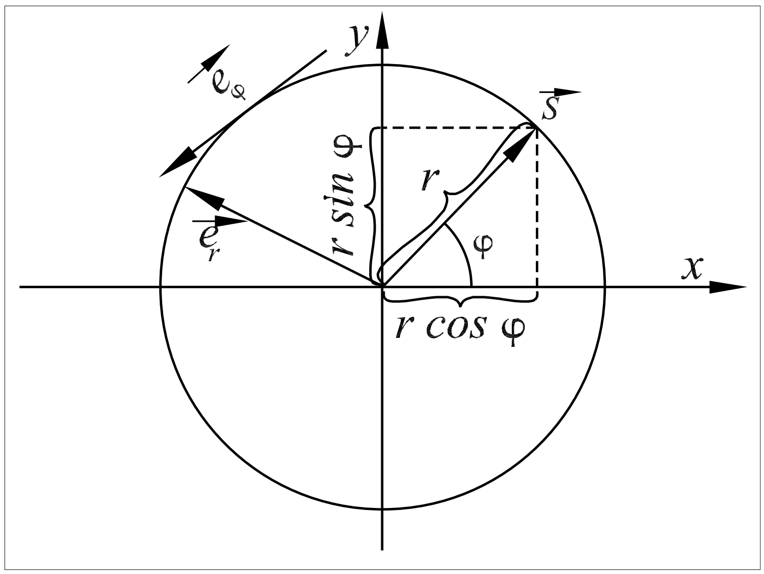


Figure 40: Polar coordinates

Example:

$$\overrightarrow{S} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{cart} = x\,\overrightarrow{e_1} + y\,\overrightarrow{e_2} \qquad \rightarrow \qquad r = \sqrt{1^2 + 1^2} = \sqrt{2} \qquad \varphi = \frac{\pi}{4} = 45^0$$

$$\overrightarrow{S} = r\,\overrightarrow{e_r} + \frac{\pi}{4}\,\overrightarrow{e_\varphi} \quad \rightarrow \quad \begin{pmatrix} \sqrt{2} \\ \frac{\pi}{4} \end{pmatrix}_{pol}$$

Note: The quantity $\begin{pmatrix} \sqrt{2} \\ \frac{\pi}{4} \end{pmatrix}_{pol}$ is *not* a vector!!!!

5.7.4 Non-orthogonal Coordinate Systems

Using a system of basis vectors that is orthogonal and normalized, i.e. $\overrightarrow{e_1} \cdot \overrightarrow{e_1} = \overrightarrow{e_2} \cdot \overrightarrow{e_2} = 0$ and $\overrightarrow{e_1} \cdot \overrightarrow{e_2} = 0$, or more general $\overrightarrow{e_i} \cdot \overrightarrow{e_j} = \delta_{ij}$ is very convenient because it is straight forward to find a certain component of a vector simply by multiplying with the corresponding basis vector

$$\vec{x} = x_1\,\vec{e_1} + x_2\,\vec{e_2} \;=\; \begin{cases} \vec{x}\cdot\vec{e_1} = x_1\,\overbrace{\vec{e_1}\cdot\vec{e_1}}^{=1} + x_2\,\overbrace{\vec{e_2}\cdot\vec{e_1}}^{=0} = x_1 \\ \vec{x}\cdot\vec{e_2} = x_1\,\underbrace{\vec{e_1}\cdot\vec{e_2}}_{=0} + x_2\,\underbrace{\vec{e_2}\cdot\vec{e_2}}_{=1} = x_2 \end{cases}$$

Sometimes, however it is necessary to represent vectors in a basis system $\vec{u}$, $\vec{v}$ that is not orthogonal. The easiest way to deal with this situation is to introduce a second set of basis vectors, the so-called *adjoint* vectors or *dual basis* $\vec{u}^\dagger$, $\vec{v}^\dagger$ such the relations

$$\vec{u}^\dagger\cdot\vec{u} = \vec{v}^\dagger\cdot\vec{v} = 1 \qquad \text{and} \qquad \vec{u}^\dagger\cdot\vec{v} = \vec{v}^\dagger\cdot\vec{u} = 0$$

are fulfilled. In components these equations read

$$u_1^\dagger\,u_1 + u_2^\dagger\,u_2 = 1 \qquad v_1^\dagger\,v_1 + v_2^\dagger\,v_2 = 1 \qquad u_1^\dagger\,v_1 + u_2^\dagger\,v_2 = 0 \qquad v_1^\dagger\,u_1 + v_2^\dagger\,u_2 = 0$$

which are four equations for the four unknowns $u_i^\dagger$, $v_i^\dagger$ and allows us to determine the adjoint vectors $\vec{u}^\dagger$ and $\vec{v}^\dagger$ if the original basis vectors $\vec{u}$ and $\vec{v}$ are linearly independent, i.e. not collinear.

Now we can express vectors in the basis $\vec{u}$ and $\vec{v}$, and determine their components by multiplying with the corresponding vectors from the adjoint basis

$$\vec{x} = a\,\vec{u} + b\,\vec{v} \;=\; \begin{cases} \vec{x}\cdot\vec{u}^\dagger = a\,\overbrace{\vec{u}\cdot\vec{u}^\dagger}^{=1} + b\,\overbrace{\vec{v}\cdot\vec{u}^\dagger}^{=0} = a \\ \vec{x}\cdot\vec{v} = a\,\underbrace{\vec{u}\cdot\vec{v}^\dagger}_{=0} + b\,\underbrace{\vec{v}\cdot\vec{v}^\dagger}_{=1} = b \end{cases}$$

Note: An ortho-normal basis system is simply the special case where the original and adjoint basis vectors are the same.

5.8 Determinants

The *determinant* is a is a descriptor of a matrix. The determinant of a 2×2 matrix is given by

$$\det \boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\,a_{22} - a_{21}\,a_{12}$$

The determinant of a 3×3 matrix it is defined as

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

Hint: The question arises, of course: "Who can remember something like this?" Well, it is actually not that difficult using the following construction. First copy the left and the middle column to the right. Then go through this scheme as indicated below: the left to right or southeast diagonals are counted positive, the right to left or southwest diagonals are counted negative, resulting in the formula for the determinant of a 3×3 matrix (unfortunately such a procedure does not exist for higher dimensional matrices and how to find their determinants is beyond the scope of this course).

$$\left| \begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \begin{array}{ccccccccc} a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\ & \searrow & & \searrow & & \searrow & & & \\ a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\ & & & \searrow & & \searrow & & \searrow & \\ a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32} \end{array} \qquad \textbf{positive: } + \\ \\ \begin{array}{ccccccccc} a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\ & & & \swarrow & & \swarrow & & \swarrow & \\ a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\ & \swarrow & & \swarrow & & \swarrow & & & \\ a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32} \end{array} \qquad \textbf{negative: } - \end{array} \right.$$

Properties:

If at least two of the column vectors are *linearly dependent* the determinant $\det \boldsymbol{A} = 0$.

$$\det(\boldsymbol{A}\,\boldsymbol{B}) = \det \boldsymbol{A}\, \det \boldsymbol{B}$$

Examples:

$$\begin{array}{rcl} x - 2y & = & \alpha_1 \\ 5\,(x - 2y) & = & \alpha_2 \end{array} \qquad \boldsymbol{A} = \begin{pmatrix} 1 & -2 \\ 5 & -10 \end{pmatrix} \qquad \vec{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

$$\det \boldsymbol{A} = (-10) \cdot 1 - (-2) \cdot 5 = 0$$

$$\left| \begin{array}{rrr|rr} 3 & 1 & 2 & 3 & 1 \\ 0 & -2 & 2 & 0 & -2 \\ 1 & 3 & 0 & 1 & 3 \end{array} \right. = 3 \cdot (-2) \cdot 0 + 1 \cdot 2 \cdot 1 + 2 \cdot 0 \cdot 3 - 2 \cdot (-2) \cdot 1 - 3 \cdot 2 \cdot 3 - 1 \cdot 0 \cdot 0 = -12$$

Note: The components of the vector product can be found from a determinant

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \vec{i}\,(a_2\,b_3 - a_3\,b_2) + \vec{j}\,(a_3\,b_1 - a_1\,b_3) + \vec{k}\,(a_1\,b_2 - a_2\,b_1)$$

5.9 The Inverse of a Matrix $\boldsymbol{A}^{-1}$

The matrix $\boldsymbol{A}$ has an inverse with the property $\boldsymbol{A}\,\boldsymbol{A}^{-1} = \boldsymbol{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ if $\det \boldsymbol{A} \neq 0$

Note:

$$\vec{y} = \boldsymbol{A}\,\vec{x} \qquad \rightarrow \qquad \vec{x} = \boldsymbol{A}^{-1}\,\vec{y}$$

$$\det(\boldsymbol{A}\,\boldsymbol{A}^{-1}) = \det \boldsymbol{I} = 1 = \det \boldsymbol{A}\,\det \boldsymbol{A}^{-1} \qquad \rightarrow \quad \det \boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}}$$

Inverse of a 2×2 matrix:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \boldsymbol{A}^{-1} = \frac{1}{\det\,\boldsymbol{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

$$\Rightarrow \quad \boldsymbol{A}\,\boldsymbol{A}^{-1} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \frac{1}{\det\,\boldsymbol{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

$$= \frac{1}{a_{11}\,a_{22} - a_{12}\,a_{21}} \begin{pmatrix} a_{11}\,a_{22} - a_{12}\,a_{21} & 0 \\ 0 & a_{11}\,a_{22} - a_{12}\,a_{21} \end{pmatrix} = \boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

5.10 Linear Systems of Equations

A system of the form

$$\begin{aligned} y_1 &= a_{11}\,x_1 + a_{12}\,x_2 + \dots + a_{1n}\,x_n \\ y_2 &= a_{21}\,x_1 + a_{22}\,x_2 + \dots + a_{2n}\,x_n \\ \vdots & \qquad\qquad \vdots \\ y_n &= a_{n1}\,x_1 + a_{n2}\,x_2 + \dots + a_{nn}\,x_n \end{aligned} \qquad \text{or} \qquad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

is called a *linear system of equations* and can be conveniently written in terms of vectors and matrices $\vec{y} = \boldsymbol{A}\,\vec{x}$. In most cases the coefficients a_{ij} and the values on the left hand side y_i are known, and one is interested in finding a solution, i.e. values for the x_i such that all equations are fulfilled. What are the conditions that such a system has solutions and what are their properties?

We distinguish two cases:

1. $\vec{y} \neq \vec{0}$, i.e. at least one of the $y_i \neq 0$. In this case the system is called *inhomogeneous* and it has a unique solution if $\det \boldsymbol{A} \neq 0$. Then the matrix $\boldsymbol{A}$ has an inverse and the solution is given by $\vec{x} = \boldsymbol{A}^{-1}\,\vec{y}$. For $\det \boldsymbol{A} = 0$ the system has either no solution or infinitely many depending on $\vec{y}$;

2. $\vec{y} = \vec{0}$, i.e. all of the $y_i = 0$. In this case the system is called *homogeneous* and it has always the solution $\vec{x} = \vec{0}$, which is called the *trivial* solution. Non-trivial solutions exist only if $\det \boldsymbol{A} = 0$ and then there are infinitely many.

Examples:

$$\begin{array}{c} 3\,x_1 + x_2 = 6 \\ 3\,x_1 - x_2 = 12 \end{array} \quad \text{inhom.,} \quad \det \boldsymbol{A} = \begin{vmatrix} 3 & 1 \\ 3 & -1 \end{vmatrix} = -6 \neq 0 \quad \rightarrow \quad \begin{array}{c} \text{unique} \\ \text{solution} \end{array} \quad \rightarrow \quad \begin{array}{c} x_1 = 3 \\ x_2 = -3 \end{array}$$

$$\begin{array}{c} 3\,x_1 + x_2 = 6 \\ 6\,x_1 + 2\,x_2 = 10 \end{array} \quad \text{inhom.,} \quad \det \boldsymbol{A} = \begin{vmatrix} 3 & 1 \\ 6 & 2 \end{vmatrix} = 0 \quad \rightarrow \quad 12 = 10 \text{ fi} \quad \rightarrow \quad \begin{array}{c} \text{no} \\ \text{solution} \end{array}$$

$$\begin{array}{c} 3\,x_1 + x_2 = 6 \\ 6\,x_1 + 2\,x_2 = 12 \end{array} \quad \text{inhom.,} \quad \det \boldsymbol{A} = \begin{vmatrix} 3 & 1 \\ 6 & 2 \end{vmatrix} = 0 \quad \rightarrow \quad x_2 = -3\,x_1 + 6 \quad \rightarrow \quad \begin{array}{c} \text{infinitely many} \\ \text{solutions} \end{array}$$

$$\begin{array}{c} 3\,x_1 + x_2 = 0 \\ 3\,x_1 - x_2 = 0 \end{array} \quad \text{hom.,} \quad \det \boldsymbol{A} = \begin{vmatrix} 3 & 1 \\ 3 & -1 \end{vmatrix} = -6 \quad \rightarrow \quad \begin{array}{c} x_2 = -3\,x_1 \\ x_2 = 3\,x_1 \end{array} \quad \rightarrow \quad \begin{array}{c} \text{trivial} \\ \text{solution} \end{array} \quad \rightarrow \quad \vec{x} = \vec{0}$$

$$\begin{array}{c} 6\,x_1 + 2\,x_2 = 0 \\ 3\,x_1 + x_2 = 0 \end{array} \quad \text{hom.,} \quad \det \boldsymbol{A} = \begin{vmatrix} 6 & 2 \\ 3 & 1 \end{vmatrix} = 0 \quad \rightarrow \quad x_2 = -3\,x_1 \quad \rightarrow \quad \begin{array}{c} \text{infinitely many} \\ \text{solution} \end{array}$$

5.11 Eigenvalues and Eigenvectors

A matrix performs a stretch, squeeze and/or rotation of a vector.The vector and matrix elements depend on the choice of the coordinate system. Since this choice is arbitrary, the question arises whether there is a special or canonical representation which is independent of the coordinate system.

There are distinguished directions [eigendirections (eigen $\sim$ self)] along which a matrix operates. Vectors pointing into these directions are only scaled but not rotated.

$$\boldsymbol{A}\vec{v} = \lambda\vec{v} \qquad \lambda \sim \text{eigenvalue} \quad \vec{v} \sim \text{eigenvector}$$

$$\text{or:} \quad \underbrace{(\boldsymbol{A} - \lambda\,\boldsymbol{I})}_{\boldsymbol{B}}\,\vec{v} = 0 \qquad \text{where} \quad \boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ is the identity matrix.}$$

The linear system of equations given by $\boldsymbol{b}\vec{v} = 0$ is homogeneous and has nontrivial solutions $\vec{v} \neq \vec{0}$ only if $\det \boldsymbol{b} = 0$. The condition for non-vanishing eigenvectors is therefore given by

$$\det(\boldsymbol{A} - \lambda\,\boldsymbol{I}) = 0$$

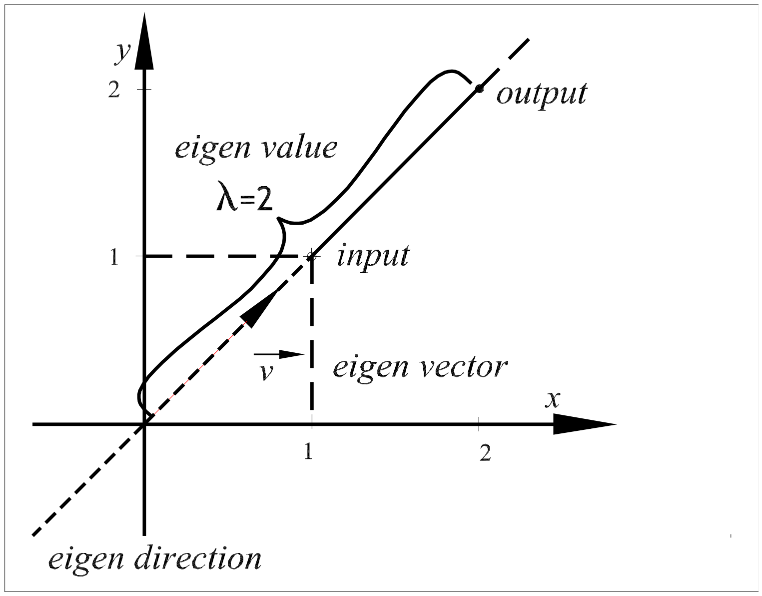


Figure 41: Eigenvalues and eigenvectors

from which the eigenvalues can be readily found. The eigenvector are then determined by solving

$$(\boldsymbol{A} - \lambda\,\boldsymbol{I})\,\vec{v} = 0$$

Examples:

$$\boldsymbol{A} = \begin{pmatrix} 13 & 4 \\ 4 & 7 \end{pmatrix} \qquad \rightarrow \quad \text{eigenvalues:} \quad \det(\boldsymbol{A} - \lambda\,\boldsymbol{I}) = \begin{vmatrix} 13-\lambda & 4 \\ 4 & 7-\lambda \end{vmatrix} = 0$$

$$\rightarrow \quad \lambda^2 - 20\lambda + 75 = 0 \quad \text{(characteristic polynomial)} \qquad \rightarrow \quad \lambda_1 = 15, \quad \lambda_2 = 5$$

$$\lambda_1 = 15: \qquad \begin{array}{rcrcl} 13\,v_1 & + & 4\,v_2 & = & 15\,v_1 \\ 4\,v_1 & + & 7\,v_2 & = & 15\,v_2 \end{array} \qquad \rightarrow \quad v_1 = 2\,v_2$$

$$\rightarrow \quad \text{choose:} \quad v_2 = 1 \quad \rightarrow \quad v_1 = 2 \quad \rightarrow \quad |\vec{v_1}| = \sqrt{5} \quad \Rightarrow \quad \vec{v_1} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 5: \quad \boldsymbol{A}\,\vec{v_2} = 5\,\vec{v_2} \quad \rightarrow \quad v_2 = -2\,v_1 \quad \rightarrow \quad \text{choose:}\ v_1 = 1 \quad \rightarrow \quad v_2 = -2 \quad \Rightarrow \quad \vec{v_2} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Note: A matrix is called *symmetric* if $a_{ij} = a_{ji}$. Symmetric Matrices have real eigenvalues and orthogonal eigenvectors $\vec{v_1} \cdot \vec{v_2} = 2 \cdot 1 + 1 \cdot (-2) = 0$.

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix} \qquad \rightarrow \quad \text{eigenvalues:} \quad \det(\boldsymbol{A} - \lambda\,\boldsymbol{I}) = \begin{vmatrix} 1-\lambda & 0 \\ 2 & 2-\lambda \end{vmatrix} = 0$$

$$\rightarrow \quad (1-\lambda)(2-\lambda) = 0 \quad \text{(characteristic polynomial)} \qquad \rightarrow \quad \lambda_1 = 1, \quad \lambda_2 = 2$$

$$\rightarrow \quad \text{eigenvectors:} \quad \boldsymbol{A}\,\overrightarrow{v} = \lambda\,\overrightarrow{v} \qquad \rightarrow \quad \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

$$\lambda_1 = 1: \qquad \begin{array}{rcl} v_1 & = & v_1 \\ 2\,v_1 + 2\,v_2 & = & v_2 \end{array} \qquad \rightarrow \quad v_1 = -\frac{1}{2} v_2$$

$$\rightarrow \quad \text{choose:} \quad v_2 = 2 \quad \rightarrow \quad v_1 = -1 \quad \rightarrow \quad |\overrightarrow{v_1}| = \sqrt{5} \quad \Rightarrow \quad \overrightarrow{v_1} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

$$\lambda_2 = 2: \quad \boldsymbol{A}\,\overrightarrow{v_2} = 2\,\overrightarrow{v_2} \quad \rightarrow \quad v_1 = 2\,v_1 \quad \rightarrow \quad v_1 = 0 \quad \rightarrow \quad \text{choose:} \; v_2 = 1 \quad \Rightarrow \quad \overrightarrow{v_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

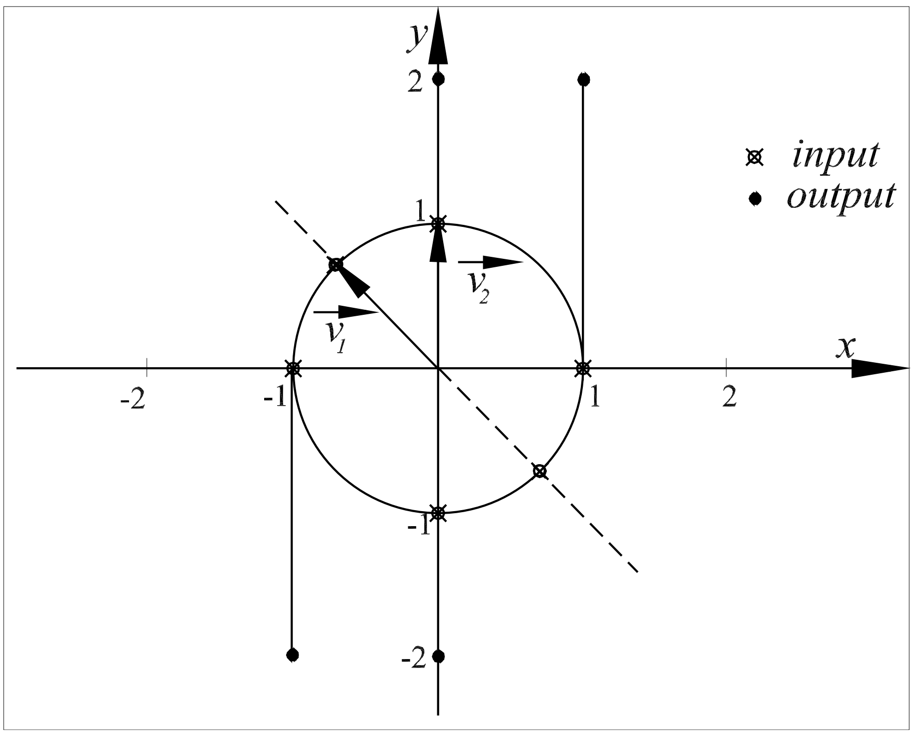


Figure 42: Determining eigenvalues and eigenvectors

6 Complex Number Theory

There are polynomials, such as $f(x) = x^2 + 1$, which do not have a root, $f(x) = 0$ for $x \in \mathbb{R}$. This is one of the reasons to extend the numbers from $x \in \mathbb{R}$ to $z \in \mathbb{C}$ where $\mathbb{C}$ is the set of complex numbers. Beyond this algebraic motivation there are many applications of complex numbers.

6.1 Representations and Basic Properties

A complex number z consists of a pair of real numbers called the real and imaginary part, respectively, and a "new" number 'i' which has the property $i^2 = -1$. While the real numbers can be represented as points on a line, a complex number is given as a point in a plane (called the complex plane) where the coordinates are its real (horizontal axis) and imaginary (vertical axis) part.

$$z = \underbrace{\overbrace{a}^{\text{real}} + i \overbrace{b}^{\text{imaginary}}}_{\text{complex number}} \qquad \text{with} \quad a, b \in \mathbb{R}$$

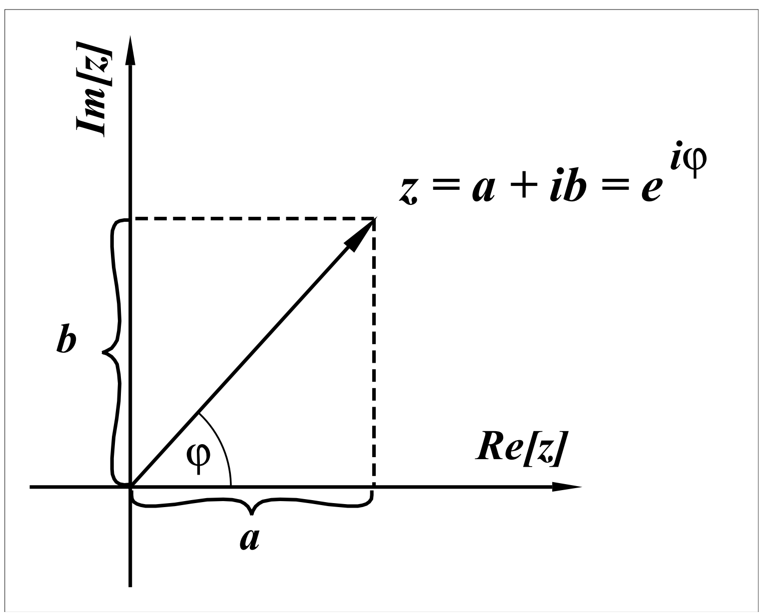


Figure 43: Representation of a complex number

The real and imaginary part, a and b, can be expressed by a distance from the origin r and an angle φ (remember polar coordinates) in terms of $a = r\cos\varphi$ and $b = r\sin\varphi$ which leads to

$$z = a + i\,b = r\,\cos\varphi + i\,r\,\sin\varphi = r\,e^{i\,\varphi} \qquad \text{with} \quad r = \sqrt{a^2 + b^2} \quad \varphi = \arctan\frac{b}{a}$$

Remarks:

- We will proof later by Taylor expansion that the relation $\cos\varphi + i\,\sin\varphi = e^{i\,\varphi}$ is true.
- All basic computations such as addition, subtraction, multiplication or division are defined for $z \in \mathbb{C}$.
- The inverse relation $\varphi = \arctan\frac{b}{a}$ is not unique because $\frac{-b}{-a} = \frac{b}{a}$. However, the first number is the third quadrant whereas the second number is in the first quadrant. Most computer languages, therefore have a second function to calculate the inverse tangent (usually called atan2 or so) which accepts two arguments, i.e. b and a and not only their ratio $\frac{b}{a}$ and returns the correct angle in the $[0, 2\pi]$ or $[-\pi, \pi]$ range.
- The number i is called the imaginary unit and is defined as $i^2 = -1$. It represents a very powerful tool to simplify calculations, in particular when trigonometric functions are involved. From its definition we find readily $i = \pm\sqrt{-1}$, $i^3 = -i$, and $i^4 = 1$.

Rules for dealing with complex numbers: $z = a + i\,b = r\,e^{i\,\varphi}$

$$\begin{aligned}
\text{Addition:} \quad (a_1 + i\,b_1) + (a_2 + i\,b_2) &= a_1 + a_2 + i\,(b_1 + b_2) \\
\text{Multiplication:} \quad (a_1 + i\,b_1)\,(a_2 + i\,b_2) &= a_1\,a_2 - b_1\,b_2 + i\,(a_1\,b_2 + a_2\,b_1) \\
r_1\,e^{i\,\varphi_1}\,r_2\,e^{i\,\varphi_2} &= r_1\,r_2\,e^{i\,(\varphi_1+\varphi_2)}
\end{aligned}$$

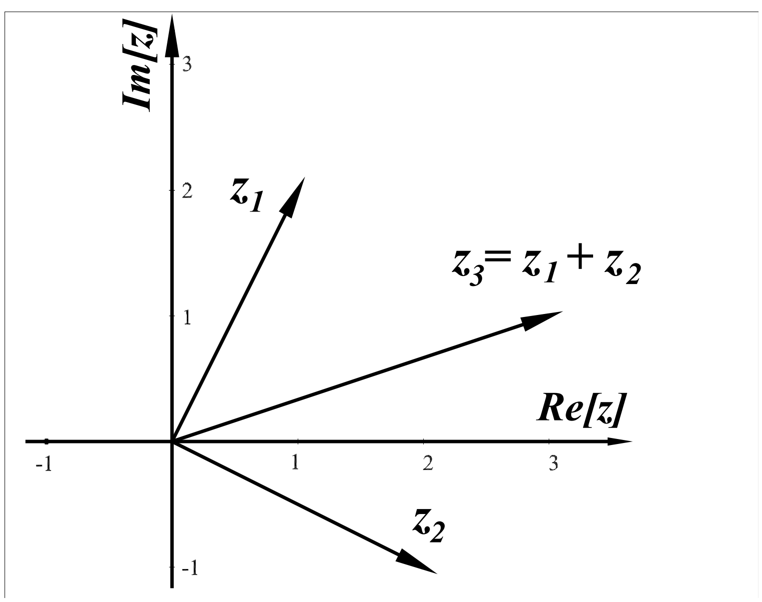


Figure 44: Adding two complex numbers

⇒ Addition of two complex numbers is done by adding their corresponding vectors

⇒ Multiplication of two complex numbers results in the product of the individual amplitudes and the sum of the phases

All the properties of real numbers are still preserved!

Examples: $z_1 = 1 + 2\,i \qquad z_2 = 2 - \,i$

$$z_1 + z_2 = 1 + 2 + i\,(2 - 1) = 3 + i \qquad\qquad z_1\,z_2 = 1 \cdot 2 - 2 \cdot (-1) + i\,(1 \cdot (-1) + 2 \cdot 2) = 4 + 3\,i$$

6.2 Complex conjugate

The complex conjugate of $z = a + i\,b$ is defined as $z^* = a - i\,b$

$$z^* = a + i\,b = r\,e^{-i\,\varphi} \qquad\qquad z\,z^* = (a + i\,b)\,(a - i\,b) = a^2 + b^2 = |\,z\,|^2 \neq z^2$$

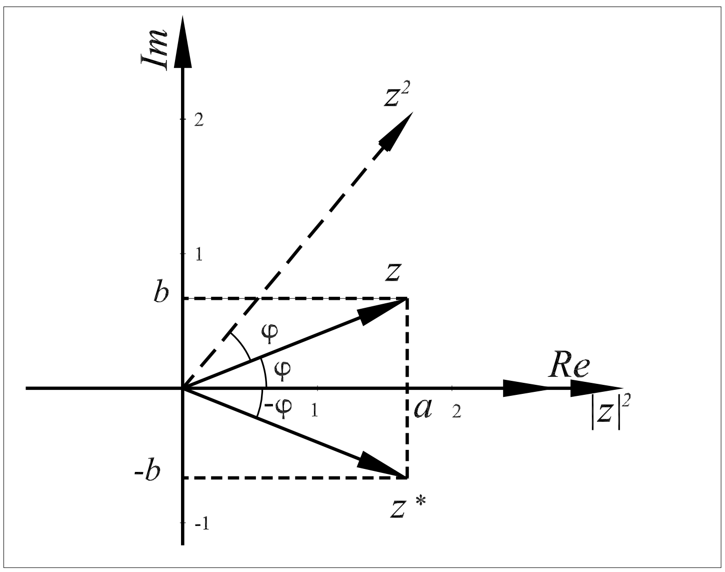


Figure 45: The complex number z and its complex conjugate z^*

Compare the following:

$$\begin{array}{ccccccccccc}
z^2 & = & z\,z & = & (a + i\,b)^2 & = & a^2 - b^2 + 2\,i\,a\,b & = & (r\,e^{i\,\varphi})^2 & = & r^2\,e^{2\,i\,\varphi} \\
|\,z\,|^2 & = & z\,z^* & = & (a + i\,b)\,(a - i\,b) & = & a^2 + b^2 & = & r\,e^{i\,\varphi}\,r\,e^{-i\,\varphi} & = & r^2
\end{array}$$

Some more rules:

$$\text{Complex division:} \quad \frac{z_1}{z_2} = \frac{a_1 + i\,b_1}{a_2 + i\,b_2} = \frac{z_1\,z_2^*}{z_2\,z_2^*} = \frac{z_1\,z_2^*}{|\,z_2\,|^2} = \frac{a_1\,a_2 + b_1\,b_2 + i\,(a_2\,b_1 - a_1\,b_2)}{a_2^2 + b_2^2}$$

$$\mathcal{R}(z) = \frac{1}{2}(z + z^*) = \frac{1}{2}(a + i\,b + a - i\,b) = a \qquad\qquad \mathcal{I}(z) = \frac{1}{2}(z - z^*) = \frac{1}{2}(a + i\,b - (a - i\,b)) = b$$

$$e^{i\,\varphi} = \cos\varphi + i\,\sin\varphi \qquad \Rightarrow \qquad \cos\varphi = \frac{1}{2}(e^{i\,\varphi} + e^{-i\,\varphi}) \qquad \sin\varphi = \frac{1}{2\,i}(e^{i\,\varphi} - e^{-i\,\varphi})$$