Geometrisches Rechnen
(WS 2020/21)

Martin Held

FB Computerwissenschaften
Universität Salzburg
A-5020 Salzburg, Austria
held@cs.sbg.ac.at

January 15, 2021
Personalia

Instructor: M. Held.
Email: held@cs.sbg.ac.at.
Base-URL: http://www.cosy.sbg.ac.at/~held.
Office: Universität Salzburg, Computerwissenschaften, Rm. 1.20, Jakob-Haringer Str. 2, 5020 Salzburg-Itzling.

Phone number (office): (0662) 8044-6304.
Phone number (secr.): (0662) 8044-6328.
Formalia

URL of course: .../teaching/geom_rechnen/geom_rechnen.html.

Lecture times: UV Thursday 8\textsuperscript{00}–10\textsuperscript{45}.

Venue: HS Christian Doppler, Jakob-Haringer Str. 2a.

Note: — PS is graded according to continuous-assessment mode!
Electronic Slides and Online Material

In addition to these slides, you are encouraged to consult the WWW home-page of this lecture:

http://www.cosy.sbg.ac.at/~held/teaching/geom_rechnen/geom_rechnen.html.

In particular, this WWW page contains links to online manuals, slides, and code.
A Few Words of Warning

I hope that these slides will serve as a practice-minded introduction to the mathematics of geometric computing. I would like to warn you explicitly not to regard these slides as the sole source of information on the topics of my course. It may and will happen that I’ll use the lecture for talking about subtle details that need not be covered in these slides! In particular, the slides won’t contain all sample calculations, proofs of theorems, demonstrations of algorithms, or solutions to problems posed during my lecture. That is, by making these slides available to you I do not intend to encourage you to attend the lecture on an irregular basis.
Acknowledgments

These slides are a revised and extended version of notes and slides originally prepared for my graphics courses. Those graphics slides were partially based on write-ups of former students, and I would like to express my thankfulness for their help with those graphics slides. This revision and extension was carried out by myself, and I am responsible for all errors.

Salzburg, August 2020

Martin Held
Legal Fine Print and Disclaimer

To the best of our knowledge, these slides do not violate or infringe upon somebody else’s copyrights. If copyrighted material appears in these slides then it was considered to be available in a non-profit manner and as an educational tool for teaching at an academic institution, within the limits of the “fair use” policy. For copyrighted material we strive to give references to the copyright holders (if known). Of course, any trademarks mentioned in these slides are properties of their respective owners.

Please note that these slides are copyrighted. The copyright holder(s) grant you the right to download and print it for your personal use. Any other use, including non-profit instructional use and re-distribution in electronic or printed form of significant portions of it, beyond the limits of “fair use”, requires the explicit permission of the copyright holder(s). All rights reserved.

These slides are made available without warrant of any kind, either express or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. In no event shall the copyright holder(s) and/or their respective employers be liable for any special, indirect or consequential damages or any damages whatsoever resulting from loss of use, data or profits, arising out of or in connection with the use of information provided in these slides.
Recommended Textbooks

G.E. Farin, D. Hansford.
*Practical Linear Algebra: A Geometry Toolbox.*

M.E. Mortenson.

*immersive linear algebra.*
# Table of Content

1. Introduction
2. Algebraic Concepts
3. Basic Linear Algebra
4. Geometric Objects
5. Basic Concepts of Topology
6. Transformations
7. Floating-Point Arithmetic and Numerical Mathematics
1 Introduction
- Motivation
- Notation
Introduction

- Motivation
- Notation
Consider the following four polynomials (in the variable $x$):

\[
p_1(x) := (1 - x)^3 \quad p_2(x) := 3x(1 - x)^2 \quad p_3(x) := 3x^2(1 - x) \quad p_4(x) := x^3
\]
Basis of a Vector Space

- Consider the following four polynomials (in the variable $x$):
  \[
p_1(x) := (1 - x)^3 \quad p_2(x) := 3x(1 - x)^2 \quad p_3(x) := 3x^2(1 - x) \quad p_4(x) := x^3
  \]

- Question: Can we write every polynomial $p(x)$ of degree at most three as
  \[
p(x) = \lambda_1 \cdot p_1(x) + \lambda_2 \cdot p_2(x) + \lambda_3 \cdot p_3(x) + \lambda_4 \cdot p_4(x)
  \]
  for suitable $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$?
Consider the following four polynomials (in the variable $x$): 

\[ p_1(x) := (1 - x)^3 \quad p_2(x) := 3x(1 - x)^2 \quad p_3(x) := 3x^2(1 - x) \quad p_4(x) := x^3 \]

**Question:** Can we write every polynomial $p(x)$ of degree at most three as 

\[ p(x) = \lambda_1 \cdot p_1(x) + \lambda_2 \cdot p_2(x) + \lambda_3 \cdot p_3(x) + \lambda_4 \cdot p_4(x) \]

for suitable $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$?

**Answer:** Yes — because $p_1(x), p_2(x), p_3(x), p_4(x)$ form a basis of the vector space of polynomials (in $x$) of degree at most three.
Basis of a Vector Space

- Consider the following four polynomials (in the variable $x$):
  \[ p_1(x) := (1 - x)^3 \quad p_2(x) := 3x(1 - x)^2 \quad p_3(x) := 3x^2(1 - x) \quad p_4(x) := x^3 \]

- Question: Can we write every polynomial $p(x)$ of degree at most three as
  \[ p(x) = \lambda_1 \cdot p_1(x) + \lambda_2 \cdot p_2(x) + \lambda_3 \cdot p_3(x) + \lambda_4 \cdot p_4(x) \]
  for suitable $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$?

- Answer: Yes — because $p_1(x), p_2(x), p_3(x), p_4(x)$ form a basis of the vector space of polynomials (in $x$) of degree at most three.

- What is a vector space? What is a basis? And what is a polynomial?
Complex Numbers for Generating Pretty Images

- How can we generate such an image?

This looks like the visualization of a Julia set. Similar to the Mandelbrot set, Julia sets can be generated via visualizing properties of series of complex numbers.
Complex Numbers for Generating Pretty Images

How can we generate such an image?

Answer: This looks like the visualization of a Julia set. Similar to the Mandelbrot set, Julia sets can be generated via visualizing properties of series of complex numbers.
How can we generate such an image?

Answer: This looks like the visualization of a Julia set. Similar to the Mandelbrot set, Julia sets can be generated via visualizing properties of series of complex numbers.

What is a complex number?
Area of a Triangle

- Consider the triangle (in the plane) with corners $(2, 1)$, $(7, 2)$ and $(3, 5)$.

![Diagram of a triangle with coordinates](image)

Question: How can we compute the area $A$ of that triangle?

The area of that triangle can be obtained by a simple determinant computation:

$$A = \frac{1}{2} \cdot \det \begin{bmatrix} 2 & 1 & 1 \\ 7 & 2 & 1 \\ 3 & 5 & 1 \end{bmatrix} = \frac{19}{2}$$

What is a determinant? And why is this claim true?
Area of a Triangle

Consider the triangle (in the plane) with corners (2, 1), (7, 2) and (3, 5).

Question: How can we compute the area $A$ of that triangle?

The area of that triangle can be obtained by a simple determinant computation:

$$A = \frac{1}{2} \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 7 & 2 & 1 \\ 3 & 5 & 1 \end{pmatrix} = \frac{19}{2}$$
Area of a Triangle

- Consider the triangle (in the plane) with corners (2, 1), (7, 2) and (3, 5).

![Diagram of a triangle with coordinates](image)

- Question: How can we compute the area $A$ of that triangle?

- The area of that triangle can be obtained by a simple determinant computation:

$$A = \frac{1}{2} \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 7 & 2 & 1 \\ 3 & 5 & 1 \end{pmatrix} = \frac{19}{2}$$
Area of a Triangle

- Consider the triangle (in the plane) with corners (2, 1), (7, 2) and (3, 5).

- Question: How can we compute the area $A$ of that triangle?
- The area of that triangle can be obtained by a simple determinant computation:

  $$A = \frac{1}{2} \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 7 & 2 & 1 \\ 3 & 5 & 1 \end{pmatrix} = \frac{19}{2}$$

- What is a determinant? And why is this claim true?
Orthogonal Frame

- Assume that the vector \( \nu_1 := (1, 2, 3) \) is a tangent vector to the curve \( \gamma \) at the point \( \gamma(6) \).
- Question: How can be quickly find two other vectors \( \nu_2 \) and \( \nu_3 \) that form an orthogonal frame with \( \nu_1 \)?

An orthogonal frame can be obtained by taking a vector cross-product of two suitable vectors:

\[
\nu_2 := \begin{pmatrix}
-2 \\
1 \\
0
\end{pmatrix}
\quad \text{and} \quad
\nu_3 := \begin{pmatrix}
-3 \\
-6 \\
5
\end{pmatrix}
\]

Then \( \nu_1 \perp \nu_2 \), \( \nu_1 \perp \nu_3 \) and \( \nu_2 \perp \nu_3 \).

By the way, what is a curve? And what does orthogonal mean?
Orthogonal Frame

- Assume that the vector $\nu_1 := (1, 2, 3)$ is a tangent vector to the curve $\gamma$ at the point $\gamma(6)$.
- Question: How can be quickly find two other vectors $\nu_2$ and $\nu_3$ that form an orthogonal frame with $\nu_1$?
- Answer: An orthogonal frame can be obtained by taking a vector cross-product of two suitable vectors:

$$
\nu_2 := \begin{pmatrix}
-2 \\
1 \\
0
\end{pmatrix}
\quad \text{and} \quad
\nu_3 := \begin{pmatrix}
1 \\
2 \\
3
\end{pmatrix} \times \begin{pmatrix}
-2 \\
1 \\
0
\end{pmatrix} = \begin{pmatrix}
-3 \\
-6 \\
5
\end{pmatrix}
$$

Then $\nu_1 \perp \nu_2$, $\nu_1 \perp \nu_3$ and $\nu_2 \perp \nu_3$. 

By the way, what is a curve? And what does orthogonal mean?
Orthogonal Frame

- Assume that the vector $\nu_1 := (1, 2, 3)$ is a tangent vector to the curve $\gamma$ at the point $\gamma(6)$.
- Question: How can be quickly find two other vectors $\nu_2$ and $\nu_3$ that form an orthogonal frame with $\nu_1$?
- Answer: An orthogonal frame can be obtained by taking a vector cross-product of two suitable vectors:

$$\nu_2 := \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \nu_3 := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -3 \\ -6 \\ 5 \end{pmatrix}$$

Then $\nu_1 \perp \nu_2$, $\nu_1 \perp \nu_3$ and $\nu_2 \perp \nu_3$.
- By the way, what is a curve? And what does orthogonal mean?
Question: How can we compute a rotation about a line $\ell$ (through the origin) with direction vector $\nu$ by an angle $\phi$?
Rotation About a Line

- **Question:** How can we compute a rotation about a line $\ell$ (through the origin) with direction vector $\nu$ by an angle $\phi$?

- **Answer:** We set up a new frame $C'$ and reduce the rotation about $\ell$ to a rotation about a coordinate axis.
Question: What is an important topological difference between the following sets?
Question: What is an important topological difference between the following sets?

- not path-connected
- path-connected, multiply-connected
Consider

\[ \sum_{i=1}^{n} \frac{1}{i} \]

for some \( n \in \mathbb{N} \).

Question: How shall we compute this sum on a computer? In particular, does it matter whether we start summing with the smallest or the largest summand?

Answer: Yes, it does matter! We'll get back to this question when we talk about floating-point arithmetic and numerical issues.
Consider
\[ \sum_{i=1}^{n} \frac{1}{i} \]
for some \( n \in \mathbb{N} \).

Question: How shall be compute this sum on a computer? In particular, does it matter whether we start summing with the smallest or the largest summand?

\[ 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1} + \frac{1}{n} \overset{?}{=} \frac{1}{n} + \frac{1}{n-1} + \ldots + \frac{1}{3} + \frac{1}{2} + 1 \]
Consider
\[ \sum_{i=1}^{n} \frac{1}{i} \]
for some \( n \in \mathbb{N} \).

Question: How shall be compute this sum on a computer? In particular, does it matter whether we start summing with the smallest or the largest summand?

\[
1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1} + \frac{1}{n} \overset{?}{=} \frac{1}{n} + \frac{1}{n-1} + \ldots + \frac{1}{3} + \frac{1}{2} + 1
\]

Answer: Yes, it does matter! We’ll get back to this question when we’ll talk about floating-point arithmetic and numerical issues.
Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in 3D).
Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in 3D).

Consider the tetrahedron $T$ formed by $p_1, p_2, p_3, p_4$. 

Answer: The probability is $1/4$ in 2D and $1/8$ in 3D.
Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in 3D).

Consider the tetrahedron $T$ formed by $p_1, p_2, p_3, p_4$.

What is the probability that the center of the sphere lies inside $T$?

Answer: The probability is $\frac{1}{4}$ in 2D and $\frac{1}{8}$ in 3D.
Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in 3D).

Consider the tetrahedron $T$ formed by $p_1, p_2, p_3, p_4$.

What is the probability that the center of the sphere lies inside $T$?

Visualization of that problem in 2D (for three random points on a circle):

![Visualization](image.png)
Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in 3D).

Consider the tetrahedron $T$ formed by $p_1, p_2, p_3, p_4$.

What is the probability that the center of the sphere lies inside $T$?

Visualization of that problem in 2D (for three random points on a circle):
Choose four points \( p_1, p_2, p_3, p_4 \) independently at random (relative to a uniform distribution) on a sphere (in 3D).

Consider the tretrahedron \( T \) formed by \( p_1, p_2, p_3, p_4 \).

What is the probability that the center of the sphere lies inside \( T \)?

Visualization of that problem in 2D (for three random points on a circle):
Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in 3D).

Consider the tetrahedron $T$ formed by $p_1, p_2, p_3, p_4$.

What is the probability that the center of the sphere lies inside $T$?

Visualization of that problem in 2D (for three random points on a circle):
Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in 3D).

- Consider the tetrahedron $T$ formed by $p_1, p_2, p_3, p_4$.
- What is the probability that the center of the sphere lies inside $T$?

Visualization of that problem in 2D (for three random points on a circle):
Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in 3D).

Consider the tetrahedron $T$ formed by $p_1, p_2, p_3, p_4$.

What is the probability that the center of the sphere lies inside $T$?

Answer: The probability is $1/4$ in 2D and $1/8$ in 3D.

Visualization of that problem in 2D (for three random points on a circle):
Consider a mountain that is shaped like a right circular cone.
Consider a mountain that is shaped like a right circular cone.

A shortest-length railroad track is supposed to start at $A$, wind around the mountain once, and end in $B$.

The height $h$ of the cone is $40\sqrt{2}$, its base radius $r$ is 20, and the distance between $A$ and $B$ is 10.

Your task:
1. Prove that the shortest-length railroad track from $A$ to $B$ that winds around the mountain once consists of an uphill portion and of a downhill portion.
2. Compute the length of the downhill portion.
Consider a mountain that is shaped like a right circular cone.

A shortest-length railroad track is supposed to start at $A$, wind around the mountain once, and end in $B$.

The height $h$ of the cone is $40\sqrt{2}$, its base radius $r$ is 20, and the distance between $A$ and $B$ is 10.
Consider a mountain that is shaped like a right circular cone.

A shortest-length railroad track is supposed to start at \( A \), wind around the mountain once, and end in \( B \).

The height \( h \) of the cone is \( 40\sqrt{2} \), its base radius \( r \) is 20, and the distance between \( A \) and \( B \) is 10.

Your task:

1. Prove that the shortest-length railroad track from \( A \) to \( B \) that winds around the mountain once consists of an uphill portion and of a downhill portion.

2. Compute the length of the downhill portion.

[Problem credit: Presh Talwalkar’s “Mind Your Decisions” YouTube Channel.]
Another Challenge Problem

- Consider an equilateral triangle and pick a random point $P$ strictly in its interior.
Another Challenge Problem

- Consider an equilateral triangle and pick a random point \( P \) strictly in its interior.
- Draw a straight-line segment from each vertex to \( P \).
Another Challenge Problem

- Consider an equilateral triangle and pick a random point $P$ strictly in its interior.
- Draw a straight-line segment from each vertex to $P$.
- Your task:
  1. Prove that these three line segments form a new triangle if rotated and translated properly.

[Problem credit: Tanya Khovanova's "Math coffin problems"].
Another Challenge Problem

- Consider an equilateral triangle and pick a random point \( P \) strictly in its interior.
- Draw a straight-line segment from each vertex to \( P \).
- Your task:
  1. Prove that these three line segments form a new triangle if rotated and translated properly.
  2. Choose any two of the three angles at \( P \) induced by these line segments, say \( \alpha \) and \( \beta \), and assume that they are known. What are the new triangle’s three interior angles in terms of \( \alpha \) and \( \beta \)?

[Problem credit: Tanya Khovanova’s “Math coffin problems”.]
1 Introduction

- Motivation
- Notation
The set \{1, 2, 3, \ldots\} of natural numbers is denoted by \( \mathbb{N} \), with \( \mathbb{N}_0 := \mathbb{N} \cup \{0\} \), while \( \mathbb{Z} \) denotes the integers (positive and negative) and \( \mathbb{R} \) the reals. The non-negative reals are denoted by \( \mathbb{R}_0^+ \), and the positive reals by \( \mathbb{R}^+ \).
Notation

- The set \( \{1, 2, 3, \ldots\} \) of natural numbers is denoted by \( \mathbb{N} \), with \( \mathbb{N}_0 := \mathbb{N} \cup \{0\} \), while \( \mathbb{Z} \) denotes the integers (positive and negative) and \( \mathbb{R} \) the reals. The non-negative reals are denoted by \( \mathbb{R}_0^+ \), and the positive reals by \( \mathbb{R}^+ \).

- Open or closed intervals \( I \subset \mathbb{R} \) are denoted using square brackets: e.g., \( I_1 = [a_1, b_1] \) or \( I_2 = [a_2, b_2] \), with \( a_1, a_2, b_1, b_2 \in \mathbb{R} \), where the right-hand “[” indicates that the value \( b_2 \) is not included in \( I_2 \).
Notation

- The set \{1, 2, 3, \ldots\} of natural numbers is denoted by \( \mathbb{N} \), with \( \mathbb{N}_0 := \mathbb{N} \cup \{0\} \), while \( \mathbb{Z} \) denotes the integers (positive and negative) and \( \mathbb{R} \) the reals. The non-negative reals are denoted by \( \mathbb{R}_0^+ \), and the positive reals by \( \mathbb{R}^+ \).
- Open or closed intervals \( I \subset \mathbb{R} \) are denoted using square brackets: e.g., \( I_1 = [a_1, b_1] \) or \( I_2 = [a_2, b_2] \), with \( a_1, a_2, b_1, b_2 \in \mathbb{R} \), where the right-hand "[" indicates that the value \( b_2 \) is not included in \( I_2 \).
- We use Greek letters like \( \lambda, \mu \) and letters in italics to denote scalar values: \( s, t \).
Notation

- The set \( \{1, 2, 3, \ldots \} \) of natural numbers is denoted by \( \mathbb{N} \), with \( \mathbb{N}_0 := \mathbb{N} \cup \{0\} \), while \( \mathbb{Z} \) denotes the integers (positive and negative) and \( \mathbb{R} \) the reals. The non-negative reals are denoted by \( \mathbb{R}_0^+ \), and the positive reals by \( \mathbb{R}^+ \).

- Open or closed intervals \( I \subset \mathbb{R} \) are denoted using square brackets: e.g., \( I_1 = [a_1, b_1] \) or \( I_2 = [a_2, b_2[ \), with \( a_1, a_2, b_1, b_2 \in \mathbb{R} \), where the right-hand “[” indicates that the value \( b_2 \) is not included in \( I_2 \).

- We use Greek letters like \( \lambda, \mu \) and letters in italics to denote scalar values: \( s, t \).

- Points are denoted by capital or lower-case letters written in italics: \( A \) or \( a \).

- We use lower-case letters for denoting vectors, including position vectors of points. (Frequently we do not distinguish between a point and its position vector.)

- The coordinates of a vector are denoted by using indices (or numbers): e.g., \( a = (a_x, a_y, a_z) \), or \( a = (a_1, a_2, \ldots, a_n) \).

- In order to state \( a \in \mathbb{R}^n \) in vector form we will mix column and row vectors freely unless a specific form is required, such as for matrix multiplication.
The term $ab$ denotes the vector from the point $A$ to the point $B$. That is, $ab := b - a$. 
Notation

- The term $ab$ denotes the vector from the point $A$ to the point $B$. That is, $ab := b - a$.
- The dot product of two vectors $a$ and $b$ is denoted by $\langle a, b \rangle$.
- The vector cross-product is denoted by a cross: $a \times b$.
- The length of a vector $a$ is denoted by $\|a\|$.
- If $a$ and $b$ are perpendicular then we will write $a \perp b$.
Notation

- The term $ab$ denotes the vector from the point $A$ to the point $B$. That is, $ab := b - a$.
- The dot product of two vectors $a$ and $b$ is denoted by $\langle a, b \rangle$.
- The vector cross-product is denoted by a cross: $a \times b$.
- The length of a vector $a$ is denoted by $\|a\|$.
- If $a$ and $b$ are perpendicular then we will write $a \perp b$.
- The straight-line segment between the points $A$ and $B$ is denoted by $\overline{AB}$.
Notation

- The term $ab$ denotes the vector from the point $A$ to the point $B$. That is, $ab := b - a$.
- The dot product of two vectors $a$ and $b$ is denoted by $\langle a, b \rangle$.
- The vector cross-product is denoted by a cross: $a \times b$.
- The length of a vector $a$ is denoted by $\|a\|$.
- If $a$ and $b$ are perpendicular then we will write $a \perp b$.
- The straight-line segment between the points $A$ and $B$ is denoted by $\overline{AB}$.
- Bold capital letters, such as $\mathbf{M}$, are reserved for matrices.
Notation

- The term $ab$ denotes the vector from the point $A$ to the point $B$. That is, $ab := b - a$.
- The dot product of two vectors $a$ and $b$ is denoted by $\langle a, b \rangle$.
- The vector cross-product is denoted by a cross: $a \times b$.
- The length of a vector $a$ is denoted by $\|a\|$.
- If $a$ and $b$ are perpendicular then we will write $a \perp b$.
- The straight-line segment between the points $A$ and $B$ is denoted by $\overline{AB}$.
- Bold capital letters, such as $\mathbf{M}$, are reserved for matrices.
- The set of all elements $x \in S$ with property $P(x)$, for some set $S$ and some predicate $P$, is denoted by

\[
\{ x \in S : P(x) \} \quad \text{or} \quad \{ x : x \in S \land P(x) \}
\]

or

\[
\{ x \in S \mid P(x) \} \quad \text{or} \quad \{ x | x \in S \land P(x) \}.
\]
The term $ab$ denotes the vector from the point $A$ to the point $B$. That is, $ab := b - a$.

The dot product of two vectors $a$ and $b$ is denoted by $\langle a, b \rangle$.

The vector cross-product is denoted by a cross: $a \times b$.

The length of a vector $a$ is denoted by $\|a\|$.

If $a$ and $b$ are perpendicular then we will write $a \perp b$.

The straight-line segment between the points $A$ and $B$ is denoted by $\overline{AB}$.

Bold capital letters, such as $M$, are reserved for matrices.

The set of all elements $x \in S$ with property $P(x)$, for some set $S$ and some predicate $P$, is denoted by

$$\{x \in S : P(x)\} \quad \text{or} \quad \{x : x \in S \land P(x)\}$$

or

$$\{x \in S | P(x)\} \quad \text{or} \quad \{x|x \in S \land P(x)\}.$$
2 Algebraic Concepts

- Algebraic Structures
- Real Numbers and Vector Space $\mathbb{R}^n$
- Complex Numbers $\mathbb{C}$
- Polynomials
Algebraic Concepts
- Algebraic Structures
  - Vector Space
  - Basis
- Real Numbers and Vector Space $\mathbb{R}^n$
- Complex Numbers $\mathbb{C}$
- Polynomials
Definition 1 (Vector space, Dt.: Vektorraum)

A set $V$ together with an “addition” $+: V \times V \rightarrow V$ and a scalar “multiplication” $\cdot: F \times V \rightarrow V$ defines a vector space over a field $F$ (with multiplicative neutral element 1) if the following conditions hold:

1. $(V, +)$ is an Abelian group.
2. Distributivity: $\lambda (a + b) = \lambda \cdot a + \lambda \cdot b$ for all $\lambda \in F$, $a, b \in V$.
3. Distributivity: $(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a$ for all $\lambda, \mu \in F$, $a \in V$.
4. Associativity: $\lambda \cdot (\mu \cdot a) = (\lambda \mu) \cdot a$ for all $\lambda, \mu \in F$, $a \in V$.
5. Neutral element: $1 \cdot a = a$ for all $a \in V$.

The multiplication sign is often dropped if the meaning is clear within a specific context: $\lambda a$ rather than $\lambda \cdot a$. 
**Vector Space**

**Definition 1 (Vector space, Dt.: Vektorraum)**

A set \( V \) together with an “addition” \( + : V \times V \to V \) and a scalar “multiplication” \( \cdot : F \times V \to V \) defines a vector space over a field \( F \) (with multiplicative neutral element 1) if the following conditions hold:

1. \( (V, +) \) is an Abelian group.
Definition 1 (Vector space, Dt.: Vektorraum)

A set $V$ together with an “addition” $+: V \times V \to V$ and a scalar “multiplication” $\cdot: F \times V \to V$ defines a vector space over a field $F$ (with multiplicative neutral element 1) if the following conditions hold:

1. $(V, +)$ is an Abelian group.
2. Distributivity: $\lambda(a + b) = \lambda \cdot a + \lambda \cdot b \quad \forall \lambda \in F, \forall a, b \in V$.
3. Distributivity: $(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a \quad \forall \lambda, \mu \in F, \forall a \in V$. 

The multiplication sign is often dropped if the meaning is clear within a specific context: $\lambda a$ rather than $\lambda \cdot a$. 

© M. Held (Univ. Salzburg)
Vector Space

**Definition 1 (Vector space, Dt.: Vektorraum)**

A set $V$ together with an “addition” $+: V \times V \rightarrow V$ and a scalar “multiplication” $\cdot: F \times V \rightarrow V$ defines a **vector space** over a field $F$ (with multiplicative neutral element $1$) if the following conditions hold:

1. $(V, +)$ is an Abelian group.
2. Distributivity: $\lambda(a + b) = \lambda \cdot a + \lambda \cdot b \quad \forall \lambda \in F, \forall a, b \in V.$
3. Distributivity: $(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a \quad \forall \lambda, \mu \in F, \forall a \in V.$
4. Associativity: $\lambda \cdot (\mu \cdot a) = (\lambda \mu) \cdot a \quad \forall \lambda, \mu \in F, \forall a \in V.$
Vector Space

Definition 1 (Vector space, Dt.: Vektorraum)

A set $V$ together with an “addition” $+: V \times V \rightarrow V$ and a scalar “multiplication” $\cdot: F \times V \rightarrow V$ defines a vector space over a field $F$ (with multiplicative neutral element 1) if the following conditions hold:

1. $(V, +)$ is an Abelian group.
2. Distributivity: $\lambda(a + b) = \lambda \cdot a + \lambda \cdot b \quad \forall \lambda \in F, \forall a, b \in V$.
3. Distributivity: $(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a \quad \forall \lambda, \mu \in F, \forall a \in V$.
4. Associativity: $\lambda \cdot (\mu \cdot a) = (\lambda \mu) \cdot a \quad \forall \lambda, \mu \in F, \forall a \in V$.
5. Neutral element: $1 \cdot a = a \quad \forall a \in V$. 

The multiplication sign is often dropped if the meaning is clear within a specific context: $\lambda a$ rather than $\lambda \cdot a$. 
A set $V$ together with an “addition” $+: V \times V \to V$ and a scalar “multiplication” $\cdot: F \times V \to V$ defines a vector space over a field $F$ (with multiplicative neutral element $1$) if the following conditions hold:

1. $(V, +)$ is an Abelian group.
2. Distributivity: $\lambda(a + b) = \lambda \cdot a + \lambda \cdot b \quad \forall \lambda \in F, \forall a, b \in V$.
3. Distributivity: $(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a \quad \forall \lambda, \mu \in F, \forall a \in V$.
4. Associativity: $\lambda \cdot (\mu \cdot a) = (\lambda \mu) \cdot a \quad \forall \lambda, \mu \in F, \forall a \in V$.
5. Neutral element: $1 \cdot a = a \quad \forall a \in V$.

The multiplication sign is often dropped if the meaning is clear within a specific context: $\lambda a$ rather than $\lambda \cdot a$. 
Vector Space $F^n$

**Definition 2 (Cartesian product, Dt.: Mengenprodukt, kartesisches Produkt)**

For a field $F$ and $n \in \mathbb{N}$, we define

$$F^n := F \times F \times \cdots \times F := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \ldots, x_n \in F \right\}.$$

Well-known sample: $\mathbb{R}^n$, i.e., $F := \mathbb{R}$. You may find it convenient to "visualize" $F^n$ as $\mathbb{R}^n$. It is trivial to generalize this definition to $F^1 \times F^2 \times \cdots \times F^n$. 

© M. Held (Univ. Salzburg)

Geometrisches Rechnen (WS 2020/21)
**Definition 2 (Cartesian product, Dt.: Mengenprodukt, kartesisches Produkt)**

For a field $F$ and $n \in \mathbb{N}$, we define

\[ F^n := F \times F \times \cdots \times F := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \ldots, x_n \in F \right\}. \]

- Well-known sample: $\mathbb{R}^n$, i.e., $F := \mathbb{R}$. You may find it convenient to “visualize” $F^n$ as $\mathbb{R}^n$. 
Vector Space $F^n$

**Definition 2 (Cartesian product, Dt.: Mengenprodukt, kartesisches Produkt)**

For a field $F$ and $n \in \mathbb{N}$, we define

$$F^n := F \times F \times \cdots \times F := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \ldots, x_n \in F \right\}.$$

- Well-known sample: $\mathbb{R}^n$, i.e., $F := \mathbb{R}$. You may find it convenient to “visualize” $F^n$ as $\mathbb{R}^n$.
- It is trivial to generalize this definition to $F_1 \times F_2 \times \cdots \times F_n$. 
Vector Space $F^n$

**Definition 3**

Let $F$ be a field. For $a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n$ and $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in F^n$, we use $\begin{pmatrix} -a_1 \\ \vdots \\ -a_n \end{pmatrix}$ as the additive inverse $-a$.
Definition 3

Let $F$ be a field. For $a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n$ and $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in F^n$, we use $\begin{pmatrix} -a_1 \\ \vdots \\ -a_n \end{pmatrix}$ as the additive inverse $-a$. Furthermore, we use $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ as zero vector $0$ of $F^n$. 

Theorem 4

Let $F$ be a field. Then $F^n$ with addition and scalar multiplication as defined above constitutes a vector space over $F$. 
Vector Space $F^n$

**Definition 3**

Let $F$ be a field. For $a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n$ and $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in F^n$, we use $\begin{pmatrix} -a_1 \\ \vdots \\ -a_n \end{pmatrix}$ as the additive inverse $-a$. Furthermore, we use $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ as zero vector $0$ of $F^n$, and define the multiplication of $a$ by a scalar $\lambda \in F$ and the addition of $a$ and $b$ as follows:

$$\lambda \cdot a := \lambda a := \begin{pmatrix} \lambda \cdot a_1 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix} \quad a + b := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$
**Vector Space $F^n$**

**Definition 3**

Let $F$ be a field. For $a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n$ and $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in F^n$, we use $\begin{pmatrix} -a_1 \\ \vdots \\ -a_n \end{pmatrix}$ as the additive inverse $-a$. Furthermore, we use $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ as zero vector $0$ of $F^n$, and define the multiplication of $a$ by a scalar $\lambda \in F$ and the addition of $a$ and $b$ as follows:

$$\lambda \cdot a := \lambda a := \begin{pmatrix} \lambda \cdot a_1 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix} \quad \text{and} \quad a + b := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

**Theorem 4**

Let $F$ be a field. Then $F^n$ with addition and scalar multiplication as defined above constitutes a vector space over $F$. 
"Exotic" Vector Spaces: Functions, Sequences

Lemma 5
The set of all real-valued functions $f : \mathbb{R} \to \mathbb{R}$ forms a vector space over $\mathbb{R}$. 

Caveats:
Subsets of functions characterized by an additional property — e.g., positive, not continuous — need not form a vector space.
Subsets of sequences characterized by an additional property — e.g., divergent sequences, monotonic sequences — need not form a vector space!
“Exotic” Vector Spaces: Functions, Sequences

Lemma 5
The set of all real-valued functions $f : \mathbb{R} \rightarrow \mathbb{R}$ forms a vector space over $\mathbb{R}$.

Lemma 6
The set of all infinite sequences $(t_n)_{n \in \mathbb{N}}$ of real numbers forms a vector space over $\mathbb{R}$.

Caveats:
Subsets of functions characterized by an additional property — e.g., positive, not continuous — need not form a vector space.
Subsets of sequences characterized by an additional property — e.g., divergent sequences, monotonic sequences — need not form a vector space!
"Exotic" Vector Spaces: Functions, Sequences

Lemma 5
The set of all real-valued functions $f : \mathbb{R} \to \mathbb{R}$ forms a vector space over $\mathbb{R}$.

Lemma 6
The set of all infinite sequences $(t_n)_{n \in \mathbb{N}}$ of real numbers forms a vector space over $\mathbb{R}$.

Caveats:
- Subsets of functions characterized by an additional property — e.g., positive, not continuous — need not form a vector space.
- Subsets of sequences characterized by an additional property — e.g., divergent sequences, monotonic sequences — need not form a vector space!
Definition 7 (Subspace, Dt.: Teilraum, Unterraum)

A subset $S$ of a vector space $V$ over a field $F$ is called a *subspace* of $V$ if

1. the zero vector belongs to $S$; i.e., $0 \in S$;
2. $\forall a, b \in S$ $a + b \in S$ ($S$ is said to be closed under vector addition);
3. $\forall a \in S \forall \lambda \in F \lambda a \in S$ ($S$ is said to be closed under scalar multiplication).

Lemma 8

The set of all continuous (real-valued) functions $f : \mathbb{R} \to \mathbb{R}$ and the set of all linear functions form subspaces of the vector space of all (real-valued) functions.
Subspace

Definition 7 (Subspace, Dt.: Teilraum, Unterraum)

A subset $S$ of a vector space $V$ over a field $F$ is called a **subspace** of $V$ if

1. the zero vector belongs to $S$; i.e., $0 \in S$;
Subspace

Definition 7 (Subspace, Dt.: Teilraum, Unterraum)

A subset $S$ of a vector space $V$ over a field $F$ is called a \textit{subspace} of $V$ if

1. the zero vector belongs to $S$; i.e., $0 \in S$;
2. $\forall a, b \in S \quad a + b \in S$ (S is said to be closed under vector addition);
Subspace

Definition 7 (Subspace, Dt.: Teilraum, Unterraum)

A subset $S$ of a vector space $V$ over a field $F$ is called a *subspace* of $V$ if

1. the zero vector belongs to $S$; i.e., $0 \in S$;
2. $\forall a, b \in S \quad a + b \in S$ (S is said to be closed under vector addition);
3. $\forall a \in S \quad \forall \lambda \in F \quad \lambda a \in S$ (S is said to be closed under scalar multiplication).
**Subspace**

**Definition 7 (Subspace, Dt.: Teilraum, Unterraum)**

A subset $S$ of a vector space $V$ over a field $F$ is called a subpace of $V$ if

1. the zero vector belongs to $S$; i.e., $0 \in S$;
2. $\forall a, b \in S \quad a + b \in S$ (S is said to be closed under vector addition);
3. $\forall a \in S \quad \forall \lambda \in F \quad \lambda a \in S$ (S is said to be closed under scalar multiplication).

**Lemma 8**

The set of all continuous (real-valued) functions $f : \mathbb{R} \to \mathbb{R}$ and the set of all linear functions form subspaces of the vector space of all (real-valued) functions.
Definition 9 (Linear combination, Dt.: Linearkombination)

Let $V$ be a vector space over $F$, and $\nu_1, \ldots, \nu_k \in V$ and $\lambda_1, \ldots, \lambda_k \in F$, for some $k \in \mathbb{N}$. The vector

$$a = \lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k$$

is called a linear combination of the vectors $\nu_1, \ldots, \nu_k$. 
Definition 9 (Linear combination, Dt.: Linearkombination)

Let $V$ be a vector space over $F$, and $\nu_1, \ldots, \nu_k \in V$ and $\lambda_1, \ldots, \lambda_k \in F$, for some $k \in \mathbb{N}$. The vector

$$a = \lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k$$

is called a \textit{linear combination} of the vectors $\nu_1, \ldots, \nu_k$.

Definition 10 (Linear hull, Dt.: lineare Hülle)

For $S \subseteq V$, with $V$ being a vector space over $F$,

$$[S] := \{\lambda_1 \nu_1 + \cdots + \lambda_k \nu_k : k \in \mathbb{N}, \nu_1, \ldots, \nu_k \in S, \lambda_1, \ldots, \lambda_k \in F\}$$

forms the \textit{linear hull} of $S$. 
Definition 9 (Linear combination, Dt.: Linearkombination)

Let $V$ be a vector space over $F$, and $\nu_1, \ldots, \nu_k \in V$ and $\lambda_1, \ldots, \lambda_k \in F$, for some $k \in \mathbb{N}$. The vector

$$a = \lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k$$

is called a *linear combination* of the vectors $\nu_1, \ldots, \nu_k$.

Definition 10 (Linear hull, Dt.: lineare Hülle)

For $S \subseteq V$, with $V$ being a vector space over $F$,

$$[S] := \{\lambda_1 \nu_1 + \cdots + \lambda_k \nu_k : k \in \mathbb{N}, \nu_1, \ldots, \nu_k \in S, \lambda_1, \ldots, \lambda_k \in F\}$$

forms the *linear hull* of $S$.

- Note: Any linear combination is formed by a finite number of vectors, even if we are allowed to pick those vectors from an infinite set!
Definition 9 (Linear combination, Dt.: Linearkombination)

Let \( V \) be a vector space over \( F \), and \( \nu_1, \ldots, \nu_k \in V \) and \( \lambda_1, \ldots, \lambda_k \in F \), for some \( k \in \mathbb{N} \). The vector

\[
a = \lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k
\]

is called a linear combination of the vectors \( \nu_1, \ldots, \nu_k \).

Definition 10 (Linear hull, Dt.: lineare Hülle)

For \( S \subseteq V \), with \( V \) being a vector space over \( F \),

\[
[S] := \{ \lambda_1 \nu_1 + \cdots + \lambda_k \nu_k : k \in \mathbb{N}, \nu_1, \ldots, \nu_k \in S, \lambda_1, \ldots, \lambda_k \in F \}
\]

forms the linear hull of \( S \).

- Note: Any linear combination is formed by a finite number of vectors, even if we are allowed to pick those vectors from an infinite set!

Lemma 11

For \( S \subseteq V \), with \( S \neq \emptyset \), the linear hull \([S]\) forms a subspace of the vector space \( V \).
Definition 12 (Linear independence, Dt.: lineare Unabhängigkeit)

The vectors \( \nu_1, \nu_2, \ldots, \nu_k \) of a vector space \( V \) over \( F \) are \textit{linearly dependent} if there exist scalars \( \lambda_1, \ldots, \lambda_k \in F \), not all zero, such that

\[
\lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k = 0.
\]
Linear Independence

Definition 12 (Linear independence, Dt.: lineare Unabhängigkeit)

The vectors $\nu_1, \nu_2, \ldots, \nu_k$ of a vector space $V$ over $F$ are linearly dependent if there exist scalars $\lambda_1, \ldots, \lambda_k \in F$, not all zero, such that

$$\lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k = 0.$$ 

Otherwise, the vectors $\nu_1, \nu_2, \ldots, \nu_k$ are linearly independent.
**Linear Independence**

**Definition 12 (Linear independence, Dt.: lineare Unabhängigkeit)**

The vectors $\nu_1, \nu_2, \ldots, \nu_k$ of a vector space $V$ over $F$ are *linearly dependent* if there exist scalars $\lambda_1, \ldots, \lambda_k \in F$, not all zero, such that

$$\lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k = 0.$$

Otherwise, the vectors $\nu_1, \nu_2, \ldots, \nu_k$ are *linearly independent*.

**Lemma 13**

If the vectors $\nu_1, \nu_2, \ldots, \nu_k$ of a vector space $V$ are linearly independent then

$$\lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k = 0 \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0.$$
Definition 12 (Linear independence, Dt.: lineare Unabhängigkeit)

The vectors \( \nu_1, \nu_2, \ldots, \nu_k \) of a vector space \( V \) over \( F \) are **linearly dependent** if there exist scalars \( \lambda_1, \ldots, \lambda_k \in F \), not all zero, such that

\[
\lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k = 0.
\]

Otherwise, the vectors \( \nu_1, \nu_2, \ldots, \nu_k \) are **linearly independent**.

Lemma 13

If the vectors \( \nu_1, \nu_2, \ldots, \nu_k \) of a vector space \( V \) are linearly independent then

\[
\lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_k \nu_k = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0.
\]

Lemma 14

The vectors \( \nu_1, \nu_2, \ldots, \nu_k \) of a vector space \( V \) are linearly independent if and only if none of them can be expressed as a linear combination of the other vectors.
Definition 15 (Basis)

The vectors $\nu_1, \nu_2, \ldots, \nu_n \in V$ form a basis of the vector space $V$ over $F$ if

1. $\nu_1, \ldots, \nu_n$ are linearly independent;
2. $\{\nu_1, \ldots, \nu_n\} = V$. 

Definition 16 (Finite dimension)

The vector space $V$ is said to have finite dimension if there exists a basis of $V$ that has finitely many vectors.

Theorem 17

Every basis of a finite vector space has the same number of basis vectors.

The number $n$ of vectors of a basis is called the dimension of the vector space.

Theorem 18

If $\nu_1, \ldots, \nu_n$ form a basis for $V$ over $F$ then for all $a \in V$ exist uniquely determined $\lambda_1, \ldots, \lambda_n \in F$ such that $a = \lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_n \nu_n$. 

Definition 15 (Basis)

The vectors $\nu_1, \nu_2, \ldots, \nu_n \in V$ form a *basis* of the vector space $V$ over $F$ if

1. $\nu_1, \ldots, \nu_n$ are linearly independent;
2. $[\{\nu_1, \ldots, \nu_n\}] = V$. 

Definition 16 (Finite dimension)

The vector space $V$ is said to have *finite dimension* if there exists a basis of $V$ that has finitely many vectors.

Theorem 17

Every basis of a finite vector space has the same number of basis vectors.

The number $n$ of vectors of a basis is called the *dimension* of the vector space.

Theorem 18

If $\nu_1, \ldots, \nu_n$ form a basis for $V$ over $F$ then for all $a \in V$ exist uniquely determined $\lambda_1, \ldots, \lambda_n \in F$ such that

$$a = \lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_n \nu_n.$$
Basis of a Vector Space

Definition 15 (Basis)
The vectors $\nu_1, \nu_2, \ldots, \nu_n \in V$ form a basis of the vector space $V$ over $F$ if

1. $\nu_1, \ldots, \nu_n$ are linearly independent;
2. $[\{\nu_1, \ldots, \nu_n\}] = V$.

Definition 16 (Finite dimension)
The vector space $V$ is said to have finite dimension if there exists a basis of $V$ that has finitely many vectors.
Basis of a Vector Space

**Definition 15 (Basis)**
The vectors $\nu_1, \nu_2, \ldots, \nu_n \in V$ form a *basis* of the vector space $V$ over $F$ if

1. $\nu_1, \ldots, \nu_n$ are linearly independent;
2. $[\{\nu_1, \ldots, \nu_n\}] = V$.

**Definition 16 (Finite dimension)**
The vector space $V$ is said to have *finite dimension* if there exists a basis of $V$ that has finitely many vectors.

**Theorem 17**
Every basis of a finite vector space has the same number of basis vectors.
Basis of a Vector Space

Definition 15 (Basis)
The vectors \( \nu_1, \nu_2, \ldots, \nu_n \in V \) form a basis of the vector space \( V \) over \( F \) if
1. \( \nu_1, \ldots, \nu_n \) are linearly independent;
2. \( \{\nu_1, \ldots, \nu_n\} = V \).

Definition 16 (Finite dimension)
The vector space \( V \) is said to have finite dimension if there exists a basis of \( V \) that has finitely many vectors.

Theorem 17
Every basis of a finite vector space has the same number of basis vectors.

- The number \( n \) of vectors of a basis is called the dimension of the vector space.
Basis of a Vector Space

**Definition 15 (Basis)**
The vectors $\nu_1, \nu_2, \ldots, \nu_n \in V$ form a basis of the vector space $V$ over $F$ if

1. $\nu_1, \ldots, \nu_n$ are linearly independent;
2. $[\{\nu_1, \ldots, \nu_n\}] = V$.

**Definition 16 (Finite dimension)**
The vector space $V$ is said to have finite dimension if there exists a basis of $V$ that has finitely many vectors.

**Theorem 17**
Every basis of a finite vector space has the same number of basis vectors.

- The number $n$ of vectors of a basis is called the dimension of the vector space.

**Theorem 18**
If $\nu_1, \ldots, \nu_n$ form a basis for $V$ over $F$ then for all $a \in V$ exist uniquely determined $\lambda_1, \ldots, \lambda_n \in F$ such that $a = \lambda_1 \nu_1 + \lambda_2 \nu_2 + \cdots + \lambda_n \nu_n$. 
Algebraic Concepts

- Algebraic Structures
- Real Numbers and Vector Space $\mathbb{R}^n$
  - Points and Vectors in $\mathbb{R}^n$
  - Canonical Basis
  - Standard Coordinate Systems
  - Convex Combinations and Convexity
- Complex Numbers $\mathbb{C}$
- Polynomials
A point is a location in a (vector) space. From a mathematical point of view it does not have any size or any other property besides its location.

A vector has a direction and a length as its main properties.

The position vector (Dt.: Ortsvektor) of a point is the vector that points from the origin of the space to the point.

It is common not to make a clean distinction between a point and its position vector.
A point is a location in a (vector) space. From a mathematical point of view it does not have any size or any other property besides its location.

A vector has a direction and a length as its main properties.

The position vector (Dt.: Ortsvektor) of a point is the vector that points from the origin of the space to the point.

It is common not to make a clean distinction between a point and its position vector.

Note that vectors can be regarded both as column matrices and as row matrices.

While it does not matter for most applications whether or not to specify a vector as a column or row matrix, there exist a few applications for which it does matter! (E.g., multiplication of a matrix and a vector.)

Thus, pay close attention to how vectors are treated when studying a textbook or using a graphics package.
Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$

- For a 2D vector $\mathbf{v}$, we may find it more convenient to use the symbols $x$ and $y$ for denoting its components, rather than $v_x$ and $v_y$:

$$\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{or} \quad \mathbf{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix}. $$

- Similarly for a 3D vector:

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{or} \quad \mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}. $$
Adding and subtracting two 2D vectors $a$ and $b$:

\[
\begin{align*}
  a + b &= \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix}
\end{align*}
\]
Adding and subtracting two 2D vectors $a$ and $b$:

$$a + b = \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix}$$
Vector Algebra

- Adding and subtracting two 2D vectors $\mathbf{a}$ and $\mathbf{b}$:

\[
\mathbf{a} + \mathbf{b} = \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix}
\]
Vector Algebra

- Adding and subtracting two 2D vectors $a$ and $b$:

$$a + b = \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix} \quad a - b := \begin{pmatrix} a_x - b_x \\ a_y - b_y \end{pmatrix}$$
Vector Algebra

- Adding and subtracting two 2D vectors \( \mathbf{a} \) and \( \mathbf{b} \):

\[
\mathbf{a} + \mathbf{b} = \left( a_x \right) + \left( b_x \right) := \left( a_x + b_x \right) \quad \mathbf{a} - \mathbf{b} := \left( a_x - b_x \right)
\]

\[
\mathbf{a} + \mathbf{b} = \left( a_x \right) + \left( b_y \right) := \left( a_y + b_y \right)
\]

\[
\mathbf{a} - \mathbf{b} := \left( a_y - b_y \right)
\]
Vector Algebra

- Adding and subtracting two 2D vectors $a$ and $b$:

\[
a + b = \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix} \quad a - b := \begin{pmatrix} a_x - b_x \\ a_y - b_y \end{pmatrix}
\]
Adding and subtracting two 2D vectors $a$ and $b$:

\[
\begin{align*}
    a + b &= \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix} \\
    a - b &= \begin{pmatrix} a_x - b_x \\ a_y - b_y \end{pmatrix}
\end{align*}
\]
Adding and subtracting two 2D vectors $a$ and $b$:

\[
\begin{align*}
a + b &= \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix} \\
a - b &= \begin{pmatrix} a_x \\ a_y \end{pmatrix} - \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x - b_x \\ a_y - b_y \end{pmatrix}
\end{align*}
\]
Vector Algebra

- Adding and subtracting two 2D vectors $a$ and $b$:

\[
a + b = \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} := \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix} \quad a - b := \begin{pmatrix} a_x - b_x \\ a_y - b_y \end{pmatrix}
\]

- Similarly for vectors in $\mathbb{R}^n$, for $n \geq 3$. 

[Diagram showing vector addition and subtraction in 2D space.]
Canonical Basis

In $\mathbb{R}^n$ we define the $n$ vectors

$$
e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \ldots, \quad e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n.
$$
Canonical Basis

- In $\mathbb{R}^n$ we define the $n$ vectors

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \ldots, \quad e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n.$$

- The vectors $e_1, \ldots, e_n$ are linearly independent since $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$ implies

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \text{i.e., } \lambda_1 = 0, \ldots, \lambda_n = 0.$$
Canonical Basis

- In $\mathbb{R}^n$ we define the $n$ vectors
  \[ e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \ldots, \quad e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n. \]

- The vectors $e_1, \ldots, e_n$ are linearly independent since $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$ implies
  \[ \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \text{ i.e., } \lambda_1 = 0, \ldots, \lambda_n = 0. \]

- Let $a \in \mathbb{R}^n$. We get
  \[ a := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \]
Canonical Basis

- In \( \mathbb{R}^n \) we define the \( n \) vectors
  \[
  e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \ldots, \quad e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n.
  \]

- The vectors \( e_1, \ldots, e_n \) are linearly independent since \( \lambda_1 e_1 + \cdots + \lambda_n e_n = 0 \) implies
  \[
  \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \text{i.e.,} \quad \lambda_1 = 0, \ldots, \lambda_n = 0.
  \]

- Let \( a \in \mathbb{R}^n \). We get
  \[
  a := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot e_n.
  \]
For $a \in \mathbb{R}^2$ we get $a = a_1 \cdot e_1 + a_2 \cdot e_2$.

E.g.:

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2e_1 + 3e_2$$

$$= 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
For $a \in \mathbb{R}^2$ we get $a = a_1 \cdot e_1 + a_2 \cdot e_2$.

E.g.:

$$
\begin{pmatrix}
2 \\
3
\end{pmatrix}
= 2e_1 + 3e_2
$$

$$
= 2 \begin{pmatrix}
1 \\
0
\end{pmatrix} + 3 \begin{pmatrix}
0 \\
1
\end{pmatrix}
$$
For $a \in \mathbb{R}^2$ we get $a = a_1 \cdot e_1 + a_2 \cdot e_2$.

E.g.:

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2e_1 + 3e_2$$

$$= 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
For \( a \in \mathbb{R}^2 \) we get \( a = a_1 \cdot e_1 + a_2 \cdot e_2 \).

E.g.:

\[
\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2e_1 + 3e_2
\]

\[
= 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\]

But this is not the only possible basis for \( \mathbb{R}^2 \).
Canonical Basis

For $a \in \mathbb{R}^2$ we get $a = a_1 \cdot e_1 + a_2 \cdot e_2$.

E.g.:

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2e_1 + 3e_2$$

$$= 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

But this is not the only possible basis for $\mathbb{R}^2$. E.g.:

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix}_{[e_1,e_2]} = 2v + w$$

$$= \begin{pmatrix} 2 \\ 1 \end{pmatrix}_{[v,w]}$$
Canonical Basis

For $a \in \mathbb{R}^2$ we get $a = a_1 \cdot e_1 + a_2 \cdot e_2$.

E.g.:

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2e_1 + 3e_2$$

$$= 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

But this is not the only possible basis for $\mathbb{R}^2$. E.g.:

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix}_{[e_1,e_2]} = 2v + w$$

$$= \begin{pmatrix} 2 \\ 1 \end{pmatrix}_{[v,w]}$$
Standard Coordinate Systems in $\mathbb{R}^2$ and $\mathbb{R}^3$

- Cartesian coordinates: $(a, b, c)$.
Standard Coordinate Systems in $\mathbb{R}^2$ and $\mathbb{R}^3$

- Cartesian coordinates: $(a, b, c)$.
- Polar coordinates (in $\mathbb{R}^2$): $(\rho, \alpha)$, with $\alpha \in [0, 2\pi]$. 

\[ (a, b, c) \]

\[ (a, b) \]

\[ \alpha \]

\[ \rho \]
Standard Coordinate Systems in $\mathbb{R}^2$ and $\mathbb{R}^3$

- Cartesian coordinates: $(a, b, c)$.
- Polar coordinates (in $\mathbb{R}^2$): $(\rho, \alpha)$, with $\alpha \in [0, 2\pi]$.
- Cylindrical coordinates: $(\rho, \alpha, c)$, with $\alpha \in [0, 2\pi]$. 
Standard Coordinate Systems in $\mathbb{R}^2$ and $\mathbb{R}^3$

- Cartesian coordinates: $(a, b, c)$.
- Polar coordinates (in $\mathbb{R}^2$): $(\rho, \alpha)$, with $\alpha \in [0, 2\pi]$.
- Cylindrical coordinates: $(\rho, \alpha, c)$, with $\alpha \in [0, 2\pi]$.
- Spherical coordinates: $(r, \alpha, \beta)$, with $\alpha \in [0, 2\pi]$ and $\beta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. 
Geographic Coordinates: Longitude and Latitude

- The z-axis of the coordinate system is aligned with the spin axis of the Earth, with the coordinate origin at the Earth’s center.
- The equator is defined as the intersection of the $xy$-plane ("fundamental plane") of this coordinate system with the Earth.
Geographic Coordinates: Longitude and Latitude

- The z-axis of the coordinate system is aligned with the spin axis of the Earth, with the coordinate origin at the Earth’s center.
- The equator is defined as the intersection of the $xy$-plane ("fundamental plane") of this coordinate system with the Earth.
- Two angles are measured from the center of the Earth: $latitude$ (Dt. “Breite”) measures the angle between any point and the equator.
Geographic Coordinates: Longitude and Latitude

- The z-axis of the coordinate system is aligned with the spin axis of the Earth, with the coordinate origin at the Earth’s center.
- The equator is defined as the intersection of the $xy$-plane (“fundamental plane”) of this coordinate system with the Earth.
- Two angles are measured from the center of the Earth: *latitude* (Dt. “Breite”) measures the angle between any point and the equator. The other angle, *longitude* (Dt. “Länge”), measures the angle along the equator from an arbitrary point on the earth. Greenwich, England, is the generally accepted zero-longitude point (Prime Meridian, Dt. “Nullmeridian”).
The z-axis of the coordinate system is aligned with the spin axis of the Earth, with the coordinate origin at the Earth’s center.

The equator is defined as the intersection of the $xy$-plane ("fundamental plane") of this coordinate system with the Earth.

Two angles are measured from the center of the Earth: $latitude$ (Dt. “Breite”) measures the angle between any point and the equator. The other angle, $longitude$ (Dt. “Länge”), measures the angle along the equator from an arbitrary point on the earth. Greenwich, England, is the generally accepted zero-longitude point (Prime Meridian, Dt. “Nullmeridian”).

A position on the Earth is specified as $\alpha$ degrees East or West, and $\beta$ degrees North or South. Thus, $\alpha \in [0, 180]$, and $\beta \in [0, 90]$. 

Lines of constant latitude are called parallels, with the equator having latitude 0. Lines of constant longitude are great circles that intersect at the poles and are called meridians. Hence, geographical coordinates are nothing but (a variant of) a spherical coordinate system.
Geographic Coordinates: Longitude and Latitude

- The z-axis of the coordinate system is aligned with the spin axis of the Earth, with the coordinate origin at the Earth’s center.
- The equator is defined as the intersection of the $xy$-plane (“fundamental plane”) of this coordinate system with the Earth.
- Two angles are measured from the center of the Earth: latitude (Dt. “Breite”) measures the angle between any point and the equator. The other angle, longitude (Dt. “Länge”), measures the angle along the equator from an arbitrary point on the earth. Greenwich, England, is the generally accepted zero-longitude point (Prime Meridian, Dt. “Nullmeridian”).
- A position on the Earth is specified as $\alpha$ degrees East or West, and $\beta$ degrees North or South. Thus, $\alpha \in [0, 180]$, and $\beta \in [0, 90]$.
- Lines of constant latitude are called parallels, with the equator having latitude 0.
- Lines of constant longitude are great circles that intersect at the poles and are called meridians.
- Hence, geographical coordinates are nothing but (a variant of) a spherical coordinate system.
Definition 19 (Affine combination, Dt.: Affinkombination)

Let $P_1, P_2, \ldots, P_k$ be $k$ points in $\mathbb{R}^n$. An affine combination of the points $P_1, \ldots, P_k$ is given by

$$\sum_{i=1}^{k} \lambda_i p_i \quad \text{with} \quad \sum_{i=1}^{k} \lambda_i = 1,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ are scalars.
Affine and Convex Combinations

**Definition 19 (Affine combination, Dt.: Affinkombination)**

Let $P_1, P_2, \ldots, P_k$ be $k$ points in $\mathbb{R}^n$. An **affine combination** of the points $P_1, \ldots, P_k$ is given by

$$
\sum_{i=1}^{k} \lambda_i p_i \quad \text{with} \quad \sum_{i=1}^{k} \lambda_i = 1,
$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ are scalars.

**Definition 20 (Convex combination, Dt.: Konvexkombination)**

Let $P_1, P_2, \ldots, P_k$ be $k$ points in $\mathbb{R}^n$. A **convex combination** of the points $P_1, \ldots, P_k$ is defined as

$$
\sum_{i=1}^{k} \lambda_i p_i \quad \text{with} \quad \sum_{i=1}^{k} \lambda_i = 1 \quad \text{and} \quad \forall (1 \leq i \leq k) \quad \lambda_i \geq 0,
$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ are scalars.
Definition 21 (Affine hull, Dt.: affine Hülle)

Let $P_1, P_2, \ldots, P_k$ be $k$ points in $\mathbb{R}^n$. The affine hull of the points $P_1, \ldots, P_k$ is the set

$$\left\{ \sum_{i=1}^{k} \lambda_i p_i : \lambda_1, \ldots, \lambda_k \in \mathbb{R} \text{ and } \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$
Definition 21 (Affine hull, Dt.: affine Hülle)

Let \( P_1, P_2, \ldots, P_k \) be \( k \) points in \( \mathbb{R}^n \). The affine hull of the points \( P_1, \ldots, P_k \) is the set

\[
\left\{ \sum_{i=1}^{k} \lambda_i p_i : \lambda_1, \ldots, \lambda_k \in \mathbb{R} \text{ and } \sum_{i=1}^{k} \lambda_i = 1 \right\}.
\]

For a set \( S \subseteq \mathbb{R}^n \) (with possibly infinitely many points), the affine hull of \( S \) is the set

\[
\left\{ \sum_{i=1}^{k} \lambda_i p_i : k \in \mathbb{N} \text{ and } P_1, P_2, \ldots, P_k \in S \text{ and } \lambda_1, \ldots, \lambda_k \in \mathbb{R} \text{ and } \sum_{i=1}^{k} \lambda_i = 1 \right\}.
\]
Definition 22 (Convex hull, Dt.: konvexe Hülle)

Let \( P_1, P_2, \ldots, P_k \) be \( k \) points in \( \mathbb{R}^n \). The **convex hull** of the points \( P_1, \ldots, P_k \) is the set

\[
\left\{ \sum_{i=1}^{k} \lambda_i p_i : \lambda_1, \ldots, \lambda_k \in \mathbb{R}_0^+ \text{ and } \sum_{i=1}^{k} \lambda_i = 1 \right\}.
\]
Definition 22 (Convex hull, Dt.: konvexe Hülle)

Let $P_1, P_2, \ldots, P_k$ be $k$ points in $\mathbb{R}^n$. The convex hull of the points $P_1, \ldots, P_k$ is the set

$$\left\{ \sum_{i=1}^{k} \lambda_i p_i : \lambda_1, \ldots, \lambda_k \in \mathbb{R}_0^+ \text{ and } \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$
Definition 22 (Convex hull, Dt.: konvexe Hülle)

Let $P_1, P_2, \ldots, P_k$ be $k$ points in $\mathbb{R}^n$. The convex hull of the points $P_1, \ldots, P_k$ is the set

$$\left\{ \sum_{i=1}^{k} \lambda_i p_i : \lambda_1, \ldots \lambda_k \in \mathbb{R}_0^+ \text{ and } \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

For a set $S \subseteq \mathbb{R}^n$ (with possibly infinitely many points), the convex hull of $S$ is the set

$$\left\{ \sum_{i=1}^{k} \lambda_i p_i : k \in \mathbb{N} \text{ and } P_1, P_2, \ldots, P_k \in S \text{ and } \lambda_1, \ldots \lambda_k \in \mathbb{R}_0^+ \text{ and } \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

The convex hull of $S$ is commonly denoted by $CH(S)$. 
Definition 23 (Convex set, Dt.: konvexe Menge)

A set \( S \subseteq \mathbb{R}^n \) is called \textit{convex} if for all \( P, Q \in S \)

\[
\overline{PQ} \subseteq S
\]

where \( \overline{PQ} \) denotes the straight-line segment between \( P \) and \( Q \).
**Definition 23 (Convex set, Dt.: konvexe Menge)**

A set \( S \subseteq \mathbb{R}^n \) is called *convex* if for all \( P, Q \in S \)

\[
\overline{PQ} \subseteq S
\]

where \( \overline{PQ} \) denotes the straight-line segment between \( P \) and \( Q \).

**Lemma 24**

For \( S \subseteq \mathbb{R}^n \), the convex hull \( CH(S) \) of \( S \) is a convex set.
Convexity

**Definition 25 (Convex superset)**

A set $B \subseteq \mathbb{R}^n$ is called a *convex superset* of a set $A \subseteq \mathbb{R}^n$ if

$$A \subseteq B \quad \text{and} \quad B \text{ is convex.}$$
Convexity

Definition 25 (Convex superset)

A set $B \subseteq \mathbb{R}^n$ is called a convex superset of a set $A \subseteq \mathbb{R}^n$ if $A \subseteq B$ and $B$ is convex.

Lemma 26

For $A \subseteq \mathbb{R}^n$, the following definitions are equivalent to Def. 22:

- $CH(A)$ is the smallest convex superset of $A$.
- $CH(A)$ is the intersection of all convex supersets of $A$. 

The definition of a convex hull (and of convexity) is readily extended from $\mathbb{R}^n$ to other vector spaces over $\mathbb{R}$. 

© M. Held (Univ. Salzburg)
Definition 25 (Convex superset)

A set $B \subseteq \mathbb{R}^n$ is called a convex superset of a set $A \subseteq \mathbb{R}^n$ if

$$A \subseteq B \quad \text{and} \quad B \text{ is convex.}$$

Lemma 26

For $A \subseteq \mathbb{R}^n$, the following definitions are equivalent to Def. 22:

- $CH(A)$ is the smallest convex superset of $A$.
- $CH(A)$ is the intersection of all convex supersets of $A$. 

The definition of a convex hull (and of convexity) is readily extended from $\mathbb{R}^n$ to other vector spaces over $\mathbb{R}$. 

© M. Held (Univ. Salzburg)
Convexity

**Definition 25 (Convex superset)**

A set $B \subseteq \mathbb{R}^n$ is called a *convex superset* of a set $A \subseteq \mathbb{R}^n$ if

$$A \subseteq B \quad \text{and} \quad B \text{ is convex.}$$

**Lemma 26**

For $A \subseteq \mathbb{R}^n$, the following definitions are equivalent to Def. 22:

- $CH(A)$ is the smallest convex superset of $A$.
- $CH(A)$ is the intersection of all convex supersets of $A$. 
Convexity

**Definition 25 (Convex superset)**

A set $B \subseteq \mathbb{R}^n$ is called a *convex superset* of a set $A \subseteq \mathbb{R}^n$ if

$A \subseteq B$ and $B$ is convex.

**Lemma 26**

For $A \subseteq \mathbb{R}^n$, the following definitions are equivalent to Def. 22:

- $CH(A)$ is the smallest convex superset of $A$.
- $CH(A)$ is the intersection of all convex supersets of $A$. 
Convexity

**Definition 25 (Convex superset)**

A set $B \subseteq \mathbb{R}^n$ is called a *convex superset* of a set $A \subseteq \mathbb{R}^n$ if

- $A \subseteq B$ and
- $B$ is convex.

**Lemma 26**

For $A \subseteq \mathbb{R}^n$, the following definitions are equivalent to Def. 22:

- $CH(A)$ is the smallest convex superset of $A$.
- $CH(A)$ is the intersection of all convex supersets of $A$.

The definition of a convex hull (and of convexity) is readily extended from $\mathbb{R}^n$ to other vector spaces over $\mathbb{R}$.
Algebraic Concepts

- Algebraic Structures
- Real Numbers and Vector Space $\mathbb{R}^n$
- Complex Numbers $\mathbb{C}$
  - Definition and Basics
  - Formulas by de Moivre and Euler
  - Mandelbrot and Julia
- Polynomials
Complex Numbers

Definition 27 (Complex numbers, Dt.: komplexe Zahlen)

The complex numbers, \( \mathbb{C} \), are formed by the set of ordered pairs of real numbers together with operations \( + : \mathbb{C} \times \mathbb{C} \to \mathbb{C} \) and \( \cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C} \) defined as follows:

\[
(a, b) + (c, d) := (a + c, b + d) \quad \forall a, b, c, d \in \mathbb{R}
\]

\[
(a, b) \cdot (c, d) := (a \cdot c - b \cdot d, b \cdot c + a \cdot d) \quad \forall a, b, c, d \in \mathbb{R}
\]

The addition and multiplication of real numbers follow standard rules of \( \mathbb{R} \).

Lemma 28: Commutativity, associativity and distributivity hold for \((\mathbb{C}, +, \cdot)\).

Alternate view: A complex number \((a, b)\) is regarded as the sum of a real and an imaginary part: \(a + bi\), with \(i^2 = -1\).

Applying standard rules of algebra used when multiplying real numbers (and the symbol \(i\)) is consistent with the definitions above:

\[
(2 + 3i) \cdot (1 - 2i) = 2 \cdot 1 + (3 \cdot 1)i - (2 \cdot 2)i^2 = (2 + 6i) + (3 - 4) = 8 - i.
\]
Definition 27 (Complex numbers, Dt.: komplexe Zahlen)

The complex numbers, \( \mathbb{C} \), are formed by the set of ordered pairs of real numbers together with operations \( + : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C} \) and \( \cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C} \) defined as follows:

- \((a, b) + (c, d) := (a + c, b + d)\) \(\forall a, b, c, d \in \mathbb{R},\)

The addition and multiplication of real numbers follow standard rules of \( \mathbb{R} \).

Lemma 28

Commutativity, associativity and distributivity hold for \( (\mathbb{C}, +, \cdot) \).

Alternate view: A complex number \((a, b)\) is regarded as the sum of a real and an imaginary part: \(a + bi\), with \(i^2 := -1\).

Applying standard rules of algebra used when multiplying real numbers (and the symbol \(i\)) is consistent with the definitions above:

\[
(2 + 3i) \cdot (1 - 2i) = 2 \cdot 1 + (3 \cdot 1)i - (2 \cdot 2)i^2 = 2 + 6i - 4(-1) = 2 + 6i + 4 = 8 + 6i
\]
Complex Numbers

Definition 27 (Complex numbers, Dt.: komplexe Zahlen)

The complex numbers, $\mathbb{C}$, are formed by the set of ordered pairs of real numbers together with operations $+ : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ and $\cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ defined as follows:

- $(a, b) + (c, d) := (a + c, b + d) \quad \forall a, b, c, d \in \mathbb{R}$,
- $(a, b) \cdot (c, d) := (a \cdot c - b \cdot d, b \cdot c + a \cdot d) \quad \forall a, b, c, d \in \mathbb{R}$,
Complex Numbers

Definition 27 (Complex numbers, Dt.: komplexe Zahlen)

The complex numbers, \( \mathbb{C} \), are formed by the set of ordered pairs of real numbers together with operations \( + : \mathbb{C} \times \mathbb{C} \to \mathbb{C} \) and \( \cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C} \) defined as follows:

- \((a, b) + (c, d) := (a + c, b + d) \quad \forall a, b, c, d \in \mathbb{R},\)
- \((a, b) \cdot (c, d) := (a \cdot c - b \cdot d, b \cdot c + a \cdot d) \quad \forall a, b, c, d \in \mathbb{R},\)

The addition and multiplication of real numbers follow standard rules of \( \mathbb{R} \).
Complex Numbers

**Definition 27 (Complex numbers, Dt.: komplexe Zahlen)**

The complex numbers, \( \mathbb{C} \), are formed by the set of ordered pairs of real numbers together with operations \( + : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C} \) and \( \cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C} \) defined as follows:

- \((a, b) + (c, d) := (a + c, b + d)\) \(\forall a, b, c, d \in \mathbb{R}\),
- \((a, b) \cdot (c, d) := (a \cdot c - b \cdot d, b \cdot c + a \cdot d)\) \(\forall a, b, c, d \in \mathbb{R}\).

The addition and multiplication of real numbers follow standard rules of \( \mathbb{R} \).

**Lemma 28**

Commutativity, associativity and distributivity hold for \((\mathbb{C}, +, \cdot)\).
Complex Numbers

**Definition 27 (Complex numbers, Dt.: komplexe Zahlen)**

The complex numbers, $\mathbb{C}$, are formed by the set of ordered pairs of real numbers together with operations $+ : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ and $\cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ defined as follows:

- $(a, b) + (c, d) := (a + c, b + d) \quad \forall a, b, c, d \in \mathbb{R}$,
- $(a, b) \cdot (c, d) := (a \cdot c - b \cdot d, b \cdot c + a \cdot d) \quad \forall a, b, c, d \in \mathbb{R}$,

The addition and multiplication of real numbers follow standard rules of $\mathbb{R}$.

**Lemma 28**

Commutativity, associativity and distributivity hold for $(\mathbb{C}, +, \cdot)$.

- Alternate view: A complex number $(a, b)$ is regarded as the sum of a real and an imaginary part: $a + bi$, with $i^2 := -1$. 
Complex Numbers

Definition 27 (Complex numbers, Dt.: komplexe Zahlen)

The complex numbers, \( \mathbb{C} \), are formed by the set of ordered pairs of real numbers together with operations \( + : \mathbb{C} \times \mathbb{C} \to \mathbb{C} \) and \( \cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C} \) defined as follows:

- \((a, b) + (c, d) := (a + c, b + d)\) \( \forall a, b, c, d \in \mathbb{R} \),
- \((a, b) \cdot (c, d) := (a \cdot c - b \cdot d, b \cdot c + a \cdot d)\) \( \forall a, b, c, d \in \mathbb{R} \),

The addition and multiplication of real numbers follow standard rules of \( \mathbb{R} \).

Lemma 28

Commutativity, associativity and distributivity hold for \((\mathbb{C}, +, \cdot)\).

- Alternate view: A complex number \((a, b)\) is regarded as the sum of a real and an imaginary part: \(a + bi\), with \(i^2 := -1\).
- Applying standard rules of algebra used when multiplying real numbers (and the symbol \(i\)) is consistent with the definitions above:

\[
(2 + 3i) \cdot (1 - 2i) = 2 \cdot 1 + (3 \cdot 1)i - (2 \cdot 2)i - (3 \cdot 2)i^2 = (2 + 6) + (3 - 4)i = 8 - i
\]
The *complex plane* is a modification of the standard Cartesian plane, with a real axis and an imaginary axis that intersect in a right angle at the point $(0, 0)$. That is, real numbers run left-right and imaginary numbers run bottom-top.
Complex Numbers

Definition 29 (Absolute value)

The *absolute value* $|z|$ (or modulus or magnitude) of a complex number $z := a + bi \in \mathbb{C}$ is given by

$$|z| := \sqrt{a^2 + b^2}.$$
Complex Numbers

**Definition 29 (Absolute value)**

The *absolute value* $|z|$ (or modulus or magnitude) of a complex number $z := a + bi \in \mathbb{C}$ is given by

$$|z| := \sqrt{a^2 + b^2}.$$

**Definition 30 (Complex conjugate, Dt.: konjugiert-komplexe Zahl)**

The complex *conjugate* $\overline{z}$ of the complex number $z := a + bi \in \mathbb{C}$ is given by

$$\overline{z} := a - bi.$$
Complex Numbers

**Definition 29 (Absolute value)**

The *absolute value* $|z|$ (or modulus or magnitude) of a complex number $z := a + bi \in \mathbb{C}$ is given by

$$|z| := \sqrt{a^2 + b^2}.$$

**Definition 30 (Complex conjugate, Dt.: konjugiert-komplexe Zahl)**

The complex *conjugate* $\bar{z}$ of the complex number $z := a + bi \in \mathbb{C}$ is given by

$$\bar{z} := a - bi.$$

**Definition 31 (Multiplicative inverse)**

The *multiplicative inverse* for $z \in \mathbb{C}$, with $z \neq 0$ is defined as

$$z^{-1} := \frac{\bar{z}}{|z|^2}.$$
Complex Numbers

Lemma 32

Easy to check for all $z_1, z_2 \in \mathbb{C}$:

\[
\begin{align*}
\overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2} \\
\overline{z_1 \cdot z_2} &= \overline{z_1} \cdot \overline{z_2} \\
\overline{z_1} &= z_1 \\
|z_1| &= |\overline{z_1}| \\
z_1 \cdot z_1^{-1} &= 1 \\
|z_1|^2 &= z_1 \cdot \overline{z_1}
\end{align*}
\]
Complex Numbers

**Lemma 32**

Easy to check for all $z_1, z_2 \in \mathbb{C}$:

\[
\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2} \quad \overline{z_1} = z_1
\]

\[
|z_1| = |\overline{z_1}| \quad z_1 \cdot z_1^{-1} = 1 \quad |z_1|^2 = z_1 \cdot \overline{z_1}
\]

**Theorem 33**

The complex numbers $(\mathbb{C}, +, \cdot)$ form a field.
A complex number $z = a + bi$ can also be written as

$$z = a + bi = r(\cos \phi + i \sin \phi),$$

with $r := |a + bi|$ and $\phi$ such that $a = r \cos \phi$ and $b = r \sin \phi$. 

By applying standard trigonometric identities, we get

$$z_1 \cdot z_2 = r_1 r_2 [\cos(\phi_1 + \phi_2) + i \sin(\phi_1 + \phi_2)],$$

$$z_1 / z_2 = r_1 / r_2 [\cos(\phi_1 - \phi_2) + i \sin(\phi_1 - \phi_2)].$$

Thus, the multiplication of one complex number with another complex number can be seen as a simultaneous rotation and stretching.

Lemma 34 (de Moivre)

Let $z := r(\cos \phi + i \sin \phi)$. Then

$$z^n = r^n (\cos n\phi + i \sin n\phi)$$

for all $n \in \mathbb{N}$. 

Complex Numbers and de Moivre’s Formula

- A complex number $z = a + bi$ can also be written as
  $$z = a + bi = r(\cos \phi + i \sin \phi),$$
  with $r := |a + bi|$ and $\phi$ such that $a = r \cos \phi$ and $b = r \sin \phi$.
- By applying standard trigonometric identities, we get
  $$z_1 \cdot z_2 = r_1 r_2 [\cos(\phi_1 + \phi_2) + i \sin(\phi_1 + \phi_2)],$$
  $$z_1 / z_2 = r_1 / r_2 [\cos(\phi_1 - \phi_2) + i \sin(\phi_1 - \phi_2)].$$
- Thus, the multiplication of one complex number with another complex number can be seen as a simultaneous rotation and stretching.
Complex Numbers and de Moivre’s Formula

- A complex number $z = a + bi$ can also be written as

  $$z = a + bi = r(\cos \phi + i \sin \phi),$$

  with $r := |a + bi|$ and $\phi$ such that $a = r \cos \phi$ and $b = r \sin \phi$.

- By applying standard trigonometric identities, we get

  $$z_1 \cdot z_2 = r_1 r_2[\cos(\phi_1 + \phi_2) + i \sin(\phi_1 + \phi_2)],$$

  $$z_1 / z_2 = r_1 / r_2[\cos(\phi_1 - \phi_2) + i \sin(\phi_1 - \phi_2)].$$

- Thus, the multiplication of one complex number with another complex number can be seen as a simultaneous rotation and stretching.

**Lemma 34 (de Moivre)**

Let $z := r(\cos \phi + i \sin \phi)$. Then

$$z^n = r^n(\cos n\phi + i \sin n\phi)$$

for all $n \in \mathbb{N}$. 
Theorem 35 (Euler)

For any $\phi \in \mathbb{R}$,

$$e^{i\phi} = \cos \phi + i \sin \phi.$$
Theorem 35 (Euler)

For any $\phi \in \mathbb{R}$,

$$e^{i\phi} = \cos \phi + i \sin \phi.$$ 

- Thus, $e^{i\phi}$ traces out the unit circle in the complex plane as $\phi$ runs from 0 to $2\pi$.
- Important application: Modeling (electric) signals that vary periodically over time.
Complex Numbers and Euler’s Formula

Theorem 35 (Euler)
For any $\phi \in \mathbb{R}$,

$$e^{i\phi} = \cos \phi + i \sin \phi.$$ 

- Thus, $e^{i\phi}$ traces out the unit circle in the complex plane as $\phi$ runs from 0 to $2\pi$.
- Important application: Modeling (electric) signals that vary periodically over time.

Corollary 36
$$e^{i\pi} = -1.$$
**Complex Numbers and Euler’s Formula**

**Sketch of Proof of Theorem 35:** The theory of Taylor/Maclaurin series tells us that, for all $x \in \mathbb{R}$:

$$
\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots
$$
Sketch of Proof of Theorem 35: The theory of Taylor/Maclaurin series tells us that, for all $x \in \mathbb{R}$:

\[
\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots
\]

\[
\sin x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots
\]
Sketch of Proof of Theorem 35: The theory of Taylor/Maclaurin series tells us that, for all $x \in \mathbb{R}$:

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots$$

$$\sin x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots$$

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \ldots$$
Complex Numbers and Euler’s Formula

**Sketch of Proof of Theorem 35:** The theory of Taylor/Maclaurin series tells us that, for all \( x \in \mathbb{R} \):

\[
\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots
\]

\[
\sin x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots
\]

\[
e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \ldots
\]

Recall that \( i^2 = -1 \).
Complex Numbers and Euler’s Formula

**Sketch of Proof of Theorem 35:** The theory of Taylor/Maclaurin series tells us that, for all \( x \in \mathbb{R} \):

\[
\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots
\]

\[
\sin x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots
\]

\[
e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \ldots
\]

Recall that \( i^2 = -1 \). Hence, \( i^3 = -i \), \( i^4 = 1 \), \( i^5 = i \), etc.
Sketch of Proof of Theorem 35: The theory of Taylor/Maclaurin series tells us that, for all \( x \in \mathbb{R} \):

\[
\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots
\]

\[
\sin x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots
\]

\[
e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \ldots
\]

Recall that \( i^2 = -1 \). Hence, \( i^3 = -i \), \( i^4 = 1 \), \( i^5 = i \), etc. If we replace \( x \) by \( ix \) in the series for \( e^x \) then we get

\[
e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{k=0}^{\infty} i^k x^k
\]
Complex Numbers and Euler’s Formula

**Sketch of Proof of Theorem 35**: The theory of Taylor/Maclaurin series tells us that, for all $x \in \mathbb{R}$:

\[
\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots
\]

\[
\sin x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots
\]

\[
e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \ldots
\]

Recall that $i^2 = -1$. Hence, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, etc. If we replace $x$ by $ix$ in the series for $e^x$ then we get

\[
e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{k=0}^{\infty} \frac{i^k x^k}{k!} = 1 + ix + \frac{i^2 x^2}{2!} + \frac{i^3 x^3}{3!} + \frac{i^4 x^4}{4!} + \frac{i^5 x^5}{5!} + \frac{i^6 x^6}{6!} + \ldots
\]
Complex Numbers and Euler’s Formula

**Sketch of Proof of Theorem 35:** The theory of Taylor/Maclaurin series tells us that, for all $x \in \mathbb{R}$:

\[
\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots
\]

\[
\sin x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots
\]

\[
e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \ldots
\]

Recall that $i^2 = -1$. Hence, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, etc. If we replace $x$ by $ix$ in the series for $e^x$ then we get

\[
e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{k=0}^{\infty} \frac{i^k x^k}{k!} = 1 + ix + \frac{i^2 x^2}{2!} + \frac{i^3 x^3}{3!} + \frac{i^4 x^4}{4!} + \frac{i^5 x^5}{5!} + \frac{i^6 x^6}{6!} + \ldots
\]

\[
= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots\right) + i \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots\right)
\]
**Complex Numbers and Euler’s Formula**

**Sketch of Proof of Theorem 35:** The theory of Taylor/Maclaurin series tells us that, for all $x \in \mathbb{R}$:

$$
\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots
$$

$$
\sin x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots
$$

$$
e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \ldots
$$

Recall that $i^2 = -1$. Hence, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, etc. If we replace $x$ by $ix$ in the series for $e^x$ then we get

$$
e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{k=0}^{\infty} \frac{i^k x^k}{k!} = 1 + ix + \frac{i^2 x^2}{2!} + \frac{i^3 x^3}{3!} + \frac{i^4 x^4}{4!} + \frac{i^5 x^5}{5!} + \frac{i^6 x^6}{6!} + \ldots
$$

$$
= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots\right) + i \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots\right)
$$

$$
= \cos x + i \sin x.
$$
The **Mandelbrot set** is the locus of complex numbers $c$ for which the sequence $(z_0, z_1, z_2, \ldots)$, with

$$z_n := \begin{cases} z_{n-1} \cdot z_{n-1} + c & \text{if } n > 0, \\ (0, 0) & \text{if } n = 0, \end{cases}$$

does not diverge.
Mandelbrot Set

- The *Mandelbrot set* is the locus of complex numbers $c$ for which the sequence $(z_0, z_1, z_2, \ldots)$, with

$$z_n := \begin{cases} z_{n-1} \cdot z_{n-1} + c & \text{if } n > 0, \\ (0, 0) & \text{if } n = 0, \end{cases}$$

does not diverge.

- If we regard the real and imaginary parts of $c$ as pixel coordinates, then pixels can be colored according to the number of iterations after which the sequence $(z_0, z_1, z_2, \ldots)$ crosses an arbitrarily chosen threshold.
Mandelbrot Set

- The *Mandelbrot set* is the locus of complex numbers $c$ for which the sequence $(z_0, z_1, z_2, \ldots)$, with

\[
z_n := \begin{cases} 
  z_{n-1} \cdot z_{n-1} + c & \text{if } n > 0, \\
  (0, 0) & \text{if } n = 0,
\end{cases}
\]

does not diverge.

- If we regard the real and imaginary parts of $c$ as pixel coordinates, then pixels can be colored according to the number of iterations after which the sequence $(z_0, z_1, z_2, \ldots)$ crosses an arbitrarily chosen threshold.

- Typically, black is used for the values of $c$ for which the sequence has not crossed the threshold after a predetermined number of iterations.
The Mandelbrot set is the locus of complex numbers $c$ for which the sequence $(z_0, z_1, z_2, \ldots)$, with

$$z_n := \begin{cases} z_{n-1} \cdot z_{n-1} + c & \text{if } n > 0, \\ (0, 0) & \text{if } n = 0, \end{cases}$$

does not diverge.

If we regard the real and imaginary parts of $c$ as pixel coordinates, then pixels can be colored according to the number of iterations after which the sequence $(z_0, z_1, z_2, \ldots)$ crosses an arbitrarily chosen threshold.

Typically, black is used for the values of $c$ for which the sequence has not crossed the threshold after a predetermined number of iterations.

[Image credit: Michael Bradshad]
Mandelbrot Set

[Image credit: https://commons.wikimedia.org/wiki/File:Mandelbrot_set_2500px.png]
A Julia set, for some constant $c \in \mathbb{C}$, is the locus of complex numbers $z$ for which the sequence $(z_0, z_1, z_2, \ldots)$, with

$$z_n := \begin{cases} 
  z_{n-1} \cdot z_{n-1} + c & \text{if } n > 0, \\
  z & \text{if } n = 0,
\end{cases}$$

does not diverge.
Julia Set

- A *Julia set*, for some constant $c \in \mathbb{C}$, is the locus of complex numbers $z$ for which the sequence $(z_0, z_1, z_2, \ldots)$, with

$$z_n := \begin{cases} z_{n-1} \cdot z_{n-1} + c & \text{if } n > 0, \\ z & \text{if } n = 0, \end{cases}$$

does not diverge.
Algebraic Concepts

- Algebraic Structures
- Real Numbers and Vector Space $\mathbb{R}^n$
- Complex Numbers $\mathbb{C}$

Polynomials
- Definition
- Arithmetic
- Roots
- Evaluation
Polynomials

Definition 37 (Monomial, Dt.: Monom)

A (real) *monomial* in $m$ variables $x_1, x_2, \ldots, x_m$ is a product of a coefficient $c \in \mathbb{R}$ and powers of the variables $x_i$ with exponents $k_i \in \mathbb{N}_0$:

$$c \prod_{i=1}^{m} x_i^{k_i} = c \cdot x_1^{k_1} \cdot x_2^{k_2} \cdot \ldots \cdot x_m^{k_m}.$$
Definition 37 (Monomial, Dt.: Monom)

A (real) monomial in \( m \) variables \( x_1, x_2, \ldots, x_m \) is a product of a coefficient \( c \in \mathbb{R} \) and powers of the variables \( x_i \) with exponents \( k_i \in \mathbb{N}_0 \):

\[
c \prod_{i=1}^{m} x_i^{k_i} = c \cdot x_1^{k_1} \cdot x_2^{k_2} \cdot \ldots \cdot x_m^{k_m}.
\]

The degree of the monomial is given by \( \sum_{i=1}^{m} k_i \).
Polynomials

Definition 37 (Monomial, Dt.: Monom)

A (real) monomial in $m$ variables $x_1, x_2, \ldots, x_m$ is a product of a coefficient $c \in \mathbb{R}$ and powers of the variables $x_i$ with exponents $k_i \in \mathbb{N}_0$:

$$c \prod_{i=1}^{m} x_i^{k_i} = c \cdot x_1^{k_1} \cdot x_2^{k_2} \cdot \ldots \cdot x_m^{k_m}.$$ 

The degree of the monomial is given by $\sum_{i=1}^{m} k_i$.

Definition 38 (Polynomial, Dt.: Polynom)

A (real) polynomial in $m$ variables $x_1, x_2, \ldots, x_m$ is a finite sum of monomials in $x_1, x_2, \ldots, x_m$.

A polynomial is univariate if $m=1$, bivariate if $m=2$, and multivariate otherwise.

Definition 39 (Degree, Dt.: Grad)

The degree of a polynomial is the maximum degree of its monomials.
Polynomials

**Definition 37 (Monomial, Dt.: Monom)**

A (real) *monomial* in $m$ variables $x_1, x_2, \ldots, x_m$ is a product of a coefficient $c \in \mathbb{R}$ and powers of the variables $x_i$ with exponents $k_i \in \mathbb{N}_0$:

$$c \prod_{i=1}^{m} x_i^{k_i} = c \cdot x_1^{k_1} \cdot x_2^{k_2} \cdot \ldots \cdot x_m^{k_m}.$$

The *degree of the monomial* is given by $\sum_{i=1}^{m} k_i$.

**Definition 38 (Polynomial, Dt.: Polynom)**

A (real) *polynomial* in $m$ variables $x_1, x_2, \ldots, x_m$ is a finite sum of monomials in $x_1, x_2, \ldots, x_m$.

A polynomial is *univariate* if $m = 1$, *bivariate* if $m = 2$, and *multivariate* otherwise.
Polynomials

Definition 37 (Monomial, Dt.: Monom)
A (real) monomial in $m$ variables $x_1, x_2, \ldots, x_m$ is a product of a coefficient $c \in \mathbb{R}$ and powers of the variables $x_i$ with exponents $k_i \in \mathbb{N}_0$:

$$c \prod_{i=1}^{m} x_i^{k_i} = c \cdot x_1^{k_1} \cdot x_2^{k_2} \cdots \cdot x_m^{k_m}.$$ 

The degree of the monomial is given by $\sum_{i=1}^{m} k_i$.

Definition 38 (Polynomial, Dt.: Polynom)
A (real) polynomial in $m$ variables $x_1, x_2, \ldots, x_m$ is a finite sum of monomials in $x_1, x_2, \ldots, x_m$.
A polynomial is univariate if $m = 1$, bivariate if $m = 2$, and multivariate otherwise.

Definition 39 (Degree, Dt.: Grad)
The degree of a polynomial is the maximum degree of its monomials.
Polynomials

- Hence, a univariate polynomial over \( \mathbb{R} \) with variable \( x \) is a term of the form

\[
a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,
\]

with coefficients \( a_0, \ldots, a_n \in \mathbb{R} \) and \( a_n \neq 0 \).

- It is a convention to drop all monomials whose coefficients are zero.
Polynomials

- Hence, a univariate polynomial over $\mathbb{R}$ with variable $x$ is a term of the form

  $$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

  with coefficients $a_0, \ldots, a_n \in \mathbb{R}$ and $a_n \neq 0$.

- It is a convention to drop all monomials whose coefficients are zero.

- Univariate polynomials are usually written according to a decreasing order of exponents of their monomials.

- In that case, the first term is the *leading term* which indicates the degree of the polynomial; its coefficient is the *leading coefficient*. 
Polynomials

- Hence, a univariate polynomial over $\mathbb{R}$ with variable $x$ is a term of the form

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

with coefficients $a_0, \ldots, a_n \in \mathbb{R}$ and $a_n \neq 0$.

- It is a convention to drop all monomials whose coefficients are zero.

- Univariate polynomials are usually written according to a decreasing order of exponents of their monomials.

- In that case, the first term is the leading term which indicates the degree of the polynomial; its coefficient is the leading coefficient.

- Univariate polynomials of degree
  - 0 are called constant polynomials,
  - 1 are called linear polynomials,
  - 2 are called quadratic polynomials,
  - 3 are called cubic polynomials,
  - 4 are called quartic polynomials,
  - 5 are called quintic polynomials.
Polynomial Arithmetic

- We define the addition of (univariate) polynomials based on the pairwise addition of corresponding coefficients:

\[
\left( \sum_{i=0}^{n} a_i x^i \right) + \left( \sum_{i=0}^{n} b_i x^i \right) := \sum_{i=0}^{n} (a_i + b_i) x^i
\]
**Polynomial Arithmetic**

- We define the addition of (univariate) polynomials based on the pairwise addition of corresponding coefficients:

  \[
  \left( \sum_{i=0}^{n} a_i x^i \right) + \left( \sum_{i=0}^{n} b_i x^i \right) := \sum_{i=0}^{n} (a_i + b_i) x^i
  \]

- The multiplication of polynomials is based on the multiplication within \(\mathbb{R}\), distributivity, and the rules

  \[ax = xa \quad \text{and} \quad x^m x^k = x^{m+k}\]

  for all \(a \in \mathbb{R}\) and \(m, k \in \mathbb{N}\):

  \[
  \left( \sum_{i=0}^{n} a_i x^i \right) \cdot \left( \sum_{j=0}^{m} b_j x^j \right) := \sum_{i=0}^{n} \sum_{j=0}^{m} (a_i b_j) x^{i+j}
  \]
Polynomial Arithmetic

- We define the addition of (univariate) polynomials based on the pairwise addition of corresponding coefficients:

\[
\left( \sum_{i=0}^{n} a_i x^i \right) + \left( \sum_{i=0}^{n} b_i x^i \right) := \sum_{i=0}^{n} (a_i + b_i) x^i
\]

- The multiplication of polynomials is based on the multiplication within \( \mathbb{R} \), distributivity, and the rules

\[
ax = xa \quad \text{and} \quad x^m x^k = x^{m+k}
\]

for all \( a \in \mathbb{R} \) and \( m, k \in \mathbb{N} \):

\[
\left( \sum_{i=0}^{n} a_i x^i \right) \cdot \left( \sum_{j=0}^{m} b_j x^j \right) := \sum_{i=0}^{n} \sum_{j=0}^{m} (a_i b_j) x^{i+j}
\]

- Elementary properties of polynomials: One can prove easily that the addition, multiplication and composition of two polynomials as well as their derivative and antiderivative (indefinite integral) again yield a polynomial.
**Polynomial Arithmetic**

- We define the addition of (univariate) polynomials based on the pairwise addition of corresponding coefficients:

  \[
  \left( \sum_{i=0}^{n} a_i x^i \right) + \left( \sum_{i=0}^{n} b_i x^i \right) := \sum_{i=0}^{n} (a_i + b_i) x^i
  \]

- The multiplication of polynomials is based on the multiplication within \( \mathbb{R} \), distributivity, and the rules

  \[ax = xa \quad \text{and} \quad x^m x^k = x^{m+k}\]

  for all \( a \in \mathbb{R} \) and \( m, k \in \mathbb{N} \):

  \[
  \left( \sum_{i=0}^{n} a_i x^i \right) \cdot \left( \sum_{j=0}^{m} b_j x^j \right) := \sum_{i=0}^{n} \sum_{j=0}^{m} (a_i b_j) x^{i+j}
  \]

- Elementary properties of polynomials: One can prove easily that the addition, multiplication and composition of two polynomials as well as their derivative and antiderivative (indefinite integral) again yield a polynomial.

- Same for multivariate polynomials.
Polynomial Arithmetic

- Instead of $\mathbb{R}$ any commutative ring $(R, +, \cdot)$ and symbols $x, y, \ldots$ that are not contained in $R$ would do. E.g.,

  $$a_{2,3}x^2y^3 + a_{1,1}xy + a_{0,1}y + a_{0,0}$$

  with $a_{2,3}, a_{1,1}, a_{0,1}, a_{0,0} \in R$. 

Polynomial Arithmetic

- Instead of \( \mathbb{R} \) any commutative ring \((R, +, \cdot)\) and symbols \(x, y, \ldots\) that are not contained in \(R\) would do. E.g.,

\[
a_{2,3}x^2y^3 + a_{1,1}xy + a_{0,1}y + a_{0,0}
\]

with \(a_{2,3}, a_{1,1}, a_{0,1}, a_{0,0} \in R\).

**Lemma 40**

The set of all polynomials with coefficients in the commutative ring \((R, +, \cdot)\) and a symbol (variable) \(x \not\in R\) forms a commutative ring, the *ring of polynomials over \(R\)*, commonly denoted by \(R[x]\).
Polynomial Arithmetic

- Instead of \( \mathbb{R} \) any commutative ring \((R, +, \cdot)\) and symbols \(x, y, \ldots\) that are not contained in \(R\) would do. E.g.,

\[
a_{2,3}x^2y^3 + a_{1,1}xy + a_{0,1}y + a_{0,0}
\]

with \(a_{2,3}, a_{1,1}, a_{0,1}, a_{0,0} \in R\).

**Lemma 40**

The set of all polynomials with coefficients in the commutative ring \((R, +, \cdot)\) and a symbol (variable) \(x \notin R\) forms a commutative ring, the *ring of polynomials over \(R\)*, commonly denoted by \(R[x]\).

- Multivariate polynomials can also be seen as univariate polynomials with coefficients out of a ring of polynomials. E.g.,

\[
a_{2,3}x^2y^3 + a_{1,1}xy + a_{0,1}y + a_{0,0} = (a_{2,3}x^2)y^3 + (a_{1,1}x + a_{0,1})y + a_{0,0}
\]

is an element of \(R[x, y] := (R[x])[y]\).
Polynomial Arithmetic

- Instead of $\mathbb{R}$ any commutative ring $(R, +, \cdot)$ and symbols $x, y, \ldots$ that are not contained in $R$ would do. E.g.,

$$a_{2,3}x^2y^3 + a_{1,1}xy + a_{0,1}y + a_{0,0} \quad \text{with } a_{2,3}, a_{1,1}, a_{0,1}, a_{0,0} \in R.$$

**Lemma 40**

The set of all polynomials with coefficients in the commutative ring $(R, +, \cdot)$ and a symbol (variable) $x \not\in R$ forms a commutative ring, the *ring of polynomials over $R$*, commonly denoted by $R[x]$.

- Multivariate polynomials can also be seen as univariate polynomials with coefficients out of a ring of polynomials. E.g.,

$$a_{2,3}x^2y^3 + a_{1,1}xy + a_{0,1}y + a_{0,0} = (a_{2,3}x^2)y^3 + (a_{1,1}x + a_{0,1})y + a_{0,0}$$

is an element of $R[x, y] := (R[x])[y]$.

**Definition 41**

Two polynomials are equal if and only if the sequences of their coefficients (arranged in some specific order) are equal.
Polynomials: Vector Space

**Theorem 42**

The univariate polynomials of \( \mathbb{R}[x] \) form an infinite *vector space* over \( \mathbb{R} \). The so-called *power basis* of this vector space is given by the monomials \( 1, x, x^2, x^3, \ldots \).
Theorem 42

The univariate polynomials of $\mathbb{R}[x]$ form an infinite vector space over $\mathbb{R}$. The so-called power basis of this vector space is given by the monomials $1, x, x^2, x^3, \ldots$.

- The $n + 1$ monomials $1, x, x^2, x^3, \ldots, x^n$ form a basis of the vector space of polynomials of degree up to $n$ over $\mathbb{R}$, for all $n \in \mathbb{N}_0$. 
Theorem 42
The univariate polynomials of $\mathbb{R}[x]$ form an infinite vector space over $\mathbb{R}$. The so-called power basis of this vector space is given by the monomials $1, x, x^2, x^3, \ldots$.

- The $n + 1$ monomials $1, x, x^2, x^3, \ldots, x^n$ form a basis of the vector space of polynomials of degree up to $n$ over $\mathbb{R}$, for all $n \in \mathbb{N}_0$.
- The power basis is not the only meaningful basis of the polynomials $\mathbb{R}[x]$: See, e.g., the Bernstein polynomials that are used to form Bézier curves.

Definition 43 (Bernstein polynomials)
The $n + 1$ Bernstein polynomials of degree $n$, for $n \in \mathbb{N}_0$, are defined as

$$B_{k,n}(x) := \binom{n}{k} x^k (1 - x)^{n-k} \quad \text{for} \ k \in \{0, 1, \ldots, n\}, \text{with} \ 0^0 := 1.$$
Theorem 42

The univariate polynomials of $\mathbb{R}[x]$ form an infinite vector space over $\mathbb{R}$. The so-called power basis of this vector space is given by the monomials $1, x, x^2, x^3, \ldots$.

- The $n + 1$ monomials $1, x, x^2, x^3, \ldots, x^n$ form a basis of the vector space of polynomials of degree up to $n$ over $\mathbb{R}$, for all $n \in \mathbb{N}_0$.
- The power basis is not the only meaningful basis of the polynomials $\mathbb{R}[x]$: See, e.g., the Bernstein polynomials that are used to form Bézier curves.

Definition 43 (Bernstein polynomials)

The $n + 1$ Bernstein polynomials of degree $n$, for $n \in \mathbb{N}_0$, are defined as

$$B_{k,n}(x) := \binom{n}{k} x^k (1 - x)^{n-k} \quad \text{for } k \in \{0, 1, \ldots, n\}, \text{ with } 0^0 := 1.$$ 

Theorem 44

The Bernstein polynomials of degree $n$ form a basis of the vector space of polynomials of degree up to $n$ over $\mathbb{R}$, for all $n \in \mathbb{N}_0$. 
Definition 45 (Polynomial equation)

A polynomial equation (aka algebraic equation) is an equation in which a polynomial is set equal to another polynomial.
Polynomials: Roots

**Definition 45 (Polynomial equation)**

A *polynomial equation* (aka *algebraic equation*) is an equation in which a polynomial is set equal to another polynomial.

**Definition 46 (Root, Dt.: Wurzel)**

The polynomial $p \in \mathbb{R}[x]$ has a *root* (aka zero) $r \in \mathbb{R}$ if $(x - r)$ divides $p$.

**Definition 47 (Multiplicity, Dt.: Vielfachheit)**

A root $r$ of a polynomial $p$ in $x$ is of *multiplicity* $k$ if $k \in \mathbb{N}$ is the maximum integer such that $(x - r)^k$ divides $p$.

**Theorem 48 (Fundamental Theorem of Algebra)**

The number of (complex) roots of a polynomial with real coefficients may not exceed its degree. It equals the degree if all roots are counted with their multiplicities.
**Definition 45 (Polynomial equation)**

A *polynomial equation* (aka *algebraic equation*) is an equation in which a polynomial is set equal to another polynomial.

**Definition 46 (Root, Dt.: Wurzel)**

The polynomial \( p \in \mathbb{R}[x] \) has a *root* (aka zero) \( r \in \mathbb{R} \) if \((x - r)\) divides \( p \).

- Hence, if \( r \) is a root of \( p \) then \( p = (x - r) \cdot p_1 \) for some \( p_1 \in \mathbb{R}[x] \), and \( p(r) = 0 \).
**Polynomials: Roots**

**Definition 45 (Polynomial equation)**

A *polynomial equation* (aka *algebraic equation*) is an equation in which a polynomial is set equal to another polynomial.

**Definition 46 (Root, Dt.: Wurzel)**

The polynomial $p \in \mathbb{R}[x]$ has a *root* (aka *zero*) $r \in \mathbb{R}$ if $(x - r)$ divides $p$.

- Hence, if $r$ is a root of $p$ then $p = (x - r) \cdot p_1$ for some $p_1 \in \mathbb{R}[x]$, and $p(r) = 0$.

**Definition 47 (Multiplicity, Dt.: Vielfachheit)**

A root $r$ of a polynomial $p$ in $x$ is of multiplicity $k$ if $k \in \mathbb{N}$ is the maximum integer such that $(x - r)^k$ divides $p$. 
Polynomials: Roots

Definition 45 (Polynomial equation)
A polynomial equation (aka algebraic equation) is an equation in which a polynomial is set equal to another polynomial.

Definition 46 (Root, Dt.: Wurzel)
The polynomial \( p \in \mathbb{R}[x] \) has a root (aka zero) \( r \in \mathbb{R} \) if \( (x - r) \) divides \( p \).

- Hence, if \( r \) is a root of \( p \) then \( p = (x - r) \cdot p_1 \) for some \( p_1 \in \mathbb{R}[x] \), and \( p(r) = 0 \).

Definition 47 (Multiplicity, Dt.: Vielfachheit)
A root \( r \) of a polynomial \( p \) in \( x \) is of multiplicity \( k \) if \( k \in \mathbb{N} \) is the maximum integer such that \( (x - r)^k \) divides \( p \).

Theorem 48 (Fundamental Theorem of Algebra)
The number of (complex) roots of a polynomial with real coefficients may not exceed its degree. It equals the degree if all roots are counted with their multiplicities.
Polynomials: Roots

Recall the quadratic formula taught in secondary school for solving second-degree polynomial equations: For $a \in \mathbb{R}\setminus\{0\}$ and $b, c \in \mathbb{R}$,

$$x_{1,2} := \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

yields the two (possibly complex) roots $x_1$ and $x_2$ of $ax^2 + bx + c$. 

Theorem 49 (Abel-Ruffini (1824))

No algebraic solution for the roots of an arbitrary polynomial of degree five or higher exists.

An algebraic solution is a closed-form expression in terms of the coefficients of the polynomial that relies only on addition, subtraction, multiplication, division, raising to integer powers, and computing $k$-th roots (square roots, cube roots, and other integer roots).

A closed-form expression is an expression that can be evaluated in a finite number of operations.
Recall the quadratic formula taught in secondary school for solving second-degree polynomial equations: For \( a \in \mathbb{R} \setminus \{0\} \) and \( b, c \in \mathbb{R} \),

\[
x_{1,2} := \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\]
yields the two (possibly complex) roots \( x_1 \) and \( x_2 \) of \( ax^2 + bx + c \).

Similar (albeit more complex) formulas exist for cubic and quartic polynomials.
Polynomials: Roots

- Recall the quadratic formula taught in secondary school for solving second-degree polynomial equations: For \( a \in \mathbb{R} \setminus \{0\} \) and \( b, c \in \mathbb{R} \),

  \[
  x_{1,2} := \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
  \]

  yields the two (possibly complex) roots \( x_1 \) and \( x_2 \) of \( ax^2 + bx + c \).
- Similar (albeit more complex) formulas exist for cubic and quartic polynomials.

Theorem 49 (Abel-Ruffini (1824))

No algebraic solution for the roots of an arbitrary polynomial of degree five or higher exists.

- An algebraic solution is a closed-form expressions in terms of the coefficients of the polynomial that relies only on addition, subtraction, multiplication, division, raising to integer powers, and computing \( k \)-th roots (square roots, cube roots, and other integer roots).
- A closed-form expression is an expression that can be evaluated in a finite number of operations.
Polynomials: Roots

Theorem 50 (Vieta’s formula)

For $a, b, c \in \mathbb{R}$, the roots $r_1, r_2$ of the quadratic polynomial $ax^2 + bx + c$ satisfy

$$r_1 + r_2 = -\frac{b}{a} \quad r_1 \cdot r_2 = \frac{c}{a}.$$
**Theorem 50 (Vieta’s formula)**

For $a, b, c \in \mathbb{R}$, the roots $r_1, r_2$ of the quadratic polynomial $ax^2 + bx + c$ satisfy

$$r_1 + r_2 = -\frac{b}{a} \quad r_1 \cdot r_2 = \frac{c}{a}.$$ 

For $a, b, c, d \in \mathbb{R}$, the roots $r_1, r_2, r_3$ of the cubic polynomial $ax^3 + bx^2 + cx + d$ satisfy

$$r_1 + r_2 + r_3 = -\frac{b}{a} \quad r_1 \cdot r_2 + r_1 \cdot r_3 + r_2 \cdot r_3 = \frac{c}{a} \quad r_1 \cdot r_2 \cdot r_3 = -\frac{d}{a}.$$ 

François Viète (Franciscus Vieta, 1540–1603). There is a more complex version of this theorem for arbitrary-degree polynomials.
Polynomials: Roots

Theorem 50 (Vieta’s formula)

For \(a, b, c \in \mathbb{R}\), the roots \(r_1, r_2\) of the quadratic polynomial \(ax^2 + bx + c\) satisfy

\[
\begin{align*}
    r_1 + r_2 &= -\frac{b}{a} & r_1 \cdot r_2 &= \frac{c}{a}.
\end{align*}
\]

For \(a, b, c, d \in \mathbb{R}\), the roots \(r_1, r_2, r_3\) of the cubic polynomial \(ax^3 + bx^2 + cx + d\) satisfy

\[
\begin{align*}
    r_1 + r_2 + r_3 &= -\frac{b}{a} & r_1 \cdot r_2 + r_1 \cdot r_3 + r_2 \cdot r_3 &= \frac{c}{a} & r_1 \cdot r_2 \cdot r_3 &= -\frac{d}{a}.
\end{align*}
\]

- François Viète (Franciscus Vieta, 1540–1603).
- There is a more complex version of this theorem for arbitrary-degree polynomials.
Definition 51 (Polynomial function; Dt.: Polynomfunktion)

A (univariate real) function \( f : I \to \mathbb{R} \), for an interval \( I \subseteq \mathbb{R} \), is a polynomial function over \( I \) if there exist \( n \in \mathbb{N}_0 \) and \( a_0, a_1, \ldots, a_n \in \mathbb{R} \) such that

\[
  f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad \text{for all } x \in I.
\]
Polynomials: Function

Definition 51 (Polynomial function; Dt.: Polynomfunktion)

A (univariate real) function $f : I \rightarrow \mathbb{R}$, for an interval $I \subseteq \mathbb{R}$, is a polynomial function over $I$ if there exist $n \in \mathbb{N}_0$ and $a_0, a_1, \ldots, a_n \in \mathbb{R}$ such that

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad \text{for all } x \in I.$$

- As usual, two (polynomial) functions over an interval $I \subseteq \mathbb{R}$ are identical if their values are identical for all arguments in $I$. 

Note: Two different polynomials may result in the same polynomial function! (E.g., over finite fields.) While some fields of mathematics (e.g., abstract algebra) make a clear distinction between polynomials and polynomial functions, we will freely mix these two terms. Also, unless noted explicitly, we will only deal with polynomials over $\mathbb{R}$. Note: Polynomial functions may come in disguise: $f(x) := \cos(2 \arccos(x))$ is a polynomial function over $[-1, 1]$ since we have $f(x) = 2x^2 - 1$ for all $x \in [-1, 1]$. 
Polynomials: Function

Definition 51 (Polynomial function; Dt.: Polynomfunktion)
A (univariate real) function \( f : I \rightarrow \mathbb{R} \), for an interval \( I \subseteq \mathbb{R} \), is a *polynomial function* over \( I \) if there exist \( n \in \mathbb{N}_0 \) and \( a_0, a_1, \ldots, a_n \in \mathbb{R} \) such that
\[
f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad \text{for all } x \in I.
\]

- As usual, two (polynomial) functions over an interval \( I \subseteq \mathbb{R} \) are identical if their values are identical for all arguments in \( I \).
- Note: Two different polynomials may result in the same polynomial function! (E.g., over finite fields.)
Polynomials: Function

**Definition 51 (Polynomial function; Dt.: Polynomfunktion)**

A (univariate real) function $f : I \rightarrow \mathbb{R}$, for an interval $I \subseteq \mathbb{R}$, is a *polynomial function* over $I$ if there exist $n \in \mathbb{N}_0$ and $a_0, a_1, \ldots, a_n \in \mathbb{R}$ such that

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad \text{for all } x \in I.$$

- As usual, two (polynomial) functions over an interval $I \subseteq \mathbb{R}$ are identical if their values are identical for all arguments in $I$.
- Note: Two different polynomials may result in the same polynomial function! (E.g., over finite fields.)
- While some fields of mathematics (e.g., abstract algebra) make a clear distinction between polynomials and polynomial functions, we will freely mix these two terms. Also, unless noted explicitly, we will only deal with polynomials over $\mathbb{R}$. 

Note: Polynomial functions may come in disguise: $f(x) := \cos(2 \arccos(x))$ is a polynomial function over $[-1, 1]$ since we have $f(x) = 2x^2 - 1$ for all $x \in [-1, 1]$. 

Polynomials: Function

Definition 51 (Polynomial function; Dt.: Polynomfunktion)

A (univariate real) function \( f : I \rightarrow \mathbb{R} \), for an interval \( I \subseteq \mathbb{R} \), is a polynomial function over \( I \) if there exist \( n \in \mathbb{N}_0 \) and \( a_0, a_1, \ldots, a_n \in \mathbb{R} \) such that

\[
f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad \text{for all } x \in I.
\]

- As usual, two (polynomial) functions over an interval \( I \subseteq \mathbb{R} \) are identical if their values are identical for all arguments in \( I \).
- Note: Two different polynomials may result in the same polynomial function! (E.g., over finite fields.)
- While some fields of mathematics (e.g., abstract algebra) make a clear distinction between polynomials and polynomial functions, we will freely mix these two terms. Also, unless noted explicitly, we will only deal with polynomials over \( \mathbb{R} \).
- Note: Polynomial functions may come in disguise: \( f(x) := \cos(2 \arccos(x)) \) is a polynomial function over \([-1, 1]\) since we have \( f(x) = 2x^2 - 1 \) for all \( x \in [-1, 1] \).
Consider a polynomial of degree \( n \) with coefficients \( a_0, a_1, \ldots, a_n \in \mathbb{R} \), with \( a_n \neq 0 \):

\[
p(x) := \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1} + a_n x^n.
\]
Polynomial Evaluation: Horner’s Algorithm

- Consider a polynomial of degree $n$ with coefficients $a_0, a_1, \ldots, a_n \in \mathbb{R}$, with $a_n \neq 0$:

\[ p(x) := \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1} + a_n x^n. \]

- A straightforward polynomial evaluation of $p$ for a given parameter $x_0$ results in $k$ multiplications for a monomial of degree $k$, plus a total of $n$ additions.
- Hence, we would get

\[ 0 + 1 + 2 + \ldots + n = \frac{n(n + 1)}{2} \]

multiplications (and $n$ additions).
Polynomial Evaluation: Horner’s Algorithm

- Consider a polynomial of degree $n$ with coefficients $a_0, a_1, \ldots, a_n \in \mathbb{R}$, with $a_n \neq 0$:

$$p(x) := \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1} + a_n x^n.$$  

- A straightforward polynomial evaluation of $p$ for a given parameter $x_0$ results in $k$ multiplications for a monomial of degree $k$, plus a total of $n$ additions.

- Hence, we would get

$$0 + 1 + 2 + \ldots + n = \frac{n(n+1)}{2}$$

multiplications (and $n$ additions).

- Obviously, we can reduce the number of multiplications to $O(n \log n)$ by resorting to exponentiation by squaring:

$$x^n = \begin{cases}  
  x \left( x^2 \right)^{\frac{n-1}{2}} & \text{if } n \text{ is odd}, \\
  \left( x^2 \right)^{\frac{n}{2}} & \text{if } n \text{ is even}.
\end{cases}$$
Polynomial Evaluation: Horner’s Algorithm

- Consider a polynomial of degree $n$ with coefficients $a_0, a_1, \ldots, a_n \in \mathbb{R}$, with $a_n \neq 0$:
  \[ p(x) := \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1} + a_n x^n. \]

- A straightforward polynomial evaluation of $p$ for a given parameter $x_0$ results in $k$ multiplications for a monomial of degree $k$, plus a total of $n$ additions.

- Hence, we would get
  \[ 0 + 1 + 2 + \ldots + n = \frac{n(n+1)}{2} \]
  multiplications (and $n$ additions).

- Obviously, we can reduce the number of multiplications to $O(n \log n)$ by resorting to exponentiation by squaring:
  \[ x^n = \begin{cases} 
  x \left( x^2 \right)^{\frac{n-1}{2}} & \text{if } n \text{ is odd}, \\
  (x^2)^{\frac{n}{2}} & \text{if } n \text{ is even}. 
  \end{cases} \]

- Can we do even better?
Polynomial Evaluation: Horner’s Algorithm

- **Horner’s Algorithm**: The idea is to rewrite the polynomial such that

\[ p(x) = a_0 + x \left( a_1 + x (a_2 + \ldots + x (a_{n-1} + x a_n) \ldots) \right) \]
Polynomial Evaluation: Horner’s Algorithm

- **Horner’s Algorithm**: The idea is to rewrite the polynomial such that

\[
p(x) = a_0 + x\left(a_1 + x(a_2 + \ldots + x(a_{n-1} + x a_n)\ldots)\right)
\]

and compute the result \( h_0 = p(x_0) \) as follows:

\[
\begin{align*}
    h_n &:= a_n \\
    h_i &:= x_0 \cdot h_{i+1} + a_i \\
    & \quad \text{for } i = 0, 1, 2, \ldots, n - 1
\end{align*}
\]
Polynomial Evaluation: Horner’s Algorithm

- **Horner’s Algorithm**: The idea is to rewrite the polynomial such that

\[ p(x) = a_0 + x \left( a_1 + x \left( a_2 + \ldots + x \left( a_{n-1} + x \ a_n \right) \ldots \right) \right) \]

and compute the result \( h_0 = p(x_0) \) as follows:

\[
\begin{align*}
    h_n &:= a_n \\
    h_i &:= x_0 \cdot h_{i+1} + a_i & \text{for } i = 0, 1, 2, \ldots, n - 1
\end{align*}
\]

**Lemma 52**

Horner’s Algorithm consumes \( n \) multiplications and \( n \) additions to evaluate a polynomial of degree \( n \).
Polynomial Evaluation: Horner’s Algorithm

- **Horner’s Algorithm**: The idea is to rewrite the polynomial such that

\[ p(x) = a_0 + x \left( a_1 + x \left( a_2 + \ldots + x (a_{n-1} + x a_n) \ldots \right) \right) \]

and compute the result \( h_0 = p(x_0) \) as follows:

\[
\begin{align*}
    h_n &:= a_n \\
    h_i &:= x_0 \cdot h_{i+1} + a_i \\
    \text{for } i &= 0, 1, 2, \ldots, n - 1
\end{align*}
\]

**Lemma 52**

Horner’s Algorithm consumes \( n \) multiplications and \( n \) additions to evaluate a polynomial of degree \( n \).

**Caveat**

Subtractive cancellation could occur at any time, and there is no easy way to determine a priori whether and for which data it will indeed occur.
Basic Linear Algebra

- Matrices
- Linear Equations
- Determinants
- Eigenvalues and Eigenvectors
- Dot Product and Norm
- Vector Cross-Product
- Quaternions
Basic Linear Algebra

- Matrices
  - Basic Definitions
  - Matrix Algebra
  - Inversion and Transpose
  - Special Matrices
  - Fast Matrix Multiplication
- Linear Equations
- Determinants
- Eigenvalues and Eigenvectors
- Dot Product and Norm
- Vector Cross-Product
- Quaternions
Matrices

**Definition 53 (Matrix)**

For $m, n \in \mathbb{N}$, an $m \times n$ matrix $A$ is a scheme of $m \cdot n$ numbers $a_{ij}$ from a field $F$, with $1 \leq i \leq m, 1 \leq j \leq n$, arranged as follows:

$$A := \begin{pmatrix}
    a_{11} & \cdots & a_{1n} \\
    \vdots & \ddots & \vdots \\
    a_{m1} & \cdots & a_{mn}
\end{pmatrix}$$
Matrices

**Definition 53 (Matrix)**

For $m, n \in \mathbb{N}$, an $m \times n$ matrix $A$ is a scheme of $m \cdot n$ numbers $a_{ij}$ from a field $F$, with $1 \leq i \leq m, 1 \leq j \leq n$, arranged as follows:

$$A := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

The numbers $a_{ij}$ are called the *coefficients* of the matrix $A$. 

© M. Held (Univ. Salzburg)  Geometrisches Rechnen  (WS 2020/21)
Matrices

**Definition 53 (Matrix)**

For $m, n \in \mathbb{N}$, an $m \times n$ matrix $A$ is a scheme of $m \cdot n$ numbers $a_{ij}$ from a field $F$, with $1 \leq i \leq m, 1 \leq j \leq n$, arranged as follows:

$$A := \begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn}
\end{pmatrix}$$

The numbers $a_{ij}$ are called the *coefficients* of the matrix $A$. The $m$ horizontal $n$-tuples $(a_{i1} \cdots a_{in})$ are called *rows* of the matrix, while the $n$ vertical $m$-tuples $(a_{ij} \cdots a_{mj})$ are called *columns* of the matrix.
Matrices

Definition 53 (Matrix)

For \( m, n \in \mathbb{N} \), an \( m \times n \) matrix \( \mathbf{A} \) is a scheme of \( m \cdot n \) numbers \( a_{ij} \) from a field \( F \), with \( 1 \leq i \leq m, 1 \leq j \leq n \), arranged as follows:

\[
\mathbf{A} := \begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn}
\end{pmatrix}
\]

The numbers \( a_{ij} \) are called the coefficients of the matrix \( \mathbf{A} \). The \( m \) horizontal \( n \)-tuples \((a_{i1} \cdots a_{in})\) are called rows of the matrix, while the \( n \) vertical \( m \)-tuples \((a_{ij} \cdots a_{mj})\) are called columns of the matrix.

- The collection of all \( m \times n \) matrices over \( F \) is denoted by \( M_{m \times n}(F) \), or simply by \( M_{m \times n} \) if the field is obvious or irrelevant. Short-hand notation: \( \mathbf{A} = [a_{ij}]_{i=1,j=1}^{m,n} \), or simply \( \mathbf{A} = [a_{ij}] \).
**Definition 53 (Matrix)**

For $m, n \in \mathbb{N}$, an $m \times n$ matrix $A$ is a scheme of $m \cdot n$ numbers $a_{ij}$ from a field $F$, with $1 \leq i \leq m, 1 \leq j \leq n$, arranged as follows:

$$
A := \begin{pmatrix}
    a_{11} & \cdots & a_{1n} \\
    \vdots & \ddots & \vdots \\
    a_{m1} & \cdots & a_{mn}
\end{pmatrix}
$$

The numbers $a_{ij}$ are called the *coefficients* of the matrix $A$.

The $m$ horizontal $n$-tuples $(a_{i1} \cdots a_{in})$ are called *rows* of the matrix, while the $n$ vertical $m$-tuples $(a_{ij} \cdots a_{mj})$ are called *columns* of the matrix.

- The collection of all $m \times n$ matrices over $F$ is denoted by $M_{m \times n}(F)$, or simply by $M_{m \times n}$ if the field is obvious or irrelevant. Short-hand notation: $A = [a_{ij}]_{i=1,j=1}^{m,n}$, or simply $A = [a_{ij}]$.

**Definition 54 (Size)**

The numbers $m$ and $n$ in Def. 53 describe the *size* of the matrix $A$. 
**Definition 53 (Matrix)**

For $m, n \in \mathbb{N}$, an $m \times n$ matrix $A$ is a scheme of $m \cdot n$ numbers $a_{ij}$ from a field $F$, with $1 \leq i \leq m$, $1 \leq j \leq n$, arranged as follows:

$$A := \begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn}
\end{pmatrix}$$

The numbers $a_{ij}$ are called the **coefficients** of the matrix $A$. The $m$ horizontal $n$-tuples $(a_{i1} \cdots a_{in})$ are called **rows** of the matrix, while the $n$ vertical $m$-tuples $(a_{ij} \cdots a_{mj})$ are called **columns** of the matrix.

- The collection of all $m \times n$ matrices over $F$ is denoted by $M_{m\times n}(F)$, or simply by $M_{m\times n}$ if the field is obvious or irrelevant. Short-hand notation: $A = [a_{ij}]_{i=1,j=1}^{m,n}$, or simply $A = [a_{ij}]$.

**Definition 54 (Size)**

The numbers $m$ and $n$ in Def. 53 describe the **size** of the matrix $A$. The matrix $A$ is **square** if $m = n$. 
Definition 55 (Zero matrix, Dt.: Nullmatrix)

For $m, n \in \mathbb{N}$, the matrix in $M_{m \times n}(F)$ with all elements equal to 0 is called the *zero matrix* (of size $m \times n$), and is denoted by the symbol $0$. 
Matrices

Definition 55 (Zero matrix, Dt.: Nullmatrix)
For $m, n \in \mathbb{N}$, the matrix in $M_{m \times n}(F)$ with all elements equal to 0 is called the zero matrix (of size $m \times n$), and is denoted by the symbol $0$.

- E.g., for $4 \times 4$ matrices we have

$$0 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}$$
Matrices

Definition 55 (Zero matrix, Dt.: Nullmatrix)

For \( m, n \in \mathbb{N} \), the matrix in \( M_{m \times n}(F) \) with all elements equal to 0 is called the zero matrix (of size \( m \times n \)), and is denoted by the symbol \( 0 \).

Definition 56 (Identity matrix, Dt.: Einheitsmatrix)

For \( n \in \mathbb{N} \), the \( n \times n \) matrix \( I := [\delta_{ij}] \), defined by \( \delta_{ij} := 1 \) if \( i = j \) and \( \delta_{ij} := 0 \) otherwise, is called the \( n \times n \) identity matrix.

- E.g., for \( 4 \times 4 \) matrices we have

\[
0 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\quad \text{and} \quad
I = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}.
\]
Matrices

Definition 55 (Zero matrix, Dt.: Nullmatrix)
For $m, n \in \mathbb{N}$, the matrix in $M_{m \times n}(F)$ with all elements equal to 0 is called the zero matrix (of size $m \times n$), and is denoted by the symbol $0$.

Definition 56 (Identity matrix, Dt.: Einheitsmatrix)
For $n \in \mathbb{N}$, the $n \times n$ matrix $I := [\delta_{ij}]$, defined by $\delta_{ij} := 1$ if $i = j$ and $\delta_{ij} := 0$ otherwise, is called the $n \times n$ identity matrix.

- Of course, the elements 0 and 1 are the additive and multiplicative neutral elements of $F$.
- E.g., for $4 \times 4$ matrices we have

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

and

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$
Definition 57 (Matrix identity)

Two matrices $A$ and $B$ over the same field $F$ are said to be equal if $A$ and $B$ have the same size and if corresponding elements are equal; that is, $A, B \in M_{m \times n}(F)$ and $A = [a_{ij}]$, $B = [b_{ij}]$, with $a_{ij} = b_{ij}$ for $1 \leq i \leq m, 1 \leq j \leq n$. 
Definition 57 (Matrix identity)

Two matrices $A$ and $B$ over the same field $F$ are said to be equal if $A$ and $B$ have the same size and if corresponding elements are equal; that is, $A, B \in M_{m \times n}(F)$ and $A = [a_{ij}], B = [b_{ij}]$, with $a_{ij} = b_{ij}$ for $1 \leq i \leq m, 1 \leq j \leq n$.

Definition 58 (Sparse, Dt.: dünn besetzt)

For $m, n \in \mathbb{N}$, the $m \times n$ matrix $A$ is called sparse if $k \ll m \cdot n$ holds for the number $k$ of non-zero coefficients of $A$. 
Definition 57 (Matrix identity)

Two matrices $\mathbf{A}$ and $\mathbf{B}$ over the same field $F$ are said to be equal if $\mathbf{A}$ and $\mathbf{B}$ have the same size and if corresponding elements are equal; that is, $\mathbf{A}, \mathbf{B} \in M_{m \times n}(F)$ and $\mathbf{A} = [a_{ij}], \mathbf{B} = [b_{ij}]$, with $a_{ij} = b_{ij}$ for $1 \leq i \leq m, 1 \leq j \leq n$.

Definition 58 (Sparse, Dt.: dünn besetzt)

For $m, n \in \mathbb{N}$, the $m \times n$ matrix $\mathbf{A}$ is called sparse if $k \ll m \cdot n$ holds for the number $k$ of non-zero coefficients of $\mathbf{A}$.

- Note: Storing an $n \times n$ matrix consumes $O(n^2)$ space, unless special precautions are taken (e.g., in the case of sparse matrices)!
### Definition 59 (Matrix addition)

Let \( A, B \in M_{m \times n}(F) \) be matrices of the same size. Then \( A + B \) is the matrix obtained by adding corresponding elements of \( A \) and \( B \); that is,

\[
A + B = [a_{ij}] + [b_{ij}] := \begin{pmatrix}
  a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\
a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn}
\end{pmatrix}.
\]
Matrix Algebra

**Definition 59 (Matrix addition)**

Let \( A, B \in M_{m \times n}(F) \) be matrices of the same size. Then \( A + B \) is the matrix obtained by adding corresponding elements of \( A \) and \( B \); that is,

\[
A + B = [a_{ij}] + [b_{ij}] := \begin{pmatrix}
  a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\
  a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn}
\end{pmatrix}.
\]

**Definition 60 (Scalar multiplication)**

Consider a matrix \( A \in M_{m \times n}(F) \) and \( \lambda \in F \). (Thus, \( \lambda \) is a scalar.) Then \( \lambda A \) is the matrix obtained by multiplying all elements of \( A \) by \( \lambda \); that is,

\[
\lambda A = \lambda [a_{ij}] := \begin{pmatrix}
  \lambda a_{11} & \cdots & \lambda a_{1n} \\
  \lambda a_{21} & \cdots & \lambda a_{2n} \\
  \vdots & \ddots & \vdots \\
  \lambda a_{m1} & \cdots & \lambda a_{mn}
\end{pmatrix}.
\]
Matrix Algebra

**Theorem 61**

\[ M_{m \times n}(F), \text{ with addition and scalar multiplication as defined in Defs. 59+60, forms a vector space over } F \text{ for all } m, n \in \mathbb{N}. \]
Theorem 61

$M_{m \times n}(F)$, with addition and scalar multiplication as defined in Defs. 59+60, forms a vector space over $F$ for all $m, n \in \mathbb{N}$.

Definition 62 (Additive inverse)

Consider a matrix $A \in M_{m \times n}(F)$. Then

$$
-A = [-a_{ij}] := \begin{pmatrix}
-a_{11} & \cdots & -a_{1n} \\
-a_{21} & \cdots & -a_{2n} \\
\vdots & \ddots & \vdots \\
-a_{m1} & \cdots & -a_{mn}
\end{pmatrix}
$$

is taken as the additive inverse of $A$. 
Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, $A$, $B$, $C$ are matrices of the same size over the same field $F$, and $\lambda$, $\mu$ are scalars out of $F$.)
Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, \( A, B, C \) are matrices of the same size over the same field \( F \), and \( \lambda, \mu \) are scalars out of \( F \).)

1. **Associativity:** \( (A + B) + C = A + (B + C) \);
Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, \( A, B, C \) are matrices of the same size over the same field \( F \), and \( \lambda, \mu \) are scalars out of \( F \).)

1. Associativity: \((A + B) + C = A + (B + C)\);
2. Commutativity: \( A + B = B + A \);
Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, $A$, $B$, $C$ are matrices of the same size over the same field $F$, and $\lambda$, $\mu$ are scalars out of $F$.)

1. Associativity: $(A + B) + C = A + (B + C)$;
2. Commutativity: $A + B = B + A$;
3. Neutral element: $0 + A = A$;
Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, A, B, C are matrices of the same size over the same field $F$, and $\lambda, \mu$ are scalars out of $F$.)

1. **Associativity:** $(A + B) + C = A + (B + C)$;
2. **Commutativity:** $A + B = B + A$;
3. **Neutral element:** $0 + A = A$;
4. **Inverse element:** $A + (-A) = 0$;
Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are matrices of the same size over the same field $F$, and $\lambda$, $\mu$ are scalars out of $F$.)

1. Associativity: $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$;
2. Commutativity: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$;
3. Neutral element: $\mathbf{0} + \mathbf{A} = \mathbf{A}$;
4. Inverse element: $\mathbf{A} + (\mathbf{A}) = \mathbf{0}$;
5. Distributivity: $(\lambda + \mu)\mathbf{A} = \lambda\mathbf{A} + \mu\mathbf{A}$;
6. Distributivity: $\lambda(\mathbf{A} + \mathbf{B}) = \lambda\mathbf{A} + \lambda\mathbf{B}$;
Matrix Algebra

Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, $A$, $B$, $C$ are matrices of the same size over the same field $F$, and $\lambda$, $\mu$ are scalars out of $F$.)

1. Associativity: $(A + B) + C = A + (B + C)$;
2. Commutativity: $A + B = B + A$;
3. Neutral element: $0 + A = A$;
4. Inverse element: $A + (−A) = 0$;
5. Distributivity: $(\lambda + \mu)A = \lambda A + \mu A$;
6. Distributivity: $\lambda(A + B) = \lambda A + \lambda B$;
7. Associativity: $\lambda(\mu A) = (\lambda \mu)A$;
Matrix Algebra

Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, \( A, B, C \) are matrices of the same size over the same field \( F \), and \( \lambda, \mu \) are scalars out of \( F \).)

1. **Associativity:** \((A + B) + C = A + (B + C)\);
2. **Commutativity:** \(A + B = B + A\);
3. **Neutral element:** \(0 + A = A\);
4. **Inverse element:** \(A + (-A) = 0\);
5. **Distributivity:** \((\lambda + \mu)A = \lambda A + \mu A\);
6. **Distributivity:** \(\lambda(A + B) = \lambda A + \lambda B\);
7. **Associativity:** \(\lambda(\mu A) = (\lambda \mu)A\);
8. **1 times matrix:** \(1A = A\);
9. **0 times matrix:** \(0A = 0\);
10. **Negation:** \((-1)A = -A\).
Matrix Algebra

Lemma 63

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, $A$, $B$, $C$ are matrices of the same size over the same field $F$, and $\lambda$, $\mu$ are scalars out of $F$.)

1. **Associativity:** $(A + B) + C = A + (B + C)$;
2. **Commutativity:** $A + B = B + A$;
3. **Neutral element:** $0 + A = A$;
4. **Inverse element:** $A + (-A) = 0$;
5. **Distributivity:** $(\lambda + \mu)A = \lambda A + \mu A$;
6. **Distributivity:** $\lambda(A + B) = \lambda A + \lambda B$;
7. **Associativity:** $\lambda(\mu A) = (\lambda \mu)A$;
8. $1A = A$;
9. $0A = 0$;
10. $(-1)A = -A$;
11. $\lambda A = 0 \Rightarrow \lambda = 0$ or $A = 0$. 
Definition 64 (Matrix multiplication)

Let $A$ be a matrix of size $m \times n$ and $B$ be a matrix of size $n \times p$ over the same field $F$; that is, the number of columns of $A$ equals the number of rows of $B$. Then $A \cdot B$, or $AB$ for sake of brevity, is the $m \times p$ matrix $C = [c_{ik}]$ whose $(i, k)$-th element is defined as

$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + \cdots + a_{in} b_{nk}.$$
Matrix Algebra

**Definition 64 (Matrix multiplication)**

Let $A$ be a matrix of size $m \times n$ and $B$ be a matrix of size $n \times p$ over the same field $F$; that is, the number of columns of $A$ equals the number of rows of $B$. Then $A \cdot B$, or $AB$ for sake of brevity, is the $m \times p$ matrix $C = [c_{ik}]$ whose $(i, k)$-th element is defined as

$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + \cdots + a_{in} b_{nk}.$$

**Lemma 65**

Matrix multiplication obeys the standard laws of arithmetic except for commutativity:

1. $(AB)C = A(BC)$ if $A, B, C$ are $m \times n, n \times p, p \times q$, respectively;
Matrix Algebra

**Definition 64 (Matrix multiplication)**

Let \( A \) be a matrix of size \( m \times n \) and \( B \) be a matrix of size \( n \times p \) over the same field \( F \); that is, the number of columns of \( A \) equals the number of rows of \( B \). Then \( A \cdot B \), or \( AB \) for sake of brevity, is the \( m \times p \) matrix \( C = [c_{ik}] \) whose \((i,k)\)-th element is defined as

\[
c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + \cdots + a_{in} b_{nk}.
\]

**Lemma 65**

Matrix multiplication obeys the standard laws of arithmetic except for commutativity:

1. \((AB)C = A(BC)\) if \( A, B, C \) are \( m \times n, n \times p, p \times q \), respectively;
2. \(\lambda(AB) = (\lambda A)B = A(\lambda B)\) if \( A, B \) are \( m \times n, n \times p \), respectively;
Definition 64 (Matrix multiplication)

Let $A$ be a matrix of size $m \times n$ and $B$ be a matrix of size $n \times p$ over the same field $F$; that is, the number of columns of $A$ equals the number of rows of $B$. Then $A \cdot B$, or $AB$ for sake of brevity, is the $m \times p$ matrix $C = [c_{ik}]$ whose $(i, k)$-th element is defined as

$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk}.$$ 

Lemma 65

Matrix multiplication obeys the standard laws of arithmetic except for commutativity:

1. $(AB)C = A(BC)$ if $A$, $B$, $C$ are $m \times n$, $n \times p$, $p \times q$, respectively;
2. $\lambda(AB) = (\lambda A)B = A(\lambda B)$ if $A$, $B$ are $m \times n$, $n \times p$, respectively;
3. $A(−B) = (−A)B = −(AB)$ if $A$, $B$ are $m \times n$, $n \times p$, respectively;
Matrix Algebra

Definition 64 (Matrix multiplication)
Let \( A \) be a matrix of size \( m \times n \) and \( B \) be a matrix of size \( n \times p \) over the same field \( F \); that is, the number of columns of \( A \) equals the number of rows of \( B \). Then \( A \cdot B \), or \( AB \) for sake of brevity, is the \( m \times p \) matrix \( C = [c_{ik}] \) whose \((i, k)\)-th element is defined as

\[
c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + \cdots + a_{in} b_{nk}.
\]

Lemma 65
Matrix multiplication obeys the standard laws of arithmetic except for commutativity:

1. \((AB)C = A(BC)\) if \( A, B, C \) are \( m \times n, n \times p, p \times q \), respectively;
2. \(\lambda(AB) = (\lambda A)B = A(\lambda B)\) if \( A, B \) are \( m \times n, n \times p \), respectively;
3. \(A(-B) = (-A)B = -(AB)\) if \( A, B \) are \( m \times n, n \times p \), respectively;
4. \((A + B)C = AC + BC\) if \( A, B \) are \( m \times n \) and \( C \) is \( n \times p \);
5. \(D(A + B) = DA + DB\) if \( A, B \) are \( m \times n \) and \( D \) is \( p \times m \).
Matrix Algebra

**Definition 64 (Matrix multiplication)**
Let $A$ be a matrix of size $m \times n$ and $B$ be a matrix of size $n \times p$ over the same field $F$; that is, the number of columns of $A$ equals the number of rows of $B$. Then $A \cdot B$, or $AB$ for sake of brevity, is the $m \times p$ matrix $C = [c_{ik}]$ whose $(i, k)$-th element is defined as

$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + \cdots + a_{in} b_{nk}.$$

**Lemma 65**
Matrix multiplication obeys the standard laws of arithmetic except for commutativity:

1. $(AB)C = A(BC)$ if $A$, $B$, $C$ are $m \times n$, $n \times p$, $p \times q$, respectively;
2. $\lambda(AB) = (\lambda A)B = A(\lambda B)$ if $A$, $B$ are $m \times n$, $n \times p$, respectively;
3. $A(−B) = (−A)B = −(AB)$ if $A$, $B$ are $m \times n$, $n \times p$, respectively;
4. $(A + B)C = AC + BC$ if $A$, $B$ are $m \times n$ and $C$ is $n \times p$;
5. $D(A + B) = DA + DB$ if $A$, $B$ are $m \times n$ and $D$ is $p \times m$.

- Note: $AB \neq BA$ even if $A$, $B$ are square. Also, $AB = 0 \not\Rightarrow [A = 0 \text{ or } B = 0]$. 
Definition 66 (Invertible, Dt.: invertierbar)

An \( n \times n \) matrix \( A \) is \textit{invertible} (or \textit{non-singular}) if there exists an \( n \times n \) matrix \( B \) such that

\[
AB = BA = I.
\]
Inversion of a Matrix

**Definition 66 (Invertible, Dt.: invertierbar)**

An $n \times n$ matrix $A$ is *invertible* (or *non-singular*) if there exists an $n \times n$ matrix $B$ such that

$$AB = BA = I.$$ 

If $A$ is invertible then the inverse matrix is denoted by $A^{-1}$. 

---

**Theorem 67**

If $A$ has inverse matrices $B$, $C$ then $B = C$.

Note that $A^{-1}$ can be obtained (if it exists) by solving $Ax_i = e_i$ for $1 \leq i \leq n$; the vectors $x_i$ form the columns of $A^{-1}$. 

**Theorem 68**

If $A$, $B$ are non-singular matrices of the same size then $AB$ is not singular, and $(AB)^{-1} = B^{-1}A^{-1}$, i.e., the inverse of the product equals the product of the inverses in the reverse order.
Definition 66 (Invertible, Dt.: invertierbar)

An \( n \times n \) matrix \( A \) is invertible (or non-singular) if there exists an \( n \times n \) matrix \( B \) such that

\[
AB = BA = I.
\]

If \( A \) is invertible then the inverse matrix is denoted by \( A^{-1} \).

Theorem 67

If \( A \) has inverse matrices \( B, C \) then \( B = C \).
Inversion of a Matrix

Definition 66 (Invertible, Dt.: invertierbar)

An $n \times n$ matrix $A$ is *invertible* (or *non-singular*) if there exists an $n \times n$ matrix $B$ such that

$$AB = BA = I.$$

If $A$ is invertible then the inverse matrix is denoted by $A^{-1}$.

Theorem 67

If $A$ has inverse matrices $B, C$ then $B = C$.

- Note that $A^{-1}$ can be obtained (if it exists) by solving $Ax_i = e_i$ for $1 \leq i \leq n$; the vectors $x_i$ form the columns of $A^{-1}$. 

Inversion of a Matrix

**Definition 66 (Invertible, Dt.: invertierbar)**

An \( n \times n \) matrix \( A \) is *invertible* (or *non-singular*) if there exists an \( n \times n \) matrix \( B \) such that

\[
AB = BA = I.
\]

If \( A \) is invertible then the inverse matrix is denoted by \( A^{-1} \).

**Theorem 67**

If \( A \) has inverse matrices \( B, C \) then \( B = C \).

- Note that \( A^{-1} \) can be obtained (if it exists) by solving \( Ax_i = e_i \) for \( 1 \leq i \leq n \); the vectors \( x_i \) form the columns of \( A^{-1} \).

**Theorem 68**

If \( A, B \) are non-singular matrices of the same size then \( AB \) is not singular, and

\[
(AB)^{-1} = B^{-1}A^{-1},
\]

i.e., the inverse of the product equals the product of the inverses in the reverse order.
Definition 69 (Transpose, Dt.: transponiert)

Consider an $m \times n$ matrix $A$. The *transpose* of $A$ is the matrix $A^t$ obtained by interchanging the rows and columns of $A$. 

Consequently, $A^t$ is an $n \times m$ matrix:

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}^t
= \begin{bmatrix}
1 & 4 \\
2 & 5 \\
3 & 6
\end{bmatrix}
\]

Lemma 70

The transpose operation has the following properties for all matrices $A$, $B$ of suitable sizes:

1. $(A^t)^t = A$
2. $(A + B)^t = A^t + B^t$
3. $(\lambda A)^t = \lambda A^t$ for a scalar $\lambda$
4. $(AB)^t = B^t A^t$
5. If $A$ is non-singular then $A^t$ is also non-singular and we have $(A^t)^{-1} = (A^{-1})^t$. 

Definition 69 (Transpose, Dt.: transponiert)

Consider an $m \times n$ matrix $A$. The transpose of $A$ is the matrix $A^t$ obtained by interchanging the rows and columns of $A$.

Consequently, $A^t$ is an $n \times m$ matrix:
Transpose of a Matrix

**Definition 69 (Transpose, Dt.: transponiert)**

Consider an \( m \times n \) matrix \( A \). The *transpose* of \( A \) is the matrix \( A^t \) obtained by interchanging the rows and columns of \( A \).

- Consequently, \( A^t \) is an \( n \times m \) matrix: 
  \[
  \begin{pmatrix} 
  1 & 2 & 3 \\
  4 & 5 & 6 \\
  \end{pmatrix}^t = 
  \begin{pmatrix} 
  1 & 4 \\
  2 & 5 \\
  3 & 6 \\
  \end{pmatrix}.
  \]
Transpose of a Matrix

Definition 69 (Transpose, Dt.: transponiert)

Consider an \( m \times n \) matrix \( A \). The \textit{transpose} of \( A \) is the matrix \( A^t \) obtained by interchanging the rows and columns of \( A \).

- Consequently, \( A^t \) is an \( n \times m \) matrix:

\[
\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{pmatrix}^t =
\begin{pmatrix}
1 & 4 \\
2 & 5 \\
3 & 6
\end{pmatrix}.
\]

Lemma 70

The transpose operation has the following properties for all matrices \( A, B \) of suitable sizes:

1. \((A^t)^t = A\);
Transpose of a Matrix

**Definition 69 (Transpose, Dt.: transponiert)**

Consider an $m \times n$ matrix $A$. The *transpose* of $A$ is the matrix $A^t$ obtained by interchanging the rows and columns of $A$.

Consequently, $A^t$ is an $n \times m$ matrix:

$$
\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{pmatrix}^t =
\begin{pmatrix}
1 & 4 \\
2 & 5 \\
3 & 6
\end{pmatrix}.
$$

**Lemma 70**

The transpose operation has the following properties for all matrices $A, B$ of suitable sizes:

1. $(A^t)^t = A$;
2. $(A + B)^t = A^t + B^t$;
Transpose of a Matrix

Definition 69 (Transpose, Dt.: transponiert)

Consider an $m \times n$ matrix $A$. The transpose of $A$ is the matrix $A^t$ obtained by interchanging the rows and columns of $A$.

Consequently, $A^t$ is an $n \times m$ matrix:

$$
\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{pmatrix}^t =
\begin{pmatrix}
1 & 4 \\
2 & 5 \\
3 & 6
\end{pmatrix}.
$$

Lemma 70

The transpose operation has the following properties for all matrices $A, B$ of suitable sizes:

1. $(A^t)^t = A$;
2. $(A + B)^t = A^t + B^t$;
3. $(\lambda A)^t = \lambda A^t$ for a scalar $\lambda$. 

**Transpose of a Matrix**

### Definition 69 (Transpose, Dt.: transponiert)

Consider an $m \times n$ matrix $A$. The *transpose* of $A$ is the matrix $A^t$ obtained by interchanging the rows and columns of $A$.

- Consequently, $A^t$ is an $n \times m$ matrix: 
  \[
  \begin{pmatrix}
  1 & 2 & 3 \\
  4 & 5 & 6 \\
  \end{pmatrix}^t = \begin{pmatrix}
  1 & 4 \\
  2 & 5 \\
  3 & 6 \\
  \end{pmatrix}.
  \]

### Lemma 70

The transpose operation has the following properties for all matrices $A, B$ of suitable sizes:

1. $(A^t)^t = A$;
2. $(A + B)^t = A^t + B^t$;
3. $(\lambda A)^t = \lambda A^t$ for a scalar $\lambda$.
4. $(AB)^t = B^t A^t$;
Transpoe of a Matrix

**Definition 69 (Transpose, Dt.: transponiert)**

Consider an $m \times n$ matrix $A$. The *transpose* of $A$ is the matrix $A^t$ obtained by interchanging the rows and columns of $A$.

- Consequently, $A^t$ is an $n \times m$ matrix: 
  $$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

**Lemma 70**

The transpose operation has the following properties for all matrices $A, B$ of suitable sizes:

1. $(A^t)^t = A$;
2. $(A + B)^t = A^t + B^t$;
3. $(\lambda A)^t = \lambda A^t$ for a scalar $\lambda$.
4. $(AB)^t = B^t A^t$;
5. If $A$ is non-singular then $A^t$ is also non-singular and we have $(A^t)^{-1} = (A^{-1})^t$. 
Definition 71 (Symmetric, Dt.: symmetrisch)

A matrix $A$ is called \textit{symmetric} if $A^t = A$. 

Definition 72 (Diagonal matrix, Dt.: Diagonalmatrix)

A square matrix $A$ is called \textit{diagonal} if $a_{ij} = 0$ for $i \neq j$.

Definition 73 (Upper-triangular, Dt.: obere Dreiecksmatrix)

A square matrix $A$ is called \textit{upper-triangular} if $a_{ij} = 0$ for $i > j$.

Definition 74 (Orthogonal, Dt.: orthogonal)

A square matrix $A$ is called \textit{orthogonal} if $A \cdot A^t = I = A^t \cdot A$.

Lemma 75

If a square matrix $A$ is orthogonal then $A^{-1} = A^t$. 

Special Matrices

Definition 71 (Symmetric, Dt.: symmetrisch)

A matrix $A$ is called symmetric if $A^t = A$.

Definition 72 (Diagonal matrix, Dt.: Diagonalmatrix)

A square matrix $A$ is called diagonal if $a_{ij} = 0$ for $i \neq j$.
Special Matrices

Definition 71 (Symmetric, Dt.: symmetrisch)
A matrix $A$ is called symmetric if $A^t = A$.

Definition 72 (Diagonal matrix, Dt.: Diagonalmatrix)
A square matrix $A$ is called diagonal if $a_{ij} = 0$ for $i \neq j$.

Definition 73 (Upper-triangular, Dt.: obere Dreiecksmatrix)
A square matrix $A$ is called upper-triangular if $a_{ij} = 0$ for $i > j$. 
Special Matrices

Definition 71 (Symmetric, Dt.: symmetrisch)
A matrix $A$ is called symmetric if $A^t = A$.

Definition 72 (Diagonal matrix, Dt.: Diagonalmatrix)
A square matrix $A$ is called diagonal if $a_{ij} = 0$ for $i \neq j$.

Definition 73 (Upper-triangular, Dt.: obere Dreiecksmatrix)
A square matrix $A$ is called upper-triangular if $a_{ij} = 0$ for $i > j$.

Definition 74 (Orthogonal, Dt.: orthogonal)
A square matrix $A$ is called orthogonal if $A \cdot A^t = I = A^t \cdot A$. 
Special Matrices

Definition 71 (Symmetric, Dt.: symmetrisch)
A matrix $A$ is called symmetric if $A^t = A$.

Definition 72 (Diagonal matrix, Dt.: Diagonalmatrix)
A square matrix $A$ is called diagonal if $a_{ij} = 0$ for $i \neq j$.

Definition 73 (Upper-triangular, Dt.: obere Dreiecksmatrix)
A square matrix $A$ is called upper-triangular if $a_{ij} = 0$ for $i > j$.

Definition 74 (Orthogonal, Dt.: orthogonal)
A square matrix $A$ is called orthogonal if $A \cdot A^t = I = A^t \cdot A$.

Lemma 75
If a square matrix $A$ is orthogonal then $A^{-1} = A^t$. 
Block Matrices

Definition 76 (Block matrix)

Let \( m, n \in \mathbb{N} \) and \( A, B, C, D \in M_{m \times n}(F) \). Then the \( 2m \times 2n \) matrix \( X \) with

\[
x_{i,j} := \begin{cases} 
    a_{i,j} & \text{if } 1 \leq i \leq m, 1 \leq j \leq n, \\
    b_{i,j-n} & \text{if } 1 \leq i \leq m, n+1 \leq j \leq 2n, \\
    c_{i-m,j} & \text{if } m+1 \leq i \leq 2m, 1 \leq j \leq n, \\
    d_{i-m,j-n} & \text{if } m+1 \leq i \leq 2m, n+1 \leq j \leq 2n 
\end{cases}
\]

is the block matrix with component matrices \( A, B, C, D \).
Definition 76 (Block matrix)

Let \( m, n \in \mathbb{N} \) and \( A, B, C, D \in M_{m \times n}(F) \). Then the \( 2m \times 2n \) matrix \( X \) with

\[
x_{i,j} := \begin{cases} 
  a_{i,j} & \text{if } 1 \leq i \leq m, 1 \leq j \leq n, \\
  b_{i,j-n} & \text{if } 1 \leq i \leq m, n+1 \leq j \leq 2n, \\
  c_{i-m,j} & \text{if } m+1 \leq i \leq 2m, 1 \leq j \leq n, \\
  d_{i-m,j-n} & \text{if } m+1 \leq i \leq 2m, n+1 \leq j \leq 2n 
\end{cases}
\]

is the block matrix with component matrices \( A, B, C, D \).

It is common to regard \( A, B, C, D \) as “coefficients” of \( X \) and write

\[
X = \begin{pmatrix} 
  A & B \\
  C & D 
\end{pmatrix},
\]

or simply

\[
X = \begin{pmatrix} 
  A & B \\
  C & D 
\end{pmatrix}.
\]
Lemma 77

For $m, n, p \in \mathbb{N}$, let $A_{11}, A_{12}, A_{21}, A_{22} \in M_{m \times n}(F)$, $B_{11}, B_{12}, B_{21}, B_{22} \in M_{n \times p}(F)$, and

$$A := \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$
Lemma 77

For $m, n, p \in \mathbb{N}$, let $A_{11}, A_{12}, A_{21}, A_{22} \in M_{m \times n}(F)$, $B_{11}, B_{12}, B_{21}, B_{22} \in M_{n \times p}(F)$, and

$$A := \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$ 

Then

$$A \cdot B = \begin{pmatrix} A_{11} \cdot B_{11} + A_{12} \cdot B_{21} & A_{11} \cdot B_{12} + A_{12} \cdot B_{12} \\ A_{21} \cdot B_{11} + A_{22} \cdot B_{21} & A_{21} \cdot B_{12} + A_{22} \cdot B_{22} \end{pmatrix}.$$
Lemma 78

Let $m, n \in \mathbb{N}$ and $A, B, D \in M_{m \times n}(F)$. Then the $2m \times 2n$ matrix $X$ with

$$X := \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

is invertible if and only if $A$ and $D$ are invertible. In this case we get

$$X^{-1} = \begin{pmatrix} A^{-1} & -A^{-1} \cdot B \cdot D^{-1} \\ 0 & D^{-1} \end{pmatrix}.$$
Fast Matrix Multiplication

- Standard multiplication of two $n \times n$ matrices results in $O(n^3)$ many multiplications.
Fast Matrix Multiplication

- Standard multiplication of two $n \times n$ matrices results in $O(n^3)$ many multiplications.

**Theorem 79 (Strassen 1969)**

Seven multiplications of scalars suffice to compute the multiplication of two $2 \times 2$ matrices. In general, $O(n^{\log_2 7}) \approx O(n^{2.807\ldots})$ arithmetic operations suffice for $n \times n$ matrices.
Fast Matrix Multiplication

- Standard multiplication of two $n \times n$ matrices results in $O(n^3)$ many multiplications.

**Theorem 79 (Strassen 1969)**

Seven multiplications of scalars suffice to compute the multiplication of two $2 \times 2$ matrices. In general, $O(n^{\log_2 7}) \approx O(n^{2.807\ldots})$ arithmetic operations suffice for $n \times n$ matrices.

- Strassen’s algorithm is more complex and numerically less stable than the standard naïve algorithm. But it is considerably more efficient for large $n$, i.e., roughly when $n > 100$, and it is very useful for large matrices over finite fields.
- It does not assume multiplication to be commutative and, thus, works over arbitrary rings.
Fast Matrix Multiplication

- Standard multiplication of two $n \times n$ matrices results in $O(n^3)$ many multiplications.

**Theorem 79 (Strassen 1969)**

Seven multiplications of scalars suffice to compute the multiplication of two $2 \times 2$ matrices. In general, $O(n^{\log_2 7}) \approx O(n^{2.807\ldots})$ arithmetic operations suffice for $n \times n$ matrices.

- Strassen’s algorithm is more complex and numerically less stable than the standard naïve algorithm. But it is considerably more efficient for large $n$, i.e., roughly when $n > 100$, and it is very useful for large matrices over finite fields.
- It does not assume multiplication to be commutative and, thus, works over arbitrary rings.

**Theorem 80 (Coppersmith&Winograd 1990)**

$O(n^{2.375\ldots})$ arithmetic operations suffice for multiplying two $n \times n$ matrices.


$O(n^{2.37286\ldots})$ arithmetic operations suffice for multiplying two $n \times n$ matrices.
Fast Matrix Multiplication

- Standard multiplication of two $n \times n$ matrices results in $O(n^3)$ many multiplications.

**Theorem 79 (Strassen 1969)**

Seven multiplications of scalars suffice to compute the multiplication of two $2 \times 2$ matrices. In general, $O(n^{\log_2 7}) \approx O(n^{2.807\ldots})$ arithmetic operations suffice for $n \times n$ matrices.

- Strassen’s algorithm is more complex and numerically less stable than the standard naïve algorithm. But it is considerably more efficient for large $n$, i.e., roughly when $n > 100$, and it is very useful for large matrices over finite fields.
- It does not assume multiplication to be commutative and, thus, works over arbitrary rings.

**Theorem 80 (Coppersmith&Winograd 1990)**

$O(n^{2.375\ldots})$ arithmetic operations suffice for multiplying two $n \times n$ matrices.


$O(n^{2.37286\ldots})$ arithmetic operations suffice for multiplying two $n \times n$ matrices.
Fast Matrix Multiplication

Strassen: For $A, B \in M_{2 \times 2}$, we compute $C = A \cdot B$ via

$$p_1 := (a_{1,2} - a_{2,2})(b_{2,1} + b_{2,2})$$
$$p_2 := (a_{1,1} + a_{2,2})(b_{1,1} + b_{2,2})$$
$$p_3 := (a_{1,1} - a_{2,1})(b_{1,1} + b_{1,2})$$
$$p_4 := (a_{1,1} + a_{1,2})b_{2,2}$$
$$p_5 := a_{1,1}(b_{1,2} - b_{2,2})$$
$$p_6 := a_{2,2}(b_{2,1} - b_{1,1})$$
$$p_7 := (a_{2,1} + a_{2,2})b_{1,1}$$
Fast Matrix Multiplication

- **Strassen**: For $A, B \in M_{2 \times 2}$, we compute $C = A \cdot B$ via

  
  
  
  
  
  
  
  
  and set

  
  
  
  
  
  
  
  
  
  
  
  
  
  
  Use block matrices to apply this concept for $n > 2$. 

Basic Linear Algebra

- Matrices
- Linear Equations
  - Linear Equations and Matrices
  - Solving Systems of Linear Equations
  - Gauss-Jordan Algorithm
  - Application: Bernstein Polynomials as Basis
- Determinants
- Eigenvalues and Eigenvectors
- Dot Product and Norm
- Vector Cross-Product
- Quaternions
Definition 82 (Linear equation, Dt.: lineare Gleichung)

A *linear equation* in $n$ unknowns $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b,$$

where $a_1, \ldots, a_n, b$ are given (real) numbers.
Linear Equations

**Definition 82 (Linear equation, Dt.: lineare Gleichung)**

A *linear equation* in $n$ unknowns $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b,$$

where $a_1, \ldots, a_n, b$ are given (real) numbers.

**Definition 83 (System of linear equations, Dt.: lineares Gleichungssystem)**

A *system of $m$ linear equations* in $n$ unknowns $x_1, x_2, \ldots, x_n$ is a family of linear equations

$$a_{11} x_1 + \cdots + a_{1n} x_n = b_1,$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$a_{m1} x_1 + \cdots + a_{mn} x_n = b_m,$$

where $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m$ are given (real) numbers.
**Linear Equations**

**Definition 82 (Linear equation, Dt.: lineare Gleichung)**

A *linear equation* in \( n \) unknowns \( x_1, x_2, \ldots, x_n \) is an equation of the form

\[
a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b,
\]

where \( a_1, \ldots, a_n, b \) are given (real) numbers.

**Definition 83 (System of linear equations, Dt.: lineares Gleichungssystem)**

A *system of \( m \) linear equations* in \( n \) unknowns \( x_1, x_2, \ldots, x_n \) is a family of linear equations

\[
a_{11} x_1 + \cdots + a_{1n} x_n = b_1, \\
\vdots \quad \vdots \\
a_{m1} x_1 + \cdots + a_{mn} x_n = b_m,
\]

where \( a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m \) are given (real) numbers.

The system is called *homogeneous* if \( b_1 = b_2 = \cdots = b_m = 0 \).
Matrices and Linear Equations

- Of course, a system of \( m \) linear equations in \( n \) unknowns \( x_1, x_2, \ldots, x_n \),

\[
\begin{align*}
    a_{11}x_1 &+ a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
    a_{21}x_1 &+ a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
    \vdots &+ \vdots + \vdots \quad \vdots \\
    a_{m1}x_1 &+ a_{m2}x_2 + \cdots + a_{mn}x_n = b_m
\end{align*}
\]
Matrices and Linear Equations

Of course, a system of \( m \) linear equations in \( n \) unknowns \( x_1, x_2, \ldots, x_n \),

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
    \vdots & \vdots \vdots \vdots \\
    a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{align*}
\]

can also be seen as one vector-valued equation:

\[
\begin{pmatrix}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
    \vdots & \vdots \vdots \vdots \\
    a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\end{pmatrix}
= 
\begin{pmatrix}
    b_1 \\
    b_2 \\
    \vdots \\
    b_m
\end{pmatrix}
\]
Matrices and Linear Equations

- Of course, a system of $m$ linear equations in $n$ unknowns $x_1, x_2, \ldots, x_n$,
  
  \[ a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n = b_1 \\
  a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n = b_2 \\
  \vdots \quad \vdots \quad \vdots \quad \vdots \\
  a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n = b_m 
  
  can also be seen as one vector-valued equation:

  \[
  \begin{pmatrix}
  a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n \\
  a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n \\
  \vdots \\
  a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n 
  \end{pmatrix}
  =
  \begin{pmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m 
  \end{pmatrix}
  
  With $A := [a_{ij}]_{i=1,j=1}^{m,n}$, $b := (b_1, \ldots, b_m) \in \mathbb{R}^m$ and $x := (x_1, \ldots, x_n) \in \mathbb{R}^n$,
Matrices and Linear Equations

- Of course, a system of \( m \) linear equations in \( n \) unknowns \( x_1, x_2, \ldots, x_n \),
  
  \[
  \begin{align*}
  a_{11} x_1 & + a_{12} x_2 & + \cdots & + a_{1n} x_n & = b_1 \\
  a_{21} x_1 & + a_{22} x_2 & + \cdots & + a_{2n} x_n & = b_2 \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  a_{m1} x_1 & + a_{m2} x_2 & + \cdots & + a_{mn} x_n & = b_m
  \end{align*}
  \]

  can also be seen as one vector-valued equation:

  \[
  \begin{pmatrix}
  a_{11} x_1 & + a_{12} x_2 & + \cdots & + a_{1n} x_n \\
  a_{21} x_1 & + a_{22} x_2 & + \cdots & + a_{2n} x_n \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} x_1 & + a_{m2} x_2 & + \cdots & + a_{mn} x_n 
  \end{pmatrix}
  \begin{pmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
  \end{pmatrix}
  =
  \begin{pmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
  \end{pmatrix}
  \]

- With \( A := [a_{ij}]_{i=1,j=1}^{m,n}, \ b := (b_1, \ldots, b_m) \in \mathbb{R}^m \) and \( x := (x_1, \ldots, x_n) \in \mathbb{R}^n \), this system can be written concisely as \( Ax = b \):

  \[
  b = 
  \begin{pmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
  \end{pmatrix}
  = 
  \begin{pmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
  \end{pmatrix}
  \begin{pmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
  \end{pmatrix}
  = Ax
  \]
Matrices and Linear Equations

So, we have

\[
A \mathbf{x} = \begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
a_{21} & \cdots & a_{2n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn}
\end{pmatrix} \cdot \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix} = \begin{pmatrix}
b_1 \\
b_2 \\
\vdots \\
b_m
\end{pmatrix} = b.
\]
Matrices and Linear Equations

So, we have

\[
A \mathbf{x} = \begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
a_{21} & \cdots & a_{2n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn}
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix} = \begin{pmatrix}
b_1 \\
b_2 \\
\vdots \\
b_m
\end{pmatrix} = \mathbf{b}.
\]

The matrix

\[
\begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
a_{21} & \cdots & a_{2n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn}
\end{pmatrix}
\]

is called the coefficient matrix of the system.
Matrices and Linear Equations

So, we have

\[ A \mathbf{x} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = b. \]

The matrix

\[ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \]

is called the coefficient matrix of the system.

The matrix

\[ \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix} \]

is called the augmented matrix of the system.
A system of $m$ linear equations in $n$ unknowns can be interpreted as follows:

- We seek the intersection of $m$ lines (for $n = 2$) or hyper-planes (for $n > 2$) in $\mathbb{R}^n$, where the $i$-th line/plane is given by the equation

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i.$$  

See Slide 156.
A system of $m$ linear equations in $n$ unknowns can be interpreted as follows:

- We seek the intersection of $m$ lines (for $n = 2$) or hyper-planes (for $n > 2$) in $\mathbb{R}^n$, where the $i$-th line/plane is given by the equation

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i.$$

See Slide 156.

- We regard the $m \times n$ matrix $A$ as a transformation matrix and seek that vector $x \in \mathbb{R}^n$ which gets mapped to the vector $b \in \mathbb{R}^m$:

$$Ax = b$$

See Slide 223.
A system of linear equations in $n$ unknowns is called *consistent* if it has a solution, i.e., if there exist (real) numbers $x_1, x_2, \ldots, x_n$ that satisfy all equations simultaneously.
Solutions of Linear Equations

**Definition 84**

A system of linear equations in \( n \) unknowns is called *consistent* if it has a solution, i.e., if there exist (real) numbers \( x_1, x_2, \ldots, x_n \) that satisfy all equations simultaneously.

- A homogeneous system is always consistent, since \( x_1 = x_2 = \cdots = x_n = 0 \)
  always is a solution, which is called *trivial* solution. Any other solution of a homogeneous system is called a *non-trivial* solution.
Solutions of Linear Equations

**Definition 84**

A system of linear equations in $n$ unknowns is called *consistent* if it has a solution, i.e., if there exist (real) numbers $x_1, x_2, \ldots, x_n$ that satisfy all equations simultaneously.

- A homogeneous system is always consistent, since $x_1 = x_2 = \cdots = x_n = 0$ always is a solution, which is called *trivial* solution. Any other solution of a homogeneous system is called a *non-trivial* solution.

**Theorem 85**

A homogeneous system of $m$ linear equations in $n$ unknowns always has a non-trivial solution if $m < n$. 
Definition 86 (Rank, Dt.: Rang)

The (column) rank of a matrix $A$, denoted by $\text{rank}(A)$, is the number of linearly independent columns of $A$. 
Solutions of Linear Equations

**Definition 86 (Rank, Dt.: Rang)**

The (column) *rank* of a matrix $A$, denoted by $\text{rank}(A)$, is the number of linearly independent columns of $A$.

**Theorem 87**

The system $Ax = b$ is consistent if and only if the rank of the coefficient matrix equals the rank of the augmented matrix.
Solutions of Linear Equations

Definition 86 (Rank, Dt.: Rang)
The (column) rank of a matrix $A$, denoted by $\text{rank}(A)$, is the number of linearly independent columns of $A$.

Theorem 87
The system $Ax = b$ is consistent if and only if the rank of the coefficient matrix equals the rank of the augmented matrix.

Theorem 88
Assume that the system $Ax = b$ is consistent. This system has a unique solution if and only if the number of linearly independent columns of the coefficient matrix (i.e., its rank) equals the number of unknowns.
Elementary Row Operations

Lemma 89

The following three types of elementary row operations may be performed on a matrix without changing its rank:

1. Interchanging two rows;
Lemma 89

The following three types of *elementary row operations* may be performed on a matrix without changing its rank:

1. Interchanging two rows;
2. Multiplying a row by a nonzero scalar;
Lemma 89

The following three types of *elementary row operations* may be performed on a matrix without changing its rank:

1. Interchanging two rows;
2. Multiplying a row by a nonzero scalar;
3. Adding a multiple of one row to another row.
Elementary Row Operations

Lemma 89

The following three types of elementary row operations may be performed on a matrix without changing its rank:

1. Interchanging two rows;
2. Multiplying a row by a nonzero scalar;
3. Adding a multiple of one row to another row.

Definition 90

A matrix $A$ is row-equivalent to a matrix $B$ if $B$ is obtained from $A$ by a sequence of elementary row operations.
Lemma 89

The following three types of elementary row operations may be performed on a matrix without changing its rank:

1. Interchanging two rows;
2. Multiplying a row by a nonzero scalar;
3. Adding a multiple of one row to another row.

Definition 90

A matrix $A$ is row-equivalent to a matrix $B$ if $B$ is obtained from $A$ by a sequence of elementary row operations.

Theorem 91

If $A$ and $B$ are row-equivalent augmented matrices of two systems of linear equations, then the two systems have the same solution sets.
Definition 92 (Reduced row-echelon form, Dt.: Treppennormalform)

A matrix is in *reduced row-echelon form* if

1. all zero rows (if any) are at the bottom of the matrix;
2. if two successive rows are nonzero then the second row starts with more zeros than the first (moving from left to right and top to bottom);
3. the leading (leftmost nonzero) entry in each nonzero row is 1;
4. all other elements of the column in which the leading entry 1 occurs are zeros.
Definition 92 (Reduced row-echelon form, Dt.: Treppennormalform)

A matrix is in *reduced row-echelon form* if

1. all zero rows (if any) are at the bottom of the matrix;
2. if two successive rows are nonzero then the second row starts with more zeros than the first (moving from left to right and top to bottom);
Elementary Row Operations

Definition 92 (Reduced row-echelon form, Dt.: Treppennormalform)

A matrix is in reduced row-echelon form if

1. all zero rows (if any) are at the bottom of the matrix;
2. if two successive rows are nonzero then the second row starts with more zeros than the first (moving from left to right and top to bottom);
3. the leading (leftmost nonzero) entry in each nonzero row is 1;
Elementary Row Operations

**Definition 92 (Reduced row-echelon form, Dt.: Treppennormalform)**

A matrix is in *reduced row-echelon form* if

1. all zero rows (if any) are at the bottom of the matrix;
2. if two successive rows are nonzero then the second row starts with more zeros than the first (moving from left to right and top to bottom);
3. the leading (leftmost nonzero) entry in each nonzero row is 1;
4. all other elements of the column in which the leading entry 1 occurs are zeros.
Elementary Row Operations

Definition 92 (Reduced row-echelon form, Dt.: Treppennormalform)

A matrix is in reduced row-echelon form if

1. all zero rows (if any) are at the bottom of the matrix;
2. if two successive rows are nonzero then the second row starts with more zeros than the first (moving from left to right and top to bottom);
3. the leading (leftmost nonzero) entry in each nonzero row is 1;
4. all other elements of the column in which the leading entry 1 occurs are zeros.

Sample matrix in reduced row-echelon form:

\[
\begin{pmatrix}
0 & 1 & * & 0 & 0 & * & * & 0 & * \\
0 & 0 & 0 & 1 & 0 & * & * & 0 & * \\
0 & 0 & 0 & 0 & 1 & * & * & 0 & * \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix $A$ into a matrix $A'$ that is in reduced row-echelon form, using elementary row operations:

$$
\begin{align*}
&k := 1. \\
&\text{If the rows } k, \ldots, m \text{ all are zero then the matrix is in reduced row-echelon form.} \\
&\text{Otherwise, suppose that the first column which has a non-zero element in the rows below the first } k-1 \text{ rows is column } c_k. \\
&\text{By interchanging the rows below the first } k-1 \text{ rows, if necessary, we ensure that the element } a_{k,c_k} \text{ is non-zero.} \\
&\text{Convert } a_{k,c_k} \text{ to 1. By adding suitable multiples of row } k \text{ to the remaining rows, where necessary, we ensure that all remaining elements in column } c_k \text{ are zero.} \\
&\text{If } k < m, \text{ repeat this process for } k := k + 1. \\
&\text{This process will eventually stop after } r \text{ steps, either because we run out of rows (if } k = m), \text{ or because we run out of non-zero columns.} \\
&\text{In general, the final matrix } A' \text{ will be in reduced row-echelon form and will have } r \text{ non-zero rows, with leading entries } 1 \text{ in columns } c_1, \ldots, c_r, \text{ respectively.} \\
&\text{By swapping columns (and updating the solution vector } x \text{ accordingly) we can guarantee that the } r \text{ non-zero rows have their leading 1's in columns } 1, \ldots, r. \\
\end{align*}
$$
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix \( A \) into a matrix \( A' \) that is in reduced row-echelon form, using elementary row operations:
  - Initially, \( k := 1 \).
  - If the rows \( k, \ldots, m \) all are zero then the matrix is in reduced row-echelon form.

By swapping columns (and updating the solution vector \( x \) accordingly) we can guarantee that the \( r \) non-zero rows have their leading 1's in columns 1, \ldots, \( r \).
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix $A$ into a matrix $A'$ that is in reduced row-echelon form, using elementary row operations:
  - Initially, $k := 1$.
  - If the rows $k, \ldots, m$ all are zero then the matrix is in reduced row-echelon form.
  - Otherwise, suppose that the first column which has a non-zero element in the rows below the first $k - 1$ rows is column $c_k$.
    - By interchanging the rows below the first $k - 1$ rows, if necessary, we ensure that the element $a_{k, c_k}$ is non-zero.
    - Convert $a_{k, c_k}$ to 1. By adding suitable multiples of row $k$ to the remaining rows, where necessary, we ensure that all remaining elements in column $c_k$ are zero.
    - If $k < m$, repeat this process for $k := k + 1$.

This process will eventually stop after $r$ steps, either because we run out of rows (if $k = m$), or because we run out of non-zero columns. In general, the final matrix $A'$ will be in reduced row-echelon form and will have $r$ non-zero rows, with leading entries 1 in columns $c_1, \ldots, c_r$, respectively.

By swapping columns (and updating the solution vector $x$ accordingly) we can guarantee that the $r$ non-zero rows have their leading 1's in columns 1, $\ldots$, $r$. 
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix $A$ into a matrix $A'$ that is in reduced row-echelon form, using elementary row operations:
  - Initially, $k := 1$.
  - If the rows $k, \ldots, m$ all are zero then the matrix is in reduced row-echelon form.
  - Otherwise, suppose that the first column which has a non-zero element in the rows below the first $k - 1$ rows is column $c_k$. By interchanging the rows below the first $k - 1$ rows, if necessary, we ensure that the element $a_{k,c_k}$ is nonzero.
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix $A$ into a matrix $A'$ that is in reduced row-echelon form, using elementary row operations:
  - Initially, $k := 1$.
  - If the rows $k, \ldots, m$ all are zero then the matrix is in reduced row-echelon form.
  - Otherwise, suppose that the first column which has a non-zero element in the rows below the first $k - 1$ rows is column $c_k$. By interchanging the rows below the first $k - 1$ rows, if necessary, we ensure that the element $a_{k,c_k}$ is nonzero. Convert $a_{k,c_k}$ to 1. By adding suitable multiples of row $k$ to the remaining rows, where necessary, we ensure that all remaining elements in column $c_k$ are zero.
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix $A$ into a matrix $A'$ that is in reduced row-echelon form, using elementary row operations:
  - Initially, $k := 1$.
  - If the rows $k, \ldots, m$ all are zero then the matrix is in reduced row-echelon form.
  - Otherwise, suppose that the first column which has a non-zero element in the rows below the first $k - 1$ rows is column $c_k$. By interchanging the rows below the first $k - 1$ rows, if necessary, we ensure that the element $a_{k,c_k}$ is nonzero. Convert $a_{k,c_k}$ to 1. By adding suitable multiples of row $k$ to the remaining rows, where necessary, we ensure that all remaining elements in column $c_k$ are zero.
  - If $k < m$, repeat this process for $k := k + 1$. 

This process will eventually stop after $r$ steps, either because we run out of rows (if $k = m$), or because we run out of non-zero columns. In general, the final matrix $A'$ will be in reduced row-echelon form and will have $r$ non-zero rows, with leading entries 1 in columns $c_1, \ldots, c_r$, respectively. By swapping columns (and updating the solution vector $x$ accordingly) we can guarantee that the $r$ non-zero rows have their leading 1's in columns 1, \ldots, $r$.
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix $A$ into a matrix $A'$ that is in reduced row-echelon form, using elementary row operations:
  - Initially, $k := 1$.
  - If the rows $k, \ldots, m$ all are zero then the matrix is in reduced row-echelon form.
  - Otherwise, suppose that the first column which has a non-zero element in the rows below the first $k - 1$ rows is column $c_k$. By interchanging the rows below the first $k - 1$ rows, if necessary, we ensure that the element $a_{k,c_k}$ is nonzero. Convert $a_{k,c_k}$ to 1. By adding suitable multiples of row $k$ to the remaining rows, where necessary, we ensure that all remaining elements in column $c_k$ are zero.
  - If $k < m$, repeat this process for $k := k + 1$.
  - This process will eventually stop after $r$ steps, either because we run out of rows (if $k = m$), or because we run out of non-zero columns.
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix $A$ into a matrix $A'$ that is in reduced row-echelon form, using elementary row operations:
  - Initially, $k := 1$.
  - If the rows $k, \ldots, m$ all are zero then the matrix is in reduced row-echelon form.
  - Otherwise, suppose that the first column which has a non-zero element in the rows below the first $k - 1$ rows is column $c_k$. By interchanging the rows below the first $k - 1$ rows, if necessary, we ensure that the element $a_{k,c_k}$ is nonzero. Convert $a_{k,c_k}$ to 1. By adding suitable multiples of row $k$ to the remaining rows, where necessary, we ensure that all remaining elements in column $c_k$ are zero.
  - If $k < m$, repeat this process for $k := k + 1$.
- This process will eventually stop after $r$ steps, either because we run out of rows (if $k = m$), or because we run out of non-zero columns.
- In general, the final matrix $A'$ will be in reduced row-echelon form and will have $r$ non-zero rows, with leading entries 1 in columns $c_1, \ldots, c_r$, respectively.
Gauss-Jordan Algorithm

- The following algorithm transforms an augmented matrix $A$ into a matrix $A'$ that is in reduced row-echelon form, using elementary row operations:
  - Initially, $k := 1$.
  - If the rows $k, \ldots, m$ all are zero then the matrix is in reduced row-echelon form.
  - Otherwise, suppose that the first column which has a non-zero element in the rows below the first $k - 1$ rows is column $c_k$. By interchanging the rows below the first $k - 1$ rows, if necessary, we ensure that the element $a_{k,c_k}$ is nonzero. Convert $a_{k,c_k}$ to 1. By adding suitable multiples of row $k$ to the remaining rows, where necessary, we ensure that all remaining elements in column $c_k$ are zero.
  - If $k < m$, repeat this process for $k := k + 1$.
  - This process will eventually stop after $r$ steps, either because we run out of rows (if $k = m$), or because we run out of non-zero columns.
  - In general, the final matrix $A'$ will be in reduced row-echelon form and will have $r$ non-zero rows, with leading entries 1 in columns $c_1, \ldots, c_r$, respectively.
  - By swapping columns (and updating the solution vector $x$ accordingly) we can guarantee that the $r$ non-zero rows have their leading 1's in columns 1, $\ldots$, $r$. 
Gauss-Jordan Algorithm

Thus, the Gauss-Jordan algorithm transforms an augmented matrix \( A \) into a matrix \( A' \) of the following form:

\[
\begin{pmatrix}
1 & 0 & a'_{1,r+1} & \cdots & a'_{1n} & b'_1 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 1 & a'_{r,r+1} & \cdots & a'_{rn} & b'_r \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & \cdots & 0 & b'_{r+1} \\
\end{pmatrix}
\]

If \( r = n + 1 \) then the system is inconsistent. (The last row reads 0 \( \cdot x'_{1} + 0 \cdot x'_{2} + \cdots + 0 \cdot x'_{n} = 1, \) which has no solutions.) If \( r \leq n \) then the system is inconsistent unless \( b'_{r+1} = b'_{r+2} = \cdots = b'_m = 0. \) If \( r = n \) and \( b'_{r+1} = b'_{r+2} = \cdots = b'_m = 0, \) then there exists a unique solution \( x'_{1} = b'_1, x'_{2} = b'_2, \ldots, x'_{n} = b'_n. \)
Gauss-Jordan Algorithm

Thus, the Gauss-Jordan algorithm transforms an augmented matrix $A$ into a matrix $A'$ of the following form:

$$
\begin{pmatrix}
1 & 0 & a'_{1,r+1} & \cdots & a'_{1n} & b'_1 \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 1 & a'_{r,r+1} & \cdots & a'_{rn} & b'_r \\
\end{pmatrix}
$$

If $r = n + 1$ then the system is inconsistent. (The last row reads $0 \cdot x'_1 + 0 \cdot x'_2 + \ldots + 0 \cdot x'_n = 1$, which has no solutions.)
Gauss-Jordan Algorithm

Thus, the Gauss-Jordan algorithm transforms an augmented matrix $\mathbf{A}$ into a matrix $\mathbf{A'}$ of the following form:

$$
\begin{pmatrix}
1 & 0 & a'_{1,r+1} & \cdots & a'_{1n} & b'_1 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 1 & a'_{r,r+1} & \cdots & a'_{rn} & b'_r \\
& & & 0 & b'_{r+1} & \vdots \\
\end{pmatrix}
$$

- If $r = n + 1$ then the system is inconsistent. (The last row reads $0 \cdot x'_1 + 0 \cdot x'_2 + \ldots + 0 \cdot x'_n = 1$, which has no solutions.)
- If $r \leq n$ then the system is inconsistent unless $b'_{r+1} = b'_{r+2} = \ldots = b'_m = 0$. 
Gauss-Jordan Algorithm

Thus, the Gauss-Jordan algorithm transforms an augmented matrix $A$ into a matrix $A'$ of the following form:

$$
\begin{pmatrix}
1 & 0 & a'_{1,r+1} & \cdots & a'_{1n} & b'_1 \\
\vdots & & \vdots & & \vdots & \vdots \\
0 & 1 & a'_{r,r+1} & \cdots & a'_{rn} & b'_r \\
& & & & & b'_{r+1} \\
0 & & & & & b'_m
\end{pmatrix}
$$

If $r = n + 1$ then the system is inconsistent. (The last row reads $0 \cdot x'_1 + 0 \cdot x'_2 + \ldots + 0 \cdot x'_n = 1$, which has no solutions.)

If $r \leq n$ then the system is inconsistent unless $b'_{r+1} = b'_{r+2} = \ldots = b'_m = 0$.

If $r = n$ and $b'_{r+1} = b'_{r+2} = \ldots = b'_m = 0$, then there exists a unique solution $x'_1 = b'_1, x'_2 = b'_2, \ldots, x'_n = b'_n$. 
Gauss-Jordan Algorithm

- Thus, the Gauss-Jordan algorithm transforms an augmented matrix $A$ into a matrix $A'$ of the following form:

\[
\begin{pmatrix}
1 & 0 & a'_{1,r+1} & \cdots & a'_{1,n} & b'_1 \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 1 & a'_{r,r+1} & \cdots & a'_{r,n} & b'_r \\
\end{pmatrix}
\]

- If $r < n$ and $b'_{r+1} = b'_{r+2} = \ldots = b'_m = 0$, then there are infinitely many solutions:

\[
\begin{align*}
x'_1 &= b'_1 - a'_{1,r+1}x'_{r+1} - a'_{1,r+2}x'_{r+2} - \ldots - a'_{1,n}x'_n, \\
\vdots \\
x'_r &= b'_r - a'_{r,r+1}x'_{r+1} - a'_{r,r+2}x'_{r+2} - \ldots - a'_{r,n}x'_n.
\end{align*}
\]

The independent unknowns $x'_{r+1}, \ldots, x'_n$ may take on arbitrary values.
Sample Linear System

\[
\begin{align*}
  x_1 &+ x_2 + 2x_3 + 3x_4 = 4 \\
  2x_1 &+ 2x_2 + 3x_3 + 4x_4 = 5
\end{align*}
\]
Sample Linear System

\[
\begin{align*}
\left\{ \begin{array}{cccccc}
  x_1 & + & x_2 & + & 2x_3 & + & 3x_4 & = & 4 \\
  2x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 5
\end{array} \right. \\
\end{align*}
\]

\[(A|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} \]
Sample Linear System

\[
\begin{cases}
    x_1 + x_2 + 2x_3 + 3x_4 = 4 \\
    2x_1 + 2x_2 + 3x_3 + 4x_4 = 5
\end{cases}
\]

\[
(A|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} + I \cdot (-2) \sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix}
\]
Sample Linear System

\[
\begin{align*}
\begin{cases}
    x_1 + x_2 + 2x_3 + 3x_4 &= 4 \\
    2x_1 + 2x_2 + 3x_3 + 4x_4 &= 5
\end{cases}
\end{align*}
\]

\[
(A|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} + I \cdot (-2) \sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix} \cdot (-1)
\]

\[
\sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix}
\]
Sample Linear System

\[
\begin{align*}
\begin{cases}
x_1 + x_2 + 2x_3 + 3x_4 &= 4 \\
2x_1 + 2x_2 + 3x_3 + 4x_4 &= 5
\end{cases}
\end{align*}
\]

\[
(A|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} + I \cdot (-2) \sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix} \cdot (-1)
\]

\[
\sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix} x_2 \leftrightarrow x_3 \sim \begin{pmatrix} 1 & 2 & 1 & 3 & 4 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix}
\]
Sample Linear System

\[
\begin{align*}
\{ & x_1 + x_2 + 2x_3 + 3x_4 = 4 \\
& 2x_1 + 2x_2 + 3x_3 + 4x_4 = 5
\end{align*}
\]

\[(A|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} + I \cdot (-2) \sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix} \cdot (-1) \]

\[
\begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix} \leftrightarrow x_2 \leftrightarrow x_3 \sim \begin{pmatrix} 1 & 2 & 1 & 3 & 4 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} + II \cdot (-2) \]

\[
\begin{pmatrix} 1 & 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix}
\]
Sample Linear System

\[
\begin{align*}
\begin{cases}
 x_1 + x_2 + 2x_3 + 3x_4 &= 4 \\
 2x_1 + 2x_2 + 3x_3 + 4x_4 &= 5
\end{cases}
\end{align*}
\]

\[
(A|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} + I \cdot (-2) \sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix} \cdot (-1)
\]

\[
\begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix} x_2 \leftrightarrow x_3 \sim \begin{pmatrix} 1 & 2 & 1 & 3 & 4 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} + II \cdot (-2)
\]

\[
\begin{pmatrix} 1 & 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} \sim \begin{cases}
 x_1 + x_2 - x_4 = -2 \\
 x_3 + 2x_4 = 3
\end{cases}
\]
Sample Linear System

\[
\begin{align*}
\left\{ \begin{array}{l}
x_1 + x_2 + 2x_3 + 3x_4 = 4 \\
2x_1 + 2x_2 + 3x_3 + 4x_4 = 5
\end{array} \right.
\end{align*}
\]

\[
(A|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} + I \cdot (-2) \sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix} \cdot (-1)
\]

\[
\sim \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix} x_2 \leftrightarrow x_3 \sim \begin{pmatrix} 1 & 2 & 1 & 3 & 4 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} + II \cdot (-2)
\]

\[
\sim \begin{pmatrix} 1 & 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} \sim \left\{ \begin{array}{l}
x_1 + x_2 - x_4 = -2 \\
x_3 + 2x_4 = 3
\end{array} \right.
\]

\[
\sim \left\{ \begin{array}{l}
x_1 = -2 - x_2 + x_4 \\
x_3 = 3 - 2x_4
\end{array} \right.
\]
Sample Linear System

\[
\begin{align*}
\begin{cases}
  x_1 + x_2 + 2x_3 + 3x_4 &= 4 \\
  2x_1 + 2x_2 + 3x_3 + 4x_4 &= 5 
\end{cases}
\end{align*}
\]

\[
(A|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} + I \cdot (-2) \quad \sim \quad \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix} \cdot (-1)
\]

\[
\begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix} x_2 \leftrightarrow x_3 \quad \sim \quad \begin{pmatrix} 1 & 2 & 1 & 3 & 4 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} + ll \cdot (-2)
\]

\[
\begin{pmatrix} 1 & 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} \quad \sim \quad \begin{cases} x_1 + x_2 - x_4 = -2 \\
  x_3 + 2x_4 = 3
\end{cases}
\]

\[
\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \end{pmatrix} - 2x_4
\]

\[
\sim \quad \text{Solution:} \quad \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} : \lambda_1, \lambda_2 \in \mathbb{R}
\]
Application: Bernstein Polynomials as Basis

Proof of Theorem 44 for $n:=3$:

- The four Bernstein polynomials are given by

\[ B_{0,3}(x) := (1-x)^3 \quad B_{1,3}(x) := 3x(1-x)^2 \quad B_{2,3}(x) := 3x^2(1-x) \quad B_{3,3}(x) := x^3. \]
Application: Bernstein Polynomials as Basis

Proof of Theorem 44 for $n:=3$:

- The four Bernstein polynomials are given by

\[ B_{0,3}(x) := (1-x)^3 \quad B_{1,3}(x) := 3x(1-x)^2 \quad B_{2,3}(x) := 3x^2(1-x) \quad B_{3,3}(x) := x^3. \]

- We get the following relation:

\[
\begin{pmatrix}
1 & -3 & 3 & -1 \\
0 & 3 & -6 & 3 \\
0 & 0 & 3 & -3 \\
0 & 0 & 0 & 1 \\
\end{pmatrix}
\cdot
\begin{pmatrix}
1 \\
x \\
x^2 \\
x^3 \\
\end{pmatrix}
=
\begin{pmatrix}
B_{0,3}(x) \\
B_{1,3}(x) \\
B_{2,3}(x) \\
B_{3,3}(x) \\
\end{pmatrix}
\]
Application: Bernstein Polynomials as Basis

Proof of Theorem 44 for \( n := 3 \):

- The four Bernstein polynomials are given by
  \[
  B_{0,3}(x) := (1-x)^3 \quad B_{1,3}(x) := 3x(1-x)^2 \quad B_{2,3}(x) := 3x^2(1-x) \quad B_{3,3}(x) := x^3.
  \]

- We get the following relation:
  \[
  \begin{pmatrix}
  1 & -3 & 3 & -1 \\
  0 & 3 & -6 & 3 \\
  0 & 0 & 3 & -3 \\
  0 & 0 & 0 & 1
  \end{pmatrix}
  \begin{pmatrix}
  1 \\
  x \\
  x^2 \\
  x^3
  \end{pmatrix}
  =
  \begin{pmatrix}
  B_{0,3}(x) \\
  B_{1,3}(x) \\
  B_{2,3}(x) \\
  B_{3,3}(x)
  \end{pmatrix}
  \]

- Inversion of that matrix yields
  \[
  \begin{pmatrix}
  1 & 1 & 1 & 1 \\
  0 & \frac{1}{3} & \frac{2}{3} & 1 \\
  0 & 0 & \frac{1}{3} & 1 \\
  0 & 0 & 0 & 1
  \end{pmatrix}
  \begin{pmatrix}
  B_{0,3}(x) \\
  B_{1,3}(x) \\
  B_{2,3}(x) \\
  B_{3,3}(x)
  \end{pmatrix}
  =
  \begin{pmatrix}
  1 \\
  x \\
  x^2 \\
  x^3
  \end{pmatrix},
  \]

  i.e., the fact that \( 1, x, x^2, x^3 \) of the power basis can be expressed in terms of
  \( B_{0,3}(x), B_{1,3}(x), B_{2,3}(x), B_{3,3}(x) \).
Basic Linear Algebra
- Matrices
- Linear Equations
- Determinants
  - Definition and Laplace Expansion
  - 2 × 2 and 3 × 3 Determinants
  - Properties of Determinants
  - Calculating Determinants
  - Determinants and Linear Systems
  - Geometric Interpretation of Determinants
- Eigenvalues and Eigenvectors
- Dot Product and Norm
- Vector Cross-Product
- Quaternions
Determinants

**Definition 93 (Submatrix, Dt.: Untermatrix)**

Let $A \in M_{n \times n}$, with $n \geq 2$. Let $A_{ij}(A)$, or simply $A_{ij}$ if there is no ambiguity, denote the $(n - 1) \times (n - 1)$ submatrix of $A$ formed by deleting the $i$-th row and $j$-th column of $A$. 

**Example:**

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 4 & 4 \end{pmatrix}$$

$A_{12} = \begin{pmatrix} 2 & 2 \\ 0 & 4 \end{pmatrix}$

$A_{33} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$

**Definition 94 (Determinant)**

The determinant, $\det(A)$, of an $n \times n$ matrix $A \in M_{n \times n}(\mathbb{R})$, for $n \in \mathbb{N}$, is defined recursively by the so-called first-row Laplace expansion:

$$\det(A) := \begin{cases} a_{11} & \text{if } n = 1, \\ \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \cdot \det(A_{1j}) & \text{if } n > 1. \end{cases}$$
Determinants

Definition 93 (Submatrix, Dt.: Untermatrix)

Let \( A \in M_{n \times n} \), with \( n \geq 2 \). Let \( A_{ij}(A) \), or simply \( A_{ij} \) if there is no ambiguity, denote the \((n - 1) \times (n - 1)\) submatrix of \( A \) formed by deleting the \( i \)-th row and \( j \)-th column of \( A \).

Example:

\[
A := \begin{pmatrix}
1 & 0 & 1 \\
2 & 1 & 2 \\
0 & 4 & 4
\end{pmatrix}
\]

\[
A_{12} = \begin{pmatrix}
2 & 2 \\
0 & 4
\end{pmatrix}
\]

\[
A_{33} = \begin{pmatrix}
1 & 0 \\
2 & 1
\end{pmatrix}
\]
Determinants

**Definition 93 (Submatrix, Dt.: Untermatrix)**

Let \( A \in M_{n \times n} \), with \( n \geq 2 \). Let \( A_{ij}(A) \), or simply \( A_{ij} \) if there is no ambiguity, denote the \((n - 1) \times (n - 1)\) submatrix of \( A \) formed by deleting the \( i \)-th row and \( j \)-th column of \( A \).

- **Example:**
  
  \[
  A := \begin{pmatrix}
  1 & 0 & 1 \\
  2 & 1 & 2 \\
  0 & 4 & 4 \\
  \end{pmatrix} \quad A_{12} := \begin{pmatrix}
  2 & 2 \\
  0 & 4 \\
  \end{pmatrix} \quad A_{33} := \begin{pmatrix}
  1 & 0 \\
  2 & 1 \\
  \end{pmatrix}
  \]

**Definition 94 (Determinant)**

The determinant, \( \det(A) \), of an \( n \times n \) matrix \( A \in M_{n \times n}(\mathbb{R}) \), for \( n \in \mathbb{N} \), is defined recursively by the so-called *first-row Laplace expansion*:

\[
\det(A) := \begin{cases} 
  a_{11} & \text{if } n = 1, \\
  \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \cdot \det(A_{1j}) & \text{if } n > 1.
\end{cases}
\]
Determinants

- Note that the term $|A|$ is also commonly used for denoting the determinant of an $n \times n$ matrix $A$, for $n \in \mathbb{N}$. 
Determinants

- Note that the term $|A|$ is also commonly used for denoting the determinant of an $n \times n$ matrix $A$, for $n \in \mathbb{N}$.
- E.g., it is common to write

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

instead of

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$
Laplace Expansion

- One can prove (albeit the proof is not entirely straightforward) that a determinant can be obtained by using any row or column for expansion if the following chessboard-like pattern is used for determining the signs of the summands:

\[
\begin{bmatrix}
+ & - & + & \cdots \\
- & + & - & \cdots \\
+ & - & + & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\]
Laplace Expansion

- One can prove (albeit the proof is not entirely straightforward) that a determinant can be obtained by using any row or column for expansion if the following chessboard-like pattern is used for determining the signs of the summands:

\[
\begin{bmatrix}
+ & - & + & \cdots \\
- & + & - & \cdots \\
+ & - & + & \cdots \\
\vdots & \vdots & \vdots & \ddots \\
\end{bmatrix}
\]

- E.g.,

\[
\text{det}(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \cdot \text{det}(A_{1j})
\]

\[
= \sum_{j=1}^{n} (-1)^{j} a_{2j} \cdot \text{det}(A_{2j})
\]

\[
= \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \cdot \text{det}(A_{i1})
\]
2 × 2 and 3 × 3 Determinants

Lemma 95

Determinant of a 2 × 2 matrix: For all $a, b, c, d \in \mathbb{R}$,

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$
Lemma 95

Determinant of a 2 × 2 matrix: For all $a, b, c, d \in \mathbb{R}$, 
\[
\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.
\]

Determinant of a 3 × 3 matrix: For all $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33} \in \mathbb{R}$, 
\[
\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \cdot \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}
\]
Lemma 95

Determinant of a $2 \times 2$ matrix: For all $a, b, c, d \in \mathbb{R}$,

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Determinant of a $3 \times 3$ matrix: For all $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33} \in \mathbb{R}$,

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \cdot \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22})$$
2 × 2 and 3 × 3 Determinants

**Lemma 95**

Determinant of a 2 × 2 matrix: For all \( a, b, c, d \in \mathbb{R} \),

\[
\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.
\]

Determinant of a 3 × 3 matrix: For all \( a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33} \in \mathbb{R} \),

\[
\begin{align*}
\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} & = a_{11} \cdot \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \\
& = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}) \\
& = a_{11}a_{22}a_{33} + a_{21}a_{13}a_{32} + a_{31}a_{12}a_{23} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} - a_{31}a_{13}a_{22}.
\end{align*}
\]
Mnemonic for Computing $3 \times 3$ Determinants (Sarrus)

\[
\begin{vmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}
\]

\[
= a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{31} a_{22} a_{13} - a_{32} a_{23} a_{11} - a_{33} a_{21} a_{12}.
\]
Mnemonic for Computing $3 \times 3$ Determinants (Sarrus)

\[
det \begin{pmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{pmatrix}
= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.
\]
Properties of Determinants

**Lemma 96**
If two columns or rows of a matrix are interchanged, then the determinant changes sign (if it is not zero), but its absolute value does not change.
Properties of Determinants

Lemma 96
If two columns or rows of a matrix are interchanged, then the determinant changes sign (if it is not zero), but its absolute value does not change.

Lemma 97
If a row (or column) of a matrix is zero, then its determinant is zero.
Properties of Determinants

Lemma 96
If two columns or rows of a matrix are interchanged, then the determinant changes sign (if it is not zero), but its absolute value does not change.

Lemma 97
If a row (or column) of a matrix is zero, then its determinant is zero.

Lemma 98
The determinant is a linear function of each row and each column.
Properties of Determinants

**Lemma 96**
If two columns or rows of a matrix are interchanged, then the determinant changes sign (if it is not zero), but its absolute value does not change.

**Lemma 97**
If a row (or column) of a matrix is zero, then its determinant is zero.

**Lemma 98**
The determinant is a linear function of each row and each column.

**Lemma 99**
If a multiple of a row is added to another row, then the value of the determinant remains unchanged. Same for columns.
Properties of Determinants

Lemma 96
If two columns or rows of a matrix are interchanged, then the determinant changes sign (if it is not zero), but its absolute value does not change.

Lemma 97
If a row (or column) of a matrix is zero, then its determinant is zero.

Lemma 98
The determinant is a linear function of each row and each column.

Lemma 99
If a multiple of a row is added to another row, then the value of the determinant remains unchanged. Same for columns.

Lemma 100
If two rows or columns of a matrix are equal then the determinant is zero.
Properties of Determinants

Lemma 101

The determinant of the product of two (square) matrices is the product of the determinants of the matrices:

\[ \det(AB) = \det(A) \det(B) \]

for all \( A, B \in M_{n \times n} \).
Properties of Determinants

Lemma 101
The determinant of the product of two (square) matrices is the product of the determinants of the matrices:

\[ \det(AB) = \det(A) \det(B) \]

for all \( A, B \in M_{n \times n} \).

Lemma 102
A matrix and its transpose have equal determinants, i.e., for all (square) \( A \),

\[ \det(A^t) = \det(A). \]
Properties of Determinants

Lemma 101
The determinant of the product of two (square) matrices is the product of the determinants of the matrices:

\[ \det(AB) = \det(A) \det(B) \]

for all \( A, B \in M_{n \times n} \).

Lemma 102
A matrix and its transpose have equal determinants, i.e., for all (square) \( A \),

\[ \det(A^t) = \det(A). \]

Lemma 103
The determinant of an orthogonal matrix is \( \pm 1 \).
Properties of Determinants

Lemma 101
The determinant of the product of two (square) matrices is the product of the determinants of the matrices:

$$\det(AB) = \det(A) \det(B)$$

for all $A, B \in M_{n \times n}$.

Lemma 102
A matrix and its transpose have equal determinants, i.e., for all (square) $A$,

$$\det(A^t) = \det(A).$$

Lemma 103
The determinant of an orthogonal matrix is $\pm 1$.

Theorem 104
The (square) matrix $A$ is not singular if and only if $\det(A) \neq 0$. 
Lemma 105

The determinant of an upper-triangular matrix

\[
A = \begin{pmatrix}
a_{11} & * & \cdots & \cdots & * \\
0 & a_{22} & \cdots & \cdots & \\
\vdots & \vdots & \ddots & \cdots & \\
\vdots & \vdots & \ddots & * & \\
0 & \cdots & \cdots & 0 & a_{nn}
\end{pmatrix}
\]

is given by the product of its diagonal elements: \( \det(A) = \prod_{i=1}^{n} a_{ii} \).
Properties of Determinants

Lemma 106

Let $n \in \mathbb{N}$ and $A, B, D \in M_{n \times n}(F)$. Then the determinant $\det(X)$ of the $2n \times 2n$ matrix $X$ with

$$X := \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

is given by

$$\det(X) = \det(A) \cdot \det(D).$$
Calculating Determinants Manually

- Make sure to make good use of the lemmas stated on the previous slides!
Calculating Determinants Manually

- Make sure to make good use of the lemmas stated on the previous slides!

\[
\begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
\]

\[
\text{det} = \begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
\]

\[
= (1 \cdot 4 \cdot 2 - (-1) \cdot 0 \cdot 1) - (2 \cdot 4 \cdot 1 - 3 \cdot 0 \cdot 2) + (-1 \cdot 4 \cdot 0 - 3 \cdot 0 \cdot 1)
\]

\[
= (8 + 0 - 0) = 8
\]

\[
= \begin{vmatrix}
0 & 2 & -3 \\
0 & 1 & 4 \\
1 & 0 & 2
\end{vmatrix}
\]

\[
= -\begin{vmatrix}
0 & -3 & 4 \\
2 & 4 & 2 \\
1 & 0 & 1
\end{vmatrix}
\]

\[
= -((-3) \cdot 2 \cdot 1 - 4 \cdot 4 \cdot 1) = -(-6 - 16) = -22
\]

\[
= \begin{vmatrix}
0 & -2 & -3 \\
1 & 0 & 2 \\
1 & 1 & 0
\end{vmatrix}
\]

\[
= -\begin{vmatrix}
-2 & -3 & 4 \\
1 & 2 & 1 \\
1 & 0 & 1
\end{vmatrix}
\]

\[
= -((-2) \cdot 1 \cdot 1 - (-3) \cdot 2 \cdot 1) = -(-2 + 6) = 4
\]

\[
= \begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
\]

\[
= -1 \cdot 4 \cdot 2 + 2 \cdot 0 \cdot 1 - (-1) \cdot 0 \cdot 1
\]

\[
= -8
\]

\[
= \begin{vmatrix}
0 & 2 & -3 \\
0 & 1 & 4 \\
1 & 0 & 2
\end{vmatrix}
\]

\[
= -\begin{vmatrix}
0 & -3 & 4 \\
2 & 4 & 2 \\
1 & 0 & 1
\end{vmatrix}
\]

\[
= -((-3) \cdot 2 \cdot 1 - 4 \cdot 4 \cdot 1) = -(-6 - 16) = -22
\]

\[
= \begin{vmatrix}
0 & -2 & -3 \\
1 & 0 & 2 \\
1 & 1 & 0
\end{vmatrix}
\]

\[
= -\begin{vmatrix}
-2 & -3 & 4 \\
1 & 2 & 1 \\
1 & 0 & 1
\end{vmatrix}
\]

\[
= -((-2) \cdot 1 \cdot 1 - (-3) \cdot 2 \cdot 1) = -(-2 + 6) = 4
\]

\[
= \begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
\]

\[
= 1 \cdot 4 \cdot 2 - (-1) \cdot 0 \cdot 1 + (3) \cdot 0 \cdot 2
\]

\[
= 8
\]

\[
= \begin{vmatrix}
0 & 2 & -3 \\
0 & 1 & 4 \\
1 & 0 & 2
\end{vmatrix}
\]

\[
= -\begin{vmatrix}
0 & -3 & 4 \\
2 & 4 & 2 \\
1 & 0 & 1
\end{vmatrix}
\]

\[
= -((-3) \cdot 2 \cdot 1 - 4 \cdot 4 \cdot 1) = -(-6 - 16) = -22
\]

\[
= \begin{vmatrix}
0 & -2 & -3 \\
1 & 0 & 2 \\
1 & 1 & 0
\end{vmatrix}
\]

\[
= -\begin{vmatrix}
-2 & -3 & 4 \\
1 & 2 & 1 \\
1 & 0 & 1
\end{vmatrix}
\]

\[
= -((-2) \cdot 1 \cdot 1 - (-3) \cdot 2 \cdot 1) = -(-2 + 6) = 4
\]

\[
= \begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
\]

\[
= 1 \cdot 4 \cdot 2 - (-1) \cdot 0 \cdot 1 + (3) \cdot 0 \cdot 2
\]

\[
= 8
\]
Calculating Determinants Manually

- Make sure to make good use of the lemmas stated on the previous slides!

\[
\begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
= 
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]
Calculating Determinants Manually

- Make sure to make good use of the lemmas stated on the previous slides!

\[
\begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
\]

\[
\begin{align*}
\det
\begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
& \equiv
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix} \\
& \equiv
\begin{vmatrix}
2 & -3 & 2 \\
1 & 4 & 2 \\
1 & 0 & 4
\end{vmatrix}
\end{align*}
\]

\[
= (-1)^{1+4} \cdot 1 \cdot \det
\begin{vmatrix}
2 & -3 & 2 \\
1 & 4 & 2 \\
1 & 0 & 4
\end{vmatrix}
\]

Expansion by first column
Calculating Determinants Manually

- Make sure to make good use of the lemmas stated on the previous slides!

\[
\begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix} \overset{I \leftarrow IV}{=} \begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

Expansion by first column

\[
= (-1)^{1+4} \cdot 1 \cdot \begin{vmatrix}
2 & -3 & 2 \\
1 & 4 & 2 \\
1 & 0 & 4 \\
\end{vmatrix} = - \begin{vmatrix}
0 & -3 & -6 \\
0 & 4 & -2 \\
1 & 0 & 4 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
0 & -3 & -6 \\
0 & 4 & -2 \\
1 & 0 & 4 \\
\end{vmatrix}
\]

Calculating Determinants Manually

- Make sure to make good use of the lemmas stated on the previous slides!

\[
\begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
2 & -3 & 2 \\
1 & 4 & 2 \\
1 & 0 & 4 \\
\end{vmatrix}
\begin{vmatrix}
0 & -3 & -6 \\
0 & 4 & -2 \\
1 & 0 & 4 \\
\end{vmatrix}
\]

\[
(-1)^{1+4} \cdot 1 \cdot \det
\begin{vmatrix}
2 & -3 & 2 \\
1 & 4 & 2 \\
1 & 0 & 4 \\
\end{vmatrix}
= - \det
\begin{vmatrix}
0 & -3 & -6 \\
0 & 4 & -2 \\
1 & 0 & 4 \\
\end{vmatrix}
\]

\[
= -(-1)^{1+3} \cdot 1 \cdot \det
\begin{vmatrix}
-3 & -6 \\
4 & -2 \\
\end{vmatrix}
\]
Calculating Determinants Manually

- Make sure to make good use of the lemmas stated on the previous slides!

\[
\begin{vmatrix}
1 & 2 & -1 & 3 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
\quad \overset{I\rightarrow IV}{\longrightarrow}
\begin{vmatrix}
0 & 2 & -3 & 2 \\
0 & 1 & 4 & 2 \\
0 & 1 & 0 & 4 \\
1 & 0 & 2 & 1
\end{vmatrix}
\]

Expansion by first column

\[
= (-1)^{1+4} \cdot 1 \cdot \det
\begin{vmatrix}
2 & -3 & 2 \\
1 & 4 & 2 \\
1 & 0 & 4
\end{vmatrix}
= - \det
\begin{vmatrix}
0 & -3 & -6 \\
0 & 4 & -2 \\
1 & 0 & 4
\end{vmatrix}
\]

\[
= -(-1)^{1+3} \cdot 1 \cdot \det
\begin{vmatrix}
-3 & -6 \\
4 & -2
\end{vmatrix}
= -((-3 \cdot (-2)) - (-6 \cdot 4)) = -30.
\]
The recursive formula results in a horrendous algorithmic complexity: If $T(n)$ denotes the number of multiplications needed for computing the determinant of an $n \times n$ matrix, with $T(2) := 2$, then $T(n) = n + n \cdot T(n - 1)$ and, thus, $T(n) > n!$. Hence, the recursive formula is not suitable for anything but small matrices. Standard alternative: Apply Gaussian elimination in order to transform the input matrix into an upper-triangular matrix, at a cost of $\Theta(n^3)$ operations. Unfortunately, this transformation introduces divisions. Bird (IPL 111(21–22), 2011) presents a simple method that requires $O(n M(n))$ additions and multiplications for an $n \times n$ matrix, where $M(n)$ is the number of such operations needed for matrix multiplication. If naïve matrix multiplication is used then we get $\Theta(n^4)$. No $\Theta(n^3)$ division-free determinant calculation is known.
Implementing Determinant Calculations

- The recursive formula results in a horrendous algorithmic complexity: If $T(n)$ denotes the number of multiplications needed for computing the determinant of an $n \times n$ matrix, with $T(2) := 2$, then $T(n) = n + n \cdot T(n-1)$ and, thus, $T(n) > n!$.
- Hence, the recursive formula is not suitable for anything but small matrices.
- Standard alternative: Apply Gaussian elimination in order to transform the input matrix into an upper-triangular matrix, at a cost of $\Theta(n^3)$ operations.
- Unfortunately, this transformation introduces divisions.
Implementing Determinant Calculations

- The recursive formula results in a horrendous algorithmic complexity: If $T(n)$ denotes the number of multiplications needed for computing the determinant of an $n \times n$ matrix, with $T(2) := 2$, then $T(n) = n + n \cdot T(n-1)$ and, thus, $T(n) > n!$.
- Hence, the recursive formula is not suitable for anything but small matrices.
- Standard alternative: Apply Gaussian elimination in order to transform the input matrix into an upper-triangular matrix, at a cost of $\Theta(n^3)$ operations.
- Unfortunately, this transformation introduces divisions.
- Bird (IPL 111(21–22), 2011) presents a simple method that requires $O(nM(n))$ additions and multiplications for an $n \times n$ matrix, where $M(n)$ is the number of such operations needed for matrix multiplication.
Implementing Determinant Calculations

- The recursive formula results in a horrendous algorithmic complexity: If \( T(n) \) denotes the number of multiplications needed for computing the determinant of an \( n \times n \) matrix, with \( T(2) := 2 \), then \( T(n) = n + n \cdot T(n - 1) \) and, thus, \( T(n) > n! \).
- Hence, the recursive formula is not suitable for anything but small matrices.
- Standard alternative: Apply Gaussian elimination in order to transform the input matrix into an upper-triangular matrix, at a cost of \( \Theta(n^3) \) operations.
- Unfortunately, this transformation introduces divisions.
- Bird (IPL 111(21–22), 2011) presents a simple method that requires \( O(nM(n)) \) additions and multiplications for an \( n \times n \) matrix, where \( M(n) \) is the number of such operations needed for matrix multiplication.
- If naïve matrix multiplication is used then we get \( \Theta(n^4) \).
- No \( \Theta(n^3) \) division-free determinant calculation is known.
Lemma 107

The linear system $Ax = b$, with $A \in M_{n \times n}$, has a unique solution if and only if $\det(A) \neq 0$. 

Lemma 108 (Cramer's Rule)

If $\det(A) \neq 0$, for $A \in M_{n \times n}(\mathbb{R})$, then the solution of $Ax = b$ is given by

$$
x_1 = \frac{\det(A_1)}{\det(A)}, \quad x_2 = \frac{\det(A_2)}{\det(A)}, \ldots, \quad x_n = \frac{\det(A_n)}{\det(A)},
$$

where $A_i$ is the matrix formed by replacing the $i$-th column of the coefficient matrix $A$ by the right-hand side $b$. 

© M. Held (Univ. Salzburg)  
Geometrisches Rechnen (WS 2020/21)
Determinants and Linear Systems

Lemma 107
The linear system $Ax = b$, with $A \in M_{n \times n}$, has a unique solution if and only if $\det(A) \neq 0$.

Lemma 108 (Cramer’s Rule)
If $\det(A) \neq 0$, for $A \in M_{n \times n}(\mathbb{R})$, then the solution of $Ax = b$ is given by

$$x_1 = \frac{\det(A_1)}{\det(A)}, \quad x_2 = \frac{\det(A_2)}{\det(A)}, \ldots, \quad x_n = \frac{\det(A_n)}{\det(A)},$$

where $A_i$ is the matrix formed by replacing the $i$-th column of the coefficient matrix $A$ by the right-hand side $b$. 
Theorem 109

Let \(a, b, c, d \in \mathbb{R}\). Consider the vectors

\[
v_1 := \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad v_2 := \begin{pmatrix} b \\ d \end{pmatrix} \quad \text{and let} \quad T := \begin{pmatrix} a & b \\ c & d \end{pmatrix}.
\]

Then \(\det(T)\) gives the signed area of the parallelogram spanned by \(v_1, v_2\). The determinant is positive if \(v_1, v_2\) form a right-handed coordinate system for \(\mathbb{R}^2\), zero if they are collinear, and negative otherwise.
Theorem 109

Let \( a, b, c, d \in \mathbb{R} \). Consider the vectors

\[
\begin{align*}
\mathbf{v}_1 & := \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 := \begin{pmatrix} b \\ d \end{pmatrix} \quad \text{and let} \quad \mathbf{T} := \begin{pmatrix} a & b \\ c & d \end{pmatrix}.
\end{align*}
\]

Then \( \det(\mathbf{T}) \) gives the signed area of the parallelogram spanned by \( \mathbf{v}_1, \mathbf{v}_2 \). The determinant is positive if \( \mathbf{v}_1, \mathbf{v}_2 \) form a right-handed coordinate system for \( \mathbb{R}^2 \), zero if they are collinear, and negative otherwise.

**Proof:** Let \( \mathbf{v}_1, \mathbf{v}_2 \) form a right-handed coordinate system. We have \( \det(\mathbf{T}) = ad - bc \).
Geometric Interpretation of Determinants: Orientation and Area

Theorem 109

Let \( a, b, c, d \in \mathbb{R} \). Consider the vectors

\[
\begin{align*}
  v_1 &:= \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad v_2 := \begin{pmatrix} b \\ d \end{pmatrix}
\end{align*}
\]

and let \( T := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \).

Then \( \det(T) \) gives the signed area of the parallelogram spanned by \( v_1, v_2 \). The determinant is positive if \( v_1, v_2 \) form a right-handed coordinate system for \( \mathbb{R}^2 \), zero if they are collinear, and negative otherwise.

Proof: Let \( v_1, v_2 \) form a right-handed coordinate system. We have \( \det(T) = ad - bc \).

Now consider the parallelogram defined by \( v_1 \) and \( v_2 \).
Theorem 109

Let $a, b, c, d \in \mathbb{R}$. Consider the vectors

$$v_1 := \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad v_2 := \begin{pmatrix} b \\ d \end{pmatrix} \quad \text{and let} \quad T := \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$ 

Then $\det(T)$ gives the signed area of the parallelogram spanned by $v_1, v_2$. The determinant is positive if $v_1, v_2$ form a right-handed coordinate system for $\mathbb{R}^2$, zero if they are collinear, and negative otherwise.

Proof: Let $v_1, v_2$ form a right-handed coordinate system. We have $\det(T) = ad - bc$.

Now consider the parallelogram defined by $v_1$ and $v_2$ and observe that its area $A$ equals $ad - bc$: 
Geometric Interpretation of Determinants: Orientation and Area

**Theorem 109**

Let $a, b, c, d \in \mathbb{R}$. Consider the vectors

$$v_1 := \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad v_2 := \begin{pmatrix} b \\ d \end{pmatrix} \quad \text{and let} \quad T := \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$ 

Then $\det(T)$ gives the signed area of the parallelogram spanned by $v_1, v_2$. The determinant is positive if $v_1, v_2$ form a right-handed coordinate system for $\mathbb{R}^2$, zero if they are collinear, and negative otherwise.

**Proof:** Let $v_1, v_2$ form a right-handed coordinate system. We have $\det(T) = ad - bc$.

Now consider the parallelogram defined by $v_1$ and $v_2$ and observe that its area $A$ equals $ad - bc$:

$$A = (a + b)(c + d) - ac - bd - 2bc$$

$$= ad - bc.$$
Geometric Interpretation of Determinants: Orientation and Area

**Theorem 109**

Let \( a, b, c, d \in \mathbb{R} \). Consider the vectors

\[
v_1 := \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad v_2 := \begin{pmatrix} b \\ d \end{pmatrix}
\]

and let \( T := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \).

Then \( \det(T) \) gives the signed area of the parallelogram spanned by \( v_1, v_2 \). The determinant is positive if \( v_1, v_2 \) form a right-handed coordinate system for \( \mathbb{R}^2 \), zero if they are collinear, and negative otherwise.

**Proof:** Let \( v_1, v_2 \) form a right-handed coordinate system. We have \( \det(T) = ad - bc \).

Now consider the parallelogram defined by \( v_1 \) and \( v_2 \) and observe that its area \( A \) equals \( ad - bc \):

\[
A = (a + b)(c + d) - ac - bd - 2bc \\
= ad - bc.
\]

Interchanging \( v_1 \) and \( v_2 \) flips their handedness and changes the sign of the determinant.
Lemma 110

For points $p_1 := (x_1, y_1)$ and $p_2 := (x_2, y_2)$ in $\mathbb{R}^2$,

$$\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$$

is positive if the triangle formed by the origin $O = (0, 0)$ and the points $p_1$ and $p_2$ has counter-clockwise (CCW) orientation.

It is negative for a clockwise (CW) orientation.

This determinant is zero if $p_1$, $p_2$ and $O$ are collinear.
**Lemma 110**

For points $p_1 := (x_1, y_1)$ and $p_2 := (x_2, y_2)$ in $\mathbb{R}^2$,

$$\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$$

is positive if the triangle formed by the origin $O = (0, 0)$ and the points $p_1$ and $p_2$ has counter-clockwise (CCW) orientation. It is negative for a clockwise (CW) orientation. This determinant is zero if $p_1$, $p_2$ and $O$ are collinear.
Geometric Interpretation of Determinants: Orientation

Lemma 110

For points \( p_1 := (x_1, y_1) \) and \( p_2 := (x_2, y_2) \) in \( \mathbb{R}^2 \),

\[
\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}
\]

is positive if the triangle formed by the origin \( O = (0, 0) \) and the points \( p_1 \) and \( p_2 \) has counter-clockwise (CCW) orientation. It is negative for a clockwise (CW) orientation. This determinant is zero if \( p_1, p_2 \) and \( O \) are collinear.
Geometric Interpretation of Determinants: Orientation

Lemma 110
For points $p_1 := (x_1, y_1)$ and $p_2 := (x_2, y_2)$ in $\mathbb{R}^2$,

$$\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$$

is positive if the triangle formed by the origin $O = (0, 0)$ and the points $p_1$ and $p_2$ has counter-clockwise (CCW) orientation. It is negative for a clockwise (CW) orientation. This determinant is zero if $p_1$, $p_2$ and $O$ are collinear.

Lemma 111
For points $p_1 := (x_1, y_1)$, $p_2 := (x_2, y_2)$ and $p_3 := (x_3, y_3)$ in $\mathbb{R}^2$,

$$\det \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix}$$

is positive if the triangle $\Delta(p_1, p_2, p_3)$ formed by $p_1$, $p_2$, $p_3$ has counter-clockwise (CCW) orientation.
Geometric Interpretation of Determinants: Orientation

**Lemma 110**

For points $p_1 := (x_1, y_1)$ and $p_2 := (x_2, y_2)$ in $\mathbb{R}^2$,

$$\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$$

is positive if the triangle formed by the origin $O = (0, 0)$ and the points $p_1$ and $p_2$ has counter-clockwise (CCW) orientation. It is negative for a clockwise (CW) orientation. This determinant is zero if $p_1$, $p_2$ and $O$ are collinear.

**Lemma 111**

For points $p_1 := (x_1, y_1)$, $p_2 := (x_2, y_2)$ and $p_3 := (x_3, y_3)$ in $\mathbb{R}^2$,

$$\det \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix}$$

is positive if the triangle $\Delta(p_1, p_2, p_3)$ formed by $p_1$, $p_2$, $p_3$ has counter-clockwise (CCW) orientation. It is negative for a clockwise (CW) orientation.
Lemma 110
For points $p_1 := (x_1, y_1)$ and $p_2 := (x_2, y_2)$ in $\mathbb{R}^2$,

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$$

is positive if the triangle formed by the origin $O = (0, 0)$ and the points $p_1$ and $p_2$ has counter-clockwise (CCW) orientation. It is negative for a clockwise (CW) orientation. This determinant is zero if $p_1$, $p_2$ and $O$ are collinear.

Lemma 111
For points $p_1 := (x_1, y_1)$, $p_2 := (x_2, y_2)$ and $p_3 := (x_3, y_3)$ in $\mathbb{R}^2$,

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

is positive if the triangle $\Delta(p_1, p_2, p_3)$ formed by $p_1$, $p_2$, $p_3$ has counter-clockwise (CCW) orientation. It is negative for a clockwise (CW) orientation. This determinant is zero if $p_1$, $p_2$ and $p_3$ are collinear.
Lemma 112

Apart from the sign,

\[
\frac{1}{2} \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}
\]

corresponds to the area of the triangle \( \Delta(O, p_1, p_2) \).
Geometric Interpretation of Determinants: Area

**Lemma 112**

Apart from the sign,

\[ \frac{1}{2} \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \]

corresponds to the area of the triangle \( \Delta(O, p_1, p_2) \).

**Lemma 113**

Apart from the sign,

\[ \frac{1}{2} \det \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix} \]

corresponds to the area of the triangle \( \Delta(p_1, p_2, p_3) \).
Geometric Interpretation of Determinants: Area

- Consider the triangle (in the plane) with corners (2, 1), (7, 2) and (3, 5).

```
\[ A = \frac{1}{2} \cdot \det \begin{bmatrix} 2 & 1 & 1 \\ 7 & 2 & 1 \\ 3 & 5 & 1 \end{bmatrix} = \frac{1}{2} \cdot \det \begin{bmatrix} 2 & 1 & 1 \\ 5 & 1 & 0 \\ 1 & 4 & 0 \end{bmatrix} = \frac{1}{2} \cdot (5 \cdot 4 - 1 \cdot 1) = \frac{19}{2}. \]
```
Geometric Interpretation of Determinants: Area

Consider the triangle (in the plane) with corners (2, 1), (7, 2) and (3, 5).

The area of that triangle is given by

\[
A = \frac{1}{2} \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 7 & 2 & 1 \\ 3 & 5 & 1 \end{pmatrix} = \frac{1}{2} \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 5 & 1 & 0 \\ 1 & 4 & 0 \end{pmatrix} = \frac{1}{2} \cdot (5 \cdot 4 - 1 \cdot 1) = \frac{19}{2}.
\]
Lemma 114

For points \( p_1 := (x_1, y_1, z_1), p_2 := (x_2, y_2, z_2), p_3 := (x_3, y_3, z_3) \) and \( p_4 := (x_4, y_4, z_4) \) in \( \mathbb{R}^3 \),

\[
\frac{1}{6} \left| \begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} \right|
\]

corresponds to the volume of the tetrahedron with corners \( p_1, p_2, p_3, p_4 \).
**Lemma 114**

For points $p_1 := (x_1, y_1, z_1), p_2 := (x_2, y_2, z_2), p_3 := (x_3, y_3, z_3)$ and $p_4 := (x_4, y_4, z_4)$ in $\mathbb{R}^3$,

$$\frac{1}{6} \left| \begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} \right|$$

Corresponds to the volume of the tetrahedron with corners $p_1, p_2, p_3, p_4$.

**Lemma 115**

For points $p_1 := (x_1, y_1, z_1), p_2 := (x_2, y_2, z_2), p_3 := (x_3, y_3, z_3)$ in $\mathbb{R}^3$,

$$\frac{1}{6} \left| \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} \right|$$

Corresponds to the volume of the tetrahedron with corners $p_1, p_2, p_3$ and the origin as fourth corner.
Lemma 116

Let $a, b, c \in \mathbb{R}^3$. Then

$$\left| \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} \right|$$

corresponds to the volume of the parallelepiped spanned by the three vectors $a, b, c$. 
Basic Linear Algebra

- Matrices
- Linear Equations
- Determinants
- Eigenvalues and Eigenvectors
- Dot Product and Norm
- Vector Cross-Product
- Quaternions
Definition 117 (Eigenvalue, Dt.: Eigenwert)

Consider a square $n \times n$ matrix $A$. A scalar $\lambda$ is called eigenvalue of $A$ if a vector $v \in \mathbb{R}^n$ exists such that

$$A v = \lambda v \quad \text{and} \quad v \neq 0.$$ 

Such a vector $v$ is called eigenvector of $A$. 

A scalar $\lambda$ is an eigenvalue of matrix $A$ if and only if the homogeneous linear system of equations $(A - \lambda I)v = 0$ has a non-zero solution. This is the case if and only if $(A - \lambda I)$ is singular, that is, if $\det(A - \lambda I) = 0$.

Thus, the eigenvalues of a matrix $A$ are the zeros of the characteristic polynomial $p_A(\lambda) := \det(A - \lambda I)$.

While this approach works for any $n \times n$ matrix, it becomes tedious for $n > 4$.

An $n \times n$ matrix can have at most $n$ eigenvalues.

Sample application of eigenvalues and eigenvectors: Principal Components Analysis.
Eigenvalues and Eigenvectors

Definition 117 (Eigenvalue, Dt.: Eigenwert)

Consider a square $n \times n$ matrix $A$. A scalar $\lambda$ is called eigenvalue of $A$ if a vector $v \in \mathbb{R}^n$ exists such that

$$A v = \lambda v \quad \text{and} \quad v \neq 0.$$  

Such a vector $v$ is called eigenvector of $A$.

- A scalar $\lambda$ is an eigenvalue of matrix $A$ if and only if the homogeneous linear system of equations $(A - \lambda I) v = 0$ has a non-zero solution. This is the case if and only if $(A - \lambda I)$ is singular, that is, if $\det(A - \lambda I) = 0$. 

Sample application of eigenvalues and eigenvectors: Principal Components Analysis.
Definition 117 (Eigenvalue, Dt.: Eigenwert)

Consider a square $n \times n$ matrix $A$. A scalar $\lambda$ is called \textit{eigenvalue} of $A$ if a vector $v \in \mathbb{R}^n$ exists such that

$$Av = \lambda v \quad \text{and} \quad v \neq 0.$$ 

Such a vector $v$ is called \textit{eigenvector} of $A$.

- A scalar $\lambda$ is an eigenvalue of matrix $A$ if and only if the homogeneous linear system of equations $(A - \lambda I)v = 0$ has a non-zero solution. This is the case if and only if $(A - \lambda I)$ is singular, that is, if $\det(A - \lambda I) = 0$.
- Thus, the eigenvalues of a matrix $A$ are the zeros of the \textit{characteristic polynomial}

$$p_A(\lambda) := \det(A - \lambda I).$$

- While this approach works for any $n \times n$ matrix, it becomes tedious for $n > 4$. 

Definition 117 (Eigenvalue, Dt.: Eigenwert)

Consider a square $n \times n$ matrix $A$. A scalar $\lambda$ is called eigenvalue of $A$ if a vector $v \in \mathbb{R}^n$ exists such that

$$Av = \lambda v \quad \text{and} \quad v \neq 0.$$ 

Such a vector $v$ is called eigenvector of $A$.

- A scalar $\lambda$ is an eigenvalue of matrix $A$ if and only if the homogeneous linear system of equations $(A - \lambda I)v = 0$ has a non-zero solution. This is the case if and only if $(A - \lambda I)$ is singular, that is, if $\det(A - \lambda I) = 0$.
- Thus, the eigenvalues of a matrix $A$ are the zeros of the characteristic polynomial

$$p_A(\lambda) := \det(A - \lambda I).$$

- While this approach works for any $n \times n$ matrix, it becomes tedious for $n > 4$.
- An $n \times n$ matrix can have at most $n$ eigenvalues.
Eigenvalues and Eigenvectors

Definition 117 (Eigenvalue, Dt.: Eigenwert)

Consider a square $n \times n$ matrix $A$. A scalar $\lambda$ is called eigenvalue of $A$ if a vector $v \in \mathbb{R}^n$ exists such that

$$A v = \lambda v \quad \text{and} \quad v \neq 0.$$ 

Such a vector $v$ is called eigenvector of $A$.

- A scalar $\lambda$ is an eigenvalue of matrix $A$ if and only if the homogeneous linear system of equations $(A - \lambda I)v = 0$ has a non-zero solution. This is the case if and only if $(A - \lambda I)$ is singular, that is, if $\det(A - \lambda I) = 0$.
- Thus, the eigenvalues of a matrix $A$ are the zeros of the characteristic polynomial $p_A(\lambda) := \det(A - \lambda I)$.
- While this approach works for any $n \times n$ matrix, it becomes tedious for $n > 4$.
- An $n \times n$ matrix can have at most $n$ eigenvalues.
- Sample application of eigenvalues and eigenvectors: Principal Components Analysis.
Sample problem: Suppose that you are given a cloud of points in $\mathbb{R}^3$. Somebody tells you that all those points lie inside of an (unknown) ellipsoid. How would you rotate and translate those points such that the main axes of the ellipsoid coincide with the coordinate axes?

Roughly, Principal Components Analysis (PCA, Dt.: Hauptkomponentenanalyse) is a statistical method for finding "structure" in such a point cloud. PCA starts with subtracting the mean of all points from every point. This is equivalent to translating the point cloud such that its centroid coincides with the origin.

Then, PCA chooses the first PCA axis as that line which goes through the centroid of the point cloud, but also minimizes the square of the distance of each point to that line. Thus, the line is as close to all of the points as possible. Equivalently, the line goes through the maximum variation in the point cloud.

The second PCA axis also goes through the centroid, and also goes through the maximum variation in the points in a direction that is orthogonal to the first axes. Similarly for the third axes.
Principal Components Analysis (PCA)

- Sample problem: Suppose that you are given a cloud of points in $\mathbb{R}^3$. Somebody tells you that all those points lie inside of an (unknown) ellipsoid. How would you rotate and translate those points such that the main axes of the ellipsoid coincide with the coordinate axes?

- Roughly, Principal Components Analysis (PCA, Dt.: Hauptkomponentenanalyse) is a statistical method for finding “structure” in such a point cloud.

- PCA starts with subtracting the mean of all points from every point. This is equivalent to translating the point cloud such that its centroid coincides with the origin.
Principal Components Analysis (PCA)

- Sample problem: Suppose that you are given a cloud of points in $\mathbb{R}^3$. Somebody tells you that all those points lie inside of an (unknown) ellipsoid. How would you rotate and translate those points such that the main axes of the ellipsoid coincide with the coordinate axes?

- Roughly, Principal Components Analysis (PCA, Dt.: Hauptkomponentenanalyse) is a statistical method for finding “structure” in such a point cloud.

- PCA starts with subtracting the mean of all points from every point. This is equivalent to translating the point cloud such that its centroid coincides with the origin.

- Then, PCA chooses the first PCA axis as that line which goes through the centroid of the point cloud, but also minimizes the square of the distance of each point to that line. Thus, the line is as close to all of the points as possible. Equivalently, the line goes through the maximum variation in the point cloud.
Principal Components Analysis (PCA)

- Sample problem: Suppose that you are given a cloud of points in $\mathbb{R}^3$. Somebody tells you that all those points lie inside of an (unknown) ellipsoid. How would you rotate and translate those points such that the main axes of the ellipsoid coincide with the coordinate axes?

- Roughly, Principal Components Analysis (PCA, Dt.: Hauptkomponentenanalyse) is a statistical method for finding “structure” in such a point cloud.

- PCA starts with subtracting the mean of all points from every point. This is equivalent to translating the point cloud such that its centroid coincides with the origin.

- Then, PCA chooses the first PCA axis as that line which goes through the centroid of the point cloud, but also minimizes the square of the distance of each point to that line. Thus, the line is as close to all of the points as possible. Equivalently, the line goes through the maximum variation in the point cloud.

- The second PCA axis also goes through the centroid, and also goes through the maximum variation in the points in a direction that is orthogonal to the first axes.

- Similarly for the third axes.
Principal Components Analysis (PCA)

- In $d$ dimensions, PCA can be thought of as fitting a $d$-dimensional (hyper-)ellipsoid to the data such that each axis of the ellipsoid represents a principal component.

- If some axis of the ellipsoid is short then the variance along that axis is also small. Hence, one would lose only a rather small amount of information if one would omit that axis and its corresponding principal component from the representation of the dataset.
Principal Components Analysis (PCA)

- In $d$ dimensions, PCA can be thought of as fitting a $d$-dimensional (hyper-)ellipsoid to the data such that each axis of the ellipsoid represents a principal component.
- If some axis of the ellipsoid is short then the variance along that axis is also small.
- Hence, one would lose only a rather small amount of information if one would omit that axis and its corresponding principal component from the representation of the dataset.
Principal Components Analysis (PCA)

- Consider $n$ points $p_i := (x_i, y_i, z_i) \in \mathbb{R}^3$.

- For our sample application, the PCA axes can be computed by finding the eigenvalues and eigenvectors of the covariance matrix $\text{Cov}$ of the coordinates of the $n$ points:

$$
\text{Cov}(x, y, z) := \begin{pmatrix}
\text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\
\text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\
\text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z)
\end{pmatrix},
$$

where

$$
\text{cov}(x, y) := \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n - 1} \quad \text{and} \quad \bar{x} := \frac{1}{n} \sum_{i=1}^{n} x_i \quad \text{and} \quad \bar{y} := \frac{1}{n} \sum_{i=1}^{n} y_i.
$$

Similarly for the other entries of the covariance matrix.
Principal Components Analysis (PCA)

- Consider \( n \) points \( p_i := (x_i, y_i, z_i) \in \mathbb{R}^3 \).
- For our sample application, the PCA axes can be computed by finding the eigenvalues and eigenvectors of the covariance matrix \( \text{Cov} \) of the coordinates of the \( n \) points:

\[
\text{Cov}(x, y, z) := \begin{pmatrix}
\text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\
\text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\
\text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z)
\end{pmatrix},
\]

where

\[
\text{cov}(x, y) := \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}
\]

and \( \bar{x} := \frac{1}{n} \sum_{i=1}^{n} x_i \) and \( \bar{y} := \frac{1}{n} \sum_{i=1}^{n} y_i \).

Similarly for the other entries of the covariance matrix.
- The origin of the PCA axes is given by the mean point \((\bar{x}, \bar{y}, \bar{z})\).
Basic Linear Algebra

- Matrices
- Linear Equations
- Determinants
- Eigenvalues and Eigenvectors
- Dot Product and Norm
  - Dot Product
  - Norm
  - Standard Dot Product on $\mathbb{R}^n$
  - Angle and Projection
- Vector Cross-Product
- Quaternions $\mathbb{H}$
Dot Product

Definition 118 (Dot product, Dt.: Skalarprodukt, inneres Produkt)

Consider a vector space $V$ over a field $F$, where $F$ is either $\mathbb{R}$ or $\mathbb{C}$. A mapping

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow F$$

$$(a, b) \mapsto \langle a, b \rangle$$

is called a *dot product* (or *inner product*) on $V$ if

1. $\langle \lambda_1 a + \lambda_2 b, c \rangle = \lambda_1 \langle a, c \rangle + \lambda_2 \langle b, c \rangle$,
2. $\langle a, b \rangle = \langle b, a \rangle$,
3. $\langle a, a \rangle \geq 0$,
4. $\langle a, a \rangle = 0 \Rightarrow a = 0$.

Note that Condition 2 ensures that $\langle a, a \rangle \in \mathbb{R}$.

If $F$ is $\mathbb{R}$ then commutativity holds. (In the sequel we will assume $F$ to be $\mathbb{R}$.)

Be warned that the notation is not uniform: $a \cdot b$ and $(a | b)$ are two other common notations for denoting the dot product of $a$ and $b$. Note the difference between $a \cdot b$ for $a, b \in V$, and $\lambda \cdot a$ for $\lambda \in F$ and $a \in V$!
Dot Product

Definition 118 (Dot product, Dt.: Skalarprodukt, inneres Produkt)

Consider a vector space $V$ over a field $F$, where $F$ is either $\mathbb{R}$ or $\mathbb{C}$. A mapping

$$\langle, \rangle : V \times V \rightarrow F$$

$$(a, b) \mapsto \langle a, b \rangle$$

is called a dot product (or inner product) on $V$ if for all $a, b, c \in V$ and all $\lambda_1, \lambda_2 \in F$

1. $$\langle \lambda_1 a + \lambda_2 b, c \rangle = \lambda_1 \langle a, c \rangle + \lambda_2 \langle b, c \rangle;$$
Definition 118 (Dot product, Dt.: Skalarprodukt, inneres Produkt)

Consider a vector space $V$ over a field $F$, where $F$ is either $\mathbb{R}$ or $\mathbb{C}$. A mapping

$$\langle , \rangle : V \times V \rightarrow F$$

$$(a, b) \mapsto \langle a, b \rangle$$

is called a dot product (or inner product) on $V$ if for all $a, b, c \in V$ and all $\lambda_1, \lambda_2 \in F$

1. $\langle \lambda_1 a + \lambda_2 b, c \rangle = \lambda_1 \langle a, c \rangle + \lambda_2 \langle b, c \rangle$;
2. $\langle a, b \rangle = \langle b, a \rangle$;
Definition 118 (Dot product, Dt.: Skalarprodukt, inneres Produkt)

Consider a vector space $V$ over a field $F$, where $F$ is either $\mathbb{R}$ or $\mathbb{C}$. A mapping

$$\langle, \rangle : V \times V \rightarrow F$$

$$(a, b) \mapsto \langle a, b \rangle$$

is called a dot product (or inner product) on $V$ if for all $a, b, c \in V$ and all $\lambda_1, \lambda_2 \in F$

1. $\langle \lambda_1 a + \lambda_2 b, c \rangle = \lambda_1 \langle a, c \rangle + \lambda_2 \langle b, c \rangle$;
2. $\langle a, b \rangle = \overline{\langle b, a \rangle}$;
3. $\langle a, a \rangle \geq 0$;
Dot Product

**Definition 118 (Dot product, Dt.: Skalarprodukt, inneres Produkt)**

Consider a vector space $V$ over a field $F$, where $F$ is either $\mathbb{R}$ or $\mathbb{C}$. A mapping

$$\langle \cdot, \cdot \rangle : V \times V \to F$$

$$(a, b) \mapsto \langle a, b \rangle$$

is called a dot product (or inner product) on $V$ if for all $a, b, c \in V$ and all $\lambda_1, \lambda_2 \in F$

1. $\langle \lambda_1 a + \lambda_2 b, c \rangle = \lambda_1 \langle a, c \rangle + \lambda_2 \langle b, c \rangle$;
2. $\langle a, b \rangle = \overline{\langle b, a \rangle}$;
3. $\langle a, a \rangle \geq 0$;
4. $\langle a, a \rangle = 0 \implies a = 0$. 

Note that Condition 2 ensures that $\langle a, a \rangle \in \mathbb{R}$. If $F$ is $\mathbb{R}$ then commutativity holds. (In the sequel we will assume $F$ to be $\mathbb{R}$.) Be warned that the notation is not uniform: $a \cdot b$ and $(a | b)$ are two other common notations for denoting the dot product of $a$ and $b$. Note the difference between $a \cdot b$ for $a, b \in V$, and $\lambda \cdot a$ for $\lambda \in F$ and $a \in V$!
Dot Product

**Definition 118 (Dot product, Dt.: Skalarprodukt, inneres Produkt)**

Consider a vector space $V$ over a field $F$, where $F$ is either $\mathbb{R}$ or $\mathbb{C}$. A mapping

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow F$$

$$(a, b) \mapsto \langle a, b \rangle$$

is called a *dot product* (or *inner product*) on $V$ if for all $a, b, c \in V$ and all $\lambda_1, \lambda_2 \in F$

1. $\langle \lambda_1 a + \lambda_2 b, c \rangle = \lambda_1 \langle a, c \rangle + \lambda_2 \langle b, c \rangle$;
2. $\langle a, b \rangle = \overline{\langle b, a \rangle}$;
3. $\langle a, a \rangle \geq 0$;
4. $\langle a, a \rangle = 0 \Rightarrow a = 0$.

- Note that Condition 2 ensures that $\langle a, a \rangle \in \mathbb{R}$. 
Dot Product

Definition 118 (Dot product, Dt.: Skalarprodukt, inneres Produkt)
Consider a vector space $V$ over a field $F$, where $F$ is either $\mathbb{R}$ or $\mathbb{C}$. A mapping

$$\langle , \rangle : V \times V \rightarrow F$$

$$(a, b) \mapsto \langle a, b \rangle$$

is called a dot product (or inner product) on $V$ if for all $a, b, c \in V$ and all $\lambda_1, \lambda_2 \in F$

1. $\langle \lambda_1 a + \lambda_2 b, c \rangle = \lambda_1 \langle a, c \rangle + \lambda_2 \langle b, c \rangle$;
2. $\langle a, b \rangle = \overline{\langle b, a \rangle}$;
3. $\langle a, a \rangle \geq 0$;
4. $\langle a, a \rangle = 0 \Rightarrow a = 0$.

- Note that Condition 2 ensures that $\langle a, a \rangle \in \mathbb{R}$.
- If $F$ is $\mathbb{R}$ then commutativity holds. (In the sequel we will assume $F$ to be $\mathbb{R}$.)
**Dot Product**

**Definition 118 (Dot product, Dt.: Skalarprodukt, inneres Produkt)**

Consider a vector space $V$ over a field $F$, where $F$ is either $\mathbb{R}$ or $\mathbb{C}$. A mapping

$$
\langle \cdot, \cdot \rangle : V \times V \to F
$$

$$(a, b) \mapsto \langle a, b \rangle$$

is called a *dot product* (or *inner product*) on $V$ if for all $a, b, c \in V$ and all $\lambda_1, \lambda_2 \in F$

1. $\langle \lambda_1 a + \lambda_2 b, c \rangle = \lambda_1 \langle a, c \rangle + \lambda_2 \langle b, c \rangle$;
2. $\langle a, b \rangle = \langle b, a \rangle$;
3. $\langle a, a \rangle \geq 0$;
4. $\langle a, a \rangle = 0 \Rightarrow a = 0$.

- Note that Condition 2 ensures that $\langle a, a \rangle \in \mathbb{R}$.
- If $F$ is $\mathbb{R}$ then commutativity holds. (In the sequel we will assume $F$ to be $\mathbb{R}$.)
- Be warned that the notation is not uniform: $a \cdot b$ and $(a \mid b)$ are two other common notations for denoting the dot product of $a$ and $b$.
- Note the difference between $a \cdot b$ for $a, b \in V$, and $\lambda \cdot a$ for $\lambda \in F$ and $a \in V$!
Norm and Triangle Inequality

Definition 119 (Length)

Based on a dot product on $V$ (over $\mathbb{R}$), we can define the length (or norm) of a vector $a \in V$ induced by that dot product as the following mapping $\| \cdot \|$ from $V$ to $\mathbb{R}$:

$$\|a\| := \sqrt{\langle a, a \rangle}.$$
Norm and Triangle Inequality

**Definition 119 (Length)**

Based on a dot product on $V$ (over $\mathbb{R}$), we can define the *length* (or *norm*) of a vector $a \in V$ induced by that dot product as the following mapping $\| \cdot \|$ from $V$ to $\mathbb{R}$:

$$\|a\| := \sqrt{\langle a, a \rangle}.$$

**Definition 120 (Unit vector, Dt.: Einheitsvektor)**

A vector $a$ is said to be a *unit vector* if $\|a\| = 1$. 

**Lemma 121**

We get the following standard properties of a norm for $\| \cdot \|$ for all $a, b \in V$:

1. $\|a\| \geq 0$;
2. $\|a\| = 0 \iff a = 0$;
3. $\|\lambda a\| = |\lambda| \cdot \|a\| \forall \lambda \in \mathbb{R}$;
4. Triangle Inequality (Dt.: Dreiecksungleichung): $\|a + b\| \leq \|a\| + \|b\|$.
Norm and Triangle Inequality

**Definition 119 (Length)**

Based on a dot product on $V$ (over $\mathbb{R}$), we can define the *length* (or *norm*) of a vector $a \in V$ induced by that dot product as the following mapping $\| \cdot \|$ from $V$ to $\mathbb{R}$:

$$\|a\| := \sqrt{\langle a, a \rangle}.$$ 

**Definition 120 (Unit vector, Dt.: Einheitsvektor)**

A vector $a$ is said to be a *unit vector* if $\|a\| = 1$.

**Lemma 121**

We get the following standard properties of a norm for $\| \cdot \|$ for all $a, b \in V$:

1. $\|a\| \geq 0$;
Norm and Triangle Inequality

Definition 119 (Length)

Based on a dot product on $V$ (over $\mathbb{R}$), we can define the length (or norm) of a vector $a \in V$ induced by that dot product as the following mapping $\|\cdot\|$ from $V$ to $\mathbb{R}$:

$$\|a\| := \sqrt{\langle a, a \rangle}.$$

Definition 120 (Unit vector, Dt.: Einheitsvektor)

A vector $a$ is said to be a unit vector if $\|a\| = 1$.

Lemma 121

We get the following standard properties of a norm for $\|\cdot\|$ for all $a, b \in V$:

1. $\|a\| \geq 0$;
2. $\|a\| = 0 \implies a = 0$;
Norm and Triangle Inequality

**Definition 119 (Length)**

Based on a dot product on \( V \) (over \( \mathbb{R} \)), we can define the *length* (or *norm*) of a vector \( a \in V \) induced by that dot product as the following mapping \( \| \cdot \| \) from \( V \) to \( \mathbb{R} \):

\[
\|a\| := \sqrt{\langle a, a \rangle}.
\]

**Definition 120 (Unit vector, Dt.: Einheitsvektor)**

A vector \( a \) is said to be a *unit vector* if \( \|a\| = 1 \).

**Lemma 121**

We get the following standard properties of a norm for \( \| \cdot \| \) for all \( a, b \in V \):

1. \( \|a\| \geq 0 \);
2. \( \|a\| = 0 \implies a = 0 \);
3. \( \|\lambda a\| = |\lambda| \cdot \|a\| \quad \forall \lambda \in \mathbb{R} \);
Norm and Triangle Inequality

**Definition 119 (Length)**

Based on a dot product on $V$ (over $\mathbb{R}$), we can define the *length* (or *norm*) of a vector $a \in V$ induced by that dot product as the following mapping $\| \cdot \|$ from $V$ to $\mathbb{R}$:

$$\|a\| := \sqrt{\langle a, a \rangle}.$$

**Definition 120 (Unit vector, Dt.: Einheitsvektor)**

A vector $a$ is said to be a *unit vector* if $\|a\| = 1$.

**Lemma 121**

We get the following standard properties of a norm for $\| \cdot \|$ for all $a, b \in V$:

1. $\|a\| \geq 0$;
2. $\|a\| = 0 \implies a = 0$;
3. $\|\lambda a\| = |\lambda| \cdot \|a\| \quad \forall \lambda \in \mathbb{R}$;
4. **Triangle Inequality** (Dt.: Dreiecksungleichung):
   $$\|a + b\| \leq \|a\| + \|b\|.$$
Cauchy-Schwarz Inequality

Lemma 122 (Cauchy-Schwarz Inequality)

∀ a, b ∈ V  |⟨a,b⟩| ≤ ∥a∥ · ∥b∥.
Cauchy-Schwarz Inequality

Lemma 122 (Cauchy-Schwarz Inequality)

\[ \forall a, b \in V \quad |\langle a, b \rangle| \leq \|a\| \cdot \|b\|. \]

Note that, for \( a, b \neq 0 \), the Cauchy-Schwarz inequality implies

\[ -1 \leq \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|} \leq 1. \]

We will make use of this fact when defining angles between vectors.
Cauchy-Schwarz Inequality

Lemma 122 (Cauchy-Schwarz Inequality)

\[ \forall a, b \in V \quad |\langle a, b \rangle| \leq \|a\| \cdot \|b\|. \]

Note that, for \( a, b \neq 0 \), the Cauchy-Schwarz inequality implies

\[ -1 \leq \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|} \leq 1. \]

We will make use of this fact when defining angles between vectors.

Lemma 123 (Pythagoras)

For \( a, b \in V \),

\[ \langle a, b \rangle = 0 \quad \Rightarrow \quad \|a + b\|^2 = \|a\|^2 + \|b\|^2. \]
Cauchy-Schwarz Inequality

**Lemma 122 (Cauchy-Schwarz Inequality)**

\[ \forall a, b \in V \quad |\langle a, b \rangle| \leq \|a\| \cdot \|b\| . \]

- Note that, for \( a, b \neq 0 \), the Cauchy-Schwarz inequality implies
  \[ -1 \leq \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|} \leq 1. \]

  We will make use of this fact when defining angles between vectors.

**Lemma 123 (Pythagoras)**

For \( a, b \in V \),
\[ \langle a, b \rangle = 0 \quad \Rightarrow \quad \|a + b\|^2 = \|a\|^2 + \|b\|^2. \]

**Proof:** Let \( a, b \in V \) with \( \langle a, b \rangle = 0 \). Then
\[ \|a + b\|^2 \]
Cauchy-Schwarz Inequality

**Lemma 122 (Cauchy-Schwarz Inequality)**

$\forall a, b \in V \quad |\langle a, b \rangle| \leq \|a\| \cdot \|b\|.$

- Note that, for $a, b \neq 0$, the Cauchy-Schwarz inequality implies

  $$-1 \leq \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|} \leq 1.$$ 

  We will make use of this fact when defining angles between vectors.

**Lemma 123 (Pythagoras)**

For $a, b \in V$,

$$\langle a, b \rangle = 0 \quad \Rightarrow \quad \|a + b\|^2 = \|a\|^2 + \|b\|^2.$$

**Proof**: Let $a, b \in V$ with $\langle a, b \rangle = 0$. Then

$$\|a + b\|^2 = \langle a + b, a + b \rangle$$
Cauchy-Schwarz Inequality

**Lemma 122 (Cauchy-Schwarz Inequality)**

\[ \forall a, b \in V \quad |\langle a, b \rangle| \leq \|a\| \cdot \|b\|. \]

- Note that, for \(a, b \neq 0\), the Cauchy-Schwarz inequality implies

\[
-1 \leq \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|} \leq 1.
\]

We will make use of this fact when defining angles between vectors.

**Lemma 123 (Pythagoras)**

For \(a, b \in V\),

\[ \langle a, b \rangle = 0 \quad \Rightarrow \quad \|a + b\|^2 = \|a\|^2 + \|b\|^2. \]

**Proof:** Let \(a, b \in V\) with \(\langle a, b \rangle = 0\). Then

\[
\|a + b\|^2 = \langle a + b, a + b \rangle = \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle
\]
**Cauchy-Schwarz Inequality**

**Lemma 122 (Cauchy-Schwarz Inequality)**

\[ \forall a, b \in V \quad |\langle a, b \rangle| \leq \|a\| \cdot \|b\|. \]

Note that, for \( a, b \neq 0 \), the Cauchy-Schwarz inequality implies

\[ -1 \leq \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|} \leq 1. \]

We will make use of this fact when defining angles between vectors.

**Lemma 123 (Pythagoras)**

For \( a, b \in V \),

\[ \langle a, b \rangle = 0 \quad \Rightarrow \quad \|a + b\|^2 = \|a\|^2 + \|b\|^2. \]

**Proof:** Let \( a, b \in V \) with \( \langle a, b \rangle = 0 \). Then

\[
\|a + b\|^2 = \langle a + b, a + b \rangle = \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle
\]
\[
= \langle a, a \rangle + \langle b, b \rangle = \|a\|^2 + \|b\|^2.
\]
Standard Dot Product and Standard Norm on $\mathbb{R}^n$

For $V := \mathbb{R}^n$ for some $n \in \mathbb{N}$, and $a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$ and $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$, it is easy to prove that

$$\langle a, b \rangle := \sum_{i=1}^{n} a_i \cdot b_i = a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_n \cdot b_n$$

does indeed yield a dot product on $\mathbb{R}^n$.  

In the sequel, unless stated otherwise, we will always use this dot product when referring to "the dot product" on $\mathbb{R}^n$ or writing $\langle a, b \rangle$ for $a, b \in \mathbb{R}^n$.

Note that this definition of a dot product and its corresponding norm on $\mathbb{R}^n$ matches our intuitive notion of the distance, $d(p, q)$, of two points $p$ and $q$ in $\mathbb{R}^n$: Their distance is given by the length of the vector from $p$ to $q$, i.e.,

$$d(p, q) := \|q - p\| = \sqrt{\langle q - p, q - p \rangle} = \sqrt{n \sum_{i=1}^{n} (q_i - p_i)^2} = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \ldots + (q_n - p_n)^2}.$$
Standard Dot Product and Standard Norm on \( \mathbb{R}^n \)

- For \( V := \mathbb{R}^n \) for some \( n \in \mathbb{N} \), and \( a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n \) and \( b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n \), it is easy to prove that

\[
\langle a, b \rangle := \sum_{i=1}^{n} a_i \cdot b_i = a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_n \cdot b_n
\]

does indeed yield a dot product on \( \mathbb{R}^n \).

- In the sequel, unless stated otherwise, we will always use this dot product when referring to “the dot product” on \( \mathbb{R}^n \) or writing \( \langle a, b \rangle \) for \( a, b \in \mathbb{R}^n \).
Standard Dot Product and Standard Norm on $\mathbb{R}^n$

- For $V := \mathbb{R}^n$ for some $n \in \mathbb{N}$, and $a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$ and $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$, it is easy to prove that

$$\langle a, b \rangle := \sum_{i=1}^{n} a_i \cdot b_i = a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_n \cdot b_n$$

does indeed yield a dot product on $\mathbb{R}^n$.

- In the sequel, unless stated otherwise, we will always use this dot product when referring to “the dot product” on $\mathbb{R}^n$ or writing $\langle a, b \rangle$ for $a, b \in \mathbb{R}^n$.

- Note that this definition of a dot product and its corresponding norm on $\mathbb{R}^n$ matches our intuitive notion of the distance, $d(p, q)$, of two points $p$ and $q$ in $\mathbb{R}^n$: Their distance is given by the length of the vector from $p$ to $q$, i.e.,

$$d(p, q) := \|q - p\| = \sqrt{\langle q - p, q - p \rangle}$$
Standard Dot Product and Standard Norm on $\mathbb{R}^n$

- For $V := \mathbb{R}^n$ for some $n \in \mathbb{N}$, and $a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$ and $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$, it is easy to prove that

$$\langle a, b \rangle := \sum_{i=1}^{n} a_i \cdot b_i = a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_n \cdot b_n$$

does indeed yield a dot product on $\mathbb{R}^n$.

- In the sequel, unless stated otherwise, we will always use this dot product when referring to “the dot product” on $\mathbb{R}^n$ or writing $\langle a, b \rangle$ for $a, b \in \mathbb{R}^n$.

- Note that this definition of a dot product and its corresponding norm on $\mathbb{R}^n$ matches our intuitive notion of the distance, $d(p, q)$, of two points $p$ and $q$ in $\mathbb{R}^n$: Their distance is given by the length of the vector from $p$ to $q$, i.e.,

$$d(p, q) := \|q - p\| = \sqrt{\langle q - p, q - p \rangle} = \sqrt{\sum_{i=1}^{n} (q_i - p_i) \cdot (q_i - p_i)}$$

$$= \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}.$$
Other Widely Used Norms on $\mathbb{R}^n$

- The norm

$$\|a - b\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$$

is also called $L_2$-norm and then denoted by $\|a - b\|_2$, 

The norm

\[ \|a - b\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2} \]

is also called $L_2$-norm and then denoted by $\|a - b\|_2$, 
Other Widely Used Norms on $\mathbb{R}^n$

- The norm

$$\|a - b\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$$

is also called $L_2$-norm and then denoted by $\|a - b\|_2$, in order to distinguish it from other well-known norms on $\mathbb{R}^n$, such as the $L_1$-norm (Manhattan metric)

$$\|a - b\|_1 := |a_1 - b_1| + |a_2 - b_2| + \cdots + |a_n - b_n|,$$
Other Widely Used Norms on $\mathbb{R}^n$

- The norm
  \[ \|a - b\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2} \]
  is also called $L_2$-norm and then denoted by $\|a - b\|_2$, in order to distinguish it from other well-known norms on $\mathbb{R}^n$, such as the $L_1$-norm (Manhattan metric)
  \[ \|a - b\|_1 := |a_1 - b_1| + |a_2 - b_2| + \cdots + |a_n - b_n|, \]
Other Widely Used Norms on $\mathbb{R}^n$

- The norm
  \[ \|a - b\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2} \]

  is also called $L_2$-norm and then denoted by $\|a - b\|_2$, in order to distinguish it from other well-known norms on $\mathbb{R}^n$, such as the $L_1$-norm (Manhattan metric)
  \[ \|a - b\|_1 := |a_1 - b_1| + |a_2 - b_2| + \cdots + |a_n - b_n|, \]

  or the $L_\infty$-norm (maximum norm)
  \[ \|a - b\|_\infty := \max_{1 \leq i \leq n} |a_i - b_i|. \]
Other Widely Used Norms on $\mathbb{R}^n$

- The norm
  \[ \|a - b\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2} \]

  is also called $L_2$-norm and then denoted by $\|a - b\|_2$, in order to distinguish it from other well-known norms on $\mathbb{R}^n$, such as the $L_1$-norm (Manhattan metric)
  \[ \|a - b\|_1 := |a_1 - b_1| + |a_2 - b_2| + \cdots + |a_n - b_n|, \]

  or the $L_\infty$-norm (maximum norm)
  \[ \|a - b\|_\infty := \max_{1 \leq i \leq n} |a_i - b_i|. \]
Definition 124 (Angle between vectors)

The angle, $\alpha$, between non-zero vectors $a, b \in \mathbb{R}^n$ is given by

$$\cos \alpha := \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|}.$$
Angle

Definition 124 (Angle between vectors)
The angle, \( \alpha \), between non-zero vectors \( a, b \in \mathbb{R}^n \) is given by
\[
\cos \alpha := \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|}.
\]

Definition 125 (Perpendicular, Dt.: senkrecht)
The vectors \( a, b \in \mathbb{R}^n \) are said to be perpendicular (or orthogonal), denoted by \( a \perp b \), if
\[
\langle a, b \rangle = 0.
\]
**Angle**

**Definition 124 (Angle between vectors)**

The *angle*, $\alpha$, *between non-zero vectors* $a, b \in \mathbb{R}^n$ is given by

$$\cos \alpha := \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|}.$$

**Definition 125 (Perpendicular, Dt.: senkrecht)**

The vectors $a, b \in \mathbb{R}^n$ are said to be *perpendicular* (or *orthogonal*), denoted by $a \perp b$, if

$$\langle a, b \rangle = 0.$$
Definition 126 (Parallel)

The non-zero vectors $a, b \in \mathbb{R}^n$ are said to be parallel, denoted by $a \parallel b$, if there exists $\lambda \in \mathbb{R}$ such that

$$a = \lambda b.$$
**Angle and Projection**

**Definition 126 (Parallel)**

The non-zero vectors \( a, b \in \mathbb{R}^n \) are said to be *parallel*, denoted by \( a \parallel b \), if there exists \( \lambda \in \mathbb{R} \) such that

\[
a = \lambda b.
\]

**Lemma 127**

The length of the orthogonal projection of a vector \( b \) onto a non-zero vector \( a \) is given by

\[
\frac{\langle a, b \rangle}{\|a\|}.
\]
Angle and Projection

Definition 126 (Parallel)
The non-zero vectors \( a, b \in \mathbb{R}^n \) are said to be parallel, denoted by \( a \parallel b \), if there exists \( \lambda \in \mathbb{R} \) such that 
\[
a = \lambda b.
\]

Lemma 127
The length of the orthogonal projection of a vector \( b \) onto a non-zero vector \( a \) is given by 
\[
\frac{\langle a, b \rangle}{\| a \|}.
\]

We have 
\[
\langle a, b \rangle = \| a \| \cdot a_1 = \| b \| \cdot b_1.
\]
Definition 126 (Parallel)

The non-zero vectors $a, b \in \mathbb{R}^n$ are said to be parallel, denoted by $a \parallel b$, if there exists $\lambda \in \mathbb{R}$ such that

$$a = \lambda b.$$ 

Lemma 127

The length of the orthogonal projection of a vector $b$ onto a non-zero vector $a$ is given by

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

We have

$$\langle a, b \rangle = \|a\| \cdot a_1 = \|b\| \cdot b_1.$$ 

This symmetry is obvious for vectors of the same length, but it holds even for vectors of different lengths: Scaling one vector scales either its length or its projection! See Slide 231.
Orthonormal Basis of a Vector Space

Definition 128 (Orthogonal basis)

The vectors \( a_1, \ldots, a_n \) form an *orthogonal basis* of a vector space \( V \) (over \( \mathbb{R} \), such as \( \mathbb{R}^n \)) if

1. the vectors \( a_1, \ldots, a_n \) form a basis of \( V \);
2. \( \forall (1 \leq i, j \leq n) \ [i \neq j \Rightarrow \langle a_i, a_j \rangle = 0] \).

Definition 129 (Orthonormal basis)

The vectors \( a_1, \ldots, a_n \) form an *orthonormal basis* of a vector space \( V \) (over \( \mathbb{R} \), such as \( \mathbb{R}^n \)) if

1. the vectors \( a_1, \ldots, a_n \) form a basis of \( V \);
2. \( \forall (1 \leq i, j \leq n) \quad \langle a_i, a_j \rangle = \delta_{ij} \).

The algorithm by Gram-Schmidt can be used to transform an arbitrary basis into an orthonormal basis.

Lemma 130

An \( n \times n \) matrix \( A \in \mathbb{M}_{n \times n}(\mathbb{R}) \) is orthogonal if and only if its columns form an orthonormal basis of \( \mathbb{R}^n \).
Orthonormal Basis of a Vector Space

**Definition 128 (Orthogonal basis)**

The vectors $a_1, \ldots, a_n$ form an *orthogonal basis* of a vector space $V$ (over $\mathbb{R}$, such as $\mathbb{R}^n$) if

1. the vectors $a_1, \ldots, a_n$ form a basis of $V$;
Definition 128 (Orthogonal basis)

The vectors $a_1, \ldots, a_n$ form an orthogonal basis of a vector space $V$ (over $\mathbb{R}$, such as $\mathbb{R}^n$) if

1. the vectors $a_1, \ldots, a_n$ form a basis of $V$;
2. $\forall (1 \leq i, j \leq n) \ [i \neq j \Rightarrow \langle a_i, a_j \rangle = 0]$.

Definition 129 (Orthonormal basis)

The vectors $a_1, \ldots, a_n$ form an orthonormal basis of a vector space $V$ (over $\mathbb{R}$, such as $\mathbb{R}^n$) if

1. the vectors $a_1, \ldots, a_n$ form a basis of $V$;
2. $\forall (1 \leq i, j \leq n) \ [i \neq j \Rightarrow \langle a_i, a_j \rangle = \delta_{ij}]$.

The algorithm by Gram-Schmidt can be used to transform an arbitrary basis into an orthonormal basis.
**Orthonormal Basis of a Vector Space**

**Definition 128 (Orthogonal basis)**

The vectors $a_1, \ldots, a_n$ form an *orthogonal basis* of a vector space $V$ (over $\mathbb{R}$, such as $\mathbb{R}^n$) if

1. the vectors $a_1, \ldots, a_n$ form a basis of $V$;
2. $\forall (1 \leq i, j \leq n) \ [i \neq j \implies \langle a_i, a_j \rangle = 0]$.

**Definition 129 (Orthonormal basis)**

The vectors $a_1, \ldots, a_n$ form an *orthonormal basis* of a vector space $V$ (over $\mathbb{R}$, such as $\mathbb{R}^n$) if

1. the vectors $a_1, \ldots, a_n$ form a basis of $V$;
2. $\forall (1 \leq i, j \leq n) \ \langle a_i, a_j \rangle = \delta_{ij}$. 

The algorithm by Gram-Schmidt can be used to transform an arbitrary basis into an orthonormal basis.
Orthonormal Basis of a Vector Space

**Definition 128 (Orthogonal basis)**

The vectors $a_1, \ldots, a_n$ form an *orthogonal basis* of a vector space $V$ (over $\mathbb{R}$, such as $\mathbb{R}^n$) if

1. the vectors $a_1, \ldots, a_n$ form a basis of $V$;
2. $\forall (1 \leq i, j \leq n) \ [i \neq j \Rightarrow \langle a_i, a_j \rangle = 0]$.

**Definition 129 (Orthonormal basis)**

The vectors $a_1, \ldots, a_n$ form an *orthonormal basis* of a vector space $V$ (over $\mathbb{R}$, such as $\mathbb{R}^n$) if

1. the vectors $a_1, \ldots, a_n$ form a basis of $V$;
2. $\forall (1 \leq i, j \leq n) \ [\langle a_i, a_j \rangle = \delta_{ij}]$.

The algorithm by Gram-Schmidt can be used to transform an arbitrary basis into an orthonormal basis.
Orthonormal Basis of a Vector Space

**Definition 128 (Orthogonal basis)**
The vectors \(a_1, \ldots, a_n\) form an *orthogonal basis* of a vector space \(V\) (over \(\mathbb{R}\), such as \(\mathbb{R}^n\)) if
1. the vectors \(a_1, \ldots, a_n\) form a basis of \(V\);
2. \(\forall (1 \leq i, j \leq n) \; [i \neq j \Rightarrow \langle a_i, a_j \rangle = 0]\).

**Definition 129 (Orthonormal basis)**
The vectors \(a_1, \ldots, a_n\) form an *orthonormal basis* of a vector space \(V\) (over \(\mathbb{R}\), such as \(\mathbb{R}^n\)) if
1. the vectors \(a_1, \ldots, a_n\) form a basis of \(V\);
2. \(\forall (1 \leq i, j \leq n) \; \langle a_i, a_j \rangle = \delta_{ij}\).

- The algorithm by Gram-Schmidt can be used to transform an arbitrary basis into an orthonormal basis.

**Lemma 130**
An \(n \times n\) matrix \(A \in M_{n \times n}(\mathbb{R})\) is orthogonal if and only if its columns form an orthonormal basis of \(\mathbb{R}^n\).
Basic Linear Algebra

- Matrices
- Linear Equations
- Determinants
- Eigenvalues and Eigenvectors
- Dot Product and Norm
- Vector Cross-Product
- Quaternions
Definition 131 (Cross-product, Dt.: Kreuzprodukt)

Let \( a = (a_x, a_y, a_z), b = (b_x, b_y, b_z) \in \mathbb{R}^3 \). The (vector) cross-product of \( a \) and \( b \) is given by

\[
\begin{align*}
\begin{pmatrix}
\begin{vmatrix}
a_y & b_y \\
a_z & b_z \\
\end{vmatrix} \\
-\begin{vmatrix}
a_x & b_x \\
a_z & b_z \\
\end{vmatrix} \\
\begin{vmatrix}
a_x & b_x \\
a_y & b_y \\
\end{vmatrix}
\end{pmatrix}
&= \begin{pmatrix}
a_y \cdot b_z - a_z \cdot b_y \\
a_z \cdot b_x - a_x \cdot b_z \\
a_x \cdot b_y - a_y \cdot b_x
\end{pmatrix}.
\end{align*}
\]

- This cross-product is only defined in \( \mathbb{R}^3 \)!
Vector Cross-Product in $\mathbb{R}^3$

Definition 131 (Cross-product, Dt.: Kreuzprodukt)

Let $a = (a_x, a_y, a_z), b = (b_x, b_y, b_z) \in \mathbb{R}^3$. The (vector) cross-product of $a$ and $b$ is given by

$$a \times b := \begin{pmatrix}
det \begin{pmatrix} a_y & b_y \\ a_z & b_z \end{pmatrix} \\
- \det \begin{pmatrix} a_x & b_x \\ a_z & b_z \end{pmatrix} \\
det \begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix}
\end{pmatrix} = \begin{pmatrix} a_y \cdot b_z - a_z \cdot b_y \\ a_z \cdot b_x - a_x \cdot b_z \\ a_x \cdot b_y - a_y \cdot b_x \end{pmatrix}.$$

- This cross-product is only defined in $\mathbb{R}^3$!
- Some authors like to define a “cross-product” for two vectors $a, b \in \mathbb{R}^2$, with $a := (a_x, a_y)$ and $b := (b_x, b_y)$, as follows:

$$a \times b := \det \begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix} = a_x \cdot b_y - a_y \cdot b_x$$

- Note, however, that its properties are different from those of Definition 131.
Properties of the Cross-Product: Orientation of the Resulting Vector

**Right-hand rule (Dt.: Drei-Finger-Regel)**

The orientation of the vector $a \times b$ can be memorized by the *right-hand rule*: Point the forefinger of your right hand into direction $a$ and point the middle finger into direction $b$. Then your thumb will point into the direction of $a \times b$.

[Image credit: Wikipedia.]
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of $2 \times 2$ and $3 \times 3$ determinants:

1. $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$;
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of $2 \times 2$ and $3 \times 3$ determinants:

1. $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$;
2. $a \times a = 0$;
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of $2 \times 2$ and $3 \times 3$ determinants:

1. $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$;
2. $a \times a = 0$;
3. $a \times b = -(b \times a) = -b \times a$;
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of $2 \times 2$ and $3 \times 3$ determinants:

1. $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$, $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$;
2. $\mathbf{a} \times \mathbf{a} = 0$;
3. $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a}) = -\mathbf{b} \times \mathbf{a}$;
4. $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$;
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of $2 \times 2$ and $3 \times 3$ determinants:

1. $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$;
2. $a \times a = 0$;
3. $a \times b = -(b \times a) = -b \times a$;
4. $a \times (b + c) = a \times b + a \times c$;
5. $(\lambda a) \times (\mu b) = \lambda \mu (a \times b)$;
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of $2 \times 2$ and $3 \times 3$ determinants:

1. $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$;
2. $a \times a = 0$;
3. $a \times b = -(b \times a) = -b \times a$;
4. $a \times (b + c) = a \times b + a \times c$;
5. $(\lambda a) \times (\mu b) = \lambda \mu (a \times b)$;
6. $\langle a, b \times c \rangle = \det \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} = \langle a \times b, c \rangle$;
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of $2 \times 2$ and $3 \times 3$ determinants:

1. $e_1 \times e_2 = e_3, \ e_2 \times e_3 = e_1, \ e_3 \times e_1 = e_2$;
2. $a \times a = 0$;
3. $a \times b = -(b \times a) = -b \times a$;
4. $a \times (b + c) = a \times b + a \times c$;
5. $(\lambda a) \times (\mu b) = \lambda \mu (a \times b)$;
6. $\langle a, b \times c \rangle = \det \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} = \langle a \times b, c \rangle$;
7. $\langle a, a \times b \rangle = 0 = \langle b, a \times b \rangle$;
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of \(2 \times 2\) and \(3 \times 3\) determinants:

1. \(e_1 \times e_2 = e_3,\ e_2 \times e_3 = e_1,\ e_3 \times e_1 = e_2;\)
2. \(a \times a = 0;\)
3. \(a \times b = -(b \times a) = -b \times a;\)
4. \(a \times (b + c) = a \times b + a \times c;\)
5. \((\lambda a) \times (\mu b) = \lambda \mu (a \times b);\)
6. \(\langle a, b \times c \rangle = \det \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} = \langle a \times b, c \rangle;\)
7. \(\langle a, a \times b \rangle = 0 = \langle b, a \times b \rangle;\)
8. \(\|a \times b\| = \sqrt{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2};\)
Properties of the Cross-Product

Lemma 132

The following properties of the vector cross-product follow from the properties of $2 \times 2$ and $3 \times 3$ determinants:

1. $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$;
2. $a \times a = 0$;
3. $a \times b = -(b \times a) = -b \times a$;
4. $a \times (b + c) = a \times b + a \times c$;
5. $(\lambda a) \times (\mu b) = \lambda \mu (a \times b)$;
6. $\langle a, b \times c \rangle = \det \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} = \langle a \times b, c \rangle$;
7. $\langle a, a \times b \rangle = 0 = \langle b, a \times b \rangle$;
8. $\|a \times b\| = \sqrt{\|a\|^2 \|b\|^2 - (\langle a, b \rangle)^2}$;
9. For non-zero vectors $a, b$, if $\alpha$ is the angle between $a$ and $b$, then

$$\sin \alpha = \frac{\|a \times b\|}{\|a\| \cdot \|b\|}.$$
Properties of the Cross-Product

- In particular, $a \times b$ is perpendicular on both $a$ and $b$!
Properties of the Cross-Product

- In particular, $a \times b$ is perpendicular on both $a$ and $b$!

**Lemma 133**

If $u, v, w$ are distinct non-collinear points in $\mathbb{R}^3$, then the area of the triangle $\Delta(u, v, w)$ equals

$$\frac{1}{2} \left\| uv \times uw \right\|.$$
Properties of the Cross-Product

In particular, $a \times b$ is perpendicular on both $a$ and $b$!

**Lemma 133**

If $u, v, w$ are distinct non-collinear points in $\mathbb{R}^3$, then the area of the triangle $\Delta(u, v, w)$ equals

$$
\frac{1}{2} \| uv \times uw \|.
$$

This is not completely surprising since, for points in $\mathbb{R}^2$ with $u_z = v_z = w_z := 0$, this is nothing but a re-statement of Theorem 109. We will later on resort to linear transformations to shed some additional light onto this claim.
Properties of the Cross-Product

- In particular, $a \times b$ is perpendicular on both $a$ and $b$!

**Lemma 133**

If $u, v, w$ are distinct non-collinear points in $\mathbb{R}^3$, then the area of the triangle $\triangle(u, v, w)$ equals

$$\frac{1}{2} \|uv \times uw\|.$$

- This is not completely surprising since, for points in $\mathbb{R}^2$ with $u_z = v_z = w_z = 0$, this is nothing but a re-statement of Theorem 109. We will later on resort to linear transformations to shed some additional light onto this claim.

**Lemma 134**

If $u, v, w$ are distinct non-collinear points in $\mathbb{R}^3$, then the distance $d$ from $w$ to the line through $u$ and $v$ is given by

$$d = \frac{\|uv \times uw\|}{\|uv\|}.$$
Orthogonal Frame

- Assume that the vector $\nu_1 := (1, 2, 3)$ is a tangent vector to the curve $\gamma$ at the point $\gamma(6)$.

An orthogonal frame can be obtained by taking a vector cross-product of two suitable vectors:

$$\nu_2 := \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

$$\nu_3 := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

Then $\nu_1 \perp \nu_2$, $\nu_1 \perp \nu_3$ and $\nu_2 \perp \nu_3$. 

Orthogonal Frame

- Assume that the vector $\nu_1 := (1, 2, 3)$ is a tangent vector to the curve $\gamma$ at the point $\gamma(6)$.
- An orthogonal frame can be obtained by taking a vector cross-product of two suitable vectors:

  $\nu_2 := \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$

Then $\nu_1 \perp \nu_2$, $\nu_1 \perp \nu_3$ and $\nu_2 \perp \nu_3$. 
Orthogonal Frame

- Assume that the vector $\nu_1 := (1, 2, 3)$ is a tangent vector to the curve $\gamma$ at the point $\gamma(6)$.
- An orthogonal frame can be obtained by taking a vector cross-product of two suitable vectors:

\[ \nu_2 := \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \]

\[ \nu_3 := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ -3 & 0 & 1 \\ 1 & -2 & 2 \end{pmatrix} = \begin{pmatrix} -3 \\ -6 \\ 5 \end{pmatrix} \]
Orthogonal Frame

- Assume that the vector \( \nu_1 := (1, 2, 3) \) is a tangent vector to the curve \( \gamma \) at the point \( \gamma(6) \).
- An orthogonal frame can be obtained by taking a vector cross-product of two suitable vectors:

\[
\nu_2 := \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}
\]

\[
\nu_3 := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} -3 \\ -6 \\ 5 \end{pmatrix}
\]

- Then \( \nu_1 \perp \nu_2 \), \( \nu_1 \perp \nu_3 \) and \( \nu_2 \perp \nu_3 \).
Basic Linear Algebra

- Matrices
- Linear Equations
- Determinants
- Eigenvalues and Eigenvectors
- Dot Product and Norm
- Vector Cross-Product
- Quaternions
Definition 135 (Quaternions)

The set of quaternions, \( \mathbb{H} \), is given by quadrupels of real numbers together with operations \( + : \mathbb{H} \times \mathbb{H} \to \mathbb{H} \) and \( \cdot : \mathbb{H} \times \mathbb{H} \to \mathbb{H} \) defined as follows for all \( \mathcal{P}_1, \mathcal{P}_2 \in \mathbb{H} \), with \( \mathcal{P}_1 := (s_1, v_1) \) and \( \mathcal{P}_2 := (s_2, v_2) \) where \( s_1, s_2 \in \mathbb{R} \) and \( v_1, v_2 \in \mathbb{R}^3 \):

\[
\mathcal{P}_1 + \mathcal{P}_2 := (s_1 + s_2, v_1 + v_2),
\]

\[
\mathcal{P}_1 \cdot \mathcal{P}_2 := (s_1 s_2 - \langle v_1, v_2 \rangle, s_1 v_2 + s_2 v_1 + v_1 \times v_2).
\]
Definition 135 (Quaternions)

The set of *quaternions*, \( \mathbb{H} \), is given by quadrupels of real numbers together with operations \( + : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H} \) and \( \cdot : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H} \) defined as follows for all \( P_1, P_2 \in \mathbb{H} \), with \( P_1 := (s_1, v_1) \) and \( P_2 := (s_2, v_2) \) where \( s_1, s_2 \in \mathbb{R} \) and \( v_1, v_2 \in \mathbb{R}^3 \):

\[
P_1 + P_2 := (s_1 + s_2, v_1 + v_2),
\]

Definition 136 (Pure quaternion)

A quaternion \( (s, v) \), with \( s \in \mathbb{R} \) and \( v \in \mathbb{R}^3 \), is called *pure* if its real part \( s \) equals zero.

We identify the set \( \{ (s, 0) \in \mathbb{H} : s \in \mathbb{R} \} \) with \( \mathbb{R} \), and \( \{ (0, v) \in \mathbb{H} : v \in \mathbb{R}^3 \} \) with \( \mathbb{R}^3 \).

Discovered by William R. Hamilton in 1843 at Dublin, Ireland: Here as he walked by on the 16th of October 1843, Sir William Rowan Hamilton in a flash of genius discovered the fundamental formula for quaternion multiplication, \( i^2 = j^2 = k^2 = ijk = -1 \), and cut it on a stone of this bridge.
**Definition 135 (Quaternions)**

The set of *quaternions*, \( \mathbb{H} \), is given by quadrupels of real numbers together with operations \( + : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H} \) and \( \cdot : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H} \) defined as follows for all \( \mathcal{P}_1, \mathcal{P}_2 \in \mathbb{H} \), with \( \mathcal{P}_1 := (s_1, \mathbf{v}_1) \) and \( \mathcal{P}_2 := (s_2, \mathbf{v}_2) \) where \( s_1, s_2 \in \mathbb{R} \) and \( \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3 \):

\[
\mathcal{P}_1 + \mathcal{P}_2 := (s_1 + s_2, \mathbf{v}_1 + \mathbf{v}_2),
\]

\[
\mathcal{P}_1 \cdot \mathcal{P}_2 := (s_1 s_2 - \langle \mathbf{v}_1, \mathbf{v}_2 \rangle, s_1 \mathbf{v}_2 + s_2 \mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2).
\]
Definition 135 (Quaternions)

The set of *quaternions*, \(\mathbb{H}\), is given by quadrupels of real numbers together with operations \(+: \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H}\) and \(\cdot: \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H}\) defined as follows for all \(\mathcal{P}_1, \mathcal{P}_2 \in \mathbb{H}\), with \(\mathcal{P}_1 := (s_1, \mathbf{v}_1)\) and \(\mathcal{P}_2 := (s_2, \mathbf{v}_2)\) where \(s_1, s_2 \in \mathbb{R}\) and \(\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3\):

\[
\mathcal{P}_1 + \mathcal{P}_2 := (s_1 + s_2, \mathbf{v}_1 + \mathbf{v}_2),
\]

\[
\mathcal{P}_1 \cdot \mathcal{P}_2 := (s_1s_2 - \langle \mathbf{v}_1, \mathbf{v}_2 \rangle, s_1 \mathbf{v}_2 + s_2 \mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2).
\]

Definition 136 (Pure quaternion)

A quaternion \((s, \mathbf{v})\), with \(s \in \mathbb{R}\) and \(\mathbf{v} \in \mathbb{R}^3\), is called *pure* if its real part \(s\) equals zero.
**Quaternions** $\mathbb{H}$

**Definition 135 (Quaternions)**
The set of *quaternions*, $\mathbb{H}$, is given by quadrupels of real numbers together with operations $+: \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H}$ and $\cdot: \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H}$ defined as follows for all $\mathcal{P}_1, \mathcal{P}_2 \in \mathbb{H}$, with $\mathcal{P}_1 := (s_1, \mathbf{v}_1)$ and $\mathcal{P}_2 := (s_2, \mathbf{v}_2)$ where $s_1, s_2 \in \mathbb{R}$ and $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$:

- $\mathcal{P}_1 + \mathcal{P}_2 := (s_1 + s_2, \mathbf{v}_1 + \mathbf{v}_2)$,
- $\mathcal{P}_1 \cdot \mathcal{P}_2 := (s_1 s_2 - \langle \mathbf{v}_1, \mathbf{v}_2 \rangle, s_1 \mathbf{v}_2 + s_2 \mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2)$.

**Definition 136 (Pure quaternion)**
A quaternion $(s, \mathbf{v})$, with $s \in \mathbb{R}$ and $\mathbf{v} \in \mathbb{R}^3$, is called *pure* if its real part $s$ equals zero.

- We identify the set $\{(s, 0) \in \mathbb{H} : s \in \mathbb{R}\}$ with $\mathbb{R}$, and $\{(0, \mathbf{v}) \in \mathbb{H} : \mathbf{v} \in \mathbb{R}^3\}$ with $\mathbb{R}^3$. 

Discovered by William R. Hamilton in 1843 at Dublin, Ireland: Here as he walked by on the 16th of October 1843, Sir William Rowan Hamilton in a flash of genius discovered the fundamental formula for quaternion multiplication, $i^2 = j^2 = k^2 = ijk = -1$, and cut it on a stone of this bridge.
Definition 135 (Quaternions)

The set of quaternions, \( \mathbb{H} \), is given by quadrupels of real numbers together with operations \( + : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H} \) and \( \cdot : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H} \) defined as follows for all \( \mathcal{P}_1, \mathcal{P}_2 \in \mathbb{H} \), with \( \mathcal{P}_1 := (s_1, v_1) \) and \( \mathcal{P}_2 := (s_2, v_2) \) where \( s_1, s_2 \in \mathbb{R} \) and \( v_1, v_2 \in \mathbb{R}^3 \): 

\[
\mathcal{P}_1 + \mathcal{P}_2 := (s_1 + s_2, v_1 + v_2),
\]

\[
\mathcal{P}_1 \cdot \mathcal{P}_2 := (s_1 s_2 - \langle v_1, v_2 \rangle, s_1 v_2 + s_2 v_1 + v_1 \times v_2).
\]

Definition 136 (Pure quaternion)

A quaternion \((s, v)\), with \( s \in \mathbb{R} \) and \( v \in \mathbb{R}^3 \), is called pure if its real part \( s \) equals zero.

- We identify the set \( \{(s, 0) \in \mathbb{H} : s \in \mathbb{R}\} \) with \( \mathbb{R} \), and \( \{(0, v) \in \mathbb{H} : v \in \mathbb{R}^3\} \) with \( \mathbb{R}^3 \).
- Discovered by William R. Hamilton in 1843 at Dublin, Ireland:

  Here as he walked by on the 16th of October 1843, Sir William Rowan Hamilton in a flash of genius discovered the fundamental formula for quaternion multiplication, \( i^2 = j^2 = k^2 = ijk = -1 \), and cut it on a stone of this bridge.
Lemma 137

A quaternion $\mathcal{P}$ can also be regarded as an extension of complex numbers as follows:

$$\mathcal{P} := s + ia + jb + kc, \quad \text{with} \quad s, a, b, c \in \mathbb{R},$$
Lemma 137

A quaternion $\mathcal{P}$ can also be regarded as an extension of complex numbers as follows:

$$\mathcal{P} := s + ia + jb + kc, \quad \text{with} \quad s, a, b, c \in \mathbb{R},$$

where standard arithmetic for real numbers is applied and where the multiplication of the imaginary elements $i, j, \text{and } k$ is defined as

$$i^2 = j^2 = k^2 := -1 \quad \text{and} \quad ijk := -1.$$
Quaternions

Lemma 137
A quaternion \( P \) can also be regarded as an extension of complex numbers as follows:

\[
P := s + ia + jb + kc, \quad \text{with} \quad s, a, b, c \in \mathbb{R},
\]

where standard arithmetic for real numbers is applied and where the multiplication of the imaginary elements \( i, j, \) and \( k \) is defined as

\[
i^2 = j^2 = k^2 := -1 \quad \text{and} \quad ijk := -1.
\]

Lemma 138
Lemma 137 implies for \( i, j, k \) that

\[
jk = -kj = i \quad \text{and} \quad ki = -ik = j \quad \text{and} \quad ij = -ji = k.
\]
Lemma 137

A quaternion $\mathcal{P}$ can also be regarded as an extension of complex numbers as follows:

$$\mathcal{P} := s + ia + jb + kc, \quad \text{with} \quad s, a, b, c \in \mathbb{R},$$

where standard arithmetic for real numbers is applied and where the multiplication of the imaginary elements $i, j, k$ is defined as

$$i^2 = j^2 = k^2 := -1 \quad \text{and} \quad ijk := -1.$$

Lemma 138

Lemma 137 implies for $i, j, k$ that

$$jk = -kj = i \quad \text{and} \quad ki = -ik = j \quad \text{and} \quad ij = -ji = k.$$ 

- Hence, a quaternion $\mathcal{P}$ can be seen as either $(s, (a, b, c))$ or $s + ia + jb + kc$, with $s, a, b, c \in \mathbb{R}$.
- It is common to switch between the two notations depending on which one is more suitable for a particular application.
Definition 139 (Conjugate, Dt.: konjugiertes Quaternion)

The *conjugate* of a quaternion $\mathcal{P} = (s, v) = (s, (a, b, c)) \in \mathbb{H}$ is defined as

$$\overline{\mathcal{P}} := (s, -v) = s - ia - jb - kc.$$
Quaternions

**Definition 139 (Conjugate, Dt.: konjugiertes Quaternion)**

The *conjugate* of a quaternion $P = (s, v) = (s, (a, b, c)) \in \mathbb{H}$ is defined as

$$
\overline{P} := (s, -v) = s - ia - jb - kc.
$$

**Definition 140 (Unit quaternion, Dt.: Einheitsquaternion)**

The *norm* of a quaternion $P = (s, v) = (s, (a, b, c)) \in \mathbb{H}$ is defined as

$$
\|P\| := \sqrt{s^2 + \|v\|^2} = \sqrt{s^2 + a^2 + b^2 + c^2}.
$$

A *unit quaternion* is a quaternion whose norm is 1.
**Quaternions**

**Definition 139 (Conjugate, Dt.: konjugiertes Quaternion)**

The *conjugate* of a quaternion $\mathcal{P} = (s, v) = (s, (a, b, c)) \in \mathbb{H}$ is defined as

$$
\overline{\mathcal{P}} := (s, -v) = s - ia - jb - kc.
$$

**Definition 140 (Unit quaternion, Dt.: Einheitsquaternion)**

The *norm* of a quaternion $\mathcal{P} = (s, v) = (s, (a, b, c)) \in \mathbb{H}$ is defined as

$$
\|\mathcal{P}\| := \sqrt{s^2 + \|v\|^2} = \sqrt{s^2 + a^2 + b^2 + c^2}.
$$

A *unit quaternion* is a quaternion whose norm is 1.

**Definition 141 (Multiplicative inverse)**

The *multiplicative inverse* $\mathcal{P}^{-1}$ of a quaternion $\mathcal{P} = (s, v) \in \mathbb{H}$, with $\mathcal{P} \neq 0$, is defined as

$$
\mathcal{P}^{-1} := \frac{\overline{\mathcal{P}}}{\|\mathcal{P}\|^2} = \frac{1}{\|\mathcal{P}\|^2} (s, -v).
$$
Lemma 142

For all $P, Q \in \mathbb{H}$, we have

$$\overline{P} = P \quad \text{and} \quad \overline{P + Q} = \overline{Q + P} \quad \text{and} \quad \overline{P \cdot Q} = \overline{Q \cdot P}.$$
Lemma 142
For all $P, Q \in \mathbb{H}$, we have
\[
\overline{P} = P \quad \text{and} \quad P + Q = \overline{Q} + \overline{P} \quad \text{and} \quad \overline{P \cdot Q} = \overline{P} \cdot \overline{Q}.
\]

Lemma 143
For all $P, Q \in \mathbb{H}$ with $P, Q \neq 0$, we have
\[
(P^{-1})^{-1} = P \quad \text{and} \quad (P \cdot Q)^{-1} = Q^{-1} \cdot P^{-1}.
\]
Quaternion Algebra

**Lemma 142**
For all $P, Q \in \mathbb{H}$, we have

$$\overline{P} = P \quad \text{and} \quad \overline{P + Q} = \overline{Q + P} \quad \text{and} \quad \overline{P \cdot Q} = \overline{Q} \cdot \overline{P}.$$  

**Lemma 143**
For all $P, Q \in \mathbb{H}$ with $P, Q \neq 0$, we have

$$(P^{-1})^{-1} = P \quad \text{and} \quad (P \cdot Q)^{-1} = Q^{-1} \cdot P^{-1}.$$  

**Lemma 144**
The inverse of a unit quaternion and the product of unit quaternions are themselves unit quaternions.

Note: The multiplication of quaternions is associative but not commutative!

A unit quaternion can be represented by $(\cos \phi, u \sin \phi)$, where $u \in \mathbb{R}^3$ with $\|u\| = 1$.

Important application in graphics: Modeling and interpolating spatial rotations.
Quaternion Algebra

Lemma 142
For all $P, Q \in \mathbb{H}$, we have

$$\overline{P} = P \quad \text{and} \quad \overline{P + Q} = \overline{Q + P} \quad \text{and} \quad \overline{P \cdot Q} = \overline{Q} \cdot \overline{P}.$$  

Lemma 143
For all $P, Q \in \mathbb{H}$ with $P, Q \neq 0$, we have

$$(P^{-1})^{-1} = P \quad \text{and} \quad (P \cdot Q)^{-1} = Q^{-1} \cdot P^{-1}.$$  

Lemma 144
The inverse of a unit quaternion and the product of unit quaternions are themselves unit quaternions.

- Note: The multiplication of quaternions is associative but not commutative!
- A unit quaternion can be represented by $(\cos \phi, u \sin \phi)$, where $u \in \mathbb{R}^3$ with $\|u\| = 1$.
- Important application in graphics: Modeling and interpolating spatial rotations.
Geometric Objects
- Lines and Planes
- Circles and Spheres
- Conics
- Curves and Surfaces
- Polygons and Polyhedra
Geometric Objects

- Lines and Planes
- Circles and Spheres
- Conics
- Curves and Surfaces
- Polygons and Polyhedra
Definition 145 (Straight line, Dt.: Gerade)

For two distinct points \( p, q \in \mathbb{R}^n \), the *straight line* defined by \( p, q \) is the set

\[
\ell(p, q) := \{ p + \lambda \cdot pq : \lambda \in \mathbb{R} \}.
\]
Lines and Straight-Line Segments

**Definition 145 (Straight line, Dt.: Gerade)**

For two distinct points \( p, q \in \mathbb{R}^n \), the *straight line* defined by \( p, q \) is the set

\[
\ell(p, q) := \{ p + \lambda \cdot pq : \lambda \in \mathbb{R} \}.
\]

- Recall that \( pq := q - p \).
- \( p + \lambda \cdot pq \) is the so-called *parametric representation* of \( \ell(p, q) \).
Lines and Straight-Line Segments

**Definition 145 (Straight line, Dt.: Gerade)**

For two distinct points \( p, q \in \mathbb{R}^n \), the *straight line* defined by \( p, q \) is the set

\[
\ell(p, q) := \{ p + \lambda \cdot pq : \lambda \in \mathbb{R} \}.
\]

- Recall that \( pq := q - p \).
- \( p + \lambda \cdot pq \) is the so-called *parametric representation* of \( \ell(p, q) \).

**Definition 146 (Ray, Dt.: Strahl, Halbgerade)**

For two distinct points \( p, q \in \mathbb{R}^n \), the *ray* starting at \( p \) through \( q \) is the set

\[
\{ p + \lambda \cdot pq : \lambda \in \mathbb{R}_0^+ \}.
\]
Definition 147 (Straight-line segment, Dt.: Geradensegment, Strecke)

For two distinct points \( p, q \in \mathbb{R}^n \), the (closed) straight-line segment defined by \( p, q \) is the set

\[
\overline{pq} := \{ p + \lambda \cdot pq : \lambda \in [0, 1] \}.
\]
Lines and Straight-Line Segments

**Definition 147 (Straight-line segment, Dt.: Geradensegment, Strecke)**

For two distinct points $p, q \in \mathbb{R}^n$, the *(closed)* straight-line segment defined by $p, q$ is the set

$$pq := \{p + \lambda \cdot pq : \lambda \in [0, 1]\}.$$

**Definition 148 (Open straight-line segment)**

For two distinct points $p, q \in \mathbb{R}^n$, the *open* straight-line segment defined by $p, q$ is the set

$$\{p + \lambda \cdot pq : \lambda \in ]0, 1[\}.$$
Lines and Straight-Line Segments

Definition 147 (Straight-line segment, Dt.: Geradensegment, Strecke)

For two distinct points \( p, q \in \mathbb{R}^n \), the \textit{(closed) straight-line segment} defined by \( p, q \) is the set

\[
\overline{pq} := \{ p + \lambda \cdot pq : \lambda \in [0, 1] \}.
\]

Definition 148 (Open straight-line segment)

For two distinct points \( p, q \in \mathbb{R}^n \), the \textit{open straight-line segment} defined by \( p, q \) is the set

\[
\{ p + \lambda \cdot pq : \lambda \in ]0, 1[ \}.
\]

- Hence, the endpoints are excluded from an open straight-line segment.
Lemma 149

For every pair of distinct points $p, q \in \mathbb{R}^2$, there exist $n \in \mathbb{R}^2$ and $c \in \mathbb{R}$ such that

$$\ell(p, q) = \{ u \in \mathbb{R}^2 : \langle u, n \rangle = c \}.$$
Lemma 149

For every pair of distinct points \( p, q \in \mathbb{R}^2 \), there exist \( n \in \mathbb{R}^2 \) and \( c \in \mathbb{R} \) such that

\[
\ell(p, q) = \{ u \in \mathbb{R}^2 : \langle u, n \rangle = c \}.
\]

\( \langle u, n \rangle = c \) is the so-called *equational representation* of \( \ell(p, q) \).
Lemma 149

For every pair of distinct points \( p, q \in \mathbb{R}^2 \), there exist \( n \in \mathbb{R}^2 \) and \( c \in \mathbb{R} \) such that

\[
\ell(p, q) = \{ u \in \mathbb{R}^2 : \langle u, n \rangle = c \}.
\]

- \( \langle u, n \rangle = c \) is the so-called *equational representation* of \( \ell(p, q) \).
- Standard formulation according to high school math:

\[
a \cdot x + b \cdot y = c, \quad \text{with} \quad n := \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{and} \quad u := \begin{pmatrix} x \\ y \end{pmatrix}.
\]
Lines in $\mathbb{R}^2$

**Lemma 149**

For every pair of distinct points $p, q \in \mathbb{R}^2$, there exist $n \in \mathbb{R}^2$ and $c \in \mathbb{R}$ such that

$$\ell(p, q) = \{u \in \mathbb{R}^2 : \langle u, n \rangle = c \}.$$ 

- $\langle u, n \rangle = c$ is the so-called *equational representation* of $\ell(p, q)$.
- Standard formulation according to high school math:

$$a \cdot x + b \cdot y = c,$$

with $n := \begin{pmatrix} a \\ b \end{pmatrix}$ and $u := \begin{pmatrix} x \\ y \end{pmatrix}$.

- Note that $\langle n, pq \rangle = 0$ holds for every such $n$. That is, the vector $n$ is a normal vector of $\ell(p, q)$. We have

$$n = \lambda \begin{pmatrix} -pq_y \\ pq_x \end{pmatrix}$$

for some non-zero scalar $\lambda \in \mathbb{R}.$
Lines in $\mathbb{R}^2$

**Definition 150 (Hessian normal form, Dt.: Hessische Normalform)**

A line equation $\langle u, n \rangle = c$ for $\ell(p, q)$ is said to be in *Hessian normal form* if $n$ (as specified in Lem. 149) is a unit vector.
**Definition 150 (Hessian normal form, Dt.: Hessische Normalform)**

A line equation \( \langle u, n \rangle = c \) for \( \ell(p, q) \) is said to be in *Hessian normal form* if \( n \) (as specified in Lem. 149) is a unit vector.

**Lemma 151**

The (signed) minimum distance \( d \) of a point \( A \in \mathbb{R}^2 \) from \( \ell(p, q) \), with \( \ell(p, q) = \{ u \in \mathbb{R}^2 : \langle u, n \rangle = c \} \), is given by

\[
d = \frac{\langle a, n \rangle - c}{\|n\|}.
\]
Lines in $\mathbb{R}^2$

**Definition 150 (Hessian normal form, Dt.: Hessische Normalform)**

A line equation $\langle u, n \rangle = c$ for $\ell(p, q)$ is said to be in *Hessian normal form* if $n$ (as specified in Lem. 149) is a unit vector.

**Lemma 151**

The (signed) minimum distance $d$ of a point $A \in \mathbb{R}^2$ from $\ell(p, q)$, with $\ell(p, q) = \{ u \in \mathbb{R}^2 : \langle u, n \rangle = c \}$, is given by

$$d = \frac{\langle a, n \rangle - c}{\|n\|}.$$  

- The signed distance of $A \in \mathbb{R}^2$ from $\ell(p, q) = \{ u \in \mathbb{R}^2 : \langle u, n \rangle = c \}$ is positive if $A$ is on that side of $\ell(p, q)$ into which $n$ points.
Planes in $\mathbb{R}^3$

**Definition 152 (Plane, Dt.: Ebene)**

For three distinct and non-collinear points $p, q, r \in \mathbb{R}^3$, the *plane* defined by $p, q, r$ is the set

$$\varepsilon(p, q, r) := \{p + \lambda \cdot pq + \mu \cdot pr : \lambda, \mu \in \mathbb{R}\}.$$
Planes in \( \mathbb{R}^3 \)

**Definition 152 (Plane, Dt.: Ebene)**

For three distinct and non-collinear points \( p, q, r \in \mathbb{R}^3 \), the *plane* defined by \( p, q, r \) is the set

\[
\varepsilon(p, q, r) := \{ p + \lambda \cdot pq + \mu \cdot pr : \lambda, \mu \in \mathbb{R} \}.
\]

- \( p + \lambda \cdot pq + \mu \cdot pr \) is the so-called parametric representation of \( \varepsilon(p, q, r) \).
Planes in $\mathbb{R}^3$

**Definition 152 (Plane, Dt.: Ebene)**

For three distinct and non-collinear points $p, q, r \in \mathbb{R}^3$, the *plane* defined by $p, q, r$ is the set

$$\varepsilon(p, q, r) := \{p + \lambda \cdot pq + \mu \cdot pr : \lambda, \mu \in \mathbb{R}\}.$$  

- $p + \lambda \cdot pq + \mu \cdot pr$ is the so-called parametric representation of $\varepsilon(p, q, r)$.

**Lemma 153**

For every triple of distinct and non-collinear points $p, q, r \in \mathbb{R}^3$, there exist $n \in \mathbb{R}^3$ and $c \in \mathbb{R}$ such that

$$\varepsilon(p, q, r) = \{u \in \mathbb{R}^3 : \langle u, n \rangle = c\}.$$
Planes in $\mathbb{R}^3$

**Definition 152 (Plane, Dt.: Ebene)**

For three distinct and non-collinear points \( p, q, r \in \mathbb{R}^3 \), the plane defined by \( p, q, r \) is the set

\[
\varepsilon(p, q, r) := \{ p + \lambda \cdot pq + \mu \cdot pr : \lambda, \mu \in \mathbb{R} \}.
\]

- \( p + \lambda \cdot pq + \mu \cdot pr \) is the so-called parametric representation of \( \varepsilon(p, q, r) \).

**Lemma 153**

For every triple of distinct and non-collinear points \( p, q, r \in \mathbb{R}^3 \), there exist \( n \in \mathbb{R}^3 \) and \( c \in \mathbb{R} \) such that

\[
\varepsilon(p, q, r) = \{ u \in \mathbb{R}^3 : \langle u, n \rangle = c \}.
\]

- \( \langle u, n \rangle = c \) is the so-called equational representation of \( \varepsilon(p, q, r) \).
Planes in \( \mathbb{R}^3 \)

**Definition 152 (Plane, Dt.: Ebene)**

For three distinct and non-collinear points \( p, q, r \in \mathbb{R}^3 \), the *plane* defined by \( p, q, r \) is the set

\[
\varepsilon(p, q, r) := \{ p + \lambda \cdot pq + \mu \cdot pr : \lambda, \mu \in \mathbb{R} \}.
\]

- \( p + \lambda \cdot pq + \mu \cdot pr \) is the so-called parametric representation of \( \varepsilon(p, q, r) \).

**Lemma 153**

For every triple of distinct and non-collinear points \( p, q, r \in \mathbb{R}^3 \), there exist \( n \in \mathbb{R}^3 \) and \( c \in \mathbb{R} \) such that

\[
\varepsilon(p, q, r) = \{ u \in \mathbb{R}^3 : \langle u, n \rangle = c \}.
\]

- \( \langle u, n \rangle = c \) is the so-called equational representation of \( \varepsilon(p, q, r) \).
- Note that \( \langle n, pq \rangle = \langle n, pr \rangle = 0 \) holds for every such \( n \). That is, the vector \( n \) is a normal vector of \( \varepsilon(p, q, r) \). We have

\[
n = \lambda(pq \times pr)
\] for some non-zero scalar \( \lambda \in \mathbb{R} \).
Planes in $\mathbb{R}^3$

**Definition 154 (Hessian normal form, Dt.: Hessische Normalform)**

A plane equation $\langle u, n \rangle = c$ for $\varepsilon(p, q, r)$ is said to be in *Hessian normal form* if $n$ (as specified in Lem. 153) is a unit vector.
Planes in $\mathbb{R}^3$

**Definition 154 (Hessian normal form, Dt.: Hessische Normalform)**

A plane equation $\langle u, n \rangle = c$ for $\varepsilon(p, q, r)$ is said to be in *Hessian normal form* if $n$ (as specified in Lem. 153) is a unit vector.

**Lemma 155**

The (signed) minimum distance $d$ of a point $A \in \mathbb{R}^3$ from $\varepsilon(p, q, r)$, with $\varepsilon(p, q, r) = \{u \in \mathbb{R}^3 : \langle u, n \rangle = c\}$, is given by

$$d = \frac{\langle a, n \rangle - c}{\|n\|}.$$
Planes in $\mathbb{R}^3$

**Definition 154 (Hessian normal form, Dt.: Hessische Normalform)**

A plane equation $\langle u, n \rangle = c$ for $\varepsilon(p, q, r)$ is said to be in *Hessian normal form* if $n$ (as specified in Lem. 153) is a unit vector.

**Lemma 155**

The (signed) minimum distance $d$ of a point $A \in \mathbb{R}^3$ from $\varepsilon(p, q, r)$, with $\varepsilon(p, q, r) = \{ u \in \mathbb{R}^3 : \langle u, n \rangle = c \}$, is given by

$$d = \frac{\langle a, n \rangle - c}{\|n\|}.$$ 

- The signed distance of $A \in \mathbb{R}^3$ from $\varepsilon(p, q, r) = \{ u \in \mathbb{R}^3 : \langle u, n \rangle = c \}$ is positive if $A$ is on that side of $\varepsilon(p, q, r)$ into which $n$ points.
Lemma 156
The equation of the line through two distinct points $p$ and $q$ in $\mathbb{R}^2$ is given by
\[
\det \begin{pmatrix}
x & y & 1 \\
p_x & p_y & 1 \\
q_x & q_y & 1
\end{pmatrix} = 0.
\]
Line/Plane Equation via Determinant

**Lemma 156**
The equation of the line through two distinct points \( p \) and \( q \) in \( \mathbb{R}^2 \) is given by

\[
\det \begin{pmatrix}
x & y & 1 \\
p_x & p_y & 1 \\
q_x & q_y & 1 \\
\end{pmatrix} = 0.
\]

**Lemma 157**
The equation of the plane through three distinct and non-collinear points \( p, q, r \) in \( \mathbb{R}^3 \) is given by

\[
\det \begin{pmatrix}
x & y & z & 1 \\
p_x & p_y & p_z & 1 \\
q_x & q_y & q_z & 1 \\
r_x & r_y & r_z & 1 \\
\end{pmatrix} = 0.
\]
Half-Plane and Half-Space

- The line $\ell(p, q) = \{ u \in \mathbb{R}^2 : \langle u, n \rangle = c \}$ partitions $\mathbb{R}^2$ into three disjoint sets: the actual line and the two (open) half-planes $\{ u \in \mathbb{R}^2 : \langle u, n \rangle - c < 0 \}$ and $\{ u \in \mathbb{R}^2 : \langle u, n \rangle - c > 0 \}$.

- Similarly for a plane in $\mathbb{R}^3$ and half-spaces.
Intersections of Lines and Planes

• The intersection of two lines $a_1 x + b_1 y = c_1$ and $a_2 x + b_2 y = c_2$ in $\mathbb{R}^2$ is given by the solution(s) of the following system of two linear equations:

\[
\begin{align*}
  a_1 x + b_1 y &= c_1 \\
  a_2 x + b_2 y &= c_2 
\end{align*}
\]
Intersections of Lines and Planes

The intersection of two lines $a_1 x + b_1 y = c_1$ and $a_2 x + b_2 y = c_2$ in $\mathbb{R}^2$ is given by the solution(s) of the following system of two linear equations:

\[
\begin{align*}
    a_1 x + b_1 y &= c_1 \\
    a_2 x + b_2 y &= c_2
\end{align*}
\]

That is,

\[
Au = c \quad \text{with} \quad A := \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}, \quad u := \begin{pmatrix} x \\ y \end{pmatrix}, \quad c := \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.
\]
Intersections of Lines and Planes

- The intersection of two lines $a_1 x + b_1 y = c_1$ and $a_2 x + b_2 y = c_2$ in $\mathbb{R}^2$ is given by the solution(s) of the following system of two linear equations:

  \[
  a_1 x + b_1 y = c_1 \\
  a_2 x + b_2 y = c_2
  \]

  That is,

  \[
  \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.
  \]

- Similarly for the intersection of $m$ (hyper-)planes in $\mathbb{R}^n$:

  \[
  a_{11} x_1 + \cdots + a_{1n} x_n = b_1 \\
  \vdots \\
  a_{m1} x_1 + \cdots + a_{mn} x_n = b_m
  \]
Geometric Objects
- Lines and Planes
- Circles and Spheres
  - Definitions
  - Equations and Parametrizations
  - Putnam Problem: Points on a Sphere
- Conics
- Curves and Surfaces
- Polygons and Polyhedra
Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

**Definition 158 (Sphere, Dt.: Sphäre, Kugeloberfläche)**

A (hyper)-sphere in $\mathbb{R}^n$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^n$ is the set

$$S(c, r) := \{ u \in \mathbb{R}^n : d(u, c) = r \}.$$

Conventionally, a hyper-sphere is called a *circle* in $\mathbb{R}^2$ and a *sphere* in $\mathbb{R}^3$. 

The equational form of a hyper-sphere (in the $L_2$-norm) can be re-written as

$$(u_1 - c_1)^2 + (u_2 - c_2)^2 + \cdots + (u_n - c_n)^2 = r^2.$$
Circles in \( \mathbb{R}^2 \) and Spheres in \( \mathbb{R}^3 \)

**Definition 158 (Sphere, Dt.: Sphäre, Kugeloberfläche)**

A (hyper)-sphere in \( \mathbb{R}^n \) with radius \( r \in \mathbb{R} \) centered at point \( c \in \mathbb{R}^n \) is the set

\[
S(c, r) := \{ u \in \mathbb{R}^n : d(u, c) = r \}.
\]

Conventionally, a hyper-sphere is called a *circle* in \( \mathbb{R}^2 \) and a *sphere* in \( \mathbb{R}^3 \).

- The equational form of a hyper-sphere (in the \( L_2 \)-norm) can be re-written as

\[
(u_1 - c_1)^2 + (u_2 - c_2)^2 + \cdots + (u_n - c_n)^2 = r^2.
\]
Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

**Definition 158 (Sphere, Dt.: Sphäre, Kugeloberfläche)**

A (hyper)-sphere in $\mathbb{R}^n$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^n$ is the set

$$S(c, r) := \{ u \in \mathbb{R}^n : d(u, c) = r \}.$$ 

Conventionally, a hyper-sphere is called a *circle* in $\mathbb{R}^2$ and a *sphere* in $\mathbb{R}^3$.

- The equational form of a hyper-sphere (in the $L_2$-norm) can be re-written as
  $$(u_1 - c_1)^2 + (u_2 - c_2)^2 + \cdots + (u_n - c_n)^2 = r^2.$$ 

**Definition 159 (Disk, Dt.: Kreisscheibe)**

A (closed) disk in $\mathbb{R}^2$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^2$ is the set

$$\{ u \in \mathbb{R}^2 : d(u, c) \leq r \}.$$ 

- An open disk in $\mathbb{R}^2$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^2$ is the set
  $$\{ u \in \mathbb{R}^2 : d(u, c) < r \}.$$
Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

**Definition 158 (Sphere, Dt.: Sphäre, Kugeloberfläche)**

A (hyper)-sphere in $\mathbb{R}^n$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^n$ is the set

$$S(c, r) := \{ u \in \mathbb{R}^n : d(u, c) = r \}.$$ 

Conventionally, a hyper-sphere is called a circle in $\mathbb{R}^2$ and a sphere in $\mathbb{R}^3$.

- The equational form of a hyper-sphere (in the $L_2$-norm) can be re-written as
  $$(u_1 - c_1)^2 + (u_2 - c_2)^2 + \cdots + (u_n - c_n)^2 = r^2.$$ 

**Definition 159 (Disk, Dt.: Kreisscheibe)**

A (closed) disk in $\mathbb{R}^2$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^2$ is the set

$$\{ u \in \mathbb{R}^2 : d(u, c) \leq r \}.$$ 

**Definition 160 (Open disk)**

An open disk in $\mathbb{R}^2$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^2$ is the set

$$\{ u \in \mathbb{R}^2 : d(u, c) < r \}.$$
Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

Definition 161 (Ball, Dt.: Kugel)

A \textit{(closed) ball} in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$B(c, r) := \{ u \in \mathbb{R}^3 : d(u, c) \leq r \}.$$
Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

Definition 161 (Ball, Dt.: Kugel)

A (closed) ball in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$B(c, r) := \{ u \in \mathbb{R}^3 : d(u, c) \leq r \}.$$  

Definition 162 (Open ball)

An open ball in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$\{ u \in \mathbb{R}^3 : d(u, c) < r \}.$$
Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

**Definition 161 (Ball, Dt.: Kugel)**

A (closed) ball in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$B(c, r) := \{u \in \mathbb{R}^3 : d(u, c) \leq r\}.$$ 

**Definition 162 (Open ball)**

An open ball in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$\{u \in \mathbb{R}^3 : d(u, c) < r\}.$$ 

- Of course, these definitions can be generalized to distances other than the standard Euclidean distance (based on the $L_2$-norm).
Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

**Definition 161 (Ball, Dt.: Kugel)**

A *(closed)* ball in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$B(c, r) := \{ u \in \mathbb{R}^3 : d(u, c) \leq r \}.$$

**Definition 162 (Open ball)**

An *(open)* ball in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$\{ u \in \mathbb{R}^3 : d(u, c) < r \}.$$

- Of course, these definitions can be generalized to distances other than the standard Euclidean distance (based on the $L_2$-norm).
- In mathematics, a terminological distinction is made between a sphere, which is a two-dimensional closed surface embedded in $\mathbb{R}^3$, and a ball, which is a shape (“solid”) in $\mathbb{R}^3$ that includes the interior of its associated sphere.
Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

**Definition 161 (Ball, Dt.: Kugel)**

A *(closed)* ball in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$B(c, r) := \{ u \in \mathbb{R}^3 : d(u, c) \leq r \}.$$

**Definition 162 (Open ball)**

An *open ball* in $\mathbb{R}^3$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^3$ is the set

$$\{ u \in \mathbb{R}^3 : d(u, c) < r \}.$$

- Of course, these definitions can be generalized to distances other than the standard Euclidean distance (based on the $L_2$-norm).
- In mathematics, a terminological distinction is made between a sphere, which is a two-dimensional closed surface embedded in $\mathbb{R}^3$, and a ball, which is a shape ("solid") in $\mathbb{R}^3$ that includes the interior of its associated sphere.
- In mathematics, for $n \in \mathbb{N}$, an $n$-sphere of radius $r$ is the set of points in $(n + 1)$-dimensional Euclidean space which are at distance $r$ from the origin, with $r := 1$ for the unit $n$-sphere $S^n$. 
Circle Equation via Determinant

Lemma 163

For points $p_1 := (x_1, y_1)$, $p_2 := (x_2, y_2)$ and $p_3 := (x_3, y_3)$ in $\mathbb{R}^2$, the equation of the circle (under the Euclidean distance) through $p_1$, $p_2$ and $p_3$ is given by

$$\det \begin{pmatrix}
  x^2 + y^2 & x & y & 1 \\
  x_1^2 + y_1^2 & x_1 & y_1 & 1 \\
  x_2^2 + y_2^2 & x_2 & y_2 & 1 \\
  x_3^2 + y_3^2 & x_3 & y_3 & 1
\end{pmatrix} = 0.$$
Circle Equation via Determinant

Lemma 163

For points $p_1 := (x_1, y_1)$, $p_2 := (x_2, y_2)$ and $p_3 := (x_3, y_3)$ in $\mathbb{R}^2$, the equation of the circle (under the Euclidean distance) through $p_1$, $p_2$ and $p_3$ is given by

$$\det \begin{pmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{pmatrix} = 0.$$ 

This can be used to check whether a fourth point $p_4 := (x_4, y_4)$ lies inside the circle defined by three points $p_1$, $p_2$, $p_3$ arranged in CCW order: The point $p_4$ lies inside that circle if and only if the determinant is greater than zero (when $x$ and $y$ are replaced by $x_4$ and $y_4$).
Sphere Equation via Determinant

**Lemma 164**

For points \( p_1 := (x_1, y_1, z_1) \), \( p_2 := (x_2, y_2, z_2) \), \( p_3 := (x_3, y_3, z_3) \) and \( p_4 := (x_4, y_4, z_4) \) in \( \mathbb{R}^3 \), the equation of the sphere (under the Euclidean distance) through \( p_1 \), \( p_2 \), \( p_3 \) and \( p_4 \) is given by

\[
\begin{vmatrix}
  x^2 + y^2 + z^2 & x & y & z & 1 \\
  x_1^2 + y_1^2 + z_1^2 & x_1 & y_1 & z_1 & 1 \\
  x_2^2 + y_2^2 + z_2^2 & x_2 & y_2 & z_2 & 1 \\
  x_3^2 + y_3^2 + z_3^2 & x_3 & y_3 & z_3 & 1 \\
  x_4^2 + y_4^2 + z_4^2 & x_4 & y_4 & z_4 & 1 \\
\end{vmatrix} = 0.
\]

- This formula generalizes to any number of dimensions.
Sphere via Ratios of Distances

Lemma 165 (Appolonius of Perga)

Consider two distinct points \( p, q \in \mathbb{R}^n \) and a constant \( k \in \mathbb{R}^+ \). Then

\[
\{ u \in \mathbb{R}^n : \frac{d(u, p)}{d(u, q)} = k \}
\]

forms a (hyper-)sphere.
Parametrization of a Circle

- We know that a circle with radius \( r \in \mathbb{R}_0^+ \) centered at the point \( c \in \mathbb{R}^2 \) has the equation \((c_x - x)^2 + (c_y - y)^2 = r^2\).
Parametrization of a Circle

- We know that a circle with radius $r \in \mathbb{R}^+_0$ centered at the point $c \in \mathbb{R}^2$ has the equation $(c_x - x)^2 + (c_y - y)^2 = r^2$.

Lemma 166

Consider a circle in $\mathbb{R}^2$ with radius $r \in \mathbb{R}^+_0$ centered at point $c \in \mathbb{R}^2$. Its parametrization is given by

$$
\begin{pmatrix}
c_x + r \cos \varphi \\
c_y + r \sin \varphi
\end{pmatrix}
$$

with $\varphi \in [0, 2\pi[$.
Parametrization of a Sphere

- We know that a sphere with radius $r \in \mathbb{R}^+_{0}$ centered at the point $c \in \mathbb{R}^3$ has the equation $(c_x - x)^2 + (c_y - y)^2 + (c_z - z)^2 = r^2$. 

Lemma 167

Consider a sphere in $\mathbb{R}^3$ with radius $r \in \mathbb{R}^+_{0}$ centered at point $c \in \mathbb{R}^3$. Its parametrization is given by

$$
\begin{bmatrix}
  c_x x + r \cos \delta \cos \varphi \\
  c_y y + r \cos \delta \sin \varphi \\
  c_z z + r \sin \delta
\end{bmatrix}
$$

with $\varphi \in [0, 2\pi]$ and $\delta \in [-\pi/2, \pi/2]$. 

We know that a sphere with radius $r \in \mathbb{R}_0^+$ centered at the point $c \in \mathbb{R}^3$ has the equation $(c_x - x)^2 + (c_y - y)^2 + (c_z - z)^2 = r^2$.

**Lemma 167**

Consider a sphere in $\mathbb{R}^3$ with radius $r \in \mathbb{R}_0^+$ centered at point $c \in \mathbb{R}^3$. Its parametrization is given by

$$
\begin{pmatrix}
    c_x + r \cos \delta \cos \varphi \\
    c_y + r \cos \delta \sin \varphi \\
    c_z + r \sin \delta
\end{pmatrix}
$$

with $\varphi \in [0, 2\pi]$ and $\delta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. 
Putnam Problem: Points on a Sphere

- Choose four points $p_1, p_2, p_3, p_4$ independently at random (relative to a uniform distribution) on a sphere (in $\mathbb{R}^3$).
- Consider the tetrahedron $T$ formed by $p_1, p_2, p_3, p_4$.
- What is the probability that the center of the sphere lies inside $T$?
Putnam Problem: Points on a Sphere

- Choose four points \( p_1, p_2, p_3, p_4 \) independently at random (relative to a uniform distribution) on a sphere (in \( \mathbb{R}^3 \)).
- Consider the tetrahedron \( T \) formed by \( p_1, p_2, p_3, p_4 \).
- What is the probability that the center of the sphere lies inside \( T \)?
- We start with considering the problem in 2D: three random points on a circle.
Putnam Problem: Points on a Sphere

- W.l.o.g., the point $p_1$ is at the north pole of the circle, centered at the origin.
Putnam Problem: Points on a Sphere

- W.l.o.g., the point $p_1$ is at the north pole of the circle, centered at the origin.
- We can select $p_2$ by picking a random angle within $[0, 360]$, or by picking a random angle within $[0, 180]$ — thus fixing a line $\ell_2$ through the origin — and then flipping a coin to choose between $p_2'$ and $p_2''$. 
Putnam Problem: Points on a Sphere

- W.l.o.g., the point $p_1$ is at the north pole of the circle, centered at the origin.
- We can select $p_2$ by picking a random angle within $[0, 360]$, or by picking a random angle within $[0, 180]$ — thus fixing a line $\ell_2$ through the origin — and then flipping a coin to choose between $p'_2$ and $p''_2$.
- Same for $\ell_3$ and $p'_3$ and $p''_3$ as candidates for $p_3$. 
Putnam Problem: Points on a Sphere

- W.l.o.g., the point $p_1$ is at the north pole of the circle, centered at the origin.
- We can select $p_2$ by picking a random angle within $[0, 360]$, or by picking a random angle within $[0, 180]$ — thus fixing a line $\ell_2$ through the origin — and then flipping a coin to choose between $p'_2$ and $p''_2$.
- Same for $\ell_3$ and $p'_3$ and $p''_3$ as candidates for $p_3$.
- With probability one, we have $\ell_2 \neq \ell_3$ and $p_1 \not\in \ell_2$ and $p_1 \not\in \ell_3$. 
Putnam Problem: Points on a Sphere

- W.l.o.g., the point $p_1$ is at the north pole of the circle, centered at the origin.
- We can select $p_2$ by picking a random angle within $[0, 360]$, or by picking a random angle within $[0, 180]$ — thus fixing a line $\ell_2$ through the origin — and then flipping a coin to choose between $p_2'$ and $p_2''$.
- Same for $\ell_3$ and $p_3'$ and $p_3''$ as candidates for $p_3$.
- With probability one, we have $\ell_2 \neq \ell_3$ and $p_1 \not\in \ell_2$ and $p_1 \not\in \ell_3$.
- The four possible triangles
  \[ \Delta(p_1, p_2', p_3') \]
  \[ \Delta(p_1, p_2', p_3'') \]
  \[ \Delta(p_1, p_2'', p_3') \]
  \[ \Delta(p_1, p_2'', p_3'') \]
  are equally likely.
Putnam Problem: Points on a Sphere

- W.l.o.g., the point \( p_1 \) is at the north pole of the circle, centered at the origin.
- We can select \( p_2 \) by picking a random angle within \([0, 360]\), or by picking a random angle within \([0, 180]\) — thus fixing a line \( \ell_2 \) through the origin — and then flipping a coin to choose between \( p'_2 \) and \( p''_2 \).
- Same for \( \ell_3 \) and \( p'_3 \) and \( p''_3 \) as candidates for \( p_3 \).
- With probability one, we have \( \ell_2 \neq \ell_3 \) and \( p_1 \notin \ell_2 \) and \( p_1 \notin \ell_3 \).
- The four possible triangles
  \[ \Delta(p_1, p'_2, p'_3) \]
  \[ \Delta(p_1, p'_2, p''_3) \]
  \[ \Delta(p_1, p''_2, p'_3) \]
  \[ \Delta(p_1, p''_2, p''_3) \]
  are equally likely.
- We know that at most two vectors can be linearly independent in \( \mathbb{R}^2 \).
Putnam Problem: Points on a Sphere

- W.l.o.g., the point $p_1$ is at the north pole of the circle, centered at the origin.
- We can select $p_2$ by picking a random angle within $[0, 360]$, or by picking a random angle within $[0, 180]$ — thus fixing a line $\ell_2$ through the origin — and then flipping a coin to choose between $p'_2$ and $p''_2$.
- Same for $\ell_3$ and $p'_3$ and $p''_3$ as candidates for $p_3$.
- With probability one, we have $\ell_2 \neq \ell_3$ and $p_1 \notin \ell_2$ and $p_1 \notin \ell_3$.
- The four possible triangles
  \[ \Delta(p_1, p'_2, p'_3) \]
  \[ \Delta(p_1, p'_2, p''_3) \]
  \[ \Delta(p_1, p''_2, p'_3) \]
  \[ \Delta(p_1, p''_2, p''_3) \]
  are equally likely.
- We know that at most two vectors can be linearly independent in $\mathbb{R}^2$.
- Hence, there exist $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ such that
  \[ 0 = \lambda_1 \cdot p_1 + \lambda_2 \cdot p_2 + \lambda_3 \cdot p_3, \]
  and not all of $\lambda_1, \lambda_2, \lambda_3$ are zero.
Putnam Problem: Points on a Sphere

- Actually, we have $\lambda_1, \lambda_2, \lambda_3$ all non-zero.
Putnam Problem: Points on a Sphere

- Actually, we have $\lambda_1, \lambda_2, \lambda_3$ all non-zero.
- If
  
  $$0 = \lambda_1 \cdot p_1 + \lambda_2 \cdot p'_2 + \lambda_3 \cdot p_3$$

  then

  $$0 = \lambda_1 \cdot p_1 - \lambda_2 \cdot p''_2 + \lambda_3 \cdot p_3.$$

Hence, we get the origin as a linear combination with positive coefficients of the three corners of a triangle for exactly one of the four triangles.

If $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}^+$ then we may assume $\lambda_1 + \lambda_2 + \lambda_3 = 1$, thus obtaining a convex combination.

Hence, a random triangle contains the center of the circle with probability $\frac{1}{4}$.

Similarly, a random tetrahedron contains the center of the sphere with probability $\frac{1}{8}$. 

© M. Held (Univ. Salzburg)
Putnam Problem: Points on a Sphere

- Actually, we have $\lambda_1, \lambda_2, \lambda_3$ all non-zero.
- If
  
  \[ 0 = \lambda_1 \cdot p_1 + \lambda_2 \cdot p'_2 + \lambda_3 \cdot p_3 \]

  then
  
  \[ 0 = \lambda_1 \cdot p_1 - \lambda_2 \cdot p''_2 + \lambda_3 \cdot p_3. \]

- Hence, we get the origin as a linear combination with positive coefficients of the three corners of a triangle for exactly one of the four triangles.
Putnam Problem: Points on a Sphere

- Actually, we have $\lambda_1, \lambda_2, \lambda_3$ all non-zero.

- If

  $$0 = \lambda_1 \cdot p_1 + \lambda_2 \cdot p'_2 + \lambda_3 \cdot p_3$$

  then

  $$0 = \lambda_1 \cdot p_1 - \lambda_2 \cdot p''_2 + \lambda_3 \cdot p_3.$$  

- Hence, we get the origin as a linear combination with positive coefficients of the three corners of a triangle for exactly one of the four triangles.

- If $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}^+$ then we may assume $\lambda_1 + \lambda_2 + \lambda_3 = 1$, thus obtaining a convex combination.
Putnam Problem: Points on a Sphere

- Actually, we have $\lambda_1, \lambda_2, \lambda_3$ all non-zero.
- If

$$0 = \lambda_1 \cdot p_1 + \lambda_2 \cdot p'_2 + \lambda_3 \cdot p_3$$

then

$$0 = \lambda_1 \cdot p_1 - \lambda_2 \cdot p''_2 + \lambda_3 \cdot p_3.$$

- Hence, we get the origin as a linear combination with positive coefficients of the three corners of a triangle for exactly one of the four triangles.
- If $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}^+$ then we may assume $\lambda_1 + \lambda_2 + \lambda_3 = 1$, thus obtaining a convex combination.
- Hence, a random triangle contains the center of the circle with probability $\frac{1}{4}$.
Putnam Problem: Points on a Sphere

- Actually, we have $\lambda_1, \lambda_2, \lambda_3$ all non-zero.
- If
  
  \[ 0 = \lambda_1 \cdot p_1 + \lambda_2 \cdot p'_2 + \lambda_3 \cdot p_3 \]

  then

  \[ 0 = \lambda_1 \cdot p_1 - \lambda_2 \cdot p''_2 + \lambda_3 \cdot p_3. \]

- Hence, we get the origin as a linear combination with positive coefficients of the three corners of a triangle for exactly one of the four triangles.
- If $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}^+$ then we may assume
  
  $\lambda_1 + \lambda_2 + \lambda_3 = 1$, thus obtaining a convex combination.

- Hence, a random triangle contains the center of the circle with probability $\frac{1}{4}$.
- Similarly, a random tetrahedron contains the center of the sphere with probability $\frac{1}{8}$. 
4 Geometric Objects

- Lines and Planes
- Circles and Spheres
- Conics
  - Cone and Conics
  - Ellipse
  - Ellipsoid
- Curves and Surfaces
- Polygons and Polyhedra
Definition 168 (Cone, Dt.: Kegel)

A (right circular) cone is formed by a set of line segments (or lines) which connect a common point, called the apex, to all the points of a circular base, where the apex lies on a perpendicular through the center of the circle. This line is called axis of the cone.
Definition 168 (Cone, Dt.: Kegel)

A (right circular) cone is formed by a set of line segments (or lines) which connect a common point, called the apex, to all the points of a circular base, where the apex lies on a perpendicular through the center of the circle. This line is called axis of the cone.

- The axis is the axis of symmetry of the cone.
Definition 168 (Cone, Dt.: Kegel)

A (right circular) cone is formed by a set of line segments (or lines) which connect a common point, called the apex, to all the points of a circular base, where the apex lies on a perpendicular through the center of the circle. This line is called axis of the cone.

- The axis is the axis of symmetry of the cone.
- A cone is characterized by its height $h$ and base radius $r$. 

---

Pythagorean theorem implies $\sqrt{h^2 + r^2}$ for the slant height $s$.

The intercept theorem implies that all cross sections of a cone parallel to the base will be similar to the base, i.e., they will also be circles.
A (right circular) cone is formed by a set of line segments (or lines) which connect a common point, called the apex, to all the points of a circular base, where the apex lies on a perpendicular through the center of the circle. This line is called axis of the cone.

- The axis is the axis of symmetry of the cone.
- A cone is characterized by its height $h$ and base radius $r$.
- The Pythagorean theorem implies $\sqrt{h^2 + r^2}$ for the slant height $s$. 
Cone

**Definition 168 (Cone, Dt.: Kegel)**

A (right circular) *cone* is formed by a set of line segments (or lines) which connect a common point, called the *apex*, to all the points of a circular base, where the apex lies on a perpendicular through the center of the circle. This line is called *axis* of the cone.

- The axis is the axis of symmetry of the cone.
- A cone is characterized by its *height* $h$ and *base radius* $r$.
- The Pythagorean theorem implies $\sqrt{h^2 + r^2}$ for the slant height $s$.
- The intercept theorem implies that all cross sections of a cone parallel to the base will be similar to the base, i.e., they will also be circles.
Railroad Track on Cone Mountain

Consider a mountain that is shaped like a right circular cone.

A shortest-length railroad track is supposed to start at $A$, wind around the mountain once, and end in $B$.

The height $h$ of the cone is $40\sqrt{2}$, its base radius $r$ is 20, and the distance between $A$ and $B$ is 10.

Your task:

1. Prove that the shortest-length railroad track from $A$ to $B$ that winds around the mountain once consists of an uphill portion and of a downhill portion.

2. Compute the length of the downhill portion.
The key insight is that the lateral surface (Dt.: Mantel) of the cone forms a circular disk sector with radius $s = \sqrt{r^2 + h^2} = 60$. 

Since the base circle has a circumference of $2r\pi = 40\pi$, while a circle with radius $60$ has circumference $120\pi$, the opening angle of the disk sector is $120^\circ$. 

The shortest distance from $A$ to $B$ is a straight-line segment.

Standard high school math yields $510/\sqrt{91}$ as length of the uphill part and $510/\sqrt{91}$ as length of the downhill part of the track.
The key insight is that the lateral surface (Dt.: Mantel) of the cone forms a circular disk sector with radius $s = \sqrt{r^2 + h^2} = 60$.

Since the base circle has a circumference of $2r\pi = 40\pi$, while a circle with radius 60 has circumference $120\pi$, the opening angle of the disk sector is $120^\circ$. 

The shortest distance from $A$ to $B$ is a straight-line segment. Standard high school math yields $\frac{510}{\sqrt{91}}$ as length of the uphill part and $\frac{510}{\sqrt{91}}$ as length of the downhill part of the track.
Railroad Track on Cone Mountain

- The key insight is that the lateral surface (Dt.: Mantel) of the cone forms a circular disk sector with radius $s = \sqrt{r^2 + h^2} = 60$.
- Since the base circle has a circumference of $2r\pi = 40\pi$, while a circle with radius 60 has circumference $120\pi$, the opening angle of the disk sector is $120^\circ$.
- The shortest distance from $A$ to $B$ is a straight-line segment.
The key insight is that the lateral surface (Dt.: Mantel) of the cone forms a circular disk sector with radius $s = \sqrt{r^2 + h^2} = 60$.

Since the base circle has a circumference of $2r\pi = 40\pi$, while a circle with radius 60 has circumference $120\pi$, the opening angle of the disk sector is $120^\circ$.

The shortest distance from $A$ to $B$ is a straight-line segment.

Standard high school math yields $\frac{510}{\sqrt{91}}$ as length of the uphill part and $\frac{510}{\sqrt{91}}$ as length of the downhill part of the track.
Conics

- Conic sections (Dt.: Kegelschnitte) are formed by the intersection of a (double circular right) cone and a plane.

[Image credit: Wikipedia.]
**Definition 169 (Ellipse)**

Consider two points $F_1, F_2$ and a distance $a \in \mathbb{R}^+$ such that $2a \geq d(f_1, f_2)$. Then the *ellipse* defined by $F_1, F_2$ and $a$ is given as follows:

$$\{u \in \mathbb{R}^2 : d(u, f_1) + d(u, f_2) = 2a\}$$
Lemma 170

The standard (axis-aligned) ellipse with width $2a$ and height $2b$ has the equation

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

If $a \geq b$ then $c = \sqrt{a^2 - b^2}$. 

![Diagram of an ellipse with foci $F_1$ and $F_2$, major and minor axes, and focal distance $c$.]
Lemma 171

The standard (axis-aligned) ellipse with width $2a$ and height $2b$ can be parametrized as

$$\begin{pmatrix} a \cdot \cos \varphi \\ b \cdot \sin \varphi \end{pmatrix} \quad \text{with } \varphi \in [0, 2\pi[.$$
Ellipsoid

- An *ellipsoid* is a quadric surface in $\mathbb{R}^3$ that has three pairwise perpendicular axes of symmetry which intersect at the so-called center of the ellipsoid. The line segments that are delimited on the axes of symmetry by the ellipsoid are called the principal axes and are commonly denoted by $a$, $b$ and $c$. 

\[
\begin{pmatrix}
    a \cdot \sin \delta \cos \phi \\
    b \cdot \sin \delta \sin \phi \\
    c \cos \delta
\end{pmatrix}
\] with 

$\phi \in [0, 2\pi]$ and 

$\delta \in [-\pi/2, \pi/2]$.
Ellipsoid

- An *ellipsoid* is a quadric surface in $\mathbb{R}^3$ that has three pairwise perpendicular axes of symmetry which intersect at the so-called center of the ellipsoid. The line segments that are delimited on the axes of symmetry by the ellipsoid are called the principal axes and are commonly denoted by $a$, $b$ and $c$.

- The standard (axis-aligned) ellipsoid centered at the origin has the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$
An *ellipsoid* is a quadric surface in $\mathbb{R}^3$ that has three pairwise perpendicular axes of symmetry which intersect at the so-called center of the ellipsoid. The line segments that are delimited on the axes of symmetry by the ellipsoid are called the principal axes and are commonly denoted by $a$, $b$ and $c$.

The standard (axis-aligned) ellipsoid centered at the origin has the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$  

We get a sphere for $a = b = c$. 

---

© M. Held (Univ. Salzburg)  Geometrisches Rechnen  (WS 2020/21)
Ellipsoid

- An *ellipsoid* is a quadric surface in $\mathbb{R}^3$ that has three pairwise perpendicular axes of symmetry which intersect at the so-called center of the ellipsoid. The line segments that are delimited on the axes of symmetry by the ellipsoid are called the principal axes and are commonly denoted by $a$, $b$ and $c$.

- The standard (axis-aligned) ellipsoid centered at the origin has the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$  

We get a sphere for $a = b = c$.

- A parametrization is given by

$$\begin{pmatrix} a \cdot \sin \delta \cos \varphi \\ b \cdot \sin \delta \sin \varphi \\ c \cos \delta \end{pmatrix} \quad \text{with} \quad \varphi \in [0, 2\pi] \quad \text{and} \quad \delta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$
Geometric Objects
- Lines and Planes
- Circles and Spheres
- Conics
- Curves and Surfaces
- Polygons and Polyhedra
Curves

- Intuitively, a curve in $\mathbb{R}^2$ is generated by a continuous motion of a pencil on a sheet of paper.
- A formal mathematical definition is not entirely straightforward, and the term “curve” is associated with two closely related notions: kinematic and geometric.
Curves

- Intuitively, a curve in $\mathbb{R}^2$ is generated by a continuous motion of a pencil on a sheet of paper.
- A formal mathematical definition is not entirely straightforward, and the term “curve” is associated with two closely related notions: kinematic and geometric.
- In the kinematic setting, a (parameterized) curve is a function of one real variable.
Curves

- Intuitively, a curve in $\mathbb{R}^2$ is generated by a continuous motion of a pencil on a sheet of paper.
- A formal mathematical definition is not entirely straightforward, and the term “curve” is associated with two closely related notions: kinematic and geometric.
- In the kinematic setting, a (parameterized) curve is a function of one real variable.
- In the geometric setting, a curve, also called an arc, is a 1-dimensional subset of space.
- Both notions are related: the image of a parameterized curve describes an arc. Conversely, an arc admits a parametrization.
- Since the kinematic setting is easier to introduce, we resort to a kinematic definition of “curve”.

Note that fairly counter-intuitive curves exist: e.g., space-filling curves like the Sierpinski curve.
Curves

- Intuitively, a curve in $\mathbb{R}^2$ is generated by a continuous motion of a pencil on a sheet of paper.
- A formal mathematical definition is not entirely straightforward, and the term “curve” is associated with two closely related notions: kinematic and geometric.
- In the kinematic setting, a (parameterized) curve is a function of one real variable.
- In the geometric setting, a curve, also called an arc, is a 1-dimensional subset of space.
- Both notions are related: the image of a parameterized curve describes an arc. Conversely, an arc admits a parametrization.
- Since the kinematic setting is easier to introduce, we resort to a kinematic definition of “curve”.
- Note that fairly counter-intuitive curves exist: e.g., space-filling curves like the Sierpinski curve.
Sierpinski Curves

- Sierpinski curves are a sequence of recursively defined continuous and closed curves in $\mathbb{R}^2$.
- Sierpinski curve of order 1:
Sierpinski Curves

- Sierpinski curves are a sequence of recursively defined continuous and closed curves in $\mathbb{R}^2$.
- Sierpinski curve of order 2:
Sierpinski Curves

- Sierpinski curves are a sequence of recursively defined continuous and closed curves in $\mathbb{R}^2$.
- Sierpinski curve of order 3:
Sierpinski Curves

- Sierpinski curves are a sequence of recursively defined continuous and closed curves in $\mathbb{R}^2$.
- Sierpinski curve of order 4:
Sierpinski Curves

- Sierpinski curves are a sequence of recursively defined continuous and closed curves in $\mathbb{R}^2$.
- Sierpinski curve of order 4:

Their limit curve, the *Sierpinski curve*, is a space-filling curve: It fills the unit square completely! It is a continuous and surjective (but not injective!) mapping of $[0, 1]$ onto $[0, 1] \times [0, 1]$. 
Curves in $\mathbb{R}^n$

**Definition 172 (Curve, Dt.: Kurve)**

Let $I \subseteq \mathbb{R}$ be an interval of the real line. A continuous (vector-valued) mapping $\gamma: I \rightarrow \mathbb{R}^n$ is called a *parametrization* of $\gamma(I)$ or a *parametric curve*.

Well-known examples of parameterized curves include a straight-line segment, a circular arc, and a helix. E.g., $\gamma: [0, 1] \rightarrow \mathbb{R}^3$ with $\gamma(t) := \begin{bmatrix} px + t \cdot (qx - px) \\ py + t \cdot (qy - py) \\ pz + t \cdot (qz - pz) \end{bmatrix}$ maps $[0, 1]$ to a straight-line segment from point $p$ to $q$.

The interval $I$ is called the *domain* of $\gamma$, and $\gamma(I)$ is called the *image* (Dt.: Bild, Spur).

**Definition 173 (Plane curve, Dt.: ebene Kurve)**

For $\gamma: I \rightarrow \mathbb{R}^n$, the curve $\gamma(I)$ is *plane* if $\gamma(I) \subseteq \mathbb{R}^2$ or if $\gamma(I)$ lies within an affine/projective plane. A non-plane curve is called a *skew curve* (Dt.: Raumkurve).

An *algebraic plane curve* is the zero set of a polynomial in two variables.
Curves in $\mathbb{R}^n$

**Definition 172 (Curve, Dt.: Kurve)**

Let $I \subseteq \mathbb{R}$ be an interval of the real line. A continuous (vector-valued) mapping $\gamma: I \rightarrow \mathbb{R}^n$ is called a *parametrization* of $\gamma(I)$ or a *parametric curve*.

- Well-known examples of parameterized curves include a straight-line segment, a circular arc, and a helix.

- E.g., $\gamma: [0, 1] \rightarrow \mathbb{R}^3$ with
  \[
  \gamma(t) := \begin{pmatrix}
  p_x + t \cdot (q_x - p_x) \\
  p_y + t \cdot (q_y - p_y) \\
  p_z + t \cdot (q_z - p_z)
  \end{pmatrix}
  \]

  maps $[0, 1]$ to a straight-line segment from point $p$ to $q$. 
Curves in $\mathbb{R}^n$

**Definition 172 (Curve, Dt.: Kurve)**

Let $I \subseteq \mathbb{R}$ be an interval of the real line. A continuous (vector-valued) mapping $\gamma: I \rightarrow \mathbb{R}^n$ is called a *parametrization* of $\gamma(I)$ or a *parametric curve*.

- Well-known examples of parameterized curves include a straight-line segment, a circular arc, and a helix.
- E.g., $\gamma: [0, 1] \rightarrow \mathbb{R}^3$ with

  \[
  \gamma(t) := \begin{pmatrix}
  p_x + t \cdot (q_x - p_x) \\
  p_y + t \cdot (q_y - p_y) \\
  p_z + t \cdot (q_z - p_z)
  \end{pmatrix}
  \]

  maps $[0, 1]$ to a straight-line segment from point $p$ to $q$.

- The interval $I$ is called the *domain* of $\gamma$, and $\gamma(I)$ is called *image* (Dt.: Bild, Spur).
Curves in $\mathbb{R}^n$

**Definition 172 (Curve, Dt.: Kurve)**

Let $I \subseteq \mathbb{R}$ be an interval of the real line. A continuous (vector-valued) mapping $\gamma: I \rightarrow \mathbb{R}^n$ is called a *parametrization* of $\gamma(I)$ or a *parametric curve*.

- Well-known examples of parameterized curves include a straight-line segment, a circular arc, and a helix.
- E.g., $\gamma: [0, 1] \rightarrow \mathbb{R}^3$ with
  \[
  \gamma(t) := \begin{pmatrix}
  p_x + t \cdot (q_x - p_x) \\
  p_y + t \cdot (q_y - p_y) \\
  p_z + t \cdot (q_z - p_z)
  \end{pmatrix}
  \]
  maps $[0, 1]$ to a straight-line segment from point $p$ to $q$.
- The interval $I$ is called the *domain* of $\gamma$, and $\gamma(I)$ is called *image* (Dt.: Bild, Spur).

**Definition 173 (Plane curve, Dt.: ebene Kurve)**

For $\gamma: I \rightarrow \mathbb{R}^n$, the curve $\gamma(I)$ is *plane* if $\gamma(I) \subseteq \mathbb{R}^2$ or if $\gamma(I)$ lies within an affine/projective plane. A non-plane curve is called a *skew curve* (Dt.: Raumkurve).
Curves in $\mathbb{R}^n$

Definition 172 (Curve, Dt.: Kurve)

Let $I \subseteq \mathbb{R}$ be an interval of the real line. A continuous (vector-valued) mapping $\gamma: I \rightarrow \mathbb{R}^n$ is called a parametrization of $\gamma(I)$ or a parametric curve.

- Well-known examples of parameterized curves include a straight-line segment, a circular arc, and a helix.
- E.g., $\gamma: [0, 1] \rightarrow \mathbb{R}^3$ with

$$
\gamma(t) := \begin{pmatrix}
  p_x + t \cdot (q_x - p_x) \\
  p_y + t \cdot (q_y - p_y) \\
  p_z + t \cdot (q_z - p_z)
\end{pmatrix}
$$

maps $[0, 1]$ to a straight-line segment from point $p$ to $q$.
- The interval $I$ is called the domain of $\gamma$, and $\gamma(I)$ is called image (Dt.: Bild, Spur).

Definition 173 (Plane curve, Dt.: ebene Kurve)

For $\gamma: I \rightarrow \mathbb{R}^n$, the curve $\gamma(I)$ is plane if $\gamma(I) \subseteq \mathbb{R}^2$ or if $\gamma(I)$ lies within an affine/projective plane. A non-plane curve is called a skew curve (Dt.: Raumkurve).

- An algebraic plane curve is the zero set of a polynomial in two variables.
Curves in $\mathbb{R}^n$

**Definition 174 (Start and end point)**

If $I$ is a closed interval $[a, b]$, for some $a, b \in \mathbb{R}$, then we call $\gamma(a)$ the *start point* and $\gamma(b)$ the *end point* of the curve $\gamma : I \to \mathbb{R}^n$. 
Curves in $\mathbb{R}^n$

**Definition 174 (Start and end point)**

If $I$ is a closed interval $[a, b]$, for some $a, b \in \mathbb{R}$, then we call $\gamma(a)$ the *start point* and $\gamma(b)$ the *end point* of the curve $\gamma: I \to \mathbb{R}^n$.

**Definition 175 (Closed, Dt.: geschlossen)**

A parametrization $\gamma: I \to \mathbb{R}^n$ is said to be *closed* (or a *loop*) if $I$ is a closed interval $[a, b]$, for some $a, b \in \mathbb{R}$, and $\gamma(a) = \gamma(b)$.

Hence, if $\gamma: I \to \mathbb{R}^n$ is simple then it is injective on $\text{int}(I)$. 

© M. Held (Univ. Salzburg)
Curves in \( \mathbb{R}^n \)

**Definition 174 (Start and end point)**

If \( I \) is a closed interval \([a, b]\), for some \( a, b \in \mathbb{R} \), then we call \( \gamma(a) \) the *start point* and \( \gamma(b) \) the *end point* of the curve \( \gamma: I \to \mathbb{R}^n \).

**Definition 175 (Closed, Dt.: geschlossen)**

A parametrization \( \gamma: I \to \mathbb{R}^n \) is said to be *closed* (or a *loop*) if \( I \) is a closed interval \([a, b]\), for some \( a, b \in \mathbb{R} \), and \( \gamma(a) = \gamma(b) \).

**Definition 176 (Simple, Dt.: einfach)**

A parametrization \( \gamma: I \to \mathbb{R}^n \) is said to be *simple* if \( \gamma(t_1) = \gamma(t_2) \) for \( t_1 \neq t_2 \in I \) implies \( \{t_1, t_2\} = \{a, b\} \) and \( I = [a, b] \), for some \( a, b \in \mathbb{R} \).

- Hence, if \( \gamma: I \to \mathbb{R}^n \) is simple then it is injective on \( \text{int}(I) \).
Many properties of curves can also be stated independently of a specific parametrization. E.g., we can regard a curve $C$ to be simple if there exists one parametrization of $C$ that is simple.
Curves in $\mathbb{R}^n$

- Many properties of curves can also be stated independently of a specific parametrization. E.g., we can regard a curve $C$ to be simple if there exists one parametrization of $C$ that is simple.

- In daily math, the standard meaning of a “curve” is the image of the equivalence class of all paths under a certain equivalence relation. (Roughly, two paths are equivalent if they are identical up to re-parametrization.)

- Hence, the distinction between a curve and (one of) its parametrizations is often blurred.

- For the sake of simplicity, we will not distinguish between a curve $C$ and one of its parametrizations $\gamma$ if the meaning is clear.

- Similarly, we will frequently call $\gamma$ a curve.

- For instance, we will frequently speak about a closed curve rather than about a closed parametrization of a curve.
Definition 177 (Jordan curve, Dt.: Jordankurve)

A set $C \subset \mathbb{R}^2$ (which is not a single point) is called a Jordan curve if there exists a simple and closed parametrization $\gamma: I \rightarrow \mathbb{R}^2$ that parameterizes $C$. 
Definition 177 (Jordan curve, Dt.: Jordankurve)

A set $C \subset \mathbb{R}^2$ (which is not a single point) is called a *Jordan curve* if there exists a simple and closed parametrization $\gamma : I \to \mathbb{R}^2$ that parameterizes $C$.

Theorem 178 (Jordan 1887)

Every Jordan curve $C$ partitions $\mathbb{R}^2 \setminus C$ into two disjoint open regions, a (bounded) “interior” region and an (unbounded) “exterior” region, with $C$ as the (topological) boundary of both of them.
Jordan Curve in $\mathbb{R}^2$

**Definition 177 (Jordan curve, Dt.: Jordankurve)**

A set $C \subset \mathbb{R}^2$ (which is not a single point) is called a *Jordan curve* if there exists a simple and closed parametrization $\gamma : I \to \mathbb{R}^2$ that parameterizes $C$.

**Theorem 178 (Jordan 1887)**

Every Jordan curve $C$ partitions $\mathbb{R}^2 \setminus C$ into two disjoint open regions, a (bounded) “interior” region and an (unbounded) “exterior” region, with $C$ as the (topological) boundary of both of them.

- Although this theorem — the so-called Jordan Curve Theorem (Dt.: Jordanscher Kurvensatz) — seems obvious, a proof is not entirely trivial.
Jordan Curve in $\mathbb{R}^2$

**Definition 177 (Jordan curve, Dt.: Jordankurve)**

A set $C \subset \mathbb{R}^2$ (which is not a single point) is called a *Jordan curve* if there exists a simple and closed parametrization $\gamma : I \to \mathbb{R}^2$ that parameterizes $C$.

**Theorem 178 (Jordan 1887)**

Every Jordan curve $C$ partitions $\mathbb{R}^2 \setminus C$ into two disjoint open regions, a (bounded) “interior” region and an (unbounded) “exterior” region, with $C$ as the (topological) boundary of both of them.

- Although this theorem — the so-called Jordan Curve Theorem (Dt.: Jordanscher Kurvensatz) — seems obvious, a proof is not entirely trivial.

**Theorem 179 (Schönflies 1906)**

For every Jordan curve $C$ there exists a homeomorphism from the plane to itself that maps $C$ to the unit sphere $S^1$.

- Roughly, a homeomorphism is a bijective continuous stretching and bending of one space into another space such that the inverse function also is continuous.
Tangent Vector for a Curve in $\mathbb{R}^n$

**Definition 180 (Tangent vector, Dt.: Tangentenvektor)**

Consider a differentiable parametrization $\gamma: I \rightarrow \mathbb{R}^n$ of a curve $C$. For $t \in I$, a tangent vector at $\gamma(t)$ with respect to $\gamma$ is given by $\gamma'(t)$. Note that $\gamma'(t)$ is a vector-valued function!
Definition 180 (Tangent vector, Dt.: Tangentenvektor)
Consider a differentiable parametrization $\gamma : I \to \mathbb{R}^n$ of a curve $C$. For $t \in I$, a tangent vector at $\gamma(t)$ with respect to $\gamma$ is given by $\gamma'(t)$.

- Note that $\gamma'(t)$ is a vector-valued function!
- It is straightforward to extend the definition of a tangent vector to parametrizations that are piecewise differentiable.
Definition 181 (Parametric surface)

Let \( \Omega \subseteq \mathbb{R}^2 \). A continuous mapping \( \alpha : \Omega \rightarrow \mathbb{R}^3 \) is called a *parametrization* of \( \alpha(\Omega) \), and \( \alpha(\Omega) \) is called the (parametric) *surface* parameterized by \( \alpha \).
Definition 181 (Parametric surface)

Let \( \Omega \subseteq \mathbb{R}^2 \). A continuous mapping \( \alpha : \Omega \to \mathbb{R}^3 \) is called a *parametrization* of \( \alpha(\Omega) \), and \( \alpha(\Omega) \) is called the (parametric) *surface* parameterized by \( \alpha \).

For instance, every point on the surface of Earth can be described by the geographic coordinates longitude and latitude. Note that parametrizations of a surface (regarded as a set \( S \subseteq \mathbb{R}^3 \)) need not be unique: two different parametrizations \( \alpha \) and \( \beta \) may exist such that \( S = \alpha(\Omega_1) = \beta(\Omega_2) \). For simplicity, we will not distinguish between a surface and one of its parametrizations if the meaning is clear.
Surfaces in $\mathbb{R}^3$

**Definition 181 (Parametric surface)**

Let $\Omega \subseteq \mathbb{R}^2$. A continuous mapping $\alpha : \Omega \to \mathbb{R}^3$ is called a *parametrization* of $\alpha(\Omega)$, and $\alpha(\Omega)$ is called the (parametric) *surface* parameterized by $\alpha$.

For instance, every point on the surface of Earth can be described by the geographic coordinates longitude and latitude.
Surfaces in $\mathbb{R}^3$

**Definition 181 (Parametric surface)**

Let $\Omega \subseteq \mathbb{R}^2$. A continuous mapping $\alpha : \Omega \rightarrow \mathbb{R}^3$ is called a *parametrization* of $\alpha(\Omega)$, and $\alpha(\Omega)$ is called the (parametric) *surface* parameterized by $\alpha$.

- For instance, every point on the surface of Earth can be described by the geographic coordinates longitude and latitude.
- Note that parametrizations of a surface (regarded as a set $S \subset \mathbb{R}^3$) need not be unique: two different parametrizations $\alpha$ and $\beta$ may exist such that $S = \alpha(\Omega_1) = \beta(\Omega_2)$. 
Surfaces in $\mathbb{R}^3$

**Definition 181 (Parametric surface)**

Let $\Omega \subseteq \mathbb{R}^2$. A continuous mapping $\alpha: \Omega \rightarrow \mathbb{R}^3$ is called a *parametrization* of $\alpha(\Omega)$, and $\alpha(\Omega)$ is called the (parametric) *surface* parameterized by $\alpha$.

- For instance, every point on the surface of Earth can be described by the geographic coordinates longitude and latitude.
- Note that parametrizations of a surface (regarded as a set $S \subset \mathbb{R}^3$) need not be unique: two different parametrizations $\alpha$ and $\beta$ may exist such that $S = \alpha(\Omega_1) = \beta(\Omega_2)$.
- For simplicity, we will not distinguish between a surface and one of its parametrizations if the meaning is clear.
Sample Parametric Surface: Frustum of a Paraboloid

\[ \alpha: [0, 1] \times [0, 2\pi] \to \mathbb{R}^3 \]

\[ \alpha(u, v) := \begin{pmatrix} u \cos v \\ u \sin v \\ 2u^2 \end{pmatrix} \]
Sample Parametric Surface: Torus

\[ \alpha : [0, 2\pi]^2 \rightarrow \mathbb{R}^3 \]

\[ \alpha(u, v) := \begin{pmatrix} (2 + \cos v) \cos u \\ (2 + \cos v) \sin u \\ \sin v \end{pmatrix} \]
Surfaces in $\mathbb{R}^3$

Lemma 182

Consider a differentiable parametrization $\alpha : \Omega \to \mathbb{R}^3$ of a surface $S$. For $(s, t) \in \Omega$, tangent vectors at $\alpha(s, t)$ with respect to $\alpha$ are given by $\frac{\partial \alpha}{\partial s}(s, t)$ and $\frac{\partial \alpha}{\partial t}(s, t)$. 

Definition 183 (Normal vector, Dt.: Normalvektor)

Consider a differentiable parametrization $\alpha : \Omega \to \mathbb{R}^3$ of a surface $S$. A normal vector $n_{\alpha}(s, t)$ at $\alpha(s, t)$ with respect to $\alpha$ is given by $n_{\alpha}(s, t) := \frac{\partial \alpha}{\partial s}(s, t) \times \frac{\partial \alpha}{\partial t}(s, t)$. The vector $n_{\alpha}(s, t)$ is indeed a normal vector of the tangential plane at $\alpha(s, t)$. 

© M. Held (Univ. Salzburg)
Surfaces in $\mathbb{R}^3$

**Lemma 182**
Consider a differentiable parametrization $\alpha : \Omega \rightarrow \mathbb{R}^3$ of a surface $S$. For $(s, t) \in \Omega$, tangent vectors at $\alpha(s, t)$ with respect to $\alpha$ are given by $\frac{\partial \alpha}{\partial s}(s, t)$ and $\frac{\partial \alpha}{\partial t}(s, t)$.

**Definition 183 (Normal vector, Dt.: Normalvektor)**
Consider a differentiable parametrization $\alpha : \Omega \rightarrow \mathbb{R}^3$ of a surface $S$. A normal vector $n_\alpha(s, t)$ at $\alpha(s, t)$ with respect to $\alpha$ is given by

$$n_\alpha(s, t) := \frac{\partial \alpha}{\partial s}(s, t) \times \frac{\partial \alpha}{\partial t}(s, t).$$
Lemma 182

Consider a differentiable parametrization \( \alpha: \Omega \to \mathbb{R}^3 \) of a surface \( S \). For \((s, t) \in \Omega\), tangent vectors at \( \alpha(s, t) \) with respect to \( \alpha \) are given by \( \frac{\partial \alpha}{\partial s}(s, t) \) and \( \frac{\partial \alpha}{\partial t}(s, t) \). 

Definition 183 (Normal vector, Dt.: Normalvektor)

Consider a differentiable parametrization \( \alpha: \Omega \to \mathbb{R}^3 \) of a surface \( S \). A normal vector \( n_\alpha(s, t) \) at \( \alpha(s, t) \) with respect to \( \alpha \) is given by

\[
n_\alpha(s, t) := \frac{\partial \alpha}{\partial s}(s, t) \times \frac{\partial \alpha}{\partial t}(s, t).
\]

- The vector \( n_\alpha(s, t) \) is indeed a normal vector of the tangential plane at \( \alpha(s, t) \).
A circle in $\mathbb{R}^2$ (with center $c$ and radius $r$) can be parameterized as follows:

$$(c_x + r \cos \phi, c_y + r \sin \phi) \quad \text{with } \phi \in [0, 2\pi[.$$
- A circle in $\mathbb{R}^2$ (with center $c$ and radius $r$) can be parameterized as follows:

\[
(c_x + r \cos \phi, c_y + r \sin \phi) \quad \text{with} \quad \phi \in [0, 2\pi].
\]

- A sphere in $\mathbb{R}^3$ (with center $c$ and radius $r$) can be parameterized as follows:

\[
(c_x + r \cos \delta \cos \phi, c_y + r \cos \delta \sin \phi, c_z + r \sin \delta) \quad \text{with} \quad \phi \in [0, 2\pi], \delta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].
\]
Geometric Objects

- Lines and Planes
- Circles and Spheres
- Conics
- Curves and Surfaces
- Polygons and Polyhedra
Definition 184 (Polygonal curve, Dt.: Polygonzug)

Consider the sequence of points \( P_0, P_1, P_2, \ldots, P_n \in \mathbb{R}^d \), for some \( d, n \in \mathbb{N} \). The polygonal curve (or polygonal chain, polygonal profile) specified by these points (“vertices”) is given by

\[
\bigcup_{i=0}^{n-1} \overline{p_i p_{i+1}}.
\]

Hence, a polygonal curve is a sequence of finitely many vertices connected by straight-line segments such that each segment (except for the first) starts at the end of the previous segment. Unless specified otherwise, we will always assume that all vertices of a polygonal curve are co-planar, i.e., that the polygonal curve is plane.

A polygonal curve with vertices \( P_0, P_1, P_2, \ldots, P_n \in \mathbb{R}^d \) is commonly denoted by \((P_0, P_1, \ldots, P_n-1, P_n)\).
Definition 184 (Polygonal curve, Dt.: Polygonzug)

Consider the sequence of points \( P_0, P_1, P_2, \ldots, P_n \in \mathbb{R}^d \), for some \( d, n \in \mathbb{N} \). The polygonal curve (or polygonal chain, polygonal profile) specified by these points ("vertices") is given by

\[
\bigcup_{i=0}^{n-1} p_i p_{i+1}.
\]

- Hence, a polygonal curve is a sequence of finitely many vertices connected by straight-line segments such that each segment (except for the first) starts at the end of the previous segment.
- Unless specified otherwise, we will always assume that all vertices of a polygonal curve are co-planar, i.e., that the polygonal curve is plane.
- A polygonal curve with vertices \( P_0, P_1, P_2, \ldots, P_n \in \mathbb{R}^d \) is commonly denoted by \((P_0, P_1, \ldots, P_{n-1}, P_n)\).
Definition 185 (Polygon)

A *polygon* is a polygonal curve with vertices $P_0, P_1, P_2, \ldots, P_n \in \mathbb{R}^d$ such that $P_0 = P_n$. 
A polygon is a polygonal curve with vertices $P_0, P_1, P_2, \ldots, P_n \in \mathbb{R}^d$ such that $P_0 = P_n$. 

- A polygon with vertices $P_0, P_1, P_2, \ldots, P_n \in \mathbb{R}^d$, with $P_0 = P_n$, is commonly called an $n$-gon.
- Note that a polygon need not form a simple curve!
Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon \( \mathcal{P} := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( \mathcal{P} \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \).

The signed area of \( \mathcal{P} \) is positive if and only if \( \mathcal{P} \) is oriented counter-clockwise (CCW).

Aka: Shoelace formula or surveyor’s formula in English textbooks.

If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Area and Orientation of a Polygon

Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon \( \mathcal{P} := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( \mathcal{P} \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \).

\[
\text{Signed area of } \mathcal{P} = \sum_{i=1}^{n} \frac{1}{2} \left[ (x_1 y_2 - x_2 y_1) + (x_2 y_3 - x_3 y_2) + \cdots + (x_n y_1 - x_1 y_n) \right],
\]

where \( p_i := (x_i, y_i) \).

The signed area of \( \mathcal{P} \) is positive if and only if \( \mathcal{P} \) is oriented counter-clockwise (CCW).

Aka: Shoelace formula or surveyor’s formula in English textbooks.
If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Area and Orientation of a Polygon

**Theorem 186 (Meister (1769), Gauß (1795))**

Consider a simple plane polygon \( P := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( P \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \). That is, the (signed) area of \( P \) equals

\[
\sum_{i=1}^{n} A_\Delta(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1 y_2 - x_2 y_1) + (x_2 y_3 - x_3 y_2) + \cdots + (x_n y_1 - x_1 y_n)],
\]

where \( p_i := (x_i, y_i) \).
Area and Orientation of a Polygon

**Theorem 186 (Meister (1769), Gauß (1795))**

Consider a simple plane polygon \( P := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( P \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \). That is, the (signed) area of \( P \) equals

\[
\sum_{i=1}^{n} A_\Delta(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1y_2 - x_2y_1) + (x_2y_3 - x_3y_2) + \cdots + (x_ny_1 - x_1y_n)],
\]

where \( p_i := \left( \begin{array}{c} x_i \\ y_i \end{array} \right) \). The signed area of \( P \) is positive if and only if \( P \) is oriented counter-clockwise (CCW).

[Diagram of a polygon with a point T and various triangles formed by connecting P to T]
Area and Orientation of a Polygon

Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon \( P := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( P \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \). That is, the (signed) area of \( P \) equals

\[
\sum_{i=1}^{n} A_{\Delta}(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1y_2 - x_2y_1) + (x_2y_3 - x_3y_2) + \cdots + (x_ny_1 - x_1y_n)],
\]

where \( p_i := \left( \begin{array}{c} x_i \\ y_i \end{array} \right) \). The signed area of \( P \) is positive if and only if \( P \) is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
Area and Orientation of a Polygon

**Theorem 186 (Meister (1769), Gauß (1795))**

Consider a simple plane polygon \( \mathcal{P} := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( \mathcal{P} \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \). That is, the (signed) area of \( \mathcal{P} \) equals

\[
\sum_{i=1}^{n} A_\Delta(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1 y_2 - x_2 y_1) + (x_2 y_3 - x_3 y_2) + \cdots + (x_n y_1 - x_1 y_n)],
\]

where \( p_i := \left( \frac{\bar{x}_i}{\bar{y}_i} \right) \). The signed area of \( \mathcal{P} \) is positive if and only if \( \mathcal{P} \) is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Area and Orientation of a Polygon

Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon $\mathcal{P} := (P_0, P_1, P_2, \ldots, P_n)$, with $P_0 = P_n$, and pick a point $T$ in the plane. Then the (signed) area of $\mathcal{P}$ is given by the sum of the signed areas of the individual triangles $\Delta(T, P_{i-1}, P_i)$. That is, the (signed) area of $\mathcal{P}$ equals

$$\sum_{i=1}^{n} A_{\Delta}(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1y_2 - x_2y_1) + (x_2y_3 - x_3y_2) + \cdots + (x_ny_1 - x_1y_n)],$$

where $p_i := (\frac{x_i}{y_i})$. The signed area of $\mathcal{P}$ is positive if and only if $\mathcal{P}$ is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Area and Orientation of a Polygon

**Theorem 186 (Meister (1769), Gauß (1795))**

Consider a simple plane polygon $P := (P_0, P_1, P_2, \ldots, P_n)$, with $P_0 = P_n$, and pick a point $T$ in the plane. Then the (signed) area of $P$ is given by the sum of the signed areas of the individual triangles $\Delta(T, P_{i-1}, P_i)$. That is, the (signed) area of $P$ equals

$$\sum_{i=1}^{n} A_{\Delta}(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1 y_2 - x_2 y_1) + (x_2 y_3 - x_3 y_2) + \cdots + (x_n y_1 - x_1 y_n)],$$

where $p_i := (\frac{x_i}{y_i})$. The signed area of $P$ is positive if and only if $P$ is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon \( P := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( P \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \). That is, the (signed) area of \( P \) equals

\[
\sum_{i=1}^{n} A_{\Delta}(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1 y_2 - x_2 y_1) + (x_2 y_3 - x_3 y_2) + \cdots + (x_n y_1 - x_1 y_n)],
\]

where \( p_i := \left( \begin{array}{c} x_i \\ y_i \end{array} \right) \). The signed area of \( P \) is positive if and only if \( P \) is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Area and Orientation of a Polygon

Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon \( P := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( P \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \). That is, the (signed) area of \( P \) equals

\[
\sum_{i=1}^{n} A_{\Delta}(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1y_2 - x_2y_1) + (x_2y_3 - x_3y_2) + \cdots + (x_ny_1 - x_1y_n)],
\]

where \( p_i := (x_i, y_i) \). The signed area of \( P \) is positive if and only if \( P \) is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Area and Orientation of a Polygon

Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon \( \mathcal{P} := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( \mathcal{P} \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \). That is, the (signed) area of \( \mathcal{P} \) equals

\[
\sum_{i=1}^{n} A_{\Delta}(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1y_2 - x_2y_1) + (x_2y_3 - x_3y_2) + \cdots + (x_ny_1 - x_1y_n)],
\]

where \( p_i := \left( \begin{array}{c} x_i \\ y_i \end{array} \right) \). The signed area of \( \mathcal{P} \) is positive if and only if \( \mathcal{P} \) is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Area and Orientation of a Polygon

Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon $P := (P_0, P_1, P_2, \ldots, P_n)$, with $P_0 = P_n$, and pick a point $T$ in the plane. Then the (signed) area of $P$ is given by the sum of the signed areas of the individual triangles $\Delta(T, P_{i-1}, P_i)$. That is, the (signed) area of $P$ equals

$$\sum_{i=1}^{n} A_\Delta(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1 y_2 - x_2 y_1) + (x_2 y_3 - x_3 y_2) + \cdots + (x_n y_1 - x_1 y_n)],$$

where $p_i := \left(\frac{x_i}{y_i}\right)$. The signed area of $P$ is positive if and only if $P$ is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Theorem 186 (Meister (1769), Gauß (1795))

Consider a simple plane polygon \( \mathcal{P} := (P_0, P_1, P_2, \ldots, P_n) \), with \( P_0 = P_n \), and pick a point \( T \) in the plane. Then the (signed) area of \( \mathcal{P} \) is given by the sum of the signed areas of the individual triangles \( \Delta(T, P_{i-1}, P_i) \). That is, the (signed) area of \( \mathcal{P} \) equals

\[
\sum_{i=1}^{n} A_\Delta(T, P_{i-1}, P_i) = \frac{1}{2} \cdot \left[ (x_1y_2 - x_2y_1) + (x_2y_3 - x_3y_2) + \cdots + (x_ny_1 - x_1y_n) \right],
\]

where \( p_i := \left( \frac{x_i}{y_i} \right) \). The signed area of \( \mathcal{P} \) is positive if and only if \( \mathcal{P} \) is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Area and Orientation of a Polygon

**Theorem 186 (Meister (1769), Gauß (1795))**

Consider a simple plane polygon $\mathcal{P} := (P_0, P_1, P_2, \ldots, P_n)$, with $P_0 = P_n$, and pick a point $T$ in the plane. Then the (signed) area of $\mathcal{P}$ is given by the sum of the signed areas of the individual triangles $\Delta(T, P_{i-1}, P_i)$. That is, the (signed) area of $\mathcal{P}$ equals

$$\sum_{i=1}^{n} A_\Delta(T, P_{i-1}, P_i) = \frac{1}{2} \cdot [(x_1y_2 - x_2y_1) + (x_2y_3 - x_3y_2) + \cdots + (x_ny_1 - x_1y_n)],$$

where $p_i := (\frac{x_i}{y_i})$. The signed area of $\mathcal{P}$ is positive if and only if $\mathcal{P}$ is oriented counter-clockwise (CCW).

- Aka: Shoelace formula or surveyor’s formula in English textbooks.
- If multiple polygons bound a polygonal domain then all contours need to be oriented consistently!
Polyhedra

- In solid modeling, a *solid* describes a closed object that could exist in $\mathbb{R}^3$ and, at least theoretically, be manufactured. Well-known examples of solids include balls, cubes, cylinders, and cones.
Polyhedra

- In solid modeling, a *solid* describes a closed object that could exist in $\mathbb{R}^3$ and, at least theoretically, be manufactured. Well-known examples of solids include balls, cubes, cylinders, and cones.

- A *polyhedron* describes a solid that is bounded by planar faces formed by polygons, joined at their edges, such that every edge is shared by an even number of faces.
Polyhedra

- In solid modeling, a *solid* describes a closed object that could exist in $\mathbb{R}^3$ and, at least theoretically, be manufactured. Well-known examples of solids include balls, cubes, cylinders, and cones.

- A *polyhedron* describes a solid that is bounded by planar faces formed by polygons, joined at their edges, such that every edge is shared by an even number of faces.

- In order to guarantee that (the surface of) the polyhedron is water-tight and forms a 2-manifold it is common to demand that
  1. every edge of every face belongs to exactly one other face, and
  2. the faces that share a vertex form a cyclic chain of polygons in which every pair of consecutive polygons shares an edge.

Recall that Euler's Formula $v - e + f = 2$ holds for the vertices, edges and faces of a polyhedron.

Note that the word “polyhedron” has slightly different meanings in solid modeling and graphics, on one hand, and combinatorial geometry and algebraic topology, on the other hand.
Polyhedra

- In solid modeling, a *solid* describes a closed object that could exist in $\mathbb{R}^3$ and, at least theoretically, be manufactured. Well-known examples of solids include balls, cubes, cylinders, and cones.

- A *polyhedron* describes a solid that is bounded by planar faces formed by polygons, joined at their edges, such that every edge is shared by an even number of faces.

- In order to guarantee that (the surface of) the polyhedron is water-tight and forms a 2-manifold it is common to demand that
  1. every edge of every face belongs to exactly one other face, and
  2. the faces that share a vertex form a cyclic chain of polygons in which every pair of consecutive polygons shares an edge.

The first requirement precludes T-junctions, while the second requirement precludes the case of two pyramids touching at a vertex.
Polyhedra

- In solid modeling, a *solid* describes a closed object that could exist in $\mathbb{R}^3$ and, at least theoretically, be manufactured. Well-known examples of solids include balls, cubes, cylinders, and cones.

- A *polyhedron* describes a solid that is bounded by planar faces formed by polygons, joined at their edges, such that every edge is shared by an even number of faces.

- In order to guarantee that (the surface of) the polyhedron is water-tight and forms a 2-manifold it is common to demand that
  1. every edge of every face belongs to exactly one other face, and
  2. the faces that share a vertex form a cyclic chain of polygons in which every pair of consecutive polygons shares an edge.

The first requirement precludes T-junctions, while the second requirement precludes the case of two pyramids touching at a vertex.

- Recall that Euler’s Formula $v - e + f = 2$ holds for the vertices, edges and faces of a polyhedron.

- Note that the word “polyhedron” has slightly different meanings in solid modeling and graphics, on one hand, and combinatorial geometry and algebraic topology, on the other hand.
Basic Concepts of Topology

- Connectedness
- Metric Space
- Topological Properties of Sets
- Topological Properties of Surfaces and Solids
- Triangulations
Basic Concepts of Topology

- Connectedness
- Metric Space
- Topological Properties of Sets
- Topological Properties of Surfaces and Solids
- Triangulations
Definition 187 (Path-connected, Dt.: wegzusammenhängend)

A set $S \subset \mathbb{R}^n$ is path-connected if for every pair of points $P, Q \in S$ there exists a curve that is completely contained in $S$ and that links $P$ and $Q$. 
Definition 187 (Path-connected, Dt.: wegzusammenhängend)

A set $S \subset \mathbb{R}^n$ is *path-connected* if for every pair of points $P, Q \in S$ there exists a curve that is completely contained in $S$ and that links $P$ and $Q$. 

```
not path-connected
```

```
path-connected
```
Definition 187 (Path-connected, Dt.: wegzusammenhängend)

A set $S \subset \mathbb{R}^n$ is path-connected if for every pair of points $P, Q \in S$ there exists a curve that is completely contained in $S$ and that links $P$ and $Q$.

Definition 188 (Simply-connected and multiply-connected)

A path-connected set $S \subset \mathbb{R}^2$ is simply-connected if every simple closed curve entirely contained within $S$ encloses only points of $S$. Otherwise, $S$ is called multiply-connected (or not simply-connected).
Basic Concepts of Topology

- Connectedness
- Metric Space
- Topological Properties of Sets
- Topological Properties of Surfaces and Solids
- Triangulations
Definition 189 (Metric space, Dt.: metrischer Raum)

A metric space is a set of points $\mathcal{X}$ with an associated distance function (aka metric) $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that the following conditions hold for all $x, y, z \in \mathcal{X}$:

1. $d(x, y) \geq 0$.
2. Identity of indiscernibles: $d(x, y) = 0 \Rightarrow x = y$.
3. Reflexivity: $d(x, x) = 0$.
4. Symmetry: $d(x, y) = d(y, x)$.
5. Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

Easy to check: $\mathbb{R}^n$, i.e., $\mathbb{R}^n$ with the Euclidean distance, is a metric space.

Easy to check: Every normed vector space is a metric space by defining $d(x, y) := ||x - y||$. 

Definition 190 (Open ball, Dt.: offene Kugel)

Consider a metric space $\mathcal{X}$ with metric $d$. For $x \in \mathcal{X}$ and $r \in \mathbb{R}^+$ we define the (generalized) open ball (relative to the metric $d$) with radius $r$ centered at $x$ as $B(x, r) := \{ y \in \mathcal{X} : d(x, y) < r \}$. 
Definition 189 (Metric space, Dt.: metrischer Raum)

A metric space is a set of points $\mathcal{X}$ with an associated distance function (aka metric) $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that the following conditions hold for all $x, y, z \in \mathcal{X}$:

1. $d(x, y) \geq 0$. 
2. Identity of indiscernibles: $d(x, y) = 0 \Rightarrow x = y$.
3. Reflexivity: $d(x, x) = 0$.
4. Symmetry: $d(x, y) = d(y, x)$.
5. Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

Easy to check: $\mathbb{E}^n$, i.e., $\mathbb{R}^n$ with the Euclidean distance, is a metric space.

Easy to check: Every normed vector space is a metric space by defining $d(x, y) := ||x − y||$. 

Definition 190 (Open ball, Dt.: offene Kugel)

Consider a metric space $\mathcal{X}$ with metric $d$. For $x \in \mathcal{X}$ and $r \in \mathbb{R}^+$ we define the (generalized) open ball (relative to the metric $d$) with radius $r$ centered at $x$ as $B(x, r) := \{y \in \mathcal{X} : d(x, y) < r\}$. 
Metric Space and Open Ball

**Definition 189 (Metric space, Dt.: metrischer Raum)**

A *metric space* is a set of points $\mathcal{X}$ with an associated distance function (aka *metric*) $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that the following conditions hold for all $x, y, z \in \mathcal{X}$:

1. $d(x, y) \geq 0$.
2. Identity of indiscernibles: $d(x, y) = 0 \implies x = y$.
Metric Space and Open Ball

**Definition 189 (Metric space, Dt.: metrischer Raum)**

A *metric space* is a set of points \( \mathcal{X} \) with an associated distance function (aka *metric*) \( d : \mathcal{X} \times \mathcal{X} \to \mathbb{R} \) such that the following conditions hold for all \( x, y, z \in \mathcal{X} \):

1. \( d(x, y) \geq 0 \).
2. Identity of indiscernibles: \( d(x, y) = 0 \implies x = y \).
3. Reflexivity: \( d(x, x) = 0 \).
Definition 189 (Metric space, Dt.: metrischer Raum)

A *metric space* is a set of points $\mathcal{X}$ with an associated distance function (aka *metric*) $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that the following conditions hold for all $x, y, z \in \mathcal{X}$:

1. $d(x, y) \geq 0$.
2. **Identity of indiscernibles:** $d(x, y) = 0 \Rightarrow x = y$.
3. **Reflexivity:** $d(x, x) = 0$.
4. **Symmetry:** $d(x, y) = d(y, x)$.

Easy to check: $\mathbb{R}^n$, i.e., $\mathbb{R}^n$ with the Euclidean distance, is a metric space.

Easy to check: Every normed vector space is a metric space by defining $d(x, y) := ||x - y||$. 

Definition 190 (Open ball, Dt.: offene Kugel)

Consider a metric space $\mathcal{X}$ with metric $d$. For $x \in \mathcal{X}$ and $r \in \mathbb{R}^+$ we define the (generalized) open ball (relative to the metric $d$) with radius $r$ centered at $x$ as $B(x, r) := \{y \in \mathcal{X} : d(x, y) < r\}$. 

Metric Space and Open Ball

**Definition 189 (Metric space, Dt.: metrischer Raum)**

A *metric space* is a set of points $\mathcal{X}$ with an associated distance function (aka *metric*) $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that the following conditions hold for all $x, y, z \in \mathcal{X}$:

1. $d(x, y) \geq 0$.
2. Identity of indiscernibles: $d(x, y) = 0 \implies x = y$.
3. Reflexivity: $d(x, x) = 0$.
4. Symmetry: $d(x, y) = d(y, x)$.
5. Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

Easy to check: $\mathbb{E}^n$, i.e., $\mathbb{R}^n$ with the Euclidean distance, is a metric space.

Easy to check: Every normed vector space is a metric space by defining $d(x, y) := ||x - y||$. 

**Definition 190 (Open ball, Dt.: offene Kugel)**

Consider a metric space $\mathcal{X}$ with metric $d$. For $x \in \mathcal{X}$ and $r \in \mathbb{R}^+$ we define the (generalized) open ball (relative to the metric $d$) with radius $r$ centered at $x$ as $B(x, r) := \{ y \in \mathcal{X} : d(x, y) < r \}$. 

© M. Held (Univ. Salzburg)
Definition 189 (Metric space, Dt.: metrischer Raum)

A *metric space* is a set of points $\mathcal{X}$ with an associated distance function (aka *metric*) $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that the following conditions hold for all $x, y, z \in \mathcal{X}$:

1. $d(x, y) \geq 0$.
2. Identity of indiscernibles: $d(x, y) = 0 \implies x = y$.
3. Reflexivity: $d(x, x) = 0$.
4. Symmetry: $d(x, y) = d(y, x)$.
5. Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

- Easy to check: $\mathbb{R}^n$, i.e., $\mathbb{R}^n$ with the Euclidean distance, is a metric space.
- Easy to check: Every normed vector space is a metric space by defining $d(x, y) := \|x - y\|$. 

**Metric Space and Open Ball**
Metric Space and Open Ball

**Definition 189 (Metric space, Dt.: metrischer Raum)**

A *metric space* is a set of points $\mathcal{X}$ with an associated distance function (aka *metric*) $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ such that the following conditions hold for all $x, y, z \in \mathcal{X}$:

1. $d(x, y) \geq 0$.
2. Identity of indiscernibles: $d(x, y) = 0 \Rightarrow x = y$.
3. Reflexivity: $d(x, x) = 0$.
4. Symmetry: $d(x, y) = d(y, x)$.
5. Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

- Easy to check: $\mathbb{E}^n$, i.e., $\mathbb{R}^n$ with the Euclidean distance, is a metric space.
- Easy to check: Every normed vector space is a metric space by defining $d(x, y) := ||x - y||$.

**Definition 190 (Open ball, Dt.: offene Kugel)**

Consider a metric space $\mathcal{X}$ with metric $d$. For $x \in \mathcal{X}$ and $r \in \mathbb{R}^+$ we define the *(generalized)* open ball (relative to the metric $d$) with radius $r$ centered at $x$ as

$$B(x, r) := \{y \in \mathcal{X} : d(x, y) < r\}.$$
Basic Concepts of Topology

- Connectedness
- Metric Space
- Topological Properties of Sets
- Topological Properties of Surfaces and Solids
- Triangulations
Consider a space $\mathcal{X}$ that has a metric, and a set $S \subseteq \mathcal{X}$. (E.g., $\mathbb{R}^n$ and the Euclidean metric, and any subset $S$ of $\mathbb{R}^n$.)
Consider a space $\mathcal{X}$ that has a metric, and a set $S \subseteq \mathcal{X}$. (E.g., $\mathbb{R}^n$ and the Euclidean metric, and any subset $S$ of $\mathbb{R}^n$.)

**Definition 191 (Interior point, Dt.: innerer Punkt)**

A point $x \in \mathcal{X}$ is an *interior point* of $S$ if there exists a radius $r > 0$ such that the open ball with center $x$ and radius $r$ is completely contained in $S$, i.e., $B(x, r) \subseteq S$. 

**Definition 192 (Interior, Dt.: Inneres)**

The set of all interior points of $S$ is the *interior* of $S$, often denoted by $\text{int}(S)$ or $S^\circ$.

**Lemma 193**

We have $\text{int}(S) \subseteq S$ for all $S \subseteq \mathcal{X}$.

**Lemma 194**

For all $x \in \mathcal{X}$, the interior of an open ball $B(x, r) \subseteq \mathcal{X}$ is the open ball itself.
Interior, Exterior and Closure

Consider a space $\mathcal{X}$ that has a metric, and a set $S \subseteq \mathcal{X}$. (E.g., $\mathbb{R}^n$ and the Euclidean metric, and any subset $S$ of $\mathbb{R}^n$.)

**Definition 191 (Interior point, Dt.: innerer Punkt)**

A point $x \in \mathcal{X}$ is an *interior point* of $S$ if there exists a radius $r > 0$ such that the open ball with center $x$ and radius $r$ is completely contained in $S$, i.e., $B(x, r) \subseteq S$.

**Definition 192 (Interior, Dt.: Inneres)**

The set of all interior points of $S$ is the *interior* of $S$, often denoted by $\text{int}(S)$ or $S^\circ$.
Consider a space $\mathcal{X}$ that has a metric, and a set $S \subseteq \mathcal{X}$. (E.g., $\mathbb{R}^n$ and the Euclidean metric, and any subset $S$ of $\mathbb{R}^n$.)

**Definition 191 (Interior point, Dt.: innerer Punkt)**

A point $x \in \mathcal{X}$ is an *interior point* of $S$ if there exists a radius $r > 0$ such that the open ball with center $x$ and radius $r$ is completely contained in $S$, i.e., $B(x, r) \subseteq S$.

**Definition 192 (Interior, Dt.: Inneres)**

The set of all interior points of $S$ is the *interior* of $S$, often denoted by $\text{int}(S)$ or $S^\circ$.

**Lemma 193**

We have $\text{int}(S) \subseteq S$ for all $S \subseteq \mathcal{X}$. 
Consider a space $X$ that has a metric, and a set $S \subseteq X$. (E.g., $\mathbb{R}^n$ and the Euclidean metric, and any subset $S$ of $\mathbb{R}^n$.)

**Definition 191 (Interior point, Dt.: innerer Punkt)**

A point $x \in X$ is an *interior point* of $S$ if there exists a radius $r > 0$ such that the open ball with center $x$ and radius $r$ is completely contained in $S$, i.e., $B(x, r) \subseteq S$.

**Definition 192 (Interior, Dt.: Inneres)**

The set of all interior points of $S$ is the *interior* of $S$, often denoted by $\text{int}(S)$ or $S^\circ$.

**Lemma 193**

We have $\text{int}(S) \subseteq S$ for all $S \subseteq X$.

**Lemma 194**

For all $x \in X$, the interior of an open ball $B(x, r) \subseteq X$ is the open ball itself.
Interior, Exterior and Closure

**Definition 195 (Exterior point, Dt.: äußerer Punkt)**

A point $y \in \mathcal{X}$ is an *exterior point* of $S$ if there exists a radius $r > 0$ such that the open ball with center $y$ and radius $r$ is completely contained in the complement of $S$ (with respect to $\mathcal{X}$), i.e., $B(y, r) \subseteq (\mathcal{X} \setminus S)$. 

**Definition 196 (Exterior, Dt.: Äußeres)**

The set of all exterior points of $S$ is the *exterior* of $S$, denoted by $\text{ext}(S)$.

**Definition 197 (Boundary, Dt.: Rand)**

All points of $\mathcal{X}$ that are neither in the interior nor in the exterior of $S$ form the *boundary*, $\partial S$, of $S$. 
Definition 195 (Exterior point, Dt.: äußerer Punkt)
A point \( y \in \mathcal{X} \) is an **exterior point** of \( S \) if there exists a radius \( r > 0 \) such that the open ball with center \( y \) and radius \( r \) is completely contained in the complement of \( S \) (with respect to \( \mathcal{X} \)), i.e., \( B(y, r) \subseteq (\mathcal{X} \setminus S) \).

Definition 196 (Exterior, Dt.: Äußeres)
The set of all exterior points of \( S \) is the **exterior** of \( S \), denoted by \( \text{ext}(S) \).
Definition 195 (Exterior point, Dt.: äußerer Punkt)
A point \( y \in \mathcal{X} \) is an exterior point of \( S \) if there exists a radius \( r > 0 \) such that the open ball with center \( y \) and radius \( r \) is completely contained in the complement of \( S \) (with respect to \( \mathcal{X} \)), i.e., \( B(y, r) \subseteq (\mathcal{X} \setminus S) \).

Definition 196 (Exterior, Dt.: Äußeres)
The set of all exterior points of \( S \) is the exterior of \( S \), denoted by \( \text{ext}(S) \).

Definition 197 (Boundary, Dt.: Rand)
All points of \( \mathcal{X} \) that are neither in the interior nor in the exterior of \( S \) form the boundary, \( \partial S \), of \( S \).
In the figure, relative to the standard Euclidean distance in $\mathbb{R}^2$, $A$ is an interior point, $B$ is on the boundary, and $C$ is an exterior point.
In the figure, relative to the standard Euclidean distance in $\mathbb{R}^2$, $A$ is an interior point, $B$ is on the boundary, and $C$ is an exterior point.

**Lemma 198**

For all $S \subseteq X$, the union of the interior, the exterior and the boundary of $S$ constitutes the whole space $X$. 
**Definition 199 (Closure, Dt.: Abschluß)**

The *closure* $\overline{S}$ of a set $S$ is the union of the interior and the boundary of $S$. 

**Lemma 200**

The closure $\overline{S}$ of a set $S$ is given by all points of $X$ that are not in the exterior of $S$. 

**Definition 201 (Open, Dt.: offen)**

A set $S \subseteq X$ is called *open* if $\text{int}(S) = S$. 

**Definition 202 (Closed, Dt.: abgeschlossen)**

A set $S \subseteq X$ is called *closed* if the complement of $S$ (relative to $X$) is open. 

Note that there exist spaces $X$ and subsets $S \subset X$ such that the interior or the exterior or the boundary of $S$ are empty. 

Warning: Intuition may easily misguide one's judgement once general spaces or metrics are studied!
**Definition 199 (Closure, Dt.: Abschluß)**

The closure $\overline{S}$ of a set $S$ is the union of the interior and the boundary of $S$.

**Lemma 200**

The closure $\overline{S}$ of a set $S$ is given by all points of $\mathcal{X}$ that are not in the exterior of $S$. 

**Note:**

There exist spaces $\mathcal{X}$ and subsets $S \subseteq \mathcal{X}$ such that the interior or the exterior or the boundary of $S$ are empty.

**Warning:**

Intuition may easily misguide one's judgement once general spaces or metrics are studied!
Definition 199 (Closure, Dt.: Abschluß)
The closure $\overline{S}$ of a set $S$ is the union of the interior and the boundary of $S$.

Lemma 200
The closure $\overline{S}$ of a set $S$ is given by all points of $X$ that are not in the exterior of $S$.

Definition 201 (Open, Dt.: offen)
A set $S \subseteq X$ is called open if $\text{int}(S) = S$. 

Note that there exist spaces $X$ and subsets $S \subseteq X$ such that the interior or the exterior or the boundary of $S$ are empty.

Warning: Intuition may easily misguide one's judgement once general spaces or metrics are studied!
Interior, Exterior and Closure

**Definition 199 (Closure, Dt.: Abschluß)**

The *closure* \( \overline{S} \) of a set \( S \) is the union of the interior and the boundary of \( S \).

**Lemma 200**

The closure \( \overline{S} \) of a set \( S \) is given by all points of \( X \) that are not in the exterior of \( S \).

**Definition 201 (Open, Dt.: offen)**

A set \( S \subseteq X \) is called *open* if \( \text{int}(S) = S \).

**Definition 202 (Closed, Dt.: abgeschlossen)**

A set \( S \subseteq X \) is called *closed* if the complement of \( S \) (relative to \( X \)) is open.

Note that there exist spaces \( X \) and subsets \( S \subseteq X \) such that the interior or the exterior or the boundary of \( S \) are empty.

Warning: Intuition may easily misguide one's judgement once general spaces or metrics are studied!
Interior, Exterior and Closure

**Definition 199 (Closure, Dt.: Abschluß)**

The closure $\overline{S}$ of a set $S$ is the union of the interior and the boundary of $S$.

**Lemma 200**

The closure $\overline{S}$ of a set $S$ is given by all points of $\mathcal{X}$ that are not in the exterior of $S$.

**Definition 201 (Open, Dt.: offen)**

A set $S \subseteq \mathcal{X}$ is called open if $\text{int}(S) = S$.

**Definition 202 (Closed, Dt.: abgeschlossen)**

A set $S \subseteq \mathcal{X}$ is called closed if the complement of $S$ (relative to $\mathcal{X}$) is open.

- Note that there exist spaces $\mathcal{X}$ and subsets $S \subseteq \mathcal{X}$ such that the interior or the exterior or the boundary of $S$ are empty.
- Warning: Intuition may easily misguide one’s judgement once general spaces or metrics are studied!
Consider a ball in $\mathbb{E}^3$ with radius $r$ centered at the origin:

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$
Consider a ball in $\mathbb{E}^3$ with radius $r$ centered at the origin:

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$  

The interior of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < r^2\}.$$  

The closure of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$  

The exterior of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 > r^2\}.$$  

The boundary of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = r^2\}.$$
Consider a ball in $\mathbb{E}^3$ with radius $r$ centered at the origin:

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$

The interior of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < r^2\}.$$

The closure of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$
Consider a ball in $\mathbb{E}^3$ with radius $r$ centered at the origin:

\[ \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\} . \]

The interior of the ball is

\[ \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < r^2\} . \]

The closure of the ball is

\[ \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\} . \]

The exterior of the ball is

\[ \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 > r^2\} . \]
Interior, Exterior and Closure

- Consider a ball in $\mathbb{E}^3$ with radius $r$ centered at the origin:
  $$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$

- The interior of the ball is
  $$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < r^2\}.$$

- The closure of the ball is
  $$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$

- The exterior of the ball is
  $$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 > r^2\}.$$

- The boundary of the ball is
  $$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = r^2\}.$$
Basic Concepts of Topology
- Connectedness
- Metric Space
- Topological Properties of Sets
- Topological Properties of Surfaces and Solids
- Triangulations
Manifolds

- Informally speaking, 2-manifolds are surfaces in 3D that are locally two-dimensional, i.e., that locally (at each point of the manifold) resemble a “bent copy of a rubber plane”.

**Definition 203 (Manifold, Dt.: Mannigfaltigkeit)**

A set \( S \subseteq \mathbb{R}^3 \) is a 2-manifold (or simply a “manifold”) if for every point \( x \in S \) there exists an open neighborhood of \( x \) in \( S \) which is homeomorphic to an open disk.

Roughly, a homeomorphism is a bijective function between two spaces that is continuous and that also has a continuous inverse. It establishes a “topological equivalence” between the spaces and, by a continuous stretching and bending, between their objects.
Manifolds

- Informally speaking, 2-manifolds are surfaces in 3D that are locally two-dimensional, i.e., that locally (at each point of the manifold) resemble a “bent copy of a rubber plane”.

**Definition 203 (Manifold, Dt.: Mannigfaltigkeit)**

A set $S \subset \mathbb{R}^3$ is a 2-manifold (or simply a “manifold”) if for every point $x \in S$ there exists an open neighborhood of $x$ in $S$ which is homeomorphic to an open disk.

Roughly, a homeomorphism is a bijective function between two spaces that is continuous and that also has a continuous inverse. It establishes a "topological equivalence" between the spaces and, by a continuous stretching and bending, between their objects.
Informally speaking, 2-manifolds are surfaces in 3D that are locally two-dimensional, i.e., that locally (at each point of the manifold) resemble a “bent copy of a rubber plane”.

**Definition 203 (Manifold, Dt.: Mannigfaltigkeit)**

A set $S \subset \mathbb{R}^3$ is a 2-*manifold* (or simply a “manifold”) if for every point $x \in S$ there exists an open neighborhood of $x$ in $S$ which is homeomorphic to an open disk.

Roughly, a homeomorphism is a bijective function between two spaces that is continuous and that also has a continuous inverse. It establishes a “topological equivalence” between the spaces and, by a continuous stretching and bending, between their objects.
Informally speaking, 2-manifolds are surfaces in 3D that are locally two-dimensional, i.e., that locally (at each point of the manifold) resemble a “bent copy of a rubber plane”.

**Definition 203 (Manifold, Dt.: Mannigfaltigkeit)**

A set \( S \subset \mathbb{R}^3 \) is a 2-\emph{manifold} (or simply a “manifold”) if for every point \( x \in S \) there exists an open neighborhood of \( x \) in \( S \) which is homeomorphic to an open disk.

Roughly, a homeomorphism is a bijective function between two spaces that is continuous and that also has a continuous inverse. It establishes a “topological equivalence” between the spaces and, by a continuous stretching and bending, between their objects.
Orientable Surface

Definition 204 (Orientable, Dt.: orientierbar)

A 2-manifold is *orientable* if a unit normal vector can be defined consistently for every point on the surface such that it varies continuously over the surface.

Gluing the ends of a strip of paper together after a twist yields a one-sided surface called a *Möbius strip* (Dt.: Möbiusband), which is not orientable.

See [https://www.youtube.com/watch?v=AmgkSdhK4K8](https://www.youtube.com/watch?v=AmgkSdhK4K8) for a cool application of topology and, in particular, of Möbius strips.
Orientable Surface

Definition 204 (Orientable, Dt.: orientierbar)
A 2-manifold is *orientable* if a unit normal vector can be defined consistently for every point on the surface such that it varies continuously over the surface.

- Gluing the ends of a strip of paper together after a twist yields a one-sided surface called a *Möbius strip* (Dt.: Möbiusband), which is not *orientable*.

- See [https://www.youtube.com/watch?v=AmgkSdhK4K8](https://www.youtube.com/watch?v=AmgkSdhK4K8) for a cool application of topology and, in particular, of Möbius strips.
The topologically simplest connected closed 2-manifold in 3D is a sphere.

By adding a “handle” to the sphere we get a torus.

It is well-known that every manifold surface can be obtained by adding a certain number of handles to the sphere.
Genus

- The topologically simplest connected closed 2-manifold in 3D is a sphere.
- By adding a “handle” to the sphere we get a torus.
- It is well-known that every manifold surface can be obtained by adding a certain number of handles to the sphere.

Definition 205 (Genus, Dt.: Geschlecht)

A connected orientable manifold surface is said to have genus $k$ if it can be cut along $k$ non-intersecting closed simple curves without causing the resultant manifold to become disconnected.
Genus

- The topologically simplest connected closed 2-manifold in 3D is a sphere.
- By adding a “handle” to the sphere we get a torus.
- It is well-known that every manifold surface can be obtained by adding a certain number of handles to the sphere.

**Definition 205 (Genus, Dt.: Geschlecht)**

A connected orientable manifold surface is said to have *genus* $k$ if it can be cut along $k$ non-intersecting closed simple curves without causing the resultant manifold to become disconnected.

- Equivalently, a manifold of genus $k$ can be obtained by adding $k$ handles to the sphere.
Genus

- The topologically simplest connected closed 2-manifold in 3D is a sphere.
- By adding a “handle” to the sphere we get a torus.
- It is well-known that every manifold surface can be obtained by adding a certain number of handles to the sphere.

**Definition 205 (Genus, Dt.: Geschlecht)**

A connected orientable manifold surface is said to have genus $k$ if it can be cut along $k$ non-intersecting closed simple curves without causing the resultant manifold to become disconnected.

- Equivalently, a manifold of genus $k$ can be obtained by adding $k$ handles to the sphere.
- Note that a general surface can also be obtained by “punching holes” through a sphere.
- However, it is not difficult to see that, topologically, adding a handle is equivalent to opening a hole on a surface.
Basic Concepts of Topology

- Connectedness
- Metric Space
- Topological Properties of Sets
- Topological Properties of Surfaces and Solids
- Triangulations
Definition 206 (Triangulation)

Let $S = \{P_1, P_2, \ldots, P_k\}$ be a set of $k$ points in $\mathbb{R}^2$. A structure $T$ is called a triangulation of $S$ if $T$ is (the straight-line embedding of) a connected planar graph such that

- all bounded faces of $T$ are triangles,
- the union of the bounded triangular faces forms the convex hull of $S$. 

---

![Image of triangulated points](image.png)
Definition 206 (Triangulation)

Let $S = \{P_1, P_2, \ldots, P_k\}$ be a set of $k$ points in $\mathbb{R}^2$. A structure $T$ is called a \textit{triangulation} of $S$ if $T$ is (the straight-line embedding of) a connected planar graph such that

- $S$ is (the embedding of) the vertex set of $T$,
Definition 206 (Triangulation)

Let $S = \{P_1, P_2, \ldots, P_k\}$ be a set of $k$ points in $\mathbb{R}^2$. A structure $T$ is called a triangulation of $S$ if $T$ is (the straight-line embedding of) a connected planar graph such that

- $S$ is (the embedding of) the vertex set of $T$,
- all bounded faces of $T$ are triangles,
Definition 206 (Triangulation)

Let \( S = \{P_1, P_2, \ldots, P_k\} \) be a set of \( k \) points in \( \mathbb{R}^2 \). A structure \( T \) is called a \textit{triangulation} of \( S \) if \( T \) is (the straight-line embedding of) a connected planar graph such that

- \( S \) is (the embedding of) the vertex set of \( T \),
- all bounded faces of \( T \) are triangles,
- the union of the bounded triangular faces forms the convex hull of \( S \).
Constrained Triangulation

Definition 207 (Constrained triangulation)

Let $S = \{P_1, P_2, \ldots, P_k\}$ be a set of $k$ points in $\mathbb{R}^2$, and $E$ be a set of line segments that link points of $S$ and that do not intersect pairwise except at common end points. A structure $T$ is called a **constrained triangulation** of $S$ if $T$ is (the straight-line embedding of) a connected planar graph such that
Definition 207 (Constrained triangulation)

Let $S = \{P_1, P_2, \ldots, P_k\}$ be a set of $k$ points in $\mathbb{R}^2$, and $E$ be a set of line segments that link points of $S$ and that do not intersect pairwise except at common end points. A structure $T$ is called a \textit{constrained triangulation} of $S$ if $T$ is (the straight-line embedding of) a connected planar graph such that

1. $S$ is (the embedding of) the vertex set of $T$,
2. all bounded faces of $T$ are triangles,
3. the union of the bounded triangular faces forms the convex hull of $S$,
4. all segments of $E$ are edges of $T$. 
Constrained Triangulation

**Definition 207 (Constrained triangulation)**

Let $S = \{P_1, P_2, \ldots, P_k\}$ be a set of $k$ points in $\mathbb{R}^2$, and $E$ be a set of line segments that link points of $S$ and that do not intersect pairwise except at common end points. A structure $T$ is called a *constrained triangulation* of $S$ if $T$ is (the straight-line embedding of) a connected planar graph such that

- $S$ is (the embedding of) the vertex set of $T$,
- all bounded faces of $T$ are triangles,
- the union of the bounded triangular faces forms the convex hull of $S$, 

![Constrained Triangulation Diagram](image)
Definition 207 (Constrained triangulation)

Let $S = \{P_1, P_2, \ldots, P_k\}$ be a set of $k$ points in $\mathbb{R}^2$, and $E$ be a set of line segments that link points of $S$ and that do not intersect pairwise except at common end points. A structure $T$ is called a constrained triangulation of $S$ if $T$ is (the straight-line embedding of) a connected planar graph such that:

- $S$ is (the embedding of) the vertex set of $T$,
- all bounded faces of $T$ are triangles,
- the union of the bounded triangular faces forms the convex hull of $S$,
- all segments of $E$ are edges of $T$. 

![Constrained Triangulation Diagram](image-url)
Transformations

- Linear Transformations
- Classification of Transformations
- Coordinate Transformations in $\mathbb{R}^2$
- Coordinate Transformations in $\mathbb{R}^3$
- Transformation of Coordinate Systems
- Applications of Coordinate (System) Transformations
- Rotations Revisited
- Projections
Transformations
- Linear Transformations
  - Linear Transformations and Matrices
  - Linear Transformations and Determinants
  - Image and Kernel
  - Linear Transformations and Dot Product
  - Linear Transformations and Cross Products
- Classification of Transformations
- Coordinate Transformations in $\mathbb{R}^2$
- Coordinate Transformations in $\mathbb{R}^3$
- Transformation of Coordinate Systems
- Applications of Coordinate (System) Transformations
- Rotations Revisited
- Projections
Linear Transformations

**Definition 208 (Linear transformation, Dt.: lineare Abbildung)**

Let $V, W$ be vector spaces over $\mathbb{R}$. A transformation $g: V \rightarrow W$ is called a **linear transformation**

1. $g(v_1 + v_2) = g(v_1) + g(v_2)$ for all $v_1, v_2 \in V$.
2. $g(\lambda v) = \lambda g(v)$ for all $v \in V$, $\lambda \in \mathbb{R}$.

E.g., $V := \mathbb{R}^n$ and $W := \mathbb{R}^m$ for some $m, n \in \mathbb{N}$.

**Lemma 209**

A linear transformation maps every line to a line, the coordinate origin of $V$ to the coordinate origin of $W$.

**Sketch of Proof:**

A line \( \{p + \lambda v : \lambda \in \mathbb{R}\} \) is mapped as follows:

\[
g(\{p + \lambda v : \lambda \in \mathbb{R}\}) = \{g(p) + \lambda g(v) : \lambda \in \mathbb{R}\}
\]
Linear Transformations

**Definition 208 (Linear transformation, Dt.: lineare Abbildung)**

Let $V, W$ be vector spaces over $\mathbb{R}$. A transformation $g : V \to W$ is called a linear transformation if

1. $g(v_1 + v_2) = g(v_1) + g(v_2)$ \quad $\forall v_1, v_2 \in V,$
Linear Transformations

**Definition 208 (Linear transformation, Dt.: lineare Abbildung)**

Let $V$, $W$ be vector spaces over $\mathbb{R}$. A transformation $g : V \to W$ is called a *linear transformation*

1. $g(v_1 + v_2) = g(v_1) + g(v_2) \quad \forall v_1, v_2 \in V,$
2. $g(\lambda v) = \lambda g(v) \quad \forall v \in V, \forall \lambda \in \mathbb{R}.$
Linear Transformations

**Definition 208 (Linear transformation, Dt.: lineare Abbildung)**

Let $V, W$ be vector spaces over $\mathbb{R}$. A transformation $g : V \rightarrow W$ is called a *linear transformation*

1. $g(v_1 + v_2) = g(v_1) + g(v_2)$ \quad $\forall v_1, v_2 \in V$,
2. $g(\lambda v) = \lambda g(v)$ \quad $\forall v \in V, \forall \lambda \in \mathbb{R}$.

- E.g., $V := \mathbb{R}^n$ and $W := \mathbb{R}^m$ for some $m, n \in \mathbb{N}$. 
Linear Transformations

Definition 208 (Linear transformation, Dt.: lineare Abbildung)
Let $V, W$ be vector spaces over $\mathbb{R}$. A transformation $g : V \rightarrow W$ is called a linear transformation if:

1. $g(v_1 + v_2) = g(v_1) + g(v_2)$ for all $v_1, v_2 \in V$,
2. $g(\lambda v) = \lambda g(v)$ for all $v \in V, \lambda \in \mathbb{R}$.

Examples:
- $V := \mathbb{R}^n$ and $W := \mathbb{R}^m$ for some $m, n \in \mathbb{N}$.

Lemma 209
A linear transformation maps
- every line to a line,
- the coordinate origin of $V$ to the coordinate origin of $W$. 

Sketch of Proof:
A line $\{p + \lambda v : \lambda \in \mathbb{R}\}$ is mapped as follows:

$$g(\{p + \lambda v : \lambda \in \mathbb{R}\}) = \{g(p + \lambda v) : \lambda \in \mathbb{R}\} = \{g(p) + \lambda g(v) : \lambda \in \mathbb{R}\}$$
**Definition 208 (Linear transformation, Dt.: lineare Abbildung)**

Let $V$, $W$ be vector spaces over $\mathbb{R}$. A transformation $g : V \rightarrow W$ is called a linear transformation if:

1. $g(v_1 + v_2) = g(v_1) + g(v_2)$ for all $v_1, v_2 \in V$,
2. $g(\lambda v) = \lambda g(v)$ for all $v \in V$ and $\lambda \in \mathbb{R}$.

**Example:** Let $V := \mathbb{R}^n$ and $W := \mathbb{R}^m$ for some $m, n \in \mathbb{N}$.

**Lemma 209**

A linear transformation maps

- every line to a line,
- the coordinate origin of $V$ to the coordinate origin of $W$.

**Sketch of Proof:** A line $\{p + \lambda v : \lambda \in \mathbb{R}\}$ is mapped as follows:

$$g(\{p + \lambda v : \lambda \in \mathbb{R}\}) = \{g(p + \lambda v) : \lambda \in \mathbb{R}\} = \{g(p) + \lambda g(v) : \lambda \in \mathbb{R}\}$$
Hence, a transformation from $V$ to $W$ is linear if and only if

1. every regular grid in $V$ gets mapped to a regular grid in $W$,
2. the coordinate origin of $V$ lands on the coordinate origin of $W$. 
Hence, a transformation from $V$ to $W$ is linear if and only if

1. every regular grid in $V$ gets mapped to a regular grid in $W$,
2. the coordinate origin of $V$ lands on the coordinate origin of $W$. 
Hence, a transformation from $V$ to $W$ is linear if and only if

1. every regular grid in $V$ gets mapped to a regular grid in $W$,
2. the coordinate origin of $V$ lands on the coordinate origin of $W$. 
Linear Transformations

Theorem 210

Let $e_1, \ldots, e_n$ be a basis of $V$, and $e'_1, \ldots, e'_m$ be a basis of $W$. A linear transformation $g: V \rightarrow W$ is uniquely determined by the images of the basis vectors $e_j$ relative to $e'_i$. It has a corresponding $m \times n$ transformation matrix whose $n$ columns are given by the images of the basis vectors $e_1, \ldots, e_n$. 

Sketch of Proof:

For $v := \sum_{j=1}^n v_j e_j$ and $w := \sum_{i=1}^m w_i e'_i$, with $w = g(v)$, we get

$$w = g(v) = g\left(\sum_{j=1}^n v_j e_j\right) = \sum_{j=1}^n v_j g(e_j) = \sum_{j=1}^n v_j \left(\sum_{i=1}^m a_{ij} e'_i\right) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} v_j\right) e'_i = A v,$$

where $A = [a_{ij}]_{m \times n}$, and $a_{ij}$ equals the $i$-th coordinate of $g(e_j)$. 
Linear Transformations

**Theorem 210**

Let $e_1, \ldots, e_n$ be a basis of $V$, and $e'_1, \ldots, e'_m$ be a basis of $W$. A linear transformation $g: V \to W$ is uniquely determined by the images of the basis vectors $e_j$ relative to $e'_i$. It has a corresponding $m \times n$ transformation matrix whose $n$ columns are given by the images of the basis vectors $e_1, \ldots, e_n$.

**Sketch of Proof:** For $v := \sum_{j=1}^n v_j e_j$ and $w := \sum_{i=1}^m w_i e'_i$, with $w = g(v)$, we get

$$w = g(v) = g\left(\sum_{j=1}^n v_j e_j\right) = \sum_{j=1}^n v_j g(e_j) = \sum_{j=1}^n v_j \left(\sum_{i=1}^m a_{ij} e'_i\right) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} v_j\right) e'_i$$

$$= A v,$$

where $A = [a_{ij}]_{i=1,j=1}^{m,n}$ and $a_{ij}$ equals the $i$-th coordinate of $g(e_j)$. \hfill \square
Linear Transformations

- Suppose that we know that a linear transformation $g$ maps $e_1$ of $\mathbb{R}^2$ to the vector \[ \begin{pmatrix} 2 \\ 0 \end{pmatrix} \] of $\mathbb{R}^2$, and $e_2$ to the vector \[ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \].

Thus, $g$ has the following matrix:

\[ \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \]
Linear Transformations

- Suppose that we know that a linear transformation $g$ maps $e_1$ of $\mathbb{R}^2$ to the vector $(2, 0)$ of $\mathbb{R}^2$, and $e_2$ to the vector $(1, 1)$.

- The transformation $g$ maps the point $(1, 2)$ to the point $(4, 2)$:

$$g((1, 2)) = g(1 \cdot (2, 0) + 2 \cdot (1, 1)) = 1 \cdot g((2, 0)) + 2 \cdot g((1, 1)) = (2, 0) + (1, 1) = (3, 1)$$
Linear Transformations

- Suppose that we know that a linear transformation $g$ maps $e_1$ of $\mathbb{R}^2$ to the vector $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ of $\mathbb{R}^2$, and $e_2$ to the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

- The transformation $g$ maps the point $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ to the point $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$:

$$g \left( \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right) = g \left( 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = 1 \cdot g \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) + 2 \cdot g \left( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

$$= 1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$
Linear Transformations

- Suppose that we know that a linear transformation $g$ maps $e_1$ of $\mathbb{R}^2$ to the vector $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ of $\mathbb{R}^2$, and $e_2$ to the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

- The transformation $g$ maps the point $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ to the point $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$:

  $$
g \left( \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right) = g \left( 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = 1 \cdot g \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) + 2 \cdot g \left( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)
  = 1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}
  = \begin{pmatrix} 4 \\ 2 \end{pmatrix}
  = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}
$$

Thus, $g$ has the following matrix:

$$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$
Linear Transformations

- Suppose that we know that a linear transformation $g$ maps $e_1$ of $\mathbb{R}^2$ to the vector $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ of $\mathbb{R}^2$, and $e_2$ to the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

- The transformation $g$ maps the point $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ to the point $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$:

  \[
g \left( \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right) = g \left( 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = 1 \cdot g \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) + 2 \cdot g \left( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)
  \]

  \[
  = 1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}
  \]

  \[
  = \begin{pmatrix} 4 \\ 2 \end{pmatrix}
  \]

  \[
  = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}
  \]

- Thus, $g$ has the following matrix:

  $\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$
Linear Transformations

Theorem 211
Every linear transformation has a corresponding matrix transformation, and it maps a linear combination of vectors to the linear combination of the images of the vectors.

Sample linear transformations in $\mathbb{R}^2$: rotation about origin, stretching, reflection (about coordinate axis or origin), shear transformation.

Note: Translation is not linear!

Lemma 212
If a linear transformation has an inverse transformation then the inverse transformation is also linear.

Lemma 213
If a linear transformation $g$ has an inverse transformation then the matrix which corresponds to $g$ is invertible.
Linear Transformations

Theorem 211
Every linear transformation has a corresponding matrix transformation, and it maps a linear combination of vectors to the linear combination of the images of the vectors.

- Sample linear transformations in $\mathbb{R}^2$: rotation about origin, stretching, reflection (about coordinate axis or origin), shear transformation.
- Note: Translation is not linear!
Linear Transformations

**Theorem 211**
Every linear transformation has a corresponding matrix transformation, and it maps a linear combination of vectors to the linear combination of the images of the vectors.

- Sample linear transformations in $\mathbb{R}^2$: rotation about origin, stretching, reflection (about coordinate axis or origin), shear transformation.
- Note: Translation is not linear!

**Lemma 212**
If a linear transformation has an inverse transformation then the inverse transformation is also linear.
Linear Transformations

Theorem 211

Every linear transformation has a corresponding matrix transformation, and it maps a linear combination of vectors to the linear combination of the images of the vectors.

- Sample linear transformations in $\mathbb{R}^2$: rotation about origin, stretching, reflection (about coordinate axis or origin), shear transformation.
- Note: Translation is not linear!

Lemma 212

If a linear transformation has an inverse transformation then the inverse transformation is also linear.

Lemma 213

If a linear transformation $g$ has an inverse transformation then the matrix which corresponds to $g$ is invertible.
Composition of Linear Transformations

Definition 214 (Composition, Dt.: Zusammensetzung)

Consider two linear transformations \( g: U \to V \) and \( h: V \to W \). The composition \( h \circ g \) is a transformation from \( U \) to \( W \) such that every \( u \in U \) is mapped to \( h(g(u)) \in W \).
Composition of Linear Transformations

Definition 214 (Composition, Dt.: Zusammensetzung)
Consider two linear transformations $g: U \to V$ and $h: V \to W$. The composition $h \circ g$ is a transformation from $U$ to $W$ such that every $u \in U$ is mapped to $h(g(u)) \in W$.

Warning
There is absolutely no consensus in the literature on whether $(h \circ g)(x)$ shall mean $h(g(x))$ or $g(h(x))$!
Composition of Linear Transformations

Definition 214 (Composition, Dt.: Zusammensetzung)
Consider two linear transformations $g: U \rightarrow V$ and $h: V \rightarrow W$. The composition $h \circ g$ is a transformation from $U$ to $W$ such that every $u \in U$ is mapped to $h(g(u)) \in W$.

Warning
There is absolutely no consensus in the literature on whether $(h \circ g)(x)$ shall mean $h(g(x))$ or $g(h(x))$!

Lemma 215
The composition of two linear transformations is a linear transformation.
Combining Matrix Transformations

Suppose that $p'$ is obtained by applying the matrix transformation $T_1$ to $p$, and $p''$ is obtained from $p'$ via $T_2$, and so on till $p^{(n)}$:

\[
\begin{align*}
(x') &= T_1 \cdot (x) \\
(y') &= \quad \quad T_2 \cdot (x') \\
(x'') &= \quad \quad T_3 \cdot (x'') \\
(y'') &= \quad \quad \vdots \\
(x^{(n)}) &= T_n \cdot (x^{(n-1)}) \\
(y^{(n)}) &= \quad \quad \vdots \\
\end{align*}
\]

Note the order of the matrix multiplications!

Recall that matrix multiplication is associative but not commutative!
Combining Matrix Transformations

Suppose that $p'$ is obtained by applying the matrix transformation $T_1$ to $p$, and $p''$ is obtained from $p'$ via $T_2$, and so on till $p^{(n)}$:

\[
\begin{pmatrix} x' \\ y' \end{pmatrix} = T_1 \cdot \begin{pmatrix} x \\ y \end{pmatrix} \quad \begin{pmatrix} x'' \\ y'' \end{pmatrix} = T_2 \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \quad \ldots \quad \begin{pmatrix} x^{(n)} \\ y^{(n)} \end{pmatrix} = T_n \cdot \begin{pmatrix} x^{(n-1)} \\ y^{(n-1)} \end{pmatrix}.
\]

Then the dependence of $p^{(n)}$ on $p$ can be expressed as

\[
\begin{pmatrix} x^{(n)} \\ y^{(n)} \end{pmatrix} = T_n \cdot \left( T_{n-1} \cdot \left( \ldots \left( T_2 \cdot \left( T_1 \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right) \right) \right) \right).
\]
Combining Matrix Transformations

Suppose that $p'$ is obtained by applying the matrix transformation $T_1$ to $p$, and $p''$ is obtained from $p'$ via $T_2$, and so on till $p^{(n)}$:

\[
\begin{align*}
\begin{pmatrix} x' \\ y' \end{pmatrix} &= T_1 \cdot \begin{pmatrix} x \\ y \end{pmatrix} \\
\begin{pmatrix} x'' \\ y'' \end{pmatrix} &= T_2 \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \\
& \quad \cdots \\
\begin{pmatrix} x^{(n)} \\ y^{(n)} \end{pmatrix} &= T_n \cdot \begin{pmatrix} x^{(n-1)} \\ y^{(n-1)} \end{pmatrix}.
\end{align*}
\]

Then the dependence of $p^{(n)}$ on $p$ can be expressed as

\[
\begin{align*}
\begin{pmatrix} x^{(n)} \\ y^{(n)} \end{pmatrix} &= T_n \cdot \left( T_{n-1} \cdot \left( \cdots \left( T_2 \cdot \left( T_1 \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right) \right) \right) \right) = \\
&= (T_n \cdot T_{n-1} \cdot \cdots \cdot T_2 \cdot T_1) \cdot \begin{pmatrix} x \\ y \end{pmatrix}.
\end{align*}
\]
Combining Matrix Transformations

Suppose that $p'$ is obtained by applying the matrix transformation $T_1$ to $p$, and $p''$ is obtained from $p'$ via $T_2$, and so on till $p^{(n)}$:

$$
\begin{pmatrix} x' \\ y' \end{pmatrix} = T_1 \cdot \begin{pmatrix} x \\ y \end{pmatrix} \quad \begin{pmatrix} x'' \\ y'' \end{pmatrix} = T_2 \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \quad \ldots \quad \begin{pmatrix} x^{(n)} \\ y^{(n)} \end{pmatrix} = T_n \cdot \begin{pmatrix} x^{(n-1)} \\ y^{(n-1)} \end{pmatrix}.
$$

Then the dependence of $p^{(n)}$ on $p$ can be expressed as

$$
\begin{pmatrix} x^{(n)} \\ y^{(n)} \end{pmatrix} = T_n \cdot \left( T_{n-1} \cdot \left( \ldots (T_2 \cdot (T_1 \cdot \begin{pmatrix} x \\ y \end{pmatrix})))) \right) =
$$

$$
= (T_n \cdot T_{n-1} \cdot \ldots \cdot T_2 \cdot T_1) \cdot \begin{pmatrix} x \\ y \end{pmatrix} =
$$

$$
= T \cdot \begin{pmatrix} x \\ y \end{pmatrix},
$$

where $T := T_n \cdot T_{n-1} \cdot \ldots \cdot T_2 \cdot T_1$. 

Caveats

Note the order of the matrix multiplications!
Recall that matrix multiplication is associative but not commutative!
Combining Matrix Transformations

- Suppose that \( p' \) is obtained by applying the matrix transformation \( T_1 \) to \( p \), and \( p'' \) is obtained from \( p' \) via \( T_2 \), and so on till \( p^{(n)} \):

\[
\begin{pmatrix}
 x' \\
 y'
\end{pmatrix} = T_1 \cdot \begin{pmatrix}
 x \\
 y
\end{pmatrix} \\
\begin{pmatrix}
 x'' \\
 y''
\end{pmatrix} = T_2 \cdot \begin{pmatrix}
 x' \\
 y'
\end{pmatrix} \\
\ldots \\
\begin{pmatrix}
 x^{(n)} \\
 y^{(n)}
\end{pmatrix} = T_n \cdot \begin{pmatrix}
 x^{(n-1)} \\
 y^{(n-1)}
\end{pmatrix}.
\]

Then the dependence of \( p^{(n)} \) on \( p \) can be expressed as

\[
\begin{pmatrix}
 x^{(n)} \\
 y^{(n)}
\end{pmatrix} = T_n \cdot \left( T_{n-1} \cdot \left( \ldots \left( T_2 \cdot \left( T_1 \cdot \begin{pmatrix}
 x \\
 y
\end{pmatrix} \right) \right) \right) \right) =
\]

\[
= \left( T_n \cdot T_{n-1} \cdot \ldots \cdot T_2 \cdot T_1 \right) \cdot \begin{pmatrix}
 x \\
 y
\end{pmatrix} =
\]

\[
= T \cdot \begin{pmatrix}
 x \\
 y
\end{pmatrix},
\]

where \( T := T_n \cdot T_{n-1} \cdot \ldots \cdot T_2 \cdot T_1 \).

Caveats

- Note the order of the matrix multiplications!
- Recall that matrix multiplication is associative but not commutative!
Order of Transformations Matters

- $T$: Translate by $(5, 0)$;  
- $R$: Rotate about origin by $\pi/4$.
Linear Transformations and Linear Equations

- So far we were concerned with determining \( g(x) \) for a linear transformation \( g \) and a vector \( x \), i.e., the image vector of \( x \) under the linear transformation \( g \).
- If \( A \) is the matrix that represents \( g \) then, via matrix multiplication,

\[
g(x) = Ax.
\]
Linear Transformations and Linear Equations

- So far we were concerned with determining $g(x)$ for a linear transformation $g$ and a vector $x$, i.e., the image vector of $x$ under the linear transformation $g$.

- If $A$ is the matrix that represents $g$ then, via matrix multiplication,

$$g(x) = Ax.$$  

- However, we can also revert the question and specify the image vector $b$, and seek the vector $x$ which gets mapped to $b$ by $g$.

- Then the answer is provided by solving the following system of linear equations:

$$Ax = b.$$
Geometric Interpretation of the Determinant of a Transformation Matrix

Consider the linear transformation $g$ with transformation matrix

$$T := \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}.$$
Geometric Interpretation of the Determinant of a Transformation Matrix

- Consider the linear transformation $g$ with transformation matrix

$$T := \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}.$$  

- Remember that its columns represent the images of the unit vectors.

![Diagram showing the transformation of a unit square to a parallelogram of twice the area](image-url)
Consider the linear transformation $g$ with transformation matrix

$$T := \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}.$$ 

Remember that its columns represent the images of the unit vectors. Hence, the unit square gets mapped by $g$ to a parallelogram of twice the area.
Consider the linear transformation $g$ with transformation matrix

\[
T := \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}.
\]

Remember that its columns represent the images of the unit vectors.

Hence, the unit square gets mapped by $g$ to a parallelogram of twice the area.

Now note that $\det(T) = 2$. 

\[\begin{array}{c}
\begin{array}{ccc}
\hline
y
\
\hline
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{ccc}
\hline
x
\
\hline
\end{array}
\end{array}
\]
Geometric Interpretation of the Determinant of a Transformation Matrix

• Now consider the linear transformation $g$ with transformation matrix

\[
T := \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}.
\]

Remember that its columns represent the images of the unit vectors. Hence, the unit square gets mapped by $g$ to a parallelogram of twice the area.

Now note that $\det(T) = -2$, and that $g$ changed the orientation/handedness of the unit vectors.
Now consider the linear transformation $g$ with transformation matrix

$$T := \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}.$$  

Remember that its columns represent the images of the unit vectors.  

Hence, the unit square gets mapped by $g$ to a parallelogram of twice the area.  

Now note that $\det(T) = -2$, and that $g$ changed the orientation/handedness of the unit vectors.
Now consider the linear transformation $g$ with transformation matrix

$$T := \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}.$$

Remember that its columns represent the images of the unit vectors. Hence, the unit square gets mapped by $g$ to a parallelogram of twice the area.
Now consider the linear transformation $g$ with transformation matrix

$$
T := \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}.
$$

Remember that its columns represent the images of the unit vectors. Hence, the unit square gets mapped by $g$ to a parallelogram of twice the area.

Now note that $\det(T) = -2$, and that $g$ changed the orientation/handedness of the unit vectors.
Theorem 216

The absolute value of the determinant of a (square) transformation matrix $A$ gives the scale factor for the area/volume of the image of the unit (hyper-)cube. If $\det(A)$ is negative then the handedness of the unit vectors has changed, i.e., the orientation of space has been inverted.
Geometric Interpretation of the Determinant of a Transformation Matrix

**Theorem 216**

The absolute value of the determinant of a (square) transformation matrix $A$ gives the scale factor for the area/volume of the image of the unit (hyper-)cube. If $\det(A)$ is negative then the handedness of the unit vectors has changed, i.e., the orientation of space has been inverted.

*Sketch of Proof:* Theorem 109 settles this claim for $2 \times 2$ matrices.
Geometric Interpretation of the Determinant of a Transformation Matrix

**Theorem 216**

The absolute value of the determinant of a (square) transformation matrix \( A \) gives the scale factor for the area/volume of the image of the unit (hyper-)cube. If \( \det(A) \) is negative then the handedness of the unit vectors has changed, i.e., the orientation of space has been inverted.

**Sketch of Proof:** Theorem 109 settles this claim for \( 2 \times 2 \) matrices. If the matrix \( A \) is a diagonal matrix then the \( i \)-th side of the unit (hyper-)cube gets scaled by the factor \( a_{ii} \). Hence, its volume changes by the factor \( \prod_{i=1}^{n} a_{ii} = \det(A) \).

![Graphical representation of the transformation matrix](image-url)
Geometric Interpretation of the Determinant of a Transformation Matrix

**Theorem 216**

The absolute value of the determinant of a (square) transformation matrix $A$ gives the scale factor for the area/volume of the image of the unit (hyper-)cube. If $\det(A)$ is negative then the handedness of the unit vectors has changed, i.e., the orientation of space has been inverted.

**Sketch of Proof:** Theorem 109 settles this claim for $2 \times 2$ matrices. If the matrix $A$ is a diagonal matrix then the $i$-th side of the unit (hyper-)cube gets scaled by the factor $a_{ii}$. Hence, its volume changes by the factor $\prod_{i=1}^{n} a_{ii} = \det(A)$. If $A$ is an upper-triangular matrix then we get a shear transformation, but its determinant still equals $\prod_{i=1}^{n} a_{ii}$. And the shear does not change the volume!
Geometric Interpretation of the Determinant of a Transformation Matrix

- Recall Theorem 104: A square matrix $A$ is invertible if and only if $\det(A) \neq 0$. 
Recall Theorem 104: A square matrix $A$ is invertible if and only if $\det(A) \neq 0$.

Now regard the square matrix $A$ as the $n \times n$ matrix of a linear transformation $g$ (of $\mathbb{R}^n$). If $A$ is invertible then, for every vector $u \in \mathbb{R}^n$,

$$A^{-1}w = u \quad \text{for} \quad w := Au.$$
Recall Theorem 104: A square matrix $A$ is invertible if and only if $\det(A) \neq 0$.

Now regard the square matrix $A$ as the $n \times n$ matrix of a linear transformation $g$ (of $\mathbb{R}^n$). If $A$ is invertible then, for every vector $u \in \mathbb{R}^n$,

$$A^{-1}w = u \quad \text{for} \quad w := Au.$$ 

Of course, mapping the image $g(u) = w$ of $u$ back to $u$ can only work if and only if $g$ maps $\mathbb{R}^n$ to all of $\mathbb{R}^n$ rather than to some subspace of $\mathbb{R}^n$, like a line or (hyper-)plane. (Otherwise, we would have to restore $\mathbb{R}^n$ from, say, a line!)
Geometric Interpretation of the Determinant of a Transformation Matrix

- Recall Theorem 104: A square matrix $A$ is invertible if and only if $\det(A) \neq 0$.

- Now regard the square matrix $A$ as the $n \times n$ matrix of a linear transformation $g$ (of $\mathbb{R}^n$). If $A$ is invertible then, for every vector $u \in \mathbb{R}^n$,
  \[ A^{-1}w = u \quad \text{for} \quad w := Au. \]

- Of course, mapping the image $g(u) = w$ of $u$ back to $u$ can only work if and only if $g$ maps $\mathbb{R}^n$ to all of $\mathbb{R}^n$ rather than to some subspace of $\mathbb{R}^n$, like a line or (hyper-)plane. (Otherwise, we would have to restore $\mathbb{R}^n$ from, say, a line!)

- This bijection from $\mathbb{R}^n$ to all of $\mathbb{R}^n$ happens precisely if $g$ maps no basis vector of $\mathbb{R}^n$ to a linear combination of images of other basis vectors.
Geometric Interpretation of the Determinant of a Transformation Matrix

- Recall Theorem 104: A square matrix $A$ is invertible if and only if $\det(A) \neq 0$.
- Now regard the square matrix $A$ as the $n \times n$ matrix of a linear transformation $g$ (of $\mathbb{R}^n$). If $A$ is invertible then, for every vector $u \in \mathbb{R}^n$,
  $$A^{-1}w = u \quad \text{for} \quad w := Au.$$  
- Of course, mapping the image $g(u) = w$ of $u$ back to $u$ can only work if and only if $g$ maps $\mathbb{R}^n$ to all of $\mathbb{R}^n$ rather than to some subspace of $\mathbb{R}^n$, like a line or (hyper-)plane. (Otherwise, we would have to restore $\mathbb{R}^n$ from, say, a line!)
- This bijection from $\mathbb{R}^n$ to all of $\mathbb{R}^n$ happens precisely if $g$ maps no basis vector of $\mathbb{R}^n$ to a linear combination of images of other basis vectors.
- And precisely in this case the unit (hyper-)cube transformed by $g$ has a non-zero volume.
Geometric Interpretation of the Determinant of a Transformation Matrix

- Recall Theorem 104: A square matrix $A$ is invertible if and only if $\det(A) \neq 0$.
- Now regard the square matrix $A$ as the $n \times n$ matrix of a linear transformation $g$ (of $\mathbb{R}^n$). If $A$ is invertible then, for every vector $u \in \mathbb{R}^n$,
  \[ A^{-1}w = u \quad \text{for} \quad w := Au. \]
- Of course, mapping the image $g(u) = w$ of $u$ back to $u$ can only work if and only if $g$ maps $\mathbb{R}^n$ to all of $\mathbb{R}^n$ rather than to some subspace of $\mathbb{R}^n$, like a line or (hyper-)plane. (Otherwise, we would have to restore $\mathbb{R}^n$ from, say, a line!)
- This bijection from $\mathbb{R}^n$ to all of $\mathbb{R}^n$ happens precisely if $g$ maps no basis vector of $\mathbb{R}^n$ to a linear combination of images of other basis vectors.
- And precisely in this case the unit (hyper-)cube transformed by $g$ has a non-zero volume.
- Now recall that the volume of the transformed (hyper-)cube is given by $\det(A)$.
- We understand that $A$ is invertible if and only if $\det(A) \neq 0$. 
Recall Theorem 104: A square matrix $A$ is invertible if and only if $\det(A) \neq 0$.

Now regard the square matrix $A$ as the $n \times n$ matrix of a linear transformation $g$ (of $\mathbb{R}^n$). If $A$ is invertible then, for every vector $u \in \mathbb{R}^n$,

$$A^{-1}w = u \quad \text{for} \quad w := Au.$$ 

Of course, mapping the image $g(u) = w$ of $u$ back to $u$ can only work if and only if $g$ maps $\mathbb{R}^n$ to all of $\mathbb{R}^n$ rather than to some subspace of $\mathbb{R}^n$, like a line or (hyper-)plane. (Otherwise, we would have to restore $\mathbb{R}^n$ from, say, a line!)

This bijection from $\mathbb{R}^n$ to all of $\mathbb{R}^n$ happens precisely if $g$ maps no basis vector of $\mathbb{R}^n$ to a linear combination of images of other basis vectors.

And precisely in this case the unit (hyper-)cube transformed by $g$ has a non-zero volume.

Now recall that the volume of the transformed (hyper-)cube is given by $\det(A)$.

We understand that $A$ is invertible if and only if $\det(A) \neq 0$.

If $\det(A) = 0$ then a solution to the linear equation $Au = b$ exists if and only if $b$ lies within the subspace $g(\mathbb{R}^n)$ of $\mathbb{R}^n$. 

© M. Held (Univ. Salzburg)
Definition 217 (Image, Dt.: Bild)

The *image* (or *column space*) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $Au$ for $u \in \mathbb{R}^n$, i.e., it equals $g(\mathbb{R}^n) \subset \mathbb{R}^m$. 

A solution to the linear equation $Au = b$ exists if and only if $b$ lies within the image of $A$.

Recall Definition 86: The rank of an $m \times n$ matrix $A$ is the number of linearly independent columns of $A$.

1. If $g$ squashes $\mathbb{R}^n$ to a line then the rank of $A$ equals 1.
2. If $g$ squashes $\mathbb{R}^n$ to a plane then the rank of $A$ equals 2.
3. ... 

Hence, the rank of $A$ equals the dimension of the image of $A$. Note that the image $g(\mathbb{R}^n)$ forms a subspace of $\mathbb{R}^m$. 
Definition 217 (Image, Dt.: Bild)

The *image* (or *column space*) of an \( m \times n \) matrix \( A \) (of a linear transformation \( g \)) is the set of all vectors \( Au \) for \( u \in \mathbb{R}^n \), i.e., it equals \( g(\mathbb{R}^n) \subset \mathbb{R}^m \).

- A solution to the linear equation \( Au = b \) exists if and only if \( b \) lies within the image of \( A \).
Geometric Interpretation of the Rank of a Transformation Matrix

Definition 217 (Image, Dt.: Bild)
The *image* (or *column space*) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $Au$ for $u \in \mathbb{R}^n$, i.e., it equals $g(\mathbb{R}^n) \subset \mathbb{R}^m$.

- A solution to the linear equation $Au = b$ exists if and only if $b$ lies within the image of $A$.
- Recall Definition 86: The rank of an $m \times n$ matrix $A$ is the number of linearly independent columns of $A$. 

\[ \text{If } g \text{ squashes } \mathbb{R}^n \text{ to a line then the rank of } A \text{ equals } 1. \]
\[ \text{If } g \text{ squashes } \mathbb{R}^n \text{ to a plane then the rank of } A \text{ equals } 2. \]
\[ \text{... } \]
\[ \text{Hence, the rank of } A \text{ equals the dimension of the image of } A. \]
\[ \text{Note that the image } g(\mathbb{R}^n) \text{ forms a subspace of } \mathbb{R}^m. \]
### Geometric Interpretation of the Rank of a Transformation Matrix

**Definition 217 (Image, Dt.: Bild)**

The *image* (or *column space*) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $Au$ for $u \in \mathbb{R}^n$, i.e., it equals $g(\mathbb{R}^n) \subset \mathbb{R}^m$.

- A solution to the linear equation $Au = b$ exists if and only if $b$ lies within the image of $A$.
- Recall Definition 86: The rank of an $m \times n$ matrix $A$ is the number of linearly independent columns of $A$.
  1. If $g$ squashes $\mathbb{R}^n$ to a line then the rank of $A$ equals 1.
Geometric Interpretation of the Rank of a Transformation Matrix

**Definition 217 (Image, Dt.: Bild)**

The *image* (or *column space*) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $Au$ for $u \in \mathbb{R}^n$, i.e., it equals $g(\mathbb{R}^n) \subset \mathbb{R}^m$.

- A solution to the linear equation $Au = b$ exists if and only if $b$ lies within the image of $A$.
- Recall Definition 86: The rank of an $m \times n$ matrix $A$ is the number of linearly independent columns of $A$.
  1. If $g$ squashes $\mathbb{R}^n$ to a line then the rank of $A$ equals 1.
  2. If $g$ squashes $\mathbb{R}^n$ to a plane then the rank of $A$ equals 2.
  3. ...
Geometric Interpretation of the Rank of a Transformation Matrix

**Definition 217 (Image, Dt.: Bild)**

The *image* (or *column space*) of an \( m \times n \) matrix \( A \) (of a linear transformation \( g \)) is the set of all vectors \( Au \) for \( u \in \mathbb{R}^n \), i.e., it equals \( g(\mathbb{R}^n) \subset \mathbb{R}^m \).

- A solution to the linear equation \( Au = b \) exists if and only if \( b \) lies within the image of \( A \).
- Recall Definition 86: The rank of an \( m \times n \) matrix \( A \) is the number of linearly independent columns of \( A \).
  1. If \( g \) squashes \( \mathbb{R}^n \) to a line then the rank of \( A \) equals 1.
  2. If \( g \) squashes \( \mathbb{R}^n \) to a plane then the rank of \( A \) equals 2.
  3. \( \ldots \)
- Hence, the rank of \( A \) equals the dimension of the image of \( A \).
A solution to the linear equation $Au = b$ exists if and only if $b$ lies within the image of $A$.

Recall Definition 86: The rank of an $m \times n$ matrix $A$ is the number of linearly independent columns of $A$.

1. If $g$ squashes $\mathbb{R}^n$ to a line then the rank of $A$ equals 1.
2. If $g$ squashes $\mathbb{R}^n$ to a plane then the rank of $A$ equals 2.
3. ... 

Hence, the rank of $A$ equals the dimension of the image of $A$.

Note that the image $g(\mathbb{R}^n)$ forms a subspace of $\mathbb{R}^m$. 

---

**Definition 217 (Image, Dt.: Bild)**

The *image* (or *column space*) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $Au$ for $u \in \mathbb{R}^n$, i.e., it equals $g(\mathbb{R}^n) \subset \mathbb{R}^m$. 

---

© M. Held (Univ. Salzburg)

Geometrisches Rechnen (WS 2020/21)
Definition 218 (Kernel, Dt.: Kern)

The *kernel* (or *null space*) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $u \in \mathbb{R}^n$ which get mapped by $g$ to the zero vector of $\mathbb{R}^m$. 

Hence, if $u_0 \in \mathbb{R}^n$ is a solution of $A u = b$ then $u_0 + w$ is also a solution of $A u = b$ for all $w$ in the kernel of $A$. 

The kernel of an $m \times n$ matrix forms a subspace of $\mathbb{R}^n$. 

Definition 219 (Corank, Dt.: Defekt)

The *corank* (nullity) of an $m \times n$ matrix $A$, denoted by $\text{corank}(A)$, is the dimension of the kernel of $A$. 

Theorem 220 (Rank-nullity theorem, Dt.: Rangsatz, Dimensionssatz)

Consider an $m \times n$ matrix $A$. Then $\text{rank}(A) + \text{corank}(A) = n$. 

---

© M. Held (Univ. Salzburg)
Definition 218 (Kernel, Dt.: Kern)

The *kernel* (or *null space*) of an \(m \times n\) matrix \(A\) (of a linear transformation \(g\)) is the set of all vectors \(u \in \mathbb{R}^n\) which get mapped by \(g\) to the zero vector of \(\mathbb{R}^m\).

- Hence, if \(u_0 \in \mathbb{R}^n\) is a solution of \(Au = b\) then \(u_0 + w\) is also a solution of \(Au = b\) for all \(w\) in the kernel of \(A\).
Definition 218 (Kernel, Dt.: Kern)

The kernel (or null space) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $u \in \mathbb{R}^n$ which get mapped by $g$ to the zero vector of $\mathbb{R}^m$.

- Hence, if $u_0 \in \mathbb{R}^n$ is a solution of $Au = b$ then $u_0 + w$ is also a solution of $Au = b$ for all $w$ in the kernel of $A$.
- The kernel of an $m \times n$ matrix forms a subspace of $\mathbb{R}^n$. 
Rank, Image and Kernel of a Transformation Matrix

Definition 218 (Kernel, Dt.: Kern)

The kernel (or null space) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $u \in \mathbb{R}^n$ which get mapped by $g$ to the zero vector of $\mathbb{R}^m$.

- Hence, if $u_0 \in \mathbb{R}^n$ is a solution of $Au = b$ then $u_0 + w$ is also a solution of $Au = b$ for all $w$ in the kernel of $A$.
- The kernel of an $m \times n$ matrix forms a subspace of $\mathbb{R}^n$.

Definition 219 (Corank, Dt.: Defekt)

The corank (nullity) of an $m \times n$ matrix $A$, denoted by $\text{corank}(A)$, is the dimension of the kernel of $A$. 
Rank, Image and Kernel of a Transformation Matrix

**Definition 218 (Kernel, Dt.: Kern)**

The *kernel* (or *null space*) of an $m \times n$ matrix $A$ (of a linear transformation $g$) is the set of all vectors $u \in \mathbb{R}^n$ which get mapped by $g$ to the zero vector of $\mathbb{R}^m$.

- Hence, if $u_0 \in \mathbb{R}^n$ is a solution of $Au = b$ then $u_0 + w$ is also a solution of $Au = b$ for all $w$ in the kernel of $A$.
- The kernel of an $m \times n$ matrix forms a subspace of $\mathbb{R}^n$.

**Definition 219 (Corank, Dt.: Defekt)**

The *corank* (nullity) of an $m \times n$ matrix $A$, denoted by $\text{corank}(A)$, is the dimension of the kernel of $A$.

**Theorem 220 (Rank-nullity theorem, Dt.: Rangsatz, Dimensionssatz)**

Consider an $m \times n$ matrix $A$. Then

$$\text{rank}(A) + \text{corank}(A) = n.$$
Geometric Interpretation of the Dot Product

- Recall that $\langle a, b \rangle := a_x \cdot b_x + a_y \cdot b_y + \ldots + a_n \cdot b_n$ for $a, b \in \mathbb{R}^n$. 
Geometric Interpretation of the Dot Product

- Recall that $\langle a, b \rangle := a_x \cdot b_x + a_y \cdot b_y + \ldots + a_n \cdot b_n$ for $a, b \in \mathbb{R}^n$.

- In Lemma 127 we claimed that the length of the orthogonal projection of a vector $b$ onto a non-zero vector $a$ is given by

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

Geometric Interpretation of the Dot Product

- Recall that $\langle a, b \rangle := a_x \cdot b_x + a_y \cdot b_y + \ldots + a_n \cdot b_n$ for $a, b \in \mathbb{R}^n$.
- In Lemma 127 we claimed that the length of the orthogonal projection of a vector $b$ onto a non-zero vector $a$ is given by
  $$\frac{\langle a, b \rangle}{\|a\|}.$$
- We consider $n := 2$. Let $a \in \mathbb{R}^2$ be arbitrary but fixed, with $\|a\| = 1$. 

Geometric Interpretation of the Dot Product

- Recall that $\langle a, b \rangle := a_x \cdot b_x + a_y \cdot b_y + \ldots + a_n \cdot b_n$ for $a, b \in \mathbb{R}^n$.
- In Lemma 127 we claimed that the length of the orthogonal projection of a vector $b$ onto a non-zero vector $a$ is given by
  \[
  \frac{\langle a, b \rangle}{\|a\|}.
  \]
- We consider $n := 2$. Let $a \in \mathbb{R}^2$ be arbitrary but fixed, with $\|a\| = 1$.
- Then we can regard $\langle a, b \rangle$ as a linear transformation by a $1 \times 2$ matrix $A$ that maps every $b \in \mathbb{R}^2$ to a value in $\mathbb{R}$:
  \[
  \langle a, b \rangle = a_x \cdot b_x + a_y \cdot b_y = (a_x \ a_y) \cdot \begin{pmatrix} b_x \\ b_y \end{pmatrix} = A \cdot b \quad \text{with} \quad A := \begin{pmatrix} a_x & a_y \end{pmatrix}
  \]
Geometric Interpretation of the Dot Product

- Recall that \( \langle a, b \rangle := a_x \cdot b_x + a_y \cdot b_y + \ldots + a_n \cdot b_n \) for \( a, b \in \mathbb{R}^n \).
- In Lemma 127 we claimed that the length of the orthogonal projection of a vector \( b \) onto a non-zero vector \( a \) is given by \( \frac{\langle a, b \rangle}{\|a\|} \).
- We consider \( n := 2 \). Let \( a \in \mathbb{R}^2 \) be arbitrary but fixed, with \( \|a\| = 1 \).
- Then we can regard \( \langle a, b \rangle \) as a linear transformation by a \( 1 \times 2 \) matrix \( A \) that maps every \( b \in \mathbb{R}^2 \) to a value in \( \mathbb{R} \):
  \[
  \langle a, b \rangle = a_x \cdot b_x + a_y \cdot b_y = (a_x \ a_y) \cdot \begin{pmatrix} b_x \\ b_y \end{pmatrix} = A \cdot b \quad \text{with} \quad A := \begin{pmatrix} a_x \\ a_y \end{pmatrix}
  \]
- We know that a linear transformation is fully specified by the images of the unit vectors.
- So, how do the unit vectors \( e_1, e_2 \) of \( \mathbb{R}^2 \) get mapped by this transformation? And what is the geometric interpretation of this transformation? That is, what is the geometric interpretation of the dot product?
Geometric Interpretation of the Dot Product

- Observe that the length $s$ of the orthogonal projection of the unit vector $e_1$ (of the $x$-axis) onto $a$ equals the $x$-coordinate of $a$: 

\[ s = a_x \]

By the same argument, the length of the orthogonal projection of the unit vector $e_2$ (of the $y$-axis) onto $a$ equals the $y$-coordinate of $a$.

We conclude that

\[
\langle a, e_1 \rangle = a_x \\
\langle a, e_2 \rangle = a_y
\]

It remains to observe that the length $d$ of the projection of $b$ onto $a$ equals the sum of the lengths of the projections of $b_x \cdot e_1$ and $b_y \cdot e_2$ onto $a$.

Hence, for $\|a\| = 1$,

\[
d = \langle a, b_x \cdot e_1 \rangle + \langle a, b_y \cdot e_2 \rangle = b_x \cdot \langle a, e_1 \rangle + b_y \cdot \langle a, e_2 \rangle = b_x \cdot a_x + b_y \cdot a_y = \langle a, b \rangle.
\]
Observe that the length $s$ of the orthogonal projection of the unit vector $e_1$ (of the $x$-axis) onto $a$ equals the $x$-coordinate of $a$:

\[ \langle a, e_1 \rangle = a_x \]

By the same argument, the length of the orthogonal projection of the unit vector $e_2$ (of the $y$-axis) onto $a$ equals the $y$-coordinate of $a$.

We conclude that
\[ \langle a, e_1 \rangle = a_x \quad \text{and} \quad \langle a, e_2 \rangle = a_y. \]

It remains to observe that the length $d$ of the projection of $b$ onto $a$ equals the sum of the lengths of the projections of $b_x \cdot e_1$ and $b_y \cdot e_2$ onto $a$.

Hence, for $\|a\| = 1$,
\[ d = \langle a, b_x \cdot e_1 \rangle + \langle a, b_y \cdot e_2 \rangle = b_x \cdot \langle a, e_1 \rangle + b_y \cdot \langle a, e_2 \rangle = b_x \cdot a_x + b_y \cdot a_y = \langle a, b \rangle. \]
Geometric Interpretation of the Dot Product

- Observe that the length \( s \) of the orthogonal projection of the unit vector \( e_1 \) (of the \( x \)-axis) onto \( a \) equals the \( x \)-coordinate of \( a \): Due to symmetry, \( s = a_x \)!
Geometric Interpretation of the Dot Product

- Observe that the length $s$ of the orthogonal projection of the unit vector $e_1$ (of the $x$-axis) onto $a$ equals the $x$-coordinate of $a$: Due to symmetry, $s = a_x$!
- By the same argument, the length of the orthogonal projection of the unit vector $e_2$ (of the $y$-axis) onto $a$ equals the $y$-coordinate $a_y$ of $a$. 

![Diagram showing geometric interpretation of the dot product](image)
Geometric Interpretation of the Dot Product

- Observe that the length \( s \) of the orthogonal projection of the unit vector \( e_1 \) (of the \( x \)-axis) onto \( a \) equals the \( x \)-coordinate of \( a \): Due to symmetry, \( s = a_x \)!
- By the same argument, the length of the orthogonal projection of the unit vector \( e_2 \) (of the \( y \)-axis) onto \( a \) equals the \( y \)-coordinate \( a_y \) of \( a \).
Geometric Interpretation of the Dot Product

- Observe that the length $s$ of the orthogonal projection of the unit vector $e_1$ (of the $x$-axis) onto $a$ equals the $x$-coordinate of $a$: Due to symmetry, $s = a_x$!
- By the same argument, the length of the orthogonal projection of the unit vector $e_2$ (of the $y$-axis) onto $a$ equals the $y$-coordinate $a_y$ of $a$.
- We conclude that $\langle a, e_1 \rangle = a_x$ and $\langle a, e_2 \rangle = a_y$. 
Geometric Interpretation of the Dot Product

- Observe that the length $s$ of the orthogonal projection of the unit vector $e_1$ (of the $x$-axis) onto $a$ equals the $x$-coordinate of $a$: Due to symmetry, $s = a_x$!
- By the same argument, the length of the orthogonal projection of the unit vector $e_2$ (of the $y$-axis) onto $a$ equals the $y$-coordinate $a_y$ of $a$.
- We conclude that $\langle a, e_1 \rangle = a_x$ and $\langle a, e_2 \rangle = a_y$.
- It remains to observe that the length $d$ of the projection of $b$ onto $a$ equals the sum of the lengths of the projections of $b_x \cdot e_1$ and $b_y \cdot e_2$ onto $a$. 

![Diagram of vectors and projections](image-url)
Geometric Interpretation of the Dot Product

- Observe that the length $s$ of the orthogonal projection of the unit vector $e_1$ (of the $x$-axis) onto $a$ equals the $x$-coordinate of $a$: Due to symmetry, $s = a_x$!
- By the same argument, the length of the orthogonal projection of the unit vector $e_2$ (of the $y$-axis) onto $a$ equals the $y$-coordinate $a_y$ of $a$.
- We conclude that $\langle a, e_1 \rangle = a_x$ and $\langle a, e_2 \rangle = a_y$.
- It remains to observe that the length $d$ of the projection of $b$ onto $a$ equals the sum of the lengths of the projections of $b_x \cdot e_1$ and $b_y \cdot e_2$ onto $a$. 

\[ d = \langle a, b_x \cdot e_1 \rangle + \langle a, b_y \cdot e_2 \rangle = b_x \cdot \langle a, e_1 \rangle + b_y \cdot \langle a, e_2 \rangle = b_x \cdot a_x + b_y \cdot a_y = \langle a, b \rangle. \]
Geometric Interpretation of the Dot Product

- Observe that the length $s$ of the orthogonal projection of the unit vector $e_1$ (of the $x$-axis) onto $a$ equals the $x$-coordinate of $a$: Due to symmetry, $s = a_x$!
- By the same argument, the length of the orthogonal projection of the unit vector $e_2$ (of the $y$-axis) onto $a$ equals the $y$-coordinate $a_y$ of $a$.
- We conclude that $\langle a, e_1 \rangle = a_x$ and $\langle a, e_2 \rangle = a_y$.
- It remains to observe that the length $d$ of the projection of $b$ onto $a$ equals the sum of the lengths of the projections of $b_x \cdot e_1$ and $b_y \cdot e_2$ onto $a$.
- Hence, for $\|a\| = 1$,

$$d = \langle a, b_x \cdot e_1 \rangle + \langle a, b_y \cdot e_2 \rangle = b_x \cdot \langle a, e_1 \rangle + b_y \cdot \langle a, e_2 \rangle = b_x \cdot a_x + b_y \cdot a_y = \langle a, b \rangle.$$
Duality: Vector and Linear Transformation

- Note the duality between vectors in $\mathbb{R}^n$ and linear transformations from $\mathbb{R}^n$ to $\mathbb{R}$ by $1 \times n$ matrices!
- Every linear transformation $g : \mathbb{R}^n \to \mathbb{R}$ that maps a vector of $\mathbb{R}^n$ to $\mathbb{R}$—i.e., to a scalar value—has a corresponding dual vector out of $\mathbb{R}^n$, and vice versa:
Duality: Vector and Linear Transformation

- Note the duality between vectors in \( \mathbb{R}^n \) and linear transformations from \( \mathbb{R}^n \) to \( \mathbb{R} \) by \( 1 \times n \) matrices!

- Every linear transformation \( g: \mathbb{R}^n \rightarrow \mathbb{R} \) that maps a vector of \( \mathbb{R}^n \) to \( \mathbb{R} \) — i.e., to a scalar value — has a corresponding dual vector out of \( \mathbb{R}^n \), and vice versa:
  - Let \( A \) be the matrix of the linear transformation \( g \).
Duality: Vector and Linear Transformation

- Note the duality between vectors in $\mathbb{R}^n$ and linear transformations from $\mathbb{R}^n$ to $\mathbb{R}$ by $1 \times n$ matrices!

- Every linear transformation $g: \mathbb{R}^n \rightarrow \mathbb{R}$ that maps a vector of $\mathbb{R}^n$ to $\mathbb{R}$ — i.e., to a scalar value — has a corresponding dual vector out of $\mathbb{R}^n$, and vice versa:
  - Let $A$ be the matrix of the linear transformation $g$.
  - Then $A \in M_{1 \times n}$, i.e.,
    \[
    A = [a_{11} \ a_{12} \ldots \ a_{1n}].
    \]
Duality: Vector and Linear Transformation

- Note the duality between vectors in $\mathbb{R}^n$ and linear transformations from $\mathbb{R}^n$ to $\mathbb{R}$ by $1 \times n$ matrices!
- Every linear transformation $g : \mathbb{R}^n \to \mathbb{R}$ that maps a vector of $\mathbb{R}^n$ to $\mathbb{R}$ — i.e., to a scalar value — has a corresponding dual vector out of $\mathbb{R}^n$, and vice versa:
  - Let $A$ be the matrix of the linear transformation $g$.
  - Then $A \in M_{1 \times n}$, i.e.,
    
    \[
    A = [a_{11} \ a_{12} \ldots \ a_{1n}].
    \]
  - Hence, we may consider $g$ to be dual to
    
    \[
    a := \begin{pmatrix}
    a_{11} \\
    a_{12} \\
    \vdots \\
    a_{1n}
    \end{pmatrix} \in \mathbb{R}^n
    \]
    
    since $g(u) = Au = \langle a, u \rangle$. 

Duality: Vector and Linear Transformation

- Note the duality between vectors in $\mathbb{R}^n$ and linear transformations from $\mathbb{R}^n$ to $\mathbb{R}$ by $1 \times n$ matrices!

- Every linear transformation $g: \mathbb{R}^n \to \mathbb{R}$ that maps a vector of $\mathbb{R}^n$ to $\mathbb{R}$ — i.e., to a scalar value — has a corresponding dual vector out of $\mathbb{R}^n$, and vice versa:
  - Let $A$ be the matrix of the linear transformation $g$.
  - Then $A \in M_{1 \times n}$, i.e.,
    \[
    A = [a_{11} \ a_{12} \ldots \ a_{1n}].
    \]
  - Hence, we may consider $g$ to be dual to
    \[
    a := \begin{pmatrix}
    a_{11} \\
    a_{12} \\
    \vdots \\
    a_{1n}
    \end{pmatrix} \in \mathbb{R}^n
    \]
    since $g(u) = Au = \langle a, u \rangle$.
  - On the other hand, every vector of $\mathbb{R}^n$ induces a dot product and, thus, corresponds to a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}$. 

Geometric Interpretation of the Cross Product

- Consider $\|a \times b\|$ for two vectors $a, b \in \mathbb{R}^3$. We will
  - define a linear transformation $g: \mathbb{R}^3 \to \mathbb{R}$ that involves $a$ and $b$,
  - consider its dual vector $c$, and
  - explain why $c$ equals $a \times b$, thus getting a geometric insight into $\|a \times b\|$. 

Remember Lemma 116: This determinant equals the (signed) volume of the parallelepiped spanned by the three vectors $u, a, b \in \mathbb{R}^3$. Note that $g$ is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}$ for every pair of fixed vectors $a, b \in \mathbb{R}^3$. By duality, there exists a vector $c$ such that

$$
\det \begin{bmatrix}
  u_x & a_x & b_x \\
  u_y & a_y & b_y \\
  u_z & a_z & b_z 
\end{bmatrix} = g(u) = \begin{bmatrix}
  c_x \\
  c_y \\
  c_z 
\end{bmatrix} \cdot \begin{bmatrix}
  u_x \\
  u_y \\
  u_z 
\end{bmatrix} = \langle c, u \rangle.
$$
Geometric Interpretation of the Cross Product

- Consider $\|a \times b\|$ for two vectors $a, b \in \mathbb{R}^3$. We will
  - define a linear transformation $g : \mathbb{R}^3 \to \mathbb{R}$ that involves $a$ and $b$,
  - consider its dual vector $c$, and
  - explain why $c$ equals $a \times b$, thus getting a geometric insight into $\|a \times b\|$.

- We define the transformation $g : \mathbb{R}^3 \to \mathbb{R}$ as

$$g(u) := \det \begin{pmatrix} u_x & a_x & b_x \\ u_y & a_y & b_y \\ u_z & a_z & b_z \end{pmatrix}.$$
Geometric Interpretation of the Cross Product

- Consider $\|a \times b\|$ for two vectors $a, b \in \mathbb{R}^3$. We will define a linear transformation $g: \mathbb{R}^3 \rightarrow \mathbb{R}$ that involves $a$ and $b$,
  - consider its dual vector $c$, and
  - explain why $c$ equals $a \times b$, thus getting a geometric insight into $\|a \times b\|$.

We define the transformation $g: \mathbb{R}^3 \rightarrow \mathbb{R}$ as

$$g(u) := \det \begin{pmatrix} u_x & a_x & b_x \\ u_y & a_y & b_y \\ u_z & a_z & b_z \end{pmatrix}.$$ 

- Remember Lemma 116: This determinant equals the (signed) volume of the parallelepiped spanned by the three vectors $u, a, b \in \mathbb{R}^3$. 

Geometric Interpretation of the Cross Product

- Consider $\|a \times b\|$ for two vectors $a, b \in \mathbb{R}^3$. We will
  - define a linear transformation $g: \mathbb{R}^3 \to \mathbb{R}$ that involves $a$ and $b$,
  - consider its dual vector $c$, and
  - explain why $c$ equals $a \times b$, thus getting a geometric insight into $\|a \times b\|$.

- We define the transformation $g: \mathbb{R}^3 \to \mathbb{R}$ as
  \[
g(u) := \det \begin{pmatrix} u_x & a_x & b_x \\ u_y & a_y & b_y \\ u_z & a_z & b_z \end{pmatrix}.
  \]

- Remember Lemma 116: This determinant equals the (signed) volume of the parallelepiped spanned by the three vectors $u, a, b \in \mathbb{R}^3$.

- Note that $g$ is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}$ for every pair of fixed vectors $a, b \in \mathbb{R}^3$. 
Geometric Interpretation of the Cross Product

- Consider $\|a \times b\|$ for two vectors $a, b \in \mathbb{R}^3$. We will
  - define a linear transformation $g: \mathbb{R}^3 \to \mathbb{R}$ that involves $a$ and $b$,
  - consider its dual vector $c$, and
  - explain why $c$ equals $a \times b$, thus getting a geometric insight into $\|a \times b\|$.

- We define the transformation $g: \mathbb{R}^3 \to \mathbb{R}$ as

$$g(u) := \det \begin{pmatrix} u_x & a_x & b_x \\ u_y & a_y & b_y \\ u_z & a_z & b_z \end{pmatrix}.$$ 

- Remember Lemma 116: This determinant equals the (signed) volume of the parallelepiped spanned by the three vectors $u, a, b \in \mathbb{R}^3$.

- Note that $g$ is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}$ for every pair of fixed vectors $a, b \in \mathbb{R}^3$.

- By duality, there exists a vector $c$ such that

$$\det \begin{pmatrix} u_x & a_x & b_x \\ u_y & a_y & b_y \\ u_z & a_z & b_z \end{pmatrix} = g(u) = [c_x c_y c_z] \cdot \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = \langle c, u \rangle.$$
Geometric Interpretation of the Cross Product

Hence, for all $u \in \mathbb{R}^3$,

$$c_x \cdot u_x + c_y \cdot u_y + c_z \cdot u_z = u_x \cdot (a_y \cdot b_z - a_z \cdot b_y) + u_y \cdot (a_z \cdot b_x - a_x \cdot b_z) + u_z \cdot (a_x \cdot b_y - a_y \cdot b_x),$$

which implies

$$c = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} = \begin{pmatrix} a_y \cdot b_z - a_z \cdot b_y \\ a_z \cdot b_x - a_x \cdot b_z \\ a_x \cdot b_y - a_y \cdot b_x \end{pmatrix} = a \times b.$$
Hence, for all $u \in \mathbb{R}^3$,

$$c_x u_x + c_y u_y + c_z u_z = u_x (a_y b_z - a_z b_y) + u_y (a_z b_x - a_x b_z) + u_z (a_x b_y - a_y b_x),$$

which implies

$$c = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix} \overset{\text{Def. 131}}{=} a \times b.$$
Geometric Interpretation of the Cross Product

- Hence, for all $u \in \mathbb{R}^3$, 
  \[
c_x \cdot u_x + c_y \cdot u_y + c_z \cdot u_z = u_x \cdot (a_y \cdot b_z - a_z \cdot b_y) + u_y \cdot (a_z \cdot b_x - a_x \cdot b_z) + u_z \cdot (a_x \cdot b_y - a_y \cdot b_x),
\]
  which implies 
  \[
c = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} = \begin{pmatrix} a_y \cdot b_z - a_z \cdot b_y \\ a_z \cdot b_x - a_x \cdot b_z \\ a_x \cdot b_y - a_y \cdot b_x \end{pmatrix} \text{ Def. } \overset{131}{=} a \times b.
\]

- Elementary geometry tells us that the volume $V$ of the parallelepiped spanned by $a, b$ and a third vector $u$ can be obtained in the following way:
Geometric Interpretation of the Cross Product

Hence, for all \( u \in \mathbb{R}^3 \),

\[
c_x \cdot u_x + c_y \cdot u_y + c_z \cdot u_z = u_x \cdot (a_y \cdot b_z - a_z \cdot b_y) + u_y \cdot (a_z \cdot b_x - a_x \cdot b_z) + u_z \cdot (a_x \cdot b_y - a_y \cdot b_x),
\]

which implies

\[
c = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} = \begin{pmatrix} a_y \cdot b_z - a_z \cdot b_y \\ a_z \cdot b_x - a_x \cdot b_z \\ a_x \cdot b_y - a_y \cdot b_x \end{pmatrix} \overset{\text{Def. 131}}{=} a \times b.
\]

Elementary geometry tells us that the volume \( V \) of the parallelepiped spanned by \( a, b \) and a third vector \( u \) can be obtained in the following way: Multiply the area \( A \) of the parallelogram spanned by \( a, b \) with the height of the parallelepiped, i.e., with the length of that component of \( u \) that is perpendicular onto \( a, b \). Hence,

\[
V = A \cdot \frac{\langle a \times b, u \rangle}{\|a \times b\|} = \frac{A}{\|a \times b\|} \cdot \langle a \times b, u \rangle.
\]
Geometric Interpretation of the Cross Product

- Hence, for all $u \in \mathbb{R}^3$,
  
  \[ c_x \cdot u_x + c_y \cdot u_y + c_z \cdot u_z = u_x \cdot (a_y \cdot b_z - a_z \cdot b_y) + u_y \cdot (a_z \cdot b_x - a_x \cdot b_z) + u_z \cdot (a_x \cdot b_y - a_y \cdot b_x), \]

  which implies

  \[
  c = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} = \begin{pmatrix} a_y \cdot b_z - a_z \cdot b_y \\ a_z \cdot b_x - a_x \cdot b_z \\ a_x \cdot b_y - a_y \cdot b_x \end{pmatrix} \text{ Def. } 131 = a \times b.
  \]

- Elementary geometry tells us that the volume $V$ of the parallelepiped spanned by $a, b$ and a third vector $u$ can be obtained in the following way: Multiply the area $A$ of the parallelogram spanned by $a, b$ with the height of the parallelepiped, i.e., with the length of that component of $u$ that is perpendicular onto $a, b$. Hence,

  \[
  V = A \cdot \frac{\langle a \times b, u \rangle}{\| a \times b \|} = \frac{A}{\| a \times b \|} \cdot \langle a \times b, u \rangle.
  \]

- On the other hand, we derived $g(u) = V = \langle c, u \rangle = \langle a \times b, u \rangle$. 
Geometric Interpretation of the Cross Product

- Hence, for all $u \in \mathbb{R}^3$,
  \[ c_x \cdot u_x + c_y \cdot u_y + c_z \cdot u_z = u_x \cdot (a_y \cdot b_z - a_z \cdot b_y) + u_y \cdot (a_z \cdot b_x - a_x \cdot b_z) + u_z \cdot (a_x \cdot b_y - a_y \cdot b_x), \]
  which implies
  \[ \mathbf{c} = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} = \begin{pmatrix} a_y \cdot b_z - a_z \cdot b_y \\ a_z \cdot b_x - a_x \cdot b_z \\ a_x \cdot b_y - a_y \cdot b_x \end{pmatrix} \overset{\text{Def. 131}}{=} \mathbf{a} \times \mathbf{b}. \]

- Elementary geometry tells us that the volume $V$ of the parallelepiped spanned by $\mathbf{a}$, $\mathbf{b}$ and a third vector $\mathbf{u}$ can be obtained in the following way: Multiply the area $A$ of the parallelogram spanned by $\mathbf{a}$, $\mathbf{b}$ with the height of the parallelepiped, i.e., with the length of that component of $\mathbf{u}$ that is perpendicular onto $\mathbf{a}$, $\mathbf{b}$. Hence,
  \[ V = A \cdot \frac{\langle \mathbf{a} \times \mathbf{b}, \mathbf{u} \rangle}{\| \mathbf{a} \times \mathbf{b} \|} = \frac{A}{\| \mathbf{a} \times \mathbf{b} \|} \cdot \langle \mathbf{a} \times \mathbf{b}, \mathbf{u} \rangle. \]

- On the other hand, we derived $g(u) = V = \langle \mathbf{c}, \mathbf{u} \rangle = \langle \mathbf{a} \times \mathbf{b}, \mathbf{u} \rangle$.

- We conclude that
  \[ A = \| \mathbf{a} \times \mathbf{b} \|, \]
  i.e., that the length of $\mathbf{a} \times \mathbf{b}$ equals the area of the parallelogram spanned by $\mathbf{a}, \mathbf{b}$.
Transformations

- Linear Transformations
- Classification of Transformations
- Coordinate Transformations in $\mathbb{R}^2$
- Coordinate Transformations in $\mathbb{R}^3$
- Transformation of Coordinate Systems
- Applications of Coordinate (System) Transformations
- Rotations Revisited
- Projections
Classification of Transformations

- Consider a mapping \( g : \mathbb{R}^n \rightarrow \mathbb{R}^n \) and a distance metric \( d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \).
- E.g., take \( n = 2 \) and the standard Euclidean distance
  \[
d(p, q) := \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}.
  \]
Classification of Transformations

- Consider a mapping $g : \mathbb{R}^n \to \mathbb{R}^n$ and a distance metric $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.
- E.g., take $n = 2$ and the standard Euclidean distance

$$d(p, q) := \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}.$$ 

**Definition 221 (Isometry, Dt.: Isometrie)**

A mapping $g : \mathbb{R}^n \to \mathbb{R}^n$ is called an *isometry* if it maps pairs of points to points the same distance apart. That is,

$$\forall (p, q \in \mathbb{R}^n) \quad d(g(p), g(q)) = d(p, q).$$
Classification of Transformations

- Consider a mapping \( g : \mathbb{R}^n \to \mathbb{R}^n \) and a distance metric \( d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \).
- E.g., take \( n = 2 \) and the standard Euclidean distance

\[
d(p, q) := \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}.
\]

**Definition 221 (Isometry, Dt.: Isometrie)**

A mapping \( g : \mathbb{R}^n \to \mathbb{R}^n \) is called an isometry if it maps pairs of points to points the same distance apart. That is,

\[
\forall (p, q \in \mathbb{R}^n) \quad d(g(p), g(q)) = d(p, q).
\]

- Another widely-used term for characterizing an isometry is distance-preserving transformation.
- In planar Euclidean geometry such a mapping is also called a congruence, and two objects \( A \) and \( B \) are said to be congruent if there exists an isometry that maps \( A \) to \( B \).
- E.g., two triangles which are congruent have corresponding sides of equal length.
Classification of Transformations

**Definition 222 (Rigid motion, Dt.: Bewegung)**

An isometry \( g \) is called a *rigid motion* if it preserves handedness.
Classification of Transformations

Definition 222 (Rigid motion, Dt.: Bewegung)

An isometry $g$ is called a *rigid motion* if it preserves handedness.

- Two objects $A$ and $B$ are said to be *equal* if there exists a rigid motion that maps $A$ to $B$. 
Classification of Transformations

**Definition 222 (Rigid motion, Dt.: Bewegung)**

An isometry \( g \) is called a *rigid motion* if it preserves handedness.

- Two objects \( A \) and \( B \) are said to be *equal* if there exists a rigid motion that maps \( A \) to \( B \).

**Caveat**

Several authors regard “rigid motion” as a synonym for “isometry”.

---

© M. Held (Univ. Salzburg)

*Geometrisches Rechnen (WS 2020/21)*

237/357
Classification of Transformations

Definition 222 (Rigid motion, Dt.: Bewegung)
An isometry $g$ is called a *rigid motion* if it preserves handedness.

- Two objects $A$ and $B$ are said to be *equal* if there exists a rigid motion that maps $A$ to $B$.

Caveat
Several authors regard “rigid motion” as a synonym for “isometry”.

- But there is a difference also when seen from a practical point of view: A rigid motion preserves the shape of an object, while an isometry may change the shape: Left glove versus right glove!
Definition 223 (Orthogonal transformation, Dt.: orthogonale Transformation)
A linear mapping that preserves distance is called \textit{orthogonal transformation}. (And the class of all such transformations on $\mathbb{R}^n$ forms the \textit{orthogonal group} of $\mathbb{R}^n$.)
Classification of Transformations

Definition 223 (Orthogonal transformation, Dt.: orthogonale Transformation)
A linear mapping that preserves distance is called *orthogonal transformation*. (And the class of all such transformations on $\mathbb{R}^n$ forms the *orthogonal group* of $\mathbb{R}^n$.)

- Hence, an orthogonal transformation is a special isometry.
Classification of Transformations

**Definition 223 (Orthogonal transformation, Dt.: orthogonale Transformation)**

A linear mapping that preserves distance is called *orthogonal transformation*. (And the class of all such transformations on $\mathbb{R}^n$ forms the *orthogonal group* of $\mathbb{R}^n$.)

- Hence, an orthogonal transformation is a special isometry.

**Lemma 224**

The group of all isometries on $\mathbb{R}^n$ is given by composites of a translation and an orthogonal transformation.
Classification of Transformations

**Definition 223 (Orthogonal transformation, Dt.: orthogonale Transformation)**

A linear mapping that preserves distance is called *orthogonal transformation*. (And the class of all such transformations on $\mathbb{R}^n$ forms the *orthogonal group* of $\mathbb{R}^n$.)

- Hence, an orthogonal transformation is a special isometry.

**Lemma 224**

The group of all isometries on $\mathbb{R}^n$ is given by composites of a translation and an orthogonal transformation.

**Lemma 225**

The group of all rigid motions on $\mathbb{R}^n$ is given by composites of a translation and a rotation.
Classification of Transformations

Lemma 226

With respect to an orthonormal basis of $\mathbb{R}^n$, an orthogonal transformation has a corresponding *orthogonal matrix*, i.e., a matrix whose columns and rows are orthonormal vectors.
Classification of Transformations

Lemma 226
With respect to an orthonormal basis of $\mathbb{R}^n$, an orthogonal transformation has a corresponding *orthogonal matrix*, i.e., a matrix whose columns and rows are orthonormal vectors.

Corollary 227
With respect to an orthonormal basis of $\mathbb{R}^n$, an orthogonal transformation is invertible: If its matrix is $A$ then the inverse transformation has matrix $A^t$. Furthermore, $\det A = \pm 1$. 
Classification of Transformations

**Lemma 226**

With respect to an orthonormal basis of $\mathbb{R}^n$, an orthogonal transformation has a corresponding *orthogonal matrix*, i.e., a matrix whose columns and rows are orthonormal vectors.

**Corollary 227**

With respect to an orthonormal basis of $\mathbb{R}^n$, an orthogonal transformation is invertible: If its matrix is $A$ then the inverse transformation has matrix $A^t$. Furthermore, $\det A = \pm 1$.

**Lemma 228**

A $2 \times 2$ orthogonal matrix $A$ is the matrix of a rotation about the origin if and only if $\det A = 1$. If $\det A = -1$ then it is the matrix of a reflection.
Classification of Transformations

Lemma 226
With respect to an orthonormal basis of $\mathbb{R}^n$, an orthogonal transformation has a corresponding **orthogonal matrix**, i.e., a matrix whose columns and rows are orthonormal vectors.

Corollary 227
With respect to an orthonormal basis of $\mathbb{R}^n$, an orthogonal transformation is invertible: If its matrix is $A$ then the inverse transformation has matrix $A^t$. Furthermore, $\det A = \pm 1$.

Lemma 228
A $2 \times 2$ orthogonal matrix $A$ is the matrix of a rotation about the origin if and only if $\det A = 1$. If $\det A = -1$ then it is the matrix of a reflection.

Lemma 229
A $3 \times 3$ orthogonal matrix $A$ is the matrix of a rotation about a straight line through the origin if and only if $\det A = 1$. 
Classification of Transformations

Definition 230 (Similarity mapping, Dt.: Ähnlichkeitsabbildung)

A mapping $g$ is called a *similarity mapping* if it preserves angles.
Definition 230 (Similarity mapping, Dt.: Ähnlichkeitsabbildung)

A mapping $g$ is called a *similarity mapping* if it preserves angles.

- E.g., two triangles which are similar have identical angles, and their sides are "in proportion".
Classification of Transformations

**Definition 230 (Similarity mapping, Dt.: Ähnlichkeitsabbildung)**

A mapping \( g \) is called a *similarity mapping* if it preserves angles.

- E.g., two triangles which are similar have identical angles, and their sides are "in proportion".

**Lemma 231**

A distance-preserving transformation is a similarity mapping, i.e., it preserves angles.
Definition 232 (Affine transformation, Dt.: affine Abbildung)

A mapping $g$ is called *affine transformation* (or *affinity*) if it is a composite of a translation and a linear transformation.
Classification of Transformations

Definition 232 (Affine transformation, Dt.: affine Abbildung)

A mapping $g$ is called *affine transformation* (or *affinity*) if it is a composite of a translation and a linear transformation.

- Affine transformations need not preserve distance, angle, area or volume.
Classification of Transformations

**Definition 232 (Affine transformation, Dt.: affine Abbildung)**

A mapping $g$ is called *affine transformation* (or *affinity*) if it is a composite of a translation and a linear transformation.

- Affine transformations need not preserve distance, angle, area or volume.

**Lemma 233**

If $g$ is an affine transformation and $p, q, r$ are collinear, then $g(p), g(q), g(r)$ are collinear. That is, affine transformations preserve lines.

**Corollary 234**

An affine transformation maps parallel lines to parallel lines.

**Lemma 235**

An affine transformation preserves ratios of lengths of intervals on any line.
Classification of Transformations

Definition 232 (Affine transformation, Dt.: affine Abbildung)
A mapping $g$ is called *affine transformation* (or *affinity*) if it is a composite of a translation and a linear transformation.

- Affine transformations need not preserve distance, angle, area or volume.

Lemma 233
If $g$ is an affine transformation and $p, q, r$ are collinear, then $g(p), g(q), g(r)$ are collinear. That is, affine transformations preserve lines.

Corollary 234
An affine transformation maps parallel lines to parallel lines.
Classification of Transformations

Definition 232 (Affine transformation, Dt.: affine Abbildung)
A mapping $g$ is called **affine transformation** (or **affinity**) if it is a composite of a translation and a linear transformation.

- Affine transformations need not preserve distance, angle, area or volume.

Lemma 233
If $g$ is an affine transformation and $p, q, r$ are collinear, then $g(p), g(q), g(r)$ are collinear. That is, affine transformations preserve lines.

Corollary 234
An affine transformation maps parallel lines to parallel lines.

Lemma 235
An affine transformation preserves ratios of lengths of intervals on any line.
Group Hierarchy of Transformations

- **orthogonal**
- **isometry**
- **similarity**
- **affine**
- **projective**
Transformations
- Linear Transformations
- Classification of Transformations
- Coordinate Transformations in $\mathbb{R}^2$
  - Rotation in $\mathbb{R}^2$
  - Stretching in $\mathbb{R}^2$
  - Shear Transformation in $\mathbb{R}^2$
  - Reflection in $\mathbb{R}^2$
  - Translation in $\mathbb{R}^2$
- Homogeneous Coordinates
- Transformation Matrices Based on Homogeneous Coordinates
- Coordinate Transformations in $\mathbb{R}^3$
- Transformation of Coordinate Systems
- Applications of Coordinate (System) Transformations
- Rotations Revisited
- Projections
Rotation in $\mathbb{R}^2$

- Rotation of point $p$ by $\theta$ about the origin yields point $p'$.

Polar coordinates:

\[
\begin{align*}
  p_x &:= r \cos \phi, \\
  p_y &:= r \sin \phi.
\end{align*}
\]

\[
\begin{align*}
  p'_x &= r \cos(\theta + \phi) = r \cos \theta \cos \phi - r \sin \theta \sin \phi = p_x \cos \theta - p_y \sin \theta, \\
  p'_y &= r \sin(\theta + \phi) = p_x \sin \theta + p_y \cos \theta.
\end{align*}
\]
Rotation of point $p$ by $\theta$ about the origin yields point $p'$. 

Polar coordinates: $p_x := r \cos \varphi$, $p_y := r \sin \varphi$. 

\[ p'_x = r \cos(\theta + \varphi) = r \cos \theta \cos \varphi - r \sin \theta \sin \varphi = p_x \cos \theta - p_y \sin \theta. \]

\[ p'_y = r \sin(\theta + \varphi) = p_x \sin \theta + p_y \cos \theta. \]
Rotation in $\mathbb{R}^2$

- Rotation of point $p$ by $\theta$ about the origin yields point $p'$.

Polar coordinates: $p_x := r \cos \varphi, \quad p_y := r \sin \varphi$.

- $p'_x = r \cos(\theta + \varphi)$
  $= r \cos \theta \cos \varphi - r \sin \theta \sin \varphi$
  $= p_x \cos \theta - p_y \sin \theta$.

- $p'_y = r \sin(\theta + \varphi)$
  $= p_x \sin \theta + p_y \cos \theta$. 
Rotation as a Matrix Transformation

- We have

\[
\begin{pmatrix}
  p'_x \\
  p'_y
\end{pmatrix} = \begin{pmatrix}
  p_x \cos \theta - p_y \sin \theta \\
  p_x \sin \theta + p_y \cos \theta
\end{pmatrix}
\]

for a rotation about the origin by the angle \( \theta \).

Lemma 236

Rotation matrices are orthogonal:

\( \text{Rot}(\theta)^{-1} = \text{Rot}(\theta)^T \).
Rotation as a Matrix Transformation

- We have
  \[
  \begin{pmatrix}
  p'_x \\
  p'_y
  \end{pmatrix} = \begin{pmatrix}
  p_x \cos \theta - p_y \sin \theta \\
  p_x \sin \theta + p_y \cos \theta
  \end{pmatrix}
  \]
  for a rotation about the origin by the angle \( \theta \).

- This relation can also be expressed by means of a rotation matrix \( \text{Rot}(\theta) \):
  \[
  \begin{pmatrix}
  p'_x \\
  p'_y
  \end{pmatrix} = \begin{pmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
  \end{pmatrix} \cdot \begin{pmatrix}
  p_x \\
  p_y
  \end{pmatrix};
  \]
  that is
  \[
  \text{Rot}(\theta) := \begin{pmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
  \end{pmatrix}.
  \]
Rotation as a Matrix Transformation

- We have

\[
\begin{pmatrix}
p_x' \\
p_y'
\end{pmatrix} = \begin{pmatrix}
p_x \cos \theta - p_y \sin \theta \\
p_x \sin \theta + p_y \cos \theta
\end{pmatrix}
\]

for a rotation about the origin by the angle \( \theta \).

- This relation can also be expressed by means of a rotation matrix \( \text{Rot}(\theta) \):

\[
\begin{pmatrix}
p_x' \\
p_y'
\end{pmatrix} = \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix} \cdot \begin{pmatrix}
p_x \\
p_y
\end{pmatrix};
\]

that is

\[
\text{Rot}(\theta) := \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}.
\]

Lemma 236

Rotation matrices are orthogonal: \( \text{Rot}(\theta)^{-1} = \text{Rot}(\theta)^t \).
General Rotation in $\mathbb{R}^2$

- Rotation of point $p$ by $\theta$ about point $a$, with $a := \begin{pmatrix} a_x \\ a_y \end{pmatrix}$, yields point $p'$.

$$p_x = a_x + r \cos \varphi \quad \text{thus,} \quad r \cos \varphi = p_x - a_x$$
$$p_y = a_y + r \sin \varphi \quad \text{thus,} \quad r \sin \varphi = p_y - a_y$$

$$p'_x = a_x + r \cos(\theta + \varphi)$$
$$= a_x + r \cos \theta \cos \varphi - r \sin \theta \sin \varphi$$
$$= a_x + (p_x - a_x) \cos \theta - (p_y - a_y) \sin \theta$$
$$p'_y = a_y + (p_x - a_x) \sin \theta + (p_y - a_y) \cos \theta$$
Stretching in $\mathbb{R}^2$

\[
\begin{pmatrix}
  p'_x \\
  p'_y
\end{pmatrix} = \begin{pmatrix}
  \lambda_1 & 0 \\
  0 & \lambda_2
\end{pmatrix} \cdot \begin{pmatrix}
  p_x \\
  p_y
\end{pmatrix}.
\]

If $\lambda_1 = \lambda_2$: uniform scaling;

If $\lambda_1 \neq \lambda_2$: non-uniform scaling or stretching.
Stretching in $\mathbb{R}^2$

\[
\begin{pmatrix}
 p'_x \\
 p'_y
\end{pmatrix} = \begin{pmatrix}
 \lambda_1 & 0 \\
 0 & \lambda_2
\end{pmatrix} \cdot \begin{pmatrix}
 p_x \\
 p_y
\end{pmatrix}.
\]

- If $\lambda_1 = \lambda_2$: (uniform) scaling;
- If $\lambda_1 \neq \lambda_2$: non-uniform scaling or stretching.
Shear Transformation in $\mathbb{R}^2$

- Suppose that we want to map a point $p$ to a point $p'$ such that

\[ p'_x = p_x + a \cdot p_y \quad \text{and} \quad p'_y = p_y. \]

Hence, a horizontal segment at height $y$ is shifted in the $x$-direction by $ay$.

\[ \text{SH}_x(0.5) \]
Shear Transformation in $\mathbb{R}^2$

- Suppose that we want to map a point $p$ to a point $p'$ such that
  
  $$p'_x = p_x + a \cdot p_y \quad \text{and} \quad p'_y = p_y.$$ 

  Hence, a horizontal segment at height $y$ is shifted in the $x$-direction by $ay$.

- The corresponding transformation matrix is given by
  
  $$\text{SH}_x(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$ 

Reflection in $\mathbb{R}^2$

- Reflection about $x$-axis:

\[
\begin{pmatrix}
    p'_x \\
    p'_y
\end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} p_x \\ p_y \end{pmatrix}.
\]
Reflection in $\mathbb{R}^2$

- Reflection about $x$-axis:
  \[
  \begin{pmatrix}
  p'_x \\
  p'_y
  \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} p_x \\
  p_y \end{pmatrix}.
  \]

- Reflection about $y$-axis:
  \[
  \begin{pmatrix}
  p'_x \\
  p'_y
  \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} p_x \\
  p_y \end{pmatrix}.
  \]

That is, a reflection about the origin is identical to a rotation about the origin by $\pi$. 
Reflection in $\mathbb{R}^2$

- Reflection about $x$-axis:
  \[
  \begin{pmatrix}
  p'_x \\
  p'_y
  \end{pmatrix}
  =
  \begin{pmatrix}
  1 & 0 \\
  0 & -1
  \end{pmatrix}
  \cdot
  \begin{pmatrix}
  p_x \\
  p_y
  \end{pmatrix}.
  \]

- Reflection about $y$-axis:
  \[
  \begin{pmatrix}
  p'_x \\
  p'_y
  \end{pmatrix}
  =
  \begin{pmatrix}
  -1 & 0 \\
  0 & 1
  \end{pmatrix}
  \cdot
  \begin{pmatrix}
  p_x \\
  p_y
  \end{pmatrix}.
  \]

- Reflection about origin:
  \[
  \begin{pmatrix}
  p'_x \\
  p'_y
  \end{pmatrix}
  =
  \begin{pmatrix}
  -1 & 0 \\
  0 & -1
  \end{pmatrix}
  \cdot
  \begin{pmatrix}
  p_x \\
  p_y
  \end{pmatrix}.
  \]

That is, a reflection about the origin is identical to a rotation about the origin by $\pi$. 
Translation in $\mathbb{R}^2$

- Translation: Move a point $p$ along a vector $v$ from its original location $p$ to its new location $p'$.

\[ p := \begin{pmatrix} p_x \\ p_y \end{pmatrix}, \quad v := \begin{pmatrix} v_x \\ v_y \end{pmatrix}, \quad p' := \begin{pmatrix} p'_x \\ p'_y \end{pmatrix}. \]

\[ p'_x = p_x + v_x, \quad p'_y = p_y + v_y, \quad p' = p + v. \]

\[ \begin{pmatrix} p'_x \\ p'_y \end{pmatrix} = \begin{pmatrix} p_x \\ p_y \end{pmatrix} + \begin{pmatrix} v_x \\ v_y \end{pmatrix}. \]
Translating a Rigid Body in $\mathbb{R}^n$

- Translate every point of $\Delta$ by $v$:

$$\Delta' = \{ p + v : p \in \Delta \}.$$

- For polygons and polytopes it suffices to translate the vertices.
Question

What is the matrix of a translation?
### Question

What is the matrix of a translation?

### Answer

No $n \times n$ matrix is the matrix of a (non-trivial) translation in $\mathbb{R}^n$!

- Why?
Translation as a Matrix Transformation

**Question**

What is the matrix of a translation?

**Answer**

No $n \times n$ matrix is the matrix of a (non-trivial) translation in $\mathbb{R}^n$!

- **Why?** Since the fixed point set of every matrix transformation includes the origin, but the origin is not invariant under a translation.
- **We will resort to homogeneous coordinates, which is a concept borrowed from projective geometry.**
Homogeneous Coordinates: Motivation

- A rational number $\frac{x}{y}$ is an equivalence class of appropriate pairs $(x', y')$.

- $2 \simeq (2, 1), (4, 2), \ldots$

- $1/3 \simeq (1/3, 1), (1, 3), (2, 6), \ldots$
Homogeneous Coordinates: Motivation

- A rational number $\frac{x}{y}$ is an equivalence class of appropriate pairs $(x', y')$.

- $2 \simeq (2, 1), (4, 2), \ldots$
- $1/3 \simeq (1/3, 1), (1, 3), (2, 6), \ldots$
- Not a unique representation: All points on a particular line through the origin represent the same rational number.
Homogeneous Coordinates: Motivation

- A rational number $\frac{x}{y}$ is an equivalence class of appropriate pairs $(x', y')$.

- $2 \simeq (2, 1), (4, 2), \ldots$
- $1/3 \simeq (1/3, 1), (1, 3), (2, 6), \ldots$
- Not a unique representation: All points on a particular line through the origin represent the same rational number.
- Canonical representative at the intersection of that line with the line $y = 1$. 
Homogeneous Coordinates: Motivation

- A rational number \( \frac{x}{y} \) is an equivalence class of appropriate pairs \((x', y')\).

\[
\frac{2}{1} \simeq (2, 1), (4, 2), \ldots
\]

\[
\frac{1}{3} \simeq (1/3, 1), (1, 3), (2, 6), \ldots
\]

- Not a unique representation: All points on a particular line through the origin represent the same rational number.
- Canonical representative at the intersection of that line with the line \( y = 1 \).
- Infinity does not need to be treated separately:
  \[
  \infty \simeq (1, 0), (2, 0), \ldots
  \]
Homogeneous Coordinates in $\mathbb{R}^2$

We identify the point $(x, y) \in \mathbb{R}^2$ with
$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \in \mathbb{R}^3$$
or with
$$\begin{pmatrix} w \\ w \cdot x \\ w \cdot y \end{pmatrix} \in \mathbb{R}^3$$
for $w \neq 0$.

Same for other points.
Homogeneous Coordinates in $\mathbb{R}^2$

- $\mathbb{R}^2$ is embedded into $\mathbb{R}^3$ by identifying it with the plane $z = 1$.

$\mathbb{R}^2$ is embedded into $\mathbb{R}^3$ by identifying it with the plane $z = 1$. 

Homogeneous Coordinates in $\mathbb{R}^2$

- $\mathbb{R}^2$ is embedded into $\mathbb{R}^3$ by identifying it with the plane $z = 1$.
- We identify the point $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ with $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \in \mathbb{R}^3$.

![Diagram showing the embedding of $\mathbb{R}^2$ into $\mathbb{R}^3$ with the $(z=1)$-plane.]
Homogeneous Coordinates in $\mathbb{R}^2$

- $\mathbb{R}^2$ is embedded into $\mathbb{R}^3$ by identifying it with the plane $z = 1$.
- We identify the point $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ with $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \in \mathbb{R}^3$ or with $\begin{pmatrix} w \cdot x \\ w \cdot y \\ w \end{pmatrix} \in \mathbb{R}^3$ for $w \neq 0$. 

$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in $\mathbb{R}^3$ representing a point in $\mathbb{R}^2$ with $z = 1$.
Homogeneous Coordinates in $\mathbb{R}^2$

- $\mathbb{R}^2$ is embedded into $\mathbb{R}^3$ by identifying it with the plane $z = 1$.
- We identify the point $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ with $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \in \mathbb{R}^3$ or with $\begin{pmatrix} x \\ y \\ w \end{pmatrix} \in \mathbb{R}^3$ for $w \neq 0$.

Same for other points.

(z=1)-plane

$0$
Homogeneous Coordinates in $\mathbb{R}^2$

- All points on a particular line through the origin in $\mathbb{R}^3$ represent the same point in $\mathbb{R}^2$.

\begin{equation*}
\begin{bmatrix}
x \\
y \\
0 \\
\end{bmatrix}
\end{equation*}
can be regarded as the point at infinity on the line through \begin{equation*}
\begin{bmatrix}
x \\
y \\
1 \\
\end{bmatrix}
\end{equation*}.

Homogeneous coordinates allow us to express translation, rotation and scaling in $\mathbb{R}^2$ by means of one $3 \times 3$ transformation matrix.

Homogeneous coordinates support scaling in a natural way, and build the basis of projective geometry.

Note that the plane $z = 1$ of $\mathbb{R}^3$ is invariant under matrix transformations of the form:

\begin{equation*}
\begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
0 & 0 & 1 \\
\end{bmatrix}
\end{equation*}.
Homogeneous Coordinates in $\mathbb{R}^2$

- All points on a particular line through the origin in $\mathbb{R}^3$ represent the same point in $\mathbb{R}^2$.
- $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ can be regarded as the point at infinity on the line through $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$.
Homogeneous Coordinates in $\mathbb{R}^2$

- All points on a particular line through the origin in $\mathbb{R}^3$ represent the same point in $\mathbb{R}^2$.
- $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ can be regarded as the point at infinity on the line through $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$.
- Homogeneous coordinates allow us to express translation, rotation, and scaling in $\mathbb{R}^2$ by means of one $3 \times 3$ transformation matrix.
- Homogeneous coordinates support scaling in a natural way, and build the basis of projective geometry.
Homogeneous Coordinates in $\mathbb{R}^2$

- All points on a particular line through the origin in $\mathbb{R}^3$ represent the same point in $\mathbb{R}^2$.
- $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ can be regarded as the point at infinity on the line through $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$.
- Homogeneous coordinates allow us to express translation, rotation and scaling in $\mathbb{R}^2$ by means of one $3 \times 3$ transformation matrix.
- Homogeneous coordinates support scaling in a natural way, and build the basis of projective geometry.
- Note that the plane $z = 1$ of $\mathbb{R}^3$ is invariant under matrix transformations of the form

$$
\begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
0 & 0 & 1
\end{pmatrix}.
$$
Definition 237 (Homogeneous coordinates, Dt.: homogene Koordinaten)

Homogeneous coordinates of \((x, y) \in \mathbb{R}^2\) are given by \(\begin{pmatrix} w \cdot x \\ w \\ w \end{pmatrix} \in \mathbb{R}^3\), for \(w \neq 0\), while the inhomogeneous coordinates of \(\begin{pmatrix} x \\ y \\ w \end{pmatrix} \in \mathbb{R}^3\) are given by \(\begin{pmatrix} x/w \\ y/w \end{pmatrix} \in \mathbb{R}^2\).
Homogeneous Coordinates in $\mathbb{R}^2$

**Definition 237 (Homogeneous coordinates, Dt.: homogene Koordinaten)**

Homogeneous coordinates of $(x, y) \in \mathbb{R}^2$ are given by \( \begin{pmatrix} w \cdot x \\ w \cdot y \\ w \end{pmatrix} \in \mathbb{R}^3 \), for $w \neq 0$, while the inhomogeneous coordinates of \( \begin{pmatrix} x \\ y \\ w \end{pmatrix} \in \mathbb{R}^3 \) are given by \( \begin{pmatrix} x/w \\ y/w \end{pmatrix} \in \mathbb{R}^2 \).

Thus, for $w \neq 0$, \( \begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathbb{R}^3 \) are homogeneous coordinates of \( \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \), and \( \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \) are the inhomogeneous coordinates of \( \begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathbb{R}^3 \).

\[ \iff x = \frac{u}{w} \text{ and } y = \frac{v}{w}. \]
Homogeneous Coordinates in $\mathbb{R}^2$

**Definition 237 (Homogeneous coordinates, Dt.: homogene Koordinaten)**

Homogeneous coordinates of $(x\ y) \in \mathbb{R}^2$ are given by $(w \cdot x \ w \cdot y \ w) \in \mathbb{R}^3$, for $w \neq 0$, while the inhomogeneous coordinates of $(x\ y \ w) \in \mathbb{R}^3$ are given by $(x/w \ y/w) \in \mathbb{R}^2$.

- Thus, for $w \neq 0$, $(u\ v \ w) \in \mathbb{R}^3$ are homogeneous coordinates of $(x\ y) \in \mathbb{R}^2$, and $(x\ y) \in \mathbb{R}^2$ are the inhomogeneous coordinates of $(u\ v \ w) \in \mathbb{R}^3$.

  $\Leftrightarrow \ x = \frac{u}{w}$ and $y = \frac{v}{w}$.

- We will find it convenient to assume $w = 1$. 

© M. Held (Univ. Salzburg)  Geometrisches Rechnen (WS 2020/21) 256/357
Transformation Matrices Based on Homogeneous Coordinates for $\mathbb{R}^2$

Translation:

\[
\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & v_x \\ 0 & 1 & v_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.
\]

We get $\text{Trans}(v_x, v_y)^{-1} = \text{Trans}(-v_x, -v_y)$. 

Stretching:

\[
\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.
\]

We get $\text{S}(\lambda_1, \lambda_2)^{-1} = \text{S}(1/\lambda_1, 1/\lambda_2)$. 

Transformation Matrices Based on Homogeneous Coordinates for $\mathbb{R}^2$

Translation:

$$
\begin{pmatrix}
  x' \\
  y' \\
  1
\end{pmatrix}
= 
\begin{pmatrix}
  1 & 0 & v_x \\
  0 & 1 & v_y \\
  0 & 0 & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
  x \\
  y \\
  1
\end{pmatrix}
\cdot
\text{Trans}(v_x, v_y)
$$

We get

$$
\text{Trans}(v_x, v_y)^{-1} = \text{Trans}(-v_x, -v_y).
$$

Stretching:

$$
\begin{pmatrix}
  x' \\
  y' \\
  1
\end{pmatrix}
= 
\begin{pmatrix}
  \lambda_1 & 0 & 0 \\
  0 & \lambda_2 & 0 \\
  0 & 0 & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
  x \\
  y \\
  1
\end{pmatrix}
\cdot
\text{S}(\lambda_1, \lambda_2)
$$

We get

$$
\text{S}(\lambda_1, \lambda_2)^{-1} = \text{S}(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}).
$$
Transformation Matrices Based on Homogeneous Coordinates for $\mathbb{R}^2$

Rotation:

\[
\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.
\]

We get $\text{Rot}(\theta)^{-1} = \text{Rot}(-\theta) = \text{Rot}(\theta)^t$. 

Rotation involves either trigonometric functions or square roots. Power series may be used to approximate the terms of a rotation matrix for small values of $\theta$. 

© M. Held (Univ. Salzburg) 
Geometrisches Rechnen (WS 2020/21)
Rotation:

\[
\begin{pmatrix}
  x' \\
  y' \\
  1
\end{pmatrix}
= \begin{pmatrix}
  \cos \theta & -\sin \theta & 0 \\
  \sin \theta & \cos \theta & 0 \\
  0 & 0 & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
  x \\
  y \\
  1
\end{pmatrix}
\]

\text{Rot}(\theta)

We get \(\text{Rot}(\theta)^{-1} = \text{Rot}(-\theta) = \text{Rot}(\theta)^t\).

- Rotation involves either trigonometric functions or square roots.
- Power series may be used to approximate the terms of a rotation matrix for small values of \(\theta\).
Transformations

- Linear Transformations
- Classification of Transformations
- Coordinate Transformations in $\mathbb{R}^2$
- Coordinate Transformations in $\mathbb{R}^3$
  - Rotation in $\mathbb{R}^3$
  - Transformation Matrices for $\mathbb{R}^3$
  - Linear Transformations and Eigenvectors
- Transformation of Coordinate Systems
- Applications of Coordinate (System) Transformations
- Rotations Revisited
- Projections
Homogeneous Coordinates and Transformations in $\mathbb{R}^3$

- Homogeneous coordinates in $\mathbb{R}^3$:

\[(x, y, z, w) \sim \left( \frac{x}{w}, \frac{y}{w}, \frac{z}{w} \right).\]
Homogeneous Coordinates and Transformations in $\mathbb{R}^3$

- Homogeneous coordinates in $\mathbb{R}^3$:
  \[(x, y, z, w) \sim \left( \frac{x}{w}, \frac{y}{w}, \frac{z}{w} \right).\]

- For a right-hand coordinate system the positive (CCW) rotation about a coordinate axis is defined as follows:
  - Look along the axis towards the origin from $+\infty$;
  - Counter-clockwise rotation about axis by angle $\pi/2$ transforms one axis to another, obeying the cyclic order $x \rightarrow y \rightarrow z \rightarrow x$. 

![Diagram of coordinate system with rotations indicated]
Rotation about $z$-Axis

A rotation about the $z$-axis can be regarded as a rotation in $\mathbb{R}^2$ about the origin that is extended to $\mathbb{R}^3$. That is,

\[
x' = x \cos \theta - y \sin \theta,
\]
\[
y' = x \sin \theta + y \cos \theta,
\]
\[
z' = z.
\]
Rotation about $x$-Axis

Rotation about the $x$-axis: Substitute $x \rightarrow y$, $y \rightarrow z$, $z \rightarrow x$ in the equations for the rotation about $z$.

\[
y' = y \cos \theta - z \sin \theta, \\
z' = y \sin \theta + z \cos \theta, \\
x' = x.
\]
Rotation about $y$-Axis

- Similarly for a rotation about the $y$-axis: Substitute $x \rightarrow y$, $y \rightarrow z$, $z \rightarrow x$ in the previous equations.

\[
\begin{align*}
    z' &= z \cos \theta - x \sin \theta, \\
    x' &= z \sin \theta + x \cos \theta, \\
    y' &= y.
\end{align*}
\]
Transformation Matrices for $\mathbb{R}^3$

Rotation (about $x$-Axis):

\[
\begin{pmatrix}
x' \\
y' \\
z' \\
1
\end{pmatrix} =
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \phi & -\sin \phi & 0 \\
0 & \sin \phi & \cos \phi & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
x \\
y \\
z \\
1
\end{pmatrix}
\]

Rotation (about $y$-Axis):

\[
\begin{pmatrix}
x' \\
y' \\
z' \\
1
\end{pmatrix} =
\begin{pmatrix}
\cos \phi & 0 & \sin \phi & 0 \\
0 & 1 & 0 & 0 \\
-\sin \phi & 0 & \cos \phi & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
x \\
y \\
z \\
1
\end{pmatrix}
\]

Rotation (about $z$-Axis):

\[
\begin{pmatrix}
x' \\
y' \\
z' \\
1
\end{pmatrix} =
\begin{pmatrix}
\cos \phi & -\sin \phi & 0 & 0 \\
\sin \phi & \cos \phi & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
x \\
y \\
z \\
1
\end{pmatrix}
\]
Transformation Matrices for $\mathbb{R}^3$

Translation:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & v_x \\ 0 & 1 & 0 & v_y \\ 0 & 0 & 1 & v_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Stretching/Scaling:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$
Linear Transformations and Eigenvectors

Question: How can we find the axis of rotation (through the origin) if we only know the rotation matrix $T$?

Answer: Since all points on the axis of rotation are invariant under the rotation, it suffices to look for a non-zero vector $v$ such that $T v = v$, i.e., for an eigenvector of $T$ with eigenvalue 1 since rotations never stretch or squish anything.

Question: How can we determine the plane of reflection (through the origin) if we only know the transformation matrix $T$?

Answer: It suffices to look for two (linearly independent) eigenvectors $u$, $v$. These two vectors span the plane sought.
Question: How can we find the axis of rotation (through the origin) if we only know the rotation matrix $T$?

Answer: Since all points on the axis of rotation are invariant under the rotation, it suffices to look for a non-zero vector $v$ such that

$$Tv = v,$$

i.e., for an eigenvector of $T$ with eigenvalue 1 since rotations never stretch or squish anything.
Linear Transformations and Eigenvectors

- Question: How can we find the axis of rotation (through the origin) if we only know the rotation matrix $T$?

- Answer: Since all points on the axis of rotation are invariant under the rotation, it suffices to look for a non-zero vector $v$ such that

$$Tv = v,$$

i.e., for an eigenvector of $T$ with eigenvalue 1 since rotations never stretch or squish anything.

- Question: How can we determine the plane of reflection (through the origin) if we only know the transformation matrix $T$?
Question: How can we find the axis of rotation (through the origin) if we only know the rotation matrix $T$?

Answer: Since all points on the axis of rotation are invariant under the rotation, it suffices to look for a non-zero vector $v$ such that

$$Tv = v,$$

i.e., for an eigenvector of $T$ with eigenvalue 1 since rotations never stretch or squish anything.

Question: How can we determine the plane of reflection (through the origin) if we only know the transformation matrix $T$?

Answer: It suffices to look for two (linearly independent) eigenvectors $u, v$. These two vectors span the plane sought.
Transformations

- Linear Transformations
- Classification of Transformations
- Coordinate Transformations in \( \mathbb{R}^2 \)
- Coordinate Transformations in \( \mathbb{R}^3 \)
- Transformation of Coordinate Systems
  - Mathematics of Coordinate System Transformations
  - Inverse Transformation
  - Sample Transformation
- Applications of Coordinate (System) Transformations
- Rotations Revisited
- Projections
Transformation of Coordinate Systems

- Space has no intrinsic coordinate system!
- Basis vectors need not have unit length.
Transformation of Coordinate Systems

- Space has no intrinsic coordinate system!
- Basis vectors need not have unit length.
- Hence, a point will have different coordinates in different coordinate systems of the same vector space.
Transformation of Coordinate Systems

- Space has no intrinsic coordinate system!
- Basis vectors need not have unit length.
- Hence, a point will have different coordinates in different coordinate systems of the same vector space.
- E.g., $C := [e_1, e_2]$ is not the only possible basis for $\mathbb{R}^2$: 

$$C := [e_1, e_2]$$

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix}$$

Our next task is to convert between different coordinate systems.
Transformation of Coordinate Systems

- Space has no intrinsic coordinate system!
- Basis vectors need not have unit length.
- Hence, a point will have different coordinates in different coordinate systems of the same vector space.

E.g., \( C := [e_1, e_2] \) is not the only possible basis for \( \mathbb{R}^2 \):

\[
\begin{pmatrix}
2 \\
3
\end{pmatrix}
_{[e_1, e_2]} =
\begin{pmatrix}
2 \\
1
\end{pmatrix}
_{[v, w]}
\]
Transformation of Coordinate Systems

- Space has no intrinsic coordinate system!
- Basis vectors need not have unit length.
- Hence, a point will have different coordinates in different coordinate systems of the same vector space.
- E.g., $C := [e_1, e_2]$ is not the only possible basis for $\mathbb{R}^2$: \[
\begin{pmatrix}
2 \\
3
\end{pmatrix}
[e_1, e_2] = \begin{pmatrix}
2 \\
1
\end{pmatrix}[v, w]
\]
- Our next task is to convert between different coordinate systems.
Transformation of Coordinate Systems

- So, what are the coordinates $p_{c'} := \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$ of a point $p_c := \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ relative to a new coordinate system $C'$?
We assume that the mapping from $C$ to $C'$ is an isometry.
Transformation of Coordinate Systems

- We assume that the mapping from $C$ to $C'$ is an isometry.
- Consider an untranslated copy $C''$ of $C'$ whose axes vectors are identical but whose origin $0''$ is at the origin of $C$. That is, $x'' \parallel x'$ and $y'' \parallel y'$ and $z'' \parallel z'$. 

\[
\begin{array}{c}
\begin{pmatrix}
e'_{1} & e'_{2} & e'_{3} \\
\delta & 0 & 0 \\
0 & 0 & 1
\end{pmatrix}
\end{array}
\]

where $e'_{1}$ represents the unit vector of the $x''$-axis of $C''$ in terms of $C$. Of course, $e'_{1}$ is also the unit vector of the $x'$-axis of $C'$. Analogously for $e'_{2}$, $e'_{3}$.
Transformation of Coordinate Systems

- We assume that the mapping from $C$ to $C'$ is an isometry.
- Consider an untranslated copy $C''$ of $C'$ whose axes vectors are identical but whose origin $0''$ is at the origin of $C$. That is, $x'' \parallel x'$ and $y'' \parallel y'$ and $z'' \parallel z'$.
- We construct the matrix

$$T_C := \begin{pmatrix} e'_1 & e'_2 & e'_3 & \delta \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $e'_1$ represents the unit vector of the $x''$-axis of $C''$ in terms of $C$. Of course, $e'_1$ is also the unit vector of the $x'$-axis of $C'$. Analogously for $e'_2, e'_3$. 

\[ z'' \]
\[ C'' \]
\[ x'' \]
\[ y'' \]
\[ z' \]
\[ C' \]
\[ x' \]
\[ y' \]
\[ \delta \]
\[ 0'' \]
\[ 0' \]
Transformation of Coordinate Systems

- We assume that the mapping from \( C \) to \( C' \) is an isometry.
- Consider an untranslated copy \( C'' \) of \( C' \) whose axes vectors are identical but whose origin \( 0'' \) is at the origin of \( C \). That is, \( x'' \parallel x' \) and \( y'' \parallel y' \) and \( z'' \parallel z' \).
- We construct the matrix

\[
T_C := \begin{pmatrix}
e'_1 & e'_2 & e'_3 & \delta \\
0 & 0 & 0 & 1
\end{pmatrix},
\]

where \( e'_1 \) represents the unit vector of the \( x'' \)-axis of \( C'' \) in terms of \( C \). Of course, \( e'_1 \) is also the unit vector of the \( x' \)-axis of \( C' \). Analogously for \( e'_2, e'_3 \).
- We know that \([e'_1, e'_2, e'_3]\) is an orthogonal matrix if \( e_1, e_2, e_3 \) are orthonormal.
Transformation of Coordinate Systems

- We have
  \[
  \begin{pmatrix}
  1 \\
  0 \\
  0 \\
  1 
  \end{pmatrix}
  \mapsto
  \begin{pmatrix}
  \mathbf{e}'_1 + \delta \\
  1
  \end{pmatrix},
  \]

  \[
  \begin{pmatrix}
  0 \\
  1 \\
  0 \\
  1 
  \end{pmatrix}
  \mapsto
  \begin{pmatrix}
  \mathbf{e}'_2 + \delta \\
  1
  \end{pmatrix},
  \]

  \[
  \begin{pmatrix}
  0 \\
  0 \\
  1 \\
  1 
  \end{pmatrix}
  \mapsto
  \begin{pmatrix}
  \mathbf{e}'_3 + \delta \\
  1
  \end{pmatrix},
  \]

  \[
  \begin{pmatrix}
  0 \\
  0 \\
  0 \\
  1 
  \end{pmatrix}
  \mapsto
  \begin{pmatrix}
  \delta \\
  1
  \end{pmatrix},
  \]

  that is
  \[
  \begin{pmatrix}
  x' \\
  y' \\
  z' \\
  1
  \end{pmatrix}
  \mapsto
  \begin{pmatrix}
  x' \mathbf{e}'_1 + y' \mathbf{e}'_2 + z' \mathbf{e}'_3 + \delta \\
  1
  \end{pmatrix},
  \]

  \[
  =
  \begin{pmatrix}
  x \\
  y \\
  z \\
  1
  \end{pmatrix}
  \]
Transformation of Coordinate Systems

We have

\[
\begin{pmatrix}
1 \\
0 \\
0 \\
1
\end{pmatrix} \mapsto_{T_C} \begin{pmatrix} e_1' + \delta \\
1 \end{pmatrix},
\]

\[
\begin{pmatrix}
0 \\
1 \\
0 \\
1
\end{pmatrix} \mapsto_{T_C} \begin{pmatrix} e_2' + \delta \\
1 \end{pmatrix},
\]

that is

\[
\begin{pmatrix}
x' \\
y' \\
z' \\
1
\end{pmatrix} \mapsto_{T_C} (x'e_1' + ye_2' + ze_3' + \delta_1) =: \begin{pmatrix}
x \\
y \\
z \\
1
\end{pmatrix}.
\]

We understand that the coordinates of a point specified relative to \(C'\) are converted by \(T_C\) to coordinates relative to \(C\):

**Theorem 238**

With \(T_C\) as defined on the previous slide, we get

\[
p_C = T_C \cdot p_{C'}
\]

and

\[
p_{C'} = T_{C}^{-1} \cdot p_C.
\]
Transformation of Coordinate Systems

We have

\[
\begin{pmatrix}
1 & 0 \\
0 & 0 \\
0 & 1
\end{pmatrix}
\mapsto_{T_C}
\begin{pmatrix}
e'_1 + \delta \\
1
\end{pmatrix},
\begin{pmatrix}
0 & 1 \\
0 & 0 \\
1 & 1
\end{pmatrix}
\mapsto_{T_C}
\begin{pmatrix}
e'_2 + \delta \\
1
\end{pmatrix},
\begin{pmatrix}
0 & 0 \\
0 & 1 \\
1 & 1
\end{pmatrix}
\mapsto_{T_C}
\begin{pmatrix}
e'_3 + \delta \\
1
\end{pmatrix},
\begin{pmatrix}
0 & 0 \\
0 & 1 \\
1 & 1
\end{pmatrix}
\mapsto_{T_C}
\begin{pmatrix}
\delta \\
1
\end{pmatrix},
\]

We understand that the coordinates of a point specified relative to \(C'\) are converted by \(T_C\) to coordinates relative to \(C\):

Theorem 238
With \(T_C\) as defined on the previous slide, we get

\[p_C = T_C \cdot p_{C'}\]
and

\[p_{C'} = T_{C^{-1}} \cdot p_C.\]
Transformation of Coordinate Systems

We have

\[
\begin{pmatrix}
1 \\
0 \\
0 \\
1
\end{pmatrix}
\xrightarrow{T_C} \begin{pmatrix}
e'_1 + \delta \\
1
\end{pmatrix},
\begin{pmatrix}
0 \\
1 \\
0 \\
1
\end{pmatrix}
\xrightarrow{T_C} \begin{pmatrix}
e'_2 + \delta \\
1
\end{pmatrix},
\begin{pmatrix}
0 \\
0 \\
1 \\
1
\end{pmatrix}
\xrightarrow{T_C} \begin{pmatrix}
e'_3 + \delta \\
1
\end{pmatrix},
\begin{pmatrix}
0 \\
0 \\
0 \\
1
\end{pmatrix}
\xrightarrow{T_C} \begin{pmatrix}
\delta \\
1
\end{pmatrix},
\]

that is

\[
\begin{pmatrix}
x' \\
y' \\
z'
\end{pmatrix}
\xrightarrow{T_C} \begin{pmatrix}
x' e'_1 + y' e'_2 + z' e'_3 + \delta \\
1
\end{pmatrix} =: \begin{pmatrix}
x \\
y \\
z
\end{pmatrix}_C.
\]
Transformation of Coordinate Systems

- We have
  \[
  \begin{pmatrix}
  1 & 0 \\
  0 & 0 \\
  0 & 1 \\
  1 & 
  \end{pmatrix}
  \xrightarrow{T_C} \begin{pmatrix}
  e'_1 + \delta \\
  1 
  \end{pmatrix},
  \begin{pmatrix}
  0 & 1 \\
  0 & 0 \\
  0 & 1 \\
  1 & 
  \end{pmatrix}
  \xrightarrow{T_C} \begin{pmatrix}
  e'_2 + \delta \\
  1 
  \end{pmatrix},
  \begin{pmatrix}
  0 & 0 \\
  0 & 1 \\
  0 & 1 \\
  1 & 
  \end{pmatrix}
  \xrightarrow{T_C} \begin{pmatrix}
  e'_3 + \delta \\
  1 
  \end{pmatrix},
  \begin{pmatrix}
  0 & 0 \\
  0 & 1 \\
  0 & 1 \\
  1 & 
  \end{pmatrix}
  \xrightarrow{T_C} \begin{pmatrix}
  \delta \\
  1 
  \end{pmatrix},
  \]

- That is
  \[
  \begin{pmatrix}
  x' \\
  y' \\
  z' \\
  1 
  \end{pmatrix}
  \xrightarrow{T_C} \begin{pmatrix}
  x' e'_1 + y' e'_2 + z' e'_3 + \delta \\
  1 
  \end{pmatrix} =: \begin{pmatrix}
  x \\
  y \\
  z \\
  1 
  \end{pmatrix}_{C}.
  \]

- We understand that the coordinates of a point specified relative to \(C'\) are converted by \(T_C\) to coordinates relative to \(C\):
Transformation of Coordinate Systems

- We have
  \[
  \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \xrightarrow{T_C} \begin{pmatrix} e'_1 + \delta \\ 1 \end{pmatrix},
  \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \xrightarrow{T_C} \begin{pmatrix} e'_2 + \delta \\ 1 \end{pmatrix},
  \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \xrightarrow{T_C} \begin{pmatrix} e'_3 + \delta \\ 1 \end{pmatrix},
  \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \xrightarrow{T_C} \begin{pmatrix} \delta \\ 1 \end{pmatrix},
  \]

that is
  \[
  \begin{pmatrix} \chi' \\ y' \\ z' \\ 1 \end{pmatrix} \xrightarrow{T_C} \begin{pmatrix} x'e'_1 + y'e'_2 + z'e'_3 + \delta \\ 1 \end{pmatrix} =: \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}_C.
  \]

- We understand that the coordinates of a point specified relative to \( C' \) are converted by \( T_C \) to coordinates relative to \( C \):

**Theorem 238**

With \( T_C \) as defined on the previous slide, we get

\[ p_C = T_C \cdot p_C' \quad \text{and} \quad p_C' = T_C^{-1} \cdot p_C. \]
Inverse Transformation

If $T$ is the matrix of an isometry then, by Lem 224,

$$
T = \begin{pmatrix}
1 & 0 & 0 & | & d \\
0 & 1 & 0 & | & 0 \\
0 & 0 & 1 & | & 0 \\
0 & 0 & 0 & | & 1
\end{pmatrix} \cdot \begin{pmatrix}
R & | & 0 \\
0 & | & 0 \\
0 & | & 0 \\
0 & | & 1
\end{pmatrix},
$$

where $R$ is an orthogonal matrix, and $d$ describes the translation.
Inverse Transformation

- If $T$ is the matrix of an isometry then, by Lem 224,
  
  \[
  T = \begin{pmatrix}
  1 & 0 & 0 & d \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1
  \end{pmatrix} \cdot \begin{pmatrix}
  R & 0 \\
  0 & 0 \\
  0 & 0 \\
  0 & 0
  \end{pmatrix},
  \]

  where $R$ is an orthogonal matrix, and $d$ describes the translation.

- Since $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$, we get
  
  \[
  T^{-1} = \begin{pmatrix}
  R^{-1} & 0 \\
  0 & 0 \\
  0 & 0 \\
  0 & 0
  \end{pmatrix} \cdot \begin{pmatrix}
  1 & 0 & 0 & -d \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1
  \end{pmatrix}.
  \]
Inverse Transformation

- If $T$ is the matrix of an isometry then, by Lem 224,
  \[
  T = \begin{pmatrix}
    1 & 0 & 0 & d \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1 \\
  \end{pmatrix}
  \cdot
  \begin{pmatrix}
    R & 0 \\
    0 & 0 \\
    0 & 0 \\
    0 & 0 \\
  \end{pmatrix},
  \]
  where $R$ is an orthogonal matrix, and $d$ describes the translation.

- Since $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$, we get
  \[
  T^{-1} = \begin{pmatrix}
    R^{-1} & 0 \\
    0 & 0 \\
    0 & 0 \\
    0 & 0 \\
  \end{pmatrix}
  \cdot
  \begin{pmatrix}
    1 & 0 & 0 & -d \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1 \\
  \end{pmatrix}.
  \]

- Since $R$ is orthogonal, we have $R^{-1} = R^t$ and get
  \[
  T^{-1} = \begin{pmatrix}
    R^t & 0 \\
    0 & 0 \\
    0 & 0 \\
    0 & 0 \\
  \end{pmatrix}
  \cdot
  \begin{pmatrix}
    1 & 0 & 0 & -d \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1 \\
  \end{pmatrix}.
  \]
Inverse Transformation

**Theorem 239**

If \([n, o, a]\) is orthogonal then we get

\[
T^{-1} = \begin{pmatrix}
  n_x & n_y & n_z & -\langle d, n \rangle \\
  o_x & o_y & o_z & -\langle d, o \rangle \\
  a_x & a_y & a_z & -\langle d, a \rangle \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]

for

\[
T := \begin{pmatrix}
  n_x & o_x & a_x & d_x \\
  n_y & o_y & a_y & d_y \\
  n_z & o_z & a_z & d_z \\
  0 & 0 & 0 & 1
\end{pmatrix}.
\]
Inverse Transformation

**Theorem 239**

If \([n, o, a]\) is orthogonal then we get

\[
T^{-1} = \begin{pmatrix}
  n_x & n_y & n_z & -\langle d, n \rangle \\
  o_x & o_y & o_z & -\langle d, o \rangle \\
  a_x & a_y & a_z & -\langle d, a \rangle \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]

for

\[
T := \begin{pmatrix}
  n_x & o_x & a_x & d_x \\
  n_y & o_y & a_y & d_y \\
  n_z & o_z & a_z & d_z \\
  0 & 0 & 0 & 1
\end{pmatrix}.
\]

- Recall that the matrix of a general affine transformation is not orthogonal!
Sample Coordinate System Transformation

- For the scenario shown below we get

\[ T = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 4 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and, thus,} \quad T^{-1} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & -3\sqrt{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \]
Sample Coordinate System Transformation

For the scenario shown below we get

\[
T = \begin{pmatrix}
\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 4 \\
\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 2 \\
0 & 0 & 1 \\
\end{pmatrix}
\quad \text{and, thus,} \quad
T^{-1} = \begin{pmatrix}
\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & -3\sqrt{2} \\
-\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\
0 & 0 & 1 \\
\end{pmatrix}.
\]

Hence,

\[
T^{-1} \cdot p_C = T^{-1} \cdot \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 1 \end{pmatrix} = p_{C'} \quad \text{and} \quad
T \cdot p_{C'} = T \cdot \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ 1 \end{pmatrix} = p_C.
\]
Transformations

- Linear Transformations
- Classification of Transformations
- Coordinate Transformations in $\mathbb{R}^2$
- Coordinate Transformations in $\mathbb{R}^3$
- Transformation of Coordinate Systems

Applications of Coordinate (System) Transformations

- Rotation About a General Axis
- Local Coordinate Systems
- Kinematics
- Rotations Revisited
- Projections
Rotation About a General Axis

- What is the matrix of the rotation about a line \( \ell \) (through the origin) with direction vector \( u \) by an angle \( \phi \)?
Rotation About a General Axis

- What is the matrix of the rotation about a line $\ell$ (through the origin) with direction vector $u$ by an angle $\phi$?

- We set up a new frame $\mathcal{C}' = [e'_1, e'_2, e'_3]$ such that
  - $0 = 0'$,
Rotation About a General Axis

- What is the matrix of the rotation about a line $\ell$ (through the origin) with direction vector $u$ by an angle $\phi$?

- We set up a new frame $C' = [e'_1, e'_2, e'_3]$ such that
  - $0 = 0'$,
  - $e'_3 = u/\|u\|$,
Rotation About a General Axis

- What is the matrix of the rotation about a line $\ell$ (through the origin) with direction vector $u$ by an angle $\phi$?

- We set up a new frame $C' = [e'_1, e'_2, e'_3]$ such that
  - $0 = 0'$,
  - $e'_3 = u/\|u\|$, 
  - $\langle e'_2, e'_3 \rangle = 0$ and $\|e'_2\| = 1$, 

Rotation About a General Axis

- What is the matrix of the rotation about a line $\ell$ (through the origin) with direction vector $u$ by an angle $\phi$?

![Diagram of rotation about a general axis]

- We set up a new frame $C' = [e'_1, e'_2, e'_3]$ such that
  - $0 = 0'$,
  - $e'_3 = u/||u||$,
  - $\langle e'_2, e'_3 \rangle = 0$ and $||e'_2|| = 1$,
  - $e'_1 := e'_2 \times e'_3$. 

Rotation About a General Axis

What is the matrix of the rotation about a line $\ell$ (through the origin) with direction vector $u$ by an angle $\phi$?

We set up a new frame $C' = [e'_1, e'_2, e'_3]$ such that

- $0 = 0'$,
- $e'_3 = u/||u||$,
- $\langle e'_2, e'_3 \rangle = 0$ and $||e'_2|| = 1$,
- $e'_1 := e'_2 \times e'_3$.

We know that $||e'_1|| = 1$
Rotation About a General Axis

- What is the matrix of the rotation about a line $\ell$ (through the origin) with direction vector $u$ by an angle $\phi$?

- We set up a new frame $C' = [e'_1, e'_2, e'_3]$ such that
  - $0 = 0'$,
  - $e'_3 = u/||u||$,
  - $\langle e'_2, e'_3 \rangle = 0$ and $||e'_2|| = 1$,
  - $e'_1 := e'_2 \times e'_3$.
- We know that $||e'_1|| = 1$ and consider the transformation matrix

$$T := \begin{pmatrix}
  e'_1 & e'_2 & e'_3 & 0 \\
  0 & 0 & 0 & 1
\end{pmatrix}.$$
Rotation About a General Axis

We know that
\[
\begin{pmatrix}
x' \\
y' \\
z'
\end{pmatrix} = T^{-1} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}.
\]
Rotation About a General Axis

We know that
\[
\begin{pmatrix}
x' \\
y' \\
z' \\
1
\end{pmatrix} = T^{-1} \cdot \begin{pmatrix}
x \\
y \\
z \\
1
\end{pmatrix}.
\]

Thus, we get the following decomposition for \(\text{Rot}(u, \phi)\):
\[
\text{Rot}(u, \phi) = T \cdot \begin{pmatrix}
\cos \phi & -\sin \phi & 0 & 0 \\
\sin \phi & \cos \phi & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix} \cdot T^{-1}
\]
from \(C'\) to \(C\)
rotation about \(z'\)-axis
Rotation About a General Axis

- We know that \[
\begin{pmatrix}
    x' \\
    y' \\
    z' \\
    1
\end{pmatrix} = T^{-1} \cdot \begin{pmatrix}
    x \\
    y \\
    z \\
    1
\end{pmatrix}.
\]

- Thus, we get the following decomposition for \(\text{Rot}(u, \phi)\):

\[
\text{Rot}(u, \phi) = \underbrace{T}_\text{from } C' \text{ back to } C \cdot \begin{pmatrix}
    \cos \phi & -\sin \phi & 0 & 0 \\
    \sin \phi & \cos \phi & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
\end{pmatrix} \cdot \underbrace{T^{-1}}_\text{from } C \text{ to } C'.
\]

- Simple algebraic operations yield

\[
\text{Rot}(u, \phi) := \begin{pmatrix}
    u_x u_x \text{ vers } \phi + \cos \phi & u_y u_x \text{ vers } \phi - u_z \sin \phi & u_z u_x \text{ vers } \phi + u_y \sin \phi & 0 \\
    u_x u_y \text{ vers } \phi + u_z \sin \phi & u_y u_y \text{ vers } \phi + \cos \phi & u_z u_y \text{ vers } \phi - u_x \sin \phi & 0 \\
    u_x u_z \text{ vers } \phi - u_y \sin \phi & u_y u_z \text{ vers } \phi + u_x \sin \phi & u_z u_z \text{ vers } \phi + \cos \phi & 0 \\
    0 & 0 & 0 & 1
\end{pmatrix}
\]

where \(\text{vers } \phi := 1 - \cos \phi\).
Rotation About a General Axis

- Given an (orthogonal) rotation matrix $T$, how can we find an axis $u$ through the origin and an angle $\phi$ such that $\text{Rot}(u, \phi) = T$?

$$\text{Rot}(u, \phi) = T := \begin{pmatrix} n_x & o_x & a_x & 0 \\ n_y & o_y & a_y & 0 \\ n_z & o_z & a_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$
Rotation About a General Axis

Given an (orthogonal) rotation matrix $T$, how can we find an axis $u$ through the origin and an angle $\phi$ such that $\text{Rot}(u, \phi) = T$?

\[
\text{Rot}(u, \phi) \equiv T := \begin{pmatrix}
    n_x & o_x & a_x & 0 \\
    n_y & o_y & a_y & 0 \\
    n_z & o_z & a_z & 0 \\
    0 & 0 & 0 & 1
\end{pmatrix}.
\]

Some calculations yield

\[
\tan \phi = \frac{\sqrt{(o_z - a_y)^2 + (a_x - n_z)^2 + (n_y - o_x)^2}}{n_x + o_y + a_z - 1},
\]

which defines $\phi$ within $[0, \pi]$. 
Rotation About a General Axis

Furthermore,

\[ u_x = \text{sign}(o_z - a_y) \sqrt{\frac{n_x - \cos \phi}{1 - \cos \phi}}, \]

\[ u_y = \text{sign}(a_x - n_z) \sqrt{\frac{o_y - \cos \phi}{1 - \cos \phi}}, \]

\[ u_z = \text{sign}(n_y - o_x) \sqrt{\frac{a_z - \cos \phi}{1 - \cos \phi}}. \]
Local Coordinate Systems

- Typically, objects are not modeled in world coordinates. Rather, *local coordinate systems* are used.
Local Coordinate Systems

- Typically, objects are not modeled in world coordinates. Rather, *local coordinate systems* are used.
- In order to transform the object it suffices to fix the position and orientation of the local coordinate system relative to the world coordinate system, or relative to some other system.
Kinematics

- We consider an articulated mechanism that consists of rigid links connected by joints.
- Every joint connects exactly two links, and describes the motion of one link relative to the other link.
Kinematics

- We consider an articulated mechanism that consists of rigid links connected by joints.
- Every joint connects exactly two links, and describes the motion of one link relative to the other link.
- The most important joints are prismatic and rotatory joints.
A mechanism can be represented as a graph, a so-called *kinematic chain*, where
- the links form the nodes, and
- the joints from the edges.

A mechanism is called an *open kinematic chain* if this graph has no cycles; *closed kinematic chain*, otherwise.
Kinematic Chain

- A mechanism can be represented as a graph, a so-called *kinematic chain*, where
  - the links form the nodes, and
  - the joints from the edges.

- A mechanism is called an *open kinematic chain* if this graph has no cycles; *closed kinematic chain*, otherwise.

- Depending on how detailed a human is modeled, a human skeleton represents either an open or a closed kinematic chain.
Local Coordinate Frames

- It is common to assign two local coordinate frames $F_{i1}$ and $F_{i2}$ to link $i$ such that
  - the $z$-axis coincides with the joint axis,
  - the $x$-axis coincides with the link axis, and
  - the $y$-axis is chosen appropriately to form a right-handed frame.
Denavit-Hartenberg Parameters

- Find a transformation matrix $i_{i-1}A$ to express a point of $F_{i,2}$ in terms of $F_{i-1,2}$.
Denavit-Hartenberg Parameters

- Find a transformation matrix $i^{-1}A$ to express a point of $F_{i,2}$ in terms of $F_{i-1,2}$.

- **A-Matrix:**

$$
\begin{align*}
{i^{-1}A} &:= \text{Rot}(z, \theta) \cdot \text{Trans}(0, 0, d) \cdot \text{Trans}(a, 0, 0) \cdot \text{Rot}(x, \alpha) \\
&= \begin{pmatrix}
\cos \theta & -\sin \theta \cos \alpha & \sin \theta \sin \alpha & a \cos \theta \\
\sin \theta & \cos \theta \cos \alpha & -\cos \theta \sin \alpha & a \sin \theta \\
0 & \sin \alpha & \cos \alpha & d \\
0 & 0 & 0 & 1
\end{pmatrix},
\end{align*}
$$

where

- $a$ \ldots link length,
- $\alpha$ \ldots link twist,
- $d$ \ldots link offset,
- $\theta$ \ldots link angle,

\textit{Denavit-Hartenberg Parameters.}
Forward and Inverse Kinematics

Forward Kinematics:
- Given: joint vector.
- Compute: Frame $T$ of the end-effector relative to the base frame.
- Solution:

$$T = 0A \cdot 1A \cdot \ldots \cdot n^{-1}A.$$
Forward and Inverse Kinematics

Forward Kinematics:
- Given: joint vector.
- Compute: Frame $\mathbf{T}$ of the end-effector relative to the base frame.
- Solution:

\[
\mathbf{T} = 0^1 \mathbf{A} \cdot 1^2 \mathbf{A} \cdot \ldots \cdot n^{-1} \mathbf{A}.
\]

Inverse Kinematics:
- Given: Frame $\mathbf{T}$ of the end-effector relative to the base frame.
- Compute: all admissible joint vectors.
- Solution: not trivial, requires solving a set of non-linear equations!
Symbolic solution preferred over numerical solution.
Inverse Kinematics

- Truly all admissible joint vectors have to be computed!
Transformations

- Linear Transformations
- Classification of Transformations
- Coordinate Transformations in $\mathbb{R}^2$
- Coordinate Transformations in $\mathbb{R}^3$
- Transformation of Coordinate Systems
- Applications of Coordinate (System) Transformations

- Rotations Revisited
  - Linear Transformations and Eigenvectors
  - Rotation Group
  - Quaternions and Rotations

- Projections
Geometric Interpretation of Eigenvectors

- Recall Def. 117: A vector $v \in \mathbb{R}^n$ is an eigenvector of the $n \times n$ matrix $A$ if
  \[ A v = \lambda v \quad \text{and} \quad v \neq 0. \]
Geometric Interpretation of Eigenvectors

- Recall Def. 117: A vector $v \in \mathbb{R}^n$ is an eigenvector of the $n \times n$ matrix $A$ if
  
  $$Av = \lambda v \quad \text{and} \quad v \neq 0.$$

- Geometrically, the vector $Au$ is some vector of $\mathbb{R}^n$ obtained by applying a linear transformation $g$, whose matrix equals $A$, to $u$. 
Geometric Interpretation of Eigenvectors

- Recall Def. 117: A vector \( v \in \mathbb{R}^n \) is an eigenvector of the \( n \times n \) matrix \( A \) if
  \[
  A v = \lambda v \quad \text{and} \quad v \neq 0.
  \]
- Geometrically, the vector \( A u \) is some vector of \( \mathbb{R}^n \) obtained by applying a linear transformation \( g \), whose matrix equals \( A \), to \( u \).
- Hence, \( v \neq 0 \) is an eigenvector of \( A \) if and only if \( g(v) \) equals \( v \) up to scaling, where the scale factor is given by the corresponding eigenvalue.
Geometric Interpretation of Eigenvectors

- Recall Def. 117: A vector $v \in \mathbb{R}^n$ is an eigenvector of the $n \times n$ matrix $A$ if
  \[ A v = \lambda v \text{ and } v \neq 0. \]

- Geometrically, the vector $A u$ is some vector of $\mathbb{R}^n$ obtained by applying a linear transformation $g$, whose matrix equals $A$, to $u$.

- Hence, $v \neq 0$ is an eigenvector of $A$ if and only if $g(v)$ equals $v$ up to scaling, where the scale factor is given by the corresponding eigenvalue.

- That is, $v \neq 0$ is an eigenvector of $A$ if and only if $g(v)$ lies within the span of $v$, i.e., the line that passes through its origin and its tip.
Geometric Interpretation of Eigenvectors

- Recall Def. 117: A vector $v \in \mathbb{R}^n$ is an eigenvector of the $n \times n$ matrix $A$ if
  \[ A v = \lambda v \quad \text{and} \quad v \neq 0. \]

- Geometrically, the vector $Au$ is some vector of $\mathbb{R}^n$ obtained by applying a linear transformation $g$, whose matrix equals $A$, to $u$.

- Hence, $v \neq 0$ is an eigenvector of $A$ if and only if $g(v)$ equals $v$ up to scaling, where the scale factor is given by the corresponding eigenvalue.

- That is, $v \neq 0$ is an eigenvector of $A$ if and only if $g(v)$ lies within the span of $v$, i.e., the line that passes through its origin and its tip.

- Linearity of the transformation implies that every other (non-zero) vector within the span of $v$ also forms an eigenvector of $A$. 

Geometric Interpretation of Eigenvectors

- Recall Def. 117: A vector $v \in \mathbb{R}^n$ is an eigenvector of the $n \times n$ matrix $A$ if
  \[ A v = \lambda v \quad \text{and} \quad v \neq 0. \]
- Geometrically, the vector $Au$ is some vector of $\mathbb{R}^n$ obtained by applying a linear transformation $g$, whose matrix equals $A$, to $u$.
- Hence, $v \neq 0$ is an eigenvector of $A$ if and only if $g(v)$ equals $v$ up to scaling, where the scale factor is given by the corresponding eigenvalue.
- That is, $v \neq 0$ is an eigenvector of $A$ if and only if $g(v)$ lies within the span of $v$, i.e., the line that passes through its origin and its tip.
- Linearity of the transformation implies that every other (non-zero) vector within the span of $v$ also forms an eigenvector of $A$.
- Note that $A$ might have just one eigenvalue while all vectors of $\mathbb{R}^n$ are eigenvectors: E.g., let $A$ be the $n \times n$ diagonal matrix with all diagonal elements equal to 2.
Geometric Interpretation of Eigenvectors

- Recall Def. 117: A vector \( \mathbf{v} \in \mathbb{R}^n \) is an eigenvector of the \( n \times n \) matrix \( \mathbf{A} \) if
  \[
  \mathbf{A} \mathbf{v} = \lambda \mathbf{v} \quad \text{and} \quad \mathbf{v} \neq 0.
  \]
- Geometrically, the vector \( \mathbf{A} \mathbf{u} \) is some vector of \( \mathbb{R}^n \) obtained by applying a linear transformation \( g \), whose matrix equals \( \mathbf{A} \), to \( \mathbf{u} \).
- Hence, \( \mathbf{v} \neq 0 \) is an eigenvector of \( \mathbf{A} \) if and only if \( g(\mathbf{v}) \) equals \( \mathbf{v} \) up to scaling, where the scale factor is given by the corresponding eigenvalue.
- That is, \( \mathbf{v} \neq 0 \) is an eigenvector of \( \mathbf{A} \) if and only if \( g(\mathbf{v}) \) lies within the span of \( \mathbf{v} \), i.e., the line that passes through its origin and its tip.
- Linearity of the transformation implies that every other (non-zero) vector within the span of \( \mathbf{v} \) also forms an eigenvector of \( \mathbf{A} \).
- Note that \( \mathbf{A} \) might have just one eigenvalue while all vectors of \( \mathbb{R}^n \) are eigenvectors: E.g., let \( \mathbf{A} \) be the \( n \times n \) diagonal matrix with all diagonal elements equal to 2.
- A matrix need not have even just one eigenvalue: E.g., consider the matrix that corresponds to a rotation by 90° about the origin in \( \mathbb{R}^2 \).
Definition 240 (2D rotation group, Dt.: Kreisgruppe)

The 2D rotation group, which is often denoted by $SO(2)$, is the set of all rotations about the origin of $\mathbb{R}^2$ under the operation of composition.

Definition 241 (3D rotation group, Dt.: Drehgruppe)

The 3D rotation group, which is often denoted by $SO(3)$, is the set of all rotations about the origin of $\mathbb{R}^3$ under the operation of composition.

Lemma 242

The rotation groups $SO(n)$ are non-Abelian groups for $n \geq 3$, while $SO(2)$ is Abelian.

Recall that rotations are linear transformations of $\mathbb{R}^3$ which (relative to an orthonormal base of $\mathbb{R}^3$) can be represented by orthogonal $3 \times 3$ matrices with determinant 1. Hence, the group $SO(3)$ can be identified with the group of these matrices under matrix multiplication. These matrices are known as “special orthogonal matrices”, thus explaining the term $SO(3)$. 
Rotation Group

**Definition 240 (2D rotation group, Dt.: Kreisgruppe)**

The *2D rotation group*, which is often denoted by $SO(2)$, is the set of all rotations about the origin of $\mathbb{R}^2$ under the operation of composition.

**Definition 241 (3D rotation group, Dt.: Drehgruppe)**

The *3D rotation group*, which is often denoted by $SO(3)$, is the set of all rotations about the origin of $\mathbb{R}^3$ under the operation of composition.
Rotation Group

**Definition 240 (2D rotation group, Dt.: Kreisgruppe)**

The **2D rotation group**, which is often denoted by $SO(2)$, is the set of all rotations about the origin of $\mathbb{R}^2$ under the operation of composition.

**Definition 241 (3D rotation group, Dt.: Drehgruppe)**

The **3D rotation group**, which is often denoted by $SO(3)$, is the set of all rotations about the origin of $\mathbb{R}^3$ under the operation of composition.

**Lemma 242**

The rotation groups $SO(n)$ are non-Abelian groups for $n \geq 3$, while $SO(2)$ is Abelian.
Rotation Group

Definition 240 (2D rotation group, Dt.: Kreisgruppe)

The 2D rotation group, which is often denoted by SO(2), is the set of all rotations about the origin of $\mathbb{R}^2$ under the operation of composition.

Definition 241 (3D rotation group, Dt.: Drehgruppe)

The 3D rotation group, which is often denoted by SO(3), is the set of all rotations about the origin of $\mathbb{R}^3$ under the operation of composition.

Lemma 242

The rotation groups SO(n) are non-Abelian groups for $n \geq 3$, while SO(2) is Abelian.

- Recall that rotations are linear transformations of $\mathbb{R}^3$ which (relative to an orthonormal base of $\mathbb{R}^3$) can be represented by orthogonal $3 \times 3$ matrices with determinant 1.
- Hence, the group SO(3) can be identified with the group of these matrices under matrix multiplication.
- These matrices are known as “special orthogonal matrices”, thus explaining the term SO(3).
Lemma 243 (Soccer Ball Lemma)

Suppose that a soccer ball is a ball with a perfectly spherical surface. Then in every soccer game and for every pair of consecutive (ideally perfect) placements of the soccer ball at the kick-off point there exist two points on the surface of the soccer ball which have the same coordinates relative to some coordinate system of the soccer field.

Proof:

We note that we may ignore any tumbling motion and focus just on the finitely many points in time when the ball does not move. Hence, the movement of a soccer ball during the game can be modelled as a sequence of finitely many rotations (about its center).

Since rotations belong to $SO(3)$, a sequence of finitely many rotations can be modelled by one rotation:

$$ R = R_n \cdots R_2 \cdot R_1 $$

We will now show that there exists a vector $v$ such that $Rv = v$. We see that the vector $v$ must be an eigenvector of the matrix $R$ with eigenvalue $\lambda = 1$. Since this requires $(R - I)v = 0$, we know that $\det(R - I) = 0$ is required.
Euler’s Rotation Theorem

Lemma 243 (Soccer Ball Lemma)
Suppose that a soccer ball is a ball with a perfectly spherical surface. Then in every soccer game and for every pair of consecutive (ideally perfect) placements of the soccer ball at the kick-off point there exist two points on the surface of the soccer ball which have the same coordinates relative to some coordinate system of the soccer field.

Proof: We note that we may ignore any tumbling motion and focus just on the finitely many points in time when the ball does not move. Hence, the movement of a soccer ball during the game can be modelled as a sequence of finitely many rotations (about its center).
Euler’s Rotation Theorem

Lemma 243 (Soccer Ball Lemma)
Suppose that a soccer ball is a ball with a perfectly spherical surface. Then in every soccer game and for every pair of consecutive (ideally perfect) placements of the soccer ball at the kick-off point there exist two points on the surface of the soccer ball which have the same coordinates relative to some coordinate system of the soccer field.

Proof: We note that we may ignore any tumbling motion and focus just on the finitely many points in time when the ball does not move. Hence, the movement of a soccer ball during the game can be modelled as a sequence of finitely many rotations (about its center). Since rotations belong to $SO(3)$, a sequence of finitely many rotations can be modelled by one rotation:

$$R := R_n \cdot \ldots \cdot R_2 \cdot R_1$$
Lemma 243 (Soccer Ball Lemma)

Suppose that a soccer ball is a ball with a perfectly spherical surface. Then in every soccer game and for every pair of consecutive (ideally perfect) placements of the soccer ball at the kick-off point there exist two points on the surface of the soccer ball which have the same coordinates relative to some coordinate system of the soccer field.

Proof: We note that we may ignore any tumbling motion and focus just on the finitely many points in time when the ball does not move. Hence, the movement of a soccer ball during the game can be modelled as a sequence of finitely many rotations (about its center).

Since rotations belong to $SO(3)$, a sequence of finitely many rotations can be modelled by one rotation:

$$R := R_n \cdot \ldots \cdot R_2 \cdot R_1$$

We will now show that there exists a vector $v$ such that $Rv = v$. We see that the vector $v$ must be an eigenvector of the matrix $R$ with eigenvalue $\lambda = 1$. Since this requires $(R - I)v = 0$, we know that $\det(R - I) = 0$ is required.
Euler’s Rotation Theorem

Proof of Lem. 243 (cont’d): We use

\[ \det(- (R - I)) = - \det(R - I) \quad \text{and} \quad \det(R^{-1}) = 1 \]
Euler’s Rotation Theorem

Proof of Lem. 243 (cont’d): We use

\[\det(-\mathbf{(R - I)}) = -\det(R - I) \quad \text{and} \quad \det(R^{-1}) = 1\]

and obtain

\[\det(R - I) = \det((R - I)^t)\]
Euler’s Rotation Theorem

Proof of Lem. 243 (cont’d): We use

\[ \det(- (R - I)) = - \det(R - I) \quad \text{and} \quad \det(R^{-1}) = 1 \]

and obtain

\[ \det(R - I) = \det \left( (R - I)^t \right) = \det \left( R^t - I \right) \]
Euler’s Rotation Theorem

Proof of Lem. 243 (cont’d): We use

\[ \det(- (R - I)) = - \det(R - I) \quad \text{and} \quad \det(R^{-1}) = 1 \]

and obtain

\[ \det(R - I) = \det((R - I)^t) = \det(R^t - I) = \det(R^{-1} - R^{-1}R) \]
Euler’s Rotation Theorem

**Proof of Lem. 243 (cont’d):** We use

$$\det(- (R - I)) = - \det(R - I) \quad \text{and} \quad \det(R^{-1}) = 1$$

and obtain

$$\det(R - I) = \det((R - I)^t) = \det(R^t - I) = \det(R^{-1} - R^{-1}R)$$

$$= \det(R^{-1}(I - R))$$
Euler’s Rotation Theorem

Proof of Lem. 243 (cont’d): We use

\[ \det(-(\mathbf{R} - \mathbf{I})) = -\det(\mathbf{R} - \mathbf{I}) \quad \text{and} \quad \det(\mathbf{R}^{-1}) = 1 \]

and obtain

\[
\det(\mathbf{R} - \mathbf{I}) = \det\left( (\mathbf{R} - \mathbf{I})^t \right) = \det\left( \mathbf{R}^t - \mathbf{I} \right) = \det\left( \mathbf{R}^{-1} - \mathbf{R}^{-1} \mathbf{R} \right)
\]

\[
= \det\left( \mathbf{R}^{-1}(\mathbf{I} - \mathbf{R}) \right) = \det\left( \mathbf{R}^{-1} \right) \det(-(\mathbf{R} - \mathbf{I}))
\]
Euler’s Rotation Theorem

**Proof of Lem. 243 (cont’d):** We use

\[ \det(-(\mathbf{R} - \mathbf{I})) = -\det(\mathbf{R} - \mathbf{I}) \quad \text{and} \quad \det(\mathbf{R}^{-1}) = 1 \]

and obtain

\[ \det(\mathbf{R} - \mathbf{I}) = \det\left((\mathbf{R} - \mathbf{I})^t\right) = \det\left(\mathbf{R}^t - \mathbf{I}\right) = \det\left(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{R}\right) \]
\[ = \det\left(\mathbf{R}^{-1}(\mathbf{I} - \mathbf{R})\right) = \det\left(\mathbf{R}^{-1}\right) \det(-(\mathbf{R} - \mathbf{I})) = -\det(\mathbf{R} - \mathbf{I}). \]
Euler’s Rotation Theorem

Proof of Lem. 243 (cont’d): We use

\[ \det(-(R - I)) = - \det(R - I) \quad \text{and} \quad \det(R^{-1}) = 1 \]

and obtain

\[
\det(R - I) = \det\left((R - I)^t\right) = \det\left(R^t - I\right) = \det\left(R^{-1} - R^{-1}R\right) \\
= \det\left(R^{-1}(I - R)\right) = \det\left(R^{-1}\right) \det(-(R - I)) = - \det(R - I).
\]

Thus, \( \det(R - I) = 0. \)
Euler’s Rotation Theorem

Proof of Lem. 243 (cont’d): We use

$$\det(-(R-I)) = -\det(R-I) \quad \text{and} \quad \det(R^{-1}) = 1$$

and obtain

$$\det(R - I) = \det\left((R - I)^t\right) = \det\left(R^t - I\right) = \det\left(R^{-1} - R^{-1}R\right)$$
$$= \det\left(R^{-1}(I - R)\right) = \det\left(R^{-1}\right) \det(-(R - I)) = -\det(R - I).$$

Thus, $\det(R - I) = 0$. Hence, there is at least one non-zero vector $v$ such that $Rv = v$. The intersection points of the soccer ball with the line through its center with direction vector $v$ are the two points claimed to remain invariant. \(\square\)
Euler’s Rotation Theorem

Proof of Lem. 243 (cont’d): We use

\[ \det(- (R - I)) = - \det(R - I) \quad \text{and} \quad \det(R^{-1}) = 1 \]

and obtain

\[
\det(R - I) = \det\left((R - I)^t\right) = \det\left(R^t - I\right) = \det\left(R^{-1} - R^{-1}R\right) \\
= \det\left(R^{-1}(I - R)\right) = \det\left(R^{-1}\right) \det(- (R - I)) = - \det(R - I).
\]

Thus, \( \det(R - I) = 0 \). Hence, there is at least one non-zero vector \( v \) such that \( Rv = v \). The intersection points of the soccer ball with the line through its center with direction vector \( v \) are the two points claimed to remain invariant. \( \square \)

Theorem 244 (Euler’s Rotation Theorem 1775)

Every displacement of a rigid body such that a point on the rigid body is kept fixed is equivalent to a single rotation about some axis that runs through the fixed point.
Lemma 245

Let $Q$ be a quaternion that is not zero and $P$ be a pure quaternion. Then $P' := QPQ^{-1}$ is a pure quaternion, too.
Quaternions and Rotation

Lemma 245
Let $Q$ be a quaternion that is not zero and $P$ be a pure quaternion. Then $P' := QPQ^{-1}$ is a pure quaternion, too.

- This quaternion operation maps the set of all pure quaternions onto itself.
- This set forms a 3-dimensional space among the space of the quaternions.
Quaternions and Rotation

Lemma 245

Let \( Q \) be a quaternion that is not zero and \( P \) be a pure quaternion. Then \( P' := QPQ^{-1} \) is a pure quaternion, too.

- This quaternion operation maps the set of all pure quaternions onto itself.
- This set forms a 3-dimensional space among the space of the quaternions.

Theorem 246

Let \( p \) be a point in \( \mathbb{R}^3 \) and consider an axis through the origin with direction vector \( u \), with \( \|u\| = 1 \). Let \( p' \) denote the rotation of \( p \) about that axis by the angle \( 2\phi \). Now consider the pure quaternions \( P := (0, p) \) and \( P' := (0, p') \). We have

\[ P' = QPQ^{-1} \quad \text{for} \quad Q := (\cos \phi, u \sin \phi). \]
**Quaternions and Rotation**

**Lemma 245**
Let $Q$ be a quaternion that is not zero and $P$ be a pure quaternion. Then $P' := QPQ^{-1}$ is a pure quaternion, too.

- This quaternion operation maps the set of all pure quaternions onto itself.
- This set forms a 3-dimensional space among the space of the quaternions.

**Theorem 246**
Let $p$ be a point in $\mathbb{R}^3$ and consider an axis through the origin with direction vector $u$, with $\|u\| = 1$. Let $p'$ denote the rotation of $p$ about that axis by the angle $2\phi$. Now consider the pure quaternions $P := (0, p)$ and $P' := (0, p')$. We have

$$P' = QPQ^{-1} \quad \text{for} \quad Q := (\cos \phi, u \sin \phi).$$

**Lemma 247**
Consider the setting of Theorem 246 and let $s := \cos \phi$, $v := u \sin \phi$. Then

$$p' = s^2 p + \langle p, v \rangle v + 2s (v \times p) + v \times (v \times p).$$
Quaternions and Rotation

- We conclude that every rotation about an axis in $\mathbb{R}^3$ corresponds to a unit quaternion.
- Conversely, every unit quaternion represents a rotation about an axis in $\mathbb{R}^3$. 

Theorem 248
There is a one-to-one correspondence between unit quaternions and rotations about axes in $\mathbb{R}^3$.

Lemma 249
The inverse quaternion models the opposite rotation.
Proof:
We have $Q^{-1} (QPQ^{-1}) Q = P$.
Geometric interpretation of this fact: Since $Q^{-1} = (s, -u)$ for a unit quaternion $Q = (s, u)$, the inverse of $Q$ rotates by the same angle, but the rotation axis points in the opposite direction. Hence, by inverting the axis, the direction of rotation is reversed!
Quaternions and Rotation

- We conclude that every rotation about an axis in $\mathbb{R}^3$ corresponds to a unit quaternion.
- Conversely, every unit quaternion represents a rotation about an axis in $\mathbb{R}^3$.

**Theorem 248**

There is a one-to-one correspondence between unit quaternions and rotations about axes in $\mathbb{R}^3$. 

**Lemma 249**

The inverse quaternion models the opposite rotation.

Proof:

We have $Q^{-1}(QPQ^{-1})Q = P$. Geometric interpretation of this fact: Since $Q^{-1} = (s, -u)$ for a unit quaternion $Q := (s, u)$, the inverse of $Q$ rotates by the same angle, but the rotation axis points in the opposite direction. Hence, by inverting the axis, the direction of rotation is reversed!
Quaternions and Rotation

- We conclude that every rotation about an axis in $\mathbb{R}^3$ corresponds to a unit quaternion.
- Conversely, every unit quaternion represents a rotation about an axis in $\mathbb{R}^3$.

**Theorem 248**
There is a one-to-one correspondence between unit quaternions and rotations about axes in $\mathbb{R}^3$.

**Lemma 249**
The inverse quaternion models the opposite rotation.
Quaternions and Rotation

- We conclude that every rotation about an axis in $\mathbb{R}^3$ corresponds to a unit quaternion.
- Conversely, every unit quaternion represents a rotation about an axis in $\mathbb{R}^3$.

**Theorem 248**
There is a one-to-one correspondence between unit quaternions and rotations about axes in $\mathbb{R}^3$.

**Lemma 249**
The inverse quaternion models the opposite rotation.

*Proof:* We have

$$Q^{-1}(QPQ^{-1})Q = P.$$
Quaternions and Rotation

- We conclude that every rotation about an axis in $\mathbb{R}^3$ corresponds to a unit quaternion.
- Conversely, every unit quaternion represents a rotation about an axis in $\mathbb{R}^3$.

**Theorem 248**

There is a one-to-one correspondence between unit quaternions and rotations about axes in $\mathbb{R}^3$.

**Lemma 249**

The inverse quaternion models the opposite rotation.

**Proof:** We have

$$Q^{-1}(Q\mathcal{P}Q^{-1})Q = \mathcal{P}.$$

- Geometric interpretation of this fact: Since $Q^{-1} = (s, -u)$ for a unit quaternion $Q := (s, u)$, the inverse of $Q$ rotates by the same angle, but the rotation axis points in the opposite direction. Hence, by inverting the axis, the direction of rotation is reversed!
Lemma 250

If $Q$ is a unit quaternion then $Q$ and $-Q$ represent the same rotation.

Sketch of Proof:
A rotation about the axis $u$ by the angle $2\phi$ equals a rotation about the (inversely oriented) axis $-u$ by the angle $-2\phi$.
Quaternions and Rotation

Lemma 250
If \( Q \) is a unit quaternion then \( Q \) and \( -Q \) represent the same rotation.

\textit{Sketch of Proof}: A rotation about the axis \( u \) by the angle \( 2\varphi \) equals a rotation about the (inversely oriented) axis \( -u \) by the angle \( -2\varphi \).
Quaternions and Rotation

**Lemma 250**

If \( Q \) is a unit quaternion then \( Q \) and \(-Q\) represent the same rotation.

*Sketch of Proof:* A rotation about the axis \( u \) by the angle \( 2\varphi \) equals a rotation about the (inversely oriented) axis \(-u\) by the angle \(-2\varphi\).

**Lemma 251**

The square \( Q^2 \) of a unit quaternion \( Q \) is a rotation by twice the angle about the same axis.
Quaternions and Rotation

Lemma 250

If $Q$ is a unit quaternion then $Q$ and $-Q$ represent the same rotation.

*Sketch of Proof:* A rotation about the axis $u$ by the angle $2\varphi$ equals a rotation about the (inversely oriented) axis $-u$ by the angle $-2\varphi$.

Lemma 251

The square $Q^2$ of a unit quaternion $Q$ is a rotation by twice the angle about the same axis.

Lemma 252

The orthogonal matrix that corresponds to a rotation by the unit quaternion $Q = (s, (a, b, c))$ is given by

$$
\begin{pmatrix}
    s^2 + a^2 - b^2 - c^2 & 2ab - 2sc & 2ac + 2sb \\
    2ab + 2sc & s^2 - a^2 + b^2 - c^2 & 2bc - 2sa \\
    2ac - 2sb & 2bc + 2sa & s^2 - a^2 - b^2 + c^2
\end{pmatrix}.
$$
Suppose that we are given two unit quaternions $Q_0$, $Q_1$ and would like to interpolate the rotations specified by these quaternions linearly.

Theorem 253 (Shoemake 1985)

Consider two unit quaternions $Q_0 := (s_0, (a_0, b_0, c_0))$ and $Q_1 := (s_1, (a_1, b_1, c_1))$. Let $\Theta$ such that

$$\cos \Theta = s_0 \cdot s_1 + a_0 \cdot a_1 + b_0 \cdot b_1 + c_0 \cdot c_1.$$ 

Then, for $t \in [0, 1]$, SLERP($Q_0, Q_1, t$) :=

$$\frac{\sin \Theta}{\sin((1 - t)\Theta)} Q_0 + \frac{\sin(t \Theta)}{\sin(t \Theta)} Q_1$$

corresponds to the interpolated quaternion at time $t \in [0, 1]$. The SLERP interpolation function achieves constant angular velocity.
Suppose that we are given two unit quaternions \( Q_0, Q_1 \) and would like to interpolate the rotations specified by these quaternions linearly.

Recall that a unit quaternion can be regarded as a point on the unit sphere in \( \mathbb{R}^4 \).

Hence, a natural approach to a linear interpolation of two quaternions is a spherical linear interpolation (Slerp) along the shorter arc of the great circle defined by \( Q_0 := (s_0, (a_0, b_0, c_0)) \) and \( Q_1 := (s_1, (a_1, b_1, c_1)) \):

\[
\text{Slerp}(Q_0, Q_1, t) = \frac{\sin \Theta}{\sin((1-t)\Theta)} Q_0 + \frac{\sin(t\Theta)}{\sin((1-t)\Theta)} Q_1
\]

This function achieves constant angular velocity.
Quaternions and Rotation: SLERP

- Suppose that we are given two unit quaternions $Q_0, Q_1$ and would like to interpolate the rotations specified by these quaternions linearly.
- Recall that a unit quaternion can be regarded as a point on the unit sphere in $\mathbb{R}^4$.
- Hence, a natural approach to a linear interpolation of two quaternions is a spherical linear interpolation (Slerp) along the shorter arc of the great circle defined by $Q_0 := (s_0, (a_0, b_0, c_0))$ and $Q_1 := (s_1, (a_1, b_1, c_1))$:

**Theorem 253 (Shoemake 1985)**
Consider two unit quaternions $Q_0 := (s_0, (a_0, b_0, c_0))$ and $Q_1 := (s_1, (a_1, b_1, c_1))$. Let $\Theta$ such that

$$
\cos \Theta = s_0 \cdot s_1 + a_0 \cdot a_1 + b_0 \cdot b_1 + c_0 \cdot c_1.
$$

Then, for $t \in [0, 1]$,

$$
\text{Slerp}(Q_0, Q_1, t) := \frac{1}{\sin \Theta} \left( \sin((1 - t)\Theta) Q_0 + \sin(t\Theta) Q_1 \right)
$$

corresponds to the interpolated quaternion at time $t \in [0, 1]$. The Slerp interpolation function achieves constant angular velocity.
Transformations

- Linear Transformations
- Classification of Transformations
- Coordinate Transformations in $\mathbb{R}^2$
- Coordinate Transformations in $\mathbb{R}^3$
- Transformation of Coordinate Systems
- Applications of Coordinate (System) Transformations
- Rotations Revisited

Projections

- Basics of Projections
- Perspective Projection
- Parallel Projection
- Projecting Curved Objects
- Perspective Normalization
Projections

- Virtually all output devices are two-dimensional.
- To draw a 3D scene, the scene has to be projected onto a 2D viewing plane.
Projections: History

- Plan from Mesopotamia, \( \approx 2000\text{BC} \).
- Early Greeks: *Agatharchus* (\( \approx 500\text{BC} \)), *Apollonius* of Perga (\( \approx 262\text{BC till } \approx 190\text{BC} \)) studied projections of quadrics.
- Romans: *Vitruvius* wrote *De Architectura*, published specifications of plan and elevation drawings, and perspective.
Projections: History

- Plan from Mesopotamia, \( \approx 2000BC \).
- Early Greeks: Agatharchus (\( \approx 500BC \)), Apollonius of Perga (\( \approx 262BC \) till \( \approx 190BC \)) studied projections of quadrics.
- Romans: Vitruvius wrote *De Architectura*, published specifications of plan and elevation drawings, and perspective.
- Early Renaissance period: Emphasis on point of view, interpretation of world.
  - Dürer
  - Giotto,
  - Mossacio,
  - Raphael,
  - Vinci.
Projections: History

- Plan from Mesopotamia, ≈ 2000BC.
- Early Greeks: Agatharchus (≈ 500BC), Apollonius of Perga (≈ 262BC till ≈ 190BC) studied projections of quadrics.
- Romans: Vitruvius wrote De Architectura, published specifications of plan and elevation drawings, and perspective.
- Early Renaissance period: Emphasis on point of view, interpretation of world.
  - Dürer
  - Giotto,
  - Mossacio,
  - Raphael,
  - Vinci.
- Leon Battista Alberti wrote the first treatise on perspective, “Della Pittura”, in 1435.
  “A painting is the intersection of a visual pyramid at a given distance, with a fixed center and a definite position of light, represented by art with lines and colors on a given surface.”
Given a scene, we have to specify how we want to view the scene:

- **Camera Position**: Point $E$.
- **Object Point at Center of Window**: Point $R$. Also called “to”-point.
- **Center of View Plane**: Point $O$.
- **“Up” direction**: Vector $V$; must not be parallel to the line through $E$ and $R$. 
Camera Set-Up

- Given a scene, we have to specify how we want to view the scene:
  - **Camera Position**: Point \( E \).
  - **Object Point at Center of Window**: Point \( R \). Also called “to”-point.
  - **Center of View Plane**: Point \( O \).
  - **“Up” direction**: Vector \( V \); must not be parallel to the line through \( E \) and \( R \).

We may also need to specify the viewing angle, and other camera-related values.
Geometric Projections

- **Projection plane**: Plane \( \Pi \).
- **Projectors**: Rays emanating from the center of projection and passing through points of the object.
- **Projection**: Intersection of projectors with plane \( \Pi \).
- Non-geometric projections used in cartography.
Different Types of Geometric Projections

- **Perspective:**
  - Center of projection is at a finite distance from $\Pi$.
  - *Perspective foreshortening* occurs.
Different Types of Geometric Projections

- **Perspective:**
  - Center of projection is at a finite distance from $\Pi$.
  - *Perspective foreshortening* occurs.

- **Parallel:**
  - Center of projection is at $\infty$.
  - Defined by the direction $v$.
  - Directions correspond to points at infinity.
Different Types of Geometric Projections

Planar geometric projection

Parallel

Orthographic

Top (plan)

Front elevation

Axonometric

Isometric

Side elevation

Oblique

Cabinet

Axonometric

Isometric

Orthographic

Top (plan)

Front elevation

Perspective

One-point

Two-point

Three-point

Oblique

Cavalier

Other
Three-Dimensional View Volume

When formulating the mathematics of projections it is customary to place the viewpoint at \((0, 0, -d)\), in the case of a perspective projection, and to assume that the projection plane \(\Pi\) is the \(xy\)-plane.
Perspective Projection

- Perspective foreshortening gives a realistic view of 3D objects.
- Used for advertising, fine art, architecture.
Perspective Projection

- Perspective foreshortening gives a realistic view of 3D objects.
- Used for advertising, fine art, architecture.
- Foreshortening is not uniform.
- Parallel edges do not remain parallel; angles, scales and other geometric properties are not preserved.

A vanishing point (1/d): Fluchtpunkt) is a point in the image plane where the projections of lines parallel to one of the three principal axes of the object converge: one-point perspective, two-point perspective, or three-point perspective.
Perspective Projection

- Perspective foreshortening gives a realistic view of 3D objects.
- Used for advertising, fine art, architecture.
- Foreshortening is not uniform.
- Parallel edges do not remain parallel; angles, scales and other geometric properties are not preserved.
- A vanishing point (Dt.: Fluchtpunkt) is a point in the image plane where the projections of lines parallel to one of the three principal axes of the object converge: one-point perspective, two-point perspective, or three-point perspective.
One Vanishing Point

- Π parallel to two principal axes of the cube: one vanishing point.
Two Vanishing Points

- \( \Pi \) is parallel to only one principal axis of the cube: two vanishing points.
Three Vanishing Points

- \( \Pi \) is not parallel to any principal axis of the cube: three vanishing points.
Due to the similarity of the triangles $\triangle(Z, O, P'_{xz})$ and $\triangle(Z, P_z, P_{xz})$ we get

\[ x' : d = x : (z + d), \text{ i.e., } x' = \frac{d \cdot x}{z + d}. \]
Mathematics of Perspective Projection

- Due to the similarity of the triangles $\triangle(Z, O, P'_{xz})$ and $\triangle(Z, P_z, P_{xz})$ we get
  \[ x' : d = x : (z + d), \quad \text{i.e.,} \quad x' = \frac{d \cdot x}{z + d}. \]

- Analogously,
  \[ y' = \frac{d \cdot y}{z + d}. \]
Let $P := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{d} & 1 \end{pmatrix}$. 

We get $P \cdot p = \begin{pmatrix} p_x \\ p_y \\ 0 \\ p_z + d \end{pmatrix} \equiv \begin{pmatrix} d \cdot p_x \\ d \cdot p_y \\ p_z + d \\ 0 \end{pmatrix} =: \begin{pmatrix} p'_x \\ p'_y \\ 0 \end{pmatrix}$. 

Apply transformation of coordinate system if the projection plane differs from $z = 0$, or if the eye point is not at $(0, 0, -d)$.
Matrix of a Perspective Projection

Let $P := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{d} & 1 \end{pmatrix}$.

We get

$$P \cdot p = \begin{pmatrix} p_x \\ p_y \\ 0 \\ \frac{p_z + d}{d} \end{pmatrix} \equiv \begin{pmatrix} \frac{d \cdot p_x}{p_z + d} \\ \frac{d \cdot p_y}{p_z + d} \\ 0 \\ 1 \end{pmatrix} =: \begin{pmatrix} p'_x \\ p'_y \\ 0 \\ 1 \end{pmatrix}.$$
Matrix of a Perspective Projection

Let \( P := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{d} & 1 \end{pmatrix} \).

We get

\[
P \cdot p = \begin{pmatrix} p_x \\ p_y \\ 0 \\ \frac{p_z + d}{d} \end{pmatrix} \equiv \begin{pmatrix} \frac{d \cdot p_x}{p_z + d} \\ \frac{d \cdot p_y}{p_z + d} \\ 0 \\ 1 \end{pmatrix} =: \begin{pmatrix} p'_x \\ p'_y \\ 0 \\ 1 \end{pmatrix}.
\]

Apply transformation of coordinate system if the projection plane differs from \( z = 0 \), or if the eye point is not at \((0, 0, -d)\).
Parallel Projection: Orthographic

- **Orthographic**: Projectors are perpendicular to the projection plane.

\[ \vec{P}_{xy} := \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}. \]
Parallel Projection: Orthographic

- **Orthographic**: Projectors are perpendicular to the projection plane.

\[ \rightarrow P_{xy} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \]

- **Front, top, side views**: Projectors parallel to one of the principal axes.


**Parallel Projection: Oblique**

- **Oblique**: Projectors not perpendicular to the projection plane.

$$\begin{align*}
\mathbf{P} &= \begin{pmatrix}
1 & 0 & d \cos \alpha \\
0 & 1 & d \sin \alpha \\
0 & 0 & 0
\end{pmatrix}
\end{align*}$$
Parallel Projection: Oblique

- **Oblique**: Projectors not perpendicular to the projection plane.
- With $d := \cot \beta$ we get

  $$x' = x + z \cdot d \cos \alpha,$$

  $$y' = y + z \cdot d \sin \alpha,$$

  $$z' = 0.$$
Parallel Projection: Oblique

- **Oblique**: Projectors not perpendicular to the projection plane.
- With $d := \cot \beta$ we get
  
  \[
  x' = x + z \cdot d \cos \alpha, \\
  y' = y + z \cdot d \sin \alpha, \\
  z' = 0.
  \]

  Thus,
  
  
  \[
  P := \begin{pmatrix}
  1 & 0 & d \cos \alpha & 0 \\
  0 & 1 & d \sin \alpha & 0 \\
  0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 1
  \end{pmatrix}.
  \]
Special Oblique Projections

- **Cavalier projection:**
  - Angle $\beta$ between projectors and projection plane is $45^\circ$; i.e., $d = 1$.
  - The length of a segment normal to the projection plane equals the length of the projection of that segment.

- **Cabinet projection:**
  - Angle $\beta$ between projectors and projection plane is $\tan^{-1} 2 \approx 63.4^\circ$; i.e., $d = \frac{1}{2}$.
  - The length of a segment normal to the projection plane equals twice the length of the projection of that segment.
Projecting Curved Objects

- It does not suffice to project the vertices and edges of an object if the object is bounded by curved surfaces.

A silhouette curve consists of all those points of the object such that the line through the point and the center of projection is tangential to the object. Note that the silhouette curves need not lie in one plane!
Projecting Curved Objects

- It does not suffice to project the vertices and edges of an object if the object is bounded by curved surfaces.

- Rather, we also have to project the silhouette curves of the object.

- A silhouette curve consists of all those points of the object such that the line through the point and the center of projection is tangential to the object.
Projecting Curved Objects

- It does not suffice to project the vertices and edges of an object if the object is bounded by curved surfaces.

- Rather, we also have to project the *silhouette curves* of the object.

- A silhouette curve consists of all those points of the object such that the line through the point and the center of projection is tangential to the object.

- Note that the silhouette curves need not lie in one plane!
Perspective Normalization

- For computing silhouette curves, hidden-surface elimination, ray tracing, and many other algorithms, it is convenient to transform the view pyramid into a view box, while maintaining the depth ordering.
Perspective Normalization

- For computing silhouette curves, hidden-surface elimination, ray tracing, and many other algorithms, it is convenient to transform the view pyramid into a view box, while maintaining the depth ordering.
  \[
  N = \begin{pmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & \frac{1}{d} & 1
  \end{pmatrix}
  \]

- Consider \( N = \begin{pmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & \frac{1}{d} & 1
  \end{pmatrix} \) that is, the \( xy \)-plane is invariant under \( N \).

The center of projection is mapped to the point at infinity on the negative \( z \)-axis:

\[
N \cdot \begin{pmatrix}
0 \\
0 \\
-1 \\
1
\end{pmatrix} = \begin{pmatrix}
0 \\
0 \\
-1 \\
0
\end{pmatrix}.
\]
Perspective Normalization

- For computing silhouette curves, hidden-surface elimination, ray tracing, and many other algorithms, it is convenient to transform the view pyramid into a view box, while maintaining the depth ordering.

- Consider $\mathbf{N} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{d} & 1 \end{pmatrix}$.

- We get

$$\mathbf{N} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}; \quad \mathbf{N} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \quad \mathbf{N} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix};$$

that is, the $xy$-plane is invariant under $\mathbf{N}$.
**Perspective Normalization**

- For computing silhouette curves, hidden-surface elimination, ray tracing, and many other algorithms, it is convenient to transform the view pyramid into a view box, while maintaining the depth ordering.

  - Consider \( \mathbf{N} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{d} & 1 \end{pmatrix} \).

  - We get
    
    \[ \mathbf{N} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}; \quad \mathbf{N} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}; \quad \mathbf{N} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}; \]

    that is, the \( xy \)-plane is invariant under \( \mathbf{N} \).

  - The center of projection is mapped to the point at infinity on the negative \( z \)-axis:
    
    \[ \mathbf{N} \cdot \begin{pmatrix} 0 \\ 0 \\ -d \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -d \\ 0 \end{pmatrix}. \]
Perspective Normalization

Summarizing, we get

\[ \mathbf{O} \cdot \mathbf{N} = \mathbf{P}, \]

where \( \mathbf{O} \) is the matrix of an orthogonal projection, and \( \mathbf{P} \) is the matrix of the corresponding perspective projection.
Perspective Normalization

- Summarizing, we get

\[ \mathbf{O} \cdot \mathbf{N} = \mathbf{P}, \]

where \( \mathbf{O} \) is the matrix of an orthogonal projection, and \( \mathbf{P} \) is the matrix of the corresponding perspective projection.

- \( \mathbf{N} \) maps

  - cylinder, cone \( \rightarrow \) cylinder or cone (possibly with non-circular cross-section),
  - line \( \rightarrow \) line,
  - plane \( \rightarrow \) plane,
  - sphere \( \rightarrow \) ellipsoid, elliptical paraboloid, two-sheet hyperboloid,
  - quadric \( \rightarrow \) quadric.
Perspective Normalization

- Summarizing, we get
  \[ \mathbf{O} \cdot \mathbf{N} = \mathbf{P}, \]
  where \( \mathbf{O} \) is the matrix of an orthogonal projection, and \( \mathbf{P} \) is the matrix of the corresponding perspective projection.

- \( \mathbf{N} \) maps
  - cylinder, cone \( \rightarrow \) cylinder or cone (possibly with non-circular cross-section),
  - line \( \rightarrow \) line,
  - plane \( \rightarrow \) plane,
  - sphere \( \rightarrow \) ellipsoid, elliptical paraboloid, two-sheet hyperboloid,
  - quadric \( \rightarrow \) quadric.

- We can modify \( \mathbf{N} \) such that all z-coordinates are scaled to lie between 0 and 1.
Floating-Point Arithmetic and Numerical Mathematics
- Floating-Point Computations
- Iterative Algorithms for Solving Non-Linear Equations
- Iterative Algorithms for Solving Linear Equations
- Numerical Integration
Floating-Point Arithmetic and Numerical Mathematics

- Floating-Point Computations
  - Numerical Errors on IEEE 754 Arithmetic
  - Compiler Dependence
  - Common Manifestations of Floating-Point Errors
  - Comparisons of Floating-Point Numbers
  - Sample Robustness Problems
  - Improving the Reliability of Floating-Point Computations

- Iterative Algorithms for Solving Non-Linear Equations
- Iterative Algorithms for Solving Linear Equations
- Numerical Integration
Floating-Point Computations

- Virtually all modern computers employ floating-point (fp) arithmetic to perform real arithmetic. Typically, the IEEE-754 standard is used, which demands 32-bit ‘single’ precision and 64-bit ‘double’ precision representations.
- No matter how many bits are used, fp-arithmetic represents a number by a fixed-length binary mantissa and an exponent of fixed size.
- Thus, only a finite number of values within a finite sub-interval of $\mathbb{R}$ can be represented accurately; all other values have to be rounded to the closest number that is representable.
Floating-Point Computations

- Virtually all modern computers employ floating-point (fp) arithmetic to perform real arithmetic. Typically, the IEEE-754 standard is used, which demands 32-bit ‘single’ precision and 64-bit ‘double’ precision representations.

- No matter how many bits are used, fp-arithmetic represents a number by a fixed-length binary mantissa and an exponent of fixed size.

- Thus, only a finite number of values within a finite sub-interval of $\mathbb{R}$ can be represented accurately; all other values have to be rounded to the closest number that is representable.

- More generally, there are two sources of error for fp-computations: input error and round-off error.

  **Input error:** It is well-known that $\frac{1}{3}$ cannot be represented by a finite sum of powers of 10. Similarly, 0.1 cannot be represented by a finite sum of powers of 2!
Floating-Point Computations

- Virtually all modern computers employ floating-point (fp) arithmetic to perform real arithmetic. Typically, the IEEE-754 standard is used, which demands 32-bit ‘single’ precision and 64-bit ‘double’ precision representations.

- No matter how many bits are used, fp-arithmetic represents a number by a fixed-length binary mantissa and an exponent of fixed size.

- Thus, only a finite number of values within a finite sub-interval of $\mathbb{R}$ can be represented accurately; all other values have to be rounded to the closest number that is representable.

- More generally, there are two sources of error for fp-computations: input error and round-off error.

**Input error:** It is well-known that $\frac{1}{3}$ cannot be represented by a finite sum of powers of 10. Similarly, 0.1 cannot be represented by a finite sum of powers of 2!

**Round-off error:** It arises from rounding results of fp-computations during an algorithm. E.g., $\sqrt{2}$ cannot be represented exactly (by a finite sum of powers of 10 or 2) since $\sqrt{2}$ is an irrational number.
Real-World Example of Round-Off Error

During the First Gulf War (1990/91), an Iraqi Scud got through the Patriot anti-missile system (AMS) and hit a barracks of the Pennsylvania National Guard in Dhahran, Saudi Arabia, killing 28 people.
Real-World Example of Round-Off Error

- During the First Gulf War (1990/91), an Iraqi Scud got through the Patriot anti-missile system (AMS) and hit a barracks of the Pennsylvania National Guard in Dhahran, Saudi Arabia, killing 28 people.

- To track the Scud, the AMS had to determine the interval between tracking times by subtracting two values of a timer. The times in tenths of a second were stored in integer registers; a stored value of 35 would be equivalent to 3.5 seconds.

- To compute the interval, the values in the registers were converted to fp-representation by multiplying them by 0.1.

As stated previously, 0.1 has a non-terminating binary expansion. Consequently, the time interval was computed with error. The larger the value in the timer, the larger the error.

At the time of the incident, the AMS had been operating for over 100 hours, resulting in an error of 0.34 seconds in the timer, causing the system to look in the wrong place for the incoming Scud.
Real-World Example of Round-Off Error

- During the First Gulf War (1990/91), an Iraqi Scud got through the Patriot anti-missile system (AMS) and hit a barracks of the Pennsylvania National Guard in Dhahran, Saudi Arabia, killing 28 people.
- To track the Scud, the AMS had to determine the interval between tracking times by subtracting two values of a timer. The times in tenths of a second were stored in integer registers; a stored value of 35 would be equivalent to 3.5 seconds.
- To compute the interval, the values in the registers were converted to fp-representation by multiplying them by 0.1.
- As stated previously, 0.1 has a non-terminating binary expansion. Consequently, the time interval was computed with error.
- The larger the value in the timer, the larger the error.
- At the time of the incident, the AMS had been operating for over 100 hours, resulting in an error of 0.34 seconds in the timer, causing the system to look in the wrong place for the incoming Scud.
Machine Precision

- The round-off error is bounded in terms of the *machine precision*, $\varepsilon$, which is the smallest value satisfying

\[ |fp(a \circ b) - (a \circ b)| \leq \varepsilon |a \circ b| \]

for all all fp-numbers $a, b$ and any of the four operations $+, -, \cdot, /$ instead of $\circ$, for which $a \circ b$ does not cause an underflow or an overflow.
Machine Precision

- The round-off error is bounded in terms of the *machine precision*, $\varepsilon$, which is the smallest value satisfying

  $$|fp(a \circ b) - (a \circ b)| \leq \varepsilon |a \circ b|$$

  for all all fp-numbers $a, b$ and any of the four operations $+, -, \cdot, /$ instead of $\circ$, for which $a \circ b$ does not cause an underflow or an overflow.

- On IEEE-754 machines, $\varepsilon = 2^{-23} \approx 1.19 \cdot 10^{-7}$ for floats, and $\varepsilon = 2^{-52} \approx 2.22 \cdot 10^{-16}$ for doubles.

- On some exotic platform, $\varepsilon$ can be determined approximately by finding the smallest positive value $x$ such that $1 + x \neq 1$. 

Note: Some compilers promote floats to doubles!
Note: Some platforms employ extended representations, or use registers longer than standard words for intermediate results! The sad truth is that hardware vendors still prefer to stick to their own standards...
Machine Precision

- The round-off error is bounded in terms of the *machine precision*, $\varepsilon$, which is the smallest value satisfying
  \[ |\text{fp}(a \circ b) - (a \circ b)| \leq \varepsilon |a \circ b| \]
  for all fp-numbers $a, b$ and any of the four operations $+, -, \cdot, /$ instead of $\circ$, for which $a \circ b$ does not cause an underflow or an overflow.

- On IEEE-754 machines, $\varepsilon = 2^{-23} \approx 1.19 \cdot 10^{-7}$ for floats, and $\varepsilon = 2^{-52} \approx 2.22 \cdot 10^{-16}$ for doubles.

- On some exotic platform, $\varepsilon$ can be determined approximately by finding the smallest positive value $x$ such that $1 + x \neq 1$.

- While one can instruct the C command `printf` to print, say, 57 digits after the decimal separator, one will “only” get the digits of the closest value that is representable:
  
  \[
  \frac{1}{3} = 0.333333333333333314829616256247390992939472198486328125000 \\
  \frac{1}{10} = 0.100000000000000000000005551115123125782702118158340454101562500
  \]
Machine Precision

- The round-off error is bounded in terms of the *machine precision*, $\varepsilon$, which is the smallest value satisfying

$$|fp(a \circ b) - (a \circ b)| \leq \varepsilon |a \circ b|$$

for all all fp-numbers $a, b$ and any of the four operations $+, -, \cdot, /$ instead of $\circ$, for which $a \circ b$ does not cause an underflow or an overflow.

- On IEEE-754 machines, $\varepsilon = 2^{-23} \approx 1.19 \cdot 10^{-7}$ for floats, and $\varepsilon = 2^{-52} \approx 2.22 \cdot 10^{-16}$ for doubles.

- On some exotic platform, $\varepsilon$ can be determined approximately by finding the smallest positive value $x$ such that $1 + x \neq 1$.

- While one can instruct the C command `printf` to print, say, 57 digits after the decimal separator, one will “only” get the digits of the closest value that is representable:

$$1/3 = 0.333333333333333314829616256247390992939472198486328125000$$
$$1/10 = 0.10000000000000000000005551115123125782702118158340454101562500$$

- Note: Some compilers promote floats to doubles!

- Note: Some platforms employ extended representations, or use registers longer than standard words for intermediate results! The sad truth is that hardware vendors still prefer to stick to their own standards . . .
Floating-Point Arithmetic and Compilers

- Old 387 floating-point units on x86 processors used 80bit registers and operators, while standard “double” variables were stored in 64bit memory cells.
- Hence, rounding to a lower precision was necessary whenever a floating-point variable is transferred from register to memory.

As a consequence, on my PC, \[\sum_{i=1}^{1000000} 0.001 = 1000.0000000009095\] with `gcc -O2 -mfpmath=387`, \[\sum_{i=1}^{1000000} 0.001 = 999.9999999832650701\] with `gcc -O0 -mfpmath=387`. Newer chips also support the SSE instruction set, and the default option `-mfpmath=sse` avoids this problem for x86-64 compilers.

Warning: The result of fp-computations may depend on the compile-time options!

Accumulation: Adding 0.001 for 1000000 times need not yield exactly 1000. Random errors tend to cancel on a large scale, and accumulate on small scale.
Floating-Point Arithmetic and Compilers

- Old 387 floating-point units on x86 processors used 80bit registers and operators, while standard “double” variables were stored in 64bit memory cells.
- Hence, rounding to a lower precision was necessary whenever a floating-point variable is transferred from register to memory.
- As a consequence, on my PC,

\[
\sum_{i=1}^{10000000} 0.001 = 1000.0000000000009095 \quad \text{with } gcc \ -O2 \ -mfpmath=387,
\]

\[
\sum_{i=1}^{10000000} 0.001 = 999.999999832650701 \quad \text{with } gcc \ -O0 \ -mfpmath=387.
\]
Old 387 floating-point units on x86 processors used 80bit registers and operators, while standard “double” variables were stored in 64bit memory cells. Hence, rounding to a lower precision was necessary whenever a floating-point variable is transferred from register to memory.

As a consequence, on my PC,

\[
\sum_{i=1}^{10000000} 0.001 = 1000.000000000009095 \quad \text{with } gcc \ -O2 \ -mfpmath=387,
\]

\[
\sum_{i=1}^{10000000} 0.001 = 999.9999999832650701 \quad \text{with } gcc \ -O0 \ -mfpmath=387.
\]

Newer chips also support the SSE instruction set, and the default option -mfpmath=sse avoids this problem for x86-64 compilers.

Warning: The result of fp-computations may depend on the compile-time options!
Old 387 floating-point units on x86 processors used 80bit registers and operators, while standard “double” variables were stored in 64bit memory cells.

Hence, rounding to a lower precision was necessary whenever a floating-point variable is transferred from register to memory.

As a consequence, on my PC, 

\[
\sum_{i=1}^{10000000} 0.001 = 1000.0000000000009095 \quad \text{with } gcc \ -O2 \ -mfpmath=387,
\]

\[
\sum_{i=1}^{10000000} 0.001 = 999.999999832650701 \quad \text{with } gcc \ -O0 \ -mfpmath=387.
\]

Newer chips also support the SSE instruction set, and the default option -mfpmath=sse avoids this problem for x86-64 compilers.

**Warning**

The result of fp-computations may depend on the compile-time options!
Floating-Point Arithmetic and Compilers

- Old 387 floating-point units on x86 processors used 80bit registers and operators, while standard “double” variables were stored in 64bit memory cells.
- Hence, rounding to a lower precision was necessary whenever a floating-point variable is transferred from register to memory.
- As a consequence, on my PC,

\[
\sum_{i=1}^{10000000} 0.001 = 1000.0000000000009095 \quad \text{with } \text{gcc -O2 -mfpmath=387},
\]

\[
\sum_{i=1}^{10000000} 0.001 = 999.999999832650701 \quad \text{with } \text{gcc -O0 -mfpmath=387}.
\]

- Newer chips also support the SSE instruction set, and the default option -mfpmath=sse avoids this problem for x86-64 compilers.

**Warning**

The result of fp-computations may depend on the compile-time options!

- Accumulation: Adding 0.001 for 1000000 times need not yield exactly 1000.
- Random errors tend to cancel on a large scale, and accumulate on small scale.
Common Manifestations of Floating-Point Errors

- Cancellation: Subtracting two numbers of almost equal magnitudes may cause a drastic loss in the number of significant digits.
- With exact arithmetic, we would have

\[(0.1234567890123456 - 0.12345678901234) = 0.56 \cdot 10^{-14} = \frac{56}{100} \cdot 10^{-14}.\]
Common Manifestations of Floating-Point Errors

- Cancellation: Subtracting two numbers of almost equal magnitudes may cause a drastic loss in the number of significant digits.
- With exact arithmetic, we would have

\[
(0.1234567890123456 - 0.12345678901234) = 0.56 \cdot 10^{-14} = \frac{56}{100} \cdot 10^{-14}.
\]

- Taking \(\frac{56}{100}\) as result of the subtraction, we would get

\[
(0.1234567890123456 - 0.12345678901234) \cdot 10^{14} = 0.560000000000000532\ldots
\]
Common Manifestations of Floating-Point Errors

- Cancellation: Subtracting two numbers of almost equal magnitudes may cause a drastic loss in the number of significant digits.

- With exact arithmetic, we would have

\[
(0.1234567890123456 - 0.12345678901234) = 0.56 \cdot 10^{-14} = \frac{56}{100} \cdot 10^{-14}.
\]

- Taking \( \frac{56}{100} \) as result of the subtraction, we would get

\[
(0.1234567890123456 - 0.12345678901234) \cdot 10^{14} = 0.56000000000000000532\ldots
\]

but when doing all computations on a floating-point arithmetic we get

\[
(0.1234567890123456 - 0.12345678901234) \cdot 10^{14} \approx 0.5592748486549226\ldots
\]
Common Manifestations of Floating-Point Errors

- Cancellation: Subtracting two numbers of almost equal magnitudes may cause a drastic loss in the number of significant digits.

With exact arithmetic, we would have

\[(0.1234567890123456 - 0.12345678901234) = 0.56 \cdot 10^{-14} = \frac{56}{100} \cdot 10^{-14}.\]

- Taking \(\frac{56}{100}\) as result of the subtraction, we would get

\[(0.1234567890123456 - 0.12345678901234) \cdot 10^{14} = 0.5600000000000000532 \ldots\]

but when doing all computations on a floating-point arithmetic we get

\[(0.1234567890123456 - 0.12345678901234) \cdot 10^{14} \approx 0.5592748486549226 \ldots\]

- Absorption due to adding/subtracting small and large numbers: the un-normalizing required to line up the decimal point may cause truncation. E.g., adding \(2^{40} = 1099511627776\) and \(2^{-14} = 0.0000610352\) yields \(1099511627776\) with double-precision arithmetic. As a consequence,

\[0 = 2^{40} - (2^{40} - 2^{-14}) \neq (2^{40} - 2^{40}) + 2^{-14} = 2^{-14}.\]
Common Manifestations of Floating-Point Errors

- Underflow: Occurs if (absolute) value of an expression is too small to be represented as a normalized number. An expression that results in an underflow may evaluate to zero, without returning an error!

- Implementations that conform to IEEE 754-2008 try to avoid the underflow gap by resorting to “subnormal” numbers, that is, they allow leading zeros in the significand.
Common Manifestations of Floating-Point Errors

- **Underflow**: Occurs if (absolute) value of an expression is too small to be represented as a normalized number. An expression that results in an underflow may evaluate to zero, without returning an error!

- Implementations that conform to IEEE 754-2008 try to avoid the underflow gap by resorting to “subnormal” numbers, that is, they allow leading zeros in the significand.

- **Overflow**: Occurs if (absolute) value of a number is too large to be represented. The evaluation of an expression that results in an overflow will raise an error flag; the actual value of the expression is positive or negative “Inf”.

- Divisions by zero will generate positive or negative “Inf”, too.
Common Manifestations of Floating-Point Errors

- **Underflow**: Occurs if (absolute) value of an expression is too small to be represented as a normalized number. An expression that results in an underflow may evaluate to zero, without returning an error!

- Implementations that conform to IEEE 754-2008 try to avoid the underflow gap by resorting to “subnormal” numbers, that is, they allow leading zeros in the significand.

- **Overflow**: Occurs if (absolute) value of a number is too large to be represented. The evaluation of an expression that results in an overflow will raise an error flag; the actual value of the expression is positive or negative “Inf”.

- Divisions by zero will generate positive or negative “Inf”, too.

- **Not a Number**: $0/0$ or $\sqrt{-1}$ will generate a special value, “NaN”.

- E.g., $\sqrt{((1 + 10^{-20}) - 1) - 10^{-20}}$ does not yield $0$, but results in an NaN error: The truncation and subsequent cancellation lets us compute $\sqrt{-10^{-20}}$. 


Common Manifestations of Floating-Point Errors

- **Underflow**: Occurs if (absolute) value of an expression is too small to be represented as a normalized number. An expression that results in an underflow may evaluate to zero, without returning an error!

- Implementations that conform to IEEE 754-2008 try to avoid the underflow gap by resorting to “subnormal” numbers, that is, they allow leading zeros in the significand.

- **Overflow**: Occurs if (absolute) value of a number is too large to be represented. The evaluation of an expression that results in an overflow will raise an error flag; the actual value of the expression is positive or negative “Inf”.

- Divisions by zero will generate positive or negative “Inf”, too.

- **Not a Number**: $0/0$ or $\sqrt{-1}$ will generate a special value, “NaN”.

- E.g., $\sqrt{((1 + 10^{-20}) - 1) - 10^{-20}}$ does not yield 0, but results in an NaN error: The truncation and subsequent cancellation lets us compute $\sqrt{-10^{-20}}$.

- Those special numbers propagate through subsequent calculations.
Floating-Point Versus Exact Real Computations

Connectivity  ×      All points are isolated
Completeness  ×      No converging sequences
Density       ×
## Floating-Point Versus Exact Real Computations

<table>
<thead>
<tr>
<th>Property</th>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectivity</td>
<td>✗</td>
<td>All points are isolated</td>
</tr>
<tr>
<td>Completeness</td>
<td>✗</td>
<td>No converging sequences</td>
</tr>
<tr>
<td>Density</td>
<td>✗</td>
<td></td>
</tr>
<tr>
<td>Closure under addition</td>
<td>✗</td>
<td>We may have overflow/underflow</td>
</tr>
<tr>
<td>Associativity of addition</td>
<td>✗</td>
<td>((a + b) + c \neq a + (b + c))</td>
</tr>
<tr>
<td>Additive commutativity</td>
<td>✓</td>
<td>(a + b = b + a)</td>
</tr>
<tr>
<td>Unique neutral element</td>
<td>✓</td>
<td>(a + 0.0 = a)</td>
</tr>
<tr>
<td>Unique additive inverse</td>
<td>✓</td>
<td>(a + (-a) = 0.0)</td>
</tr>
</tbody>
</table>
Floating-Point Versus Exact Real Computations

Connectivity  
- All points are isolated

Completeness  
- No converging sequences

Density  

Closure under addition  
- We may have overflow/underflow

Associativity of addition  
- $(a + b) + c \neq a + (b + c)$

Additive commutativity  
- $a + b = b + a$

Unique neutral element  
- $a + 0.0 = a$

Unique additive inverse  
- $a + (−a) = 0.0$

Closure under multiplication  
- We may have overflow/underflow

Associativity of multiplication  
- $(a \cdot b) \cdot c \neq a \cdot (b \cdot c)$

Multiplicative commutativity  
- $a \cdot b = b \cdot a$

Unique unit element  
- $a \cdot 1.0 = a$

Unique multiplicative inverse  
- $a \cdot (1.0/a) = 1.0$
<table>
<thead>
<tr>
<th>Property</th>
<th>Floating-Point</th>
<th>Exact Real Computations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectivity</td>
<td>×</td>
<td>All points are isolated</td>
</tr>
<tr>
<td>Completeness</td>
<td>×</td>
<td>No converging sequences</td>
</tr>
<tr>
<td>Density</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Closure under addition</td>
<td>×</td>
<td>We may have overflow/underflow</td>
</tr>
<tr>
<td>Associativity of addition</td>
<td>×</td>
<td>$(a + b) + c \neq a + (b + c)$</td>
</tr>
<tr>
<td>Additive commutativity</td>
<td>✓</td>
<td>$a + b = b + a$</td>
</tr>
<tr>
<td>Unique neutral element</td>
<td>✓</td>
<td>$a + 0.0 = a$</td>
</tr>
<tr>
<td>Unique additive inverse</td>
<td>✓</td>
<td>$a + (-a) = 0.0$</td>
</tr>
<tr>
<td>Closure under multiplication</td>
<td>×</td>
<td>We may have overflow/underflow</td>
</tr>
<tr>
<td>Associativity of multiplication</td>
<td>×</td>
<td>$(a \cdot b) \cdot c \neq a \cdot (b \cdot c)$</td>
</tr>
<tr>
<td>Multiplicative commutativity</td>
<td>✓</td>
<td>$a \cdot b = b \cdot a$</td>
</tr>
<tr>
<td>Unique unit element</td>
<td>✓</td>
<td>$a \cdot 1.0 = a$</td>
</tr>
<tr>
<td>Unique multiplicative inverse</td>
<td>✓</td>
<td>$a \cdot (1.0/a) = 1.0$</td>
</tr>
<tr>
<td>Distributivity</td>
<td>×</td>
<td>$a \cdot (b + c) \neq (a \cdot b) + (a \cdot c)$</td>
</tr>
</tbody>
</table>
### Floating-Point Versus Exact Real Computations

<table>
<thead>
<tr>
<th>Property</th>
<th>Exact Real</th>
<th>Floating-Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectivity</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Completeness</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Density</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Closure under addition</td>
<td>×</td>
<td>We may have overflow/underflow</td>
</tr>
<tr>
<td>Associativity of addition</td>
<td>×</td>
<td>$(a + b) + c \neq a + (b + c)$</td>
</tr>
<tr>
<td>Additive commutativity</td>
<td>✓</td>
<td>$a + b = b + a$</td>
</tr>
<tr>
<td>Unique neutral element</td>
<td>✓</td>
<td>$a + 0.0 = a$</td>
</tr>
<tr>
<td>Unique additive inverse</td>
<td>✓</td>
<td>$a + (-a) = 0.0$</td>
</tr>
<tr>
<td>Closure under multiplication</td>
<td>×</td>
<td>We may have overflow/underflow</td>
</tr>
<tr>
<td>Associativity of multiplication</td>
<td>×</td>
<td>$(a \cdot b) \cdot c \neq a \cdot (b \cdot c)$</td>
</tr>
<tr>
<td>Multiplicative commutativity</td>
<td>✓</td>
<td>$a \cdot b = b \cdot a$</td>
</tr>
<tr>
<td>Unique unit element</td>
<td>✓</td>
<td>$a \cdot 1.0 = a$</td>
</tr>
<tr>
<td>Unique multiplicative inverse</td>
<td>✓</td>
<td>$a \cdot (1.0/a) = 1.0$</td>
</tr>
<tr>
<td>Distributivity</td>
<td>×</td>
<td>$a \cdot (b + c) \neq (a \cdot b) + (a \cdot c)$</td>
</tr>
<tr>
<td>Completeness of order</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Transitivity</td>
<td>✓</td>
<td>$(a \leq b) \land (b \leq c) \Rightarrow (a \leq c)$</td>
</tr>
<tr>
<td>Translation invariance</td>
<td>×</td>
<td>$(a &lt; b) \nRightarrow ((c + a) &lt; (c + b))$</td>
</tr>
</tbody>
</table>
Floating-Point Comparisons and Precision Thresholds

- Topological decisions in geometry are based on the results of floating-point (fp) computations, which are prone to round-off errors.
- Comparing two fp-numbers $a$ and $b$ by means of $a = b$ will hardly ever yield true.
Floating-Point Comparisons and Precision Thresholds

- Topological decisions in geometry are based on the results of floating-point (fp) computations, which are prone to round-off errors.
- Comparing two fp-numbers $a$ and $b$ by means of $a = b$ will hardly ever yield true.
- Threshold-based comparison:

  $$(a =_\varepsilon b) \iff (|a - b| \leq \varepsilon),$$

  for some positive value of $\varepsilon$. 

Note: $|x - y| \leq \varepsilon$ rather than $|x - y| < \varepsilon$!

Caveat: $=_\varepsilon$ is no longer transitive: $a =_\varepsilon b$ and $b =_\varepsilon c$ need not imply $a =_\varepsilon c$.

Note: fp-numbers are “denser” close to zero than far away from zero.

Note: $|x - y| \leq \varepsilon$ need not imply $|\alpha \cdot x - \alpha \cdot y| \leq \varepsilon$.

Thus, use relative errors or scale the data appropriately.

Obvious disadvantage of scaling: Unless only shifts by two are performed, new errors may be introduced.
Floating-Point Comparisons and Precision Thresholds

- Topological decisions in geometry are based on the results of floating-point (fp) computations, which are prone to round-off errors.
- Comparing two fp-numbers $a$ and $b$ by means of $a = b$ will hardly ever yield true.
- Threshold-based comparison:

\[(a \approx \epsilon \ b) \iff (|a - b| \leq \epsilon),\]

for some positive value of $\epsilon$.
- Note: $|a - b| \leq \epsilon$ rather than $|a - b| < \epsilon$!
Floating-Point Comparisons and Precision Thresholds

- Topological decisions in geometry are based on the results of floating-point (fp) computations, which are prone to round-off errors.
- Comparing two fp-numbers \( a \) and \( b \) by means of \( a = b \) will hardly ever yield true.
- Threshold-based comparison:
  \[
  (a =_\varepsilon b) \iff (|a - b| \leq \varepsilon),
  \]
  for some positive value of \( \varepsilon \).
- Note: \( |a - b| \leq \varepsilon \) rather than \( |a - b| < \varepsilon \)!
- Caveat: \( =_\varepsilon \) is no longer transitive: \( a =_\varepsilon b \) and \( b =_\varepsilon c \) need not imply \( a =_\varepsilon c \).

Note: fp-numbers are “denser” close to zero than far away from zero.
Floating-Point Comparisons and Precision Thresholds

- Topological decisions in geometry are based on the results of floating-point (fp) computations, which are prone to round-off errors.
- Comparing two fp-numbers \(a\) and \(b\) by means of \(a = b\) will hardly ever yield true.
- Threshold-based comparison:

\[
(a =_\varepsilon b) :\iff (|a - b| \leq \varepsilon),
\]

for some positive value of \(\varepsilon\).
- Note: \(|a - b| \leq \varepsilon\) rather than \(|a - b| < \varepsilon\)!
- Caveat: \(=_\varepsilon\) is no longer transitive: \(a =_\varepsilon b\) and \(b =_\varepsilon c\) need not imply \(a =_\varepsilon c\).
- Note: fp-numbers are “denser” close to zero than far away from zero.
- Note: \(|x - y| \leq \varepsilon\) need not imply \(|\alpha \cdot x - \alpha \cdot y| \leq \varepsilon\).
Floating-Point Comparisons and Precision Thresholds

- Topological decisions in geometry are based on the results of floating-point (fp) computations, which are prone to round-off errors.
- Comparing two fp-numbers $a$ and $b$ by means of $a = b$ will hardly ever yield true.
- Threshold-based comparison:
  \[ (a =_\varepsilon b) \iff (|a - b| \leq \varepsilon), \]
  for some positive value of $\varepsilon$.
- Note: $|a - b| \leq \varepsilon$ rather than $|a - b| < \varepsilon$!
- Caveat: $=_\varepsilon$ is no longer transitive: $a =_\varepsilon b$ and $b =_\varepsilon c$ need not imply $a =_\varepsilon c$.
- Note: fp-numbers are “denser” close to zero than far away from zero.
- Note: $|x - y| \leq \varepsilon$ need not imply $|\alpha \cdot x - \alpha \cdot y| \leq \varepsilon$.
- Thus, use relative errors or scale the data appropriately.
- Obvious disadvantage of scaling: Unless only shifts by two are performed, new errors may be introduced.
Sample Robustness Problem: Lack of Convergence

- Theory tells us that we can approximate the first derivative $f'$ of a function $f$ at the point $x_0$ by evaluating $\frac{f(x_0+h) - f(x_0)}{h}$ for sufficiently small values of $h$. 

By considering the function $f(x) := x^3$ and $x_0 := 10$, we have the following approximations for $f'(10)$:

- $h := 10^{-1}$: $f'(10) \approx 303.0099999$
- $h := 10^{-3}$: $f'(10) \approx 300.0300009$
- $h := 10^{-5}$: $f'(10) \approx 300.0002999$
- $h := 10^{-7}$: $f'(10) \approx 300.0000003$
- $h := 10^{-9}$: $f'(10) \approx 300.0000010$
- $h := 10^{-11}$: $f'(10) \approx 299.9854586$
- $h := 10^{-13}$: $f'(10) \approx 298.9963832$
- $h := 10^{-15}$: $f'(10) \approx 568.4341886$
- $h := 10^{-17}$: $f'(10) \approx 0.000000000$

The cancellation error increases as the step size, $h$, decreases. On the other hand, the truncation error decreases as $h$ decreases. These two opposing effects result in a minimum error (and “best” step size $h$) that is high above the machine precision!
### Sample Robustness Problem: Lack of Convergence

- Theory tells us that we can approximate the first derivative $f'$ of a function $f$ at the point $x_0$ by evaluating $\frac{f(x_0+h)-f(x_0)}{h}$ for sufficiently small values of $h$.
- Consider $f(x) := x^3$ and $x_0 := 10$:

<table>
<thead>
<tr>
<th>$h$</th>
<th>$f'(10)$</th>
<th>$h$</th>
<th>$f'(10)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^0$</td>
<td>$331.00000000$</td>
<td>$10^{-1}$</td>
<td>$303.00999999$</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>$300.3000999$</td>
<td>$10^{-3}$</td>
<td>$300.0300009$</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>$300.0030000$</td>
<td>$10^{-5}$</td>
<td>$300.0002999$</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>$300.000298$</td>
<td>$10^{-7}$</td>
<td>$300.0000003$</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>$300.000219$</td>
<td>$10^{-9}$</td>
<td>$300.0000106$</td>
</tr>
<tr>
<td>$10^{-10}$</td>
<td>$300.002379$</td>
<td>$10^{-11}$</td>
<td>$299.9854586$</td>
</tr>
<tr>
<td>$10^{-12}$</td>
<td>$300.1332515$</td>
<td>$10^{-13}$</td>
<td>$298.9963832$</td>
</tr>
<tr>
<td>$10^{-14}$</td>
<td>$318.3231456$</td>
<td>$10^{-15}$</td>
<td>$568.4341886$</td>
</tr>
<tr>
<td>$10^{-16}$</td>
<td>$0.000000000$</td>
<td>$10^{-17}$</td>
<td>$0.000000000$</td>
</tr>
</tbody>
</table>

The cancellation error increases as the step size, $h$, decreases. On the other hand, the truncation error decreases as $h$ decreases. These two opposing effects result in a minimum error (and "best" step size $h$) that is high above the machine precision!
Sample Robustness Problem: Lack of Convergence

- Theory tells us that we can approximate the first derivative $f'$ of a function $f$ at the point $x_0$ by evaluating $\frac{f(x_0 + h) - f(x_0)}{h}$ for sufficiently small values of $h$.
- Consider $f(x) := x^3$ and $x_0 := 10$:

  $h := 10^0 : \quad f'(10) \approx 331.0000000 \quad h := 10^{-1} : \quad f'(10) \approx 303.0099999$

  $h := 10^{-2} : \quad f'(10) \approx 300.3000999 \quad h := 10^{-3} : \quad f'(10) \approx 300.0300009$

  $h := 10^{-4} : \quad f'(10) \approx 300.0030000 \quad h := 10^{-5} : \quad f'(10) \approx 300.0002999$

  $h := 10^{-6} : \quad f'(10) \approx 300.0000298 \quad h := 10^{-7} : \quad f'(10) \approx 300.0000003$

  $h := 10^{-8} : \quad f'(10) \approx 300.0000219 \quad h := 10^{-9} : \quad f'(10) \approx 300.0000106$

  $h := 10^{-10} : \quad f'(10) \approx 300.0002379 \quad h := 10^{-11} : \quad f'(10) \approx 299.9854586$

  $h := 10^{-12} : \quad f'(10) \approx 300.1332515 \quad h := 10^{-13} : \quad f'(10) \approx 298.9963832$

  $h := 10^{-14} : \quad f'(10) \approx 318.3231456 \quad h := 10^{-15} : \quad f'(10) \approx 568.4341886$

  $h := 10^{-16} : \quad f'(10) \approx 0.000000000 \quad h := 10^{-17} : \quad f'(10) \approx 0.000000000$

- The cancellation error increases as the step size, $h$, decreases. On the other hand, the truncation error decreases as $h$ decreases.
- These two opposing effects result in a minimum error (and “best” step size $h$) that is high above the machine precision!
The quartic equation $x^4 + 4x^3 + 6x^2 + 4x + 1 = 0$ has the quadruple root $x = -1$.
Sample Robustness Problem: Ill-Conditioned Equations

- The quartic equation $x^4 + 4x^3 + 6x^2 + 4x + 1 = 0$ has the quadruple root $x = -1$.
- Changing the coefficient of $x$ to $4.00000001$ drastically affects the solution: Now we get $x = -1.01002496875 \ldots$ and $x = -0.99002503124 \ldots$ as the only real roots of the equation.
Sample Robustness Problem: Ill-Conditioned Equations

- The quartic equation $x^4 + 4x^3 + 6x^2 + 4x + 1 = 0$ has the quadruple root $x = -1$.
- Changing the coefficient of $x$ to 4.00000001 drastically affects the solution: Now we get $x = -1.01002496875\ldots$ and $x = -0.99002503124\ldots$ as the only real roots of the equation.
- Similarly, the linear system
  \[
  \begin{align*}
  x + 2y & = 3 \\
  0.48x + 0.99y & = 1.47
  \end{align*}
  \]
  has the exact solution $x = 1, y = 1$,
The quartic equation \( x^4 + 4x^3 + 6x^2 + 4x + 1 = 0 \) has the quadruple root \( x = -1 \).

Changing the coefficient of \( x \) to 4.00000001 drastically affects the solution: Now we get \( x = -1.01002496875 \ldots \) and \( x = -0.99002503124 \ldots \) as the only real roots of the equation.

Similarly, the linear system
\[
\begin{align*}
    x + 2y &= 3 \\
    0.48x + 0.99y &= 1.47
\end{align*}
\]
has the exact solution \( x = 1, y = 1 \), while the system
\[
\begin{align*}
    x + 2y &= 3 \\
    0.49x + 0.99y &= 1.47
\end{align*}
\]
has the exact solution \( x = 3, y = 0 \).
Sample Robustness Problem: Ill-Conditioned Equations

- The quartic equation \( x^4 + 4x^3 + 6x^2 + 4x + 1 = 0 \) has the quadruple root \( x = -1 \).
- Changing the coefficient of \( x \) to \( 4.00000001 \) drastically affects the solution: Now we get \( x = -1.01002496875 \ldots \) and \( x = -0.99002503124 \ldots \) as the only real roots of the equation.
- Similarly, the linear system
  \[
  \begin{align*}
  x + 2y &= 3 \\
  0.48x + 0.99y &= 1.47 
  \end{align*}
  \]
  has the exact solution \( x = 1, y = 1 \), while the system
  \[
  \begin{align*}
  x + 2y &= 3 \\
  0.49x + 0.99y &= 1.47 
  \end{align*}
  \]
  has the exact solution \( x = 3, y = 0 \).
- Note, however, that the old solution, \( x = 1, y = 1 \), also “nearly” fulfills this equation.
Sample Robustness Problem: Ill-Conditioned Equations

- The quartic equation $x^4 + 4x^3 + 6x^2 + 4x + 1 = 0$ has the quadruple root $x = -1$.
- Changing the coefficient of $x$ to $4.00000001$ drastically affects the solution: Now we get $x = -1.01002496875\ldots$ and $x = -0.99002503124\ldots$ as the only real roots of the equation.
- Similarly, the linear system
  
  \begin{align*}
  x + 2y & = 3 \\
  0.48x + 0.99y & = 1.47
  \end{align*}

  has the exact solution $x = 1, y = 1$, while the system
  
  \begin{align*}
  x + 2y & = 3 \\
  0.49x + 0.99y & = 1.47
  \end{align*}

  has the exact solution $x = 3, y = 0$.
- Note, however, that the old solution, $x = 1, y = 1$, also “nearly” fulfills this equation.
- Thus, a small change (or error!) in the coefficients can dramatically affect the solutions of an equation: *ill-conditioned* or *ill-posed*!
- An $n \times n$ system of linear equations is ill-conditioned if and only if the determinant of the coefficient matrix is close to zero.
Sample Robustness Problem: Ill-Conditioned Equations

- If an equation (or a system of equations) is ill-conditioned, then the usual procedure of checking a numerical solution by calculation of the residuals is problematic.
- Consider the $2 \times 2$ linear system

\[
\begin{align*}
1.2969x + 0.8648y &= 0.8642 \\
0.2161x + 0.1441y &= 0.1440
\end{align*}
\]

that is, \( A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \).

- The exact solution is \( x = 2 \) and \( y = -2 \).
Sample Robustness Problem: Ill-Conditioned Equations

• If an equation (or a system of equations) is ill-conditioned, then the usual procedure of checking a numerical solution by calculation of the residuals is problematic.

• Consider the $2 \times 2$ linear system

$$
1.2969x + 0.8648y = 0.8642 \\
0.2161x + 0.1441y = 0.1440
$$

that is, $A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

• The exact solution is $x = 2$ and $y = -2$.

• But we get close-to-zero residuals also for other pairs of $x$ and $y$:

$x_1 = 0.9911 \\ y_1 = -0.4870 \\ \| A \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \| \approx 10^{-8}$

$x_2 = 2.001557851 \\ y_2 = -2.002336236 \\ \| A \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \| \approx 10^{-10}$

$x_3 = -0.000004626 \\ y_3 = 0.999312976 \\ \| A \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \| \approx 10^{-9}$
Sample Robustness Problem: Incorrect Orientation Predicate

- [Kettner et alii 2006] study the standard determinant-based orientation predicate on IEEE 754 fp-arithmetic to check the sidedness of \((p_x + x \cdot u, p_y + y \cdot u)\) relative to two points \(q, r\), for \(0 \leq x, y \leq 255\) and with \(u := 2^{-53}\):

\[
sign \det \begin{pmatrix}
1 & p_x + x \cdot u & p_y + y \cdot u \\
1 & q_x & q_y \\
1 & r_x & r_y
\end{pmatrix} \begin{cases}
> & \quad 0 \\
= & \\
< & 
\end{cases}
\]

The resulting \(256 \times 256\) array of signs (as a function of \(x, y\)) is color-coded: A yellow (red, blue) pixel indicates collinear (negative, positive, resp.) orientation. The black line indicates the line through \(q\) and \(r\). Note the sign inversions!

[Image credit: www.mpi-inf.mpg.de/~kettner/proj/NonRobust/]
Sample Robustness Problem: Incorrect Orientation Predicate

- [Kettner et alii 2006] study the standard determinant-based orientation predicate on IEEE 754 fp-arithmetic to check the sidedness of \((p_x + x \cdot u, p_y + y \cdot u)\) relative to two points \(q, r\), for \(0 \leq x, y \leq 255\) and with \(u := 2^{-53}\):

\[
\text{sign det } \begin{pmatrix} 1 & p_x + x \cdot u & p_y + y \cdot u \\ 1 & q_x & q_y \\ 1 & r_x & r_y \end{pmatrix} \begin{cases} > 0 \? \\ = 0 \? \\ < 0 \? \end{cases}
\]

- The resulting \(256 \times 256\) array of signs (as a function of \(x, y\)) is color-coded: A yellow (red, blue) pixel indicates collinear (negative, positive, resp.) orientation.

- The black line indicates the line through \(q\) and \(r\).

- Note the sign inversions!

[Image credit: www.mpi-inf.mpg.de/~kettner/proj/NonRobust/]

\[\text{www.mpi-inf.mpg.de/~kettner/proj/NonRobust/}\]
Sample Robustness Problem: Incorrect Orientation Predicate

- [Kettner et alii 2006]: A yellow (red, blue) pixel indicates collinear (negative, positive, resp.) orientation.

\[ p := \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \quad q := \begin{pmatrix} 12 \\ 12 \end{pmatrix} \quad r := \begin{pmatrix} 24 \\ 24 \end{pmatrix} \]
Sample Robustness Problem: Incorrect Orientation Predicate

- [Kettner et alii 2006]: A yellow (red, blue) pixel indicates collinear (negative, positive, resp.) orientation.

\[
p := \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \quad q := \begin{pmatrix} 8.8000000000000007 \\ 8.8000000000000007 \end{pmatrix} \quad r := \begin{pmatrix} 12.1 \\ 12.1 \end{pmatrix}
\]
Floating-point computation is by nature inexact, and it is not difficult to misuse it so that the computed answers consist almost entirely of ‘noise’.

One of the principal problems of numerical analysis is to determine how accurate the results of certain numerical methods will be; a ‘credibility gap’ problem is involved here: we don’t know how much of the computer’s answers to believe.

Novice computer users solve this problem by implicitly trusting in the computer as an infallible authority; they tend to believe all digits of a printed answer are significant.

Disillusioned computer users have just the opposite approach, they are constantly afraid their answers are almost meaningless.
The gap between the theory of the reals and floating-point practice has important and severe consequences for the actual coding practice when implementing (geometric) algorithms that require floating-point arithmetic:

1. The correctness proof of the mathematical algorithm does not extend to the program, and the program can fail on seemingly appropriate input data.
Floating-Point Comparisons and Precision Thresholds

- The gap between the theory of the reals and floating-point practice has important and severe consequences for the actual coding practice when implementing (geometric) algorithms that require floating-point arithmetic:
  - The correctness proof of the mathematical algorithm does not extend to the program, and the program can fail on seemingly appropriate input data.
  - Local consistency need not imply global consistency.
Floating-Point Comparisons and Precision Thresholds

The gap between the theory of the reals and floating-point practice has important and severe consequences for the actual coding practice when implementing (geometric) algorithms that require floating-point arithmetic:

1. The correctness proof of the mathematical algorithm does not extend to the program, and the program can fail on seemingly appropriate input data.
2. Local consistency need not imply global consistency.

Numerical analysis . . .

. . . and adequate coding are a must when implementing algorithms that deal with real numbers. Otherwise, the implementation of an algorithm may turn out to be absolutely useless in practice, even if the algorithm (and even its implementation) would come with a rigorous mathematical proof of correctness!
Improving the Reliability of FP-Calculations

- Try to perform all numerical computations relative to the original input data.
Improving the Reliability of FP-Calculations

- Try to perform all numerical computations relative to the original input data.
- Make sure that different calls of the same function with the “same” input will yield exactly the same output. E.g., guaranteeing

\[
\det(P, Q, R) = \det(Q, R, P) = -\det(Q, P, R)
\]

is a must (for points \( P, Q, R \in \mathbb{R}^2 \)).
Improving the Reliability of FP-Calculations

- Try to perform all numerical computations relative to the original input data.
- Make sure that different calls of the same function with the “same” input will yield exactly the same output. E.g., guaranteeing
  \[ \det(P, Q, R) = \det(Q, R, P) = -\det(Q, P, R) \]
  is a must (for points \( P, Q, R \in \mathbb{R}^2 \)).
- Do not resort to multiple precision thresholds.
- At most two thresholds: One to avoid divisions by zero, and possibly another threshold to catch “nearly zero” numbers.
Improving the Reliability of FP-Calculations

- Try to perform all numerical computations relative to the original input data.
- Make sure that different calls of the same function with the “same” input will yield exactly the same output. E.g., guaranteeing
  \[ \det(P, Q, R) = \det(Q, R, P) = -\det(Q, P, R) \]
  is a must (for points \( P, Q, R \in \mathbb{R}^2 \)).
- Do not resort to multiple precision thresholds.
- At most two thresholds: One to avoid divisions by zero, and possibly another threshold to catch “nearly zero” numbers.
- Take a close look at your calculations: Different terms might be arithmetically identical, but their numerical behavior may be substantially different, and one term may be far better than the other!
Improving the Reliability of FP-Calculations

- Try to perform all numerical computations relative to the original input data.
- Make sure that different calls of the same function with the “same” input will yield exactly the same output. E.g., guaranteeing
  \[
  \det(P, Q, R) = \det(Q, R, P) = -\det(Q, P, R)
  \]
  is a must (for points \( P, Q, R \in \mathbb{R}^2 \)).
- Do not resort to multiple precision thresholds.
- At most two thresholds: One to avoid divisions by zero, and possibly another threshold to catch “nearly zero” numbers.
- Take a close look at your calculations: Different terms might be arithmetically identical, but their numerical behavior may be substantially different, and one term may be far better than the other!
- E.g., compute a finite series by starting with the smallest rather than with the largest summand:
  \[
  1 + \frac{1}{2} + \ldots + \frac{1}{1000000} \approx 14.3927267228649889 \\
  \frac{1}{1000000} + \ldots + \frac{1}{2} + 1 \approx 14.3927267228657723
  \]
Improving the Reliability of FP-Calculations

- Try to perform all numerical computations relative to the original input data.
- Make sure that different calls of the same function with the “same” input will yield exactly the same output. E.g., guaranteeing
  \[
  \det(P, Q, R) = \det(Q, R, P) = -\det(Q, P, R)
  \]
is a must (for points \(P, Q, R \in \mathbb{R}^2\)).
- Do not resort to multiple precision thresholds.
- At most two thresholds: One to avoid divisions by zero, and possibly another threshold to catch “nearly zero” numbers.
- Take a close look at your calculations: Different terms might be arithmetically identical, but their numerical behavior may be substantially different, and one term may be far better than the other!
- E.g., compute a finite series by starting with the smallest rather than with the largest summand:
  \[
  1 + \frac{1}{2} + \ldots + \frac{1}{1000000} \approx 14.3927267228649889
  \]
  \[
  \frac{1}{1000000} + \ldots + \frac{1}{2} + 1 \approx 14.3927267228657723
  \]
  Mathematica: \(\approx 14.39272672286572363138\ldots\)
Improving the Reliability of FP-Calculations: Quadratic Equations

- Different terms might be arithmetically identical, but their numerical behavior may be substantially different, and one term may be far better than the other!

- Mathematics tells us that the solutions of the quadratic equation $ax^2 + bx + c = 0$ are given by

$$x_{1,2} := \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$
Improving the Reliability of FP-Calculations: Quadratic Equations

- Different terms might be arithmetically identical, but their numerical behavior may be substantially different, and one term may be far better than the other!

- Mathematics tells us that the solutions of the quadratic equation 
  \[ ax^2 + bx + c = 0 \]
  are given by
  \[ x_{1,2} := \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \]

- Unfortunately, using this formula means begging for troubles if \(|a \cdot c|\) is small compared to \(|b|\), since the subtraction of \(\sqrt{b^2 - 4ac}\) from \(b\) may cause serious cancellation.
Improving the Reliability of FP-Calculations: Quadratic Equations

- Different terms might be arithmetically identical, but their numerical behavior may be substantially different, and one term may be far better than the other!

- Mathematics tells us that the solutions of the quadratic equation \( ax^2 + bx + c = 0 \) are given by
  \[
  x_{1,2} := \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.
  \]

- Unfortunately, using this formula means begging for troubles if \( |a \cdot c| \) is small compared to \( |b| \), since the subtraction of \( \sqrt{b^2 - 4ac} \) from \( b \) may cause serious cancellation.

- Better: Let
  \[
  \Delta := -\frac{1}{2} \left( b + \text{sign}(b) \sqrt{b^2 - 4ac} \right).
  \]

  Then the roots are obtained more reliably as
  \[
  x_1 := \frac{\Delta}{a} \quad \text{and} \quad x_2 := \frac{c}{\Delta}. \quad \text{(This is a consequence of Viète’s formulas.)}
  \]
Improving the Reliability of FP-Calculations: Quadratic Equations

- E.g., consider the equation $x^2 + 10^4 x + 10^{-9} = 0$.
- The classical formula yields

  $$x_1 \approx -10000.00000000000000000000000000000000000$$
  $$x_2 \approx -0.00000000000000000000000000000000000$$

According to Mathematica, the true solution is

  $$x_1 \approx -9999.9999999999000000000000$$
  $$x_2 \approx -0.00000000000000000000000000000000000$$
E.g., consider the equation $x^2 + 10^4 x + 10^{-9} = 0$.

The classical formula yields

$$x_1 \approx -10000.0000000000000000000000000000000,$$
$$x_2 \approx -0.0000000000000000000000000000000.$$

The refined approach yields

$$x_1 \approx -10000.0000000000000000000000000000000,$$
$$x_2 \approx -0.0000000000001000000000000000030.$$
E.g., consider the equation $x^2 + 10^4 x + 10^{-9} = 0$.

The classical formula yields

$$x_1 \approx -10000.00000000000000000000000000000000,$$
$$x_2 \approx -0.000000000000000000000000000030.$$

The refined approach yields

$$x_1 \approx -10000.00000000000000000000000000000000,$$
$$x_2 \approx -0.0000000000010000000000000000030.$$

According to Mathematica, the true solution is

$$x_1 \approx -9999.999999999999990000000000000000,$$
$$x_2 \approx -0.0000000000010000000000000000010.$$
Floating-Point Arithmetic and Numerical Mathematics

- Floating-Point Computations
- Iterative Algorithms for Solving Non-Linear Equations
  - Basics of Iterative Algorithms
  - Bisection
  - Regula Falsi
  - Newton-Raphson Method
  - Secant Method
- Iterative Algorithms for Solving Linear Equations
- Numerical Integration
Iterative Algorithms for Solving Non-Linear Equations

- We are interested in solving the equation $f(x) = 0$, for a function $f : \mathbb{R} \rightarrow \mathbb{R}$. This means finding all $x \in \mathbb{R}$ for which $f(x) = 0$.
- Explicit (algebraic) root-finding is possible for polynomial equations of degree less than five.
Iterative Algorithms for Solving Non-Linear Equations

- We are interested in solving the equation $f(x) = 0$, for a function $f : \mathbb{R} \to \mathbb{R}$. This means finding all $x \in \mathbb{R}$ for which $f(x) = 0$.
- Explicit (algebraic) root-finding is possible for polynomial equations of degree less than five.
- For other types of non-linear equations, dozens of iterative methods have been proposed.
- Two basic schemes:
  - Bracketing: e.g., bisection, regula falsi;
  - Polishing: e.g., Newton-Raphson method, secant method.
- Extensions to vector-valued functions are possible.
Basics of Iterative Root Finding

- We attempt to compute a sequence \((x_k)_{k=0}^\infty\), depending on some initial value(s) \(x_0\) resp. \(x_0, x_1\) and on \(f\) and its derivatives.
- Ideally, \(\lim_{k \to \infty} x_k = \bar{x}\).

Question: How shall we find a suitable initial value \(x_0\)?
Answer: Study the function \(f\).

Question: What is a suitable initial value \(x_0\)?
Answer: Whether or not \(x_0\) is suitable depends on \(f\) and on the iteration method used.

Question: Is the iteration guaranteed to converge?
Answer: Unfortunately, no – unless specific criteria are fulfilled.

Question: Is the iteration guaranteed to find all roots?
Answer: At best, an iteration method will find one root at a time.

Question: How quickly will the iteration converge?
Answer: This depends on the convergence rate of the iteration method, see later.

General advice
Do not use iteration methods on a function you do not know much about. In particular, do not use an iteration method to test whether a root exists in the neighborhood of some initial value.
Basics of Iterative Root Finding

- We attempt to compute a sequence \((x_k)_{k=0}^{\infty}\), depending on some initial value(s) \(x_0\) resp. \(x_0, x_1\) and on \(f\) and its derivatives.
- Ideally, \(\lim_{k \to \infty} x_k = \overline{x}\).
- Question: How shall we find a suitable initial value \(x_0\)?
  Answer: Study the function \(f\).

General advice
Do not use iteration methods on a function you do not know much about.
In particular, do not use an iteration method to test whether a root exists in the neighborhood of some initial value.
Basics of Iterative Root Finding

- We attempt to compute a sequence \((x_k)_{k=0}^{\infty}\), depending on some initial value(s) \(x_0\) resp. \(x_0, x_1\) and on \(f\) and its derivatives.
- Ideally, \(\lim_{k \to \infty} x_k = x\).
- Question: How shall we find a suitable initial value \(x_0\)?
  Answer: Study the function \(f\).
- Question: What is a suitable initial value \(x_0\)?
  Answer: Whether or not \(x_0\) is suitable depends on \(f\) and on the iteration method used.
Basics of Iterative Root Finding

- We attempt to compute a sequence \((x_k)_{k=0}^{\infty}\), depending on some initial value(s) \(x_0\) resp. \(x_0, x_1\) and on \(f\) and its derivatives.
- Ideally, \(\lim_{k \to \infty} x_k = \bar{x}\).
- Question: How shall we find a suitable initial value \(x_0\)?
  Answer: Study the function \(f\).
- Question: What is a suitable initial value \(x_0\)?
  Answer: Whether or not \(x_0\) is suitable depends on \(f\) and on the iteration method used.
- Question: Is the iteration guaranteed to converge?
  Answer: Unfortunately, no – unless specific criteria are fulfilled.
Basics of Iterative Root Finding

- We attempt to compute a sequence \((x_k)_{k=0}^{\infty}\), depending on some initial value(s) \(x_0\) resp. \(x_0, x_1\) and on \(f\) and its derivatives.
- Ideally, \(\lim_{k \to \infty} x_k = x\).
- Question: How shall we find a suitable initial value \(x_0\)?
  Answer: Study the function \(f\).
- Question: What is a suitable initial value \(x_0\)?
  Answer: Whether or not \(x_0\) is suitable depends on \(f\) and on the iteration method used.
- Question: Is the iteration guaranteed to converge?
  Answer: Unfortunately, no – unless specific criteria are fulfilled.
- Question: Is the iteration guaranteed to find all roots?
  Answer: At best, an iteration method will find one root at a time.
Basics of Iterative Root Finding

- We attempt to compute a sequence \((x_k)_{k=0}^{\infty}\), depending on some initial value(s) \(x_0\) resp. \(x_0, x_1\) and on \(f\) and its derivatives.
- Ideally, \(\lim_{k \to \infty} x_k = x\).
- Question: How shall we find a suitable initial value \(x_0\)?
  Answer: Study the function \(f\).
- Question: What is a suitable initial value \(x_0\)?
  Answer: Whether or not \(x_0\) is suitable depends on \(f\) and on the iteration method used.
- Question: Is the iteration guaranteed to converge?
  Answer: Unfortunately, no – unless specific criteria are fulfilled.
- Question: Is the iteration guaranteed to find all roots?
  Answer: At best, an iteration method will find one root at a time.
- Question: How quickly will the iteration converge?
  Answer: This depends on the convergence rate of the iteration method, see later.

General advice

Do not use iteration methods on a function you do not know much about.
In particular, do not use an iteration method to test whether a root exists in the neighborhood of some initial value.
Basics of Iterative Root Finding

- We attempt to compute a sequence \((x_k)_{k=0}^{\infty}\), depending on some initial value(s) \(x_0\) resp. \(x_0, x_1\) and on \(f\) and its derivatives.
- Ideally, \(\lim_{k \to \infty} x_k = x\).
- Question: How shall we find a suitable initial value \(x_0\)?
  Answer: Study the function \(f\).
- Question: What is a suitable initial value \(x_0\)?
  Answer: Whether or not \(x_0\) is suitable depends on \(f\) and on the iteration method used.
- Question: Is the iteration guaranteed to converge?
  Answer: Unfortunately, no – unless specific criteria are fulfilled.
- Question: Is the iteration guaranteed to find all roots?
  Answer: At best, an iteration method will find one root at a time.
- Question: How quickly will the iteration converge?
  Answer: This depends on the convergence rate of the iteration method, see later.

General advice

Do not use iteration methods on a function you do not know much about.
Basics of Iterative Root Finding

- We attempt to compute a sequence \((x_k)_{k=0}^{\infty}\), depending on some initial value(s) \(x_0\) resp. \(x_0, x_1\) and on \(f\) and its derivatives.
- Ideally, \(\lim_{k \to \infty} x_k = \bar{x}\).
- Question: How shall we find a suitable initial value \(x_0\)?
  Answer: Study the function \(f\).
- Question: What is a suitable initial value \(x_0\)?
  Answer: Whether or not \(x_0\) is suitable depends on \(f\) and on the iteration method used.
- Question: Is the iteration guaranteed to converge?
  Answer: Unfortunately, no – unless specific criteria are fulfilled.
- Question: Is the iteration guaranteed to find all roots?
  Answer: At best, an iteration method will find one root at a time.
- Question: How quickly will the iteration converge?
  Answer: This depends on the convergence rate of the iteration method, see later.

General advice
Do not use iteration methods on a function you do not know much about. In particular, do not use an iteration method to test whether a root exists in the neighborhood of some initial value.
Basics of Iterative Root Finding

- How can we state how rapidly a sequence \((x_k)_{k=0}^{\infty}\) converges to the root \(\bar{x}\)?
Basics of Iterative Root Finding

- How can we state how rapidly a sequence \((x_k)_{k=0}^{\infty}\) converges to the root \(\bar{x}\)?

**Definition 254 (Convergence rate, Dt.: Konvergenzrate)**

Let \((x_k)_{k=0}^{\infty}\) be a sequence that is used to approximate a root \(\bar{x}\), and let \(e_k := \bar{x} - x_k\) be the error of the \(k\)-th approximation \(x_k\) of \(\bar{x}\). The *convergence rate* of an iteration method is the largest exponent \(p\) such that

\[
\lim_{k \to \infty} \frac{|e_k|}{|e_{k-1}|^p} = c,
\]

for a suitable asymptotic error constant \(c \in \mathbb{R}^+\).
Basics of Iterative Root Finding

- How can we state how rapidly a sequence \((x_k)_{k=0}^{\infty}\) converges to the root \(\overline{x}\)?

**Definition 254 (Convergence rate, Dt.: Konvergenzrate)**

Let \((x_k)_{k=0}^{\infty}\) be a sequence that is used to approximate a root \(\overline{x}\), and let \(e_k := \overline{x} - x_k\) be the error of the \(k\)-th approximation \(x_k\) of \(\overline{x}\). The convergence rate of an iteration method is the largest exponent \(p\) such that

\[
\lim_{k \to \infty} \frac{|e_k|}{|e_{k-1}|^p} = c,
\]

for a suitable asymptotic error constant \(c \in \mathbb{R}^+\).

- If \(p = 1\) then the convergence is called *linear*.
- If \(p = 2\) then the convergence is called *quadratic*.
- If \(1 < p < 2\) then the convergence is called *super-linear*.
Basics of Iterative Root Finding

- How can we state how rapidly a sequence \((x_k)_{k=0}^\infty\) converges to the root \(\overline{x}\)?

**Definition 254 (Convergence rate, Dt.: Konvergenzrate)**

Let \((x_k)_{k=0}^\infty\) be a sequence that is used to approximate a root \(\overline{x}\), and let \(e_k := \overline{x} - x_k\) be the error of the \(k\)-th approximation \(x_k\) of \(\overline{x}\). The *convergence rate* of an iteration method is the largest exponent \(p\) such that

\[
\lim_{k \to \infty} \frac{|e_k|}{|e_{k-1}|^p} = c,
\]

for a suitable asymptotic error constant \(c \in \mathbb{R}^+\).

If \(p = 1\) then the convergence is called *linear*.

If \(p = 2\) then the convergence is called *quadratic*.

If \(1 < p < 2\) then the convergence is called *super-linear*.

- Linear convergence means that the error is reduced by a constant factor per iteration, i.e., that the number of correct digits increases by one after a constant number of iterations.

- Quadratic convergence means that the number of correct digits roughly doubles with each iteration.
Bisection

- Consider a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$, and assume that for $a, b \in \mathbb{R}$ you know $\text{sign}(f(a)) = -\text{sign}(f(b))$, with $a < b$ and $f(a) \cdot f(b) \neq 0$.

- Intermediate Value Theorem: Since we have opposite signs for $f$ at $a, b$, and $f$ is continuous, we conclude that $f$ has at least one root $\bar{x}$ in the interval $[a, b]$.

By checking the sign of $f \left(\frac{a+b}{2}\right)$ and appropriately replacing $a$ or $b$ by $\frac{a+b}{2}$, this interval is halved at each step of the iteration:

- if $\text{sign}(f(\frac{a+b}{2})) = 0$ then $x := \frac{a+b}{2}$, stop;
- if $\text{sign}(f(\frac{a+b}{2})) = \text{sign}(f(a))$ then $a := \frac{a+b}{2}$;
- if $\text{sign}(f(\frac{a+b}{2})) = \text{sign}(f(b))$ then $b := \frac{a+b}{2}$.

Since bisection traps a root, it is guaranteed to converge. However, it needs at least three iterations to achieve one additional significant digit of the root!

Caveat: Although several roots might exist within the interval $[a, b]$, only one root will be found.

Caveat: Root-bracketing is not feasible for finding even-multiplicity roots.
**Bisection**

- Consider a continuous function $f : \mathbb{R} \to \mathbb{R}$, and assume that for $a, b \in \mathbb{R}$ you know $\text{sign}(f(a)) = -\text{sign}(f(b))$, with $a < b$ and $f(a) \cdot f(b) \neq 0$.

- Intermediate Value Theorem: Since we have opposite signs for $f$ at $a, b$, and $f$ is continuous, we conclude that $f$ has at least one root $\bar{x}$ in the interval $[a, b]$.

- By checking the sign of $f\left(\frac{a+b}{2}\right)$ and appropriately replacing $a$ or $b$ by $\frac{a+b}{2}$, this interval is halved at each step of the iteration:

  \[
  \text{if } \text{sign}(f\left(\frac{a+b}{2}\right)) = 0 \quad \text{then } \bar{x} := \frac{a+b}{2}, \text{ stop;}
  \]
  \[
  \text{if } \text{sign}(f\left(\frac{a+b}{2}\right)) = \text{sign}(f(a)) \quad \text{then } a := \frac{a+b}{2};
  \]
  \[
  \text{if } \text{sign}(f\left(\frac{a+b}{2}\right)) = \text{sign}(f(b)) \quad \text{then } b := \frac{a+b}{2}.
  \]

Since bisection traps a root, it is guaranteed to converge. However, it needs at least three iterations to achieve one additional significant digit of the root!

Caveat: Although several roots might exist within the interval $[a, b]$, only one root will be found. Caveat: Root-bracketing is not feasible for finding even-multiplicity roots.
Bisection

- Consider a continuous function \( f: \mathbb{R} \rightarrow \mathbb{R} \), and assume that for \( a, b \in \mathbb{R} \) you know \( \text{sign}(f(a)) = -\text{sign}(f(b)) \), with \( a < b \) and \( f(a) \cdot f(b) \neq 0 \).

- Intermediate Value Theorem: Since we have opposite signs for \( f \) at \( a, b \), and \( f \) is continuous, we conclude that \( f \) has at least one root \( \bar{x} \) in the interval \([a, b] \).

- By checking the sign of \( f(\frac{a+b}{2}) \) and appropriately replacing \( a \) or \( b \) by \( \frac{a+b}{2} \), this interval is halved at each step of the iteration:

  \[
  \begin{align*}
  \text{if } \text{sign}(f(\frac{a+b}{2})) &= 0 \quad \text{then } \bar{x} := \frac{a+b}{2}, \text{ stop;} \\
  &= \text{sign}(f(a)) \quad \text{then } a := \frac{a+b}{2}; \\
  &= \text{sign}(f(b)) \quad \text{then } b := \frac{a+b}{2}.
  \end{align*}
  \]

- Since bisection traps a root, it is guaranteed to converge. However, it needs at least three iterations to achieve one additional significant digit of the root!
Consider a continuous function $f : \mathbb{R} \to \mathbb{R}$, and assume that for $a, b \in \mathbb{R}$ you know $\text{sign}(f(a)) = -\text{sign}(f(b))$, with $a < b$ and $f(a) \cdot f(b) \neq 0$.

Intermediate Value Theorem: Since we have opposite signs for $f$ at $a, b$, and $f$ is continuous, we conclude that $f$ has at least one root $x$ in the interval $[a, b]$.

By checking the sign of $f\left(\frac{a+b}{2}\right)$ and appropriately replacing $a$ or $b$ by $\frac{a+b}{2}$, this interval is halved at each step of the iteration:

$$\begin{align*}
\text{if } \text{sign}(f\left(\frac{a+b}{2}\right)) &= 0 & \text{then } x &= \frac{a+b}{2}, \text{ stop;}
\text{if } \text{sign}(f\left(\frac{a+b}{2}\right)) &= \text{sign}(f(a)) & \text{then } a &= \frac{a+b}{2};
\text{if } \text{sign}(f\left(\frac{a+b}{2}\right)) &= \text{sign}(f(b)) & \text{then } b &= \frac{a+b}{2}.
\end{align*}$$

Since bisection traps a root, it is guaranteed to converge. However, it needs at least three iterations to achieve one additional significant digit of the root!

Caveat: Although several roots might exist within the interval $[a, b]$, only one root will be found.

Caveat: Root-bracketing is not feasible for finding even-multiplicity roots.
Regula Falsi

- Aka “false position method” in some English literature.
- Rather than blindly testing $c := \frac{a+b}{2}$, one could also compute the $x$-intercept of the secant through $(a, f(a))$ and $(b, f(b))$:
  \[
  c := b - \frac{f(b)(b - a)}{f(b) - f(a)}.
  \]
- Now evaluate $\text{sign}(f(c))$, and keep either $a$ or $b$, just as with bisection.
Regula Falsi

- Aka “false position method” in some English literature.
- Rather than blindly testing $c := \frac{a+b}{2}$, one could also compute the $x$-intercept of the secant through $(a, f(a))$ and $(b, f(b))$:

\[ c := b - \frac{f(b)(b-a)}{f(b) - f(a)}. \]

- Now evaluate $\text{sign}(f(c))$, and keep either $a$ or $b$, just as with bisection.
- The regula falsi method shares with bisection the advantage of trapping a root and, thus, of always converging.
- However, it tends to converge faster than the bisection method if $a$ and $b$ are close together.
- This basic scheme can be improved further to achieve super-linear convergence; e.g., Brent-Dekker method or Illinois method.
Newton-Raphson Method

- Suppose that \( f \) and \( f' \) are continuous near a root \( \bar{x} \) of \( f \), and that \( x_0 \) is close to \( \bar{x} \).
- The Newton-Raphson method is based on the approximation of a function \( f \) by the straight-line tangent at \( x = x_k \):

\[
y = f(x_k) + f'(x_k)(x - x_k).
\]

An estimate \( x_{k+1} \) for the root is obtained by setting \( y := 0 \) and solving for \( x \):

\[
x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}.
\]
Newton-Raphson Method

- Suppose that \( f \) and \( f' \) are continuous near a root \( \bar{x} \) of \( f \), and that \( x_0 \) is close to \( \bar{x} \).

- The Newton-Raphson method is based on the approximation of a function \( f \) by the straight-line tangent at \( x = x_k \):

\[
y = f(x_k) + f'(x_k)(x - x_k).
\]

An estimate \( x_{k+1} \) for the root is obtained by setting \( y := 0 \) and solving for \( x \):

\[
x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}.
\]

- If \( f'(x) \) is non-zero and \( x_0 \) sufficiently close to the actual root \( \bar{x} \) then the Newton-Raphson method exhibits a quadratic convergence rate.

- That is, near a root the number of significant digits approximately doubles with each iteration.

- If the root is multiple then the rate of convergence may decrease to linear.

Caveat: The Newton-Raphson method may be unstable near a horizontal asymptote or a local minimum, and might even diverge.

Note: Global convergence is not guaranteed even for "nice" functions!
Newton-Raphson Method

- Suppose that $f$ and $f'$ are continuous near a root $\bar{x}$ of $f$, and that $x_0$ is close to $\bar{x}$.
- The Newton-Raphson method is based on the approximation of a function $f$ by the straight-line tangent at $x = x_k$:

$$y = f(x_k) + f'(x_k)(x - x_k).$$

An estimate $x_{k+1}$ for the root is obtained by setting $y := 0$ and solving for $x$:

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}.$$

- If $f'(x)$ is non-zero and $x_0$ sufficiently close to the actual root $\bar{x}$ then the Newton-Raphson method exhibits a quadratic convergence rate.
- That is, near a root the number of significant digits approximately doubles with each iteration.
- If the root is multiple then the rate of convergence may decrease to linear.
- Caveat: The Newton-Raphson method may be unstable near a horizontal asymptote or a local minimum, and might even diverge.
- Note: Global convergence is not guaranteed even for “nice” functions!
Secant Method

- If the derivative $f'(x_k)$ is too difficult to compute then the tangent may be replaced by the secant through two points $(x_{k-1}, f(x_{k-1}))$ and $(x_k, f(x_k))$:

  $$x_{k+1} = x_k - \frac{f(x_k) (x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}.$$ 

- This yields a simplification of the Newton-Raphson method which is known as Secant method.
Secant Method

- If the derivative $f'(x_k)$ is too difficult to compute then the tangent may be replaced by the secant through two points $(x_{k-1}, f(x_{k-1}))$ and $(x_k, f(x_k))$:

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}.$$  

- This yields a simplification of the Newton-Raphson method which is known as Secant method.

- The rate of convergence is super-linear, and, thus, slower than for the Newton-Raphson method.

- Note that two initial values $x_0, x_1$ are needed.
Floating-Point Arithmetic and Numerical Mathematics

- Floating-Point Computations
- Iterative Algorithms for Solving Non-Linear Equations
- Iterative Algorithms for Solving Linear Equations
  - Avoiding Gaussian Elimination
  - Jacobi Iteration
  - Gauss-Seidel Iteration
- Numerical Integration
Iterative Algorithms for Solving Linear Equations

- Recall that finding the exact solution $x$ of the system of linear equations $Ax = b$ requires $O(n^3)$ time for an $n \times n$ matrix $A$.

- A direct (and exact) solution turns out to be a waste of time if $n$ goes into the thousands or millions and if $A$ is sparse. In that case, iterative methods may be much faster than direct methods.
Iterative Algorithms for Solving Linear Equations

- Recall that finding the exact solution $x$ of the system of linear equations $Ax = b$ requires $O(n^3)$ time for an $n \times n$ matrix $A$.
- A direct (and exact) solution turns out to be a waste of time if $n$ goes into the thousands or millions and if $A$ is sparse. In that case, iterative methods may be much faster than direct methods.
- Suppose that we know the exact solution: $x$.
- If we write $x$ as $x = x' + \Delta x$ then we get
  \[ A\Delta x = Ax - A x' = b - A x'. \]
- Interpreting this equation as basis for an iterative formula $x^{(k+1)} = x^{(k)} + \Delta x$ yields
  \[ A(x^{(k+1)} - x^{(k)}) = b - Ax^{(k)}. \]
Iterative Algorithms for Solving Linear Equations

- Recall that finding the exact solution $x$ of the system of linear equations $Ax = b$ requires $O(n^3)$ time for an $n \times n$ matrix $A$.
- A direct (and exact) solution turns out to be a waste of time if $n$ goes into the thousands or millions and if $A$ is sparse. In that case, iterative methods may be much faster than direct methods.
- Suppose that we know the exact solution: $x$.
- If we write $x$ as $x = x' + \Delta x$ then we get
  $$A\Delta x = Ax - Ax' = b - Ax'.$$
- Interpreting this equation as basis for an iterative formula $x^{(k+1)} = x^{(k)} + \Delta x$ yields
  $$A(x^{(k+1)} - x^{(k)}) = b - Ax^{(k)}.$$
- So far, we would have gained little, as we would still have to solve for $x^{(k+1)}$ . . .
- Bold idea: Replace $A$ on the left-hand side of this equation by an easily invertible matrix $B$ that is “close to” $A$. 
Iterative Algorithms for Solving Linear Equations

- We get

\[ B(x^{(k+1)} - x^{(k)}) = b - Ax^{(k)}, \]

or

\[ Bx^{(k+1)} = b - (A - B)x^{(k)}. \]

- One can formulate conditions under which the solution obtained by this iterative scheme is guaranteed to converge to the exact solution of \( Ax = b. \)
Iterative Algorithms for Solving Linear Equations

- We get
  \[ B(x^{(k+1)} - x^{(k)}) = b - Ax^{(k)}, \]
  or
  \[ Bx^{(k+1)} = b - (A - B)x^{(k)}. \]

- One can formulate conditions under which the solution obtained by this iterative scheme is guaranteed to converge to the exact solution of \( Ax = b. \)

- Typical application in graphics: Iterative solution of a radiosity equation.
Jacobi Iteration

- Assume that all diagonal elements of $A$ are non-zero, and let $B$ be the diagonal matrix that contains all diagonal elements of $A$.
- Applying the iteration
  
  $$B x^{(k+1)} = b - (A - B) x^{(k)}.$$  

  is equivalent to

  $$a_{ii} x_i^{(k+1)} = b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)}$$  

  and, thus,

  $$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)} \right).$$
Jacobi Iteration

- Assume that all diagonal elements of $\mathbf{A}$ are non-zero, and let $\mathbf{B}$ be the diagonal matrix that contains all diagonal elements of $\mathbf{A}$.

- Applying the iteration

$$\mathbf{B}x^{(k+1)} = b - (\mathbf{A} - \mathbf{B})x^{(k)}.$$ 

is equivalent to

$$a_{ii} x_i^{(k+1)} = b_i - \sum_{j=1, j\neq i}^{n} a_{ij} x_j^{(k)} \quad \text{and, thus,} \quad x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1, j\neq i}^{n} a_{ij} x_j^{(k)} \right).$$

- If

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|,$$

i.e., if $\mathbf{A}$ is strictly diagonally dominant then this so-called Jacobi iteration is guaranteed to converge. (Different and less stringent conditions do also suffice.)
Gauss-Seidel Iteration

- Gauss-Seidel iteration is a modification of Jacobi iteration that can converge faster in some cases.
- Basic idea: Use the most up-to-date information available.

\[ a_{ii} x_{(k+1)}^i = b_i - \sum_{j=1}^{i-1} a_{ij} x_{(k+1)}^j - \sum_{j=i+1}^{n} a_{ij} x_{(k)}^j. \]
Gauss-Seidel Iteration

- Gauss-Seidel iteration is a modification of Jacobi iteration that can converge faster in some cases.
- Basic idea: Use the most up-to-date information available.
- If \( x_1^{(k+1)}, x_2^{(k+1)}, \ldots, x_{i-1}^{(k+1)} \) are already known, then these new values can be used for the computation of \( x_i^{(k+1)} \):

\[
a_{ii} x_i^{(k+1)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}.\]

Again, convergence is guaranteed if \( A \) is strictly diagonally dominant. Tends to converge faster than Jacobi iteration, but is significantly more difficult to parallelize.
Gauss-Seidel Iteration

- Gauss-Seidel iteration is a modification of Jacobi iteration that can converge faster in some cases.
- Basic idea: Use the most up-to-date information available.
- If \( x_1^{(k+1)}, x_2^{(k+1)}, \ldots, x_i^{(k+1)} \) are already known, then these new values can be used for the computation of \( x_i^{(k+1)} \):

\[
a_{ii} x_i^{(k+1)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}.
\]

- Again, convergence is guaranteed if \( A \) is strictly diagonally dominant.
- Tends to converge faster than Jacobi iteration, but is significantly more difficult to parallelize.
Floating-Point Arithmetic and Numerical Mathematics

- Floating-Point Computations
- Iterative Algorithms for Solving Non-Linear Equations
- Iterative Algorithms for Solving Linear Equations

- Numerical Integration
  - Integration Rules
  - Multi-dimensional Integration and Monte-Carlo Integration
Numerical Integration

- Suppose we want to compute an integral

\[ I = \int_a^b f(x) \, dx. \]
Numerical Integration

• Suppose we want to compute an integral
  \[ I = \int_{a}^{b} f(x) \, dx. \]

• The best way to compute this integral would be to solve it analytically, and get
  \[ I = \int_{a}^{b} f(x) \, dx = F(b) - F(a), \quad \text{with } F'(x) = f(x). \]
Numerical Integration

- Suppose we want to compute an integral

  \[ I = \int_{a}^{b} f(x) \, dx. \]

- The best way to compute this integral would be to solve it analytically, and get

  \[ I = \int_{a}^{b} f(x) \, dx = F(b) - F(a), \quad \text{with } F'(x) = f(x). \]

- However, there are many functions that cannot be integrated analytically. Thus, methods for approximating the integral through *quadrature rules* of the form

  \[ \hat{I} = \sum_{i=1}^{n} \omega_i f(x_i) \]

  have been devised, which is essentially a weighted sum of samples of the function \( f \) at various points \( x_i \) using weights \( \omega_i \).

- The many different quadrature rules can be distinguished by their sampling patterns and weights.
Midpoint Rule for Numerical Integration

- We divide the interval \([a, b]\) into a fixed number \(n\) of subintervals, each of size \(h = (b - a)/n\).
- We then choose one sample point at the midpoint of each subinterval:

\[
\hat{I} = h \sum_{i=1}^{n} f(a + (i - \frac{1}{2})h)
\]

\[
= h \left[ f(a + \frac{h}{2}) + f(a + \frac{3h}{2}) + \cdots + f(b - \frac{h}{2}) \right].
\]

The Midpoint Rule is exact for constant or linear functions. Otherwise, its error is bounded by \(O(n^{-2})\), provided that \(f\) has at least two continuous derivatives on \([a, b]\).
Midpoint Rule for Numerical Integration

- We divide the interval \([a, b]\) into a fixed number \(n\) of subintervals, each of size \(h = (b - a)/n\).
- We then choose one sample point at the midpoint of each subinterval:

\[
\hat{I} = h \sum_{i=1}^{n} f(a + (i - \frac{1}{2})h)
\]

\[
= h \left[ f(a + \frac{h}{2}) + f(a + \frac{3h}{2}) + \cdots + f(b - \frac{h}{2}) \right].
\]

- The Midpoint Rule is exact for constant or linear functions. Otherwise, its error is bounded by \(O(n^{-2})\), provided that \(f\) has at least two continuous derivatives on \([a, b]\).
Trapezoidal Rule for Numerical Integration

- The trapezoidal rule is similar to the midpoint rule, except that we sample the function at the ends of each subinterval, and compute the area of a trapezoid for each subinterval.

\[
\hat{I} = \sum_{i=1}^{n} \frac{h}{2} [f(a + (i - 1)h) + f(a + ih)]
\]

\[
= h \left[ \frac{1}{2} f(a) + f(a + h) + f(a + 2h) + \cdots + f(b - h) + \frac{1}{2} f(b) \right].
\]
The trapezoidal rule is similar to the midpoint rule, except that we sample the function at the ends of each subinterval, and compute the area of a trapezoid for each subinterval.

\[
\hat{I} = \sum_{i=1}^{n} \frac{h}{2} [f(a + (i - 1)h) + f(a + ih)]
\]

\[
= h \left[ \frac{1}{2} f(a) + f(a + h) + f(a + 2h) + \cdots + f(b - h) + \frac{1}{2} f(b) \right].
\]

For the trapezoid rule, the error is also bounded by \( O(n^{-2}) \).
Simpson’s Rule for Numerical Integration

- Simpson’s rule is similar to the trapezoidal rule, except that we compute the area under a quadratic polynomial approximation (instead of a linear approximation for the trapezoid). The equation is:

\[
\hat{I} = h \left[ \frac{1}{3} f(a) + \frac{4}{3} f(a + h) + \frac{2}{3} f(a + 2h) + \frac{4}{3} f(a + 3h) + \frac{2}{3} f(a + 4h) + \cdots + \frac{4}{3} f(b - h) + \frac{1}{3} f(b) \right].
\]
Simpson’s Rule for Numerical Integration

- Simpson’s rule is similar to the trapezoidal rule, except that we compute the area under a quadratic polynomial approximation (instead of a linear approximation for the trapezoid). The equation is:

\[
\hat{l} = h \left[ \frac{1}{3} f(a) + \frac{4}{3} f(a + h) + \frac{2}{3} f(a + 2h) + \frac{4}{3} f(a + 3h) + \frac{2}{3} f(a + 4h) + \cdots + \frac{4}{3} f(b - h) + \frac{1}{3} f(b) \right].
\]

- Simpson’s rule is exact for polynomial functions up to cubics. The error can be bounded by the fourth derivative, i.e., \(O(n^{-4})\).

- It converges very quickly, assuming that \(f\) has a continuous fourth derivative.
Simpson’s Rule for Numerical Integration

- Simpson’s rule is similar to the trapezoidal rule, except that we compute the area under a quadratic polynomial approximation (instead of a linear approximation for the trapezoid). The equation is:

\[
\hat{I} = h \left[ \frac{1}{3} f(a) + \frac{4}{3} f(a + h) + \frac{2}{3} f(a + 2h) + \frac{4}{3} f(a + 3h) + \frac{2}{3} f(a + 4h) + \ldots + \frac{4}{3} f(b - h) + \frac{1}{3} f(b) \right].
\]

- Simpson’s rule is exact for polynomial functions up to cubics. The error can be bounded by the fourth derivative, i.e., \(O(n^{-4})\).
- It converges very quickly, assuming that \(f\) has a continuous fourth derivative.
- There are higher-order rules that can achieve even faster convergence, but require the function to be even smoother — a very rare event in computer graphics!
A common way to extend a 1D quadrature rule to higher dimensions is to use a tensor product rule. These rules have the form

\[ \hat{I} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_s=1}^{n} \omega_{i_1} \omega_{i_2} \cdots \omega_{i_s} f(x_{i_1}, x_{i_2}, \ldots, x_{i_s}), \]

where \( s \) is the dimension, and the \( \omega_{i_k} \) and \( x_{i_k} \) are weights and sample locations for a given one-dimensional quadrature rule.
Multi-Dimensional Integration

- A common way to extend a 1D quadrature rule to higher dimensions is to use a tensor product rule. These rules have the form

\[
\hat{I} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_s=1}^{n} \omega_{i_1} \omega_{i_2} \cdots \omega_{i_s} f(x_{i_1}, x_{i_2}, \ldots, x_{i_s}),
\]

where \( s \) is the dimension, and the \( \omega_{i_k} \) and \( x_{i_k} \) are weights and sample locations for a given one-dimensional quadrature rule.

- Thus, if we start with an \( n \)-point quadrature rule in 1D, we need \( N = n^d \) sample points for a \( d \)-dimensional integral.

- In terms of the total number of samples the convergence is only \( O(N^{-r/d}) \) if the 1D rule has a convergence rate of \( O(n^{-r}) \).

- If we throw in a discontinuity in \( f \), things get even worse!
Monte Carlo Integration

- The basic Monte Carlo method is

$$\int_{a}^{b} f(x) \, dx \approx \frac{b - a}{n} \sum_{i=1}^{n} f(X_i)$$

where the points $X_i$ are chosen independently and uniformly at random within the interval $[a, b]$. 

This method has a convergence rate of $O(n^{-1/2})$, regardless of the smoothness of the function $f$. Note that the convergence rate does not deteriorate in higher dimensions, and the number of samples needed does not grow astronomically. This is particularly useful in graphics, where we often need to calculate multi-dimensional integrals of discontinuous functions, for which Newton-Cotes rules do not work well. (E.g., in distributed ray tracing.)
Monte Carlo Integration

- The basic Monte Carlo method is

\[
\int_{a}^{b} f(x) \, dx \approx \frac{b - a}{n} \sum_{i=1}^{n} f(X_i)
\]

where the points \(X_i\) are chosen independently and uniformly at random within the interval \([a, b]\).

- This method has a convergence rate of \(O(n^{-1/2})\), regardless of the smoothness of the function \(f\).

- Note that the convergence rate does not deteriorate in higher dimensions, and the number of samples needed does not grow astronomically.
Monte Carlo Integration

- The basic Monte Carlo method is

\[
\int_{a}^{b} f(x) \, dx \approx \frac{b - a}{n} \sum_{i=1}^{n} f(X_i)
\]

where the points \(X_i\) are chosen independently and uniformly at random within the interval \([a, b]\).

- This method has a convergence rate of \(O(n^{-1/2})\), regardless of the smoothness of the function \(f\).

- Note that the convergence rate does not deteriorate in higher dimensions, and the number of samples needed does not grow astronomically.

- This is particularly useful in graphics, where we often need to calculate multi-dimensional integrals of discontinuous functions, for which Newton-Cotes rules do not work well. (E.g., in distributed ray tracing.)
The End!

I hope that you enjoyed this course, and I wish you all the best for your future studies.