Lecture Notes by Jürgen Vollmer

Theoretical Mechanics

— Working Copy, Chapter 2 –

— 2021-10-07 04:50:18+02:00—

Fakultät für Physik und Geowissenschaften, Universität Leipzig

LECTURES DELIVERED AT FAKULTÄT FÜR PHYSIK UND GEOWISSENSCHAFTEN, UNIVERSITÄT LEIPZIG
http://www.uni-leipzig.de/physik/~vollmer

Contents

ts v

1 Basic Principles 1

1.1 Basic notions of mechanics 2

1.2 Dimensional analysis 6

nitude guesses 9

10

g 12

2 Balancing Forces and Torques 13

2.1 Motivation and outline: forces are vectors 14

2.2 Sets 15

2.3 Groups 21

2.4 Fields 24

2.5 Vector spaces 27

2.6 Physics application: balancing forces 32

2.7 The inner product 34

2.8 Cartesian coordinates 36

2.9 Cross products — torques 41

2.10 Worked example: Calder's mobiles 48

2.11 Problems 49

2.12 Further reading 58

3 Newton's Laws 59

3.1 Motivation and outline: What is causing motion? 60
3.2 Time derivatives of vectors 60
3.3 Newton's axioms and equations of motion (EOM) 62
3.4 Constants of motion (CM) 70
3.5 Worked example: Flight of an Earth-bound rocket 80
3.6 Problems 83
3.7 Further reading 88

4 Motion of Point Particles 91

4.1 Motivation and outline: EOM are ODEs 92
4.2 Integrating ODEs — Free flight 94
4.3 Separation of variables — Settling with Stokes drag 98
4.4 Worked example: Free flight with turbulent friction 105
4.5 Linear ODEs — Particle suspended from a spring 108
4.6 The center of mass (CM) inertial frame 115
4.7 Worked example: the Kepler problem 120
4.8 Mechanical similarity — Kepler's 3rd Law 121
4.9 Solving ODEs by coordinate transformations — Kepler's 1st law 122
4.10 Problems 126
4.11 Further reading 133

5 Impact of Spatial Extension 135

5.1 Motivation and outline: How do particles collide? 136
5.2 Collisions of hard-ball particles 138
5.3 Volume integrals — A professor falling through Earth 140
5.4 Center of mass and spin of extended objects 147
5.5 Bodies with internal degrees of freedom: Revisiting mobiles 152
5.6 Worked example: Reflection of balls 157
5.7 Problems 158

6 Integrable Dynamics 161

6.1 Motivation and Outline: How to deal with constraint motion? 162
6.2 Lagrange formalism 163

h one degree of freedom 168
6.4 Dynamics with two degrees of freedom 177
6.5 Dynamics of 2-particle systems 180
6.6 Conservation laws, symmetries, and the Lagrange formalism 180
6.7 Worked problems: spinning top and running wheel 180
6.8 Problems 181

7 Deterministic Chaos 183

Take Home Messags 185

A Physical constants, material constants, and estimates 187
A.1 Solar System 187

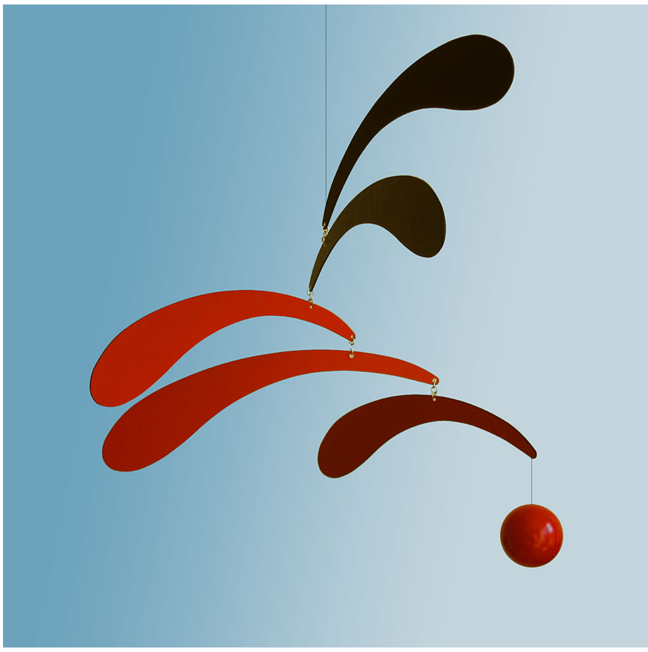
Bibliography 189

Index 191

2

Balancing Forces and Torques

In Chapter 1 we observed that positions and velocities of particles are specified by indicating their unit, magnitude and directions. Hence, they are vectors. In the present chapter we learn how vectors are defined in mathematics, and how they are used and handled in physics. In order to provide a formal definition we introduce a number of mathematical concepts, like groups, that will be revisited in forthcoming chapters. As first important application we deal with balancing forces and torques.



Mobile (sculpture) in the style of Alexander Calder
Andrew Dunn / wikimedia CC BY-SA 2.0

At the end of this chapter we will be able to determine how a mobile hangs from the ceiling.

2.1 *Motivation and outline: forces are vectors*

In mechanics we use vectors to describe forces, displacements and velocities. A displacement describes the relative position of two points in space, and the velocity can be thought of as a distance divided by the time needed to go from the initial to the final point. (A mathematically more thorough definition will be given in Chapter 3.) For forces it is of paramount importance to indicate in which direction they are acting. Similarly, in contrast to speed, a velocity can not be specified in terms of a number with a unit, e.g. 5 m/s. By its very definition one also has to specify the direction of motion. Finally, also a displacement involves a length specification and a direction.

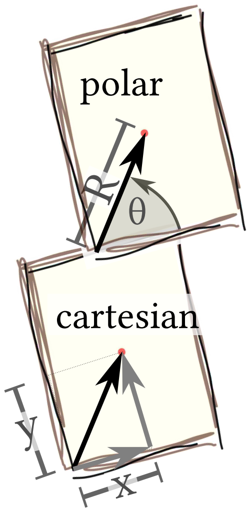


Figure 2.1: The displacement of the red point from the bottom left corner to the the middle of the page can either be specified by the direction θ and the distance R (polar coordinates, top), or by the distances x and y along the sides of the paper (Cartesian coordinates, bottom).

add more explanation

Example 2.1: Displacement of a red dot from the lower left corner to the middle of a paper

This displacement is illustrated in Figure 2.1. It can either be specified in terms of the distance R of the point from the corner and the angle θ of the line connecting the points and the lower edge of the paper (i.e. the direction of the point). Alternatively, it can be given in terms of two distances (x, y) that refer to the length x of a displacement along the edge of the paper and a displacement y in the direction vertical to the edge towards the paper. This can be viewed as result of two subsequent displacements indicated by gray arrows.

In three dimensions, one has to adopt a third direction out of the plane used for the paper, and hence three numbers, to specify a displacements—or indeed any other vector.

	displacement $\boldsymbol{x} = (x_1, x_2, x_3)$	velocity $\boldsymbol{v} = (v_1, v_2, v_3)$	force $\boldsymbol{F} = (f_1, f_2, f_3)$
unit	$[\boldsymbol{x}] = \mathrm{m}$	$[\boldsymbol{v}] = \mathrm{m\,s^{-1}}$	$[\boldsymbol{F}] = \mathrm{kg\ m/s^2}$
magnitude	$\lvert\boldsymbol{x}\rvert = \sqrt{x_1^2 + x_2^2 + x_3^2}$	$\lvert\boldsymbol{v}\rvert = \sqrt{v_1^2 + v_2^2 + v_3^2}$	$\lvert\boldsymbol{F}\rvert = \sqrt{f_1^2 + f_2^2 + f_3^2}$
direction	$\hat{\boldsymbol{x}} = \boldsymbol{x}/\lvert\boldsymbol{x}\rvert$	$\hat{\boldsymbol{v}} = \boldsymbol{v}/\lvert\boldsymbol{v}\rvert$	$\hat{\boldsymbol{F}} = \boldsymbol{F}/\lvert\boldsymbol{F}\rvert$

A basic introduction of mechanics can be given based on this heuristic account of vectors. However, for the thorough exposition that serve as a foundation of theoretical physics a more profound mathematical understanding of vectors is crucial. Hence, a large part of this chapter will be devoted to mathematical concepts.

Outline

In the first part of this chapter we introduce the mathematical notions of sets and groups that are needed to provide a mathematically sound definition of a vector space. Sets are the most fundamental structure of mathematics. It denotes a collection of elements, e.g., numbers like the digits of our number system $\{1, 2, \ldots, 9\}$ or the set of students in my class. Mathematical structures refer to sets where the elements obey certain additional properties, like in groups and vector spaces. They are expressed in terms of *operations* that take one or several elements of the set, and return a result that may or may not be part of the given set. When an operation f takes an element of a set A and returns another element of A we write $f : A \to A$. When an operation $\circ$ takes two elements of a set A and returns a single element of A we write[1] $\circ : A \times A \to A$. Equipped with the mathematical tool of vectors we will explore the physical concepts of forces and torques, and how they are balanced in systems at rest.

[1] Here $A \times A$ is the set, (a_1, a_2), of all pairs of elements $a_1, a_2 \in A$. Further details will be given in Definition 2.3 below.

2.2 *Sets*

In mathematics and physics we often wish to make statements about a collection of objects, numbers, or other distinct entities.

Definition 2.1: Set

A *set* is a gathering of well-defined, distinct objects of our perception or thoughts.
An object a that is part of a set A is an *element* of A;
we write $a \in A$.
If a set M has a finite number n of elements we say that its *cardinality* is n. We write $|M| = n$.

Remark 2.1. Notations and additional properties:

a) When a set M has a finite number of elements, e.g., $+1$ and -1, one can specify the elements by explicitly stating the elements, $M = \{+1, -1\}$. In which order they are states does not play a role, and it also does not make a difference when elements are provided several times. In other words the set M of cardinality two can be specified by any of the following statements

$$M = \{-1, +1\} = \{+1, -1\} = \{-1, 1, 1, 1, \} = \{-1, 1, +1, -1\}$$

b) If e is not an element of a set M, we write $e \notin M$. For instance $-1 \in M$ and $2 \notin M$.

c) There is exactly one set with no elements, i. e. with cardinality zero. It is denoted as empty set, $\emptyset$.

Example 2.2: Sets

- Set of capitals of German states:

  $A_C = $ {Berlin, Bremen, Hamburg, Stuttgart, Mainz, Wiesbaden, München, Magdeburg, Saarbrücken, Potsdam, Kiel, Hannover, Dresden, Schwerin, Düsseldorf, Erfurt}

- Set of small letters in German:

  A_L = {a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z, ä, ö, ü, ß}

- Set of month with 28 days:[2]

  A_M = {January, February, March, April, May, June, July, August, September, October, November, December}

The cardinalities of these sets are

$|A_C| = 16$, $|A_L| = 30$, and $|A_M| = 12$.

[2] Most of them have even more days.

Example 2.3: Sets of sets

A set can be an element of a set. For instance the set

$$M = \{1, 3, \{1, 2\}\}$$

has three elements 1, 3 and $\{1, 2\}$ such that $|M| = 3$, and

$$1 \in M\,, \qquad \{1, 2\} \in M\,, \qquad 2 \notin M \qquad \{1\} \notin M\,.$$

Often it is bulky to list all elements of a set. In obvious cases we use ellipses such as A_L = {a, b, c, ..., z, ä, ö, ü, ß} for the set given in Example 2.2. Alternatively, one can provide a set M by specifying the properties $A(x)$ of its elements x in the following form

$$\underbrace{M}_{\text{The set } M} \underbrace{=}_{\text{contains}} \underbrace{\{}_{\text{all elements}} \underbrace{x}_{x,} \underbrace{:}_{\text{with :}} \underbrace{A(x)}_{\text{properties} \dots} \}.$$

where the properties specify one of several properties of the elements. The properties are separated by commas, and must all be true for all elements of the set.

Example 2.4: Set definition by property

The set of digits $D = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ can also be defined as follows $D = \{1, \dots, 9\} = \{x : 0 < x \leq 9, x \in \mathbb{Z}\}$. In the latter definition $\mathbb{Z}$ denotes the set of all integer numbers.

In order to specify the properties in a compact form we use logical junctors as short hand notation. In the present course we adopt the notations not $\neg$, and $\wedge$, or $\vee$, implies $\Rightarrow$, and is equivalent $\Leftrightarrow$ for the relations indicated in 2.1.

The definition of the digits in Example 2.4 entails that all elements of D are also numbers in $\mathbb{Z}$: we say that D is a subset of $\mathbb{Z}$.

Definition 2.2: Subset and Superset

The set M_1 is a *subset* of M_2, if all elements of M_1 are also contained in M_2. We write[3] $M_1 \subseteq M_2$. We denote M_2 then as *superset* of M_1, writing $M_2 \supseteq M_1$.

[3] Some authors use $\subset$ instead of $\subseteq$, and $\subsetneq$ to denote proper subsets.

Table 2.1: List of the results of different junctors acting on two statements A and B. Here 0 and 1 indicate that a statement is wrong or right, respectively. In the rightmost column we state the contents of the expression in the left column in words. The final three lines provide examples of more complicated expressions.

A	0	0	1	1	
B	0	1	0	1	
$\neg A$	1	1	0	0	not A
$\neg B$	1	0	1	0	not B
$A \vee B$	0	1	1	1	A or B
$A \wedge B$	0	0	0	1	A and B
$A \Rightarrow B$	1	1	0	1	A implies B
$A \Leftrightarrow B$	1	0	0	1	A is equivalent to B
$A \vee \neg B$	1	0	1	1	A or not B
$\neg A \wedge B$	0	1	0	0	not A and B
$A \wedge \neg B$	0	0	1	0	A and not B

The set M_1 is a *proper subset* of M_2 when at least one of the elements of M_2 is not contained in M_1. In this case $|M_1| < |M_2|$, and we write $M_1 \subset M_2$ or $M_2 \supset M_1$.

Example 2.5: Subsets

- The set of month with names that end with "ber" is a subset of the set A_M of Example 2.2

$$\{\text{September, October, November, December}\} \subseteq A_M$$

- For the set M of Example 2.3 one has

$$\{1\} \subseteq M\,, \quad \{1,3\} \subseteq M\,, \quad \{1,2\} \not\subseteq M\,, \quad \{2,\{1,2\}\} \not\subseteq M\,.$$

  Note that $\{1, 2\}$ is an element of M. However, it is not a subset. The last two sets are no subsets because $2 \notin M$.

Two sets are the same when they are subsets of each other.

Theorem 2.1: Equivalence of Sets

Two sets A and B are *equal* or *equivalent*, iff

$$(A \subseteq B) \;\wedge\; (B \subseteq A)\,.$$

Remark 2.2 (iff). In mathematics "iff" indicates that something holds "if and only if". Observe its use in the following two statements: A number is an even number if it is the product of two even numbers. A number is an even number iff it is the product of an even number and another number.

Remark 2.3 (precedence of operations in logical expressions.). In logical expressions we first evaluate $\in$, $\notin$ and other set operations that are used to build logical expressions. Then we evaluate the junctor $\neg$ that is acting on a a single logical expression. Finally the other junctors $\wedge$, $\vee$, $\Rightarrow$, and $\Leftrightarrow$ are evaluated. Hence, the brackets are not required in Theorem 2.1.

Proof of Theorem 2.1.
$A \subseteq B$ implies that $a \in A \Rightarrow a \in B$.
$B \subseteq A$ implies $b \in B \Rightarrow b \in A$.
If $A \subseteq B$ and $B \subseteq A$, then we also have $a \in A \Leftrightarrow a \in B$. □

The description of sets by properties of its members, Example 2.4, suggests that one will often be interested in operations on sets. For instance the odd and even numbers are subsets of the natural numbers. Together they form this set, and one is left with the even numbers when removing the odd numbers from the natural numbers. Hence, we define the following operations on sets.

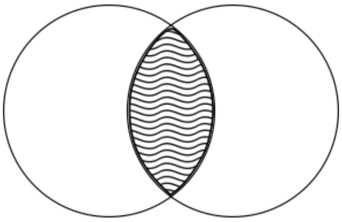


Figure 2.2: Intersection of two sets.

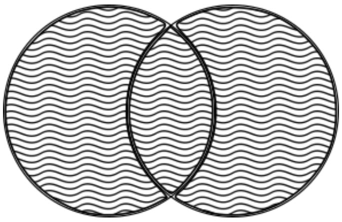


Figure 2.3: Union of two sets.

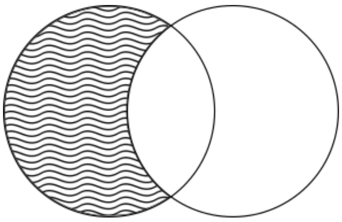


Figure 2.4: Difference of two sets.

Figure 2.5: Complement of a set.

Definition 2.3: Set Operations

For two sets M_1 and M_2 we define the following operations:

- *Intersection*: $M_1 \bigcap M_2 = \{m \,:\, m \in M_1 \wedge m \in M_2\}$,
- *Union*: $M_1 \bigcup M_2 = \{m \,:\, m \in M_1 \vee m \in M_2\}$,
- *Difference*: $M_1 \backslash M_2 = \{m \,:\, m \in M_1 \wedge m \notin M_2\}$,
- The *complement* of a set M in a *universe* U is defined for subsets $M \subseteq U$ as $M^C = \{m \in U \,:\, m \notin M\} = U \backslash M$.
- The *Cartesian product of two sets M_1 and M_2* is defined as the set of ordered pairs (a, b) of elements $a \in M_1$ and $b \in M_2$: $M_1 \times M_2 = \{(a, b) \,:\, a \in M_1, b \in M_2\}$.

A graphical illustration of the operations is provided in Figures 2.2 to 2.5.

Example 2.6: Set operations: participants in my class

Consider the set of participants P in my class. The sets of female F and male M participants of the class are proper subsets of P with an empty intersection $F \bigcap M$. The set of non-female participants is $P \backslash F$. The set of heterosexual couples in the class is a subset of the Cartesian product $F \times M$. Furthermore, the union $F \bigcup M$ is a proper subset of P, when there is a participant who is neither female nor male.

Definition 2.4: Logical quantors

A logical statement S about elements a of a set A may hold

- for all elements of a set — we write: $\forall a \in A \,:\, S$
- for some elements of a set — we write: $\exists a \in A \,:\, \mathsf{S}$
- for exactly one elements of a set — we write: $\exists! a \in A \,:\, \mathsf{S}$

name	symbol	description
natural numbers	$\mathbb{N}$	$\{1, 2, 3, \dots\}$
natural numbers with 0	$\mathbb{N}_0$	$\mathbb{N} \bigcup \{0\}$
negative numbers	$-\mathbb{N}$	$\{-n \,:\, n \in \mathbb{N}\}$
even numbers	$2\mathbb{N}$	$\{2n \,:\, n \in \mathbb{N}\}$
odd numbers	$2\,\mathbb{N} - 1$	$\{2n - 1 \,:\, n \in \mathbb{N}\}$
integer numbers	$\mathbb{Z}$	$(-\mathbb{N}) \bigcup \mathbb{N}_0$
rational numbers	$\mathbb{Q}$	$\left\{\frac{p}{q} \,:\, p \in \mathbb{Z}, q \in \mathbb{N}\right\}$
real numbers	$\mathbb{R}$	see below
complex numbers	$\mathbb{C}$	$\mathbb{R} + \mathrm{i}\mathbb{R}$, where $\mathrm{i} = \sqrt{-1}$

Table 2.2: Summary of important sets of numbers.

Example 2.7: Logical quantors and properties of set elements

Let $|m|$ denote the number of days in a month $a \in A_M$ (cf. Example 2.2). Then the following statements are true: There is exactly one month that has exactly 28 days:

$$\exists! a \in A_M \,:\, |a| = 28$$

Some months have exactly 30 days:

$$\exists a \in A_M \,:\, |a| = 30$$

All month have at least 28 days:

$$\forall a \in A_M \,:\, |a| \geq 28$$

2.2.1 *Sets of Numbers*

Many sets of numbers that are of interest in physics have infinitely many elements. We construct them in Table 2.2 based on the natural numbers

$\mathbb{N}$: check ISO norm

$\mathbb{N}$: remark on Neumann construction?

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

or the natural numbers with zero

$$\mathbb{N}_0 = \mathbb{N} \bigcup \{0\} \; .$$

Remark 2.4. Some authors adopt the convention that zero is included in the natural numbers $\mathbb{N}$. When this matters you have to check which convention is adopted.

There are many more sets of numbers. For instance, in mathematics the set of constructable numbers is relevant for certain proofs in geometry, and in physics and computer graphics quaternions are handy when it comes to problems involving three-dimensional rotations. In any case one needs intervals of numbers.

Definition 2.5: Interval of Real Numbers $\mathbb{R}$

An *interval* is a continuous subset of a set of numbers. We distinguish *open, closed, and half-open subsets*.

- closed interval: $[a,\, b] = \{x \,:\, x \geq a,\, x \leq b\}\,,$
- open interval: $(a,\, b) =]a,\, b[= \{x \,:\, x > a,\, x < b\}\,,$
- right open interval: $[a,\, b) = [a,\, b[= \{x \,:\, x \geq a,\, x < b\}\,,$
- left open interval: $(a,\, b] =]a,\, b] = \{x \,:\, x > a,\, x \leq b\}\,.$

Subsets of $\mathbb{R}$ will be denoted as real intervals.

add limits, closure, and $\mathbb{R}$ as closure of $\mathbb{Q}$.

2.2.2 *Self Test*

Problem 2.1. **Relations between sets**

Let A, B, C, and D be pairwise distinct elements. Select one of the symbols

$$\in,\quad \notin,\quad \ni,\quad \not\ni,\quad \subset,\quad \not\subset,\quad \supset,\quad \not\supset,\quad =$$

and avoid $\notin$, $\not\ni$, $\not\subset$, $\not\supset$ wherever possible.

a) $\{A, B\}\ \square\ \{A, B, C\}$,

b) $\{A\}\ \square\ B$,

c) $\{\emptyset\}\ \square\ \emptyset$,

d) $\{\{A\}\}\ \square\ \{\{A\}, \{B\}\}$,

e) $A\ \square\ \{A, B, C\}$,

f) $\{A, C, D\} \cap \{A, B\}\ \square\ \{A, B, C, D\}$,

g) $\{A, C, D\} \setminus \{A, B\}\ \square\ \{A, B, C\}$,

h) $\{A, C, D\} \cup \{A, B\}\ \square\ A$.

Problem 2.2. **Intervals**

a) Provide $[1; 17] \cap]0; 5[$ as a single interval.

b) Provide $[-1, 4] \setminus [1, 2[$ as union of two intervals.

Problem 2.3. **Sets of numbers**

Which of the following statements are true?

a) $\{6 \cdot z | z \in \mathbb{Z}\} \subset \{2 \cdot z | z \in \mathbb{Z}\}$.

b) $\{2 \cdot z | z \in \mathbb{Z}\} \cap \{3 \cdot z | z \in \mathbb{Z}\} = \{6 \cdot z | z \in \mathbb{Z}\}$.

c) Let $T(a)$ be the set of numbers that divide a. Then

$$\forall a, b \in \mathbb{N} : \quad T(a) \cup T(b) = T(a \cdot b)$$

Example: $T(2) = \{1, 2\}$, $T(3) = \{1, 3\}$, and $T(6) = \{1, 2, 3, 6\}$.

2.3 Groups

A group G refers to a set of operations $t \in G$ that are changing some data or objects. Elementary examples refer to reflections in space, turning some sides of a Rubik's cube, or translations in space, as illustrated in Figure 2.1. The subsequent action of two group elements t_1 and t_2 of G is another (typically more complicated) transformation $t_3 \in G$. Analogous to the concatenation of functions, we write $t_3 = t_2 \circ t_1$, and we say t_3 is t_2 after t_1. The set of transformations forms a group iff it obeys the following rules.

Definition 2.6: Group

A set $(G, \circ)$ is called a *group* with operation $\circ : G \times G \to G$ when the following rules apply

a) The set is *closed*: $\forall g_1, g_2 \in G : g_1 \circ g_2 \in G.$

b) The set has a *neutral element*: $\exists e \in G\ \forall g \in G : e \circ g = g.$

c) Each element has an *inverse element*:
$$\forall g \in G\ \exists i \in G : g \circ i = e.$$

d) The operation $\circ$ is *associative*:
$$\forall g_1, g_2, g_3 \in G : (g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3).$$

Wüstholz rather suggests: $\forall e \in G : e \circ f = f \Rightarrow \forall g \in G\ \exists h \in G : h \circ g = e$

Definition 2.7: Commutative Group

A group $(G, \circ)$ is called a *commutative group* when

e) the group operation is *commutative*:
$$\forall g_1, g_2 \in G : g_1 \circ g_2 = g_2 \circ g_1.$$

Remark 2.5. Commutative groups are also denoted as *Abelian groups*.

When the group has a finite number of elements the result of the group operation can explicitly be specified by a group table. We demonstrate this by the smallest groups. The empty set can not be a group because it has no neutral element. Therefore the smallest groups have a single element and two elements. Both of these groups are commutative.

Example 2.8: Smallest groups

$(\{n\}, \odot)$ comprises only the neutral element.

$\odot$	n
n	n

The smallest non-trivial group has two elements $(\{0, 1\}, \oplus)$ with

$\oplus$	0	1
0	0	1
1	1	0

It describes the turning of a piece of paper:
Not turning, 0, does not change anything (neutral element).
Turning, 1, shows the other side, and turning twice is equivalent to not turning at all (1 is its own inverse).

Remark 2.6. The group properties imply that all elements of the group must appear exactly once in each row and each column of the group table. As a consequence the smallest non-commutative group is the dihedral group of order 6 with six elements that is discussed in Problem 2.7.

Figure 2.6: Rotation of a book by multiples of $\pi/2$ around three orthogonal axes.

watch out: there is a problem with inverse & neutral elts!

Example 2.9: Non-commutative groups: rotations

The rotation of an object in space is a group. In particular this holds for the $90°$-rotations of an object around a vertical and a horizontal axis. Figure 2.6 illustrates that these rotations do not commute.

Example 2.10: Non-commutative groups: edit text fields

We consider the text fields of a fixed length n in an electronic form. Then the operations
"Put the letter L into position $\sqcup$ of the field"
with $L \in \{\_, a, \ldots, z, A, \ldots, Z\}$
and $\sqcup \in \{1, \ldots, n\}$ form a group.
Also in this case one can easily check that the order of the operations is relevant. In the left and right column the same operations are preformed for a text field of length $n = 4$:

$$
\begin{array}{rr}
|\_|\_|\_|\_| & |\_|\_|\_|\_| \\
\to |M|\_|\_|\_| & \to |P|\_|\_|\_| \\
\to |M|a|\_|\_| & \to |P|h|\_|\_| \\
\to |M|a|t|\_| & \to |P|h|y|\_| \\
\to |M|a|t|h| & \to |P|h|y|s| \\
\to |M|a|t|s| & \to |P|h|y|h| \\
\to |M|a|y|s| & \to |P|h|t|h| \\
\to |M|h|y|s| & \to |P|a|t|h| \\
\to |P|h|y|s| & \to |M|a|t|h|
\end{array}
$$

Remark 2.7. Notations and additional properties:

a) Depending of the context the inverse element is denoted as g^{-1} or as $-g$. This depends on whether the operation is considered a multiplication or rather an addition. In accordance with this choice the neutral element is denoted as 1 or 0.

b) The second property of groups, b) $\exists e \in G\ \forall g \in G : e \circ g = g$, implies that also $g \circ e = g$. The proof is provided as Problem 2.8.

c) When a group is not commutative then one must distinguish the left and right inverse. The condition $g \circ i = e$ does not imply $i \circ g = e$. However, there always is another element $j \in G$ such that $j \circ g = e$. An example is provided in Problem 2.7.

2.3.1 *Self Test*

Problem 2.4. **Checking group axioms**

Which of the following sets are groups?

a) $(\mathbb{N}, +)$

b) $(\mathbb{Z}, +)$

c) $(\mathbb{Z}, \cdot)$

d) $(\{+1, -1\}, \cdot)$

e) $(\{0\}, +)$

$(\{1, \dots, 12\}, \oplus)$

where $\oplus$ in f) revers to adding as we do it on a clock, e.g. $10 \oplus 4 = 2$.

Problem 2.5. **The group with three elements**

Let $(\mathcal{G}, \circ)$ be a group with three elements $\{n, l, r\}$, where n is the neutral element.

a) Show that there only is a single choice for the result of the group operations $a \circ b$ with $a, b \in \mathcal{G}$. Provide the group table.

b) Verify that the group describes the rotations of an equilateral triangle that interchange the positions of the angles.

c) Show that there is a bijective map $\mathrm{m} : \{n, l, r\} \to \{0, 1, 2\}$ with the following property:

$$\forall a, b \in \mathcal{G} \; : \; a \circ b = \big(\mathrm{m}(a) + \mathrm{m}(b)\big) \mathrm{mod} 3 \, .$$

We say that the group $\mathcal{G}$ is isomorphic to the natural numbers with addition addition modulo 3.[4]

[4] The natural number modulo n amount to n classes that represent the remainder of the numbers after division by n. For instance, for the natural numbers modulo two the 0 represents even numbers, and the 1 odd numbers. Similarly, for the natural numbers modulo three the 0 represents numbers that are divisible by three, and for the sum of 2 and 2 modulo 3 one obtains $(2 + 2) \mathrm{mod} 3 = 4 \mathrm{mod} 3 = 1$.

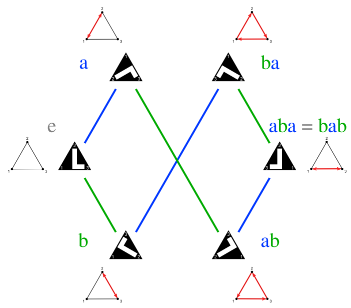
Problem 2.6. **Symmetry group of rectangles**

A polygon has a symmetry with an associated symmetry operation a when a only interchanges the vertices of the polygon. It does not alter the position. To get a grip on this concept we consider the symmetry operations of a rectangle.

a) Sketch how reflections with respect to a symmetry axis interchange the vertices of a rectangle. What happens when the reflections are repeatedly applied?

b) Show that the symmetry operations form a group with four elements. Provide a geometric interpretation for all group elements.

c) Provide the group table.



Watchduck (a.k.a. Tilman Piesk), wikimedia CC BY-SA

Figure 2.7: Reflections of equilateral triangle with respect to the three symmetry axes form a group with six elements; see Problem 2.7.

Problem 2.7. **Dihedral group of order** 6

Figure 2.7 illustrates the effect of reflections of a triangle with respect to its three symmetry axis. All group elements can be generated by repeated action of two reflections, e.g. those denoted as a and b in the figure.

a) Verify that the group properties, Definition 2.6, together with the three additional requirements

$$a \circ a = b \circ b = e \quad \text{and} \quad a \circ b \circ a = b \circ a \circ b$$

imply that the group has exactly six elements,

$$\mathcal{G} = \{e, a, b, a \circ b, b \circ a, a \circ b \circ a\}\,.$$

b) Work out the group table.

c) Verify by inspection that e is the neutral element for operation from the right *and* from the left.

d) Verify that the group is *not* commutative, and provide an example of a group element where the left inverse and the right inverse differ.

e) The group can also be represented in terms of a reflection and the rotations described in Problem 2.5. How would the graphical representation, analogous to Figure 2.7, look like in that case.

Problem 2.8. **Uniqueness of the neutral element**

Proof that the group axioms, Definition 2.6, imply that $e \circ g = g$ implies that also $g \circ e = g$.

2.4 *Fields*

Besides being of importance to characterize the action of symmetry operations like reflections or rotations, groups are also important for us because they admit further characterization of sets of numbers.

The natural numbers are not a group. For the addition they are lacking the neutral elements, and for adding and multiplications they are lacking inverse elements.

In contrast the group $(\mathbb{Z}, +)$ is a commutative group with infinitely many elements.

Example 2.11: The group $(\mathbb{Z}, +)$

The numbers $\mathbb{Z}$ with operation $+$ form a group. This is demonstrated here by checking the group axioms.

a) Addition of any two numbers provides a number:

$\forall x, y \in \mathbb{Z} \, : \, (x + y) \in \mathbb{Z}$.

b) The neutral element of the addition is 0:
$\exists\, 0 \in \mathbb{Z}\ \forall z \in \mathbb{Z} \, : \, z + 0 = z = 0 + z$.

c) For every element $z \in \mathbb{Z}$ there is an inverse $(-z) \in \mathbb{Z}$:
$\forall z \in \mathbb{Z}\ \exists (-z) \in \mathbb{Z} \, : \, z + (-z) = 0 = (-z) + z$.

d) The addition of numbers is associative:
$\forall z_1, z_2, z_3 \in \mathbb{Z} \, : \, z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$.

However, the numbers $\mathbb{Z}$ still lack inverse elements of the multiplication. The rational numbers $\mathbb{Q}$ and the real numbers $\mathbb{R}$ are commutative groups for addition and multiplication (with the special rule that multiplication with 0 has no inverse element), and their elements also obey distributivity. Such sets are called number fields.

Definition 2.8: Field

A set $(\mathbb{F}, +, \cdot)$ is called a *field* with neutral elements 0 and 1 for addition $+$ and multiplication $\cdot$, respectively, when its elements comply with the following rules

a) $(\mathbb{F}, +)$ is a commutative group,

b) $(\mathbb{F} \backslash \{0\}, \cdot)$ is a commutative group,

c) Addition and Multiplication are distributive:

$$\forall a, b, c \in \mathbb{F} : a \cdot (b + c) = a \cdot b + a \cdot c$$

Remark 2.8. For the multiplication of field elements one commonly suppresses the $\cdot$ for the multiplication, writing e.g. $a\,b$ rather than $a \cdot b$.

Example 2.12: The smallest field has two elements

The smallest field $(\{0, 1\}, \oplus, \odot)$ comprises only the neutral elements 0 of the group $(\{0, 1\}, \oplus)$ with two elements, and 1 of the group $(\{1\}, \odot)$ with one element.

Example 2.13: Complex numbers are a field

a) The sum of two complex numbers $z_1 = x_1 + \mathrm{i}y_1$ and $z_2 = x_2 + \mathrm{i}y_2$ amounts to

$$z_1 + z_2 = (x_1 + \mathrm{i}y_1) + (x_2 + \mathrm{i}y_2) = (x_1 + x_2) + \mathrm{i}(y_1 + y_2)$$

Hence, the group properties for $+$ follow from the properties of the real numbers x_1, x_2 and y_1, y_2, respectively.
b) They also entail distributivity of complex numbers.
c) The product of the complex numbers $z_1 = x_1 + \mathrm{i}y_1$ and

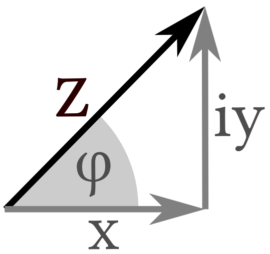


Figure 2.8: Complex numbers z can be represented as $z = x + \mathrm{i}y$ in a plane where (x, y) are the Cartesian coordinates of z. Alternatively, one can adopt a representation in terms of polar coordinates $z = R\,\mathrm{e}^{\mathrm{i}\varphi}$ where $R = \sqrt{x^2 + y^2}$ and φ is the angle with respect to the x-axis.

$z_2 = x_2 + \mathrm{i}y_2$ amounts to

$$\begin{aligned} z_1 \cdot z_2 &= (x_1 + \mathrm{i}y_1) \cdot (x_2 + \mathrm{i}y_2) \\ &= (x_1\, x_2 + \mathrm{i}y_1\, x_2 + \mathrm{i}y_1\, x_2 + \mathrm{i}^2\, y_1\, y_2) \\ &= (x_1\, x_2 - y_1\, y_2) + \mathrm{i}\,(y_1\, x_2 + x_1\, y_2) \end{aligned}$$

Checking the group axioms based on this representation of the complex numbers is tedious. One better adopts a representation in terms of polar coordinates, $z_1 = R_1\,\mathrm{e}^{\mathrm{i}\varphi_1}$ and $z_2 = R_2\,\mathrm{e}^{\mathrm{i}\varphi_2}$ (see Figure 2.8) where (cf Problem 2.10)

$$z_1 \cdot z_2 = R_1\,\mathrm{e}^{\mathrm{i}\varphi_1} \cdot R_2\,\mathrm{e}^{\mathrm{i}\varphi_2} = \big(R_1\, R_2\big)\,\mathrm{e}^{\mathrm{i}\,(\varphi_1+\varphi_2)}$$

Here, the group properties follow from those of multiplying R_1 and R_2, and adding φ_1 and φ_2.

Remark 2.9 (complex conjugation). Each complex numbers z has a complex conjugate, denoted as z^* or $\bar{z}$, that is defined as

$$\forall z = x + \mathrm{i}y = R\,\mathrm{e}^{\mathrm{i}\varphi} \in \mathbb{C}: \quad \bar{z} = x - \mathrm{i}y = R\,\mathrm{e}^{-\mathrm{i}\varphi} \tag{2.4.1}$$

Complex conjugation provides an effective way to calculate the absolute value $|z| = R$ of complex numbers

$$\begin{aligned} & z\,\bar{z} = (x + \mathrm{i}y)\,(x - \mathrm{i}y) = x^2 - \mathrm{i}^2\, y^2 = x^2 + y^2 = R^2 \\ \text{and}\quad & z\,\bar{z} = R\,\mathrm{e}^{\mathrm{i}\varphi}\; R\,\mathrm{e}^{-\mathrm{i}\varphi} = R^2\,\mathrm{e}^0 = R^2 \\ \Rightarrow\quad & |z| = \sqrt{z\,\bar{z}} = \sqrt{\bar{z}\,z} \end{aligned} \tag{2.4.2}$$

Remark 2.10. In physics complex numbers are commonly applied to describe rotations in a plane: Multiplication by $\mathrm{e}^{\mathrm{i}\theta}$ rotates a complex number z by an angle θ around the origin:

$$\forall z = R\,\mathrm{e}^{\mathrm{i}\varphi} \in \mathbb{C}: \quad z \cdot \mathrm{e}^{\mathrm{i}\theta} = R\,\mathrm{e}^{\mathrm{i}(\varphi+\theta)} \tag{2.4.3}$$

2.4.1 *Self Test*

Problem 2.9. **Checking field axioms**

Which of the following sets are fields?

a) $(\mathbb{Z}, +, \cdot)$

b) $(\{1, 2, \ldots, 12\}, +\mathrm{mod}12, \cdot\mathrm{mod}12)$
like on a clock: $11 \oplus 2 = 13\mathrm{mod}12 = 1$ and $4 \odot 5 = 20\mathrm{mod}12 = 8$.

c) $(\{0, 1, 2\}, +\mathrm{mod}3, \cdot\mathrm{mod}3)$
for instance $2 \odot 2 = 2 + 2 = 4\mathrm{mod}3 = 1$ and $2 \oplus 1 = 3\mathrm{mod}3 = 0$.

Problem 2.10. **Euler's equation and trigonometric relations**

Euler's equation $e^{ix} = \cos x + i \sin x$ relates complex values exponential functions and trigonometric functions.

a) Sketch the position of $R\,e^{ix}$ in the complex plane, and indicate how Euler's equation is related to the Theorem of Pythagoras.

b) Complex valued exponential functions obey the same rules as their real-valued cousins. In particular, for $R = 1$ one has $e^{i\,(x+y)} = e^{ix}\,e^{iy}$. Compare the real and complex parts of the expressions on both sides of this relation. What does this imply about $\sin(2x)$ and $\cos(2x)$?

2.5 *Vector spaces*

With the notions introduced in the preceding sections we can give now the formal definition of a vector space

Definition 2.9: Vector Space

A *vector space* $(\mathsf{V}, \mathbb{F}, \oplus, \odot)$ is a set of *vectors* $\boldsymbol{v} \in \mathsf{V}$ over a field $(\mathbb{F}, +, \cdot)$ with binary operations $\oplus : \mathsf{V} \times \mathsf{V} \to \mathsf{V}$ and $\odot : \mathbb{F} \times \mathsf{V} \to \mathsf{V}$ complying with the following rules

a) $(\mathsf{V}, \oplus)$ is a commutative group

b) associativity: $\forall a, b \in \mathbb{F}\ \forall \boldsymbol{v} \in \mathsf{V} : a \odot (b \odot \boldsymbol{v}) = (a \cdot b) \odot \boldsymbol{v}$

c) distributivity 1:
$$\forall a, b \in \mathbb{F}\ \forall \boldsymbol{v} \in \mathsf{V} : (a + b) \odot \boldsymbol{v} = (a \odot \boldsymbol{v}) \oplus (b \odot \boldsymbol{v})$$

d) distributivity 2:
$$\forall a \in \mathbb{F}\ \forall \boldsymbol{v}, \boldsymbol{w} \in \mathsf{V} : a \odot (\boldsymbol{v} \oplus \boldsymbol{w}) = (a \odot \boldsymbol{v}) \oplus (a \odot \boldsymbol{w})$$

Remark 2.11. It is common to use $+$ and $\cdot$ instead of $\oplus$ and $\odot$, respectively, with the understanding that it is clear from the context in the equation whether the symbols refer to operations involving vectors, only numbers, or a number and a vector.

Moreover, as for the multiplication of numbers, one commonly drops the $\odot$ for the multiplication, writing e.g. $a\,\boldsymbol{v}$ rather than $a \cdot \boldsymbol{v}$.

Example 2.14: Vector spaces: displacements in the plane

For displacements we define the operation $\oplus$ as concatenation of displacements, and $\odot$ as increasing the length of the displacement by a given factor without touching the direction.
a) The neutral element amounts to staying, one can always shift back, move between any two points in a plane, and commutativity follows form the properties of parallelo-

grams, see Figure 2.9.
b,c) The vectors select the direction. Scalar multiplication only changes the length of the vectors, and the length is a real number.
d) Is implied by the Intercept Theorem.

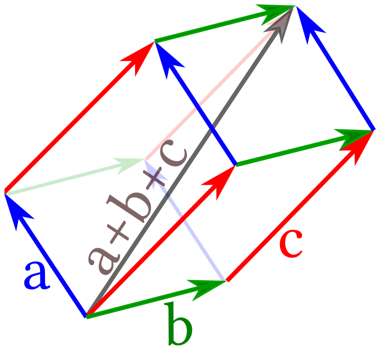


Figure 2.9: The arrows indicate displacements by three vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, as discussed in Example 2.14. Their commutativity and associativity follow from the properties of parallelograms. This holds in the plane, and also when the vectors span a three-dimensional volume.

Example 2.15: Vector spaces: $\mathbb{R}^D$

For every $D \in \mathbb{N}$ the D-fold Cartesian product $\mathbb{R}^D$ of the real numbers is a vector space over $\mathbb{R}$ when defining the operation $+$ and $\cdot$ as

$$\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^D : \boldsymbol{a} + \boldsymbol{b} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_D \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_D \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_D + b_D \end{pmatrix}$$

$$\forall s \in \mathbb{R}\ \forall \boldsymbol{a} \in \mathbb{R}^D : s \cdot \boldsymbol{a} = s \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_D \end{pmatrix} = \begin{pmatrix} s\, a_1 \\ s\, a_2 \\ \vdots \\ s\, a_D \end{pmatrix}$$

In a more compact manner this is also written as,

$$\forall \boldsymbol{a} = (a_i), \boldsymbol{b} = (b_i), s \in \mathbb{R} : \boldsymbol{a} + \boldsymbol{b} = (a_i + b_i) \ \wedge\ s\, \boldsymbol{a} = (s\, a_i)$$

Checking the properties of a vector space is given as Problem 2.11a).

Definition 2.10: $N \times M$ Matrix: $\mathbb{M}^{N\times M}(\mathbb{F})$

For $N, M \in \mathbb{N}$ we define $N \times M$ *matrices* $A, B \in \mathbb{M}^{N\times M}(\mathbb{F})$ over the field $\mathbb{F}$ as arrays, $A = (a_{ij})$, $B = (b_{ij})$, with components $a_{ij}, b_{ij}, \in \mathbb{F}$.
The indices $i \in \{1, \dots, N\}$ and $j \in \{1, \dots, M\}$ label the rows and columns of the array, respectively.
The sum of matrices and the product with a scalar are defined component-wise as

$$\forall A, B \in \mathbb{M}^{N\times M},\ c \in \mathbb{F} : A + B = (a_{ij} + b_{ij}) \ \wedge\ c \cdot A = (c\, a_{ij})$$

Example 2.16: 2×3 matrices: summation and multiplication with a scalar

To be specific we provide here the sum of two 2×3 matrices and the multiplication by a factor of π. Let

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 7 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 12 & 13 \\ 14 & 15 \\ 16 & 17 \end{pmatrix} .$$

Then

$$A + B = \begin{pmatrix} 2+12 & 3+13 \\ 4+14 & 5+15 \\ 6+16 & 7+17 \end{pmatrix} = \begin{pmatrix} 14 & 16 \\ 18 & 20 \\ 22 & 24 \end{pmatrix}$$

$$\pi\, A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 7 \end{pmatrix} = \begin{pmatrix} 2\,\pi & 3\,\pi \\ 4\,\pi & 5\,\pi \\ 6\,\pi & 7\,\pi \end{pmatrix}$$

Example 2.17: Vector spaces: $M \times N$ matrices

The $N \times M$ matrices over a field $\mathbb{F}$, $(\mathbb{M}^{N\times M}, \mathbb{F}, +, \cdot)$ form a vector space. The proof is given as Problem 2.11b).

Definition 2.11: Matrix multiplication

For matrices one defines a product as follows

$$\odot : \mathbb{M}^{N\times L} \times \mathbb{M}^{L\times M} \to \mathbb{M}^{N\times M}$$

$$\forall A \in \mathbb{M}^{N\times L}, B \in \times\mathbb{M}^{L\times M} : A \odot B = C = (c_{ij}) = \left(\sum_{k=1}^{L} a_{ik} b_{kj} \right)$$

Remark 2.12. Also for matrix multiplication one commonly suppresses the $\odot$ operator, writing $A\,B$ rather than $A \odot B$.

Remark 2.13. For square matrices $\mathbb{M}^{M\times M}$ the operation $+$ and $\odot$ define a sum and a product that take two elements of $\mathbb{M}^{M\times M}$ and return an element of $\mathbb{M}^{M\times M}$. Nevertheless, $(\mathbb{M}^{M\times M}, +, \odot)$ is *not* a field: In general, $\odot$ is not commutative and matrices do not necessarily have an inverse.

Remark 2.14. Square matrices can be used to represent reflections and rotations. In Problem 2.12 we provide an example of eight matrices that form a symmetry group.

Example 2.18: Vector spaces: Polynomials of degree 2

For a field $\mathbb{F}$ the polynomials P_2 of degree two in the variable x are defined as

$$\mathsf{P}_2 = \{ \boldsymbol{p} = [p_0 + p_1\,x + p_2\,x^2] : p_0, p_1, p_2 \in \mathbb{F} \}$$

This set is a vector space with respect to the summation

$$\begin{aligned} \boldsymbol{p} + \boldsymbol{q} &= [p_0 + p_1\,x + p_2\,x^2] + [q_0 + q_1\,x + q_2\,x^2] \\ &= \Big[(p_0 + q_0) + (p_1 + q_1)\,x + (p_2 + q_2)\,x^2 \Big] \end{aligned}$$

and the multiplication with a scalar $s \in \mathbb{F}$

$$s \cdot \boldsymbol{p} = s \cdot (p_0 + p_1\,x + p_2\,x^2) = \Big[(s\,p_0) + (s\,p_1)\,x + (s\,p_2)\,x^2) \Big]$$

Proof. Each element $\boldsymbol{p} = [p_0 + p_1\,x + p_2\,x^2]$ of this vector space is uniquely described by the three-tuple $(p_0, p_1, p_2) \in \mathbb{F}^3$ with rules for addition and scalar multiplication analogous to those discussed for $\mathbb{R}^3$ in Example 2.15. Hence, the proof for $\mathbb{R}^3$ also applies here. □

clarify

In physics we heavily make use of the correspondence evoked by the proof in Example 2.18. The relative position of two objects with respect to each other is commonly described in terms of (the sum of several) vectors. In order to gain further information about the positions, we will then recast the *geometric* problem about the positions into an *algebraic* problem stated in terms of linear equations. The latter can then be solved by straightforward analytical calculations. Vice versa, abstract findings about the solutions of sets of equations will be recast in terms of geometry in order to visualize the abstract results. The change of perspective has become a major avenue to drive theoretical physics throughout the 20th century. For mechanical problems it forms the core of the mathematical formulation of problems in robotics and computer vision. Quantum mechanics is entirely build on the principles of vector spaces and their generalization to Hilbert spaces. General relativity and quantum field theory take Noether's theorem as their common starting point, which is build upon concepts from group theory and the requirement that physical predictions must not change when taking different choices how to mathematically describe the system. An important concern of these notes is to serve as a training ground to practice the changing of mathematical perspective for the purpose of solving physics problem. As a first physical application we discuss now force balances. Then we resume the discussion of vector spaces, taking a closer look into the calculation of coordinates and distances.

2.5.1 *Self Test*

Problem 2.11. **Checking vector-space properties**

a) Verify that $\mathbb{R}^D$ with the operations defined in Example 2.15 is a vector space.

b) Verify that $N \times M$ matrices, as defined in Definition 2.10, form a vector space.

Problem 2.12. **Geometric interpretation of matrices** We explore the set of the eight matrices

$$M = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \begin{pmatrix} 0 & c \\ d & 0 \end{pmatrix}, \quad \text{with} \quad a, b, c, d \in \{\pm 1\} \right\}$$

a) Let the action $\circ$ denotes matrix multiplication. Verify that $(M, \odot)$ is a group with respect to matrix multiplication, as defined in Definition 2.11. We denote its neutral element as $\mathbb{I}$.

b) Show that the group has five non-trivial elements $s_1, \dots s_5$ that are self inverse:

$$s_i \neq \mathbb{I} \quad \wedge \quad s_i \circ s_i = \mathbb{I} \quad \text{for} \quad i \in \{1, \dots 6\} \,.$$

c) Show that the other two elements d and r obey $d \circ r = r \circ d = \mathbb{I}$, that $r = d \circ d \circ d$, and that $d = r \circ r \circ r$.

d) Show that the set of points $P = \{(1,1), (-1,1), (-1,-1), (1,-1)\}$ is mapped to P by the action of an element of the group:

$$\forall m \in M \ \wedge \ p \in P : \quad p \circ m \in P$$

Hint: The action of the matrix on the vector defined as follows

$$(v_1, v_2) \circ \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \begin{pmatrix} v_1\, m_{11} + v_2\, m_{21} \\ v_1\, m_{12} + v_2\, m_{22} \end{pmatrix}$$

e) What is the geometric interpretation of the group M? Illustrate the action of the group elements in terms of transformations of a suitably chosen geometric object.

Problem 2.13. **Polynomials of degree** N

For a field $\mathbb{F}$ the polynomials P_N of degree N in the variable x are defined as

$$\mathsf{P}_N = \left\{ \boldsymbol{p} = \left[\sum_{i=0}^{N} p_i\, x^i \right] : p_0, \dots, p_N \in \mathbb{F} \right\}$$

a) State the rules of addition and multiplication with a scalar $s \in \mathbb{F}$ in analogy to the special case of $N = 2$ discussed in Example 2.18.

b) Verify that the polynomials of degree N are a vector space.

Tug of War, Nikolay Bogdanov-Belsky, 1939
wikiart / public domain

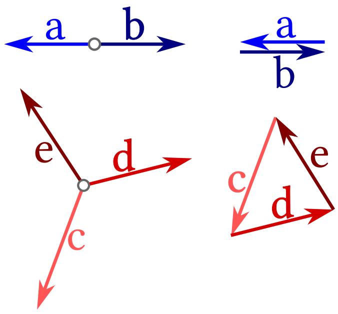


Figure 2.10: The left diagrams show two and three forces acting on a ring. To the right it is demonstrated that they add to zero.

2.6 Physics application: balancing forces

It is an experience from tug of war that nothing moves as long as forces are balanced. In this example one can add a ring to the rope. The pulling forces act in opposing directions on the ring, as illustrated in the upper left diagram in Figure 2.10. The lower left diagram shows the case, where three parties are pulling on the ring. In any case the total force on the ring amounts to the sum of the acting forces, forces are vectors, and all sums of vectors obey the same rules. As far as graphical illustrations are concerned the sum of forces looks therefore the same as the sum of displacements in Figure 2.1. For the ring the sums of the forces are illustrated in the right panels of Figure 2.10. The ring does not move when they add to zero.

Axiom 2.1: Force balance

Let N forces $\boldsymbol{F}_1, \dots, \boldsymbol{F}_N$ act on a body. The body does not move as long as the forces add to zero, i.e. iff $\boldsymbol{0} = \sum_{i=1}^{N} \boldsymbol{F}_i$.

Remark 2.15. Strictly speaking the body might turn, but its center of mass will not move. We come back to this point in Section 2.9.

explain center of mass

Example 2.19: Balancing on a slackline

A person balances on a slackline that is fixed to trees at its opposing sides. At the point where she is standing there are three forces acting:
her weight $\boldsymbol{F}_d = M g$ pushing downwards, and
forces along the slackline towards the left $\boldsymbol{F}_l$ and right $\boldsymbol{F}_r$.
She can stay at rest as long as

$$\boldsymbol{0} = \boldsymbol{F}_d + \boldsymbol{F}_l + \boldsymbol{F}_r$$

The forces $\boldsymbol{F}_l$ and $\boldsymbol{F}_r$ are counterbalanced by the trees. These forces become huge when the slackline runs almost horizontally. Every now a then a careless slackliner roots out a tree or fells a pillar.

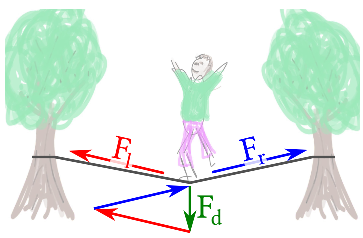


Figure 2.11: For a person balancing on a slackline, the gravitational force $\boldsymbol{F}_d$ (d for down) is balanced by forces $\boldsymbol{F}_l$ and $\boldsymbol{F}_r$ along the line that pull towards the left and right, respectively. See Example 2.19 for further discussion.

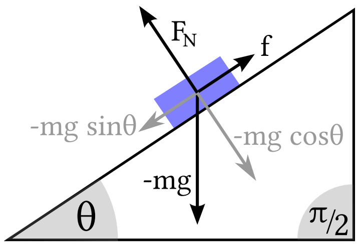
Example 2.20: Measuring the static friction coefficient

In a rough approximation static friction between two surfaces arises due to interlocking or surface irregularities. One must lift a block by a little amount to unlock the surfaces. In line with this argument dimensional analysis suggest that static friction should be proportional to the normal force between the surfaces. It is independent of the contact area, and depends on the material of the surfaces. This is indeed what is observed experimentally: The static friction force, f in Figure 2.12, can take values up to a maximum value of γ

times the normal force, $\boldsymbol{F}_N$, where γ typically takes values slightly less than one. By splitting the gravitational force, $m\boldsymbol{g}$ acting on a block on a plane into its components parallel and normal to the surface (gray arrows in Figure 2.12), one finds that in the presence of a force balance $m\boldsymbol{g} + \boldsymbol{f} + \boldsymbol{F}_N = \boldsymbol{0}$ one has

$$\left.\begin{aligned} F_N &= mg\cos\theta \\ f &= mg\sin\theta \\ f &< \gamma F_N \end{aligned}\right\} \quad \begin{aligned} &\Rightarrow \sin\theta < \gamma\cos\theta \\ &\quad \Rightarrow \theta < \theta_c = \arctan\gamma \end{aligned}$$

When θ exceeds θ_c the block starts to slide. Hence, one can infer γ from measurements of θ_c.



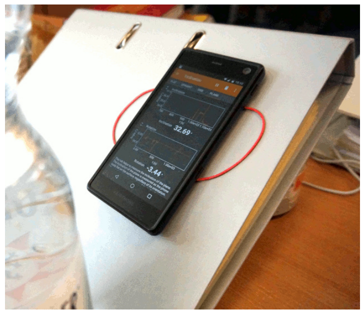
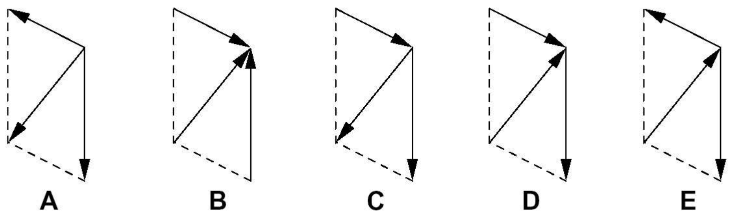


Figure 2.12: (top) As long as θ is smaller than the angle of friction the blue block does not slide. (bottom) Placing my cell phone on two rubber bands on a folder provides a maximum angle of about 33°, i. e. $\mu \simeq 0.5$. Using PhyPhox and a cell phone one can easily measure θ_c and μ for other combinations of materials.

2.6.1 *Self Test*

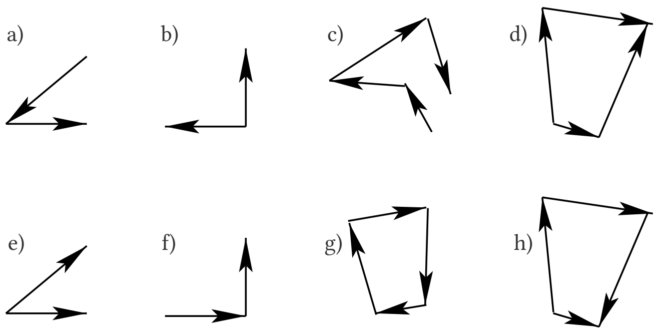
Problem 2.14. **Particles at rest**

There are three forces acting on the center of mass of a body. In which cases does it stay at rest?



Problem 2.15. **Graphical sum of vectors**

Determine the sum of the vectors. In which cases is the resulting vector vertical to the horizontal direction?



Problem 2.16. **Towing a stone**

Three Scottish muscleman[5] try to tow a stone with mass $M = 20\,\text{cwt}$ from a field. Each of them gets his own rope, and he can act a maximal force of 300 lbg as long as the ropes run in directions that differ by at least 30°.

a) Sketch the forces acting on the stone and their sum. By which ratio is the force exerted by three men larger than that of a single man?

[5] In highland games one still uses Imperial Units. A hundredweight (cwt) amounts to eight stones (stone) that each have a mass of 14 pounds (lb). A pound-force (lbg) amounts to the gravitational force acting on a pound. One can solve this problem without converting units.

b) The stone counteracts the pulling of the men by a static friction force $\mu M g$, where g is the gravitational acceleration. What is the maximum value that the friction coefficient μ may take when the men can move the stone?

2.7 *The inner product*

The position of a particle, the direction of its motion and the angle of attack of forces are constantly changing during the motion of a particle. In Chapter 3 we explore how they are related. The calculations are feasible because the involved vector spaces also have an inner product.

Definition 2.12: Inner Product of vector spaces over $\mathbb{R}$ or $\mathbb{C}$

The *inner product* on a vector space $(\mathsf{V}, \mathbb{R}, \oplus, \odot)$ defines a binary operation $\langle \_ \mid \_ \rangle : \mathsf{V} \times \mathsf{V} \to \mathbb{R}$ with the following properties for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathsf{V}$ and $c \in \mathbb{R}$

a) commutativity: $\langle \boldsymbol{v} \mid \boldsymbol{w} \rangle = \langle \boldsymbol{w} \mid \boldsymbol{v} \rangle$

b) linearity in the first argument: $\langle c\, \boldsymbol{v} \mid \boldsymbol{w} \rangle = c\, \langle \boldsymbol{v} \mid \boldsymbol{w} \rangle$
and $\langle \boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{w} \rangle = \langle \boldsymbol{u} \mid \boldsymbol{w} \rangle + \langle \boldsymbol{v} \mid \boldsymbol{w} \rangle$

c) positivity: $\langle \boldsymbol{v} \mid \boldsymbol{v} \rangle \geq 0$
where equality applies iff $\boldsymbol{v} = \boldsymbol{0}$, $\langle \boldsymbol{v} \mid \boldsymbol{v} \rangle = 0 \Leftrightarrow \boldsymbol{v} = \boldsymbol{0}$

For a vector space over $\mathbb{C}$ the requirement a) is replaced by

a) conjugate symmetry: $\langle \boldsymbol{v} \mid \boldsymbol{w} \rangle = \overline{\langle \boldsymbol{w} \mid \boldsymbol{v} \rangle}$

and the constant c is a complex number.

Remark 2.16. The idea underlying these properties is that $\sqrt{\langle \boldsymbol{v} \mid \boldsymbol{v} \rangle}$ can be interpreted as the length of the vector $\boldsymbol{v}$.

Remark 2.17. Conjugate symmetry and linearity for the first argument imply the following relations for the second argument

$$\langle \boldsymbol{v} \mid c\, \boldsymbol{w} \rangle = \overline{\langle c\, \boldsymbol{w} \mid \boldsymbol{v} \rangle} = \bar{c}\, \overline{\langle \boldsymbol{w} \mid \boldsymbol{v} \rangle} = \bar{c}\, \langle \boldsymbol{v} \mid \boldsymbol{w} \rangle$$
$$\langle \boldsymbol{u} \mid \boldsymbol{v} + \boldsymbol{w} \rangle = \overline{\langle \boldsymbol{v} + \boldsymbol{w} \mid \boldsymbol{u} \rangle} = \overline{\langle \boldsymbol{v} \mid \boldsymbol{u} \rangle} + \overline{\langle \boldsymbol{w} \mid \boldsymbol{u} \rangle} = \langle \boldsymbol{u} \mid \boldsymbol{v} \rangle + \langle \boldsymbol{u} \mid \boldsymbol{w} \rangle$$

Remark 2.18. Certain properties that hold for addition and scalar multiplication do *not* hold for the inner product.

a) There is no inverse: The information about the direction of vectors is lost upon taking the inner product. For instance, when $\langle \boldsymbol{u} \mid \boldsymbol{v} \rangle = 0$ and $\langle \boldsymbol{u} \mid \boldsymbol{w} \rangle = 0$ then one still can not tell the result of $\langle \boldsymbol{v} \mid \boldsymbol{w} \rangle$.

b) Associativity does not hold: $\langle \boldsymbol{u} \mid \boldsymbol{v} \rangle\, \boldsymbol{w} \neq \boldsymbol{u}\, \langle \boldsymbol{v} \mid \boldsymbol{w} \rangle$.

Example 2.21: Inner product for real-valued vectors

For real-valued vectors the inner product is commutative, $\langle \boldsymbol{v} \mid \boldsymbol{w} \rangle = \langle \boldsymbol{w} \mid \boldsymbol{v} \rangle$. The inner product is then also written as $\boldsymbol{v} \cdot \boldsymbol{w}$, and it obeys bilinearity

$$\boldsymbol{u} \cdot (a\boldsymbol{v} + b\boldsymbol{w}) = a(\boldsymbol{u} \cdot \boldsymbol{v}) + b(\boldsymbol{u} \cdot \boldsymbol{w})$$

Theorem 2.2: Geometric Interpretation of the Inner Product for Real-Valued Vectors

For vectors of $\mathbb{R}^D$ the inner product of two vectors $\boldsymbol{a}, \boldsymbol{b}$ takes the value

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}|\, |\boldsymbol{b}|\, \cos\theta$$

where $\theta = \angle(\boldsymbol{a}, \boldsymbol{b})$ is the angle between the two vectors, see Figure 2.13.

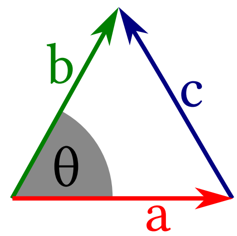


Figure 2.13: Notations for the geometric interpretation of the inner product, Theorem 2.2

Proof. The cosine theorem for triangles with sides of length a, b and c and angle θ opposite to c states that

$$c^2 = a^2 + b^2 - 2\,a\,b\,\cos\theta$$

Let now a, b, and c be the length of the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$, as shown in Figure 2.13. Then we have

$$a^2 + b^2 - 2\,a\,b\,\cos\theta = c^2 = \boldsymbol{c} \cdot \boldsymbol{c} = (\boldsymbol{a} - \boldsymbol{b}) \cdot (\boldsymbol{a} - \boldsymbol{b})$$
$$= \boldsymbol{a} \cdot \boldsymbol{a} - 2\,\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{b} = a^2 + b^2 - 2\,\boldsymbol{a} \cdot \boldsymbol{b}$$
$$\Rightarrow \quad \boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}|\, |\boldsymbol{b}|\, \cos\theta$$

□

Remark 2.19. Theorem 2.2 entails that the inner product $\boldsymbol{u} \cdot \boldsymbol{v}$ vanishes when the vectors are orthogonal, $\theta = \pi/2$. Also in general we say that

$$\boldsymbol{v} \text{ and } \boldsymbol{w} \text{ are othogonal iff} \qquad \langle \boldsymbol{v} \mid \boldsymbol{w} \rangle = 0\,.$$

Remark 2.20. The expression for the inner product provided in Theorem 2.2 does not imply that the inner product is unique. Rather it is a consequence of the cosine theorem that holds iff the geometric interpretation of the vectors applies. This is demonstrated by an example provided in Problem 2.17.

2.7.1 *Self Test*

Problem 2.17. **The inner product is not unique**

Let v_1 and v_2 be two non-orthogonal vectors in a two-dimensional vector space with an inner product $\langle \_ \mid \_ \rangle$, and let λ_1 and λ_2 two positive real numbers. Then the following relation defines another inner product $(\_ \mid \_)$:

$$(\boldsymbol{a} \mid \boldsymbol{b}) = \lambda_1 \langle \boldsymbol{a} \mid \boldsymbol{e}_1 \rangle\, \langle \boldsymbol{e}_1 \mid \boldsymbol{b} \rangle + \lambda_2 \langle \boldsymbol{a} \mid \boldsymbol{e}_2 \rangle\, \langle \boldsymbol{e}_2 \mid \boldsymbol{b} \rangle \tag{2.7.1}$$

a) Verify that the properties a) and b) of an inner product $\langle \_ \mid \_ \rangle$ as given in Definition 2.12 are also be obeyed by $(\_ \mid \_)$.

b) Verify that $(\boldsymbol{a} \mid \boldsymbol{a}) \geq 0$ iff λ_1 and λ_2 two positive real numbers.

c) Verify that $(\boldsymbol{a} \mid \boldsymbol{a}) = 0$ implies $\boldsymbol{a} = \boldsymbol{0}$ iff the vector space is two-dimensional.

Problem 2.18. **Inner products for polynomials**

Let $\boldsymbol{p} = \left[\sum_{i=0}^{D} p_i\, x^i\right]$ and $\boldsymbol{q} = \left[\sum_{i=0}^{D} q_i\, x^i\right]$ be elements of the vector space of N-dimensional polynomials. Verify that the following rules define inner products on this space.

a) $\langle \boldsymbol{p} \mid \boldsymbol{q} \rangle = \sum_{i=0}^{N} \bar{p}_i\, q_i$

b) $\langle \boldsymbol{p} \mid \boldsymbol{q} \rangle_{[a,b]} = \int_a^b \mathrm{d}x \overline{\left[\sum_{i=0}^{D} p_i\, x^i\right]} \left[\sum_{i=0}^{D} q_i\, x^i\right]$ for $a < b \in \mathbb{R}$

c) Show that $\boldsymbol{p} = [1]$ and $\boldsymbol{q} = [x]$ are orthogonal with respect to the inner product defined in a). Under which condition are they also orthogonal for the inner product defined in b)?

2.8 *Cartesian coordinates*

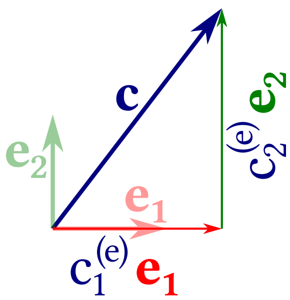


Figure 2.14: Representation of the vector $\boldsymbol{c}$ in terms of the orthogonal unit vectors $(\boldsymbol{e}_1, \boldsymbol{e}_2)$.

Theorem 2.2 entails an extremely elegant possibility to deal with vectors. We first illustrate the idea based on a two-dimensional example, Figure 2.14, and then we develop the general theory:

Let $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ be two orthogonal vectors that have unit length,

$$\langle \boldsymbol{e}_1 | \boldsymbol{e}_1 \rangle = \langle \boldsymbol{e}_2 \mid \boldsymbol{e}_2 \rangle = 1 \quad \text{and} \quad \langle \boldsymbol{e}_1 | \boldsymbol{e}_2 \rangle = 0$$

For every vector $\boldsymbol{c}$ in the plane described by these two vectors, we can then find two numbers $c_1^{(e)}$ and $c_2^{(e)}$ such that

$$\boldsymbol{c} = c_1^{(e)}\, \boldsymbol{e}_1 + c_2^{(e)}\, \boldsymbol{e}_2$$

Now the choice of the vectors $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ entails that triangle with edge $\boldsymbol{c}$, $c_1^{(e)}\, \boldsymbol{e}_1$, and $c_2^{(e)}\, \boldsymbol{e}_2$ is right-angled and that

$$\begin{aligned} c_i^{(e)} &= |\boldsymbol{c}|\ \cos\angle(\boldsymbol{c}, \boldsymbol{e}_i) = \langle \boldsymbol{c} \mid \boldsymbol{e}_i \rangle \quad \text{for } i \in \{1,2\} \\ \Rightarrow \quad \boldsymbol{c} &= \langle \boldsymbol{c} \mid \boldsymbol{e}_1 \rangle\, \boldsymbol{e}_1 + \langle \boldsymbol{c} \mid \boldsymbol{e}_2 \rangle\, \boldsymbol{e}_2 \end{aligned}$$

This strategy to represent vectors applies in all dimensions.

Definition 2.13: Basis and Coordinates

Let $\mathsf{B} = \{\boldsymbol{e}_i\},\ i \in \{1, \ldots, D\}$ be a set of D pairwise orthogonal unit vectors

$$\forall i,j \in \{1,\dots,D\}: \quad \boldsymbol{e}_i \cdot \boldsymbol{e}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

in a vector space $(\mathsf{V}, \mathbb{F}, +, \cdot)$ with inner product $\langle \_ | \_ \rangle$. We say that B forms a *basis* for a D-dimensional vector space iff

$$\forall \boldsymbol{v} \in \mathsf{V}\ \exists v_i,\ i \in \{1,\dots,D\}: \quad \boldsymbol{v} = \sum_{i=1}^{D} v_i^{(e)}\ \boldsymbol{e}_i$$

In that case we also have $v_i^{(e)} = \langle \boldsymbol{v} \mid \boldsymbol{e}_i \rangle,\ i \in \{1,\dots,D\}$ and these numbers are called the *coordinates* of the vector $\boldsymbol{v}$. The number of vectors D in the basis of the vector space is denoted as *dimension of the vector space*.

Remark 2.21. The choice of a basis, and hence also of the coordinates, is not unique. Figure 2.15 shows the representation of a vector in terms of two different bases $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ and $(\boldsymbol{n}_1, \boldsymbol{n}_2)$. We suppress the superscript that indicates the basis when the choice of the basis is clear from the context.

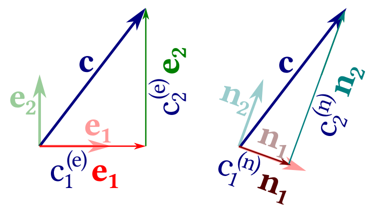


Figure 2.15: Representation of the vector $\boldsymbol{c}$ of Figure 2.14 in terms of the bases $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ and $(\boldsymbol{n}_1, \boldsymbol{n}_2)$.

Remark 2.22. For a given basis the representation in terms of coordinates is unique.

Proof. 1. The coordinates a_i of a vector $\boldsymbol{a}$ are explicitly given by $a_i = \langle a \mid \boldsymbol{e}_i \rangle$. This provides unique numbers for a given basis set.

2. Assume now that two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ have the same coordinate representation. Then the vector-space properties imply

$$\left.\begin{array}{l} \boldsymbol{a} = \sum_i c_i\, \boldsymbol{e}_i \\ \boldsymbol{b} = \sum_i c_i\, \boldsymbol{e}_i \end{array}\right\} \Rightarrow \boldsymbol{a} - \boldsymbol{b} = \left(\sum_i c_i\, \boldsymbol{e}_i\right) - \left(\sum_i c_i\, \boldsymbol{e}_i\right) = \sum_i (c_i - c_i)\, \boldsymbol{e}_i = \sum_i 0\, \boldsymbol{e}_i = \boldsymbol{0}$$

$$\Rightarrow \quad \boldsymbol{a} = \boldsymbol{b}$$

Hence, they must be identical. □

Remark 2.23 (Kronecker δ_{ij}). It is convenient to introduce the abbreviation δ_{ij} for

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

where i, j are elements of some index set. This symbol is denoted as *Kronecker* δ. With the Kronecker symbol the condition on orthogonal unit vectors of a basis can more concisely be written as

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$$

Moreover, for $i, j \in \{1,\dots,D\}$ the numbers, δ_{ij}, describe a $D \times D$ matrix which is the neutral element for multiplication with another $D \times D$ matrix, and also with a vector of $\mathbb{R}^D$, when it is interpreted as a $D \times 1$ matrix.

Theorem 2.3: Scalar product on $\mathbb{R}^D$ and $\mathbb{C}^D$

The axioms of vector spaces and the inner product imply that

$$\text{on } \mathbb{R}^D : \qquad \langle \boldsymbol{a} \mid \boldsymbol{b} \rangle = \sum_{i=1}^{D} \langle \boldsymbol{a} \mid i \rangle \, \langle i \mid \boldsymbol{b} \rangle = \sum_{i=1}^{D} a_i b_i$$

$$\text{on } \mathbb{C}^D : \qquad \langle \boldsymbol{a} \mid \boldsymbol{b} \rangle = \sum_{i=1}^{D} \langle \boldsymbol{a} \mid i \rangle \, \langle i \mid \boldsymbol{b} \rangle = \sum_{i=1}^{D} a_i \bar{b}_i$$

where the bar indicates complex conjugation of complex numbers. This can be written as follows when representing the coordinates as a 1D array of numbers

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_D \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_D \end{pmatrix} = a_1 \, \bar{b}_1 + a_2 \, \bar{b}_2 + \cdots + a_D \, \bar{b}_D$$

where the complex conjugation does not apply for real numbers. This latter form of the inner product is denoted as *scalar product*.

Proof. We first note that the case of real numbers can be interpreted as special case of the complex numbers with a vanishing complex part. Hence, we only provide the proof for the complex case.

We use the representations $\boldsymbol{a} = \sum_i \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \boldsymbol{e}_i$ and $\boldsymbol{b} = \sum_j \langle \boldsymbol{b} \mid \boldsymbol{e}_j \rangle \, \boldsymbol{e}_j$, and work step by step from the left to the aspired result:

$$\begin{aligned} \langle \boldsymbol{a} \mid \boldsymbol{b} \rangle &= \left\langle \sum_i \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \boldsymbol{e}_i \, \middle| \, \sum_j \langle \boldsymbol{b} \mid \boldsymbol{e}_j \rangle \, \boldsymbol{e}_j \right\rangle \\ &= \sum_i \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \left\langle \boldsymbol{e}_i \, \middle| \, \sum_j \langle \boldsymbol{b} \mid \boldsymbol{e}_j \rangle \, \boldsymbol{e}_j \right\rangle \\ &= \sum_i \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \sum_j \overline{\langle \boldsymbol{b} \mid \boldsymbol{e}_j \rangle} \, \langle \boldsymbol{e}_i \mid \boldsymbol{e}_j \rangle \\ &= \sum_i \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \sum_j \langle \boldsymbol{e}_j \mid \boldsymbol{b} \rangle \, \delta_{ij} \\ &= \sum_i \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \langle \boldsymbol{e}_i \mid \boldsymbol{b} \rangle \, . \end{aligned}$$

Due to $a_i = \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle$ and $\bar{b}_i = \langle \boldsymbol{e}_i \mid \boldsymbol{b} \rangle$ we therefore have

$$\langle \boldsymbol{a} \mid \boldsymbol{b} \rangle = \sum_i a_i \, \bar{b}_i$$

□

Remark 2.24. Einstein pointed out that the sums over pairs of identical indices arise ubiquitously in calculations like to proof of Theorem 2.3. He therefore adopted the convention that one always sums over pairs of identical indices, and does no longer explicitly write that down. This leads to substantially clearer representation of the calculation. For instance, the proof looks then as follows:

$$
\begin{aligned}
& \langle \boldsymbol{a} \mid \boldsymbol{b} \rangle = \Big\langle \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \boldsymbol{e}_i \,\Big|\, \langle \boldsymbol{b} \mid \boldsymbol{e}_j \rangle \, \boldsymbol{e}_j \Big\rangle = \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \Big\langle \boldsymbol{e}_i \,\Big|\, \langle \boldsymbol{b} \mid \boldsymbol{e}_j \rangle \, \boldsymbol{e}_j \Big\rangle \\
& \qquad = \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \overline{\langle \boldsymbol{b} \mid \boldsymbol{e}_j \rangle} \, \langle \boldsymbol{e}_i \mid \boldsymbol{e}_j \rangle = \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \overline{\langle \boldsymbol{b} \mid \boldsymbol{e}_j \rangle} \, \delta_{ij} \\
& \qquad = \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \, \langle \boldsymbol{e}_i \mid \boldsymbol{b} \rangle \\
\Rightarrow \quad & \langle \boldsymbol{a} \mid \boldsymbol{b} \rangle = a_i \, \bar{b}_i
\end{aligned}
$$

Remark 2.25. Dirac pointed out that the vector product $\langle \boldsymbol{a} \mid \boldsymbol{b} \rangle$ takes the form of the multiplication of a $1 \times D$ matrix for $\boldsymbol{a}$ and a $D \times 1$ matrix for $\boldsymbol{b}$. He suggested to symbolically write down these vectors as a *bra vector* $\langle a|$ and a *ket vector* $|b\rangle$. When put together as a bra-(c)-ket $\langle a|b\rangle$ one recovers the inner product, and introducing $|\boldsymbol{e}_i\rangle\langle e_i|$ and observing Einstein notation comes down to inserting a unit matrix. For instance for 2×2 vectors

$$
\langle \boldsymbol{a} \mid \boldsymbol{b} \rangle = (a_1, a_2) \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = (a_1, a_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \langle \boldsymbol{e}_i \mid \boldsymbol{b} \rangle
$$

Conceptually this is a very useful observation because it provides an easy rule to sort out what changes in the equations when one represents a problem in terms of a different basis.

Example 2.22: Changing coordinates from basis (e_i) to basis $(\boldsymbol{n}_i)$

We observe Dirac's observation that the expressions $|e_i\rangle\langle e_i|$ and $|n_i\rangle\langle n_i|$ sandwiched between a bra and a ket amounts to multiplication with one. Hence, the coordinates change according to

$$
a_i^{(n)} = \langle \boldsymbol{a} \mid \boldsymbol{n}_i \rangle = \langle \boldsymbol{a} \mid \boldsymbol{e}_j \rangle \langle \boldsymbol{e}_j | \boldsymbol{n}_i \rangle = a_j^{(e)} \langle \boldsymbol{e}_j | \boldsymbol{n}_i \rangle
$$

which amounts to multiplying the vector with entries $(a_j^{(e)}, j = 1, \dots, D)$ with the $D \times D$ matrix T with entries $t_{ji} = \langle \boldsymbol{e}_j | \boldsymbol{n}_i \rangle$.
On the other hand, for the inner products we have

$$
\begin{aligned}
a_i^{(e)} \bar{b}_i^{(e)} &= \langle \boldsymbol{a} \mid \boldsymbol{b} \rangle = \langle \boldsymbol{a} \mid \boldsymbol{e}_i \rangle \langle \boldsymbol{e}_i \mid \boldsymbol{b} \rangle \\
&= \langle \boldsymbol{a} \mid \boldsymbol{n}_j \rangle \langle \boldsymbol{n}_j \mid \boldsymbol{e}_i \rangle \langle \boldsymbol{e}_i \mid \boldsymbol{n}_k \rangle \langle \boldsymbol{n}_k | \boldsymbol{b} \rangle = \langle \boldsymbol{a} \mid \boldsymbol{n}_j \rangle \langle \boldsymbol{n}_j \mid \boldsymbol{n}_k \rangle \langle \boldsymbol{n}_k | \boldsymbol{b} \rangle \\
&= \langle \boldsymbol{a} \mid \boldsymbol{n}_j \rangle \, \delta_{jk} \, \langle \boldsymbol{n}_k | \boldsymbol{b} \rangle = \langle \boldsymbol{a} \mid \boldsymbol{n}_j \rangle \langle \boldsymbol{n}_j | \boldsymbol{b} \rangle = a_i^{(n)} \bar{b}_i^{(n)}
\end{aligned}
$$

Its value does not change, even though the coordinates take entirely different values.

add worked example for an explicit coordinate transformation

2.8.1 Self Test

Problem 2.19. **Cartesian coordinates in the plane**

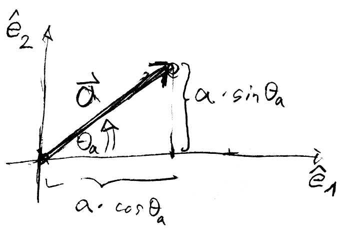
a) Mark the following points in a Cartesian coordinate system:

$$(0, 0) \quad (0, 3) \quad (2, 5) \quad (4, 3) \quad (4, 0)$$

Add the points $(0, 0)$ $(4, 3)$ $(0, 3)$ $(4, 0)$, and connect the points in the given order. What do you see?

b) What do you find when drawing a line segment connecting the following points?

$$(0, 0) \quad (1, 4) \quad (2, 0) \quad (-1, 3) \quad (3, 3) \quad (0, 0)$$



Problem 2.20. **Geometric and algebraic form of the scalar product** The sketch in the margin shows a vector $\boldsymbol{a}$ in the plane, and its representation as a linear combination of two orthonormal vectors $(\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2)$,

$$\boldsymbol{a} = a \cos\theta_a \, \hat{\boldsymbol{e}}_1 + a \sin\theta_a \, \hat{\boldsymbol{e}}_2$$

Here, a is the length of the vector $\boldsymbol{a}$, and $\theta_1 = \angle(\hat{\boldsymbol{e}}_1, \boldsymbol{a})$.

a) Analogously to $\boldsymbol{a}$ we consider another vector $\boldsymbol{b}$ with a representation

$$\boldsymbol{b} = b \cos\theta_b \, \hat{\boldsymbol{e}}_1 + b \sin\theta_b \, \hat{\boldsymbol{e}}_2$$

Employ the rules of scalar products, vector addition and multiplication with scalars to show that

$$\boldsymbol{a} \cdot \boldsymbol{b} = a \, b \, \cos(\theta_a - \theta_b)$$

Hint: Work backwards, expressing $\cos(\theta_a - \theta_b)$ in terms of $\cos\theta_a$, $\cos\theta_b$, $\sin\theta_a$, and $\sin\theta_b$.

b) As a shortcut to the explicit calculation of a) one can introduce the coordinates $a_1 = a \cos\theta_a$ and $a_2 = a \sin\theta_a$, and write $\boldsymbol{a}$ as a tuple of two numbers. Proceeding analogously for $\boldsymbol{b}$ one obtains

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

How does the product $\boldsymbol{a} \cdot \boldsymbol{b}$ look like in terms of these coordinates?

c) How do the arguments in a) and b) change for D dimensional vectors that are represented as linear combinations of a set of orthonormal basis vectors $\hat{\boldsymbol{e}}_1, \dots, \hat{\boldsymbol{e}}_D$?

What changes when the basis is not orthonormal?
What if it is not even orthogonal?

Problem 2.21. **Scalar product on $\mathbb{R}^D$**

Show that the scalar product on $\mathbb{R}^D$ takes exactly the same form as for the complex case, Theorem 2.3.

However, complex conjugation is not necessary in that case.

Problem 2.22. **Pauli matrices form a basis for a 4D vector space**

Show that the Pauli matrices

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

form a basis of the real vector space of 2×2 Hermitian matrices, $\mathbb{H}$, with

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{H} \quad \Leftrightarrow \quad a_{ij} \in \mathbb{C} \ \wedge \ a_{ij} = a_{ji}^*$$

Show to that end

a) The matrices $\sigma_0, \dots, \sigma_4$ are linearly independent.

b) $x_0, \dots, x_4 \in \mathbb{R} \quad \Rightarrow \quad \sum_{i=0}^{4} x_i\,\sigma_i \in \mathbb{H}$

c) $M \in \mathbb{H} \quad \Rightarrow \quad \exists x_0, \dots, x_4 \in \mathbb{R} : M = \sum_{i=0}^{4} x_i\,\sigma_i$

What about linear combinations with coefficients $z_1, \dots, z_4$? Is $\sum_{i=0}^{4} z_i\,\sigma_i$ Hermitian? Do these matrices form a vector space?

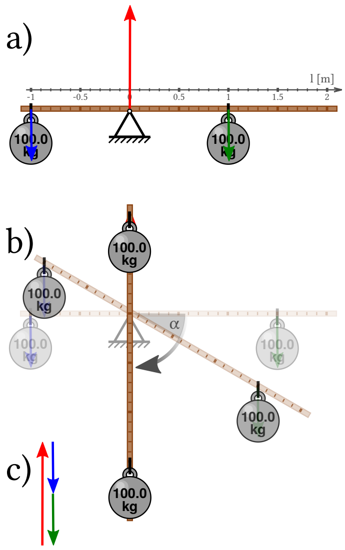
2.9 *Cross products — torques*

adapted from rachaelvoorhees from arlington, va / wikimedia CC BY 2.0

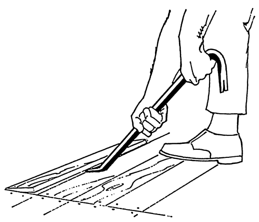
more explanation needed.

The picture in the margin shows the sign of a seesaw, a playground toy that works even for people with vastly different weight and size. Figure 2.16a) shows a balanced scale. When the forces acting on the scale do not add up to zero, we pick up the scale. It moves. The according force balance for the beam of the scale is shown in Figure 2.16c). In general the beam does not stay at rest, when the two masses are not attached at the same distance from the fulcrum. The force balance, Figure 2.16c), still hods, and the beam turns, rather than being lifted. The sum of attached forces tells us if an object is displaced. In analogy we introduce the *torque* to describe whether it turns.

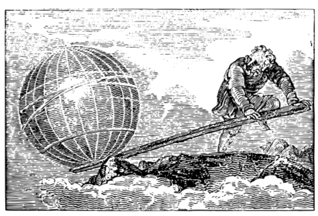
When the beam is vertical there is no torque, and it takes its maximum when the beam is horizontal. In the former case the forces act parallel to the beam, and in the latter they act in orthogonal direction. Moreover, a weight that is attached at a larger distance to the fulcrum induces a larger torque, and the torque also increases with mass. This is expressed in the lever rule.



based on from Jahobr/wikimedia CC0 1.0
Figure 2.16: a) The lever is balanced when two equal masses are attached at the same distance from the fulcrum. b) It is at (stable) rest only in a single position when equal weighs are attached at different distances. c) In all positions the sum of the forces on the beam, by the fulcrum and by the two weights, add to zero.



Pearson Scott Foresman / Public domain
Figure 2.17: Action of a crowbar.



Mechanic's Magazine cover of Vol II, Knight & Lacey, London, 1824./wikipedia, public domain
Figure 2.18: Illustration of Archimedes' remark about moving the earth.

Example 2.23: Torques on a Lever

The torque T exerted by a lever is given by the product, $T = l\,F$, of the modulus of the force F acting vertical to the lever and the distance l between the fulcrum and the point where the force is applied, which is called *length of the lever arm*.
When several forces act on the same lever, then the total torque amounts to the sum of the torques induced by the individual forces, $T = \sum_i l_i\,F_i$. For the scale in Figure 2.16a) and b) we find

$$T_a = (1\,\text{m})\;(100\,\text{kg})\;(-g) + (-1\,\text{m})\;(100\,\text{kg})\;(-g) = 0$$
$$T_b = (1.5\,\text{m})\;(100\,\text{kg})\;(-g)\;\cos\alpha + (-1\,\text{m})\;(100\,\text{kg})\;(-g)\;\cos\alpha$$
$$\simeq -500\;\cos\alpha\;\text{kg}\,\text{m}^2/\text{s}^2$$

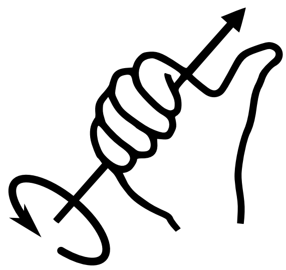
The torque vanishes only when $\alpha = \pi/2$ as shown in the figure, and for the unstable tipping point $\alpha = -\pi/2$.

Remark 2.26. Adopting a lever where force is applied on a long arm allows one to move very heavy objects or break very stable objects. Common technological applications are the crowbar and the lever. Archimedes was so impressed by this principle that he is quoted to have remarked "Δοσ μοι που στω και κινω την γην" (Archimedes, 1878), i.e. "Give me but one firm spot on which to stand, and I will move the earth" (Oxford Dictionary of Quotations, 1953)

estimate amplification of force the the crowbar

Observe the sign of the torque: In Example 2.23 it is positive for counterclockwise motion, and negative for clockwise motion. The axis of rotation is fixed by the fulcrum. However, when acting the crowbar, one applies a horizontal force to get the crowbar under the obstacle. This induces a rotation around a vertical axis. Subsequently, a vertical force is applied to lift the obstacle. It induces a rotation around a horizontal axes. The relation between the directions of the lever arm, the force, and the rotation axis is commonly illustrated by the right-hand rule (Figure 2.19): Here the arm points in the direction of the lever arm, the fingers in the direction of the applied force, and the thump along the rotation axis. This suggests to define torque as a product of two vectors, the arm ℓ and the force $\boldsymbol{F}$ that provide the torque, $\boldsymbol{T}$, which is a vector of length $|\ell|\;|\boldsymbol{F}|\;\sin\angle\ell,\boldsymbol{F}$ in a direction normal to the plane defined by ℓ and $\boldsymbol{F}$. This operation, $\boldsymbol{T} = \ell \times \boldsymbol{F}$ defines the cross product. We explore its properties in a mathematical digression.

2.9.1 *Algebraic properties of cross products*



Schorschi2 at de.wikipediaderivative work: Wizard191, public domain
Figure 2.19: Right-hand rule.

Definition 2.14: Cross product on $\mathbb{R}^3$

The *cross product* on the vector space $\mathbb{R}^3$ defines a binary operation $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ with the following properties for $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ and $c \in \mathbb{R}$

a) anti-commutativity: $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$

b) distributivity: $\boldsymbol{u} \times (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \times \boldsymbol{v} + \boldsymbol{u} \times \boldsymbol{w}$

c) compatibility with scalar multiplication:
$$(c\,\boldsymbol{u}) \times \boldsymbol{v} = \boldsymbol{u} \times (c\,\boldsymbol{v}) = c\,(\boldsymbol{u} \times \boldsymbol{v})$$

d) symmetry of scalar triple product (Jacobi identity):
$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \boldsymbol{v} \cdot (\boldsymbol{w} \times \boldsymbol{u}) = \boldsymbol{w} \cdot (\boldsymbol{u} \times \boldsymbol{v})$$

Moreover for every right-handed set of three orthonormal vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$ we require

e) normalization: $\boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \times \boldsymbol{e}_3) = 1$

Remark 2.27. The cross product of a vector with itself vanishes
$$\forall \boldsymbol{v} \in \mathbb{R}^3 \,:\, \boldsymbol{v} \times \boldsymbol{v} = \boldsymbol{0}$$

Proof. Vanishing of $\boldsymbol{v} \times \boldsymbol{v}$ is a consequence of anti-commutativity:
$$\boldsymbol{v} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{v} \quad \Rightarrow 2\,\boldsymbol{v} \times \boldsymbol{v} = \boldsymbol{0} \quad \Rightarrow \boldsymbol{v} \times \boldsymbol{v} = \boldsymbol{0}$$
□

Theorem 2.4: Right-handed orthonormal basis in $\mathbb{R}^3$

Let $\boldsymbol{e}_1, \boldsymbol{e}_2 \in \mathbb{R}^3$ be orthonormal vectors, $\boldsymbol{e}_1 \cdot \boldsymbol{e}_2 = \delta_{12}$. Then $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2$ form a right-handed orthonormal basis for $\mathbb{R}^3$, and we have
$$\boldsymbol{e}_i \cdot (\boldsymbol{e}_j \times \boldsymbol{e}_k) = \begin{cases} 1 & \text{for} \quad ijk \in \{123, 231, 312\} \\ -1 & \text{for} \quad ijk \in \{132, 213, 321\} \\ 0 & \text{else} \end{cases}$$

Remark 2.28 (Levi-Civita tensor ε_{ijk}). It is convenient to introduce the abbreviation ε_{ijk} for
$$\varepsilon_{ijk} = \begin{cases} 1 & \text{for} \quad ijk \in \{123, 231, 312\} \\ -1 & \text{for} \quad ijk \in \{132, 213, 321\} \\ 0 & \text{else} \end{cases}$$
This symbol is denoted as *Levi-Civita tensor* ε_{ijk}. With this symbol the relations between right-handed orthogonal unit vectors of a basis can more concisely be written as
$$\boldsymbol{e}_i \cdot (\boldsymbol{e}_j \times \boldsymbol{e}_k) = \varepsilon_{ijk}$$

Moreover, it immediately provides the following representation of the scalar triple product $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$ in terms of coordinates u_i, v_j, w_k, $i, j, k \in \{1, 2, 3\}$,

$$\left.\begin{aligned} \boldsymbol{u} &= \textstyle\sum_{i=1}^{3} u_i\, \boldsymbol{e}_i \\ \boldsymbol{v} &= \textstyle\sum_{j=1}^{3} v_j\, \boldsymbol{e}_j \\ \boldsymbol{w} &= \textstyle\sum_{k=1}^{3} w_k\, \boldsymbol{e}_k \end{aligned}\right\} \quad \Rightarrow \quad \boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \sum_{i,j,k=1}^{3} \varepsilon_{ijk}\, u_i\, v_j\, w_k$$

or even $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \varepsilon_{ijk}\, u_i\, v_j\, w_k$ with Einstein notation. The symmetry of the triple scalar product is an immediate consequence of the symmetry of the ε-tensor.

Proof. The identity $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \varepsilon_{ijk}\, u_i\, v_j\, w_k$ follows from the compatibility with scalar product and the relation for the basis vectors $\boldsymbol{e}_i \cdot (\boldsymbol{e}_j \times \boldsymbol{e}_k)$. The details of the proof are given as Problem 2.23. □

Proof of Theorem 2.4. We show that $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2$ form three orthonormal vectors. By assumption $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are orthonormal. Hence, we show that $\boldsymbol{e}_3$ is a unit vector that is orthogonal to $\boldsymbol{e}_1$ and $\boldsymbol{e}_3$:

$$\begin{aligned} \boldsymbol{e}_1 \cdot \boldsymbol{e}_3 &= \boldsymbol{e}_1 \cdot (\boldsymbol{e}_1 \times \boldsymbol{e}_2) = \boldsymbol{e}_2 \cdot (\boldsymbol{e}_1 \times \boldsymbol{e}_1) = \boldsymbol{e}_2 \cdot \boldsymbol{0} = 0 \\ \boldsymbol{e}_2 \cdot \boldsymbol{e}_3 &= \boldsymbol{e}_2 \cdot (\boldsymbol{e}_1 \times \boldsymbol{e}_2) = \boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \times \boldsymbol{e}_2) = \boldsymbol{e}_1 \cdot \boldsymbol{0} = 0 \\ \boldsymbol{e}_3 \cdot \boldsymbol{e}_3 &= \boldsymbol{e}_3 \cdot (\boldsymbol{e}_1 \times \boldsymbol{e}_2) = \boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \times \boldsymbol{e}_3) = 1 \end{aligned}$$

□

Remark 2.29 (bac-cab rule). The double cross product can be expressed in terms of scalar products. Commonly this relation is stated in terms of three vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c} \in \mathbb{R}^3$,

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}\,(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{c}\,(\boldsymbol{a} \cdot \boldsymbol{b})$$

and referred to as bac-cab rule.

Proof. We express the three vectors in terms of their coordinates with respect to the orthonormal basis $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$,

$$\boldsymbol{a} = \sum_{i=1}^{3} a_i\, \boldsymbol{e}_i \quad \boldsymbol{b} = \sum_{j=1}^{3} b_j\, \boldsymbol{e}_j \quad \boldsymbol{c} = \sum_{k=1}^{3} c_k\, \boldsymbol{e}_k \quad \text{with} \quad a_i, b_j, c_k \in \mathbb{R}$$

and use the rules defining the cross products and inner products

$$\begin{aligned} \boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) &= \left(\sum_{i=1}^{3} a_i\, \boldsymbol{e}_i\right) \times \left[\left(\sum_{j=1}^{3} b_j\, \boldsymbol{e}_j\right) \times \left(\sum_{k=1}^{3} c_k\, \boldsymbol{e}_k\right)\right] \\ &= \sum_{i,j,k=1}^{3} a_i\, b_j\, c_k \boldsymbol{e}_i \times (\boldsymbol{e}_j \times \boldsymbol{e}_k) \end{aligned}$$

When $j = k$ or when j and k are both different from i then the summand vanishes due to Remark 2.27. For $i = j \neq k$ one has $\boldsymbol{e}_i \times (\boldsymbol{e}_j \times \boldsymbol{e}_k) = -\boldsymbol{e}_k$, and for $i = k \neq j$ one has $\boldsymbol{e}_i \times (\boldsymbol{e}_j \times \boldsymbol{e}_k) = \boldsymbol{e}_j$. Consequently,

$$\begin{aligned} \boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) &= \sum_{i,k=1}^{3} a_i\, b_i\, c_k (-\boldsymbol{e}_k) + \sum_{i,j=1}^{3} a_i\, b_j\, c_i (\boldsymbol{e}_j) \\ &= \boldsymbol{b}\,(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{c}\,(\boldsymbol{a} \cdot \boldsymbol{b}) \end{aligned}$$

□

Remark 2.30 (Jacobi identity). The cross product obeys the Jacobi identity:

$$\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) + \boldsymbol{v} \times (\boldsymbol{w} \times \boldsymbol{u}) + \boldsymbol{w} \times (\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{0}$$

Proof. This can be verified by evaluating the triple cross products by the bac-cab rule. Details are give as Problem 2.24. □

Remark 2.31. In coordinate notation the cross product takes the form

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

Proof. For component k of $\boldsymbol{a} \times \boldsymbol{b}$ we have

$$[\boldsymbol{a} \times \boldsymbol{b}]_k = \hat{\boldsymbol{e}}_k \cdot (\boldsymbol{a} \times \boldsymbol{b}) = \hat{\boldsymbol{e}}_k \cdot \left[\left(\sum_{i=1}^{3} a_i\, \hat{\boldsymbol{e}}_i \right) \times \left(\sum_{j=1}^{3} b_j\, \hat{\boldsymbol{e}}_j \right) \right]$$

$$= \sum_{i,j=1}^{3} a_i\, b_j \;\; \hat{\boldsymbol{e}}_k \cdot \left[\hat{\boldsymbol{e}}_i \times \hat{\boldsymbol{e}}_j \right]_k = \sum_{i,j=1}^{3} a_i\, b_j\, \varepsilon_{ijk}$$

In the remark this is explicitly written out for $k \in \{1, 2, 3\}$. □

2.9.2 *Geometric interpretation of cross products*

The cross product and the scalar triple product have distinct geometrical interpretations. The geometric meaning of the cross product $\boldsymbol{a} \times \boldsymbol{b}$ can best be seen by adopting a basis where the first basis vector is parallel to $\boldsymbol{e}_1 = \boldsymbol{a}/|\boldsymbol{a}|$, and the second basis vector $\boldsymbol{e}_2$ lies orthogonal to $\boldsymbol{e}_1$ in the plane spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$. The third basis vector will then be $\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2$. The angle between $\boldsymbol{a}$ and $\boldsymbol{b}$, and hence also of $\boldsymbol{e}_1$ and $\boldsymbol{b}$ is denoted as θ. Thus, $\boldsymbol{b}$ can be written as $\boldsymbol{b} = b_1\, \boldsymbol{e}_1 + b_2\, \boldsymbol{e}_2 = |\boldsymbol{b}|\, (\cos\theta\, \boldsymbol{e}_1 + \sin\theta\, \boldsymbol{e}_2)$, cf. Figure 2.20). For this choice of the basis we find

$$\begin{aligned} \boldsymbol{a} \times \boldsymbol{b} &= |\boldsymbol{a}|\, \boldsymbol{e}_1 \times (b_1\, \boldsymbol{e}_1 + b_2\, \boldsymbol{e}_2) = |\boldsymbol{a}|\, b_1\, \boldsymbol{e}_1 \times \boldsymbol{e}_1 + |\boldsymbol{a}|\, b_2\, \boldsymbol{e}_1 \times \boldsymbol{e}_2 \\ &= |\boldsymbol{a}|\, |\boldsymbol{b}|\, \sin\theta\, \boldsymbol{e}_3 \end{aligned}$$

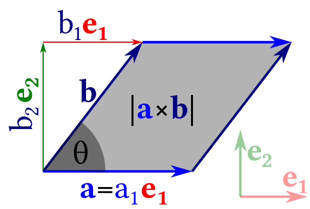


Figure 2.20: Geometric interpretation of the absolute value of the cross product.

Figure 2.20 illustrates that $|\boldsymbol{a}|\, |\boldsymbol{b}|\, \sin\theta$ amounts to the area of the parallelogram spanned by the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. Hence, the cross product amounts to a vector that is aligned vertically on the parallelogram, with a length that amounts to the area of the parallelogram.

In order to evaluate also the product $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c}$ we introduce the coordinate representation of $\boldsymbol{c}$ as $\boldsymbol{c} = c_1\, \boldsymbol{e}_1 + c_2\, \boldsymbol{e}_2 + c_3\, \boldsymbol{e}_3$ (Figure 2.21), and observe

$$\begin{aligned} (\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} &= |\boldsymbol{a} \times \boldsymbol{b}|\, \boldsymbol{e}_3 \cdot (c_1\, \boldsymbol{e}_1 + c_2\, \boldsymbol{e}_2 + c_3\, \boldsymbol{e}_3) \\ &= |\boldsymbol{a} \times \boldsymbol{b}|\, c_3 = a_1\, b_2\, c_3 \end{aligned}$$

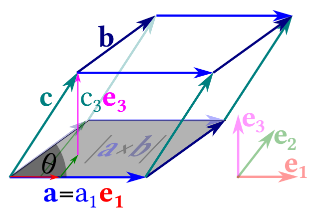


Figure 2.21: Geometric interpretation of the scalar triple product.

This amounts to the product of the area of the parallelogram spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$ multiplied by the height of the parallelepiped spanned by the vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$. Due to the special choice of the basis this volume amount to $a_1\, b_2\, c_3$ because all other contributions to the general expression $\sum_{ijk} a_i\, b_j\, c_k\, \varepsilon_{ijk}$ vanish. The symmetry of the scalar triple product, property d) in Definition 2.14, is understood from this perspective as the statement that the volume of the parallelepiped is invariant under (cyclic) renaming of the vectors that define its edges.

As a final remark, we emphasize that the geometric interpretation that we have given to the cross product holds in general — in spite of the special basis adopted in the derivation. It is a distinguishing feature of vector spaces that the scalar numbers that are derived from vectors take the same values every choice of the basis. It is up to the physicist to find the basis that admits the easiest calculations.

2.9.3 *The Torque*

The cross product equips us with the mathematical notions to define the torque on a body.

Definition 2.15: Torque

The *torque* $\boldsymbol{T}$ defines a force that is going to rotate a body around a position $\boldsymbol{q}_0$. Let $\boldsymbol{F}_i$ be the forces that attach the body at the positions $\boldsymbol{q}_i$ with respect to the considered origin. Then the torque is defined as

$$\boldsymbol{T} = \sum_i (\boldsymbol{q}_i - \boldsymbol{q}_0) \times \boldsymbol{F}_i$$

Remark 2.32. The value of the torque depends on the choice of the reference position $\boldsymbol{q}_0$.

Remark 2.33. In general, the torques induced by different forces point in different directions. They are added as vectors. We will further discuss this below in Example 2.24.

Axiom 2.2: Torque balance

Let N forces $\boldsymbol{F}_1, \dots, \boldsymbol{F}_N$ attack a body at the (body-fixed) positions $\boldsymbol{q}_i$. The body does not rotate around the position $\boldsymbol{q}_0$ as long as the sum of the torques induced by the forces add to zero, i.e. iff $\boldsymbol{0} = \boldsymbol{T} = \sum_{i=1}^{N} (\boldsymbol{q}_i - \boldsymbol{q}_0) \times \boldsymbol{F}_i$.

Example 2.24: Sailing boat

When a sailboat is going broad reach, as shown in Figure 2.22, the following forces are acting on the boat:

a) the wind in the sails generates a torque towards the bow around a horizontal axis that lies diagonal to the boat axis

b) the buoyancy of the water generates a torque along a horizontal axis parallel to the boat the counteracts heeling

c) the water drag on the hull generates a torque towards the bow around a horizontal axis that is orthogonal to the boat axis

d) the fin and the rudder generate lift forces that generate a torque around a vertical axis

e) the sailor stacks out in the trapeze to generate an additional torque in order to balance the torques

His aim is to minimize the heeling of the boat and to maximize the speed. The boat capsizes if he does not manage to balance the torques.

Gwicke commonswiki, public domain
Figure 2.22: A sailor stacking out in a trapeze in order to minimize the heeling of his sailboat.

2.9.4 *Self Test*

Problem 2.23. Fill in the details of the proof for Remark 2.28.

Problem 2.24. Fill in the details of the proof for Remark 2.30.

Problem 2.25. **Turning a wheel**

Two forces of magnitude 4 N are acting on a wheel of radius r that can freely rotate around its axis. What magnitude should a third force, $\boldsymbol{F}$, have that is attacking at a distance $r/2$ from the axis, such that there is no net torque acting on the wheel?

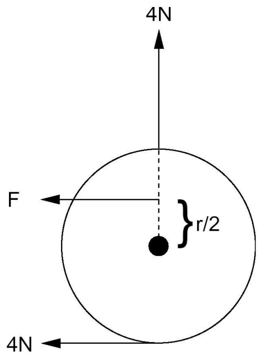


Figure 2.23: Setup for Problem 2.25.

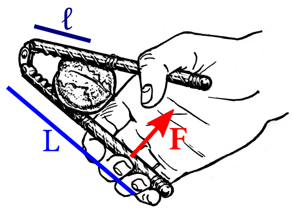
Problem 2.26. **Nutcrackers**

A common type of nutcrackers employs the principle of lever arms to crack nuts with a reasonable amount of force (see Problem 2.26). We idealize the nut as a spring with spring constant $k = 1\,\text{kN/mm}$ and assume that it breaks when it is compressed by $\Delta = 0.6\,\text{mm}$. The nut is mounted at a distance of $l = 3\,\text{cm}$ from the joint of the nutcracker and the hand exerts a force $\boldsymbol{F}$ at a distance L.

a) Demonstrate that a force of magnitude $F = \dfrac{lk\Delta}{L}$ is required to crack the nut.

b) Calculate the numerical value of F.

c) If you try to crack the nut by placing it under a heavy stone: which mass should that stone have in order to crack the nut?



based on Pearson Scott Foresman nutcracker-tool, public domain
Figure 2.24: Setup for Problem 2.26.

2.10 Worked example: Calder's mobiles

We describe here the setup of a traditional mobile where beams are supported by a string in the middle and balanced by attaching masses or further beams at their outer ends. The setup of a mobile can be laid out on a plane surface, as shown in Figure 2.25. The different parts of the mobile should not run into each other. Hence, they must not over overlap in the 2d layout.

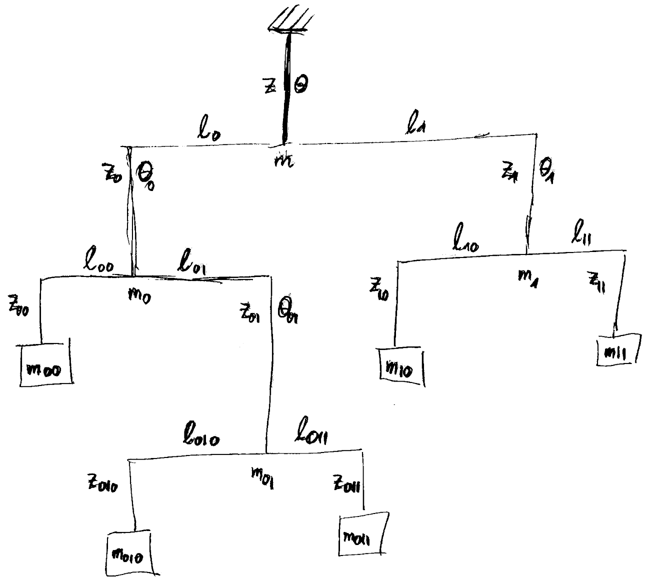


Figure 2.25: Notations for the mathematical description of the motion of a mobile. The mobile is suspended at a string of length z that holds a beam with two sections of length ℓ_0 to the left and ℓ_1 to the right, respectively. The string holds the total mass m of the mobile. When suspended, the beam can rotate by an angle θ out of the plane.

The left arm of the uppermost beam has length ℓ_0, and it holds another beam with an overall additional mass m_0 that can take an angle θ_0 out of the plane in the suspended mobile. Similarly, the right arm has length ℓ_1, and it holds another beam with an overall additional mass m_1 that can take an out-of-plane angle θ_1. The situation further down is described by hierarchical binary indices, as indicated in the figure.

The mobile can represented as a binary tree. Each beam has two arms reaching left (0) and right (1). We assume that the mass of the beams may be neglected, and reach the masses at the far ends of the mobile, by going down from the suspension and marking the track by a sequence of 0 and 1. The leftmost mass, 00, of the mobile in Figure 2.25 is reached by going left, 0, twice. The next one in counterclockwise direction by going left 0, right 1, left 0, and hence denoted as 010, and so forth. Hence, the mobile is build of beams that are labeled by some index I. They support a total mass m_I, and can rotated out of the plane by an angle θ_I. The beam has two arms of length ℓ_{I0} to the left and ℓ_{I1} to the right that support masses m_{I0} and m_{I1} attached to strings of length z_{I0} and z_{I1}. This hierarchical setup of the descriptions allows us to reduce the requirement of stability by a condition that the forces and torques acting on the beams must be balanced. For the forces this implies

$$\boldsymbol{F}_I = \boldsymbol{F}_{I0} + \boldsymbol{F}_{I1} \quad \Rightarrow \quad m_I = m_{I0} + m_{I1}$$

and for the torques we find

$$\ell_{I0} m_{I0} g = \ell_{I1} m_{I1} g \quad \Rightarrow \quad \ell_{I0} m_{I0} = \ell_{I1} m_{I1}$$

When we take all masses to take the same value m in Figure 2.25, we hence find

$$\ell_{010} = \ell_{011} \qquad \ell_{10} = \ell_{11} \qquad \ell_{00} = 2\,\ell_{01} \qquad 3\,\ell_0 = 2\,\ell_1$$

Moreover, vector calculus provides an effective means to specify the positions of the masses. We select the support of the mobile as origin of the coordinate system. The support of the uppermost beam is at position $(0, 0, -z)$. Then the far ends of the uppermost beam are at positions $\boldsymbol{l}_o = (-\ell_0 \cos\theta, -\ell_0 \sin\theta, -z)$ and $\boldsymbol{l}_1 = (\ell_1 \cos\theta, -\ell_1 \sin\theta, -z)$, respectively. Moreover, from the left end we reach the far ends of the next beam by the displacement vectors $\boldsymbol{l}_{00} = (-\ell_{00} \cos\theta_0, -\ell_{00} \sin\theta_0, -z_0)$ and $\boldsymbol{l}_{01} = (\ell_{01} \cos\theta_0, \ell_{01} \sin\theta_0, -z_0)$. Hence, the positions of the first two masses can be represented by the following sums of vectors

$$\boldsymbol{q}_{00} = \boldsymbol{l}_0 + \boldsymbol{l}_{00} - \begin{pmatrix} 0 \\ 0 \\ z_{00} \end{pmatrix} = \begin{pmatrix} -\ell_0 \cos\theta - \ell_{00} \cos\theta_0 \\ -\ell_0 \sin\theta - \ell_{00} \sin\theta_0 \\ -z - z_0 - z_{00} \end{pmatrix}$$

$$\boldsymbol{q}_{010} = \boldsymbol{l}_0 + \boldsymbol{l}_{01} + \boldsymbol{l}_{010} - \begin{pmatrix} 0 \\ 0 \\ z_{010} \end{pmatrix} = \begin{pmatrix} -\ell_0 \cos\theta + \ell_{01} \cos\theta_0 - \ell_{010} \cos\theta_{01} \\ -\ell_0 \sin\theta + \ell_{01} \sin\theta_0 - \ell_{010} \sin\theta_{01} \\ -z - z_0 - z_{01} - z_{010} \end{pmatrix}$$

We urge the reader to also work out the expressions for the positions of the other masses.

add discussion and stability analysis for bended beams

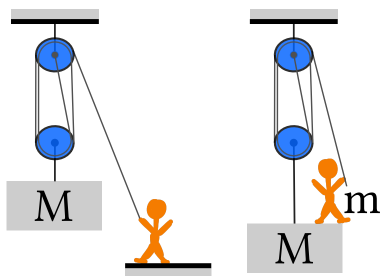
2.11 *Problems*

2.11.1 *Rehearsing Concepts*

Problem 2.27. **Tackling tackles and pulling pulleys**

a) Which forces are required to hold the balance in the left and the right sketch?

b) Let the sketched person and the weight have masses of $m = 75\,\text{kg}$ and $M = 300\,\text{kg}$, respectively. Which power is required then to haul the line at a speed of $1\,\text{m/s}$.
Hint: The power is defined here as the change of) $M g\, z(t)$ and $(M + m)\, g\, z(t)$, per unit time, respectively. Verify by dimensional analysis that this is a meaningful definition.



2.11.2 *Practicing Concepts*

Problem 2.28. **Angles between three balanced forces**

We consider three masses m_1, m_2, and m_3. With three ropes they are attached to a ring at position $\boldsymbol{q}_0$. The ropes with the attached masses hang over the edge of a table at the fixed positions $\boldsymbol{q}_1 = (x_1, 0)$, $\boldsymbol{q}_2 = (0, y_2)$, and $\boldsymbol{q}_3 = (w, y_3)$. Here, w denotes the width of the table board. We now determine the angles θ_{ij} between the ropes from $\boldsymbol{q}_0$ to $\boldsymbol{q}_i$ and $\boldsymbol{q}_j$, respectively.

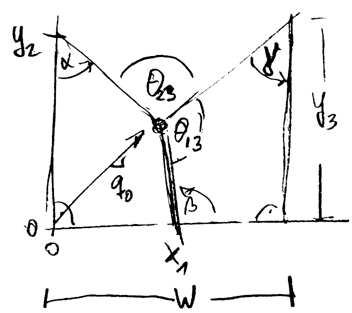


Figure 2.26: Setup of Problem 2.28.

a) Let $\hat{\boldsymbol{e}}_i = (\boldsymbol{q}_i - \boldsymbol{q}_0)/|\boldsymbol{q}_i - \boldsymbol{q}_0|$ be the unit vectors pointing from the ring to the positions where the ropes hang over the table edge, and θ_{ij} be the angle between $\hat{\boldsymbol{e}}_i$ and $\hat{\boldsymbol{e}}_j$. Argue why

$$\mathbf{0} = \sum_{i=1}^{3} m_i \, \hat{\boldsymbol{e}}_i$$

Multiplying this equation with $\hat{\boldsymbol{e}}_1, \dots \hat{\boldsymbol{e}}_3$ provides three equations that are linear in $\cos\theta_{ij}$. The first one is $0 = M_1 + M_2 \cos\theta_{12} + M_3 \cos\theta_{13}$. Find the other two equation, and solve the equations as follows.
From the equation that is given above you find $\cos\theta_{12}$ in terms of $\cos\theta_{13}$.
Inserting this into the other equation involving $\cos\theta_{12}$ (and rearranging terms) provides $\cos\theta_{23}$ in terms of $\cos\theta_{13}$.
Inserting this into the third equation provides

$$\cos\theta_{13} = \frac{M_2^2 - M_1^2 - M_3^2}{2\, M_1\, M_3}$$

b) Which angle θ_{23} do you find when $M_1 = M_2 = M_3$? The three forces have the same absolute value in this case. Which symmetry argument does then also provide the value of the angle?

c) Determine also the other two angles θ_{13} and θ_{12}. They can also be found from a symmetry argument without calculation.
Hint: The angles do not care which mass you denote as 1, 2, and 3.

d) Note that we found the angles θ_{ij} without referring to the positions $\boldsymbol{q}_1, \dots \boldsymbol{q}_3$! Make a sketch what this implies for the position of the ring, and how $\boldsymbol{q}_0$ changes qualitatively upon changing a mass.

The calculation of the position $\boldsymbol{q}_0$ can then be attacked by observing that

$$\boldsymbol{q}_0 = \boldsymbol{q}_1 + l_1 \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix} = \boldsymbol{q}_2 + l_2 \begin{pmatrix} \sin\alpha \\ -\cos\alpha \end{pmatrix} = \boldsymbol{q}_3 + l_3 \begin{pmatrix} -\sin\gamma \\ -\cos\gamma \end{pmatrix}$$

where l_i is the distance of the ring to the position where rope i hangs over the table. Further, the fact that the angles of quadrilaterals add to 2π provides

$$\alpha = \theta_{23} - \gamma \qquad \text{and} \qquad \beta = \frac{3\pi}{2} - \gamma - \theta_{13}$$

Altogether these are 8 equations to determine the two components of $\boldsymbol{q}_0$, $l_1, \dots l_3$, and the angles α, β, and γ. Determine $\boldsymbol{q}_0$.

Problem 2.29. **Torques acting on a ladder**

The sketch in the margin shows the setup of a ladder leaning to the roof of a hut. The indicated angle from the downwards vertical to the ladder is denoted as θ. There is a gravitational force of magnitude Mg acting of a ladder of mass M. At the point where it leans to the roof there is a normal force of magnitude F_r acting from the roof to the ladder. At the ladder feet there is a normal force to the ground of magnitude F_g, and a tangential friction force of magnitude γF_f. This is again the sketch to the ladder leaning to the roof of a hut. The angle from the downwards vertical to the ladder is denoted θ. There is a gravitational force of magnitude Mg acting of a ladder. At the point where it leans to the roof there is a normal force of magnitude F_r. At the ladder feet there is a normal force to the ground of magnitude F_g, and a tangential friction force of magnitude F_f.

change to problem given on homework sheet 3.

a) In principle there also is a friction force $\gamma_r F_r$ acting at the contact from the ladder to the roof. Why is it admissible to neglect this force?
Remark: There are at least two good arguments.

b) Determine the vertical and horizontal force balance for the ladder. Is there a unique solution?

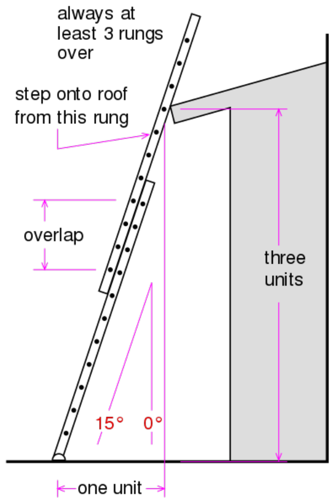
c) The feet of the ladder start sliding when F_f exceeds the maximum static friction force γF_g. What does this condition entail for the angle θ?
Assume that $\gamma \simeq 0.3$ What does this imply for the critical angle θ_c.

d) Where does the mass of the ladder enter the discussion? Do you see why?

e) Determine the torque acting on the ladder. Does it matter whether you consider the torque with respect to the contact point to the roof, the center of mass, or the foot of the ladder?

f) The ladder slides when the modulus of the friction force F_f exceeds a maximum value $\mu_S F_g$ where μ is the static friction coefficient for of the ladder feet on the ground. For metal feet on a wooden ground it takes a value of $\mu_S \simeq 2$. What does that tell about the angels where the ladder starts to slide?

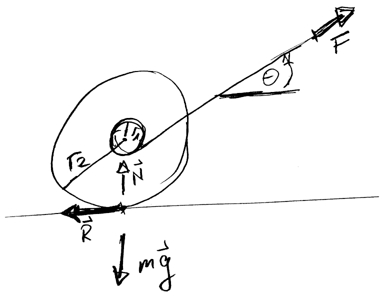
g) Why does a ladder commonly starts sliding when when a man has climbed to the top? Is there anything one can do against it? Is that even true, or just an urban legend?



original: Bradley, vector: Sarang / wikimedia public domain
Figure 2.27: Setup for Problem 2.29: leaning a ladder to a roof.

Problem 2.30. **Walking a yoyo**

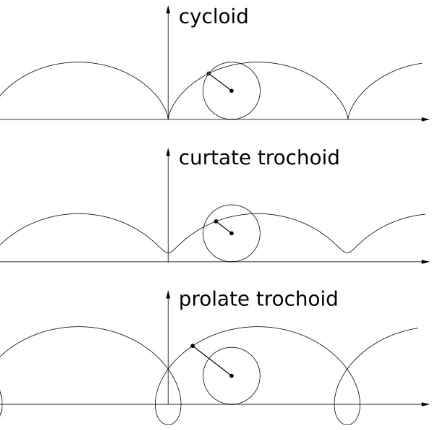
The sketch to the right shows a yoyo of mass m standing on the ground. It is held at a chord that extends to the top right. There are four forces acting on the yoyo: gravity mg, a normal force $\boldsymbol{N}$ from the ground, a friction force $\boldsymbol{R}$ at the contact to the ground, and the force $\boldsymbol{F}$ due to the chord. The chord is wrapped around an axle of radius r_1. The outer radius of the yoyo is r_2.



a) Which conditions must hold such that there is no net force acting on the center of mass of the yoyo?

b) For which angle θ does the torque vanish?

c) Perform an experiment: What happens for larger and for smaller angels θ? How does the yoyo respond when fix the height where you keep the chord and pull continuously?



based on Kmhkmh Zykloiden, CC BY 4.0

Problem 2.31. **Retro-reflector paths on bike wheels**

The more traffic you encounter when it becomes dark the more important it becomes to make your bikes visible. Retro-reflectors fixed in the sparks enhance the visibility to the sides. They trace a path of a curtate trochoid that is characterized by the ratio ρ of the reflectors distance d to the wheel axis and the wheel radius r. A small stone in the profile traces a cycloid ($\rho = 1$). Animations of the trajectories can be found at `https://en.wikipedia.org/wiki/Trochoid` and `http://katgym.by.lo-net2.de/c.wolfseher/web/zykloiden/zykloiden.html`.

A trochoid is most easily described in two steps: Let $\boldsymbol{M}(\theta)$ be the position of the center of the disk, and $\boldsymbol{D}(\theta)$ the vector from the center to the position $\boldsymbol{q}(\theta)$ that we follow (i.e. the position of the retro-reflector) such that $\boldsymbol{q}(\theta) = \boldsymbol{M}(\theta) + \boldsymbol{D}(\theta)$.

a) The point of contact of the wheel with the street at the initial time t_0 is the origin of the coordinate system. Moreover, we single out one spark and denote the change of its angle with respect to its initial position as θ. Note that negative angles θ describe forward motion of the wheel!

Sketch the setup and show that

$$\boldsymbol{M}(\theta) = \begin{pmatrix} -r\theta \\ r \end{pmatrix}, \qquad \boldsymbol{D}(\theta) = \begin{pmatrix} -d\,\sin(\varphi+\theta) \\ d\,\cos(\varphi+\theta) \end{pmatrix}.$$

What is the meaning of φ in this equation?

check signs of components of $\boldsymbol{D}$

b) The length of the track of a trochoid can be determined by integrating the modulus of its velocity over time, $L = \int_{t_0}^{t} \mathrm{d}t\ |\dot{\boldsymbol{q}}(\theta(t))|$. Show that therefore

$$L = r \int_0^\theta \mathrm{d}\theta\ \sqrt{1+\rho^2+2\rho\,\cos(\varphi+\theta)}$$

c) Consider now the case of a cycloid and use $\cos(2x) = \cos^2 x - \sin^2 x$ to show that the expression for L can then be written as

$$L = 2\,r \int_0^\theta \mathrm{d}\theta\ \left|\cos\frac{\varphi+\theta}{2}\right|$$

How long is one period of the track traced out by a stone picked up by the wheel profile?

2.11.3 *Mathematical Foundation*

Problem 2.32. **The natural numbers modulo n are a group**

We consider here groups G_n where the combined action of group elements can be represented as a sum of two numbers modulo $n \in \mathbb{N}$. In other words, for the elements of G_n can be represented by the numbers $\{0, \dots, n-1\}$, and for all $a, b \in G_n$ we define $a \circ b = (a+b) \bmod n$.

a) Show that G_n is a group.

b) Show that G_n represents the rotations that interchange the vertices of a regular n-sided polygon.

Problem 2.33. **Groups with four elements**

In Problem 2.32 we encountered the group G_n. Here, we will study another group with four elements. The neutral element will be denoted as n.

a) Show that the group has at least one non-trivial element e that is self-inverse, $e \circ e = n$.
Remark: Non-trivial means here that $e \neq n$.

b) Show that the group is isomorphic to G_4 if there is exactly one non-trivial element that is self-inverse. In other words: the group elements can be represented in that case by the numbers $\{0, \dots, 3\}$, and the operation of the group on two of its elements yields the same result as the action of G_4 on the corresponding numbers.

c) Show that the group is isomorphic to G_4 if there is at least one element that is not self-inverse.

d) Determine the group table for the case where all group elements are self-inverse. Show that it is unique, and that it is isomorphic to the symmetry group of rectangles (cf. Problem 2.6).

e) Proof that all groups with four elements are commutative by representing the group elements in terms of generating elements. Do not refer to the group table.

Problem 2.34. **Conic Sections**

A conic section describes the line of intersection of a double cone C and a plane P in three dimensions. In the margin we show the shape of conic sections for different inclinations that are characterized by the eccentricity ϵ. Depending on the inclination of the plane one observes

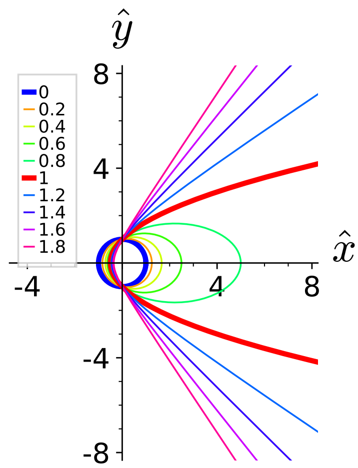


Figure 2.28: Conic sections for different eccentricity ϵ. i.e., the ratio of the slope of the plane P and the surface of the double cone, as observed in a plane that contains the axis of the double cone and is orthogonal to P.

- a circle, when the axis of the cone is orthogonal to P, i.e. for $\epsilon = 0$,
- an ellipse, when the plane is slightly tilted, $\epsilon < 1$,

- a parabola, when its inclination matches with the opening angle of the cone, $\epsilon = 1$, and
- a hyperbola, when it intersects with both sides of the double cone, $\epsilon > 1$.

a) Sketch the different types of intersection of the double cone and the plane.

b) Determine the vector $\boldsymbol{a}$ that points from the vertex of the double cone to the point where the plane intersects the axis of the double cone.

c) Describe the points in the intersection as sum of $\boldsymbol{a}$ and a vector $\boldsymbol{b}$ that lies in the plane.

d) Determine the length of the vector $\boldsymbol{b}$ as function of the angle θ that characterizes the direction of $\boldsymbol{b}$ in P. How can this expression be used to plot the functions shown in Figure 4.19.

Problem 2.35. **Linear dependence of three vectors in 2D**

In the lecture I pointed out that every vector $\boldsymbol{v} = (v_1, v_2)$ of a two-dimensional vector space can be represented as a *unique* linear combination of two linearly independent vectors $\boldsymbol{a}$ and $\boldsymbol{b}$,

$$\boldsymbol{v} = \alpha\, \boldsymbol{a} + \beta\, \boldsymbol{b}$$

In this exercise we revisit this statement for $\mathbb{R}^2$ with the standard forms of vector addition and multiplication by scalars.

a) Provide a triple of vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{v}$ such that $\boldsymbol{v}$ can *not* be represented as a scalar combination of $\boldsymbol{a}$ and $\boldsymbol{b}$.

b) To be specific we henceforth fix

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

Determine the numbers α and β such that

$$\boldsymbol{v} = \alpha\, \boldsymbol{a} + \beta\, \boldsymbol{b}$$

c) Consider now also a third vector

$$\boldsymbol{c} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and find two different choices for (α, β, γ) such that

$$\boldsymbol{v} = \alpha\, \boldsymbol{a} + \beta\, \boldsymbol{b} + \gamma \boldsymbol{c}$$

What is the general constraints on (α, β, γ) such that $\boldsymbol{v} = \alpha\, \boldsymbol{a} + \beta\, \boldsymbol{b} + \gamma \boldsymbol{c}$.
What does this imply on the number of solutions?

d) Discuss now the linear dependence of the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ by exploring the solutions of

$$\boldsymbol{0} = \alpha\, \boldsymbol{a} + \beta\, \boldsymbol{b} + \gamma \boldsymbol{c}$$

How are the constraints for the null vector related to those obtained in part c)?

Problem 2.36. **Algebraic number fields**

Consider the set $\mathbb{K} = \mathbb{Q} + \mathrm{I}\mathbb{Q}$ with $\mathrm{I}^2 \in \mathbb{Q}$. We define the operations $+$ and $\cdot$ in analogy to those of the complex numbers (cf. Example 2.13): For $z_1 = x_1 + \mathrm{I}y_1$ and $z_2 = x_2 + \mathrm{I}y_2$ we have $x_1, y_1, x_2, y_2 \in \mathbb{Q}$ and

$$\begin{aligned} \forall z_1, z_2 \in \mathbb{K} : & z_1 + z_2 = (x_1 + x_2) + \mathrm{I}\,(y_1 + y_2) \\ & z_1 \cdot z_2 = (x_1\, x_2 + \mathrm{I}^2\, y_1 y_2) + \mathrm{I}\,(x_1\, y_2 + y_1\, x_2) \end{aligned}$$

$$\forall c \in \mathbb{Q}, z = (x + \mathrm{i}y) \in \mathbb{K} : cz = c\,x + \mathrm{I}\,y$$

a) Let I be a rational number, $\mathrm{I} \in \mathbb{Q}$. Show that $\mathbb{K} = \mathbb{Q}$.

b) Consider $\mathrm{I} = \sqrt{10}$. Show that $\mathbb{K}$ is a field that is different from $\mathbb{Q}$.

c) Consider $\mathrm{I} = \sqrt{8}$. In this case $\mathbb{K}$ is *not* a field! Why?

Find the general rule: For which natural numbers n does $\mathrm{I} = \sqrt{n}$ provide a non-trivial field?
Remark: Non-trivial means here different from $\mathbb{Q}$.

Problem 2.37. **Bases for polynomials**

We consider the set of polynomials $\mathbb{P}_N$ of degree N with real coefficients p_n, $n \in \{0, \dots, N\}$,

$$\mathbb{P}_N := \left\{ \boldsymbol{p} = \left(\sum_{k=0}^{N} p_k\, x^k \right) \quad \text{mit} \;\; p_k \in \mathbb{R}, k \in \{0, \dots, N\} \right\}$$

a) Demonstrate that $(\mathbb{P}_N, \mathbb{R}, +, \cdot)$ is a vector space when one adopts the operations

$$\forall \quad \boldsymbol{p} = \left(\sum_{k=0}^{N} p_k\, x^k \right) \in \mathbb{P}_N\,, \quad \boldsymbol{q} = \left(\sum_{k=0}^{N} q_k\, x^k \right) \in \mathbb{P}_N\,, \text{ and } \; c \in \mathbb{R} \; :$$

$$\boldsymbol{p} + \boldsymbol{q} = \left(\sum_{k=0}^{N} (p_k + q_k)\, x^k \right) \quad \text{and} \quad c \cdot \boldsymbol{p} = \left(\sum_{k=0}^{N} (c\, p_k)\, x^k \right) .$$

(b) Demonstrate that

$$\boldsymbol{p} \cdot \boldsymbol{q} = \left(\int_0^1 \mathrm{d}x \left(\sum_{k=0}^{N} p_k\, x^k \right) \left(\sum_{j=0}^{N} q_j\, x^j \right) \right) ,$$

establishes a scalar product on this vector space.

(c) Demonstrate that the three polynomials $\boldsymbol{b}_0 = (1)$, $\boldsymbol{b}_1 = (x)$ and $\boldsymbol{b}_2 = (x^2)$ form a basis of the vector space $\mathbb{P}_2$: For each polynomial $\boldsymbol{p}$ in $\mathbb{P}_2$ there are real numbers x_k, $k \in \{0, 1, 2\}$, such that $\boldsymbol{p} = x_0\, \boldsymbol{b}_0 + x_1\, \boldsymbol{b}_1 + x_2\, \boldsymbol{b}_2$. However, in general we have $x_i \neq \boldsymbol{p} \cdot \boldsymbol{b}_i$. Why is that?
Hint: Is this an orthonormal basis?

(d) Demonstrate that the three vectors $\hat{\boldsymbol{e}}_0 = (1)$, $\hat{\boldsymbol{e}}_1 = \sqrt{3}\,(2\,x - 1)$ and $\hat{\boldsymbol{e}}_2 = \sqrt{5}\,(6\,x^2 - 6\,x + 1)$ are orthonormal.

(e) Demonstrate that every vector $\boldsymbol{p} \in \mathbb{P}_2$ can be written as a scalar combination of $(\hat{\boldsymbol{e}}_0, \hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2)$,

$$\boldsymbol{p} = (\boldsymbol{p} \cdot \hat{\boldsymbol{e}}_0)\, \hat{\boldsymbol{e}}_0 + (\boldsymbol{p} \cdot \hat{\boldsymbol{e}}_1)\, \hat{\boldsymbol{e}}_1 + (\boldsymbol{p} \cdot \hat{\boldsymbol{e}}_2)\, \hat{\boldsymbol{e}}_2\,.$$

Hence, $(\hat{\boldsymbol{e}}_0, \hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2)$ form an orthonormal basis of $\mathbb{P}_2$.

*(f) Find a constant c and a vector $\hat{\boldsymbol{n}}_1$, such that $\hat{\boldsymbol{n}}_0 = (c\,x)$ and $\hat{\boldsymbol{n}}_1$ form an orthonormal basis of $\mathbb{P}_1$.

Problem 2.38. **Systems of linear equations**

A system of N linear equations of M variables $x_1, \dots x_M$ comprises N equations of the form

$$\begin{aligned} b_1 &= a_{11}\, x_1 + a_{12}\, x_2 + \cdots + a_{1M}\, x_M \\ b_2 &= a_{21}\, x_1 + a_{22}\, x_2 + \cdots + a_{2M}\, x_M \\ \vdots \quad & \quad \vdots \\ b_N &= a_{N1}\, x_1 + a_{N2}\, x_2 + \cdots + a_{NM}\, x_M \end{aligned}$$

where $b_i, a_{ij} \in \mathbb{R}$ for $i \in \{1, \dots, N\}$ and $j \in \{1, \dots, M\}$.

a) Demonstrate that the linear equations $(\mathbb{L}_M, \mathbb{R}, +, \cdot)$ form a vector space when one adopts the operations

$$\begin{aligned} \forall \quad & \boldsymbol{p} = \left[p_0 = p_1\, x_1 + p_2\, x_2 + \cdots + p_M\, x_M\right] \in \mathbb{L}_N\,, \\ & \boldsymbol{q} = \left[q_0 = q_1\, x_1 + q_2\, x_2 + \cdots + q_M\, x_M\right] \in \mathbb{L}_N\,, \\ & c \in \mathbb{R} \;: \end{aligned}$$

$$\begin{aligned} \boldsymbol{p} + \boldsymbol{q} &= \left[p_0 + q_0 = (p_1 + q_1)\, x_1 + (p_2 + q_2)\, x_2 + \cdots + (p_M + q_M)\, x_M\right] \\ c \cdot \boldsymbol{p} &= \left[c\, p_0 = c\, p_1\, x_1 + c\, p_2\, x_2 + \cdots + c\, p_M\, x_M\right]. \end{aligned}$$

How do these operations relate to the operations performed in Gauss elimination to solve the system of linear equations?

b) The system of linear equations can also be stated in the following form

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{N1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{N2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1M} \\ a_{2M} \\ \vdots \\ a_{NM} \end{pmatrix} x_M$$

$$\Leftrightarrow \qquad \boldsymbol{b} = x_1\, \boldsymbol{a}_1 + x_2\, \boldsymbol{a}_2 + \cdots + x_M\, \boldsymbol{a}_M$$

where $\boldsymbol{b}$ is expressed as a linear combination of $\boldsymbol{a}_1, \ldots \boldsymbol{a}_M$ by means of the numbers $x_1, \ldots, x_M$. What do the conditions on linear independence and representation of vectors by means of a basis tell about the existence and uniqueness of the solutions of a system of linear equations.

2.11.4 *Transfer and Bonus Problems, Riddles*

Problem 2.39. **Crossing a river**

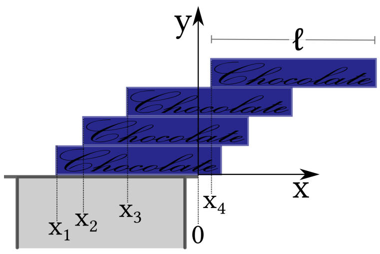
A ferry is towed at the bank of a river of width $B = 100$ m that is flowing at a velocity $v_F = 4$ m/s to the right. At time $t = 0$ s it departs and is heading with a constant velocity $v_B = 10$ km/h to the opposite bank.

a) When will it arrive at the other bank when it always heads straight to the other side? (In other words, at any time its velocity is perpendicular to the river bank.)

   How far will it drift downstream on its journey?

b) In which direction (i.e. angle of velocity relative to the downstream velocity of the river) must the ferryman head to reach exactly at the opposite side of the river?

   Determine first the general solution. What happens when you try to evaluate it for the given velocities?



Problem 2.40. **Piling bricks**

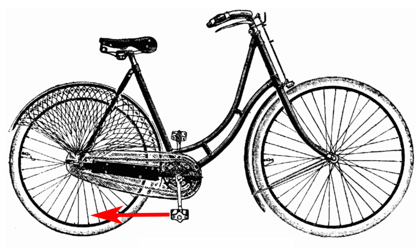
At Easter and Christmas Germans consume enormous amounts of chocolate. If you happen to come across a considerable pile of chocolate bars (or beer mats, or books, or anything else of that form) I recommend the following experiment:

a) We consider N bars of length l piled on a table. What is the maximum amount that the topmost bar can reach beyond the edge of the table.

b) The sketch above shows the special case $N = 4$.
   However, what about the limit $N \to \infty$?

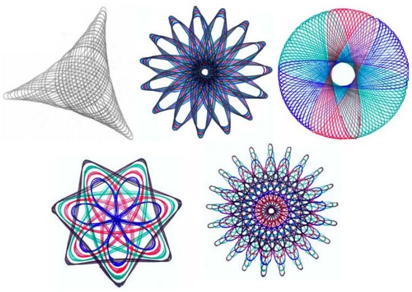
Problem 2.41. **Where does the bike go?**

Consider the picture of the bicycle to the left. The red arrow indicates a force that is acting on the paddle in backward direction.
Will the bicycle move forwards or backwards?
Take a bike and do the experiment!



adapted from picture "Damenfahrrad von 1900" in article "Fahrrad" of Lueger (1926–1931)

Problem 2.42. **Hypotrochoids, roulettes, and the Spirograph**



wikimedia, public domain

A roulette is the curve traced by a point (called the generator or pole) attached to a disk or other geometric object when that object rolls without slipping along a fixed track. A pole on the circumference of a disk that rolls on a straight line generates a cycloid. A pole inside that disk generates a trochoid. If the disk rolls along the inside or outside of a circular track it generates a hypotrochoid. The latter curves can be drawn with a spirograph, a beautiful drawing toy based on gears that illustrates the mathematical concepts of the least common multiple (LCM) and the lowest common denominator (LCD).

a) Consider the track of a pole attached to a disk with n cogs that rolls inside a circular curve with $m > n$ cogs. Why does the resulting curve form a closed line? How many revolutions does the disk make till the curve closes? What is the symmetry of the resulting roulette? (The curves to the top left is an examples with three-fold symmetry, and the one to the bottom left has seven-fold symmetry.)

b) Adapt the description for the curves developed in Problem 2.31 such that you can describe hypotrochoids.

c) Test your result by writing a Python program that plots the curves for given m and n.

2.12 *Further reading*

The second chapter of Großmann (2012) provides a clear and concise introduction to the mathematical framework of vectors with an emphasis on applications to physics problems.

A nice discussion of force and torque balances with many worked exercises can be found in Chapter 2 of Morin (2007).

Bibliography

Archimedes, 1878, in *Pappi Alexandrini Collectionis, Book VIII, c. AD 340*, edited by F. O. Hultsch (Apud Weidmannos, Berlin), p. 1060, cited following https://commons.wikimedia.org/wiki/File:Archimedes_lever,_vector_format.svg.

Arnol'd, V. I., 1992, *Ordinary Differential Equations* (Springer, Berlin).

Epstein, L. C., 2009, *Thinking Physics, Understandable Practical Reality* (Insight Press, San Francisco).

Finney, G. A., 2000, American Journal of Physics **68**, 223–227.

Gale, D. S., 1970, American Journal of Physics **38**, 1475.

Gommes, C. J., 2010, Am. J. Phys. **78**, 236.

Großmann, S., 2012, *Mathematischer Einführungskurs für die Physik* (Springer), very clear introduction of the mathematical concepts for physics students., URL https://doi.org/10.1007/978-3-8348-8347-6.

Harte, J., 1988, *Consider a spherical cow. A course in environmental problem solving* (University Science Books, Sausalito, CA).

Kagan, D., L. Buchholtz, and L. Klein, 1995, Phys.Teach. **33**, 150.

Lueger, O., 1926–1931, *Luegers Lexikon der gesamten Technik und ihrer Hilfswissenschaften* (Dt. Verl.-Anst., Stuttgart), URL https://digitalesammlungen.uni-weimar.de/viewer/resolver?urn=urn:nbn:de:gbv:wim2-g-3163268.

Mahajan, S., 2010, *Street-fighting Mathematics. The Art of Educated Guessing and Opportunistic Problem Solving* (MIT Press, Cambridge, MA), ISBN 9780262514293 152, some parts are available from the author, URL https://mitpress.mit.edu/books/street-fighting-mathematics.

Morin, D., 2007, *Introduction to Classical Mechanics* (Cambridge), comprehensive introduction with a lot of exercises—many of them with worked solutions. The present lectures cover the Chapters 1–8 of the book. I warmly recomment to study Chapters 1 and 6., URL https://scholar.harvard.edu/david-morin/classical-mechanics.

Morin, D., 2014, *Problems and Solutions in Introductory Mechanics* (CreateSpace), a more elementary introduction with a lot of solved exercises self-published at Amazon. Some Chapters can also be downloaded from the autor's home page, URL `https://scholar.harvard.edu/david-morin/mechanics-problem-book`.

Murray, J., 2002, *Mathematical Biology* (Springer).

Nordling, C., and J. Österman, 2006, *Physics Handbook for Science and Engineering* (Studentlitteratur, Lund), 8 edition, ISBN 91-44-04453-4, quoted after Wikipedia's List of humorous units of measurement, accessed on 5 May 2020.

Purcell, E. M., 1977, American Journal of Physics **45**(3).

Seifert, H. S., M. W. Mills, and M. Summerfield, 1947, American Journal of Physics **15**(3), 255.

Sommerfeld, A., 1994, *Mechanik*, volume 1 of *Vorlesungen über theoretische Physik* (Harri Deutsch, Thun, Frankfurt/M.).

Zee, A., 2020, *Fly by Night Physics: How Physicists Use the Backs of Envelopes* (Princeton UP), ISBN 9780691182544.

Index

D-tuples, 28
$\setminus$ set difference, **18**
$\bigcap$ set intersection, **18**
$\bigcup$ set union, **18**
Kronecker δ_{ij}, 37
$\emptyset$ empty set, 15
$\exists$, there is, **18**
$\mathbb{R}^D$, 28
$\subseteq$ subset, **16**
Levi-Civita tensor ε_{ijk}, 43

$\forall$, for all, **18**

axiom
  force balance, 32
  torque balance, 46

BAC-CAB rule, 44
basis
  orthonormal basis in $\mathbb{R}^3$, 43

cardinality, **15**
Cartesian product, *see* product, Cartesian
complex number
  2D rotation, 26
  complex conjugate, 26
  field, 25
coordinates
  basis change, 39
  Cartesian, 14, **36**
  polar, 14
    and complex numbers, 26
cross product
  geometric interpretation, 45
  Jacobi identity, 45
cycloid, 52

Dirac notation, 39

Einstein summation convention, 38
element, **15**

field, **25**
  complex numbers, 25
  rational numbers, 25
  real numbers, 25
  smallest, 25
for all $\forall$, **18**
force
  balance, 32
friction
  static, 32

group, **21**
  90°-rotations with respect to two vertical axis, 22
  3 elements, 23
  4 elements, 53
  Abelian, **21**
  commutative, **21**
  dihedral of order 6, 24
  edit text field, 22
  rotations of regular polygons, 53
  smallest, 21
  symmetries of rectangle, 23

hypotrochoids, 58

inner product, 34
  geometric interpretation, 35
interval, **20**

Jacobi identity, 43, 45

junctors
overview, 17
precedence, 17

Kronecker δ, 37

Levi-Civita tensor ε_{ijk}, 43
logical junctors
overview, 17
precedence, 17

matrix, **28**
multiplication, 29
vector space, 29

numbers
complex, **19**
field property, 25
even, **19**
integer, **19**
group property, 24
natural, **19**
negative, **19**
odd, **19**
rational, **19**
real, **19**

polar coordinates, 14
and complex numbers, 26
polynomials
vector space, 29, 31
product
Cartesian, **18**

rotation
in 2D by complex number, 26
roulettes, 58

set, **15**
cardinality, **15**
complement, **18**
definition by property, 16
difference ($\setminus$), **18**
element, **15**
empty $\emptyset$, 15
equivalence, 17
examples, 15
intersection ($\bigcap$), 18
logical quantors, 18
set operations, 18
sets of sets
examples, 16
subset ($\subseteq$), **16**
superset ($\supseteq$), **16**
union ($\bigcup$), **18**
set operations, **18**
slackline, 32
spirograph, 58

there is $\exists$, **18**
torque, **46**
balance, 46
lever, 42
torque balance
mobile, 48
trochoid, 52

vector
cross product, **43**
orthogonal, 35
vector product, **43**
Jacobi identity, 43
vector space, **27**
$\mathbb{R}^D$, 28
basis, **36**
coordinate transformation, 39
coordinates, **36**
displacements, 27
inner product, 34
invariance of the inner product, 39
matrix, 29
polynomials, 29, 31
scalar product, 38