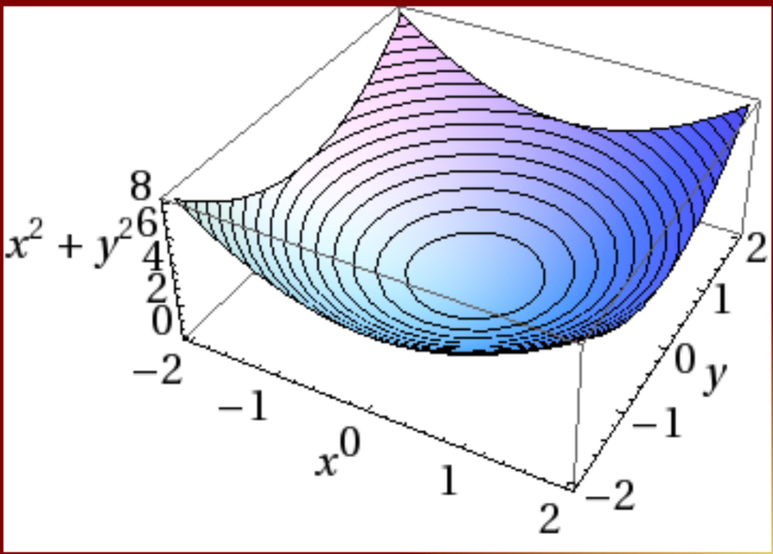
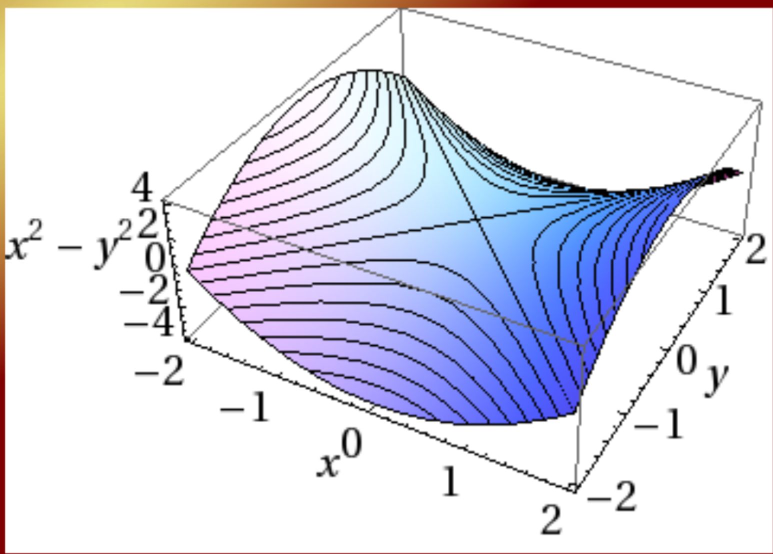
# Mathematics for Physicists II





Tobias Brandes and
Niels Walet

# Mathematics for Physicists II

Tobias Brandes and Niels Walet

Last changed on March 23, 2010

# Contents

# Chapter 1

# Complex Numbers

## 1.1 Basic Properties

### 1.1.1 Introduction

Up to now you have seen 4 major sets of numbers:

- The set $\mathbb{N}$ of natural, $0, 1, 2, ...$ (non-negative integers). The sum and product of two natural numbers are natural numbers.
- The set $\mathbb{Z}$ of all integers, $0, \pm 1, \pm 2....$ This closes under addition, subtraction and multiplication.
- The set $\mathbb{Q}$ of all rational numbers (fractions) $p/q$ like $3/5$, closes under division as well.
- To this we add the irrational numbers, like $\sqrt{2}, 3^{1/3}$
- And obtain $\mathbb{R}$ the set of real numbers includes $1, 2.34, \pi, 4/5, e = e^1$ etc

There is a lot of subtle mathematics associated with them. Are there more rational numbers than integers? More reals than rationals? We can also try to solve equations. In physics we usually mean "find a real number that solves the equation".
**Example 1.1:**

Find the zero of the function (polynomial) $p(x) = x^2 - 1$.

**Solution:**

$$p(x) = x^2 - 1 = 0 \rightsquigarrow x = \pm\sqrt{1} = \pm 1. \tag{1.1}$$

There are equations we can't solve in this way, e.g. the roots or zeroes of the polynomial $q(x) = x^2 + 1$, are not real, i.e.

$$q(x) = x^2 + 1 = 0 \rightsquigarrow x = \pm\sqrt{-1} = ????. \tag{1.2}$$

The square root of $-1$ is not defined within the real numbers. There is no real zero of $q(x)$ (look at the curve $q(x) = x^2 + 1$). We define new numbers (complex numbers) so that we can solve any equation of the kind of (1.2).

Complex numbers are defined as $z = x + iy$ with real $x$ and real $y$ and $i := +\sqrt{-1}$.

The symbol $i$ is called 'complex unit' with the property

$$i^2 = -1. \tag{1.3}$$

Question: What is $(-i)^2$?
Using $i$, we can solve

$$q(x) = x^2 + 1 = 0 \rightsquigarrow x = \pm i. \tag{1.4}$$

In $z = x + iy$, $x = \mathrm{Re}(z)$ is called real part and $y = \mathrm{Im}(z)$ is called imaginary part of the complex number $z$. Complex numbers are elements of a set called $\mathbb{C}$.

So far this may seem a bit artificial. However, we now can solve arbitrary quadratic equations, i.e., find solutions in $\mathbb{C}$:

**Example 1.2:**

Solve $z^2 - 2z + 5 = 0$.

**Solution:**

$$\begin{aligned} z^2 - 2z + 5 &= 0 \\ \rightsquigarrow\ & z_{1/2} = 1 \pm \sqrt{\frac{(-2)^2}{4} - 5} = 1 \pm \sqrt{-4} = 1 \pm \sqrt{4}\sqrt{-1} = 1 \pm 2i. \end{aligned}$$

In fact, within the complex numbers one can always find the root of a quadratic equation (and in fact all the roots of an arbitrary polynomial, i.e. solve equations like $z^{12} + 4z^4 + 17 = 0$ etc., although for this last example there is no general formula as in the quadratic case.)

Actually, complex numbers first arose in the 15th century in the solution of cubic equations of the form $z^3 + bz + c = 0$. The *general* solution of such equations are

$$\begin{aligned} z &= -\left(\frac{2^{\frac{1}{3}}\, b}{\left(-27\,c + \sqrt{108\,b^3 + 729\,c^2}\right)^{\frac{1}{3}}}\right) + \frac{\left(-27\,c + \sqrt{108\,b^3 + 729\,c^2}\right)^{\frac{1}{3}}}{3\,2^{\frac{1}{3}}}, \\ z &= \frac{\left(1 + i\sqrt{3}\right)\, b}{2^{\frac{2}{3}}\left(-27\,c + \sqrt{108\,b^3 + 729\,c^2}\right)^{\frac{1}{3}}} - \frac{\left(1 - i\sqrt{3}\right)\left(-27\,c + \sqrt{108\,b^3 + 729\,c^2}\right)^{\frac{1}{3}}}{6\,2^{\frac{1}{3}}}, \\ z &= \frac{\left(1 - i\sqrt{3}\right)\, b}{2^{\frac{2}{3}}\left(-27\,c + \sqrt{108\,b^3 + 729\,c^2}\right)^{\frac{1}{3}}} - \frac{\left(1 + i\sqrt{3}\right)\left(-27\,c + \sqrt{108\,b^3 + 729\,c^2}\right)^{\frac{1}{3}}}{6\,2^{\frac{1}{3}}}, \end{aligned} \tag{1.5}$$

and therefore inolve the complex unit $i$, even when the resulting solutions are real!

### 1.1.2 Basic Definitions

If we look at a pair of two complex numbers

$$z_1 = x_1 + iy_1, \quad z_2 = x_2 + iy_2, \tag{1.6}$$

we can define all the standard algebraic manipulations.

Equality:

$$z_1 = z_2 \Leftrightarrow x_1 = x_2 \text{ and } y_1 = y_2. \tag{1.7}$$

Addition:

$$z_1 \pm z_2 = (x_1 \pm x_2) + i(y_1 \pm y_2). \tag{1.8}$$

Multiplication:

$$\begin{aligned} z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) = x_1 x_2 + i(x_1 y_2 + x_2 y_1) + i^2 y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1). \end{aligned} \tag{1.9}$$

Zero:

$$z = x + iy = 0 \Leftrightarrow x = 0 \text{ and } y = 0. \tag{1.10}$$

Other:

$$\begin{aligned} z &= x + iy \rightsquigarrow iz = i(x + iy) = ix - y = -y + ix \\ (-i)i &= -(i^2) = -(-1) = 1 \\ i^3 &= i(i^2) = -i \\ i^4 &= (i^2)(i^2) = 1 \\ \frac{1}{i} &= \frac{i}{i^2} = \frac{i}{-1} = -i. \end{aligned} \tag{1.11}$$

The complex conjugate $\bar{z}$ of a complex number $z = x + iy$ is defined as
$z = x + iy \rightsquigarrow \bar{z} = x - iy.$

(Sometimes one writes $z^*$ instead of $\bar{z}$.) We have

$$\mathrm{Re}(z) = \mathrm{Re}(\bar{z}), \quad \mathrm{Im}(z) = -\mathrm{Im}(\bar{z}). \tag{1.12}$$

From this definition, we find

$$\mathrm{Re}(z) = x = \frac{z + \bar{z}}{2}, \quad \mathrm{Im}(z) = y = \frac{z - \bar{z}}{2i}. \tag{1.13}$$

(see next workshop for some exercises).
Division: we use a trick by calculating $z\bar{z}$:

$$z\bar{z} = (x + iy)(x - iy) = x^2 + y^2. \tag{1.14}$$

(Check this !). For $z \neq 0$,

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{x - iy}{x^2 + y^2} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}. \tag{1.15}$$

Other more complicated examples were dealt with in class and the workshop.

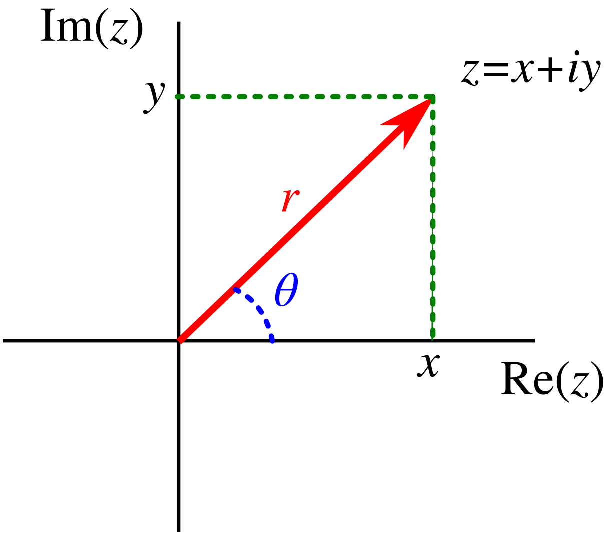



Figure 1.1: Complex number as a vector

## 1.2 Polar Form of Complex Numbers

### 1.2.1 Vector Representation

A complex number $z = x + iy$ has two real components: the real part $\mathrm{Re}(z) = x$ and the imaginary part $\mathrm{Im}(z) = y$. Let us write them in the form of a 2D vector, cf. Fig. 1.1.

$$z = x + iy \leftrightarrow \mathbf{r} = (x, y). \tag{1.16}$$

Polar Form: Any complex number $z = x + iy$ can be written in polar form, (remember that $\boldsymbol{r} = (r\cos\theta, r\sin\theta)$ in polar coordinates)

$$\begin{aligned}
z &= x + iy = r\cos(\theta) + ir\sin(\theta) \\
x &= r\cos(\theta), \quad y = r\sin(\theta) \\
r &= \sqrt{x^2 + y^2} = \sqrt{z\bar{z}} \\
\tan(\theta) &= \frac{y}{x} \rightsquigarrow \theta = \arctan\left(\frac{y}{x}\right).
\end{aligned} \tag{1.17}$$

### 1.2.2 Argument and Modulus

The length of the vector $\mathbf{r} = (x, y)$ representing the complex number $z = x + iy$,

$$|z| := r = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}} \tag{1.18}$$

is called modulus of $z$.

The angle $\theta$,

$$\arg(z) := \theta = \arctan\left(\frac{y}{x}\right) \tag{1.19}$$

is called the argument of $z$. The angle $\theta$ is usually restricted to $-\pi < \theta \leq \pi$, even though any interval of length $2\pi$ will do, and we sometimes use $0 \leq \theta < 2\pi$.

### 1.2.3 Manipulations in Vector/Polar Form

Addition:

$$\begin{aligned} (x,y) &\quad \leftrightarrow \quad z = x+iy \\ (x_1,y_1)+(x_2,y_2) = (x_1+x_2, y_1+y_2) &\quad \leftrightarrow \quad z_1+z_2 = (x_1 \pm x_2)+i(y_1 \pm y_2). \end{aligned} \tag{1.20}$$

Therefore, we have to add the vectors representing the complex numbers.
Complex Conjugate:

$$\begin{aligned} z &= x+iy = r\cos(\theta)+ir\sin(\theta) \rightsquigarrow \\ \bar{z} &= x-iy = r\cos(\theta)-ir\sin(\theta) = r\cos(-\theta)+ir\sin(-\theta). \end{aligned} \tag{1.21}$$

The angle becomes negative: check what this means geometrically!
Multiplication:

$$\begin{aligned} z_1 &= r_1\cos(\theta_1)+ir_1\sin(\theta_1), \quad z_2 = r_2\cos(\theta_2)+ir_2\sin(\theta_2) \rightsquigarrow \\ z_1z_2 &= r_1r_2[\cos(\theta_1)\cos(\theta_2)-\sin(\theta_1)\sin(\theta_2)]+ir_1r_2[\cos(\theta_1)\sin(\theta_2)+\sin(\theta_1)\cos(\theta_2)] \\ &= r_1r_2[\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2)] \end{aligned} \tag{1.22}$$

This means that

$$|z_1z_2| = |z_1||z_2|, \quad \arg(z_1z_2) = \arg(z_1)+\arg(z_2). \tag{1.23}$$

Try to sketch an example for this in the $x-y$ diagram for yourself!

### 1.2.4 Complex exponential

Let us differentiate a complex number of unit modulus, $z = \cos\theta + i\sin\theta$ w.r.t. $\theta$.

$$\frac{d}{d\theta}(\cos\theta+i\sin\theta) = -\sin\theta+i\cos\theta = i(\cos\theta+i\sin\theta). \tag{1.24}$$

In short,

$$\frac{d}{d\theta}z = iz \tag{1.25}$$

This suggests that we can write $z$ as $z = e^{i\theta}$. We shall construct further proof the consistency of this suggestion (which can also be used to define the complex exponent) below.

### 1.2.5 De Moivre's Theorem

We can easily generalise the multiplication of two complex numbers in polar form to calculcate an arbitrary power of $z$, $z^n$ (integer $n$):

$$\begin{aligned} z &= r[\cos(\theta)+i\sin(\theta)] \rightsquigarrow \\ z^2 &= r^2[\cos(2\theta)+i\sin(2\theta)] \rightsquigarrow \\ z^3 &= zz^2 = r^3[\cos(3\theta)+i\sin(3\theta)] \rightsquigarrow \\ &\vdots \\ z^n &= r^n[\cos(n\theta)+i\sin(n\theta)] \rightsquigarrow \end{aligned} \tag{1.26}$$

$$\tag{1.27}$$

Since $z^n = r^n[\cos(\theta)+i\sin(\theta)]^n$, this means

$$[\cos(\theta)+i\sin(\theta)]^n = \cos(n\theta)+i\sin(n\theta) \tag{1.28}$$

which is a useful equation for proving trigonometric identities; it is also useful for doing many integrals that occur in physics.
Example:

$$\begin{aligned}
[\cos(\theta) + i\sin(\theta)]^2 &= [\cos(2\theta) + i\sin(2\theta)] \Leftrightarrow \\
\cos^2(\theta) - \sin^2(\theta) + 2i\cos(\theta)\sin(\theta) &= \cos(2\theta) + i\sin(2\theta) \Rightarrow \\
\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta) \quad &, \quad \sin(2\theta) = 2\cos(\theta)\sin(\theta). \qquad (1.29)
\end{aligned}$$

Clearly, this is entirely consistent with the properties of the exponent: $(e^{i\theta})^n = (e^a)^b$ with $a = i\theta$ and $b = n$ is $e^{ab} = e^{in\theta}$.

## 1.3 The Exponential Function

### 1.3.1 A Power Series for $e^x$

We already know the exponential function $e^x$ for real $x$. We want to generalise it to complex arguments $z$ because it is the most important function in your (physics) life! Let's have a look at $e^x$ again. We can *define* $e^x$ as follows:

The exponential function $f(x) = \exp(x)$ is the unique solution of the first order differential equation $f'(x) = f(x), \quad f(x = 0) = 1$.

This describes, for example, the increase of the number $f$ of animals with time $x$ (strange choice of variable, I know) at unit rate 1, if at time $x = 0$ there was one animal.
**Task:** 1. Think of other, better examples.
2. Can you define for $e^{-x}$?
In this case, there are more and better examples, se last semester's course notes.

Now, the solution of the differential equation above is $f(x) = e^x$, but can we express this in a different manner? Suppose you had no exp–button on your scientific calculator, and you were a survivor on a remote planet with the task to reconstruct mathematics and physics as a part of personkind's knowledge, how would you calculate $e^x$ ? Let's try to write $f(x)$ as a 'polynomial'

$$f(x) = 1 + a_1 x + a_2 x^2 + a_3 x^3 + ..., \qquad (1.30)$$

where we have to determine the constants $a_1$, $a_2$,... from the differential equation:

$$\begin{aligned}
f'(x) &= a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + ... \\
f(x) &= 1 + a_1 x + a_2 x^2 + a_3 x^3 \qquad (1.31)
\end{aligned}$$

We now equate terms of the same power in $x$, since different powers vary in different ways as $x$ changes, so we can disentagle the uinfinite series (cf. equality of polynomials)

$$\rightsquigarrow a_1 = 1, \quad a_2 = \frac{1}{1\cdot 2}, \quad a_3 = \frac{1}{1\cdot 2\cdot 3}, .. \qquad (1.32)$$

We recognise that the general form of this *power series* is

$$f(x) = \exp(x) = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + ..., \quad n! := 1\cdot 2\cdot 3...\cdot n. \qquad (1.33)$$

The symbol $n! := 1\cdot 2\cdot 3...\cdot n$ is called the **factorial**. One defines $0! = 1$. We write the equation for $f(x)$ in another, more condensed and elegant form

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad (1.34)$$

**You have to remember this formula throughout your whole life**. Now, we generalise this 'all–your–life' formula to complex numbers $z$,

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad . \tag{1.35}$$

### 1.3.2 Power Series for $\sin(x)$ and $\cos(x)$

Next to the exponential, certainly sin and cos belong to some of the most important functions in physics. In much the same spirit as above, we can define them by their differential equations, and derive an expression in terms of a series. Here we state the result, that can be used to define sin and cos:

$$\begin{aligned} \sin(x) &:= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + ... \\ \cos(x) &:= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + ... \end{aligned} \tag{1.36}$$

**task:** Calculate the derivatives of $\sin(x)$ and $\cos(x)$ from the series definition above and check that it is consistent! (i.e., $\sin'(x) = \cos(x)$ and $\cos'(x) = -\sin(x)$.

Now we are in a position to test our series for the complex exponential,

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \tag{1.37}$$

We use it for a purely imaginary variable,

$$z = x + iy = iy, \quad (x = 0), \tag{1.38}$$

where $y$ is real (remember that now that we have both real and complex numbers, one always has to state which type of number one is talking about). Remember

$$i^0 = 1, \quad i^1 = i, \quad i^2 = -1, \quad i^3 = -i, \quad i^4 = 1, \quad i^5 = i, ... \tag{1.39}$$

This lead to

$$\exp(iy) = 1 + \frac{iy}{1!} - \frac{y^2}{2!} - \frac{iy^3}{3!} + \frac{y^4}{4!} + \frac{iy^5}{5!} - ... \tag{1.40}$$

The terms are alternately real and imaginary, so we cansplit the series

$$\exp(iy) = \left(1 - \frac{y^2}{2!} + \frac{y^4}{4!} - ...\right) + i\left(\frac{y}{1!} - \frac{y^3}{3!} + \frac{y^5}{5!} + ...\right). \tag{1.41}$$

We compare the right and side of this equation with the series for $\sin(y)$ and $\cos(y)$, we find the important **Euler's Formula**

$$\exp(iy) = \cos(y) + i\sin(y), \quad y \text{ real}. \tag{1.42}$$

Using Euler's formula for the variable $y = \theta$ (angle in our polar representation), we find for any complex number $z$ the representation

$$z = x + iy = r[\cos(\theta) + i\sin(\theta)] = r\exp(i\theta). \tag{1.43}$$

## 1.4 Euler's Formula

We recall Euler's formula

$$\exp(i\theta) = \cos(\theta) + i\sin(\theta) \tag{1.44}$$

and our polar representation of a complex number $z = x + iy$,

$$z = x + iy = r[\cos(\theta) + i\sin(\theta)] = r\exp(i\theta). \tag{1.45}$$

Let us have a look at the polar diagram: $z$ is represented in the **complex plane** $z$**–plane**, i.e. the $x$–$y$–plane, where

$$x = Re(z), \quad y = Im(z). \tag{1.46}$$

$r = \sqrt{x^2 + y^2}$ is the modulus of $z$, i.e. geometrically the length of the vector $(x, y)$. The argument $arg(z) = \theta$ is the angle between the vector $(x, y)$ and the $x$–axis.
Motion along a line: we keep $\theta$ fixed and change $r$ from small to larger values. The corresponding complex numbers move along a straight line that has a fixed angle $\theta$ with the $x$–axis.
Motion along a circle: we keep $r$ fixed and change $\theta$ from 0 to larger values. We recognize that the corresponding complex numbers move along a circle with constant radius $r$.

### 1.4.1 The Unit Circle

For $r = 1$, the complex numbers

$$z(\theta) = \cos(\theta) + i\sin(\theta) = \exp(i\theta) \tag{1.47}$$

are all situated on a circle with radius $r = 1$ around the origin in the complex $z$–plane.
Coming back to the complex exponential, we notice

$$\sin(\theta) = Im(e^{i\theta}), \quad \cos(\theta) = Re(e^{i\theta}), \tag{1.48}$$

where we wrote $e^z$ for $\exp(z)$ (both are the same, this is only a different notation). We obtain sin and cos from the exponential of imaginary argument. What happens if we calculate the exponential for a general complex argument $z = x + iy$ that has both a real and an imaginary part? We calculate

$$e^z = e^{x+iy} = e^x e^i y = e^x[\cos(y) + i\sin(y)] = e^x\cos(y) + ie^x\sin(y). \tag{1.49}$$

Please note that the real part $x$ only determines the modulus $r = e^x$ of $e^z$,

$$|e^z| = |e^{x+iy}| = e^x. \tag{1.50}$$

In particular, this means

$$|e^z| = |e^{x+iy}| = |e^x||e^{iy}| = e^x, \tag{1.51}$$

Since $|e^x| = e^x$ for real $x$, we see that

$$|e^{iy}| = 1. \tag{1.52}$$

The circle is useful to describe, e.g., the motion of a particle on a circle. If we increase the angle $\theta$ linearly as a function of time $t$,

$$\theta = \omega t, \tag{1.53}$$

where $\omega$ is a fixed *angular frequency* (in $\mathrm{s}^{-1}$), we have

$$z(t) = r\exp(i\omega t) = r\cos(\omega t) + ir\sin(\omega t), \tag{1.54}$$

where we allowed an arbitrary radius $r$ again. The real part of $z$ describes the $x$–position, the imaginary part of $z$ describes the $y$–position of the particle on the circle. The complex number $z$ itself is not a measurable quantity, but it contains useful information ($x$ and $y$–position of the particle).

Another important use of the complex exponent is in describing oscillations. This is due the fact that we can calculate the harmonic (sinusoidal) oscillations in terms of a complex exponent, when we argue that the real part is to be taken. The standard place to use such tecniques is in cicuit theory. A voltage is then represented by $V = V_0\cos(\omega t) = \mathrm{Re}(V_0 e^{i\omega t})$, and the real part in the expression will be conveniently forgotten for a while. The response of a capacitor to such an applied voltage, $I = \frac{dQ}{dt} = \frac{d}{dt}CV$ then leads to the complex equation (with $I = I_0 e^{i\omega t}$) $I_0 = i\omega C V_0$, i.e. we have a simple relation between applied potential and current, just like the $V = IR$ relation for a resistor. We can add such complex resistances, and take the rela part of the currents or volatges at the end; the phase of the complex numbers gives the phase difference betwen voltage and current.

### 1.4.2 Roots, $n$–th roots of unity

A number $w$ is called an $n$–th root of a complex number $z$ if $w^n = z$. We write carelessly $w = z^{1/n}$.

The solutions of the equation $z^n = 1$ where $n$ is a positive integer are called **$n$–th roots of unity**. We have

$$z^n = 1 \rightsquigarrow |z^n| = |z|^n = 1 \rightsquigarrow |z| = 1 \rightsquigarrow z = e^{i\theta}, \tag{1.55}$$

therefore we have to solve

$$e^{in\theta} = 1. \tag{1.56}$$

Since

$$e^{2\pi k i} = cos(2\pi k) + i\sin(2\pi k) = 1, \quad k = 0, 1, 2, ... \tag{1.57}$$

we have $n$ different solutions, i.e., $n$ solutions where the phase is between 0 and $2\pi$,

$$z = e^{2k\pi i/n}, k = 0, 1, 2, .., n-1. \tag{1.58}$$

## 1.5 Trigonometric and Hyperbolic Functions

### 1.5.1 Definitions

We recall Euler's formula for the sine and cosine,

$$\exp(ix) = \cos(x) + i\sin(x), \tag{1.59}$$

where $x$ is a real number. From this, we can express sine and cosine as

$$\begin{aligned} \cos(x) &:= \frac{1}{2}\left(e^{ix} + e^{-ix}\right) \\ \sin(x) &:= \frac{1}{2i}\left(e^{ix} - e^{-ix}\right). \end{aligned} \tag{1.60}$$

We now *define* the hyperbolic functions 'hyperbolic cosine' and 'hyperbolic sine' as

$$\begin{aligned} \cosh(x) &:= \frac{1}{2}\left(e^{x} + e^{-x}\right) \\ \sinh(x) &:= \frac{1}{2}\left(e^{x} - e^{-x}\right), \end{aligned} \tag{1.61}$$

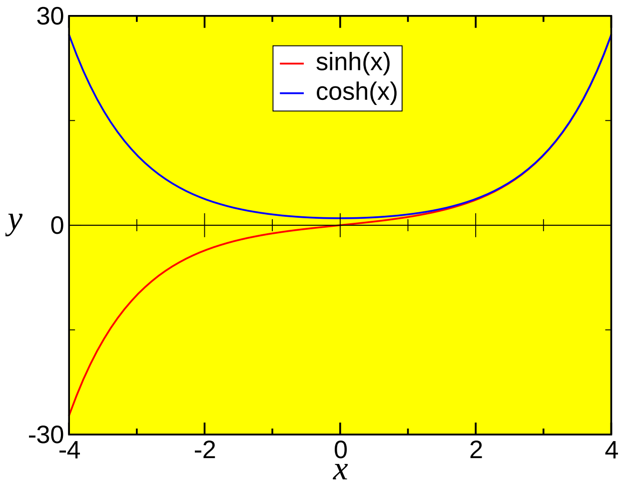



Figure 1.2: Hyperbolic functions

i.e. analogous to cosine and sine but without the imaginary unit $i$. Using $i^2 = -1$, we recognise that

$$\cosh(x) = \cos(ix), \quad \sinh(x) = -i\sin(ix), \tag{1.62}$$

which means that trigonometric and hyperbolic functions are closely related. Their behaviour as a function of $x$, however, is different: while sine and cosine are oscillatory functions, the hyperbolic functions $\cosh(x)$ and $\sinh(x)$ are not oscillatory, because they are just linear combinations of $e^x$ and $e^{-x}$ which are not oscillatory. We have the following properties:

$$\begin{aligned}
\cosh(0) &= 1, \quad \cosh(x) = \cosh(-x) \\
\cosh(x \to \infty) &\to \frac{1}{2}e^x \to_{x\to\infty} \infty, \quad \cosh(x \to -\infty) \to \frac{1}{2}e^{-x} \to_{x\to-\infty} \infty \\
\sinh(0) &= 0, \quad \sinh(x) = -\sinh(-x) \\
\sinh(x \to \infty) &\to \frac{1}{2}e^x \to_{x\to\infty} \infty, \quad \sinh(x \to -\infty) \to -\frac{1}{2}e^{-x} \to_{x\to-\infty} -\infty.
\end{aligned} \tag{1.63}$$

from which we already can sketch the two hyperbolic functions, see Fig. 1.2.

In addition, one defines the hyperbolic tangent and cotangent

$$\tanh(x) := \frac{\sinh(x)}{\cosh(x)}, \quad \coth(x) := \frac{\cosh(x)}{\sinh(x)}. \tag{1.64}$$

### 1.5.2 Inverse hyperbolic functions

Inverting

$$y = \sinh(x) \to x = \sinh^{-1}(y), \tag{1.65}$$

we find the inverse hyperbolic sine $\sinh^{-1}$ by setting

$$y = \frac{e^x - e^{-x}}{2} \rightsquigarrow e^{2x} - 2ye^x - 1 = 0. \tag{1.66}$$

This is a quadratic equation in $u = e^x$ with the solutions

$$u_\pm = y \pm \sqrt{y^2 + 1}. \tag{1.67}$$

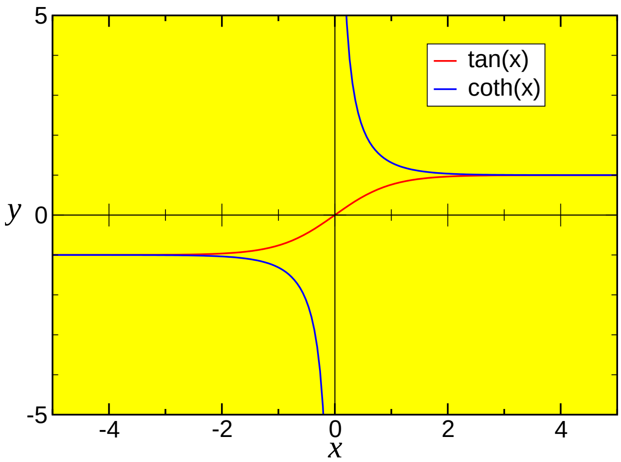



Figure 1.3: Hyperbolic tangent and cotangent.

Since $u = e^x > 0$ is positive, we must take the positive solution $u_+$ and must discard the negative solution $u_-$. Therefore,

$$e^x \equiv u = y + \sqrt{y^2+1} \rightsquigarrow x = \ln\left(y + \sqrt{y^2+1}\right), \tag{1.68}$$

which means that

$$\sinh^{-1}(y) = \ln\left(y + \sqrt{y^2+1}\right). \tag{1.69}$$

Similarly, one obtains

$$\tanh^{-1}(y) = \frac{1}{2}\ln\left[\frac{1+y}{1-y}\right]. \tag{1.70}$$

The $\cosh^{-1}$ is a bit more tricky.

### 1.5.3 Derivatives

These are obtained by going back to the definitions of the hyperbolic functions.

$$\sinh'(x) = \cosh(x), \quad \cosh'(x) = \sinh(x), \quad \tanh'(x) = 1 - \tanh^2(x). \tag{1.71}$$

### 1.5.4 Hyperbolic Identities

These also are obtained by using the definitions of cosh and sinh:

$$\begin{aligned} \cosh^2(x) - \sinh^2(x) &= 1, \quad \cosh(2x) = 1 + 2\sinh^2(x) \\ \sinh(2x) &= 2\sinh(x)\cosh(x). \end{aligned} \tag{1.72}$$

# Chapter 2

# Second Order Linear Differential Equations

## 2.1 Ordinary 2nd Order Linear Differential Equations

### 2.1.1 Origin of Differential Equations: the Harmonic Oscillator as an Example

We consider a particle of mass $m$ that is moving along a straight line in $x$–direction. At time $t$, its coordinate is $x = x(t)$. It is attached to springs with spring constant $k > 0$ so that there is a 'restoring' force $f_r(x) = -kx$ acting on the particle. At $x = 0$, the mass is in equilibrium and no force is acting. In addition, there is a friction force $f_f(v) = -\gamma v$ acting on the particle which is proportional (with friction constant $\gamma > 0$) to its velocity $v = \dot{x}(t)$, and an external force $f_e(x)$ that could have its origin in, e.g., some crazy experimentalist fiercly forcing the mass to follow her hand.

Newton's law states that $m\ddot{x}(t)$ equals the sum $f_r(x) + f_f(x) + f_e(x)$ of all forces on the particle, i.e.

$$\begin{aligned} m\ddot{x}(t) &= -kx - \gamma\dot{x}(t) + f_e(x) \Leftrightarrow \\ \ddot{x}(t) + \frac{\gamma}{m}\dot{x}(t) + \frac{k}{m}x(t) &= \frac{1}{m}f_e(x), \quad k > 0, \gamma > 0. \end{aligned} \tag{2.1}$$

To find the position $x$ of the particle at time $t$, i.e. the function $x(t)$, we have to solve the **differential equation of the forced, damped linear harmonic oscillator**, Eq. (2.1). Learn this standard form of the forced damped harmonic oscillator by heart and it will save you from much misery in the future.
CHECK: to which forces do the terms 'forced', 'damped', and 'harmonic' refer ?

Is this a well–defined task? No, in order to know $x(t)$ at all times later than, say, $t = 0$, we must specify the **initial conditions**, i.e. the initial position of the particle $x(t = 0)$ and its initial velocity $\dot{x}(t = 0)$.

Eq. (2.1)is called **2nd order differential equation** because the highest derivative appearing is a second derivative. Because Newton's law (for a general force) leads to second derivatives (acceleration term!), 2nd order differential equations belong to the most important differential equations in physics.

Eq. (2.1) is called **linear** because we don't have terms like $\ddot{x}^2(t)$ or $x^4(t)$. In general and in more complicated cases (e.g., motion in three dimensions), such terms can leads to **chaos**. The study of differential equations therefore is of paramount importance in order to understand chaos.

Eq. (2.1) is called **Ordinary** because the desired function $x$ is a function of one variable ($t$) only and not more than one variable, in which case differential equations are called partial differential equations.

### 2.1.2 Definitions

In the mathematic literature, people sometimes don't care about the physical background of equations and introduce other notations. In the following, instead of $x(t)$, $\dot{x}(t)$ etc we discuss differential equations for functions $y(x)$ of one variable $x$, with $y'(x)$ denoting the first and $y''(x)$ the second derivative, respectively.

**A 2nd order inhomogeneous linear differential equation for the function** $y(x)$ has the form

$$y''(x) + p(x)y'(x) + q(x)y(x) = f(x), \tag{2.2}$$

where $p(x)$, $q(x)$, and $f(x)$ are known functions of $x$ and $y(x)$ is the function one would like to calculate.

In general, there is no method to obtain a solution $y(x)$ of Eq. (2.2 that could be written down in a simple form, such as $y(x) = \sin(x)$ etc.

**A 2nd order homogeneous linear differential equation for the function** $y(x)$ has the form

$$y''(x) + p(x)y'(x) + q(x)y(x) = 0, \tag{2.3}$$

i.e. the term $f(x)$ is zero on the r.h.s. of Eq.(2.2).

**A 2nd order inhomogeneous linear differential equation for the function** $y(x)$ **with constant coefficients** has the form

$$y''(x) + py'(x) + qy(x) = f(x), \tag{2.4}$$

where $p$ and $q$ are real numbers, $f(x)$ is a known function of $x$, and $y(x)$ is the function one would like to calculate.

**A 2nd order homogeneous linear differential equation for the function** $y(x)$ **with constant coefficients** has the form

$$y''(x) + py'(x) + qy(x) = 0, \tag{2.5}$$

where $p$ and $q$ are real numbers, and $y(x)$ is the function one would like to calculate.

**Initial Value Problem for 2nd order differential equation for a function** $y(x)$: To solve the initial value problem for a 2nd order differential equation for a function $y(x)$ means to solve $y(x)$ for the specific, given **initial conditions**

$$y(x = x_0) = y_0, \quad y'(x = x_0) = y'_0. \tag{2.6}$$

In the example of our harmonic oscillator this means that we start the motion at $t = t_0 = 0$ at the initial position $x(t_0) = x_0$ with the initial velocity $\dot{x}(t_0) = \dot{x}_0$.

### 2.1.3 How to Solve Them

In general, there is no recipe or general method of how to solve a given differential equation. In this lecture, we only discuss the 2nd order inhomogeneous linear differential equation for the function $y(x)$ with constant coefficients, for which there is a general method. 'Differential Equations' is a difficult topic, and still today a research subject in mathematics. Generations of people have tried to solve differential equations by finding new exact solutions, developing approximation techniques etc. For example, a big problem in Einstein's theory of gravitation is that the fundamental (partial) differential equations are known, but only very few exact solutions are known. This is still a hot topic today.

To warm up a bit, we solve a few simple cases of Eq.(2.1).

EXAMPLE: a particle of mass $m$ under a constant external force $f_e(x) = f_e$ that does not depend on $x$. We have

$$\begin{aligned} \ddot{x}(t) &= \frac{1}{m} f_e \rightsquigarrow \dot{x}(t) = \frac{1}{m} f_e t + \dot{x}(0) \rightsquigarrow \\ x(t) &= \frac{1}{2m} f_e t^2 + \dot{x}(0)t + x(0). \end{aligned} \tag{2.7}$$

Here, the values $x(0)$ and $\dot{x}(t = 0)$ determine the initial condition at $t = 0$.

CHECK: go back to Pisa (Galilei) and establish the relation between this equation and the experiment of a freely falling mass $m$. In a 'Gedankenexperiment' (thought experiment), change the initial conditions $\dot{x}(t = 0)$ and $x(0)$ and discuss what changes then. What does a positive or a negative $f_e$ mean?

## 2.2 2nd order homogeneous linear differential equations with constant coefficients I

We recall that this type of equation has the form

$$y''(x) + py'(x) + qy(x) = 0, \tag{2.8}$$

where $p$ and $q$ are real numbers, and $y(x)$ is the function one would like to calculate. An example is the differential equation of the damped linear harmonic oscillator

$$\ddot{x}(t) + \frac{\gamma}{m}\dot{x}(t) + \frac{k}{m}x(t) \quad = \quad 0, \quad k > 0, \gamma > 0, \tag{2.9}$$

cf. Eq.(2.1).

### 2.2.1 Undamped oscillator

Consider

$$y''(x) + qy(x) = 0, q > 0.$$

This is the case $p = 0$ of Eq. (2.8). An example for this is the differential equation of the undamped linear harmonic oscillator

$$\ddot{x}(t) + \frac{k}{m}x(t) \quad = \quad 0, \tag{2.10}$$

where $k > 0$ here, cf. Eq.(2.1). From our physical intuition, we know that the mass point described by Eq.(2.10) performs oscillations at an **angular frequency** $\omega$. Therefore, we try sin and cos functions as solution: If we write

$$\begin{aligned} x(t) &= x_1 \sin(\omega t) \rightsquigarrow \dot{x}(t) = x_1 \omega \cos(\omega t) \\ \rightsquigarrow \ddot{x}(t) &= -x_1 \omega^2 \sin(\omega t) = -\omega^2 x(t). \end{aligned} \tag{2.11}$$

Here, $x_1$ is an arbitrary constant. The function $x(t) = x_1 \sin(\omega t)$ fulfills the differential equation Eq. (2.10), if

$$\omega^2 = \frac{k}{m}. \tag{2.12}$$

If on the other hand we write

$$\begin{aligned} x(t) &= x_2 \cos(\omega t) \rightsquigarrow \dot{x}(t) = -x_2 \omega \sin(\omega t) \\ \rightsquigarrow \ddot{x}(t) &= -x_2 \omega^2 \cos(\omega t) = -\omega^2 x(t), \end{aligned} \tag{2.13}$$

we again recognise that the function$x(t) = x_2 \cos(\omega t)$ fulfills the differential equation Eq. (2.10), if $\omega^2 = k/m$ (same as before). Again, $x_2$ is an arbitrary constant. Therefore, we find **two solutions of the second order differential equation** Eq. (2.10). Now we are a bit confused. Let us summarize what we have found so far, using our 'mathematical notation',

$$\begin{aligned} y''(x) + qy(x) &= 0, \quad q > 0 \rightsquigarrow \\ y(x) &= y_1(x) = y_1 \sin(\sqrt{q}x), \quad y(x) = y_2(x) = y_2 \cos(\sqrt{q}x). \end{aligned} \tag{2.14}$$

We now make an important observation:
**THEOREM: With two solutions $y_1(x)$ and $y_2(x)$ of a linear homogeneous differential equation, also the sum $y_1(x) + y_2(x)$ is a solution of the linear homogeneous differential equation**.

PROOF:

$$\begin{aligned} y_1''(x) + p(x)y_1'(x) + q(x)y_1(x) = 0, y_2''(x) + p(x)y_2'(x) + q(x)y_2(x) &= 0 \rightsquigarrow \\ \left[y_1''(x) + y_2''(x)\right] + p(x)\left[y_1'(x) + y_2'(x)\right] + q(x)\left[y_1(x) + y_2(x)\right] &= 0 \rightsquigarrow \\ \left[y_1 + y_2\right]''(x) + p(x)\left[y_1 + y_2\right]'(x) + q(x)\left[y_1(x) + y_2(x)\right] &= 0. \end{aligned}$$

We have used the fact that the sum of the derivatives of two functions is the derivative of the sum of the functions.

The **general solution** of $y''(x) + qy(x) = 0, q > 0$ can be written as the sum

$$y''(x) + qy(x) = 0 \rightsquigarrow y(x) = y_1 \sin(\sqrt{q}x) + y_2 \cos(\sqrt{q}x). \tag{2.15}$$

### 2.2.2 Initial Value Problem

Here we consider

$$\ddot{x}(t) + \omega^2 x(t) = 0,$$

the equation of the undamped linear harmonic oscillator. Note that we write $x(t)$ instead of $y(x)$ here. We have found the general solution as

$$x(t) = x_1 \sin(\omega t) + x_2 \cos(\omega t), \tag{2.16}$$

where $x(t)$ is the position $x$ at time $t$. As mentioned above, in order to know $x(t)$ at all times later than, say, $t = 0$, we must specify the **initial conditions**, i.e. the initial position of the particle $x_0 = x(t = 0)$ and its initial velocity $v_0 = \dot{x}(t = 0)$, i.e.

$$\begin{aligned} x_0 = x(t=0) &= x_1 \sin(\omega 0) + x_2 \cos(\omega 0) = x_2 \\ v_0 = \dot{x}(t=0) &= x_1 \omega \cos(\omega t) - x_2 \omega \sin(\omega t)|_{t=0} = x_1 \omega. \end{aligned} \tag{2.17}$$

Therefore, we can express the parameters $x_1$ and $x_2$ by the given initial values $x_0$ and $v_0$ and obtain

$$x(t) = \frac{v_0}{\omega} \sin(\omega t) + x_0 \cos(\omega t). \tag{2.18}$$

### 2.2.3 Exponential

Consider

$$y''(x) + qy(x) = 0, q < 0.$$

We notice that for $q < 0$ the argument in the sin and cos in Eq.(2.15) becomes imaginary since $\sqrt{q} = \sqrt{-|q|} = i\sqrt{|q|}$ for $q < 0$. Let us find a solution by recalling that the exponential function $f(x) = \exp(x)$ fulfills

$$f(x) = e^x \rightsquigarrow f'(x) = e^x \rightsquigarrow f''(x) = e^x \rightsquigarrow ... \tag{2.19}$$

More generally, we have

$$\begin{aligned} f(x) &= e^{\lambda x} \rightsquigarrow f'(x) = \lambda e^{\lambda x} \rightsquigarrow f''(x) = \lambda^2 e^{\lambda x} \rightsquigarrow f''(x) = \lambda^2 f(x) \qquad (2.20) \\ f(x) &= e^{-\lambda x} \rightsquigarrow f'(x) = -\lambda e^{-\lambda x} \rightsquigarrow f''(x) = (-\lambda)^2 e^{-\lambda x} \rightsquigarrow f''(x) = \lambda^2 f(x). \end{aligned}$$

Comparing this to our differential equation,

$$y''(x) - |q|y(x) = 0 \Leftrightarrow y''(x) = |q|y(x), \tag{2.21}$$

we recognize by comparing with Eq. (2.20) that two independent solutions of Eq. (2.21) are

$$\begin{aligned} y''(x) - |q|y(x) &= 0, \quad q \neq 0 \rightsquigarrow \\ y_1(x) &= y_1 e^{\sqrt{|q|}x}, \quad y_2(x) = y_2 e^{-\sqrt{|q|}x}. \end{aligned} \tag{2.22}$$

As above, the most general solution again is the sum of these two, i.e. the **linear combination** of $e^{-\sqrt{|q|}x}$ and $e^{\sqrt{|q|}x}$ with the two independent constants $y_1$ and $y_2$,

$$\begin{aligned} y''(x) - |q|y(x) &= 0 \rightsquigarrow \\ y(x) &= y_1 e^{\sqrt{|q|}x} + y_2 e^{-\sqrt{|q|}x}. \end{aligned} \tag{2.23}$$

### 2.2.4 Summary

We summarize the two pairs of solutions for $q > 0$ and $q = -|q| < 0$ of $y''(x) + qy(x) = 0$ in the following table:

| $y''(x) + qy(x) = 0, \quad q = k^2 > 0$ | $y''(x) + qy(x) = 0, \quad q = -\kappa^2 < 0$ |
|---|---|
| two solutions<br>$y_1(x) = y_1 \sin(kx), y_2(x) = y_2 \cos(kx)$ | two solutions<br>$y_1(x) = y_1 e^{\kappa x}, y_2(x) = y_2 e^{-\kappa x}$ |
| general solution<br>$y(x) = y_1 \sin(kx) + y_2 \cos(kx)$ | general solution<br>$y(x) = y_1 e^{\kappa x} + y_2 e^{-\kappa x}$ |
| character:<br>oscillatory (sin and cos) | character:<br>exponential (decreasing and incr.) |

**Note that the sign of $q$ makes all the difference!**

## 2.3 2nd order homogeneous linear differential equations with constant coefficients II

Now we attack the case of arbitrary $p$ and $q$ in our differential equation

$$y''(x) + py'(x) + qy(x) = 0. \tag{2.24}$$

Remember that for $p > 0$ and $q > 0$ this corresponds to the differential equation Eq. (2.1) of the damped linear harmonic oscillator. We already know that this system performs oscillations ($\rightsquigarrow$ sin, cos) that can be exponentially damped ($\rightsquigarrow$ exp). Therefore, we expect something related to sin, cos, exp functions. But these are all related to each other if we recall what we have learned about complex numbers:

$$\begin{aligned} \exp(ix) &= \cos(x) + i\sin(x), x \text{ real} \\ \cos(x) &= \frac{e^{ix} + e^{-ix}}{2} = \mathrm{Re}[e^{ix}] \\ \sin(x) &= \frac{e^{ix} - e^{-ix}}{2i} = \mathrm{Im}[e^{ix}]. \end{aligned} \tag{2.25}$$

Furthermore, for arbitrary complex $z = x + iy$,

$$e^z = e^{x+iy} = e^x e^{iy} = e^x[\cos(y) + i\sin(y)] = e^x \cos(y) + ie^x \sin(y). \tag{2.26}$$

The function $e^z$ with complex $z$ comprises the real exponential as well as sin and cos.

Let us therefore try an **exponential Ansatz** in Eq. (2.24),

$$y(x) = e^{zx} \rightsquigarrow y''(x) + py'(x) + qy(x) = [z^2 + pz + q]e^{zx} = 0. \tag{2.27}$$

We recognize that $y(x) = e^{zx}$ fulfills the differential equation, if the bracket [...] is zero:

$$[z^2 + pz + q] = 0. \tag{2.28}$$

This is a quadratic equation which in general has two solutions,

$$z^2 + pz + q = 0 \rightsquigarrow z_{1/2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \tag{2.29}$$

Various cases arrise for different signs of the argument of the square root (the discriminant). Let us look at each of these cases in turn.

### 2.3.1 Positive discriminant

In the case $\frac{p^2}{4} - q > 0$,

$$z_{1/2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} \tag{2.30}$$

are both real and the two solutions fulfilling Eq. (2.24) are

$$\frac{p^2}{4} - q > 0 \rightsquigarrow y_1(x) = y_1 e^{[-\frac{p}{2}+\sqrt{\frac{p^2}{4}-q}]x}, \quad y_2(x) = y_2 e^{[-\frac{p}{2}-\sqrt{\frac{p^2}{4}-q}]x} \tag{2.31}$$

The general solution is the linear combination of the two,

$$\frac{p^2}{4} - q > 0 \rightsquigarrow y(x) = y_1 e^{[-\frac{p}{2}+\sqrt{\frac{p^2}{4}-q}]x} + y_2 e^{[-\frac{p}{2}-\sqrt{\frac{p^2}{4}-q}]x}. \tag{2.32}$$

In this case there are no oscillations at all. The 'damping term' $py'(x)$ is too strong.

### 2.3.2 negative discriminant

In the case $\frac{p^2}{4} - q < 0$, the two zeros become complex:

$$z_{1/2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} = -\frac{p}{2} \pm i\sqrt{q - \frac{p^2}{4}} =: -\frac{p}{2} \pm i\Omega, \tag{2.33}$$

where we define an angular frequency $\Omega = \sqrt{q - p^2/4}$. Now, the two solutions fulfilling Eq. (2.24) are

$$\frac{p^2}{4} - q < 0 \rightsquigarrow y_1(x) = y_1 e^{[-\frac{p}{2}+i\Omega]x}, \quad y_2(x) = y_2 e^{[-\frac{p}{2}-i\Omega]x}, \quad \Omega := \sqrt{q - \frac{p^2}{4}}. \tag{2.34}$$

The general solution is the linear combination of the two,

$$\frac{p^2}{4} - q < 0 \rightsquigarrow y(x) = y_1 e^{[-\frac{p}{2}+i\Omega]x} + y_2 e^{[-\frac{p}{2}-i\Omega]x}, \quad \Omega := \sqrt{q - \frac{p^2}{4}} - q. \tag{2.35}$$

We rewrite this as

$$\begin{aligned}
y(x) &= y_1 e^{[-\frac{p}{2}+i\Omega]x} + y_2 e^{[-\frac{p}{2}-i\Omega]x} = e^{-px/2}\left\{y_1 e^{i\Omega x} + y_2 e^{-i\Omega x}\right\} \\
&= e^{-px/2}\left\{y_1[\cos(\Omega x) + i\sin(\Omega x)] + y_2[\cos(\Omega x) - i\sin(\Omega x)]\right\} \\
&= e^{-px/2}\left\{[y_1 + y_2]\cos(\Omega x) + i[y_1 - y_2]\sin(\Omega x)\right\}.
\end{aligned} \tag{2.36}$$

Now, this seems a bit odd since we have got a complex solution due to the term $i(y_1 - y_2)$. However, the constant coefficients $y_1$ and $y_2$ can be complex anyway (and still $y(x)$ is a solution of the differential equation). If we are only interested in real functions $y(x)$, we can re–define new constants $c_1 := y_1 + y_2$ and $c_2 := i[y_1 - y_2]$ such that the general solution becomes

$$\begin{aligned} y''(x) + py'(x) + qy(x) &= 0, \quad \frac{p^2}{4} - q < 0 \rightsquigarrow \\ y(x) &= e^{-px}\left\{c_1 \cos(\Omega x) + c_2 \sin(\Omega x)\right\}. \end{aligned} \tag{2.37}$$

Still $c_1$ and $c_2$ could be complex numbers, but we can choose them real if we only want real functions $y(x)$.

### 2.3.3 The Marginal Case

For the marginal case $\frac{p^2}{4} - q = 0$ we have only one solution, $z = -p/2$, and thus only one integration constant. This is clearly insufficient! Help can be found in the expression for the simplest marginal case,

$$y''(x) = 0$$

which has as solution

$$y(x) = a + bx.$$

We therefore substitute $y(x) = (a + bx)\exp(-px/2)$ and find

$$\begin{aligned} y(x) &= (a + bx)\exp(-px/2), \\ y'(x) &= b\exp(-px/2) - \frac{p}{2}(a + bx)exp(-px/2), \\ y''(x) &= -bp\exp(-px/2) + \frac{p^2}{4}(a + bx)exp(-px/2). \end{aligned} \tag{2.38}$$

Using the differential equation $y''(x) + py'(x) + qy(x) = 0$, we find that

$$y''(x) + py'(x) + qy(x) = \left(\frac{p^2}{4} - \frac{p^2}{2} + q\right)(a + bx)exp(-px/2) - bp\exp(-px/2) + bp\exp(-px/2) = 0 \tag{2.39}$$

Here we have used $q = 4p^2$. Thus

$$\boxed{y(x) = (a + bx)\exp(-px/2).}$$

### 2.3.4 Summary

Solutions of $y''(x) + py'(x) + qy(x) = 0$:

$$\begin{aligned} &p^2 > 4q \quad y(x) = \exp(-\frac{p}{2}x)\left[y_1 \exp\left(\sqrt{\frac{p^2}{4} - q}x\right) + y_2 \exp\left(-\sqrt{\frac{p^2}{4} - q}x\right)\right] \\ &p^2 < 4q \quad y(x) = \exp(-\frac{p}{2}x)\left[c_1 \cos\left(\sqrt{q - \frac{p^2}{4}}x\right) + c_2 \sin\left(\sqrt{q - \frac{p^2}{4}}x\right)\right] \\ &p^2 = 4q \quad y(x) = \exp(-px/2)(a + bx) \end{aligned}$$

## 2.4 Inhomogeneous Equations

Now we arrive at the most general case we treat here, the **second order inhomogeneous linear differential equation for the function $y(x)$ with constant coefficients**

$$y''(x) + py'(x) + qy(x) = f(x), \tag{2.40}$$

where $p$ and $q$ are real numbers, $f(x)$ is a known function of $x$, and $y(x)$ is the function one would like to calculate. In the following, we become a bit more 'physical' and discuss the differential equation of the forced, damped linear harmonic oscillator, Eq. (2.1),

$$\ddot{x}(t) + 2\gamma\dot{x}(t) + \omega^2 x(t) \quad = \quad \frac{1}{m}f(x), \quad \gamma > 0. \tag{2.41}$$

instead of Eq. (2.40). Since this means that both $p > 0$ and $q > 0$ in Eq. (2.40), we are not very general. Similar results can be obtained for the the general case.

### 2.4.1 Solution to inhomogeneous equations

For inhomogeneous equations the superposition principle is violated in a specail manner: we can easily show that the general solution of the differential equation can be written as the sum of a general solution of the related homogeneous equation, and a special solution to the inhomogeneous one. In other words if $y_{\text{special}}(x)$ satisfies the inhomogeneous equation $y''(x) + py'(x) + qy(x) = f(x)$, then $y_{\text{hom}}(x) + y_{\text{special}}(x)$ satisfies this same equation if $y_{\text{hom}}(x)$ satisfies the equation $y''(x) + py'(x) + qy(x) = 0$. This can be checked easily

$$\begin{gathered} y''_{\text{hom}}(x) + y''_{\text{special}}(x) + p(y'_{\text{hom}}(x) + y'_{\text{special}}(x)) + q(y_{\text{hom}}(x) + y_{\text{special}}(x)) = \\ y''_{\text{hom}}(x) + py'_{\text{hom}}(x) + qy_{\text{hom}}(x) + y''_{\text{special}}(x) + py'_{\text{special}}(x) + q + y_{\text{special}}(x) = 0 + f(x) \end{gathered} \tag{2.42}$$

The art of the exercise is thus in finding a special solution of the inhomogeneous problem.

## 2.5 * Green's function approach

### 2.5.1 * Initial Conditions for the Homogeneous Case

The solution for the homogeneous equation $f \equiv 0$ was obtained above in Eq. (2.40),

$$\begin{aligned} y_h(x) &= e^{-\frac{p}{2}x}\left\{c_1\cos(\Omega x) + c_2\sin(\Omega x)\right\}, \\ x_h(t) &= e^{-\gamma t}\left\{x_1\cos(\Omega t) + x_2\sin(\Omega t)\right\}, \end{aligned} \tag{2.43}$$

with $\Omega = \sqrt{\omega^2 - \gamma/2}$.

Specifying the **initial conditions**

$$x_h(t=0) = x_0, \quad \dot{x}_h(t=0) = v_0, \tag{2.44}$$

we find

$$x_h(t) \quad = \quad x_0\left\{e^{-\gamma t}\cos(\omega t) + \frac{\delta}{\omega}e^{-\gamma t}\sin(\omega t)\right\} + v_0 e^{-\gamma t}\frac{\sin(\omega t)}{\omega}. \tag{2.45}$$

Thus $x_h(t)$ describes the motion of the harmonic oscillator for $f \equiv 0$ (homogeneous case). If we choose the initial time $t = t_0$ instead of $t = 0$, we have

$$\begin{aligned} x_h(t) &= x_0\left\{e^{-\gamma[t-t_0]}\cos(\omega[t-t_0]) + \frac{\delta}{\omega}e^{-\gamma[t-t_0]}\sin(\omega[t-t_0])\right\} \\ &+ v_0 e^{-\gamma[t-t_0]}\frac{\sin(\omega[t-t_0])}{\omega}, \end{aligned} \tag{2.46}$$

i.e. everything remains the same; only the 'origin' of time $t_0$ is shifted, i.e the time scale is shifted by $t_0$.

---

EXERCISE: check that Eq. (2.45) fulfills the correct initial conditions!

---

### 2.5.2 * The Inhomogeneous Case: Effect of the External Force

Now let us discuss the additional effect of the external force, i.e. the inhomogeneous term $f(t)/m$ in Eq. (2.41). First of all, we recognize that $f(t)/m$ is an additional acceleration, $a(t) = f(t)/m$, of the mass $m$ due to the force $f(t)$, using Newton's law. What is the additional displacement, $\Delta x(t)$, of the mass due to that acceleration? In a very short time interval from time $t = t'$ to $t = t' + \delta t'$, due to the acceleration $a(t')$ the mass aquires the additional velocity

$$v(t') = a(t')\delta t' = \frac{f(t')}{m}\delta t'. \tag{2.47}$$

The subsequent additional displacement $\Delta x(t > t')$ has to be proportional to that additional velocity and can be calculated using Eq.(2.46) with 'initial' additional shift $x_0 = 0$ and 'initial' additional velocity $v_0 = v(t')$,

$$\begin{aligned} \Delta x(t > t') &= e^{-\gamma[t-t']}\frac{\sin(\omega[t-t'])}{\omega} \times v(t') \\ &= e^{-\gamma[t-t']}\frac{\sin(\omega[t-t'])}{\omega} \times \frac{f(t')}{m}\delta t', \\ &=: G(t-t') \times \frac{f(t')}{m}\delta t', \end{aligned} \tag{2.48}$$

where in the last line we introduced an abbreviation for the term $e^{-\gamma[t-t']}\sin(\omega[t-t'])/\omega$. The function $G(t-t')$ is called **response function** (**Green's function**) of the harmonic oscillator since it describes its response to an additional, infinitesimal acceleration $f(t')\delta t'/m$. Note that we have made no additional assumptions on how this force $f(t')$ actually behaves as a function of time.

The total additional shift $x_f(t)$ at time $t$ can be calculated from Eq.(2.48) by integrating the contributions from all times $t'$ with $t_0 < t' < t$,

$$x_f(t) = \int_{t_0}^{t} dt' \Delta x(t > t') = \int_{t_0}^{t} dt' G(t-t')\frac{f(t')}{m}. \tag{2.49}$$

The position $x(t)$ at time $x$ now is given by the contribution $x_h(t)$ (force $f = 0$) plus the additional shift $x_f(t)$ (force $f \neq 0$),

$$x(t) = x_h(t) + x_f(t) = x_h(t) + \int_{t_0}^{t} dt' G(t-t')\frac{f(t')}{m}. \tag{2.50}$$

Putting everything together, we find a somewhat lengthy, but very convincing expression (we set the initial time $t_0 = 0$ for simplicity),

$$\begin{aligned} x(t) &= x_0\left\{e^{-\gamma t}\cos(\omega t) + \frac{\delta}{\omega}e^{-\gamma t}\sin(\omega t)\right\} + v_0 e^{-\gamma t}\frac{\sin(\omega t)}{\omega} \\ &+ \int_0^t dt' e^{-\gamma[t-t']}\frac{\sin(\omega[t-t'])}{\omega}\frac{f(t')}{m}. \end{aligned} \tag{2.51}$$

# Chapter 3

# Functions of more than one variable

## 3.1 Functions of several variables

Definition: A real function $f(x_1, ..., x_n)$ of $n$ real variables $x_1$,...,$x_n$ is a map

$$f : R^n \to R, \quad \boldsymbol{x} := (x_1, x_2, ..., x_n) \to f(\boldsymbol{x}) = f(x_1, ..., x_n). \tag{3.1}$$

In the following, we mainly discuss functions of two variables $x_1$ and $x_2$, i.e. the case $n = 2$. Such functions can be represented by a **three–dimensional surface plot**, where the value $z = f(x, y)$ at each point $(x, y)$ in the $x$–$y$ plane is plotted in the $z$–direction over the $x$–$y$ plane. Here are two examples:

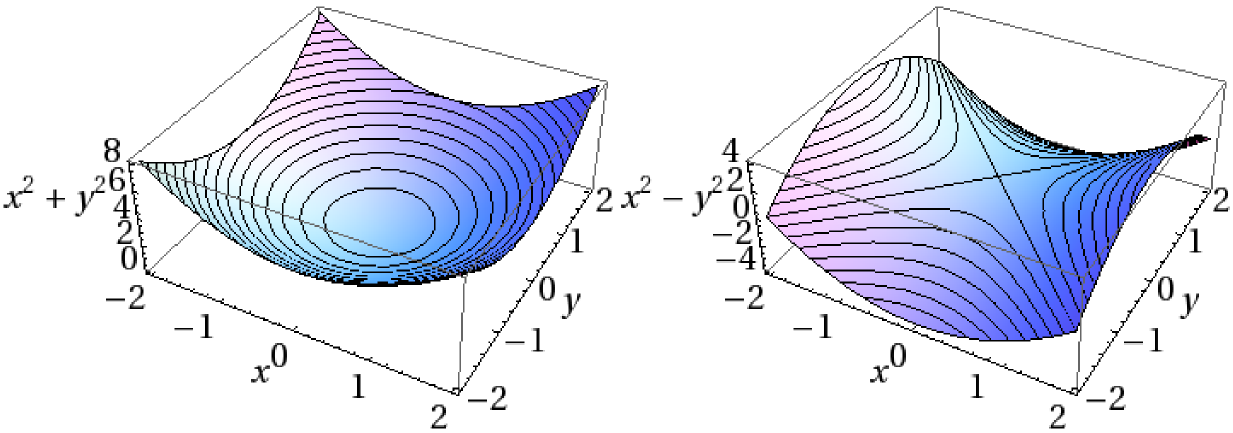



Figure 3.1: Examples of functions $f(x, y)$ of two variables $x$ and $y$: Paraboloid $f(x, y) = x^2 + y^2$ (LEFT), saddle $f(x, y) = x^2 - y^2$ (RIGHT).

### The Paraboloid $f(x, y) = x^2 + y^2$

We can understand this graph as follows:

1. If we keep $y$ fixed and change $x$, we have a parabola for each $y$, e.g.

$$\begin{aligned} f(x, 0) &= x^2 \\ f(x, -1) &= x^2 + 1 \\ f(x, 1) &= x^2 + 1 \\ f(x, 2) &= x^2 + 4 \\ &\dots \end{aligned}$$

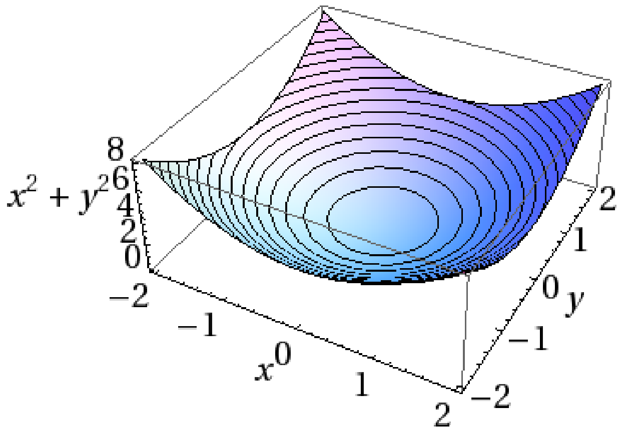



Figure 3.2: Paraboloid $f(x, y) = x^2 + y^2$.

If we keep $x$ fixed and change $y$, we have a parabola for each $x$, e.g.

$$\begin{aligned} f(0, y) &= y^2 \\ f(-1, y) &= 1 + y^2 \\ f(1, y) &= 1 + y^2 \\ f(3, y) &= 9 + y^2 \\ &\ldots \end{aligned}$$

These are cross–sections of the graph in $x$ and $y$ direction. It is important that you learn to visualise these cross–sections in your mind.

2. If we keep the value of the function fixed, i.e. $z = f(x, y) = z_0 = const > 0$, we find circles

$$z_0 = f(x, y) = x^2 + y^2. \tag{3.2}$$

Exercise: visualise these circles from the figure above. What are the radii of the circles for a given height $f(x, y) = z_0$?

### The Saddle $f(x, y) = x^2 - y^2$

We can understand this graph as follows:

1. If we keep $y$ fixed and change $x$, we have a parabola for each $y$, e.g.

$$\begin{aligned} f(x, 0) &= x^2 \\ f(x, -1) &= x^2 - 1 \\ f(x, 1) &= x^2 - 1 \\ f(x, 2) &= x^2 - 4 \\ &\ldots \end{aligned}$$

These are parabolas bending upwards but with their origin shifted to negative values.

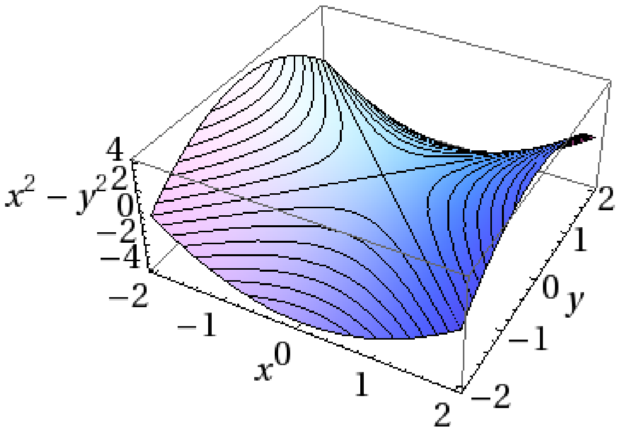



Figure 3.3: Saddle $f(x,y) = x^2 - y^2$.

If we keep $x$ fixed and change $y$, we have a parabola for each $x$, e.g.

$$\begin{aligned} f(0,y) &= -y^2 \\ f(-1,y) &= 1-y^2 \\ f(1,y) &= 1-y^2 \\ f(3,y) &= 9-y^2 \\ &\ldots \end{aligned}$$

These are parabolas bending downwards but with their origin shifted to positive values. We can built up the whole graph from these cross-sections. The interesting thing is that we understand the **global** shape of the surface $f(x,y)$, i.e. its saddle–shape, only form 'glueing' together all the cross–sections.

The most interesting points in both the paraboloid and the saddle are the extrema at $(x,y) = (0,0)$: for the paraboloid, this is a global minimum, for the saddle this is neither a minimum (it is a minimum in $x$–direction only) nor a maximum (it is a maximum in $y$–direction only): it is a **saddle–point**.

### 3.1.1 Symmetries

#### The Paraboloid $f(x,y) = x^2 + y^2$

We consider the circle of fixed radius $r$,

$$r^2 = x^2 + y^2 \tag{3.3}$$

in the $x$–$y$ plane. For all points on this circle, the function $f(x,y) = x^2 + y^2$ has the same value $f(x,y) = r^2$. A rotation of a point $(x,y)$ on this circle around the origin $(x,y) = (0,0)$ does not alter $f(x,y)$. In fact, if we rotate $f(x,y)$ continuously around the $z$–axis, $f(x,y)$ remains invariant. The function $f(x,y)$ has a continous rotation symmetry. Therefore, it is also called 'rotational paraboloid' sometimes. We can build up the whole surface $f(x,y)$ from rings with radius $r$ stacked on top of each other in the $z$–direction. Again, when dealing with functions of more than one variable, it is important that you develop this geometric kind of thinking.

The function $f(x,y)$ also has other symmetries: $f(x,y) = f(-x,y) = f(x,-y) = f(-x,-y)$ (reflection of one or two of its variables).

## 3.2 Partial Derivatives

### 3.2.1 Reminder: Derivative of a function of one variable

The derivative $f'(x)$ of a function $f(x)$ gives the slope of the function at $x$. It is defined as

$$\frac{df(x)}{dx} \equiv f'(x) := \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}. \tag{3.4}$$

Change of height: the quantity

$$f'(x)dx$$

gives the change of the height of the curve $f(x)$ (measured from the $x$–axis) at the point $x$, if we move a tiny step $dx$ along the $x$–axis.

### 3.2.2 Derivatives for functions of two variables

For a function $f(x,y)$ with two independent variables, in a certain point $(x,y)$ we can define the slope in either the $x$– or the $y$–direction. These two give rise to the **partial derivatives**

$$\begin{aligned}\frac{\partial}{\partial x} f(x,y) &:= \lim_{h\to 0} \frac{f(x+h,y)-f(x,y)}{h}\\ \frac{\partial}{\partial y} f(x,y) &:= \lim_{h\to 0} \frac{f(x,y+h)-f(x,y)}{h}.\end{aligned} \tag{3.5}$$

#### Partial derivative $\frac{\partial}{\partial x} f(x,y_0)$

The geometrical meaning of this is a follows: we keep $y = y_0$ constant and consider the surface $f(x,y)$ along the $x$–direction, i.e. the curve $f(x,y_0)$ on the surface that appears through the cross–section with the plane $y = y_0$ parallel to the $x$–$z$–plane. The partial derivative $\frac{\partial}{\partial x} f(x,y_0)$ gives the slope of this curve at $x$. In other words: the partial derivative $\frac{\partial}{\partial x} f(x,y)$ gives the slope of the surface at $(x,y)$ in $x$–direction. Change of height: the quantity

$$\frac{\partial}{\partial x} f(x,y)dx$$

gives the change of the height of the surface $f(x,y)$ (measured from the $x$–$y$–plane) at the point $(x,y)$, if we move a tiny step $dx$ along the $x$–direction.

#### Partial derivative $\frac{\partial}{\partial y} f(x_0,y)$

The geometrical meaning of this is a follows: we keep $x = x_0$ constant and consider the surface $f(x,y)$ along the $y$–direction, i.e. the curve $f(x_0,y)$ on the surface that appears through the cross–section with the plane $x = x_0$ parallel to the $y$–$z$–plane. The partial derivative $\frac{\partial}{\partial y} f(x_0,y)$ gives the slope of this curve at $y$. In other words: the partial derivative $\frac{\partial}{\partial y} f(x,y)$ gives the slope of the surface at $(x,y)$ in $y$–direction. Change of height: the quantity

$$\frac{\partial}{\partial y} f(x,y)dy$$

gives the change of the height of the surface $f(x,y)$ (measured from the $x$–$y$–plane) at the point $(x,y)$, if we move a tiny step $dy$ along the $y$–direction.

### Total change of height (total differential)

The quantity

$$df(x,y) := \frac{\partial}{\partial x} f(x,y)dx + \frac{\partial}{\partial y} f(x,y)dy \tag{3.6}$$

is called the **total differential of** $f(x,y)$ **at the point** $(x,y)$ and gives the total change of the height of the surface $f(x,y)$ (measured from the $x$–$y$–plane) at the point $(x,y)$, if we move a tiny step $dx$ along the $x$–direction and a tiny step $dy$ along the $y$–direction.

### How to calculate partial derivatives

This is very simple:

- To calculate $\frac{\partial}{\partial x} f(x,y)$, we keep $y$ fixed and differentiate $f(x,y)$ with respect to $x$. In doing so, $y$ is regarded as a fixed parameter.
- To calculate $\frac{\partial}{\partial y} f(x,y)$, we keep $x$ fixed and differentiate $f(x,y)$ with respect to $y$. In doing so, $x$ is regarded as a fixed parameter.

### Examples

$$\begin{aligned}
f(x,y) &= x^2 + y^2 \rightsquigarrow \frac{\partial}{\partial x} f(x,y) = 2x, \quad \frac{\partial}{\partial y} f(x,y) = 2y \\
f(x,y) &= x^2 y^3 \rightsquigarrow \frac{\partial}{\partial x} f(x,y) = 2xy^3, \quad \frac{\partial}{\partial y} f(x,y) = x^2 3y^2. \\
f(x,y) &= e^{-xy} \rightsquigarrow \frac{\partial}{\partial x} f(x,y) = -ye^{-xy}, \quad \frac{\partial}{\partial y} f(x,y) = -xe^{-xy}.
\end{aligned}$$

## 3.2.3 Higher Derivatives, Notation

Higher partial derivatives are easily defined: The second partial derivative $\frac{\partial^2}{\partial x^2} f(x,y)$ is the partial derivative with respect to $x$ of the partial derivative $\frac{\partial}{\partial x} f(x,y)$, etc. To simplify the notation, one often defines

$$\begin{aligned}
f_x \equiv \frac{\partial}{\partial x} f(x,y), \quad f_{xx} &\equiv \frac{\partial^2}{\partial x^2} f(x,y) := \frac{\partial}{\partial x}\frac{\partial}{\partial x} f(x,y) \\
f_y \equiv \frac{\partial}{\partial y} f(x,y), \quad f_{yy} &\equiv \frac{\partial^2}{\partial y^2} f(x,y) := \frac{\partial}{\partial y}\frac{\partial}{\partial y} f(x,y) \\
f_{xy} &\equiv \frac{\partial^2}{\partial x \partial y} f(x,y) := \frac{\partial}{\partial x}\frac{\partial}{\partial y} f(x,y) \\
f_{yx} &\equiv \frac{\partial^2}{\partial y \partial x} f(x,y) := \frac{\partial}{\partial y}\frac{\partial}{\partial x} f(x,y).
\end{aligned} \tag{3.7}$$

### Examples of Higher Partial Derivatives

$$\begin{aligned}
f(x,y) &= x^2 + y^2 \rightsquigarrow f_x(x,y) = 2x, \quad f_y(x,y) = 2y \\
f_{xx}(x,y) &= f_{yy}(x,y) = 2, \quad f_{xy}(x,y) = f_{yx}(x,y) = 0.
\end{aligned}$$

### 3.2.4 Minima, Maxima and Saddle points

Extrema (stationary points) occur when all partial derivatives are zero. Types: minima, maxima and saddlepoints (minimum in one direction, maximum in another).

## 3.3 Curves on Surfaces

### 3.3.1 Parametric curves in the $x$–$y$–plane

Definition: A curve in the $x$–$y$–plane is a map

$$R \to R^2, \quad t \to \mathbf{x}(t) := (x(t), y(t)) \tag{3.8}$$

which associates with each values of the parameter $t$ ('time' $t$) a point $(x(t), y(t))$ in the $x$–$y$–plane.

#### Examples

1. The circle around the origin,

$$(x(t), y(t)) = (r\cos(t), r\sin(t)). \tag{3.9}$$

Check that $x(t)^2 + y(t)^2 = r^2$ for all $t$.
2. The curve line

$$(x(t), y(t)) = (t^2, t). \tag{3.10}$$

Sketch this!

### 3.3.2 Parametric urves on Surfaces

Consider a function $f(x, y)$, i.e. a surface $z = f(x, y)$ above the $x$–$y$–plane. Consider a curve $(x(t), y(t))$ in the $x$–$y$–plane. This curve defines a corresponding curve

$$z(t) = f(x(t), y(t)) \tag{3.11}$$

*on* the surface. Example: for $f(x, y) = x^2 + y^2$ and $(x(t), y(t)) = (r\cos(t), r\sin(t))$, $z(t) = r^2$. The circle in the $x$–$y$–plane corresponds to a ring hovering at a distance $r^2$ above the plane, being part of the surface of the paraboloid $x^2 + y^2$. Sketch the corresponding picture (lecture)!

### 3.3.3 Change of height along a Curve

Reminder: Total change of height (total differential)

$$df(x, y) := \frac{\partial}{\partial x} f(x, y) dx + \frac{\partial}{\partial y} f(x, y) dy$$

is called the **total differential of** $f(x, y)$ **at the point** $(x, y)$ and gives the total change of the height of the surface $f(x, y)$ (measured from the $x$–$y$–plane) at the point $(x, y)$, if we move a tiny step $dx$ along the $x$–direction and a tiny step $dy$ along the $y$–direction.

From this, we can calculate the change of the height of the curve $z(t) = f(x(t), y(t))$:

$$\frac{dz(t)}{dt} = \frac{df(x(t), y(t))}{dt} = \frac{\partial}{\partial x} f(x, y) \frac{dx(t)}{dt} + \frac{\partial}{\partial y} f(x, y) \frac{dy(t)}{dt}. \tag{3.12}$$

This is a **chain rule**

$$\frac{df(x(t), y(t))}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}. \tag{3.13}$$

**Example:** $f(x,y) = x^2 + y^2$ **and** $(x(t), y(t)) = (t^2, t)$

We have

$$\frac{dz(t)}{dt} = = \frac{\partial}{\partial x} f(x,y) \frac{dx(t)}{dt} + \frac{\partial}{\partial y} f(x,y) \frac{dy(t)}{dt} = 2x(t) \cdot 2t + 2y(t) \cdot 1 = 4t^3 + 2t.$$

We can check this by direct calculation, $z(t) = t^4 + t^2 \rightsquigarrow dz(t)/dt = 4t^3 + 2t$. The general formula, however, makes it clear that there a two contributions to the change of the curve $z(t)$ on the surface: 1. the 'geometric change' (partial derivatives $f_x$, $f_y$) of the surface. 2. the 'kinematic change', i.e. the time derivatives $dx(t)/dt$, $dy(t)/dt$ that determine the speed by which we sweep along the curve $z(t)$.

**Example:** $f(x,y) = x^2 - y^2$ **and** $(x(t), y(t)) = (\cos t, \sin t)$

$$\begin{aligned} \frac{dz(t)}{dt} &= = \frac{\partial}{\partial x} f(x,y) \frac{dx(t)}{dt} + \frac{\partial}{\partial y} f(x,y) \frac{dy(t)}{dt} = 2x(t) \cdot (-\sin(t)) - 2y(t) \cdot \cos(t) \\ &= -2\cos(t)\sin(t) - 2\sin(t)\cos(t) = -2\sin(2t). \end{aligned}$$

Direct check with $z(t) = \cos^2(t) - \sin^2(t) = \cos(2t)$.

## 3.4 The Gradient

### 3.4.1 Definition of the Gradient

Definition: Let $f(x,y)$ be a real function of two variables. The **gradient** grad$f$ of $f$ in the point $(x_0, y_0)$ in the $x$–$y$–plane is the two–component *vector* of the partial derivatives $f_x$ and $f_y$ of $f$,

$$\operatorname{grad} f(x_0, y_0) \equiv \nabla f(x_0, y_0) = \left(f_x(x_0, y_0), f_y(x_0, y_0)\right). \tag{3.14}$$

The symbol $\nabla$ is called 'Nabla'–operator. Note: the gradient of $f$ in the point $(x_0, y_0)$ is a two–dimensional vector in the $x$–$y$–plane attached to that point. The map $(x,y) \to \nabla f(x,y)$ defines a **vector field**, i.e. to each point (vector) $(x,y)$ in the $x$–$y$–plane, a vector $\nabla f(x,y)$ is attached.

### 3.4.2 Examples

**Paraboloid** $f(x,y) = x^2 + y^2$

In this case,

$$\nabla f(x,y) = (2x, 2y).$$

Sketch this vector field in the $x$–$y$–plane (solution is given in the lecture).

**Hyperboloid** $f(x,y) = x^2 - y^2$

In this case,

$$\nabla f(x,y) = (2x, -2y).$$

Sketch this vector field in the $x$–$y$–plane (solution is given in the lecture).

### 3.4.3 Gradient and Differential; Geometrical Meaning

Reminder: Total change of height (total differential)

$$df(x,y) := \frac{\partial}{\partial x} f(x,y)dx + \frac{\partial}{\partial y} f(x,y)dy$$

is called the **total differential of** $f(x,y)$ **at the point** $(x,y)$ and gives the total change of the height of the surface $f(x,y)$ (measured from the $x$–$y$–plane) at the point $(x,y)$, if we move a tiny step $dx$ along the $x$–direction and a tiny step $dy$ along the $y$–direction.

Consider now a certain point $(x_0,y_0)$ in the $x$–$y$–plane, with the gradient $\nabla f(x_0,y_0)$ of the function $f(x,y)$ attached. In that point, the total change of height of the function $f(x,y)$ can be written as a scalar product,

$$df = \frac{\partial}{\partial x} f dx + \frac{\partial}{\partial y} f dy = (f_x, f_y) \cdot (dx, dy)$$

of the two vectors $(f_x, f_y) = \nabla f (\equiv \nabla f(x_0,y_0))$ and $(dx,dy)$. We now change $dx$ and $dy$ slightly, thereby changing the vector $(dx,dy)$ of the differentials. Then, for a certain values of $dx$ and $dy$, the vector $(dx,dy)$ becomes perpendicular to the gradient $(f_x,f_y)$, i.e. the scalar product $df = \nabla f \cdot (dx,dy)$ vanishes. In this direction $(dx,dy)$, the height of the surface does not change, it determines the direction of an equipotential line. Therefore, the gradient $\nabla f(x_0,y_0)$ is perpendicular to the equipotential line through $(x_0,y_0)$; it determines the direction of the steepest increase of the function $f(x,y)$.

**Example: Paraboloid** $f(x,y) = x^2 + y^2$

We have

$$\nabla f(x,y) = (2x, 2y).$$

The equipotential lines are circles $r^2 = x^2 + y^2$ in the $x$–$y$–plane. The gradient is perpendicular to these circles. Picture in the lecture.

# Chapter 4

# Series and Limits

## 4.1 Finite and Infinite Series

### 4.1.1 Finite Series of Natural Numbers (Gauß)

$$\sum_{k=1}^{n} k = 1 + 2 + 3 + ... + n = \frac{n(n+1)}{2}. \tag{4.1}$$

First Proof (C. F. Gauß) for $n = 100$:

$$\sum_{k=1}^{100} k = (1 + 100) + (2 + 99) + (3 + 98) + ... + (50 + 51) = 50 \cdot 101 = 5050. \tag{4.2}$$

General proof by **induction**:
1. Induction Start ($n = 1$):

$$\sum_{k=1}^{1} k = 1 = \frac{1(1+1)}{2} \rightsquigarrow OK. \tag{4.3}$$

2. Induction Step ($n \to n + 1$): Assume Eq. (4.1) is true for $n$, then show that it is also true for $n + 1$:

$$\sum_{k=1}^{n+1} k = \sum_{k=1}^{n} k + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2} \rightsquigarrow OK. \tag{4.4}$$

This can now be used to prove it is true for $n + 2$, etc.

### 4.1.2 Finite Geometric Progression

$$\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x}, \qquad x \neq 0. \tag{4.5}$$

Proof: Write

$$\begin{aligned} S_n = \sum_{k=0}^{n} x^k &= 1 + x + x^2 + ... + x^n \\ &= 1 + x(1 + x + x^2 + ... + x^{n-1} + x^n) - x^{n+1} \\ &= 1 + xS_n - x^{n+1} \rightsquigarrow \\ S_n &= \frac{1 - x^{n+1}}{1 - x}. \end{aligned} \tag{4.6}$$

Alternative proof by **induction**: home exercise.

### 4.1.3 Binomial

$$(x+y)^n = \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n-k} y^k. \tag{4.7}$$

Here, we define the **binomial coefficient**

$$\left( \begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdot ... (n-k+1)}{1\cdot 2 \cdot ... \cdot k}. \tag{4.8}$$

The proof of Eq.(4.7) goes again via induction $n \to n+1$. Not shown here. Examples:

$$\begin{aligned} (x+y)^2 &= x^2 + 2xy + y^3 \\ (x+y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3. \end{aligned} \tag{4.9}$$

### 4.1.4 Infinite Series

#### Definition

A series

$$S := \sum_{k=0}^{\infty} a_k = a_0 + a_1 + a_2 + ... = \lim_{n\to\infty} S_n, \quad S_n := \sum_{k=0}^{n} a_k \tag{4.10}$$

is called **infinite series**. It is the limit of the sequence of finite series $S_n$ when the upper limit $n$ tends toward infinity. (The objects $S_n$ are also called "partial sums".) In contrast to the finite series $S_n$, the infinite series $S$ can **diverge**. $S$ is said to be **convergent** is $S_n$ approaches a finite limit as $n \to \infty$.

**Example 4.1:**

Use a constant $a_k$

$$a_k = 1 \rightsquigarrow \sum_{k=0}^{\infty} a_k = 1 + 1 + 1 + 1 + .... \tag{4.11}$$

is divergent because the partial sums $S_n = n$, which clearly diverge as $n \to \infty$.

**Example 4.2:**

The geometric series

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \quad |x| < 1. \tag{4.12}$$

**learn this one by heart**.
This series converges for arbitrary (real or complex) numbers $x$ with $|x| < 1$. [Please sketch the condition $|z| < 1$ for complex $z$ as an an area in the complex plane.]

Proof of Eq. (4.33):

$$\begin{aligned} S &= \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + ... = 1 + x(1 + x + x^2 + ...) = 1 + x \cdot S \rightsquigarrow \\ S &= \frac{1}{1-x}. \end{aligned} \tag{4.13}$$

The problem with infinite series is that often it is not easy to decide if or if not they converge, e.g. for which values of $x$ in the above example.

A necessary condition for convergence of $S = \sum_{k=0}^{\infty} a_k$ is that $a_k \to 0$ as $k \to \infty$.

A sufficient condition for convergence: is the **ratio test**

> **Ratio test**: Consider the series $S := \sum_{k=0}^{\infty} a_k$ and assume $a_k \neq 0$ for all $k > k_0$. Define the ratio
>
> $$\begin{aligned} R &:= \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| \rightsquigarrow \\ R < 1 &\quad \text{series is convergent} \\ R > 1 &\quad \text{series is divergent.} \end{aligned} \tag{4.14}$$
>
> For $R = 1$, the ratio test can't decide whether the series is convergent or divergent.

## 4.2 Taylor–Series

One of the main motivations to investigate infinite series is the desired to write arbitrary functions $f(x)$ as polynomials of infinite degree, i.e.

$$\begin{aligned} f(x) &= a_0 + a_1 x + a_2 x^2 + a_3 x^4 + \ldots \\ &= \sum_{k=0}^{\infty} a_k x^k. \end{aligned} \tag{4.15}$$

For each fixed $x$, this is an infinite series of the form $S := \sum_{k=0}^{\infty} b_k$ with $b_k := a_k x^k$. An important question is, e.g., how to determine the coefficients $a_k$ for a given function $f(x)$, and to decide for which values of $x$ the series for $f(x)$ does converge.

### 4.2.1 The Exponential Function

We already know one example for such a series which is the exponential function

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{4.16}$$

**You have to remember this formula throughout your whole life**. This series converges for arbitrary values of (complex or real) $x$ since (ratio test!)

$$R := \lim_{n\to\infty} \left| \frac{x^{n+1}/(n+1)!}{x^n/n!} \right| = \lim_{n\to\infty} \frac{x}{n+1} = 0. \tag{4.17}$$

By use of the this **exponential series** one defines the famous **Euler number**

$$e := \sum_{n=0}^{\infty} \frac{1}{n!} = \exp(1). \tag{4.18}$$

### 4.2.2 Power Series for $\sin(x)$ and $\cos(x)$

We repeat our result for the series that define sin and cos:

$$\begin{aligned} \sin(x) &:= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \\ \cos(x) &:= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \end{aligned} \tag{4.19}$$

### 4.2.3 General Case

Now we treat the case of an arbitrary function

$$f(x) \quad = \quad a_0 + a_1 x + a_2 x^2 + a_3 x^4 + ... = \sum_{k=0}^{\infty} a_k x^k. \tag{4.20}$$

The above equation means that we try to represent the function by an 'infinite' polynomial. In the following, we assume that all derivatives of $f(x)$, i.e. $f'(x) =: f^{(1)}(x)$, $f''(x) =: f^{(2)}(x)$, $f'''(x) =: f^{(3)}(x)$, ... etc. exist. We write

$$\begin{aligned}
f(x=0) &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + ...\Big|_{x=0} = a_0 \\
f'(x=0) &= a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + ...\Big|_{x=0} = a_1 \\
f''(x=0) &= 2a_2 + 2\cdot 3a_3 x + 3\cdot 4a_4 x^2 + ...\Big|_{x=0} = 1\cdot 2a_2 \\
f^{(3)}(x=0) &= 2\cdot 3a_3 + 2\cdot 3\cdot 4a_4 x + ...|_{x=0} = 1\cdot 2\cdot 3a_3 \\
... &= ... \\
f^{(n)}(x=0) &= 1\cdot 2\cdot ...\cdot n\cdot a_n = n! a_n \\
\rightsquigarrow a_n &= \frac{f^{(n)}(x=0)}{n!}.
\end{aligned} \tag{4.21}$$

Collecting all terms, we find the

**Taylor expansion of $f(x)$ around $x = 0$,**

$$f(x) = \frac{f(x=0)}{0!} + \frac{f'(x=0)}{1!}x + \frac{f''(x=0)}{2!}x^2 + ... = \sum_{n=0}^{\infty} \frac{f^{(n)}(x=0)}{n!}x^n. \tag{4.22}$$

We define the truncated Taylor series

$$f_N(x) := \sum_{n=0}^{N} \frac{f^{(n)}(x=0)}{n!}x^n \rightsquigarrow f(x) = \lim_{N\to\infty} f_N(x). \tag{4.23}$$

The truncated Taylor series for finite $N$ is often used as an approximation for the function $f(x)$. For larger and larger $N$, we expect that this approximation of the function $f(x)$ by a polynomial of degree $N$ becomes better and better, if the series converges, of course. Let us look at an example to see how this works:

### 4.2.4 Example: The Exponential Function $\exp(x)$

We calculate the Taylor series of $f(x) = \exp(x)$ around $x = 0$. To do so, we have to calculate the derivatives

$$\begin{aligned}
f^{(0)}(x=0) \equiv f(x=0) &= e^x|_{x=0} = 1 \\
f^{(1)}(x=0) &= e^x|_{x=0} = 1 \\
f^{(2)}(x=0) &= e^x|_{x=0} = 1 \\
... & \quad ...
\end{aligned} \tag{4.24}$$

This is particularily simple because all the derivatives of $e^x$ are $e^x$. This means that

$$\begin{aligned} f(x) &= \lim_{N\to\infty} f_N(x) = \lim_{N\to\infty} \sum_{n=0}^{N} \frac{f^{(n)}(x=0)}{n!} x^n \\ &= \lim_{N\to\infty} \sum_{n=0}^{N} \frac{x^n}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{n!}. \end{aligned} \tag{4.25}$$

We recognise that the Taylor expansion of $f(x) = \exp(x)$ just reproduces our old result, Eq. (4.16).

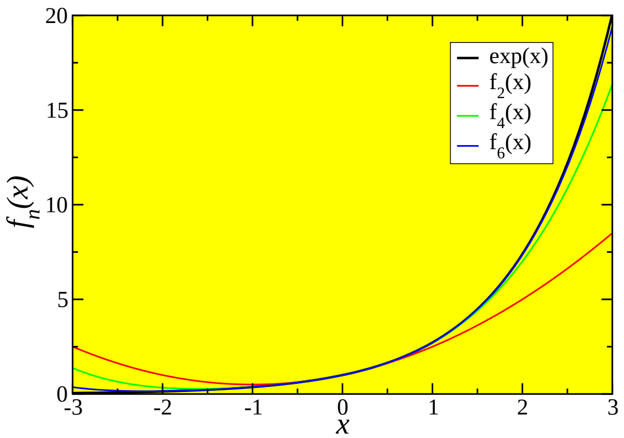



Figure 4.1:
Approximation of the function $f(x) = \exp(x)$ by the truncated Taylor Series $f_N(x)$, Eq. (4.25), for $N = 2, 4, 6$. For the interval $x \in [-3, 3]$ shown here, the approximation of $\exp(x)$ by $f_6(x)$ is already very good.

We can apply the ratio test to the series for the exponential. WIth $a_n = x^n/n!$, we find $R = \lim_{n\to\infty} x/(n+1) = 0$ for every fixed $x$. The Taylor series for the exponent thus converges for all $x$.

## 4.3 Taylor–Expansion of Functions

### 4.3.1 Convergence: Expansion of $f(x) = \ln(1+x)$

The derivatives of this function are

$$\begin{aligned} f(x) &= \ln(1+x) \rightsquigarrow f(0) = \ln(1) = 0 \\ f'(x) &= (1+x)^{-1} \\ f''(x) &= (-1)(1+x)^{-2} \\ f^{(3)}(x) &= 2(1+x)^{-3} \\ f^{(4)}(x) &= -6(1+x)^{-4} \\ \ldots & \quad \ldots \\ f^{(n)}(x) &= (-1)^{n+1}(n-1)!(1+x)^{-n} \rightsquigarrow f^{(n)}(x=0) = (-1)^{n+1}(n-1)!. \end{aligned} \tag{4.26}$$

We use this to expand $f(x)$ around $x = 0$,

$$\ln(1+x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x=0)}{n!} x^n = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n (n-1)!}{n!} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n}. \tag{4.27}$$

Now we ask: for which values of $x$ does this Taylor series actually converge? We use the ratio test and write

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}x^n}{n} &= \sum_{n=1}^{\infty} a_n \rightsquigarrow a_n = \frac{(-1)^{n+1}x^n}{n}\\ R &:= \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\frac{n}{n+1}|x| = |x|\,. \end{aligned} \tag{4.28}$$

From Eq. (4.14), we recognise that the series

- converges for $|x| < 1$.
- diverges for $|x| > 1$.

At $x = -1$ we don't expect the series to converge because $\ln(1 + (-1)) = \ln(0)$ is undefined (minus infinity). To decide what happens at $x = 1$, we have to invoke an additional convergence test:

**Leibnitz' test for alternating series:** The alternating series

$$\sum_{k=0}^{\infty}(-1)^k|a_k| \tag{4.29}$$

converges, if $|a_{k+1}| < |a_k|$ for all $k$, and $\lim_{k\to\infty} a_k = 0$.

We apply this rule to the case $x = 1$ of our series Eq. (**??**) for $\ln(1 + x)$: At $x = 1$, $|a_{n+1}| = 1/(n+1) < |a_n| = 1/n$ and $\lim_{n\to\infty}|a_n| = 0$, that means the Leibnitz' test tells us that the series converges at $x = 1$. The result gives us a famous formula for $\ln 2$,

$$\ln 2 = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - ..., \tag{4.30}$$

and we summarise our results for $f(x) = \ln(1 + x)$ as

$$\ln(1+x) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}x^n}{n}, \quad |x| < 1. \tag{4.31}$$

We say that the **radius of convergence** $R$ of this series is $R = 1$. For values of $x$ beyond that radius, the series diverges and does no longer represent the function $\ln(1 + x)$. In other words, the Taylor series Eq. (4.31) is only useful for 'small' $x$.

### 4.3.2 Alternative way to generate a Taylor Series

Let us write $\ln(1 + x)$ in a 'complicated way', i.e. as an integral:

$$\ln(1+x) = \int_0^x \frac{dt}{1+t}. \tag{4.32}$$

Now, we use our result for the geometric series, Eq.(4.33),

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad |x| < 1. \tag{4.33}$$

(**'LEARN THIS ONE BY HEART'**) with $t = -x$, which leads to

$$\begin{aligned}\ln(1+x) &= \int_0^x dt\frac{1}{1+t} = \int_0^x dt\sum_{n=0}^{\infty}(-t)^n\\ &= \int_0^x dt(1 - t + t^2 - t^3 + ...)\end{aligned} \tag{4.34}$$

We integrate this term by term, which is easy,

$$\begin{aligned}\ln(1+x) &= \int_0^x dt(1-t+t^2-t^3+...) = \sum_{n=0}^{\infty}\int_0^x dt(-t)^n = x - \frac{x^2}{2}+\frac{x^3}{3} - ... = \\ &= \sum_{n=0}^{\infty}\frac{(-1)^n x^{n+1}}{n+1} = \sum_{n=1}^{\infty}\frac{(-1)^{n-1}x^n}{n},\end{aligned} \tag{4.35}$$

which is the same as Eq. (4.31)

### 4.3.3 Taylor expansion of $f(x)$ around an arbitrary $x = a$

So far we have always expanded our functions $f(x)$ in the vicinity of $x = 0$, i.e. 'around' $x = 0$:

**Taylor expansion of $f(x)$ around $x = 0$,**

$$f(x) = \frac{f(x=0)}{0!} + \frac{f'(x=0)}{1!}x + \frac{f''(x=0)}{2!}x^2 + ... = \sum_{n=0}^{\infty}\frac{f^{(n)}(x=0)}{n!}x^n. \tag{4.36}$$

The Taylor expansion of a function $f(x)$ near $x = a$ is performed in an analogous way, but with $x = 0$ replaced by $x = a$, and $x = x - 0$ replaced by $x - a$:

**Taylor expansion of $f(x)$ around $x = a$,**

$$f(x) = \frac{f(a)}{0!} + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + ... = \sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{4.37}$$

In some books, the special case of a Taylor series around $x = 0$ is called **Maclaurin Series** .

## 4.4 Further Examples for Series and Limits

### 4.4.1 Newtonian Limit of Relativistic Energy

According to Einstein, the total energy of a particle of rest mass $m_0$ and velocity $v$ is

$$E = \frac{m_0c^2}{\sqrt{1-\left(\frac{v}{c}\right)^2}}, \tag{4.38}$$

where $c$ is the speed of light on vacuum. We would like to find an approximation of this formula for small velocities $v \ll c$, in order to compare to Newton's expression for the kinetic energy, $E_{\rm kin} = (1/2)m_0v^2$. Defining $\beta := v/c$, we recognise that our mathematical task is to ($x := \beta^2$)

**Expand $f(x) = \frac{1}{\sqrt{1-x}}$ around $x = 0$:**

We write $f(x) = (1-x)^{-1/2}$ and

$$\begin{aligned} f(0) &= 1 \\ f'(0) &= (-1)(-1/2)(1-x)^{-3/2}\Big|_{x=0} = 1/2 \\ f''(0) &= (-1)(-1/2)(-1)(-3/2)(1-x)^{-3/2}\Big|_{x=0} = 3/4 \\ &... \end{aligned} \tag{4.39}$$

With our

**Taylor expansion of $f(x)$ around $x = 0$,**

$$f(x) = \frac{f(x=0)}{0!} + \frac{f'(x=0)}{1!}x + \frac{f''(x=0)}{2!}x^2 + ... = \sum_{n=0}^{\infty} \frac{f^{(n)}(x=0)}{n!}x^n, \tag{4.40}$$

we find the first terms as

$$f(x) = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + ... \tag{4.41}$$

(note that $2! = 2$). Therefore, with $x = \beta^2 = (v/c)^2$, we obtain

$$E = m_0 c^2 \left[ 1 + \frac{1}{2}\left(\frac{v}{c}\right)^2 + \frac{3}{8}\left(\frac{v}{c}\right)^4 + O\left(\frac{v}{c}\right)^6 \right]. \tag{4.42}$$

Here, we introduced the $O$**–symbol** (speak 'order of'), i.e. $O(x)^6$ means 'terms of order $x^6$ or higher powers' like $x^7$, $x^8$ etc. This is a convenient way to express that in a Taylor expansion with the first few terms written down as above, there are higher order terms to follow that one does not care to write down explicitely here. These higher order terms in fact become smaller and smaller for $|x| < 1$.

We can take use of the $O$–symbol to write

$$E = m_0 c^2 + \frac{1}{2}m_0 v^2 + O\left(\frac{v}{c}\right)^4. \tag{4.43}$$

This shows that the first term in the total energy is a velocity–independent rest energy of the particle, and the second term is the *lowest order approximation* to its kinetic energy. The relativistic correction to the kinetic energy if of order $(v/c)^4$, i.e. very small for velocities small compared to the speed of light.

### 4.4.2 Limits

#### Expressions of the type '$0/0$'

Often we have to discuss and sketch functions like

$$f(x) = \frac{\sin(x)}{x} \tag{4.44}$$

with a seemingly ill-defined behaviour at $x = 0$. Direct substitution gives $0/0$, which is indeed not well defined. However, a closer look shows that we can make sense of the function even at $x = 0$: we expand $\sin(x)$ by its Taylor series near $x = 0$, i.e.

$$f(x) = \frac{\sin(x)}{x} = \frac{x - \frac{x^3}{3!} + O(x^5)}{x} = 1 - \frac{x^2}{3!} + O(x^4). \tag{4.45}$$

This means, that if $x$ approaches $x = 0$ we have the finite value $f(x = 0) = 1$, i.e.

$$\lim_{x \to 0} f(x) = 1. \tag{4.46}$$

The deviation from $f(x = 0) = 1$ close to $x = 0$ is described by the second term, $-x^2/6$, i.e. a quadratic decrease of the function for small values of $x$.

# Chapter 5

# Two–by–Two Matrices

## 5.1 Two-by-Two Matrices: Introduction

### 5.1.1 Linear Equations of Two Unknowns

Consider the system of linear equations for the two unknowns $x$ and $y$,

$$\begin{aligned} ax + by &= e \\ cx + dy &= f, \end{aligned} \tag{5.1}$$

where $a, b, c, d, e, f$ are constant numbers. This system can be easily solved: solve the first equation for $y$,

$$y = \frac{e - ax}{b} \tag{5.2}$$

and insert it into the second equation,

$$\begin{aligned} cx + dy &= cx + \frac{de - adx}{b} = f \rightsquigarrow (cb - ad)x = fb - de \\ x &= \frac{de - fb}{ad - cb} \\ y &= \frac{e - ax}{b} = \frac{e(ad - cb) - a(de - fb)}{b(ad - cb)} = \frac{af - ec}{ad - cb}. \end{aligned} \tag{5.3}$$

For this general solution for $x$ and $y$ to be valid, the denominator $ad - cb$ apparently has to be different from zero.

### 5.1.2 Two–by–Two Matrices: Definition

We write the two unknowns $x$ and $y$ as the components of a two–dimensional vector $\boldsymbol{x}$,

$$\boldsymbol{x} := \begin{pmatrix} x \\ y \end{pmatrix}. \tag{5.4}$$

Then, we write the two constants $e$ and $f$ as the components of a two–dimensional vector $\boldsymbol{v}$

$$\boldsymbol{v} := \begin{pmatrix} e \\ f \end{pmatrix}. \tag{5.5}$$

The two–by–two system of linear equations, Eq. (5.1), **maps the vector $\boldsymbol{x}$ onto the vector $\boldsymbol{v}$**. We write this in the following abstract form:

$$A\boldsymbol{x} = \boldsymbol{v} \Leftrightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix}, \tag{5.6}$$

where we defined the **two–by–two matrix**

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{5.7}$$

A two–by–two matrix is a quadratic scheme which, upon operating on a vector $\boldsymbol{x}$ on its right, transforms this vector into another vector $\boldsymbol{v}$ according to the rule

$$A\boldsymbol{x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = \boldsymbol{v}. \tag{5.8}$$

By comparison we recognise that this **matrix equation**, $A\boldsymbol{x} = \boldsymbol{v}$, is equivalent to the system Eq.(5.1).

### 5.1.3 Linear Mappings and Matrix Operatings

A *linear mapping* $A$ from $\mathbb{R}^2 \to \mathbb{R}^2$ maps a vector $\boldsymbol{x}$ onto the vector $A\boldsymbol{x}$. The mapping is represented by a two-by-two matrix $A$. The mapping must fulfill
$\boldsymbol{x}_1 + \boldsymbol{x}_2 \to A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = A\boldsymbol{x}_1 + A\boldsymbol{x}_2, \quad \lambda\boldsymbol{x} \to A(\lambda\boldsymbol{x}) = \lambda A\boldsymbol{x}, \quad \lambda \in \mathbb{R}.$

The above can be generalised (trivially) to complex matrices; the mapping is then from $\mathbb{C}^2 \to \mathbb{C}^2$, and $\lambda$ can also be complex

#### Examples

$$\begin{aligned}
A &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}, \quad \boldsymbol{x}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{x}_1 + \boldsymbol{x}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \\
\rightsquigarrow A\boldsymbol{x}_1 &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 3 \cdot (-2) \\ 2 \cdot 1 + (-1) \cdot (-2) \end{pmatrix} = \begin{pmatrix} -5 \\ 4 \end{pmatrix} \\
A\boldsymbol{x}_2 &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}.
\end{aligned} \tag{5.9}$$

We compare this to

$$\begin{aligned}
A &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}, \quad \boldsymbol{x}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{x}_1 + \boldsymbol{x}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\
\rightsquigarrow A(\boldsymbol{x}_1 + \boldsymbol{x}_2) &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \end{pmatrix} = \begin{pmatrix} -5 \\ 4 \end{pmatrix} + \begin{pmatrix} 4 \\ 1 \end{pmatrix} = A\boldsymbol{x}_1 + A\boldsymbol{x}_2 \rightsquigarrow \text{OK}.
\end{aligned}$$

## 5.2 Two–by–Two Matrices: Linear Mappings

The **determinant** $\det(A)$ of a two–by–two matrix $A$ is defined as

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - cb.$$

### 5.2.1 Specific Linear Mappings 1: the Unit Matrix

This is the *trivial mapping* represented by the **unit** or **identity** matrix, $I$,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.10}$$

We have $\det(I) = 1$. Check that $I\boldsymbol{x} = \boldsymbol{x}$ for any vector $\boldsymbol{x}$.

### 5.2.2 Specific Linear Mappings 2: Stretching and Shrinking

These are linear mappings $A$ represented by the multiples of the **unit matrix**, where $c$ is a real number such that

$$A = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}. \tag{5.11}$$

We have $\det(A) = c^2 > 1$. Check that in this case $A\boldsymbol{x} = c\boldsymbol{x}$ for any vector $\boldsymbol{x}$.

### 5.2.3 Specific Linear Mappings 3: Projections

These are linear mappings $A$ such as

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.12}$$

We have $\det(A) = 0$. Check that in this case, for any vector $\boldsymbol{x} = (x, y)$, $A\boldsymbol{x} = (x, 0)$: the vector is projected onto the $x$-axis.

### 5.2.4 Specific Linear Mappings 4: Rotations

These are mappings $R(\theta)$ that rotate vectors around the origin by an angle $\theta$,

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{5.13}$$

In this case, $\det(R(\theta)) = \cos^2\theta - (-\sin^2\theta) = 1$. A vector $\boldsymbol{x} = (x, y)$ is rotated into

$$R(\theta)\boldsymbol{x} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}. \tag{5.14}$$

Examples for rotations are

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}. \tag{5.15}$$

#### Special Rotations: $\theta = 0$

In this case,

$$R(\theta = 0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \quad \text{(unit matrix)}. \tag{5.16}$$

#### Special Rotations: $\theta = \frac{\pi}{2}$

In this case,

$$R\left(\theta = \frac{\pi}{2}\right) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -i\sigma_y \quad (-i \text{ times Pauli Matrix } \sigma_y). \tag{5.17}$$

### 5.2.5 Specific Linear Mappings 5: Reflections

These are mappings $S(\theta)$ that reflect a vectors at a fixed axis:

$$S(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{5.18}$$

In this case, $\det(S(\theta)) = -\cos^2\theta - \sin^2\theta = -1$. A vector $\boldsymbol{x} = (x, y)$ is transformed into

$$S(\theta)\boldsymbol{x} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\sin\theta \\ x\sin\theta - y\cos\theta \end{pmatrix}. \tag{5.19}$$

Examples:

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} \cos\frac{1}{2}\theta \\ \sin\frac{1}{2}\theta \end{pmatrix} = \begin{pmatrix} \cos\frac{1}{2}\theta\cos\theta + \sin\frac{1}{2}\theta\sin\theta \\ \cos\frac{1}{2}\theta\sin\theta - \sin\frac{1}{2}\theta\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\frac{1}{2}\theta \\ \sin\frac{1}{2}\theta \end{pmatrix}, \tag{5.20}$$

where we have a formula for trigonometric functions (CHECK). Furthermore, we have

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin\theta \\ -\cos\theta \end{pmatrix}. \tag{5.21}$$

Sketch this in the $x$-$y$-plane (lecture). We recognise that $S(\theta)$ defines a reflection at the axis defined by the direction of the vector $(\cos\frac{1}{2}\theta, \sin\frac{1}{2}\theta)$

#### Special Reflection: $\theta = 0$

In this case,

$$S(\theta = 0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_z \quad \text{(Pauli Matrix } \sigma_z\text{)}. \tag{5.22}$$

#### Special Reflection: $\theta = \frac{\pi}{2}$

In this case,

$$S\left(\theta = \frac{\pi}{2}\right) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_x \quad \text{(Pauli Matrix } \sigma_x\text{)}. \tag{5.23}$$

## 5.3 Two–by–Two Matrices: Index Notation and Multiplication

### 5.3.1 Basis Vectors and Index Notation

#### Vectors

The vectors $\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are called **basis vectors** of $\mathbb{R}^2$.

Any arbitrary vector $\boldsymbol{a} \in \mathbb{R}^2$ is written as a **linear combination**

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 = \sum_{i=1}^{2} a_i\boldsymbol{e}_i. \tag{5.24}$$

In this representation, sometimes **Einstein's summation convention** is used: We write $\boldsymbol{a} = \sum_{i=1}^{2} a_i\boldsymbol{e}_i = a_i\boldsymbol{e}_i$, omitting the sum symbol in order to simplify the notation. The sum is automatically carried out over repeated indices. Here, the index is $i$.

### Matrices

The element $A_{ij}$ of a matrix $A$ is the entry in its $i$-th row and its $j$-th column. For two-by-two matrices, this reads $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$.

Note: be very careful not to mix up the row and the column index!

### Matrix operating on vector

The result of a linear mapping $\boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x}$ can be written in index form, too:

$$\begin{aligned} & \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to A\boldsymbol{x} = \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ \longleftrightarrow \quad & y_i = \sum_{j=1}^{2} A_{ij} x_j. \end{aligned} \tag{5.25}$$

This means that the first and second components, $y_1$ and $y_2$, of $\boldsymbol{y} = A\boldsymbol{x}$ are given by

$$y_1 = \sum_{j=1}^{2} A_{1j} x_j, \quad y_2 = \sum_{j=1}^{2} A_{2j} x_j. \tag{5.26}$$

Note that the index $j$ runs over the columns of the matrix $A$.

## 5.3.2 Multiplication of a Matrix with a Scalar

This is simple,

$$\lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}. \tag{5.27}$$

## 5.3.3 Matrix Multiplication: Definition

A matrix $A$ moves a vector $\boldsymbol{x}$ into a new vector $\boldsymbol{y} = A\boldsymbol{x}$. This new vector can again be transformed into another vector $\boldsymbol{y}'$ by acting with another matrix B on it: $\boldsymbol{y}' = B\boldsymbol{y} = BA\boldsymbol{x}$. The combined operation $C = BA$ transforms the original vector $\boldsymbol{x}$ into $\boldsymbol{y}'$ in one single step. This **matrix product** is calculated according to

$$\begin{aligned} B &= \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}, \quad A = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \\ \leadsto BA &= \begin{pmatrix} a_2 a_1 + b_2 c_1 & a_2 b_1 + b_2 d_1 \\ c_2 a_1 + d_2 c_1 & c_2 b_1 + d_2 d_1 \end{pmatrix}. \end{aligned} \tag{5.28}$$

In general, the matrix product does not commmute, i.e.,

$$AB \neq BA. \tag{5.29}$$

This means that in contrast to real or complex numbers, the result of a multiplication of two matrices $A$ and $B$ depends on the order of $A$ and $B$.

The **commutator** $[A, B]$ of two matrices $A$ and $B$ is defined as $[A, B] = AB - BA$.

The commutator plays a central role in quantum mechanics, where classical variables like position $x$ and momentum $p$ are replaced by **operators**(matrices) which in general do not commute, i.e., their commutator is non–zero.
Example:

$$\begin{aligned} \sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_z\sigma_x &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \sigma_x\sigma_z = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \neq \sigma_z\sigma_x, \quad [\sigma_z, \sigma_x] = 2\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \end{aligned} \tag{5.30}$$

### 5.3.4 Matrix Multiplication: Index Notation

The abstract way to write a matrix multiplication with indices:

$$C = BA \rightsquigarrow C_{ij} = \sum_{k=1}^{2} B_{ik}A_{kj}. \quad (= B_{ik}A_{kj} \text{ in the summation convention}). \tag{5.31}$$

To get the element in the $i$th row and $j$th column of the product $BA$, take the scalar product of the $i$th row-vector of $B$ with the $j$-th column vector of $A$. This looks complicated but it is not, it is just another formulation of our definition Eq.(5.28).

## 5.4 Inverse of a Matrix

### 5.4.1 Motivation

Solving the linear two–by–two system Eq. (5.1) for the components $x$, $y$ of the vector $\boldsymbol{x}$, is equivalent to the matrix equation

$$A\boldsymbol{x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix} = \boldsymbol{v} = \begin{pmatrix} e \\ f \end{pmatrix}. \tag{5.32}$$

We recognise that in order to explicitely solving this for $\boldsymbol{x}$, we have to **invert** the operation $A$.

### 5.4.2 Definition and Theorem

The **inverse** $A^{-1}$ **of a two–by–two matrix** $A$ is defined as the matrix fulfilling

$$A^{-1}A = AA^{-1} = I, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The **determinant** $\det(A)$ of a two–by–two matrix $A$ is defined as

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - cb.$$

<u>Theorem</u> Consider the two–by–two matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{5.33}$$

If the determinant of $A$ is non–zero, i.e. $\det(A) = ad - cb \neq 0$, the inverse of $A$ exists and is given by

$$A^{-1} = \frac{1}{ad-cb}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \equiv \begin{pmatrix} \frac{d}{ad-cb} & \frac{-b}{ad-cb} \\ \frac{-c}{ad-cb} & \frac{a}{ad-cb} \end{pmatrix}. \tag{5.34}$$

For the proof of this, we just multiply $A$ with $A^{-1}$ and $A^{-1}$ with $A$:

$$AA^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \frac{1}{ad-cb}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad-cb}\begin{pmatrix} ad-bc & -ab+ba \\ cd-dc & -cb+da \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.35}$$

Exercise: Check the same for $A^{-1}A$.

### Examples

$$\begin{aligned} A &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \rightsquigarrow \det(A) = -1-6 \neq 0, \quad A^{-1} = \frac{1}{-7}\begin{pmatrix} -1 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & \frac{-1}{7} \end{pmatrix} \\ A &= \begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix} \rightsquigarrow \det(A) = 3\cdot 4 - 2\cdot 6 = 0 \rightsquigarrow A^{-1}\text{does not exist.} \end{aligned}$$

### Solving the Linear Equations (5.1)

We are now in a position to solve Eq. (5.1) by the inverse of a matrix:

$$\begin{aligned} A\boldsymbol{x} &= \boldsymbol{v} \Leftrightarrow A^{-1}A\boldsymbol{x} = A^{-1}\boldsymbol{v} \Leftrightarrow \boldsymbol{x} = A^{-1}\boldsymbol{v} \\ \rightsquigarrow \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{ad-cb}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} \frac{de-bf}{ad-bc} \\ \frac{-ce+af}{ad-bc} \end{pmatrix}. \end{aligned} \tag{5.36}$$

## 5.5 Eigenvalues and eigenvectors

In understanding the nature of a matrix (or really the linear transformation represented by it) we often would like to understand those vectors that are transformed into themselves, i.e., where

$$A\boldsymbol{e} = \lambda\boldsymbol{e}, \qquad (\lambda \text{ is a constant}). \tag{5.37}$$

This equation is called the "eigenvalue problem"; $\lambda$ is called an eigenvalue and $\boldsymbol{e}$ an eigenvector.

The question is now, obviously, how to determine $\lambda$ and $\boldsymbol{e}$. By using a simple trick, we can solve for $\lambda$ without having to know $\boldsymbol{e}$. To this end we write

$$\lambda\boldsymbol{x} = \lambda I\boldsymbol{x} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\boldsymbol{x},$$

and we can thus rewrite the eigenvalue problem as

$$(A - \lambda I)\boldsymbol{x} = \boldsymbol{0}. \tag{5.38}$$

We have argued in the previous section that this has only the trivial solution if $\det(A-\lambda I) \neq 0$, i.e., $\boldsymbol{x} = \boldsymbol{0}$. In order to avoid this we must require

$$\det(A - \lambda I) = 0 \quad . \tag{5.39}$$

For a two-by-two matrix this is a simple quadratic equation.

**Example 5.1:**

Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

**Solution:**

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)^2 - 4 = 0 \end{aligned}$$

Solutions are $\lambda = 1 \pm 2 = -1, 3$.
Now determine the eigenvalues by solving $(A - \lambda I)\boldsymbol{e} = 0$. For $\lambda_1 = -1$ we find

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \mathbf{0}$$

and thus $e_1 = -e_2$, and $\boldsymbol{e}^{(1)} = c \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. The arbitrary constant can be chosen at will. Some standard choices are 1 (simple), $1/\sqrt{2}$ (length 1), etc.

The same algebra for the other eigenvalue leads to $\boldsymbol{e}^{(2)} = d \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

If we rewrite

$$\begin{pmatrix} x \\ y \end{pmatrix} = (x+y)/2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (x-y)/2 \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

we can easily understand the importantce of the eigenvectors:

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

Thus the component parallel to $(1,1)$ is stretches by a factor of 3, and the component parallel to $(-1,1)$ is inverted (multiplied by $-1$).

### 5.5.1 A physics example

The most important physical example of the role of the eigenvalue problem can be found in the case of coupled oscillators. Consider Fig. 5.1. There we show the case of two masses, coupled by three springs. We assume that $x_1$ and $x_2$ are the distances of the masses from the equilibrium position. At that point we assume the strings are untentioned (neither stretched nor compressed).

The equations of motion take a simple form

$$\begin{aligned} m_1 \ddot{x}_1 &= -k_1 x_1 + k_2(x_2 - x_1) = -(k_1 + k_2)x_1 + k_2 x_2 \\ m_2 \ddot{x}_2 &= -k_3 x_2 - k_2(x_2 - x_1) = -(k_3 + k_2)x_2 + k_2 x_1 \end{aligned}$$

We now take the masses equal ($m_1 = m_2 = m$), and all the spring constants equal as well ($k_1 = k_2 = k_3 = m\omega^2$. We then find that

$$\ddot{\boldsymbol{x}} = \omega^2 \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \boldsymbol{x}$$

This equation can now be solved by writing the standard exponential form, $\boldsymbol{x} = \boldsymbol{e} e^{zt}$. We then get

$$z^2 \boldsymbol{e} = \omega^2 \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \boldsymbol{e}$$

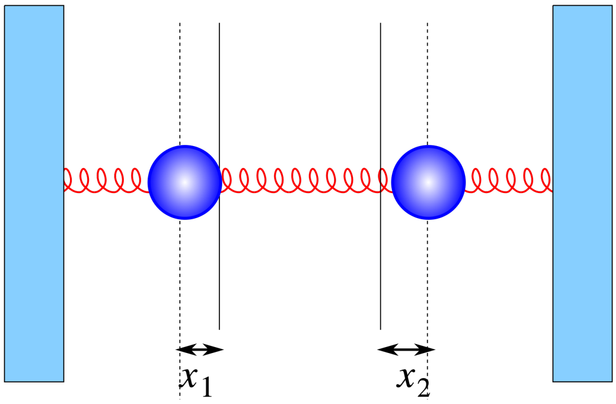



Figure 5.1: Two coupled oscillators.

Which is an eigenvalue problem. Write $z^2 = \omega^2\lambda$, and we find that $\lambda = -1, -3$. Thus $z = \pm i\omega, \pm i\sqrt{3}\omega$. The eigenvectors for these two eigenvalues are $(1, 1)$ and $(1, -1)$, respectively.

Thus, in all its generality, we find using superposition that

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} (A\cos(\omega t) + B\sin(\omega t)) + \begin{pmatrix} 1 \\ -1 \end{pmatrix} (C\cos(\sqrt{3}\omega t) + D\sin(\sqrt{3}\omega t)). \tag{5.40}$$

This general motion thus consists of the superposition of motion of the two masses in phase ($x_1 = x_2$, with frequency $\omega$) and one maximally out of phase ($x_1 = -x_1$, with frequency $\sqrt{3}\omega$).