Geometrisches Rechnen
(WS 2024/25)

Martin Held

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

January 23, 2025

UNIVERSITÄT SALZBURG
Computational Geometry and Applications Lab

Personalia

Instructor: M. Held.

Email: `held@cs.sbg.ac.at`.

Base-URL: `www.cosy.sbg.ac.at/~held`.

Office: PLUS, FB Informatik, Rm. 1.20, Jakob-Haringer Str. 2, 5020 Salzburg-Itzling.

Phone number (office): (0662) 8044-6304.

Phone number (secr.): (0662) 8044-6300.

Formalia

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

Lecture times: Thursday 7^{45}–10^{45}.

Venue: PLUS, FB Informatik, T05, Jakob-Haringer Str. 2, 5020 Salzburg-Itzling.

Note: — UV is graded according to continuous-assessment mode!

— regular attendance is compulsory!

Electronic Slides and Online Material

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

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.
- See also In Praise of Lectures by T.W. Körner.

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, July 2024 Martin Held

Legal Fine Print and Disclaimer

Recommended Textbooks

- G.E. Farin, D. Hansford.
  Practical Linear Algebra: A Geometry Toolbox.
  A K Peters/CRC Press, 4th edition, 2021; ISBN 978-0367507848.
- M.E. Mortenson.
  Mathematics for Computer Graphics Applications.
  Industrial Press, 2nd rev. edition, 1999; ISBN 978-0831131111.
- J. Ström, K. Åström, and T. Akenine-Möller.
  immersive linear algebra.
  ISBN 978-91-637-9354-7;
  `https://immersivemath.com/ila`.

Table of Content

Basis of a Vector Space

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

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

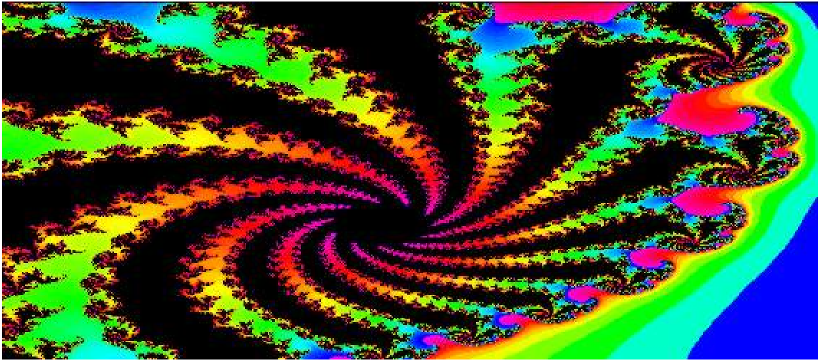
- Question: Can we write every polynomial $p(x)$ of degree at most three as

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

  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

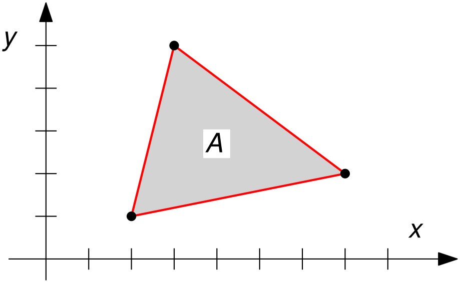
- How can we generate such an image?



- Answer: This looks like a 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 $(3,5)$.



- 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

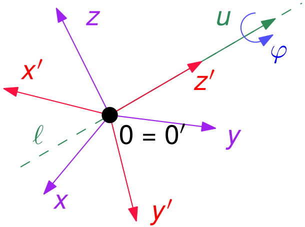
- Assume that the vector $\nu_1 := (1, 2, 3)$ is a tangent vector to the curve γ at the point $\gamma(6)$.
- Question: How can we quickly find two other vectors ν_2 and ν_3 that form an orthogonal frame with ν_1?
- Answer: An orthogonal frame can be obtained by taking a vector cross-product of two suitable vectors:

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

  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?

Rotation About a Line

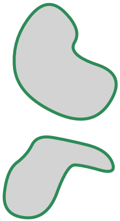
- Question: How can we compute a rotation about a line ℓ (through the origin) with direction vector u by an angle φ?



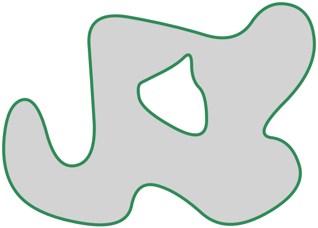
- Answer: We set up a new frame $\mathcal{C}'$ and reduce the rotation about ℓ to a rotation about a coordinate axis.

Basic Topology

- Question: What is an important topological difference between the following sets?



not path-connected



path-connected, multiply-connected

Computation with Floating-Point Arithmetic

- Consider

$$\sum_{i=1}^{n} \frac{1}{i}$$

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

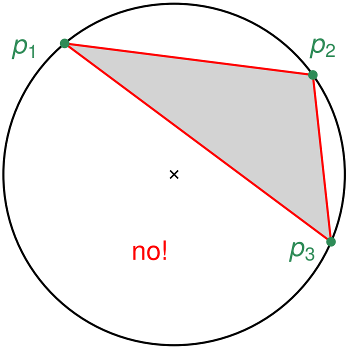
- Question: How shall be compute this sum on a computer? In particular, does it matter whether we start summing with the smallest or the largest summand?

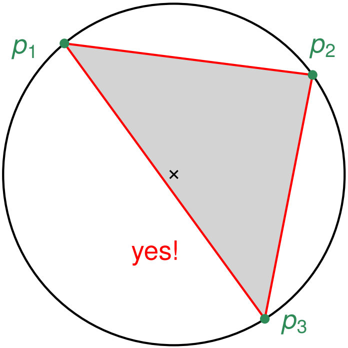
$$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1} + \frac{1}{n} \stackrel{?}{=} \frac{1}{n} + \frac{1}{n-1} + \ldots + \frac{1}{3} + \frac{1}{2} + 1$$

- Answer: Yes, it does matter! We'll get back to this question when we'll talk about floating-point arithmetic and numerical issues.

Applied Linear Algebra for Solving a Putnam Problem

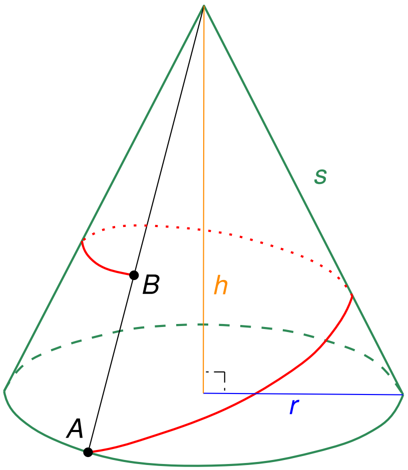
- Choose four points p_1, p_2, p_3, p_4 independently at random (relative to a uniform distribution) on a sphere (in 3D).
- Consider the tetrahedron T formed by p_1, p_2, p_3, p_4.
- What is the probability that the center of the sphere lies inside T?
- Answer: The probability is $^1/_4$ in 2D and $^1/_8$ in 3D.
- Visualization of that problem in 2D (for three random points on a circle):





Gain a Better Understanding of Geometry and the Underlying Math

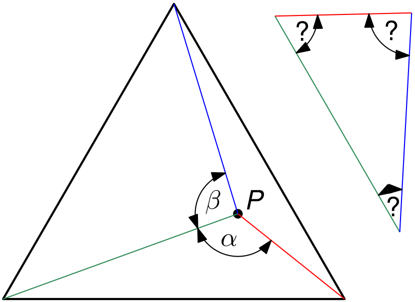
- Consider a mountain that is shaped like a (perfect) right circular cone.
- A shortest-length railroad track is supposed to start at A, wind around the mountain once, and end in B.
- The height h of the cone is $40\sqrt{2}$, its base radius r is 20, and the distance between A and B is 10.
- Your task:
  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.
- 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 α and β, and assume that they are known. What are the new triangle's three interior angles in terms of α and β?



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

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 λ, μ and letters in italics to denote scalar values: s, t.
- Points are denoted by capital or lower-case letters written in italics: e.g., A and P or a and p.
- 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., $a = (a_x, a_y, a_z)$ for $a \in \mathbb{R}^3$, or $a = (a_1, a_2, \ldots, a_n)$ for $a \in \mathbb{R}^n$.
- 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.

Notation

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

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

  or

  $$\{x \in S | \; P(x)\} \quad \text{or} \quad \{x | x \in S \wedge \; P(x)\}.$$

- Quantifiers: The universal quantifier is denoted by $\forall$, and $\exists$ denotes the existential quantifier.

Order-theoretic and Algebraic Concepts

Extreme Elements

Definition 1 (*Minimal element, Dt.: minimales Element*)

Let $(S, \preceq)$ be a poset and $T \subseteq S$. An element $a \in T$ is a *minimal element* of T if no $b \in T \setminus \{a\}$ exists such that $b \preceq a$.

Definition 2 (*Least element, Dt.: kleinstes Element, Minimum*)

Let $(S, \preceq)$ be a poset and $T \subseteq S$. An element $a \in T$ is a *least element* (or *minimum*) of T if $\quad \forall b \in T \setminus \{a\} \quad a \preceq b$.

Definition 3 (*Maximal element, Dt.: maximales Element*)

Let $(S, \preceq)$ be a poset and $T \subseteq S$. An element $a \in T$ is a *maximal element* of T if no $b \in T \setminus \{a\}$ exists such that $a \preceq b$.

Definition 4 (*Greatest element, Dt.: größtes Element, Maximum*)

Let $(S, \preceq)$ be a poset and $T \subseteq S$. An element $a \in T$ is a *greatest element* (or *maximum*) of T if $\quad \forall b \in T \setminus \{a\} \quad b \preceq a$.

Infimum

Definition 5 (*Lower bound, Dt.: untere Schranke*)

Let $(S, \preceq)$ be a partially ordered set and let $T \subseteq S$. The set T is *bounded below* if there exists an element $s \in S$, a *lower bound* of T, such that

$$\forall t \in T \quad s \preceq t.$$

Definition 6 (*Greatest lower bound, infimum, Dt.: Infimum, größte untere Schranke*)

Let $(S, \preceq)$ be a partially ordered set and let $T \subseteq S$. An element $s \in S$ is called *greatest lower bound* (or *infimum* of T), and denoted by $\inf(T)$, if

$$\forall t \in T \quad s \preceq t \quad \text{and} \quad \forall s' \in S \quad \big((\forall t \in T \;\; s' \preceq t) \;\Rightarrow\; s' \preceq s\big).$$

Mind the difference!

The terms “minimum” and “infimum” have different meanings!

Infimum and Supremum

Lemma 7

Let $(S, \preceq)$ be a partially ordered set and let $T \subseteq S$.
(1) If the infimum of T exists then it is unique.
(2) If the infimum of T belongs to T then it is also the minimum of T.

- The definitions of *upper bound* and *supremum* are obtained by replacing terms like “lower” by “upper” in these definitions.
- Similar to minimal and minimum, note the difference between *optimal* and *optimum*.

Vector Space

Definition 8 (*Vector space, Dt.: Vektorraum*)

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

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

- In the sequel we use the same symbols $+$ and $\cdot$ for both types of operations.
- Furthermore, we postulate the standard precedence rules.
- The multiplication sign is often dropped if the meaning is clear within a specific context: λa rather than $\lambda \odot a$.

Vector Space F^n

Definition 9 (*Cartesian product, Dt.: Mengenprodukt, kartesisches Produkt*)

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

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

- Named after the latinized version of the name of René Descartes (1596–1650).
- 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 $F_1 \times F_2 \times \cdots \times F_n$ for n (possibly different) fields $F_1, \ldots, F_n$.

Vector Space F^n

Definition 10

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:

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

Theorem 11

Let F be a field. Then F^n with addition and scalar multiplication as defined above constitutes a vector space over F for every $n \in \mathbb{N}$.

“Exotic” Vector Spaces: Functions, Sequences

Lemma 12

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

Lemma 13

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!

Subspace

Definition 14 (*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. $\forall a, b \in S \qquad a + b \in S$ (S is said to be closed under vector addition);
3. $\forall a \in S \;\; \forall \lambda \in F \qquad \lambda a \in S$ (S is said to be closed under scalar multiplication).

Lemma 15

The set of all continous (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.

Linear Combination

Definition 16 (*Linear combination, Dt.: Linearkombination*)

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

$$\nu := \lambda_1\nu_1 + \lambda_2\nu_2 + \dots + \lambda_k\nu_k$$

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

Definition 17 (*Linear hull, Dt.: lineare Hülle*)

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

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

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 18

For $S \subseteq V$, with $S \neq \emptyset$, the linear hull $[S]$ forms a subspace of the vector space V.

Linear Independence

Definition 19 (*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 20

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 21

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.

Basis of a Vector Space

Definition 22 (*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 23 (*Finite dimension*)

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

Theorem 24

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

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

Theorem 25

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

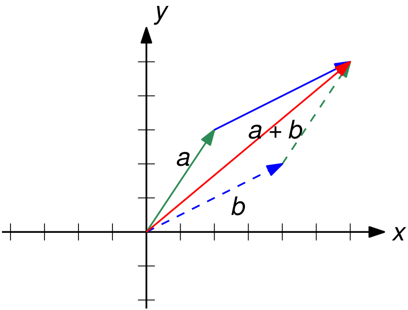
$\mathbb{R}^n$: Points and Vectors

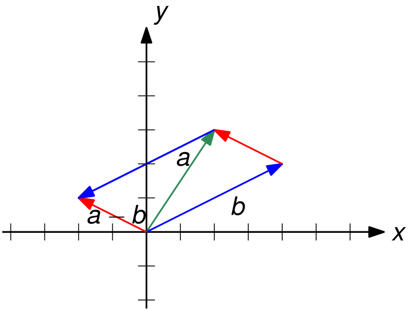
- A *point* is a location in a (vector) space. From a mathematical point of view it does not have any size or any other property besides its location.
- A *vector* has a direction and a length as its main properties.
- The *position vector* (Dt.: Ortsvektor) of a point is the vector that points from 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.

Vector Algebra

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

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





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

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

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

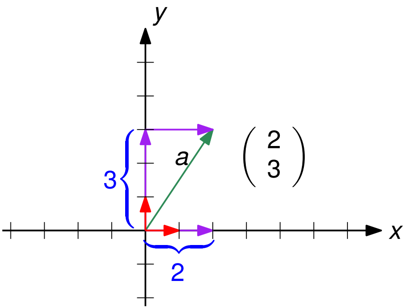
- Let $a \in \mathbb{R}^n$. We get

$$a := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot e_n.$$

Canonical Basis

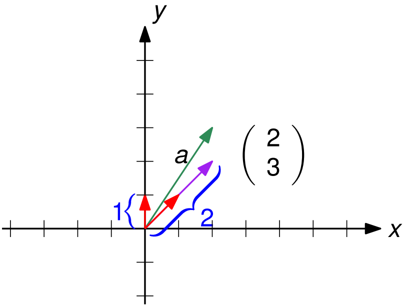
- For $a \in \mathbb{R}^2$ we get $a = a_1 \cdot e_1 + a_2 \cdot e_2$. E.g.:

$$\begin{pmatrix}2\\3\end{pmatrix} = 2e_1 + 3e_2$$
$$= 2\begin{pmatrix}1\\0\end{pmatrix} + 3\begin{pmatrix}0\\1\end{pmatrix}$$



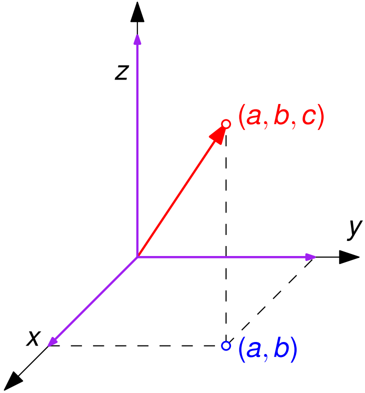
- But this is not the only possible basis for $\mathbb{R}^2$. E.g.:

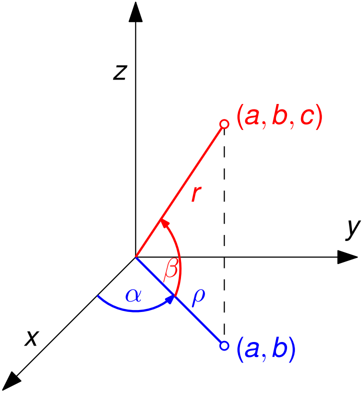
$$\begin{pmatrix}2\\3\end{pmatrix}_{[e_1,e_2]} = 2v + w$$
$$= \begin{pmatrix}2\\1\end{pmatrix}_{[v,w]}$$



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

- Cartesian coordinates: (a, b, c).
- Polar coordinates (in $\mathbb{R}^2$): (ρ, α), with $\alpha \in [0, 2\pi[$.
- Cylindrical coordinates: (ρ, α, c), with $\alpha \in [0, 2\pi[$.
- Spherical coordinates: (r, α, β), 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.
- 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 α degrees East or West, and β 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 halves of great circles that intersect at the poles; they are called *meridians*.
- Hence, geographical coordinates are nothing but (a variant of) a spherical coordinate system.

Affine and Convex Combinations

Definition 26 (*Affine combination, Dt.: Affinkombination*)

Let $p_1, p_2, \ldots, p_k$ be k points in $\mathbb{R}^n$. An *affine combination* of $p_1, \ldots, p_k$ 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 27 (*Convex combination, Dt.: Konvexkombination*)

Let $p_1, p_2, \ldots, p_k$ be k points in $\mathbb{R}^n$. A *convex combination* of $p_1, \ldots, p_k$ is given by

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

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

Affine Hull

Definition 28 (*Affine hull, Dt.: affine Hülle*)

Let $p_1, p_2, \ldots, p_k$ be k points in $\mathbb{R}^n$. The *affine hull* of $p_1, \ldots, p_k$ is the set

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

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

$$\{\sum_{i=1}^{k} \lambda_i p_i : \ k \in \mathbb{N} \text{ and } p_1, p_2, \ldots, p_k \in S \text{ and } \lambda_1, \ldots \lambda_k \in \mathbb{R} \text{ and } \sum_{i=1}^{k} \lambda_i \ = \ 1\}.$$

Convex Hull

Definition 29 (*Convex hull, Dt.: konvexe Hülle*)

Let $p_1, p_2, \dots, p_k$ be k points in $\mathbb{R}^n$. The *convex hull* of $p_1, \dots, p_k$ is the set

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

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

$$\{\sum_{i=1}^{k} \lambda_i p_i : \ k \in \mathbb{N} \text{ and } p_1, p_2, \dots, p_k \in S \text{ and } \lambda_1, \dots \lambda_k \in \mathbb{R}_0^+ \text{ and } \sum_{i=1}^{k} \lambda_i = 1\}.$$

The convex hull of S is commonly denoted by $CH(S)$.

Convexity

Definition 30 (*Convex set, Dt.: konvexe Menge*)

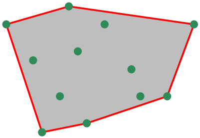
A set $S \subseteq \mathbb{R}^n$ is called *convex* if for all $p, q \in S$

$$\overline{pq} \subseteq S$$

where $\overline{pq}$ denotes the straight-line segment between p and q.

Lemma 31

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



Convexity

Definition 32 (*Convex superset*)

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

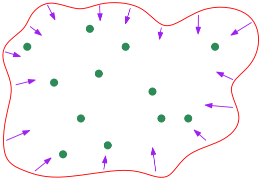
$$A \subseteq B \quad \text{and} \quad B \text{ is convex.}$$

Lemma 33

For $A \subseteq \mathbb{R}^n$, the following definitions are equivalent to Def. 29:

- $CH(A)$ 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 $\mathbb{R}$.



Complex Numbers

Definition 34 (*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:

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

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

Lemma 35

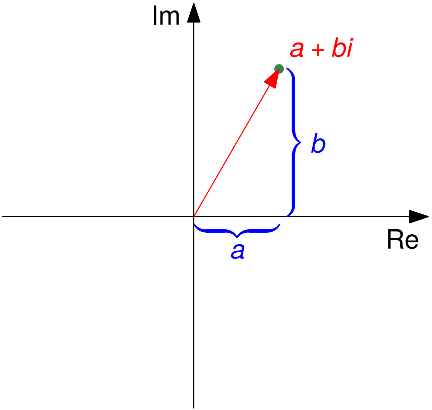
Commutativity, associativity and distributivity hold for $(\mathbb{C}, +, \cdot)$.

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

  $(2 + 3i) \cdot (1 - 2i) = 2 \cdot 1 + (3 \cdot 1)i - (2 \cdot 2)i - (3 \cdot 2)i^2 = 8 - i$

Complex Numbers and Complex Plane

- The *complex plane*, aka *Gauss 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 36 (*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 37 (*Complex conjugate, Dt.: konjugiert-komplexe Zahl*)

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

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

Definition 38 (*Multiplicative inverse*)

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

$$z^{-1} := \overline{z}|z|^{-2} = \frac{\overline{z}}{|z|^2}.$$

Complex Numbers

Lemma 39

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

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \qquad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2} \qquad \overline{\overline{z_1}} = z_1$$

$$|z_1| = |\overline{z_1}| \qquad z_1 \cdot z_1^{-1} = 1 \qquad |z_1|^2 = z_1 \cdot \overline{z_1}$$

Theorem 40

The complex numbers $(\mathbb{C}, +, \cdot)$ form a field.

Complex Numbers and de Moivre's Formula

- A complex number $z := a + bi$, for $a, b \in \mathbb{R}$, can also be written as

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

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

- By applying standard trigonometric identities, we get

$$z_1 \cdot z_2 = r_1 r_2[\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)],$$

$$z_1/z_2 = r_1/r_2[\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)].$$

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

Lemma 41 (*de Moivre*)

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

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

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

Complex Numbers and Euler's Formula

Theorem 42 (*Euler*)

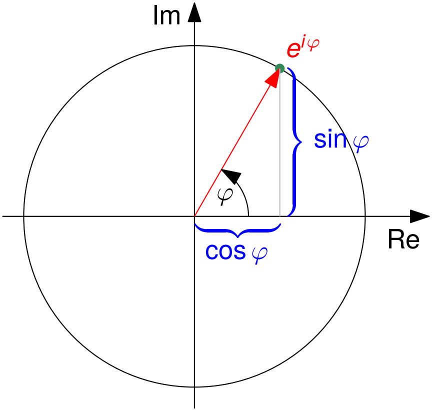
For any $\varphi \in \mathbb{R}$,

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

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

Corollary 43

$e^{i\pi} = -1.$



Complex Numbers and Euler's Formula

Sketch of Proof of Theorem 42 : 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

$$\begin{aligned} e^{ix} &= \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{k=0}^{\infty} \frac{i^k x^k}{k!} = 1 + ix + \frac{i^2 x^2}{2!} + \frac{i^3 x^3}{3!} + \frac{i^4 x^4}{4!} + \frac{i^5 x^5}{5!} + \frac{i^6 x^6}{6!} + \ldots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \ldots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots\right) \\ &= \cos x + i \sin x. \end{aligned}$$

$\square$

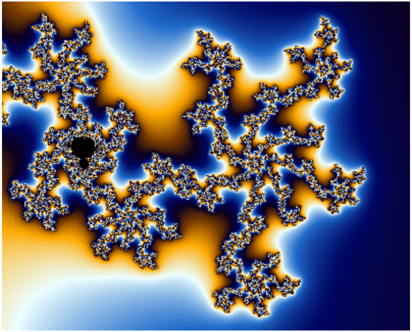
Mandelbrot Set

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

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

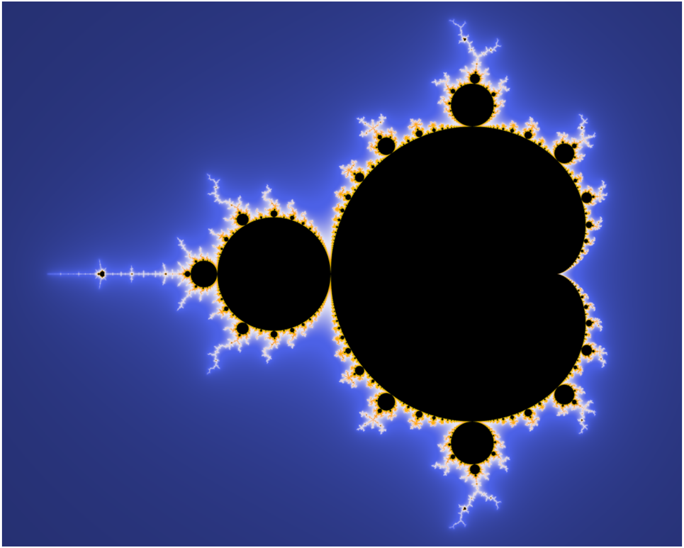
does not diverge.

- 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



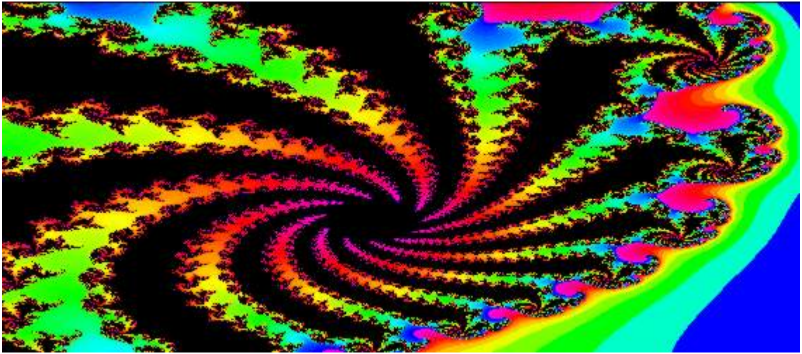
[Image credit: `https://commons.wikimedia.org/wiki/File:Mandelbrot_set_2500px.png`]

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

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

does not diverge.



Polynomials

Definition 44 (*Monomial, Dt.: Monom*)

For $m \in \mathbb{N}$, 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 45 (*Polynomial, Dt.: Polynom*)

For $m \in \mathbb{N}$, 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 46 (*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.
- 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:

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

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

$$ax = xa \quad \text{and} \quad x^m \cdot x^k = x^{m+k}$$

  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 $x, y, \dots$ that are not contained in R would do. E.g.,

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

Lemma 47

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 $R[x, y] := (R[x])[y]$.

Definition 48

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

Polynomials: Vector Space

Theorem 49

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 50 (*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 51

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 $n \in \mathbb{N}_0$.

Polynomials: Roots

Definition 52 (*Polynomial equation*)

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

Definition 53 (*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 54 (*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 55 (*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 $ax^2 + bx + c$.
- Similar (albeit more complex) formulas exist for cubic and quartic polynomials.

Theorem 56 (*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

Lemma 57

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

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

Lemma 58

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

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

- These two lemmas are special cases of a general theorem by François Viète (Franciscus Vieta, 1540–1603).

Polynomials: Function

Definition 59 (*Polynomial function; Dt.: Polynomfunktion*)

A (univariate real) function $f\colon I \to \mathbb{R}$, 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

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

- 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 branches 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 $x \in [-1, 1]$.

Polynomial Evaluation: Horner's Algorithm

- Consider a polynomial $p \in \mathbb{R}[x]$ 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).
- Can we do better?
- Obviously, we can reduce the number of multiplications to $O(n \log n)$ by resorting to exponentiation by squaring:

$$x^n = \begin{cases} x\,(x^2)^{\frac{n-1}{2}} & \text{if } n \text{ is odd,} \\ (x^2)^{\frac{n}{2}} & \text{if } n \text{ is even.} \end{cases}$$

- Can we do even better?

Polynomial Evaluation: Horner's Algorithm

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

$$p(x) = a_0 + x\Big(a_1 + x\big(a_2 + \ldots + x(a_{n-1} + x\, a_n)\ldots\big)\Big)$$

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

$$h_n := a_n$$
$$h_i := x_0 \cdot h_{i+1} + a_i \qquad\qquad \text{for } i = 0, 1, 2, ..., n-1$$

Lemma 60

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

Definition 61 (*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 62 (*Size*)

The numbers m and n in Def. 61 describe the *size* of the matrix $\mathbf{A}$. The matrix $\mathbf{A}$ is *square* if $m = n$.

Matrices

Definition 63 (*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 $\mathbf{0}$.

Definition 64 (*Identity matrix, Dt.: Einheitsmatrix*)

For $n \in \mathbb{N}$, the $n \times n$ matrix $\mathbf{I} := [\delta_{ij}]$, 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 4×4 matrices we have

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Matrices

Definition 65 (*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, $\mathbf{A}, \mathbf{B} \in M_{m \times n}(F)$ 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 66 (*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**.

- 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)!

Matrix Algebra

Definition 67 (*Matrix addition*)

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

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

Definition 68 (*Scalar multiplication*)

Consider a matrix $\mathbf{A} \in M_{m\times n}(F)$ and $\lambda \in F$. (Thus, λ is a scalar.) Then $\lambda\mathbf{A}$ is the matrix obtained by multiplying all elements of $\mathbf{A}$ by λ; that is,

$$\lambda\mathbf{A} = \lambda[a_{ij}] := \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \cdots & \lambda a_{2n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

Matrix Algebra

Definition 69 (*Additive inverse*)

Consider a matrix $\mathbf{A} \in M_{m\times n}(F)$. Then

$$-\mathbf{A} = [-a_{ij}] := \begin{pmatrix} -a_{11} & \cdots & -a_{1n} \\ -a_{21} & \cdots & -a_{2n} \\ \vdots & \ddots & \vdots \\ -a_{m1} & \cdots & -a_{mn} \end{pmatrix}$$

is taken as the additive inverse of $\mathbf{A}$.

Theorem 70

$M_{m\times n}(F)$, with addition and scalar multiplication as defined in Defs. 67+68, forms a vector space over F for all $m, n \in \mathbb{N}$.

Matrix Algebra

Lemma 71

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 λ, μ 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: $\lambda(\mathbf{A}+\mathbf{B}) = \lambda\mathbf{A}+\lambda\mathbf{B}$;
7. Associativity: $\lambda(\mu\mathbf{A}) = (\lambda\mu)\mathbf{A}$;
8. $1\mathbf{A} = \mathbf{A}$;
9. $0\mathbf{A} = \mathbf{0}$;
10. $(-1)\mathbf{A} = -\mathbf{A}$;
11. $\lambda\mathbf{A} = \mathbf{0} \quad \Rightarrow \quad \lambda = 0$ or $\mathbf{A} = \mathbf{0}$.

Matrix Algebra

Definition 72 (*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 $\mathbf{A} \cdot \mathbf{B}$, or **AB** for sake of brevity, is the $m \times p$ matrix $\mathbf{C} = [c_{ik}]$ 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 73

Matrix multiplication obeys the standard laws of arithmetic except for commutativity:

1. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ if $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are $m \times n$, $n \times p$, $p \times q$, respectively;
2. $\lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B} = \mathbf{A}(\lambda\mathbf{B})$ if $\mathbf{A}, \mathbf{B}$ are $m \times n$, $n \times p$, respectively;
3. $\mathbf{A}(-\mathbf{B}) = (-\mathbf{A})\mathbf{B} = -(\mathbf{AB})$ if $\mathbf{A}, \mathbf{B}$ are $m \times n$, $n \times p$, respectively;
4. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ if $\mathbf{A}, \mathbf{B}$ are $m \times n$ and $\mathbf{C}$ is $n \times p$;
5. $\mathbf{D}(\mathbf{A} + \mathbf{B}) = \mathbf{DA} + \mathbf{DB}$ if $\mathbf{A}, \mathbf{B}$ are $m \times n$ and $\mathbf{D}$ is $p \times m$.

- Note: $\mathbf{AB} \neq \mathbf{BA}$ even if $\mathbf{A}, \mathbf{B}$ are square. Also, $\mathbf{AB} = \mathbf{0} \not\Rightarrow [\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}]$

Inversion of a Matrix

Definition 74 (*Invertible, Dt.: invertierbar*)

An $n \times n$ matrix $\mathbf{A}$ is *invertible* (or *non-singular*) if there exists an $n \times n$ matrix $\mathbf{B}$ such that

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{I}.$$

If $\mathbf{A}$ is invertible then the inverse matrix is denoted by $\mathbf{A}^{-1}$.

Theorem 75

If $\mathbf{A}$ has inverse matrices $\mathbf{B}, \mathbf{C}$ then $\mathbf{B} = \mathbf{C}$.

- Note that $\mathbf{A}^{-1}$ can be obtained (if it exists) by solving $\mathbf{A}x_i = e_i$ for $1 \leq i \leq n$; the vectors x_i form the columns of $\mathbf{A}^{-1}$.

Theorem 76

If $\mathbf{A}, \mathbf{B}$ are invertible matrices of the same size then $\mathbf{A} \cdot \mathbf{B}$ is invertible, and

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1},$$

i.e., the inverse of the product equals the product of the inverses in the reverse order.

Transpose of a Matrix

Definition 77 (*Transpose, Dt.: transponiert*)

Consider an $m \times n$ matrix $\mathbf{A}$. The *transpose* of $\mathbf{A}$ is the matrix $\mathbf{A}^t$ obtained by interchanging the rows and columns of $\mathbf{A}$.

- Consequently, $\mathbf{A}^t$ 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 78

The transpose operation has the following properties for all matrices $\mathbf{A}, \mathbf{B}$ of suitable sizes:

1. $(\mathbf{A}^t)^t = \mathbf{A}$;
2. $(\mathbf{A} + \mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t$;
3. $(\lambda \mathbf{A})^t = \lambda \mathbf{A}^t$ for a scalar λ.
4. $(\mathbf{A} \cdot \mathbf{B})^t = \mathbf{B}^t \cdot \mathbf{A}^t$;
5. If $\mathbf{A}$ is invertible then $\mathbf{A}^t$ is also invertible and we have $(\mathbf{A}^t)^{-1} = (\mathbf{A}^{-1})^t$.

Special Matrices

Definition 79 (*Symmetric, Dt.: symmetrisch*)

A matrix $\mathbf{A}$ is called *symmetric* if $\mathbf{A}^t = \mathbf{A}$.

Definition 80 (*Diagonal matrix, Dt.: Diagonalmatrix*)

A square matrix $\mathbf{A}$ is called *diagonal* if $a_{ij} = 0$ for $i \neq j$.

Definition 81 (*Upper-triangular, Dt.: obere Dreiecksmatrix*)

A square matrix $\mathbf{A}$ is called *upper-triangular* if $a_{ij} = 0$ for $i > j$.

Definition 82 (*Orthogonal, Dt.: orthogonal*)

A square matrix $\mathbf{A}$ is called *orthogonal* if $\mathbf{A} \cdot \mathbf{A}^t = \mathbf{I} = \mathbf{A}^t \cdot \mathbf{A}$.

Lemma 83

If a square matrix $\mathbf{A}$ is *orthogonal* then $\mathbf{A}^{-1} = \mathbf{A}^t$.

Block Matrices

Definition 84 (*Block matrix*)

Let $m, n \in \mathbb{N}$ and $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D} \in M_{m\times n}(F)$. Then the $2m \times 2n$ matrix $\mathbf{X}$ with

$$x_{i,j} := \begin{cases} a_{i,j} & \text{if } 1 \le i \le m, 1 \le j \le n, \\ b_{i,j-n} & \text{if } 1 \le i \le m, n+1 \le j \le 2n, \\ c_{i-m,j} & \text{if } m+1 \le i \le 2m, 1 \le j \le n, \\ d_{i-m,j-n} & \text{if } m+1 \le i \le 2m, n+1 \le j \le 2n \end{cases}$$

is a *block matrix* with component matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$.

$$\mathbf{X} = \left(\begin{array}{ccc|ccc} a_{11} & \dots & a_{1n} & b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} & b_{m1} & \dots & b_{mn} \\ \hline c_{11} & \dots & c_{1n} & d_{11} & \dots & d_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mn} & d_{m1} & \dots & d_{mn} \end{array}\right)$$

- It is common to regard $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ as "coefficients" of $\mathbf{X}$ and write

$$\mathbf{X} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right),$$

or simply

$$\mathbf{X} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}.$$

Block Matrices

Lemma 85

For $m, n, p \in \mathbb{N}$, let $\mathbf{A}_{11}, \mathbf{A}_{12}, \mathbf{A}_{21}, \mathbf{A}_{22} \in M_{m\times n}(F)$, $\mathbf{B}_{11}, \mathbf{B}_{12}, \mathbf{B}_{21}, \mathbf{B}_{22} \in M_{n\times p}(F)$, and

$$\mathbf{A} := \left(\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right) \qquad \text{and} \qquad \mathbf{B} := \left(\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \end{array}\right).$$

Then

$$\mathbf{A}\cdot\mathbf{B} = \left(\begin{array}{c|c} \mathbf{A}_{11}\cdot\mathbf{B}_{11} + \mathbf{A}_{12}\cdot\mathbf{B}_{21} & \mathbf{A}_{11}\cdot\mathbf{B}_{12} + \mathbf{A}_{12}\cdot\mathbf{B}_{22} \\ \hline \mathbf{A}_{21}\cdot\mathbf{B}_{11} + \mathbf{A}_{22}\cdot\mathbf{B}_{21} & \mathbf{A}_{21}\cdot\mathbf{B}_{12} + \mathbf{A}_{22}\cdot\mathbf{B}_{22} \end{array}\right).$$

Block Matrices

Lemma 86

Let $n \in \mathbb{N}$ and $\mathbf{A}, \mathbf{B}, \mathbf{D} \in M_{n\times n}(F)$. Then the $2n \times 2n$ matrix $\mathbf{X}$ with

$$\mathbf{X} := \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{0} & \mathbf{D} \end{array}\right)$$

is invertible if and only if $\mathbf{A}$ and $\mathbf{D}$ are invertible. In this case we get

$$\mathbf{X}^{-1} = \left(\begin{array}{c|c} \mathbf{A}^{-1} & -\mathbf{A}^{-1} \cdot \mathbf{B} \cdot \mathbf{D}^{-1} \\ \hline \mathbf{0} & \mathbf{D}^{-1} \end{array}\right).$$

Fast Matrix Multiplication

- Standard multiplication of two $n \times n$ matrices results in $\Theta(n^3)$ many arithmetic operations.

Theorem 87 (*Strassen (1969)*)

Seven multiplications of scalars suffice to compute the multiplication of two 2×2 matrices. In general, $O(n^{\log_2 7}) \approx O(n^{2.807\cdots})$ arithmetic operations suffice for multiplying two $n \times n$ matrices.

Theorem 88 (*Coppersmith&Winograd (1990)*)

$O(n^{2.37547\cdots})$ arithmetic operations suffice for multiplying two $n \times n$ matrices.

Lemma 89 (*Williams (2011, 2012), Le Gall (2014), Alman&Williams (2021)*)

$O(n^{2.37285\cdots})$ arithmetic operations suffice for multiplying two $n \times n$ 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.
- Open problem: What is the true lower bound?

Fast Matrix Multiplication

Sketch of Proof of Theorem 87 : For $\mathbf{A}, \mathbf{B} \in M_{2\times 2}(\mathbb{R})$, we compute $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ via

$$
\begin{aligned}
p_1 &:= (a_{1,2} - a_{2,2})(b_{2,1} + b_{2,2}) \\
p_2 &:= (a_{1,1} + a_{2,2})(b_{1,1} + b_{2,2}) \\
p_3 &:= (a_{1,1} - a_{2,1})(b_{1,1} + b_{1,2}) \\
p_4 &:= (a_{1,1} + a_{1,2})b_{2,2} \\
p_5 &:= a_{1,1}(b_{1,2} - b_{2,2}) \\
p_6 &:= a_{2,2}(b_{2,1} - b_{1,1}) \\
p_7 &:= (a_{2,1} + a_{2,2})b_{1,1}
\end{aligned}
$$

and set

$$
\begin{aligned}
c_{1,1} &:= a_{1,1}b_{1,1} + a_{1,2}b_{2,1} &&= p_1 + p_2 - p_4 + p_6 \\
c_{1,2} &:= a_{1,1}b_{1,2} + a_{1,2}b_{2,2} &&= p_4 + p_5 \\
c_{2,1} &:= a_{2,1}b_{1,1} + a_{2,2}b_{2,1} &&= p_6 + p_7 \\
c_{2,2} &:= a_{2,1}b_{1,2} + a_{2,2}b_{2,2} &&= p_2 - p_3 + p_5 - p_7 .
\end{aligned}
$$

This uses seven multiplications and $O(1)$ additions/subtractions.

Use block matrices to apply this concept recursively for $n > 2$. This yields the recurrence relation $T(n) = 7 \cdot T\left(\frac{n}{2}\right) + O(n^2)$ for the time complexity T, and the bound claimed follows from the Master Theorem. $\square$

Linear Equations

Definition 90 (*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 91 (*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

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m, \end{array}$$

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 $x_1, x_2, \ldots, x_n$,

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

  can also be seen as one vector-valued equation:

$$\begin{pmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n \\ \vdots & & \vdots & & \ddots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

- With $\mathbf{A} := [a_{ij}]_{i=1,j=1}^{m,n}$, $b := (b_1, \ldots, b_m) \in \mathbb{R}^m$ and $x := (x_1, \ldots, x_n) \in \mathbb{R}^n$, this system can be written concisely as $\mathbf{A}x = b$:

$$\mathbf{A}x = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = b$$

Matrices and Linear Equations

- So, we have

$$\mathbf{A}x = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = b.$$

- 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 $\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$ is called the *augmented matrix* of the system.

Geometric Interpretation of Linear Equations

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 163.

- We regard the $m \times n$ matrix $\mathbf{A}$ as a transformation matrix and seek that vector $x \in \mathbb{R}^n$ which gets mapped to the vector $b \in \mathbb{R}^m$:

$$\mathbf{A}x = b$$

See Slide 234.

Solutions of Linear Equations

Definition 92

A system of linear equations in n unknowns is called *consistent* if it has a solution, i.e., if there exist (real) numbers $x_1, x_2, \dots, x_n$ 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 93

A homogeneous system of m linear equations in n unknowns always has a non-trivial solution if $m < n$.

Solutions of Linear Equations

Definition 94 (*Rank, Dt.: Rang*)

The (column) *rank* of a matrix $\mathbf{A}$, denoted by $\mathtt{rank}(\mathbf{A})$, is the number of linearly independent columns of $\mathbf{A}$.

Theorem 95

The system $\mathbf{A}x = b$ is consistent if and only if the rank of the coefficient matrix equals the rank of the augmented matrix.

Theorem 96

Assume that the system $\mathbf{A}x = b$ is consistent. This system has a unique solution if and only if the rank of the coefficient matrix equals the number of unknowns.

Elementary Row Operations

Lemma 97

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 98

A matrix **A** is *row-equivalent* to a matrix **B** if **B** is obtained from **A** by a sequence of elementary row operations.

Theorem 99

If **A** and **B** are row-equivalent augmented matrices of two systems of linear equations, then the two systems have the same solution sets.

Elementary Row Operations

Definition 100 (*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 $\mathbf{A}$ into a row-equivalent matrix $\mathbf{A}'$ that is in reduced row-echelon form, using elementary row operations:
  - Initially, $k := 1$.
  - If the rows $k, \dots, m$ 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 $\mathbf{A}'$ will be in reduced row-echelon form and will have r non-zero rows, with leading entries 1 in columns $c_1, \dots, c_r$, 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 $1, \dots, r$.

Gauss-Jordan Algorithm

- Thus, the Gauss-Jordan algorithm transforms an augmented matrix $\mathbf{A}$ into a matrix $\mathbf{A}'$ of the following form:

$$\left(\begin{array}{ccc|ccc|c}
1 & & 0 & a'_{1,r+1} & \cdots & a'_{1n} & b'_1 \\
 & \ddots & & \vdots & & \vdots & \vdots \\
0 & & 1 & a'_{r,r+1} & \cdots & a'_{rn} & b'_r \\
\hline
 & & & & & & b'_{r+1} \\
 & & & 0 & & & \vdots \\
 & & & & & & b'_m
\end{array}\right)$$

- 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 $\mathbf{A}$ into a matrix $\mathbf{A}'$ of the following form:

$$\left(\begin{array}{ccc|ccc|c}
1 & & 0 & a'_{1,r+1} & \cdots & a'_{1n} & b'_1 \\
 & \ddots & & \vdots & & \vdots & \vdots \\
0 & & 1 & a'_{r,r+1} & \cdots & a'_{rn} & b'_r \\
\hline
 & & & & & & b'_{r+1} \\
 & & & 0 & & & \vdots \\
 & & & & & & b'_m
\end{array}\right)$$

- If $r < n$ and $b'_{r+1} = b'_{r+2} = \ldots = b'_m = 0$, then there are infinitely many solutions:

$$\begin{aligned}
x'_1 &= b'_1 - a'_{1,r+1}x'_{r+1} - a'_{1,r+2}x'_{r+2} - \ldots - a'_{1n}x'_n, \\
&\vdots \\
x'_r &= b'_r - a'_{r,r+1}x'_{r+1} - a'_{r,r+2}x'_{r+2} - \ldots - a'_{rn}x'_n.
\end{aligned}$$

The independent unknowns $x'_{r+1}, \ldots, x'_n$ may take on arbitrary values.

Sample Linear System

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & x_2 & + & 2x_3 & + & 3x_4 & = & 4 \\ 2x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 5 \end{array}\right.$$

$$(\mathbf{A}|b) = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 & 5 \end{pmatrix} \begin{matrix} \\ + I \cdot (-2) \end{matrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix} \begin{matrix} \\ \cdot(-1) \end{matrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix} x_2 \leftrightarrow x_3 \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 4 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} \begin{matrix} + II \cdot (-2) \\ \\ \end{matrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} \rightsquigarrow \left\{\begin{array}{rcrcrcrcr} x_1 & & & + & x_2 & - & x_4 & = & -2 \\ & & x_3 & & & + & 2x_4 & = & 3 \end{array}\right.$$

$$\rightsquigarrow \left\{\begin{array}{rcrcrcr} x_1 & = & -2 & - & x_2 & + & x_4 \\ x_3 & = & 3 & & & - & 2x_4 \end{array}\right.$$

$$\rightsquigarrow \text{ Solution: } \left\{ \begin{pmatrix} -2 \\ 0 \\ 3 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix} : \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

Application: Bernstein Polynomials as Basis

Proof of Theorem 51 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:

$$\begin{pmatrix} 1 & -3 & 3 & -1 \\ 0 & 3 & -6 & 3 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} B_{0,3}(x) \\ B_{1,3}(x) \\ B_{2,3}(x) \\ B_{3,3}(x) \end{pmatrix}$$

Inversion of this matrix yields

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & \frac{1}{3} & \frac{2}{3} & 1 \\ 0 & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} B_{0,3}(x) \\ B_{1,3}(x) \\ B_{2,3}(x) \\ B_{3,3}(x) \end{pmatrix} = \begin{pmatrix} 1 \\ x \\ x^2 \\ x^3 \end{pmatrix},$$

i.e., the fact that $1, x, x^2, x^3$ of the power basis can be expressed in terms of $B_{0,3}(x), B_{1,3}(x), B_{2,3}(x), B_{3,3}(x)$. □

Determinants

Definition 101 (*Submatrix, Dt.: Untermatrix*)

Let $\mathbf{A} \in M_{n\times n}(\mathbb{R})$, with $n \geq 2$. Let $\mathbf{A}_{ij}(\mathbf{A})$, or simply $\mathbf{A}_{ij}$ if there is no ambiguity, denote the $(n-1)\times(n-1)$ *submatrix* of $\mathbf{A}$ formed by deleting the i-th row and j-th column of $\mathbf{A}$.

- Example:

$$\mathbf{A} := \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 4 & 4 \end{pmatrix} \qquad \mathbf{A}_{12} = \begin{pmatrix} 2 & 2 \\ 0 & 4 \end{pmatrix} \qquad \mathbf{A}_{33} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

Definition 102 (*Determinant*)

The *determinant*, $\det(\mathbf{A})$, of an $n \times n$ matrix $\mathbf{A} \in M_{n\times n}(\mathbb{R})$, for $n \in \mathbb{N}$, is defined recursively by the so-called *first-row Laplace expansion*:

$$\det(\mathbf{A}) := \begin{cases} a_{11} & \text{if } n = 1, \\ \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \cdot \det(\mathbf{A}_{1j}) & \text{if } n > 1. \end{cases}$$

Determinants

- Note that the term $|\mathbf{A}|$ is also commonly used for denoting the determinant of an $n \times n$ matrix $\mathbf{A}$, 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:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

- E.g.,

$$\begin{aligned} \det(\mathbf{A}) &= \textstyle\sum_{j=1}^{n} (-1)^{1+j} a_{1j} \cdot \det(\mathbf{A}_{1j}) && \ldots \text{ first row} \\ &= \textstyle\sum_{j=1}^{n} (-1)^{j} a_{2j} \cdot \det(\mathbf{A}_{2j}) && \ldots \text{ second row} \\ &= \textstyle\sum_{i=1}^{n} (-1)^{i+1} a_{i1} \cdot \det(\mathbf{A}_{i1}) && \ldots \text{ first column} \end{aligned}$$

2×2 and 3×3 Determinants

Lemma 103

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 $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33} \in \mathbb{R}$,

$$\begin{aligned}
\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\
&= a_{11} \cdot \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \cdot \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}) \\
&= a_{11}a_{22}a_{33} + a_{21}a_{13}a_{32} + a_{31}a_{12}a_{23} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} - a_{31}a_{13}a_{22}.
\end{aligned}$$

Mnemonic for Computing 3×3 Determinants (Sarrus)

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.$$

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

$$- \quad - \quad - \qquad + \quad + \quad +$$

Properties of Determinants

Lemma 104

If a row (or column) of a matrix is zero, then its determinant is zero.

Lemma 105

The determinant is a linear function of each row and each column.

Lemma 106

If a multiple of a row is added to another row, then the value of the determinant remains unchanged. Same for columns.

Lemma 107

If two rows or columns of a matrix are equal then the determinant is zero.

Lemma 108

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 109

The determinant of the product of two (square) matrices is the product of the determinants of the matrices:

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$$

for all $\mathbf{A}, \mathbf{B} \in M_{n\times n}$.

Lemma 110

A matrix and its transpose have equal determinants, i.e., for all (square) matrices $\mathbf{A}$,

$$\det(\mathbf{A}^t) = \det(\mathbf{A}).$$

Lemma 111

The determinant of an orthogonal matrix is ± 1.

Theorem 112

The (square) matrix $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.

Properties of Determinants

Lemma 113

The determinant of an upper-triangular matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & * & \cdots & \cdots & * \\ 0 & a_{22} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & * \\ 0 & \cdots & \cdots & 0 & a_{nn} \end{pmatrix}$$

is given by the product of its diagonal elements: $\det(\mathbf{A}) = \prod_{i=1}^{n} a_{ii}$.

Corollary 114

An upper-triangular matrix is invertible if and only if all its diagonal elements are non-zero.

Properties of Determinants

Lemma 115

Let $n \in \mathbb{N}$ and $\mathbf{A}, \mathbf{B}, \mathbf{D} \in M_{n \times n}(\mathbb{R})$. Then the determinant $\det(\mathbf{X})$ of the $2n \times 2n$ block matrix $\mathbf{X}$ with

$$\mathbf{X} := \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}$$

is given by

$$\det(\mathbf{X}) = \det(\mathbf{A}) \cdot \det(\mathbf{D}).$$

Corollary 116

Let $n \in \mathbb{N}$ and $\mathbf{A}, \mathbf{B}, \mathbf{D} \in M_{n \times n}(\mathbb{R})$. Then the $2n \times 2n$ block matrix $\mathbf{X}$ with

$$\mathbf{X} := \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}$$

is invertible if and only if the matrices $\mathbf{A}$ and $\mathbf{D}$ are invertible.

Calculating Determinants Manually

- Make sure to make good use of the lemmas stated on the previous slides!

$$\det\begin{pmatrix}1 & 2 & -1 & 3\\0 & 1 & 4 & 2\\0 & 1 & 0 & 4\\1 & 0 & 2 & 1\end{pmatrix} \overset{I-IV}{=} \det\begin{pmatrix}0 & 2 & -3 & 2\\0 & 1 & 4 & 2\\0 & 1 & 0 & 4\\1 & 0 & 2 & 1\end{pmatrix} \text{ Expansion by first column } =$$

$$= \quad (-1)^{1+4}\cdot 1\cdot\det\begin{pmatrix}2 & -3 & 2\\1 & 4 & 2\\1 & 0 & 4\end{pmatrix} = -\det\begin{pmatrix}0 & -3 & -6\\0 & 4 & -2\\1 & 0 & 4\end{pmatrix}$$

$$= \quad -(-1)^{1+3}\cdot 1\cdot\det\begin{pmatrix}-3 & -6\\4 & -2\end{pmatrix} = -((-3\cdot(-2)) - (-6\cdot 4)) = -30.$$

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 $O(n \cdot M(n))$ additions and multiplications for an $n \times n$ matrix, where $M(n)$ is the number of arithmetic operations consumed by multiplying two $n \times n$ matrices.
- If naïve matrix multiplication is used then we get $\Theta(n^4)$.
- No $\Theta(n^3)$ division-free determinant calculation is known.

Determinants and Linear Systems

Lemma 117

The linear system $\mathbf{A}x = b$, with $\mathbf{A} \in M_{n\times n}$, has a unique solution if and only if $\det(\mathbf{A}) \neq 0$.

Lemma 118 (*Cramer's Rule*)

If $\det(\mathbf{A}) \neq 0$, for $\mathbf{A} \in M_{n\times n}(\mathbb{R})$, then the solution of $\mathbf{A}x = b$ is given by

$$x_1 = \frac{\det(\mathbf{A}_1)}{\det(\mathbf{A})},\ x_2 = \frac{\det(\mathbf{A}_2)}{\det(\mathbf{A})},\ \ldots\ ,\ x_n = \frac{\det(\mathbf{A}_n)}{\det(\mathbf{A})},$$

where $\mathbf{A}_i$ is the matrix formed by replacing the i-th column of the coefficient matrix $\mathbf{A}$ by the right-hand side b.

Geometric Interpretation of Determinants: Orientation and Area

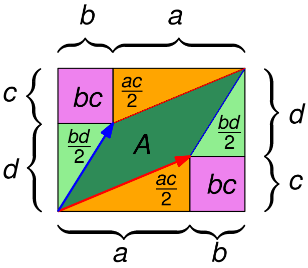
Theorem 119

Let $a, b, c, d \in \mathbb{R}$. Consider the 2D vectors

$$v_1 := \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad v_2 := \begin{pmatrix} b \\ d \end{pmatrix} \qquad \text{and let} \quad \mathbf{T} := \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then $\det(\mathbf{T})$ 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 $\det(\mathbf{T}) = ad - bc$.



Now consider the parallelogram defined by v_1 and v_2 and observe that its area A equals $ad - bc$:

$$\begin{aligned} A &= (a+b)(c+d) - ac - bd - 2bc \\ &= ad - bc. \end{aligned}$$

Interchanging v_1 and v_2 flips their handedness and changes the sign of the determinant. □

Geometric Interpretation of Determinants: Orientation

Lemma 120

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 121

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 CCW orientation. It is negative for a CW orientation, and zero if p_1, p_2 and p_3 are collinear.

Geometric Interpretation of Determinants: Area

Lemma 122

For points $p_1 := (x_1, y_1)$ and $p_2 := (x_2, y_2)$ in $\mathbb{R}^2$,

$$\frac{1}{2}\left|\det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}\right|$$

corresponds to the area of the triangle $\Delta(O, p_1, p_2)$.

Lemma 123

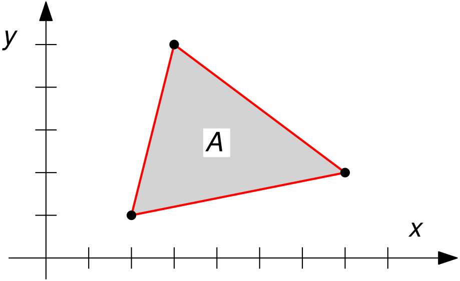
For points $p_1 := (x_1, y_1)$, $p_2 := (x_2, y_2)$ and $p_3 := (x_3, y_3)$ in $\mathbb{R}^2$,

$$\frac{1}{2}\left|\det\begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix}\right|$$

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 $(3, 5)$.



- The area of that triangle is given by

$$A = \frac{1}{2} \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 7 & 2 & 1 \\ 3 & 5 & 1 \end{pmatrix} = \frac{1}{2} \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 5 & 1 & 0 \\ 1 & 4 & 0 \end{pmatrix} = \frac{1}{2} \cdot (5 \cdot 4 - 1 \cdot 1) = \frac{19}{2}.$$

Geometric Interpretation of Determinants: Volume

Lemma 124

Let $a, b, c \in \mathbb{R}^3$. Then

$$\left| \det \begin{pmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{pmatrix} \right|$$

corresponds to the volume of the parallelepiped spanned by the three vectors a, b, c.

Geometric Interpretation of Determinants: Volume

Lemma 125

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 $\mathbb{R}^3$,

$$\frac{1}{6}\left|\det\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix}\right|$$

corresponds to the volume of the tetrahedron with corners p_1, p_2, p_3 and the origin as fourth corner.

Lemma 126

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 $\mathbb{R}^3$,

$$\frac{1}{6}\left|\det\begin{pmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{pmatrix}\right|$$

corresponds to the volume of the tetrahedron with corners p_1, p_2, p_3, p_4.

Eigenvalues and Eigenvectors

Definition 127 (*Eigenvalue, Dt.: Eigenwert*)

Consider a square $n \times n$ matrix $\mathbf{A} \in M_{n\times n}(\mathbb{R})$. A scalar $\lambda \in \mathbb{R}$ is called *eigenvalue* of $\mathbf{A}$ if a vector $v \in \mathbb{R}^n$ exists such that

$$\mathbf{A}v = \lambda v \quad \text{and} \quad v \neq 0.$$

Such a vector v is called *eigenvector* of $\mathbf{A}$.

Lemma 128

A scalar λ is an eigenvalue of matrix $\mathbf{A}$ if and only if the homogeneous linear system of equations

$$(\mathbf{A} - \lambda\mathbf{I})v = 0$$

has a non-trivial solution. This is the case if and only if $(\mathbf{A} - \lambda\mathbf{I})$ is singular, that is, if and only if

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

Eigenvalues and Eigenvectors

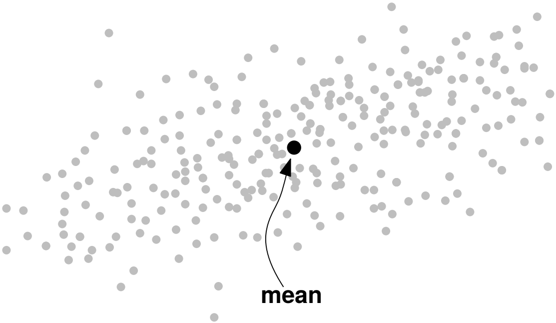
- Thus, the eigenvalues of a matrix $\mathbf{A}$ are the zeros of the *characteristic polynomial*

$$p_{\mathbf{A}}(\lambda) := \det(\mathbf{A} - \lambda \mathbf{I}).$$

- An $n \times n$ matrix can have at most n eigenvalues.
- While this approach works for any $n \times n$ matrix, it becomes tedious for $n > 4$.
- Sample application of eigenvalues and eigenvectors: Principal Components Analysis.

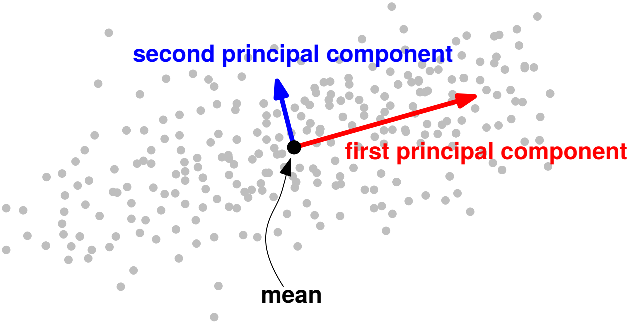
Principal Components Analysis (PCA)

- Suppose that we are given a cloud of points in $\mathbb{R}^3$. Somebody tells us that all points lie inside of an (unknown) ellipsoid. How would we rotate/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 points such that their centroid matches the origin.



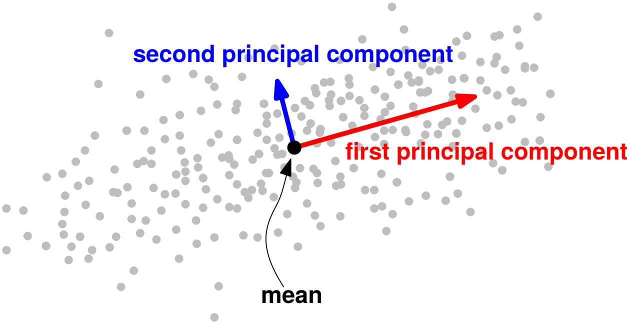
Principal Components Analysis (PCA)

- Then, PCA chooses the first PCA axis as that line which goes through the centroid of the point cloud, but also minimizes the (average) squared 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 axis.
- Similarly for the third axis.



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 $p_i := (x_i, y_i, z_i) \in \mathbb{R}^3$.
- Then the PCA axes can be computed by finding the eigenvalues and eigenvectors of the *covariance matrix* **Cov** of the coordinates of the n points:

$$\mathbf{Cov}(x, y, z) := \begin{pmatrix} \mathrm{cov}(x, x) & \mathrm{cov}(x, y) & \mathrm{cov}(x, z) \\ \mathrm{cov}(y, x) & \mathrm{cov}(y, y) & \mathrm{cov}(y, z) \\ \mathrm{cov}(z, x) & \mathrm{cov}(z, y) & \mathrm{cov}(z, z) \end{pmatrix},$$

  where

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad \bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i \quad \text{and} \quad \bar{z} := \frac{1}{n}\sum_{i=1}^{n} z_i$$

  and

$$\mathrm{cov}(x, y) := \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}.$$

  Similarly for the other entries of the covariance matrix.
- The origin of the PCA axes is given by the mean point $(\bar{x}, \bar{y}, \bar{z})$.

Dot Product

Definition 129 (*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

$$\begin{aligned} \langle,\rangle : V \times V &\rightarrow F \\ (a, b) &\mapsto \langle a,b\rangle \end{aligned}$$

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. $\langle a,a\rangle = 0 \quad \Rightarrow \quad a = 0$.

- 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 130 (*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 $\|.\|$ from V to $\mathbb{R}$:

$$\|a\| := \sqrt{\langle a, a\rangle}.$$

Definition 131 (*Unit vector, Dt.: Einheitsvektor*)

A vector a is said to be a *unit vector* if $\|a\| = 1$.

Lemma 132

We get the following standard properties of a norm for $\|.\|$ 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 133 (*Cauchy-Schwarz Inequality*)

$$\forall a, b \in V \qquad |\langle a,b\rangle| \le \|a\| \cdot \|b\|.$$

- Note that, for $a, b \neq 0$, the Cauchy-Schwarz inequality implies

$$-1 \le \frac{\langle a,b\rangle}{\|a\| \cdot \|b\|} \le 1.$$

We will make use of this fact when defining angles between vectors.

Lemma 134 (*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

$$\begin{aligned}\|a+b\|^2 &= \langle a+b,a+b\rangle = \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle \\ &= \langle a,a\rangle + \langle b,b\rangle = \|a\|^2 + \|b\|^2.\end{aligned}$$

□

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

$$\|a-b\| = \sqrt{(a_1-b_1)^2+(a_2-b_2)^2+\cdots+(a_n-b_n)^2}$$

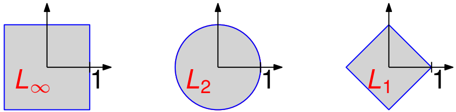
is also called L_2-norm and then denoted by $\|a-b\|_2$, in order to distinguish it from other well-known norms on $\mathbb{R}^n$, such as the L_1-norm (Manhattan metric)

$$\|a-b\|_1 := |a_1-b_1|+|a_2-b_2|+\cdots+|a_n-b_n|,$$

or the L_∞-norm (maximum norm)

$$\|a-b\|_\infty := \max_{1\le i\le n} |a_i-b_i|.$$

unit "circles"



Angle

Definition 135 (*Angle between vectors*)

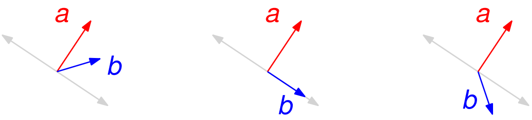
The *angle*, α, *between non-zero vectors* $a, b \in \mathbb{R}^n$ is given by

$$\cos\alpha := \frac{\langle a, b\rangle}{\|a\| \cdot \|b\|}.$$

Definition 136 (*Perpendicular, Dt.: senkrecht*)

The vectors $a, b \in \mathbb{R}^n$ are said to be *perpendicular* (or *orthogonal*), denoted by $a \perp b$, if

$$\langle a, b\rangle = 0.$$



$\langle a, b\rangle > 0$ $\langle a, b\rangle = 0$ $\langle a, b\rangle < 0$

Angle and Projection

Definition 137 (*Parallel*)

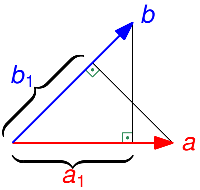
The non-zero vectors $a, b \in \mathbb{R}^n$ are said to be *parallel*, denoted by $a \parallel b$, if there exists $\lambda \in \mathbb{R}$ such that

$$a = \lambda b.$$

Lemma 138

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 242.

Orthonormal Basis of a Vector Space

Definition 139 (*Orthogonal basis*)

The vectors $a_1, \ldots, a_n$ form an *orthogonal basis* of a vector space V over $\mathbb{R}$ if

1. the vectors $a_1, \ldots, a_n$ form a basis of V;
2. $\forall (1 \leq i, j \leq n) \quad [i \neq j \quad \Rightarrow \quad \langle a_i, a_j \rangle = 0]$.

Definition 140 (*Orthonormal basis*)

The vectors $a_1, \ldots, a_n$ form an *orthonormal basis* of a vector space V over $\mathbb{R}$ 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 141

An $n \times n$ matrix $\mathbf{A} \in M_{n \times n}(\mathbb{R})$ is orthogonal if and only if its columns form an orthonormal basis of $\mathbb{R}^n$.

Vector Cross-Product in $\mathbb{R}^3$

Definition 142 (*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:

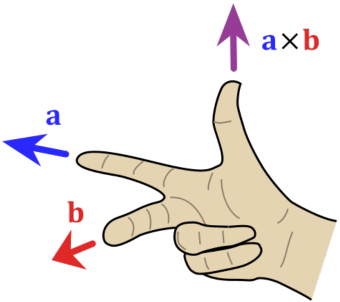
$$a \times b := \det \begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix} = a_x \cdot b_y - a_y \cdot b_x$$

- Note, however, that its properties are different from those of Definition 142.

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: `en.wikipedia.org`.]

Properties of the Cross-Product

Lemma 143

The following properties of the vector cross-product follow from the properties of 2×2 and 3×3 determinants:

1. $e_1 \times e_2 = e_3, \quad e_2 \times e_3 = e_1, \quad e_3 \times e_1 = e_2$;
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 α 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!

Lemma 144

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 119. We will later on resort to linear transformations to shed some additional light onto this claim.

Lemma 145

If u, v, w are distinct non-collinear points in $\mathbb{R}^3$, then the distance d of w from 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 a curve at the point p.
- An orthogonal frame at p can be obtained by taking a vector cross-product of two suitable vectors:

$$\nu_2 := \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

$$\nu_3 := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} 2 & 1 \\ 3 & 0 \end{vmatrix} \\ -\begin{vmatrix} 1 & -2 \\ 3 & 0 \end{vmatrix} \\ \begin{vmatrix} 1 & -2 \\ 2 & 1 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -3 \\ -6 \\ 5 \end{pmatrix}$$

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

Quaternions $\mathbb{H}$

Definition 146 (*Quaternions*)

The set of *quaternions*, $\mathbb{H}$, is given by quadrupels of real numbers together with operations $+\colon \mathbb{H} \times \mathbb{H} \to \mathbb{H}$ and $\cdot\colon \mathbb{H} \times \mathbb{H} \to \mathbb{H}$ 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 147 (*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.

Quaternions $\mathbb{H}$

Lemma 148

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, and k is defined as

$$i^2 = j^2 = k^2 := -1 \quad \text{and} \quad ijk := -1.$$

Lemma 149

Lemma 148 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.

Quaternions

Definition 150 (*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 151 (*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 152 (*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

$$\mathcal{P}^{-1} := \frac{\overline{\mathcal{P}}}{\|\mathcal{P}\|^2} = \frac{1}{\|\mathcal{P}\|^2}(s, -v).$$

Quaternion Algebra

Lemma 153

The quaternion $(1, 0) \in \mathbb{H}$ is the neutral element of quaternion multiplication.

Lemma 154

The multiplication of quaternions is associative and distributive (over addition) but not commutative.

Lemma 155

For every quaternion $\mathcal{P} \in \mathbb{H}$ with $\mathcal{P} \neq 0$, we have

$$\mathcal{P}^{-1} \cdot \mathcal{P} = (1, 0) = \mathcal{P} \cdot \mathcal{P}^{-1}.$$

Lemma 156

For all $\mathcal{P}, \mathcal{Q} \in \mathbb{H}$, we have

$$\overline{\overline{\mathcal{P}}} = \mathcal{P} \qquad \text{and} \qquad \overline{\mathcal{P} + \mathcal{Q}} = \overline{\mathcal{Q}} + \overline{\mathcal{P}} \qquad \text{and} \qquad \overline{\mathcal{P} \cdot \mathcal{Q}} = \overline{\mathcal{Q}} \cdot \overline{\mathcal{P}}.$$

Quaternion Algebra

Lemma 157

For all $\mathcal{P}, \mathcal{Q} \in \mathbb{H}$ with $\mathcal{P}, \mathcal{Q} \neq 0$, we have

$$(\mathcal{P}^{-1})^{-1} = \mathcal{P} \qquad \text{and} \qquad (\mathcal{P} \cdot \mathcal{Q})^{-1} = \mathcal{Q}^{-1} \cdot \mathcal{P}^{-1}.$$

Lemma 158

The inverse of a unit quaternion and the product of unit quaternions are themselves unit quaternions.

- 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.
- See `https://eater.net/quaternions` for a neat set of videos on "visualizing quaternions".

Geometric Objects

Lines

Definition 159 (*Straight line, Dt.: Gerade*)

For two distinct points $p, q \in \mathbb{R}^n$, the *straight line* defined by p, q is the set

$$\ell(p,q) := \{p + \lambda \cdot pq \colon \lambda \in \mathbb{R}\}.$$

- Recall that $pq := q - p$.
- Since, for all $\lambda \in \mathbb{R}$,

  $$p + \lambda \cdot pq = p + \lambda \cdot (q - p) = (1 - \lambda) \cdot p + \lambda \cdot q,$$

  we have

  $$\ell(p,q) = \{\alpha \cdot p + \beta \cdot q \colon \alpha, \beta \in \mathbb{R} \text{ with } \alpha + \beta = 1\}.$$

  Hence, $\ell(p,q)$ is the set of all affine combinations of p and q.
- $p + \lambda \cdot pq$ is the so-called *parametric representation* of $\ell(p,q)$.

Definition 160 (*Ray, Dt.: Strahl, Halbgerade*)

For two distinct points $p, q \in \mathbb{R}^n$, the *ray* starting at p through q is the set

$$\{p + \lambda \cdot pq \colon \lambda \in \mathbb{R}_0^+\}.$$

Lines and Straight-Line Segments

Definition 161 (*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

$$\overline{pq} := \{p + \lambda \cdot pq \colon \lambda \in [0, 1]\}.$$

- Since, for all $\lambda \in [0, 1]$,

  $$p + \lambda \cdot pq = (1 - \lambda) \cdot p + \lambda \cdot q,$$

  we have

  $$\overline{pq} = \{\alpha \cdot p + \beta \cdot q \colon \alpha, \beta \in \mathbb{R}_0^+ \text{ with } \alpha + \beta = 1\}.$$

  Hence, $\overline{pq}$ is the set of all convex combinations of p and q.

Definition 162 (*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

$$\{p + \lambda \cdot pq \colon \lambda \in]0, 1[\}.$$

Lines in $\mathbb{R}^2$

Lemma 163

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

$$\ell(p, q) = \{u \in \mathbb{R}^2 : \langle u, n \rangle = c\}.$$

- The equation $\langle u, n \rangle = c$ is the so-called *equational representation* of $\ell(p, q)$, aka *implicit form*.
- 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 $\lambda \in \mathbb{R}$.
- Standard formulation according to high school math:

$$a \cdot x + b \cdot y = c, \qquad \text{with} \quad n := \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{and} \quad u := \begin{pmatrix} x \\ y \end{pmatrix}.$$

Lines in $\mathbb{R}^2$

Definition 164 (*Hessian normal form, Dt.: Hessische Normalform*)

A line equation $\langle u, n \rangle = c$ for $\ell(p, q)$, as specified in Lem. 163, is said to be in *Hessian normal form* if n is a unit vector.

Lemma 165

The (signed) minimum distance d of a point $a \in \mathbb{R}^2$ from $\ell(p, q)$, with $\ell(p, q) = \{u \in \mathbb{R}^2 \colon \langle u, n \rangle = c\}$, is given by

$$d = \frac{\langle a, n \rangle - c}{\|n\|}.$$

- The signed distance of point $a \in \mathbb{R}^2$ from $\ell(p, q) = \{u \in \mathbb{R}^2 \colon \langle u, n \rangle = c\}$ is positive if a is on that side of $\ell(p, q)$ into which n points.

Planes in $\mathbb{R}^3$

Definition 166 (*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

$$\varepsilon(p, q, r) := \{p + \lambda \cdot pq + \mu \cdot pr \colon \lambda, \mu \in \mathbb{R}\}.$$

- $p + \lambda \cdot pq + \mu \cdot pr$ is the so-called parametric representation of $\varepsilon(p, q, r)$.
- Since, for all $\lambda, \mu \in \mathbb{R}$,

  $$p + \lambda \cdot pq + \mu \cdot pr = p + \lambda \cdot (q - p) + \mu \cdot (r - p) = (1 - \lambda - \mu) \cdot p + \lambda \cdot q + \mu \cdot r,$$

  we have

  $$\varepsilon(p, q, r) := \{\alpha \cdot p + \beta \cdot q + \gamma \cdot r \colon \alpha, \beta, \gamma \in \mathbb{R} \text{ with } \alpha + \beta + \gamma = 1\}.$$

  Hence, $\varepsilon(p, q, r)$ is the set of all affine combinations of p, q and r.

Planes in $\mathbb{R}^3$

Lemma 167

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

$$\varepsilon(p, q, r) = \{u \in \mathbb{R}^3 : \langle u, n \rangle = c\}.$$

- The equation $\langle u, n \rangle = c$ 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

$$n = \lambda(pq \times pr) \quad \text{for some non-zero scalar } \lambda \in \mathbb{R}.$$

Planes in $\mathbb{R}^3$

Definition 168 (*Hessian normal form, Dt.: Hessische Normalform*)

A plane equation $\langle u,n\rangle = c$ for $\varepsilon(p,q,r)$, as specified in Lem. 167, is said to be in *Hessian normal form* if n is a unit vector.

Lemma 169

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.

Line/Plane Equation via Determinant

Lemma 170

The equation of the line through two distinct points p and q in $\mathbb{R}^2$ is given by

$$\det \begin{pmatrix} x & y & 1 \\ p_x & p_y & 1 \\ q_x & q_y & 1 \end{pmatrix} = 0.$$

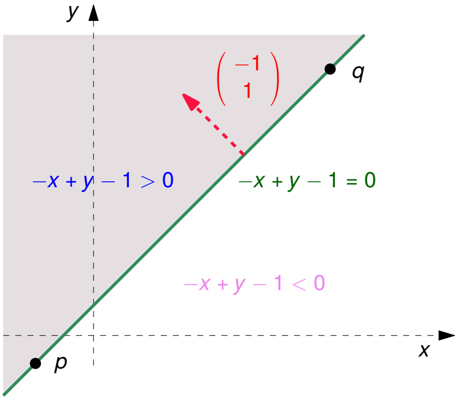
Lemma 171

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 $\ell(p,q) = \{u \in \mathbb{R}^2 \colon \langle u,n\rangle = c\}$ partitions $\mathbb{R}^2$ into three disjoint sets: the actual line and the two (open) *half-planes* $\{u \in \mathbb{R}^2 \colon \langle u,n\rangle - c < 0\}$ and $\{u \in \mathbb{R}^2 \colon \langle u,n\rangle - c > 0\}$.



- Similarly for a plane in $\mathbb{R}^3$ and *half-spaces*.

Intersections of Lines and Planes

- The intersection of two lines $a_1x + b_1y = c_1$ and $a_2x + b_2y = c_2$ in $\mathbb{R}^2$ is given by the solution(s) of the following system of two linear equations:

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

That is,

$$\mathbf{A}u = c \qquad \text{with} \quad \mathbf{A} := \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \quad u := \begin{pmatrix} x \\ y \end{pmatrix} \quad c := \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

- Similarly for the intersection of m (hyper-)planes in $\mathbb{R}^n$:

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

Circles in $\mathbb{R}^2$ and Spheres in $\mathbb{R}^3$

Definition 172 (*Sphere, Dt.: Sphäre, Kugeloberfläche*)

The *(hyper-)sphere* in $\mathbb{R}^n$ with radius $r \in \mathbb{R}$ centered at point $c \in \mathbb{R}^n$, under the Euclidean distance $d(\cdot,\cdot)$, 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 $\mathbb{R}^3$.

Definition 173 (*Disk, Dt.: Kreisscheibe*)

The *(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 174 (*Open disk*)

The *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 175 (*Ball, Dt.: Kugel*)

The *(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 176 (*Open ball*)

The *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

Lemma 177

The equation of a circle in $\mathbb{R}^2$ (under the Euclidean distance) with radius $r \in \mathbb{R}_0^+$ centered at point $c \in \mathbb{R}^2$ is given by

$$(c_x - x)^2 + (c_y - y)^2 = r^2.$$

Lemma 178

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

Lemma 179

The equation of a sphere in $\mathbb{R}^3$ (under the Euclidean distance) with radius $r \in \mathbb{R}_0^+$ centered at point $c \in \mathbb{R}^3$ is given by

$$(c_x - x)^2 + (c_y - y)^2 + (c_z - z)^2 = r^2.$$

Lemma 180

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

$$\det \begin{pmatrix} x^2 + y^2 + z^2 & x & y & z & 1 \\ x_1^2 + y_1^2 + z_1^2 & x_1 & y_1 & z_1 & 1 \\ x_2^2 + y_2^2 + z_2^2 & x_2 & y_2 & z_2 & 1 \\ x_3^2 + y_3^2 + z_3^2 & x_3 & y_3 & z_3 & 1 \\ x_4^2 + y_4^2 + z_4^2 & x_4 & y_4 & z_4 & 1 \end{pmatrix} = 0.$$

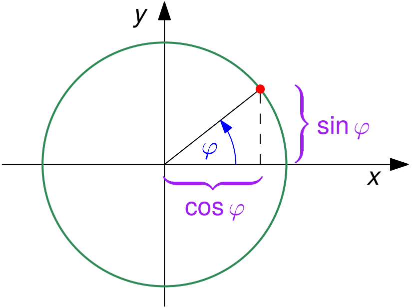
- This formula generalizes to any number of dimensions.

Parametrization of a Circle

Lemma 181

The parametrization of a circle in $\mathbb{R}^2$ with radius $r \in \mathbb{R}_0^+$ centered at point $c \in \mathbb{R}^2$ is given by

$$\begin{pmatrix} c_x + r\cos\varphi \\ c_y + r\sin\varphi \end{pmatrix} \qquad \text{with } \varphi \in [0, 2\pi[.$$

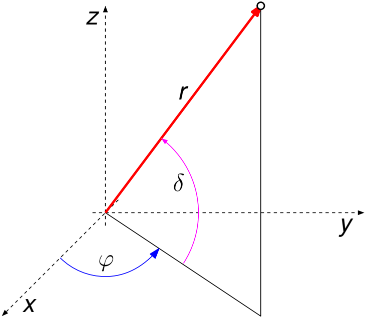


Parametrization of a Sphere

Lemma 182

The parametrization of a sphere in $\mathbb{R}^3$ with radius $r \in \mathbb{R}_0^+$ centered at point $c \in \mathbb{R}^3$ is given by

$$\begin{pmatrix} c_x + r\cos\delta\cos\varphi \\ c_y + r\cos\delta\sin\varphi \\ c_z + r\sin\delta \end{pmatrix} \qquad \text{with } \varphi \in [0, 2\pi[\text{ and } \delta \in [-\frac{\pi}{2}, \frac{\pi}{2}].$$



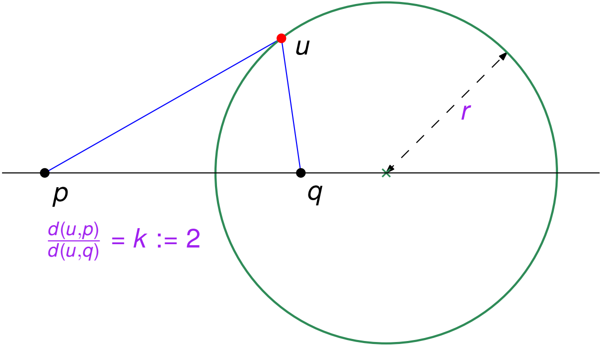
Sphere via Ratios of Distances

Lemma 183 (*Appolonius of Perga*)

Consider two distinct points $p, q \in \mathbb{R}^n$ and a constant $k \in \mathbb{R}^+$. Then

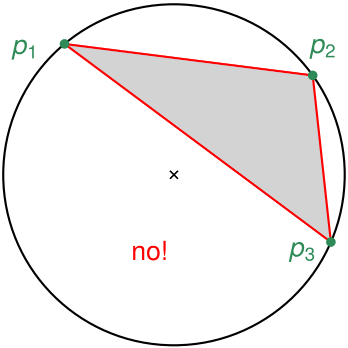
$$\{u \in \mathbb{R}^n \colon \frac{d(u,p)}{d(u,q)} = k\}$$

forms a (hyper-)sphere.



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.

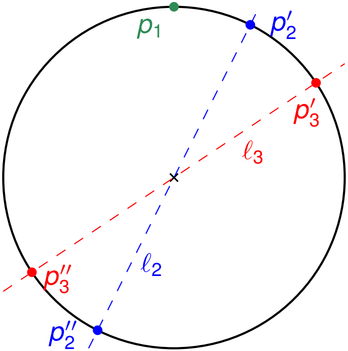


Putnam Problem: Points on a Sphere

- W.l.o.g., the point p_1 is at the north pole of the circle, centered at the origin.
- 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 ℓ_2 through the origin — and then flipping a coin to choose between p_2' and p_2''.
- Same for ℓ_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

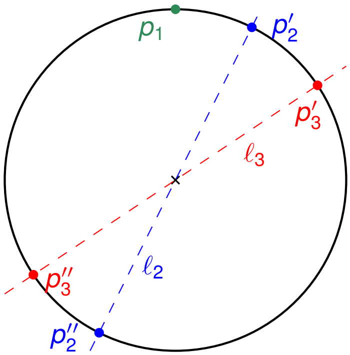
- Actually, we have $\lambda_1, \lambda_2, \lambda_3$ all non-zero. W.l.o.g., $\lambda_1 > 0$.
- If

$$0 = \lambda_1 \cdot p_1 + \lambda_2 \cdot p_2' + \lambda_3 \cdot p_3$$

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 $^1/_4$.
- Similarly, a random tetrahedron contains the center of the sphere with probability $^1/_8$.

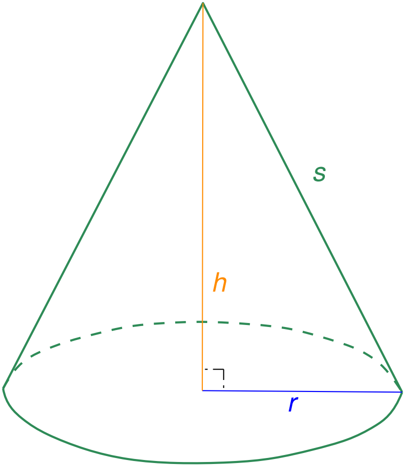


Cone

Definition 184 (*Cone, Dt.: Kegel*)

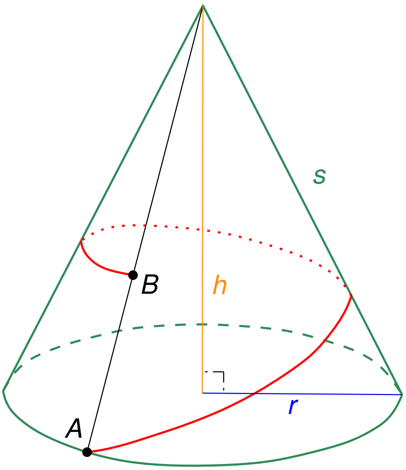
A (right circular) *cone* is formed by a set of line segments (or lines) which connect a common point, called the *apex*, to all the points of a circular base, where the apex lies on a perpendicular through the center of the circle. This line is called *axis* of the cone.

- 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.



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 120π, the opening angle of the disk sector is 120^o.
- The shortest distance from A to B is a straight-line segment.
- The law of cosines,

$$d(A,B)^2 = s^2 + (s-10)^2 + 2s(s-10)\cos 120,$$

  yields $d(A,B) = 10\sqrt{91}$.

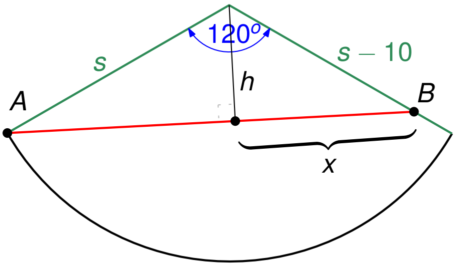
- Let x be the length of the downhill portion of the track. We have

$$(s-10)^2 = h^2 + x^2$$

  and

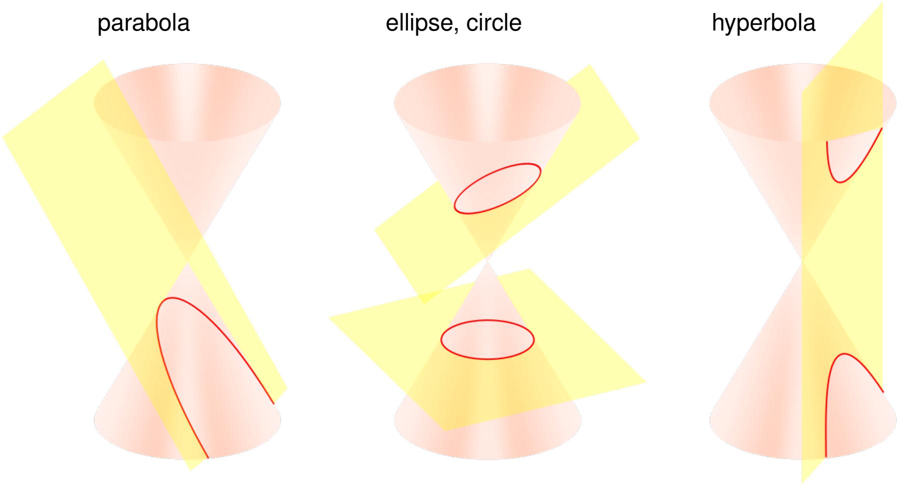
$$s^2 = h^2 + (d(A,B) - x)^2.$$

  We get $x = {}^{400}/\sqrt{91}$ as length of the downhill portion of the track.



Conics

- Conic sections (Dt.: Kegelschnitte) are formed by the intersection of a (double circular right) cone and a plane.



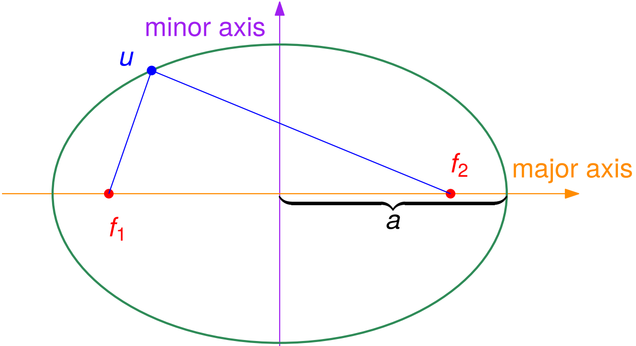
[Image credit: `https://en.wikipedia.org`.]

Ellipse

Definition 185 (*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:

$$\{u \in \mathbb{R}^2 \colon d(u, f_1) + d(u, f_2) = 2a\}$$



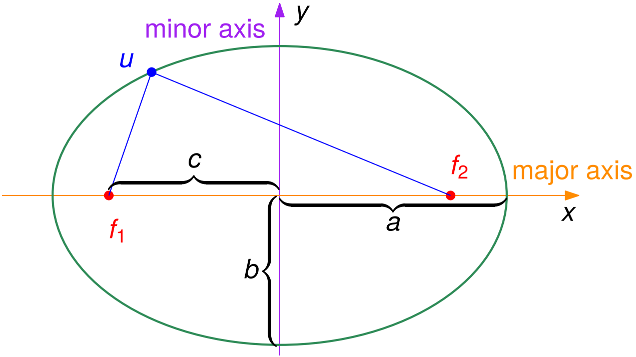
Ellipse

Lemma 186

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 $c = \sqrt{a^2 - b^2}$.

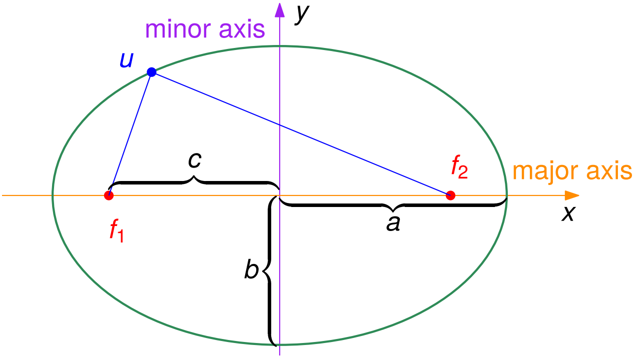


Ellipse

Lemma 187

The standard (axis-aligned) ellipse with width $2a$ and height $2b$ can be parametrized as

$$\begin{pmatrix} a \cdot \cos\varphi \\ b \cdot \sin\varphi \end{pmatrix} \qquad \text{with } \varphi \in [0, 2\pi[.$$



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.
- 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

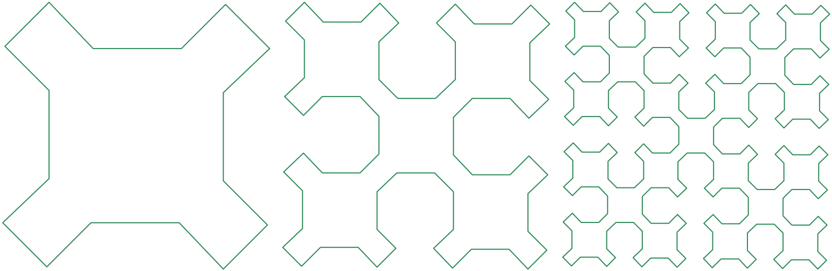
$$\begin{pmatrix} a \cdot \sin\delta\cos\varphi \\ b \cdot \sin\delta\sin\varphi \\ c\cos\delta \end{pmatrix} \qquad \text{with } \varphi \in [0, 2\pi[\text{ and } \delta \in [-\frac{\pi}{2}, \frac{\pi}{2}].$$

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 S_n in $\mathbb{R}^2$.
- Sierpinski curves of orders 1–3 :



- Their limit curve, *the Sierpinski curve*, is a space-filling curve: In the limit, for $n \to \infty$, it fills the unit square completely!
- Its length grows exponentially and unboundedly as n grows.
- Other space-filling curves exist: E.g., Peano curve, Hilbert curve.

Curves in $\mathbb{R}^n$

Definition 188 (*Curve, Dt.: Kurve*)

Let $I \subseteq \mathbb{R}$ be an interval of the real line. A continuous (vector-valued) mapping $\gamma\colon I \to \mathbb{R}^n$ 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., $\gamma\colon [0,1] \to \mathbb{R}^3$ 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 γ, and $\gamma(I)$ is called *image* (Dt.: Bild, Spur).

Definition 189 (*Plane curve, Dt.: ebene Kurve*)

For $\gamma\colon I \to \mathbb{R}^n$, the curve $\gamma(I)$ is *plane* if $\gamma(I) \subseteq \mathbb{R}^2$ or if $\gamma(I)$ lies within a 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 190 (*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 $\gamma\colon I \to \mathbb{R}^n$.

Definition 191 (*Closed, Dt.: geschlossen*)

A parametrization $\gamma\colon I \to \mathbb{R}^n$ 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 192 (*Simple, Dt.: einfach*)

A parametrization $\gamma\colon I \to \mathbb{R}^n$ 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 $\gamma\colon I \to \mathbb{R}^n$ is simple then it is injective on $]a, b[$.

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 $\mathcal{C}$ to be simple if there exists one parametrization of $\mathcal{C}$ 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 $\mathcal{C}$ and one of its parametrizations γ if the meaning is clear.
- Similarly, we will frequently call γ a curve.
- For instance, we will frequently speak about a closed curve rather than about a closed parametrization of a curve.

Convex Curve in $\mathbb{R}^2$

Definition 193 (*Supporting line, Dt.: Stützgerade*)

In $\mathbb{R}^2$, a line ℓ is a supporting line of a curve $\mathcal{C}$ if

1. ℓ passes through a point of $\mathcal{C}$,
2. $\mathcal{C}$ lies completely in one of the two closed half-planes induced by ℓ.

- There may be many supporting lines for a curve at a given point.
- If a tangent exists at a given point, then it is the unique supporting line at this point if it does not separate the curve.

Definition 194 (*Convex curve*)

In $\mathbb{R}^2$, a curve is convex if it has a supporting line through each of its points.

Lemma 195

Every convex curve is a subset of the boundary of its own convex hull.

- It is straightforward to extend the notion of convexity from $\mathbb{R}^2$ to plane curves.

Jordan Curve in $\mathbb{R}^2$

Definition 196 (*Jordan curve, Dt.: Jordankurve*)

A set $\mathcal{C} \subset \mathbb{R}^2$ (which is not a single point) is called a *Jordan curve* if there exists a simple and closed parametrization $\gamma\colon I \to \mathbb{R}^2$ that parameterizes $\mathcal{C}$.

Theorem 197 (*Jordan 1887*)

Every Jordan curve $\mathcal{C}$ partitions $\mathbb{R}^2 \setminus \mathcal{C}$ into two disjoint open regions, a (bounded) "interior" region and an (unbounded) "exterior" region, with $\mathcal{C}$ 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 198 (*Schönflies 1906*)

For every Jordan curve $\mathcal{C}$ there exists a homeomorphism from the plane to itself that maps $\mathcal{C}$ 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 199 (*Tangent vector, Dt.: Tangentenvektor*)

Consider a differentiable parametrization $\gamma\colon I \to \mathbb{R}^n$ of a curve $\mathcal{C}$. For $t \in I$, a *tangent vector* at $\gamma(t)$ with respect to γ 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.

Frenet Frame for Curves in $\mathbb{R}^3$

Definition 200 (*Frenet frame, Dt.: begleitendes Dreibein*)

Let $\gamma\colon I \to \mathbb{R}^3$ be a C^2 curve that is regular of order two. Then the *Frenet frame* (aka *moving trihedron*) at $\gamma(t)$ is defined as an orthonormal basis of vectors $T(t), N(t), B(t)$ as follows:

$$T(t) := \frac{\gamma'(t)}{\|\gamma'(t)\|} \qquad \text{unit tangent;}$$

$$N(t) := \frac{T'(t)}{\|T'(t)\|} \qquad \text{unit (principal) normal;}$$

$$B(t) := T(t) \times N(t) \qquad \text{unit binormal.}$$

Lemma 201

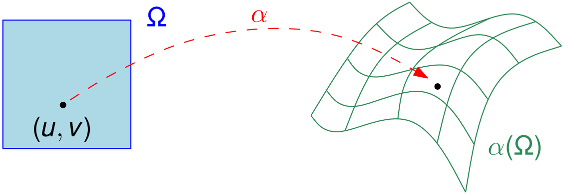
Let $\gamma\colon I \to \mathbb{R}^3$ be a C^2 curve that is regular of order two, and define T, N, B as in Def. 200. We get for all $t \in I$:

- $N(t)$ is normal to $T(t)$, and
- $B(t)$ is a unit vector.

Surfaces in $\mathbb{R}^3$

Definition 202 (*Parametric surface*)

Let $\Omega \subseteq \mathbb{R}^2$. A continuous mapping $\alpha\colon \Omega \to \mathbb{R}^3$ is called a *parametrization* of $\alpha(\Omega)$, and $\alpha(\Omega)$ is called the (parametric) *surface* parameterized by α.

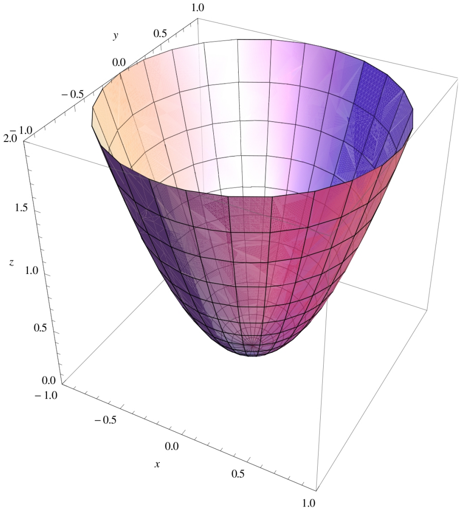


- For instance, every point on the surface of Earth can be described by the geographic coordinates longitude and latitude.
- Note that parametrizations of a surface (regarded as a set $\mathcal{S} \subset \mathbb{R}^3$) need not be unique: two different parametrizations α and β may exist such that $\mathcal{S} = \alpha(\Omega_1) = \beta(\Omega_2)$.
- 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

$$\alpha\colon [0,1] \times [0,2\pi] \to \mathbb{R}^3$$

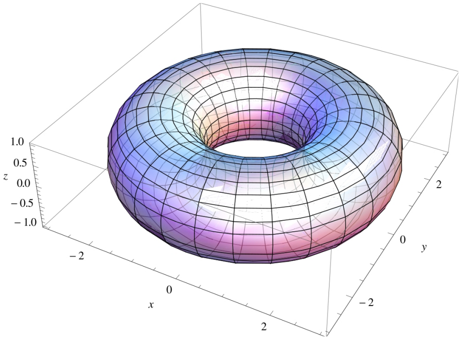
$$\alpha(u,v) := \begin{pmatrix} u\cos v \\ u\sin v \\ 2u^2 \end{pmatrix}$$



Sample Parametric Surface: Torus

$$\alpha\colon [0, 2\pi]^2 \to \mathbb{R}^3$$

$$\alpha(u, v) := \begin{pmatrix} (2 + \cos v) \cos u \\ (2 + \cos v) \sin u \\ \sin v \end{pmatrix}$$



Surfaces in $\mathbb{R}^3$

Lemma 203

Consider a differentiable parametrization $\alpha\colon \Omega \to \mathbb{R}^3$ of a surface $\mathcal{S}$. For $(s,t) \in \Omega$, tangent vectors at $\alpha(s,t)$ with respect to α are given by $\frac{\partial\alpha}{\partial s}(s,t)$ and $\frac{\partial\alpha}{\partial t}(s,t)$.

Definition 204 (*Normal vector, Dt.: Normalvektor*)

Consider a differentiable parametrization $\alpha\colon \Omega \to \mathbb{R}^3$ of a surface $\mathcal{S}$. A *normal vector* $n_\alpha(s,t)$ at $\alpha(s,t)$ with respect to α 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 $\alpha(s,t)$.

Polygonal Curve

Definition 205 (*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 $\gamma\colon [0, n] \to \mathbb{R}^d$ with

$$\gamma(t) := p_i + (t - i) \cdot (p_{i+1} - p_i) \quad \text{if } t \in [i, i+1] \text{ for some } i \in \{1, 2, \ldots, n-1\}.$$

- 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.
- It is common to extend this definition by allowing $n = 0$, in which case we get a single point.
- Unless stated otherwise, we will always assume that all vertices of a polygonal curve are co-planar, i.e., that the polygonal curve is plane. The default plane is $\mathbb{R}^2$.

Polygon

Definition 206 (*Polygon*)

For $n \in \mathbb{N}$ with $n \geq 3$, a *polygon* with vertices $p_0, p_1, p_2, \ldots, p_n \in \mathbb{R}^d$, aka *n-gon*, is a polygonal curve such that $p_0 = p_n$.

Definition 207 (*Simple polygon, Dt.: einfaches Polygon*)

A polygon is *simple* if it admits a simple parametrization.

- If a plane polygon $\mathcal{P}$ is simple then, by the Jordan Curve Theorem, it splits the plane into two regions, one of which is bounded.
- In this case it is common to be a bit liberal and use the term “polygon” for either the (simple) polygonal curve $\mathcal{P}$ or for the entire (closed) region bounded by $\mathcal{P}$; the actual meaning has to be inferred from the context.
- If $\mathcal{P}$ is regarded to be only the simple polygonal curve then the bounded region (without $\mathcal{P}$ itself) is called the polygon’s *interior*, and points within that region are said to be *inside* of $\mathcal{P}$.

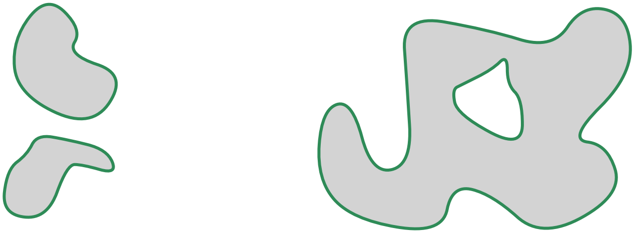
Connectedness

Definition 208 (*Path-connected, Dt.: wegzusammenhängend*)

A set $\mathcal{S} \subset \mathbb{R}^n$ is *path-connected* if for every pair of points $p, q \in \mathcal{S}$ there exists a curve that is completely contained in $\mathcal{S}$ and that links p and q.

Definition 209 (*Simply-connected and multiply-connected*)

A path-connected set $\mathcal{S} \subset \mathbb{R}^2$ is *simply-connected* if every simple closed curve entirely contained within $\mathcal{S}$ encloses only points of $\mathcal{S}$. Otherwise, $\mathcal{S}$ is called *multiply-connected* (or *not simply-connected*).



not path-connected   path-connected, multiply-connected

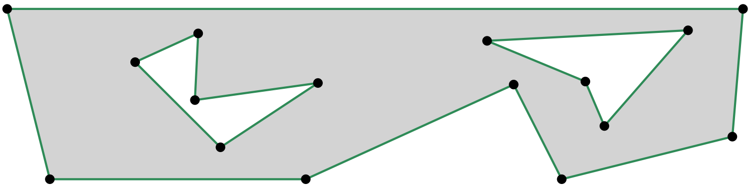
Polygonal Region

Definition 210 (*Polygonal region*)

A *polygonal region* is a (possibly) multiply-connected but connected subset of $\mathbb{R}^2$ that is bounded by k simple polygons $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_k$, for some $k \in \mathbb{N}$, such that

1. no pair of polygons (seen as curves) intersect,
2. the polygons $\mathcal{P}_2, \ldots, \mathcal{P}_k$ lie in the interior of $\mathcal{P}_1$,
3. for $2 \leq i, j \leq k$, the polygon $\mathcal{P}_i$ does not lie in the interior of the polygon $\mathcal{P}_j$.

The polygon $\mathcal{P}_1$ is called *outer polygon* and the polygons $\mathcal{P}_2, \ldots, \mathcal{P}_k$ are called *islands* or *holes*.



Area and Orientation of a Polygon

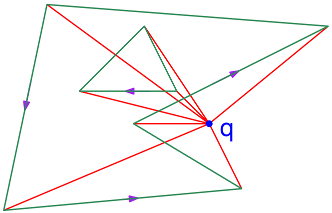
Theorem 211 (*Meister (1769), Gauß (1795)*)

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 q in the plane. Then the (signed) area of $\mathcal{P}$ is given by the sum of the signed areas of the individual triangles $\Delta(q, p_{i-1}, p_i)$. That is, the (signed) area of $\mathcal{P}$ equals

$$\sum_{i=1}^{n} A_\Delta(q, p_{i-1}, p_i) = \frac{1}{2} \cdot [(x_0 y_1 - x_1 y_0) + (x_1 y_2 - x_2 y_1) + \cdots + (x_{n-1} y_0 - x_0 y_{n-1})],$$

where $p_i := \binom{x_i}{y_i}$. The signed area of $\mathcal{P}$ is positive if and only if $\mathcal{P}$ is oriented 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!



Planar Straight-Line Graph

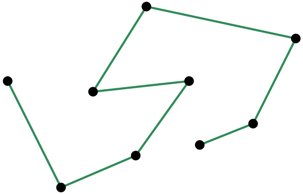
Definition 212 (*Planar straight-line graph*)

A *planar straight-line graph* (PSLG) is a finite collection of isolated vertices and straight-line segments such that

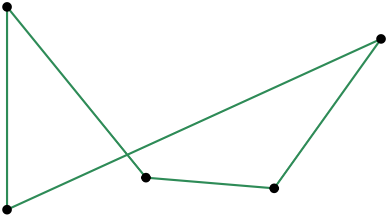
- each two segments intersect only in vertices shared by both of them,
- no segment passes through a vertex other than one of its two end-points.

- Hence, a PSLG is an embedding of a planar graph such that all its edges are drawn as straight-line segments.
- Aka: Plane geometric graph.
- Hence, simple polygonal curves and simple polygons are special PSLGs.
- Of course, Euler's Theorem applies to the faces, edges and vertices of a PSLG.

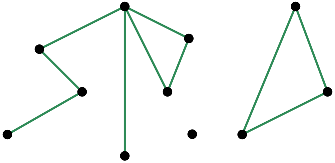
Sample Polygonal Chains and PSLGs



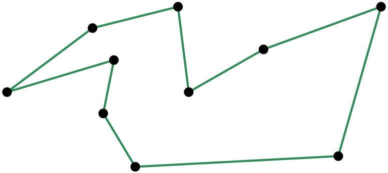
polygonal curve



polygon, not simple



planar straight-line graph



simple polygon

Polyhedron

- Unfortunately, even in $\mathbb{R}^3$ there there is no universal agreement over how to define the analogue to a polygon in $\mathbb{R}^3$ . . .

Grünbaum (1994)

"The Original Sin in the theory of polyhedra goes back to Euclid, . . . and many others, . . . at each stage . . . the writers failed to define what are the polyhedra."

Definition 213 (*Polyhedron, Dt.: Polyeder*)

A *polyhedron* in $\mathbb{R}^3$ is either

- a (possibly unbounded) solid given by the intersection of finitely many halfspaces, or
- a connected bounded solid whose boundary is formed by a finite collection of plane polygons ("faces") such that
  1. each vertex is incident to at least three edges and faces,
  2. each edge is shared by exactly two faces,
  3. each two faces intersect only in vertices and edges shared by both of them,
  4. the faces that share a vertex form a cyclic chain of polygons in which every pair of consecutive polygons shares an edge.

Polyhedron

Polyhedron versus Polytope

1. For convex solids, some authors (in some fields of mathematics) prefer to use the term "polytope" for a bounded polyhedron, whereas "polyhedron" is a generic convex object.
2. From this point of view, a polyhedron is the intersection of a finite number of halfspaces and is defined by its faces whereas a polytope is the convex hull of a finite number of points and is defined by its vertices.

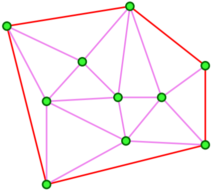
- The situation gets worse once different fields of mathematics and computer science are considered!
- Note: Plural of "polyhedron" is "polyhedra".
- Recall that Euler's Formula $v - e + f = 2$ holds for the vertices, edges and faces of a polyhedron.

Triangulation

Definition 214 (*Triangulation*)

Let $S = \{P_1, P_2, \ldots, P_k\}$ be a set of k points in $\mathbb{R}^2$. A planar straight-line graph T is called a *triangulation* of S if

- S forms 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.

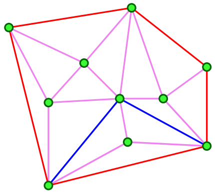


Constrained Triangulation

Definition 215 (*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 planar straight-line graph T is called a *constrained triangulation* of S if

- S forms 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.



Basic Concepts of Topology

Metric Space and Open Ball

Definition 216 (*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: $d(x, y) = 0 \quad \Rightarrow \quad x = y$.
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 217 (*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

$$B(x, r) := \{y \in \mathcal{X} : d(x, y) < r\}.$$

Interior, Exterior and Closure

- Consider a space $\mathcal{X}$ that has a metric, and a set $\mathcal{S} \subseteq \mathcal{X}$. (E.g., $\mathbb{R}^n$ and the Euclidean metric, and any subset $\mathcal{S}$ of $\mathbb{R}^n$.)

Definition 218 (*Interior point, Dt.: innerer Punkt*)

A point $x \in \mathcal{X}$ is an *interior point* of $\mathcal{S}$ if there exists a radius $r > 0$ such that the open ball with center x and radius r is completely contained in $\mathcal{S}$, i.e., $B(x,r) \subseteq \mathcal{S}$.

Definition 219 (*Interior, Dt.: Inneres*)

The set of all interior points of $\mathcal{S}$ is the *interior* of $\mathcal{S}$, often denoted by $\text{int}(\mathcal{S})$ or $\mathcal{S}^\circ$.

Lemma 220

We have $\text{int}(\mathcal{S}) \subseteq \mathcal{S}$ for all $\mathcal{S} \subseteq \mathcal{X}$.

Lemma 221

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

Definition 222 (*Exterior point, Dt.: äußerer Punkt*)

A point $y \in \mathcal{X}$ is an *exterior point* of $\mathcal{S}$ 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 $\mathcal{S}$ (with respect to $\mathcal{X}$), i.e., $B(y, r) \subseteq (\mathcal{X} \setminus \mathcal{S})$.

Definition 223 (*Exterior, Dt.: Äußeres*)

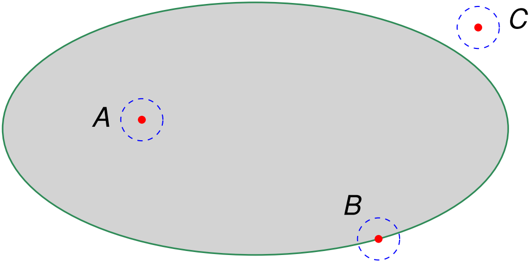
The set of all exterior points of $\mathcal{S}$ is the *exterior* of $\mathcal{S}$, denoted by $\text{ext}(\mathcal{S})$.

Definition 224 (*Boundary, Dt.: Rand*)

All points of $\mathcal{X}$ that are neither in the interior nor in the exterior of $\mathcal{S}$ form the *boundary*, $\partial\mathcal{S}$, of $\mathcal{S}$.

Interior, Exterior and Closure

- 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 225

For all $\mathcal{S} \subseteq \mathcal{X}$, the union of the interior, the exterior and the boundary of $\mathcal{S}$ constitutes the whole space $\mathcal{X}$.

Interior, Exterior and Closure

Definition 226 (*Closure, Dt.: Abschluss*)

The *closure* $\overline{\mathcal{S}}$ of a set $\mathcal{S}$ is the union of the interior and the boundary of $\mathcal{S}$.

Lemma 227

The closure $\overline{\mathcal{S}}$ of a set $\mathcal{S}$ is given by all points of $\mathcal{X}$ that are not in the exterior of $\mathcal{S}$.

Definition 228 (*Open, Dt.: offen*)

A set $\mathcal{S} \subseteq \mathcal{X}$ is called *open* if $\text{int}(\mathcal{S}) = \mathcal{S}$.

Definition 229 (*Closed, Dt.: abgeschlossen*)

A set $\mathcal{S} \subseteq \mathcal{X}$ is called *closed* if the complement of $\mathcal{S}$ (relative to $\mathcal{X}$) is open.

- Note that there exist spaces $\mathcal{X}$ and subsets $\mathcal{S} \subset \mathcal{X}$ such that the interior or the exterior or the boundary of $\mathcal{S}$ are empty.

Definition 230 (*Compact*)

A subset of a Euclidean space is called *compact* if it is bounded and closed.

Interior, Exterior and Closure

- Consider a ball in $\mathbb{E}^3$ with radius r centered at the origin:

$$\{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 + z^2 \leq r^2\}.$$

- The interior of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 + z^2 < r^2\}.$$

- The closure of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 + z^2 \leq r^2\}.$$

- The exterior of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 + z^2 > r^2\}.$$

- The boundary of the ball is

$$\{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 + z^2 = r^2\}.$$

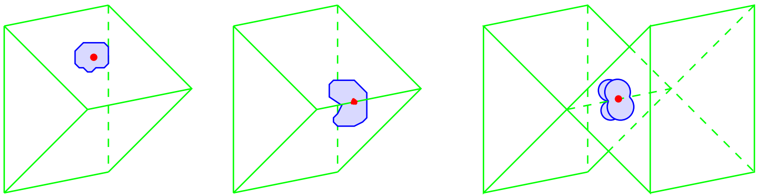
Manifolds

- 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 231 (*Manifold, Dt.: Mannigfaltigkeit*)

A set $\mathcal{S} \subset \mathbb{R}^3$ is a 2-*manifold* (or simply a "manifold") if for every point $x \in \mathcal{S}$ there exists an open neighborhood of x in $\mathcal{S}$ 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.



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 232 (*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.

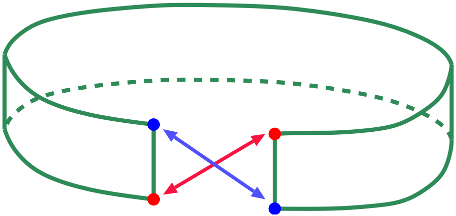
- Equivalently, a manifold of genus k can be obtained by adding k handles to the sphere.
- Note that a general surface can also be obtained by “punching holes” through a sphere.
- However, it is not difficult to see that, topologically, adding a handle is equivalent to opening a hole on a surface.

Orientable Surface

Definition 233 (*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` for a cool application of topology and, in particular, of Möbius strips.

Transformations

Linear Transformations

Definition 234 (*Linear transformation, Dt.: lineare Abbildung*)

Let V, W be vector spaces over $\mathbb{R}$. A transformation $g\colon V \to W$ is called a *linear transformation* if

1. $g(v_1 + v_2) = g(v_1) + g(v_2) \qquad \forall v_1, v_2 \in V$,
2. $g(\lambda v) = \lambda g(v) \qquad \forall v \in V, \forall \lambda \in \mathbb{R}$.

- E.g., $V := \mathbb{R}^n$ and $W := \mathbb{R}^m$ for some $m, n \in \mathbb{N}$.

Lemma 235

Every linear transformation maps

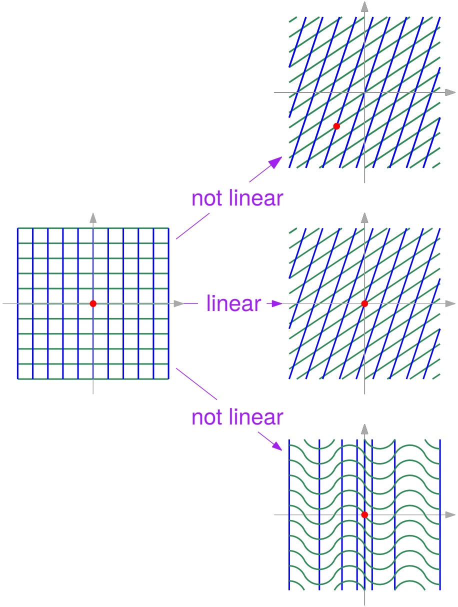
- a line to a line (or a point),
- the coordinate origin of V to the coordinate origin of W.

Sketch of Proof: A line $\{p + \lambda v\colon \lambda \in \mathbb{R}\}$ is mapped as follows:

$$g(\{p + \lambda v\colon \lambda \in \mathbb{R}\}) = \{g(p + \lambda v)\colon \lambda \in \mathbb{R}\} = \{g(p) + \lambda g(v)\colon \lambda \in \mathbb{R}\} \qquad \square$$

Linear Transformations

- 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 236

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 $g\colon V \to W$ 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

$$w = g(v) = g(\sum_{j=1}^{n} v_j e_j) = \sum_{j=1}^{n} v_j g(e_j) = \sum_{j=1}^{n} v_j (\sum_{i=1}^{m} a_{ij} e'_i) = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} v_j) e'_i$$
$$= \mathbf{A} v,$$

where $\mathbf{A} = [a_{ij}]_{i=1,j=1}^{m,n}$ and a_{ij} equals the i-th coordinate of $g(e_j)$. □

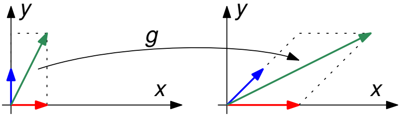
Linear Transformations

- 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 $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$:

$$\begin{aligned} g\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) &= g\left(1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 1 \cdot g\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + 2 \cdot g\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\ &= 1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 4 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{aligned}$$

- Thus, g has the following matrix:

$$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$



Linear Transformations

- 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 237

If a linear transformation has an inverse transformation then the inverse transformation is also linear.

Lemma 238

If a linear transformation g has an inverse transformation then the matrix which corresponds to g is invertible.

Composition of Linear Transformations

Definition 239 (*Composition, Dt.: Zusammensetzung*)

Consider two linear transformations $g\colon U \to V$ and $h\colon V \to W$. 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 $g(h(x))$!

Lemma 240

The composition of two linear transformations is a linear transformation.

- Hence, if $\mathbf{A}$ is the matrix of g and $\mathbf{B}$ is the matrix of h then $\mathbf{B} \cdot \mathbf{A}$ is the matrix of $h \circ g$.

Combining Matrix Transformations

- Suppose that p' is obtained by applying the matrix transformation $\mathbf{T}_1$ to p, and p'' is obtained from p' via $\mathbf{T}_2$, and so on till $p^{(n)}$:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{T}_1 \cdot \begin{pmatrix} x \\ y \end{pmatrix} \qquad \begin{pmatrix} x'' \\ y'' \end{pmatrix} = \mathbf{T}_2 \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \qquad \cdots \qquad \begin{pmatrix} x^{(n)} \\ y^{(n)} \end{pmatrix} = \mathbf{T}_n \cdot \begin{pmatrix} x^{(n-1)} \\ y^{(n-1)} \end{pmatrix}.$$

  Then the dependence of $p^{(n)}$ on p can be expressed as

$$\begin{aligned}\begin{pmatrix} x^{(n)} \\ y^{(n)} \end{pmatrix} &= \mathbf{T}_n \cdot \left(\mathbf{T}_{n-1} \cdot \left(\ldots \left(\mathbf{T}_2 \cdot \left(\mathbf{T}_1 \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right)\right)\right)\right) = \\ &= (\mathbf{T}_n \cdot \mathbf{T}_{n-1} \cdot \ldots \cdot \mathbf{T}_2 \cdot \mathbf{T}_1) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \\ &= \mathbf{T} \cdot \begin{pmatrix} x \\ y \end{pmatrix},\end{aligned}$$

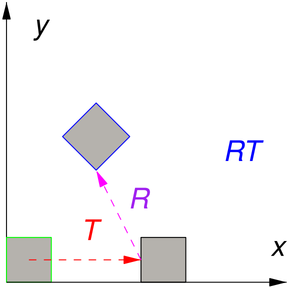
  where $\mathbf{T} := \mathbf{T}_n \cdot \mathbf{T}_{n-1} \cdot \ldots \cdot \mathbf{T}_2 \cdot \mathbf{T}_1$.

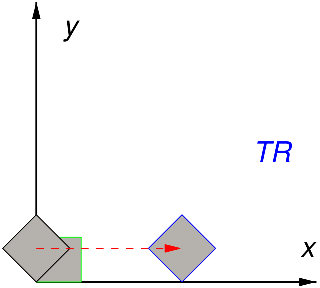
Caveats

- 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 $\mathbf{A}$ is the matrix that represents g then, via matrix multiplication,

$$g(x) = \mathbf{A}x.$$

- 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 for the unknown vector x:

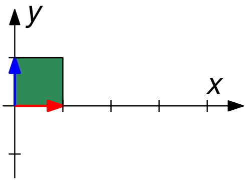
$$\mathbf{A}x = b.$$

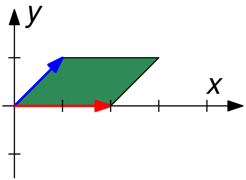
Geometric Interpretation of the Determinant of a Transformation Matrix

- Consider the linear transformation g with transformation matrix

$$\mathbf{T} := \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}.$$

- 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 $\det(\mathbf{T}) = 2$.



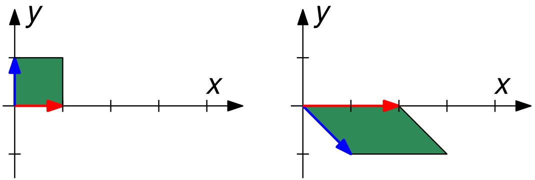


Geometric Interpretation of the Determinant of a Transformation Matrix

- Now consider the linear transformation g with transformation matrix

$$\mathbf{T} := \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}.$$

- 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 $\det(\mathbf{T}) = -2$, and that g changed the handedness of the unit vectors.

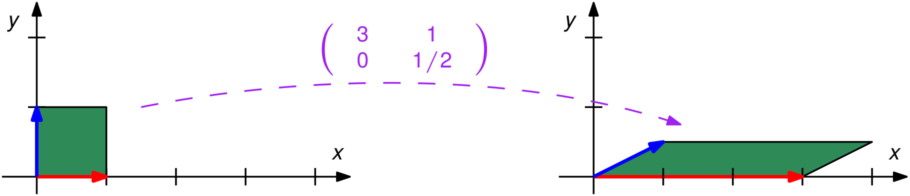


Geometric Interpretation of the Determinant of a Transformation Matrix

Theorem 241

The absolute value of the determinant of a (square) transformation matrix $\mathbf{A}$ gives the scale factor for the area/volume of the image of the unit (hyper-)cube. If $\det(\mathbf{A})$ is negative then the handedness of the unit vectors has changed, i.e., the orientation of space has been inverted.

Sketch of Proof: Theorem 119 settles this claim for 2×2 matrices. If the matrix $\mathbf{A}$ 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 $\prod_{i=1}^{n} a_{ii} = \det(\mathbf{A})$.
If $\mathbf{A}$ 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 112: A square matrix **A** is invertible if and only if $\det(\mathbf{A}) \neq 0$.
- 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 $u \in \mathbb{R}^n$,

$$\mathbf{A}^{-1} w = u \qquad \text{for} \quad w := \mathbf{A}u.$$

- Of course, mapping the image $g(u) =: w$ of u back to u can 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 $\det(\mathbf{A})$.
- We understand that **A** is invertible if and only if $\det(\mathbf{A}) \neq 0$.
- If $\det(\mathbf{A}) = 0$ then a solution to the linear equation $\mathbf{A}u = b$ exists if and only if b lies within the subspace $g(\mathbb{R}^n)$ of $\mathbb{R}^n$.

Geometric Interpretation of the Rank of a Transformation Matrix

Definition 242 (*Image, Dt.: Bild*)

The *image* (or *column space*) of an $m \times n$ matrix $\mathbf{A}$ (of a linear transformation g) is the set of all vectors $\mathbf{A}u$ for $u \in \mathbb{R}^n$, i.e., it equals $g(\mathbb{R}^n) \subset \mathbb{R}^m$.

- A solution to the linear equation $\mathbf{A}u = b$ exists if and only if b lies within the image of $\mathbf{A}$.
- Recall Definition 94: The rank of an $m \times n$ matrix $\mathbf{A}$ is the number of linearly independent columns of $\mathbf{A}$.
  1. If g squashes $\mathbb{R}^n$ to a line then the rank of $\mathbf{A}$ equals 1.
  2. If g squashes $\mathbb{R}^n$ to a plane then the rank of $\mathbf{A}$ equals 2.
  3. ...
- Hence, the rank of $\mathbf{A}$ equals the dimension of the image of $\mathbf{A}$.
- Note that the image $g(\mathbb{R}^n)$ forms a subspace of $\mathbb{R}^m$.

Rank, Image and Kernel of a Transformation Matrix

Definition 243 (*Kernel, Dt.: Kern*)

The *kernel* (or *null space*) of an $m \times n$ matrix $\mathbf{A}$ (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 $\mathbf{A}u = b$ then $u_0 + w$ is also a solution of $\mathbf{A}u = b$ for all w in the kernel of $\mathbf{A}$.
- The kernel of an $m \times n$ matrix forms a subspace of $\mathbb{R}^n$.

Definition 244 (*Corank, Dt.: Defekt*)

The *corank* (nullity) of an $m \times n$ matrix $\mathbf{A}$, denoted by $\text{corank}(\mathbf{A})$, is the dimension of the kernel of $\mathbf{A}$.

Theorem 245 (*Rank-nullity theorem, Dt.: Rangsatz, Dimensionssatz*)

Consider an $m \times n$ matrix $\mathbf{A}$. Then

$$\text{rank}(\mathbf{A}) + \text{corank}(\mathbf{A}) = n.$$

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 $a, b \in \mathbb{R}^n$.
- In Lemma 138 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 1×2 matrix $\mathbf{A}$ that maps every $b \in \mathbb{R}^2$ to a value in $\mathbb{R}$:

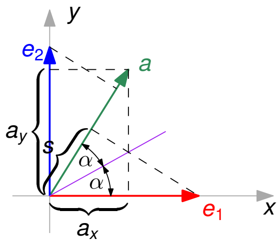
$$\begin{aligned}\langle a,b\rangle &= a_x \cdot b_x + a_y \cdot b_y = \begin{pmatrix} a_x & a_y \end{pmatrix} \cdot \begin{pmatrix} b_x \\ b_y \end{pmatrix} \\ &= \mathbf{A} \cdot b \qquad \text{with} \quad \mathbf{A} := \begin{pmatrix} a_x & a_y \end{pmatrix}\end{aligned}$$

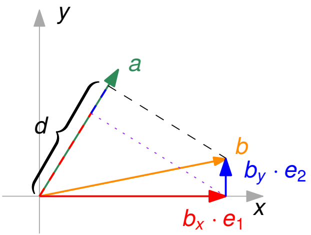
- We know that a linear transformation is fully specified by the images of the unit vectors.
- So, how do the unit vectors e_1, e_2 of $\mathbb{R}^2$ get mapped by this transformation? And what is the geometric interpretation of this transformation? That is, what is the geometric interpretation of the dot product?

Geometric Interpretation of the Dot Product

- Elementary math shows $\langle a, e_1 \rangle = a_x$ and $\langle a, \lambda \cdot e_1 \rangle = \lambda \cdot a_x$ for $\lambda \in \mathbb{R}$, where e_1 is the unit vector of the x-axis. Similarly, $\langle a, e_2 \rangle = a_y$ and $\langle a, \lambda \cdot e_2 \rangle = \lambda \cdot a_y$.
- The length s of the orthogonal projection of e_1 onto a equals the x-coordinate of a: Since $\|a\| = \|e_1\| = 1$, 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.
- 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 $\|a\| = 1$,

$$d = \langle a, b_x \cdot e_1 \rangle + \langle a, b_y \cdot e_2 \rangle = b_x \cdot a_x + b_y \cdot a_y = \langle a, b \rangle.$$





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 $g\colon \mathbb{R}^n \to \mathbb{R}$ 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 $\mathbf{A}$ be the matrix of the linear transformation g.
  - Then $\mathbf{A} \in M_{1\times n}$, i.e.,
$$\mathbf{A} = [a_{11} a_{12} \dots a_{1n}].$$
  - 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 $g(u) = \mathbf{A}u = \langle a, u \rangle$.
  - 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 $g\colon \mathbb{R}^3 \to \mathbb{R}$ 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 $g\colon \mathbb{R}^3 \to \mathbb{R}$ 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 124: 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

$$\det \begin{pmatrix} u_x & a_x & b_x \\ u_y & a_y & b_y \\ u_z & a_z & b_z \end{pmatrix} = g(u) = \begin{pmatrix} c_x & c_y & c_z \end{pmatrix} \cdot \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = \langle c, u \rangle.$$

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

$$c = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} = \begin{pmatrix} a_y \cdot b_z - a_z \cdot b_y \\ a_z \cdot b_x - a_x \cdot b_z \\ a_x \cdot b_y - a_y \cdot b_x \end{pmatrix} \overset{\text{Def. 142}}{=} a \times b.$$

- 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 and 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 $g(u) = V = \langle c, u \rangle = \langle a \times b, u \rangle$.
- We conclude that

$$A = \|a \times b\|,$$

  i.e., that the length of $a \times b$ equals the area of the parallelogram spanned by a, b.

Classification of Transformations

- Consider a mapping $g\colon \mathbb{R}^n \to \mathbb{R}^n$ and a distance metric $d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.
- 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 246 (*Isometry, Dt.: Isometrie*)

A mapping $g\colon \mathbb{R}^n \to \mathbb{R}^n$ is called an *isometry* if it maps pairs of points to points the same distance apart. That is, if

$$\forall (p, q \in \mathbb{R}^n) \;\; d(g(p), g(q)) = d(p,q).$$

- 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 247 (*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!

Classification of Transformations

Definition 248 (*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 249

The group of all isometries on $\mathbb{R}^n$ is given by composites of a translation and an orthogonal transformation.

Lemma 250

The group of all rigid motions on $\mathbb{R}^n$ is given by composites of a translation and a rotation.

Classification of Transformations

Lemma 251

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 252

With respect to an orthonormal basis of $\mathbb{R}^n$, an orthogonal transformation is invertible: If its matrix is $\mathbf{A}$ then the inverse transformation has matrix $\mathbf{A}^t$. Furthermore, $\det \mathbf{A} = \pm 1$.

Lemma 253

A 2×2 orthogonal matrix $\mathbf{A}$ is the matrix of a rotation about the origin if and only if $\det \mathbf{A} = 1$. If $\det \mathbf{A} = -1$ then it is the matrix of a reflection.

Lemma 254

A 3×3 orthogonal matrix $\mathbf{A}$ is the matrix of a rotation about a straight line through the origin if and only if $\det \mathbf{A} = 1$.

Classification of Transformations

Definition 255 (*Similarity mapping, Dt.: Ähnlichkeitsabbildung*)

A mapping g is called a *similarity mapping* if it preserves distance ratios and angles.

- E.g., two triangles which are similar have identical angles, and their sides are "in proportion".

Lemma 256

A distance-preserving transformation is a similarity mapping, i.e., it preserves angles.

Definition 257 (*Conformal, Dt.: winkeltreu*)

A mapping is called *conformal* if it preserves angles between directed curves.

Classification of Transformations

Definition 258 (*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 259

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.

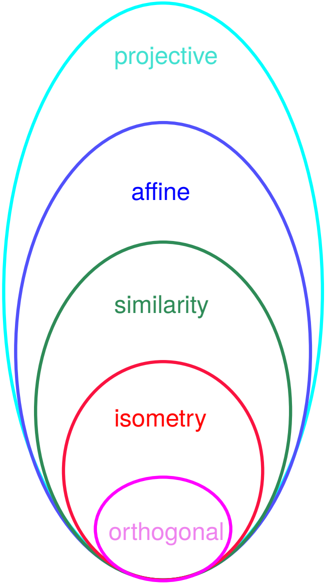
Lemma 260

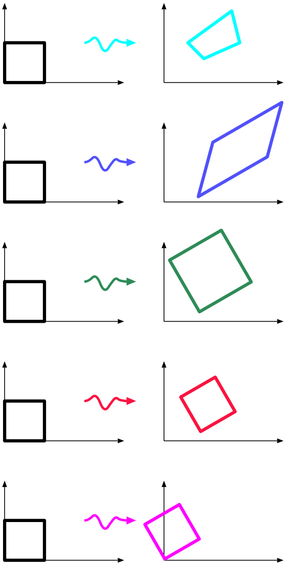
An affine transformation preserves ratios of lengths of intervals on any line.

Corollary 261

An affine transformation maps parallel lines to parallel lines.

Group Hierarchy of Transformations





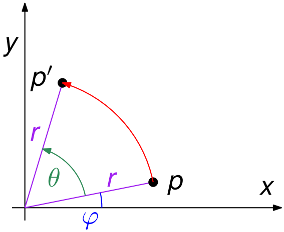
Rotation in $\mathbb{R}^2$

- Rotation of point p by θ about the origin yields point p'.

Polar coordinates: $p_x := r\cos\varphi, \quad p_y := r\sin\varphi.$

$$\begin{aligned} p'_x &= r\cos(\theta+\varphi) \\ &= r\cos\theta\cos\varphi - r\sin\theta\sin\varphi \\ &= p_x\cos\theta - p_y\sin\theta. \end{aligned}$$

$$\begin{aligned} p'_y &= r\sin(\theta+\varphi) \\ &= p_x\sin\theta + p_y\cos\theta. \end{aligned}$$



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 θ.
- This relation can also be expressed by means of a rotation matrix $\mathbf{Rot}(\theta)$:
$$\begin{pmatrix} p'_x \\ p'_y \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}}_{=:\mathbf{Rot}(\theta)} \cdot \begin{pmatrix} p_x \\ p_y \end{pmatrix};$$
that is
$$\mathbf{Rot}(\theta) := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Lemma 262

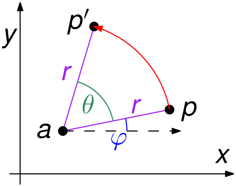
Rotation matrices are orthogonal: $\mathbf{Rot}(\theta)^{-1} = \mathbf{Rot}(\theta)^t$.

General Rotation in $\mathbb{R}^2$

- Rotation of point p by θ about point a, with $a := \begin{pmatrix} a_x \\ a_y \end{pmatrix}$, yields point p'.

$$p_x = a_x + r\cos\varphi \qquad \text{thus,} \quad r\cos\varphi = p_x - a_x$$
$$p_y = a_y + r\sin\varphi \qquad \text{thus,} \quad r\sin\varphi = p_y - a_y$$

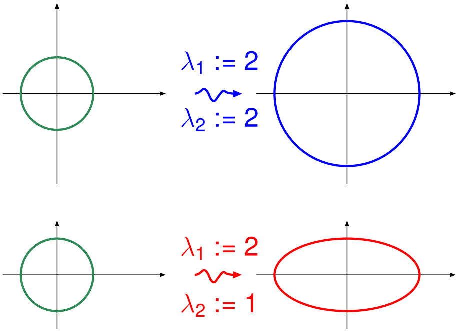
$$\begin{aligned} p'_x &= a_x + r\cos(\theta + \varphi) \\ &= a_x + r\cos\theta\cos\varphi - r\sin\theta\sin\varphi \\ &= a_x + (p_x - a_x)\cos\theta - (p_y - a_y)\sin\theta \\ p'_y &= a_y + (p_x - a_x)\sin\theta + (p_y - a_y)\cos\theta \end{aligned}$$



Stretching in $\mathbb{R}^2$

$$\begin{pmatrix} p'_x \\ p'_y \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}}_{=:\mathbf{S}(\lambda_1,\lambda_2)} \cdot \begin{pmatrix} p_x \\ p_y \end{pmatrix}$$

- If $\lambda_1 = \lambda_2$: (*uniform*) *scaling*;
- If $\lambda_1 \neq \lambda_2$: *non-uniform scaling* or *stretching*.



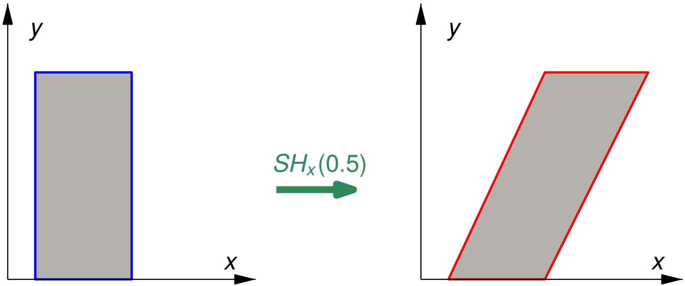
Shear Transformation in $\mathbb{R}^2$

- 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 $a \cdot y$.
- The corresponding transformation matrix is given by

$$\mathbf{SH}_x(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$



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 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 π.

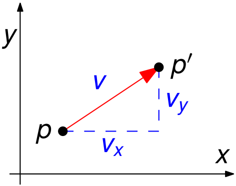
Translation in $\mathbb{R}^2$

- Translation: Move a point p along a vector v from its original location p to its new location p'.

$$p := \begin{pmatrix} p_x \\ p_y \end{pmatrix} \quad v := \begin{pmatrix} v_x \\ v_y \end{pmatrix} \quad p' := \begin{pmatrix} p'_x \\ p'_y \end{pmatrix}$$

$$p'_x = p_x + v_x, \quad p'_y = p_y + v_y, \quad p' = p + v$$

$$\begin{pmatrix} p'_x \\ p'_y \end{pmatrix} = \begin{pmatrix} p_x \\ p_y \end{pmatrix} + \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

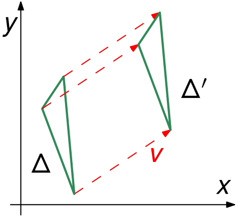


Translation of a Rigid Body

- Translate every point of Δ by v:

$$\Delta' = \{p + v \; : \; p \in \Delta\}.$$

- For polygons and polytopes it suffices to translate the vertices.



Translation as a Matrix Transformation

Question

What is the matrix of a translation?

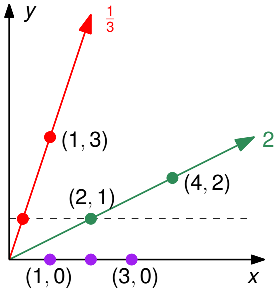
Answer

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

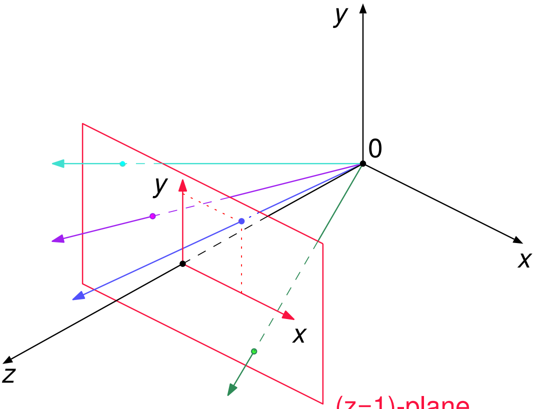
- A rational number $\frac{x}{y}$ is an equivalence class of appropriate pairs (x', y').



- $2 \simeq (2, 1), (4, 2), \ldots$
- $1/3 \simeq (1/3, 1), (1, 3), (2, 6), \ldots$
- 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 $\mathbb{R}^2$

- $\mathbb{R}^2$ 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 $w \neq 0$.
- Same for other points.
- 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 $\mathbb{R}^2$

- Homogeneous coordinates allow us to express translation, rotation and scaling in $\mathbb{R}^2$ by means of one 3×3 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

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix}.$$

Homogeneous Coordinates in $\mathbb{R}^2$

Definition 263 (*Homogeneous coordinates, Dt.: homogene Koordinaten*)

Homogeneous coordinates of $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ 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 $\begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathbb{R}^3$

  $$\Longleftrightarrow \quad x = \frac{u}{w} \text{ and } y = \frac{v}{w}.$$

- We will find it convenient to assume that $w = 1$.

Transformation Matrices Based on Homogeneous Coordinates for $\mathbb{R}^2$

Translation:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & v_x \\ 0 & 1 & v_y \\ 0 & 0 & 1 \end{pmatrix}}_{=:\mathbf{Trans}(v_x, v_y)} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

We get $\quad \mathbf{Trans}(v_x, v_y)^{-1} = \mathbf{Trans}(-v_x, -v_y)$.

Stretching:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=:\mathbf{S}(\lambda_1, \lambda_2)} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

We get $\quad \mathbf{S}(\lambda_1, \lambda_2)^{-1} = \mathbf{S}(\frac{1}{\lambda_1}, \frac{1}{\lambda_2})$.

Transformation Matrices Based on Homogeneous Coordinates for $\mathbb{R}^2$

Rotation:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} .= \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=:\mathbf{Rot}(\theta)} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

We get $\quad \mathbf{Rot}(\theta)^{-1} = \mathbf{Rot}(-\theta) = \mathbf{Rot}(\theta)^t.$

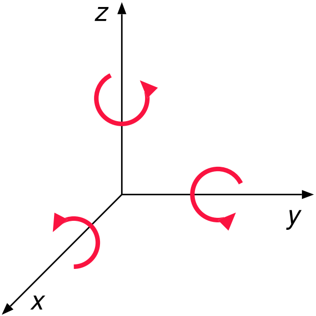
- Rotation involves either trigonometric functions or square roots.
- Power series may be used to approximate the terms of a rotation matrix for small values of θ.

Homogeneous Coordinates and Transformations in $\mathbb{R}^3$

- Homogeneous coordinates in $\mathbb{R}^3$:

$$(x, y, z, w) \simeq \left(\frac{x}{w}, \frac{y}{w}, \frac{z}{w}\right).$$

- 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 90° transforms one axis to another, obeying the cyclic order $x \to y \to z \to x$.



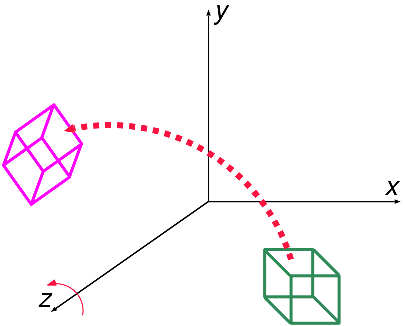
Rotation about z-Axis

- 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,

$$x' = x\cos\theta - y\sin\theta,$$

$$y' = x\sin\theta + y\cos\theta,$$

$$z' = z.$$



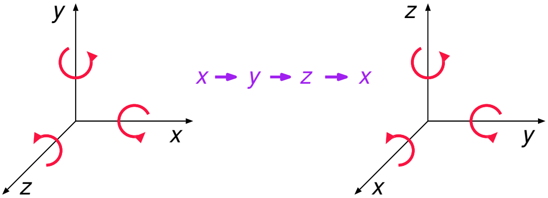
Rotation about x-Axis

- Rotation about the x-axis: Substitute $x \to y, y \to z, z \to x$ in the equations for the rotation about z.

$$y' = y\cos\theta - z\sin\theta,$$

$$z' = y\sin\theta + z\cos\theta,$$

$$x' = x.$$



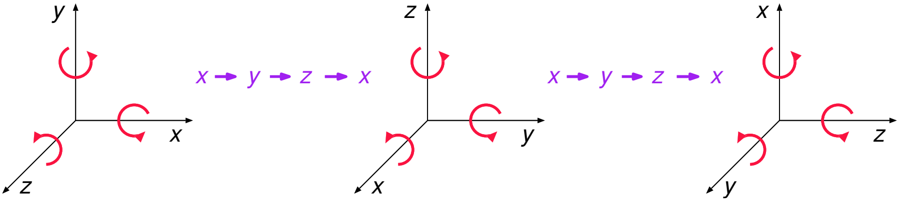
Rotation about y-Axis

- Similarly for a rotation about the y-axis: Substitute $x \to y, y \to z, z \to x$ in the previous equations.

$$z' = z\cos\theta - x\sin\theta,$$

$$x' = z\sin\theta + x\cos\theta,$$

$$y' = y.$$



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

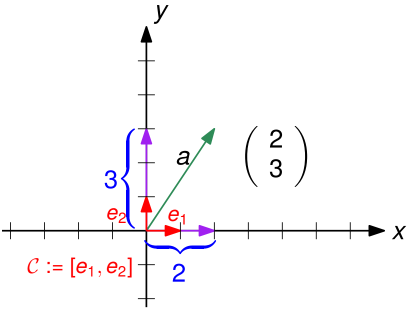
- Question: How can we find the axis of rotation (through the origin) if we only know the rotation matrix **T**?
- Answer: Since all points on the axis of rotation are invariant under the rotation, it suffices to look for a non-zero vector v such that

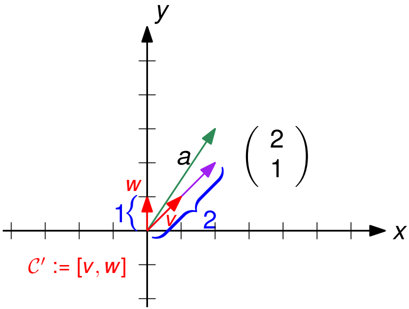
  $$\mathbf{T}v = v,$$

  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.

Transformation of Coordinate Systems

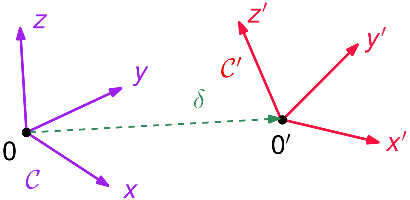
- Space has no intrinsic coordinate system!
- Basis vectors need not have unit length.
- Hence, a point will have different coordinates in different coordinate systems of the same vector space.
- E.g., $\mathcal{C} := [e_1, e_2]$ is not the only possible basis for $\mathbb{R}^2$: $\begin{pmatrix} 2 \\ 3 \end{pmatrix}_{[e_1, e_2]} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}_{[v, w]}$
- Our next task is to convert between different coordinate systems.





Transformation of Coordinate Systems

- So, what are the coordinates $p_{\mathcal{C}'} := \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\mathcal{C}'}$ of a point $p_{\mathcal{C}} := \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ relative to a new coordinate system $\mathcal{C}'$?

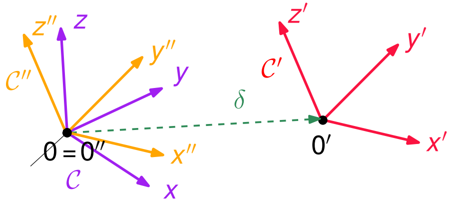


Transformation of Coordinate Systems

- We assume that the mapping from $\mathcal{C}$ to $\mathcal{C}'$ is an isometry.
- Consider an untranslated copy $\mathcal{C}''$ of $\mathcal{C}'$ whose axes vectors are identical but whose origin $0''$ is at the origin of $\mathcal{C}$. That is, $x'' \parallel x'$ and $y'' \parallel y'$ and $z'' \parallel z'$.
- We construct the matrix

$$\mathbf{T}_{\mathcal{C}} := \left(\begin{array}{c|c|c|c} e'_1 & e'_2 & e'_3 & \delta \\ \hline 0 & 0 & 0 & 1 \end{array} \right),$$

where e'_1 represents the unit vector of the x''-axis of $\mathcal{C}''$ in terms of $\mathcal{C}$. Of course, e'_1 is also the unit vector of the x'-axis of $\mathcal{C}'$. 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
$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \hline 1 \end{pmatrix} \xmapsto{\mathbf{T}_{\mathcal{C}}} \begin{pmatrix} e'_1 + \delta \\ \hline 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \hline 1 \end{pmatrix} \xmapsto{\mathbf{T}_{\mathcal{C}}} \begin{pmatrix} e'_2 + \delta \\ \hline 1 \end{pmatrix},$$
$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ \hline 1 \end{pmatrix} \xmapsto{\mathbf{T}_{\mathcal{C}}} \begin{pmatrix} e'_3 + \delta \\ \hline 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \hline 1 \end{pmatrix} \xmapsto{\mathbf{T}_{\mathcal{C}}} \begin{pmatrix} \delta \\ \hline 1 \end{pmatrix},$$
that is
$$\begin{pmatrix} x' \\ y' \\ z' \\ \hline 1 \end{pmatrix} \xmapsto{\mathbf{T}_{\mathcal{C}}} \begin{pmatrix} x'e'_1 + y'e'_2 + z'e'_3 + \delta \\ \hline 1 \end{pmatrix} =: \begin{pmatrix} x \\ y \\ z \\ \hline 1 \end{pmatrix}_{\mathcal{C}}.$$
- We understand that the coordinates of a point specified relative to $\mathcal{C}'$ are converted by $\mathbf{T}_{\mathcal{C}}$ to coordinates relative to $\mathcal{C}$:

Theorem 264

With $\mathbf{T}_{\mathcal{C}}$ as defined on the previous slide, we get
$$p_{\mathcal{C}} = \mathbf{T}_{\mathcal{C}} \cdot p_{\mathcal{C}'} \qquad \text{and} \qquad p_{\mathcal{C}'} = \mathbf{T}_{\mathcal{C}}^{-1} \cdot p_{\mathcal{C}}.$$

Inverse Transformation

- If $\mathbf{T}$ is the matrix of an isometry then, by Lemma 249,

$$\mathbf{T} = \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & \\ \hline 0 & 0 & 0 & 1 \end{array}\right) \cdot \left(\begin{array}{ccc|c} & & & 0 \\ & \mathbf{R} & & 0 \\ & & & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right),$$

where $\mathbf{R}$ is an orthogonal matrix, and v describes the translation.

- Since $(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$, we get

$$\mathbf{T}^{-1} = \left(\begin{array}{ccc|c} & & & 0 \\ & \mathbf{R}^{-1} & & 0 \\ & & & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right) \cdot \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 0 & -v \\ 0 & 0 & 1 & \\ \hline 0 & 0 & 0 & 1 \end{array}\right).$$

- Since $\mathbf{R}$ is orthogonal, we have $\mathbf{R}^{-1} = \mathbf{R}^t$ and get

$$\mathbf{T}^{-1} = \left(\begin{array}{ccc|c} & & & 0 \\ & \mathbf{R}^{t} & & 0 \\ & & & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right) \cdot \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 0 & -v \\ 0 & 0 & 1 & \\ \hline 0 & 0 & 0 & 1 \end{array}\right).$$

Inverse Transformation

Theorem 265

If $[n, o, a]$ is orthogonal then we get

$$\mathbf{T}^{-1} = \begin{pmatrix} n_x & n_y & n_z & -\langle v, n\rangle \\ o_x & o_y & o_z & -\langle v, o\rangle \\ a_x & a_y & a_z & -\langle v, a\rangle \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

for

$$\mathbf{T} := \begin{pmatrix} n_x & o_x & a_x & v_x \\ n_y & o_y & a_y & v_y \\ n_z & o_z & a_z & v_z \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- Recall that the matrix of a general affine transformation is not orthogonal!

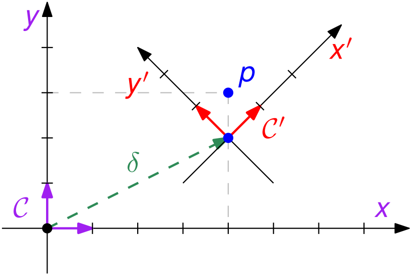
Sample Coordinate System Transformation

- For the scenario shown below we get

$$\mathbf{T} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 4 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and, thus,} \quad \mathbf{T}^{-1} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & -3\sqrt{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \sqrt{2} \\ 0 & 0 & 1 \end{pmatrix}.$$

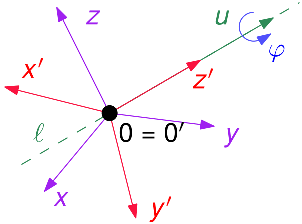
- Hence,

$$\mathbf{T}^{-1} \cdot p_{\mathcal{C}} = \mathbf{T}^{-1} \cdot \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 1 \end{pmatrix} = p_{\mathcal{C}'} \quad \text{and} \quad \mathbf{T} \cdot p_{\mathcal{C}'} = \mathbf{T} \cdot \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 1 \end{pmatrix} = p_{\mathcal{C}}.$$



Rotation About a General Axis

- What is the matrix of the rotation about a line ℓ (through the origin) with direction vector u by an angle φ?



- We set up a new frame $\mathcal{C}' = [e_1', e_2', e_3']$ 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

$$\mathbf{T} := \left(\begin{array}{ccc|c} e_1' & e_2' & e_3' & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right).$$

Rotation About a General Axis

- We know that $\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \mathbf{T}^{-1} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$.
- Thus, we get the following decomposition for $\mathbf{Rot}(u, \varphi)$:

$$\mathbf{Rot}(u,\varphi) = \underbrace{\mathbf{T}}_{\substack{\text{from } \mathcal{C}' \\ \text{back to } \mathcal{C}}} \cdot \underbrace{\begin{pmatrix} \cos\varphi & -\sin\varphi & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\text{rotation about } z'\text{-axis}} \cdot \underbrace{\mathbf{T}^{-1}}_{\substack{\text{from } \mathcal{C} \\ \text{to } \mathcal{C}'}}$$

- Simple algebraic operations yield

$$\mathbf{Rot}(u,\varphi) = \begin{pmatrix} u_x u_x \, vers\,\varphi + \cos\varphi & u_y u_x \, vers\,\varphi - u_z \sin\varphi & u_z u_x \, vers\,\varphi + u_y \sin\varphi & 0 \\ u_x u_y \, vers\,\varphi + u_z \sin\varphi & u_y u_y \, vers\,\varphi + \cos\varphi & u_z u_y \, vers\,\varphi - u_x \sin\varphi & 0 \\ u_x u_z \, vers\,\varphi - u_y \sin\varphi & u_y u_z \, vers\,\varphi + u_x \sin\varphi & u_z u_z \, vers\,\varphi + \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $vers\,\varphi := 1 - \cos\varphi$.

Rotation About a General Axis

- Given an (orthogonal) rotation matrix **T**, how can we find an axis u through the origin and an angle φ such that $\mathbf{Rot}(u, \varphi) = \mathbf{T}$?

$$\mathbf{Rot}(u, \varphi) \stackrel{?}{=} \mathbf{T} := \begin{pmatrix} n_x & o_x & a_x & 0 \\ n_y & o_y & a_y & 0 \\ n_z & o_z & a_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- Some calculations yield

$$\tan \varphi = \frac{\sqrt{(o_z - a_y)^2 + (a_x - n_z)^2 + (n_y - o_x)^2}}{n_x + o_y + a_z - 1},$$

which defines φ within $[0, \pi]$.

Rotation About a General Axis

- Furthermore,

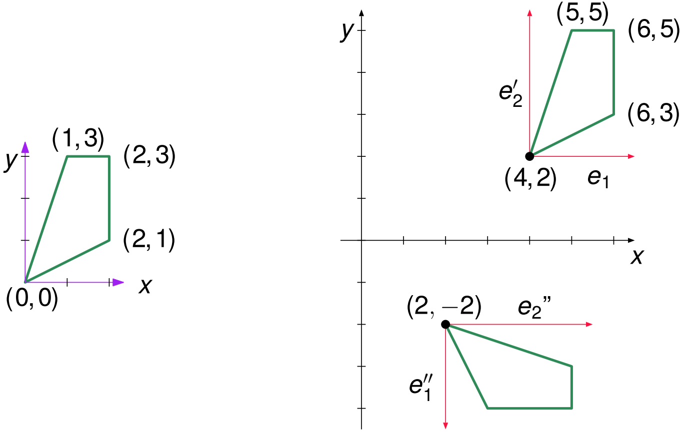
$$u_x = \operatorname{sign}(o_z - a_y)\sqrt{\frac{n_x - \cos\varphi}{1 - \cos\varphi}},$$

$$u_y = \operatorname{sign}(a_x - n_z)\sqrt{\frac{o_y - \cos\varphi}{1 - \cos\varphi}},$$

$$u_z = \operatorname{sign}(n_y - o_x)\sqrt{\frac{a_z - \cos\varphi}{1 - \cos\varphi}}.$$

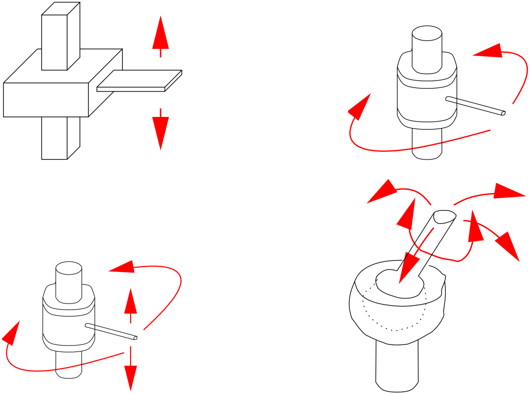
Local Coordinate Systems

- 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.



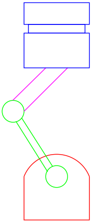
Kinematics

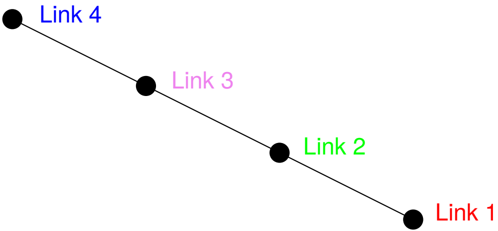
- We consider an articulated mechanism that consists of rigid links connected by joints.
- Every joint connects exactly two links, and describes the motion of one link relative to the other link.
- The most important joints are prismatic and rotatory joints.



Kinematic Chain

- A mechanism can be represented as a graph, a so-called *kinematic chain*, where
  - the links form the nodes, and
  - the joints form the edges.

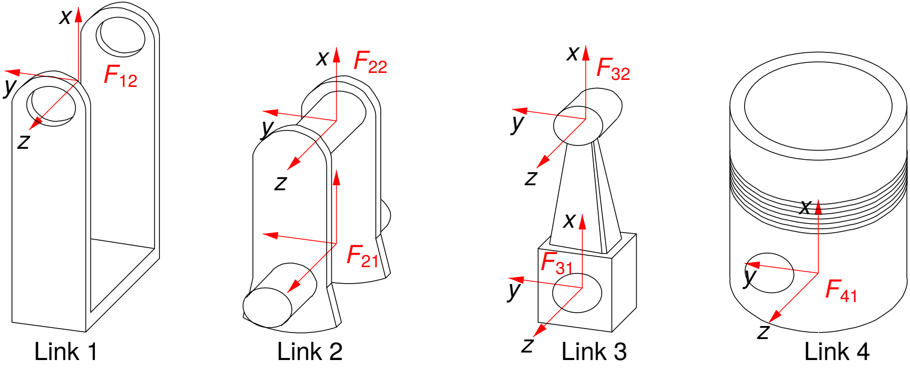




- A mechanism is called an *open kinematic chain* if this graph has no cycles; *closed kinematic chain*, otherwise.
- Depending on how detailed a human is modeled, a human skeleton represents either an open or a closed kinematic chain.

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 ${}^{i-1}_{i}\mathbf{A}$ to express a point of $F_{i,2}$ in terms of $F_{i-1,2}$.
- *A-Matrix*:

$$
\begin{aligned}
{}^{i-1}_{i}\mathbf{A} &:= \mathbf{Rot}(z,\theta)\cdot\mathbf{Trans}(0,0,d)\cdot\mathbf{Trans}(a,0,0)\cdot\mathbf{Rot}(x,\alpha)\\
&= \begin{pmatrix} \cos\theta & -\sin\theta\cos\alpha & \sin\theta\sin\alpha & a\cos\theta \\ \sin\theta & \cos\theta\cos\alpha & -\cos\theta\sin\alpha & a\sin\theta \\ 0 & \sin\alpha & \cos\alpha & d \\ 0 & 0 & 0 & 1 \end{pmatrix},
\end{aligned}
$$

$$
\text{where}\quad \left.\begin{array}{lll} a & \ldots & \text{link length,} \\ \alpha & \ldots & \text{link twist,} \\ d & \ldots & \text{link offset,} \\ \theta & \ldots & \text{link angle,} \end{array}\right\}\ \textit{Denavit-Hartenberg parameters.}
$$

Forward and Inverse Kinematics

Forward Kinematics:

- Given: joint vector.
- Compute: Frame **T** of the end-effector relative to the base frame.
- Solution:

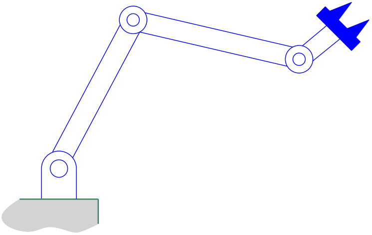
$$\mathbf{T} = {}^{0}_{1}\mathbf{A} \cdot {}^{1}_{2}\mathbf{A} \cdot \ldots \cdot {}^{n-1}_{n}\mathbf{A}.$$

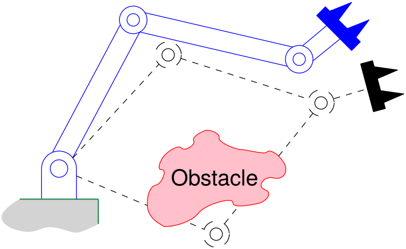
Inverse Kinematics:

- Given: Frame **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!





Geometric Interpretation of Eigenvectors

- Recall Def. 127: A vector $v \in \mathbb{R}^n$ is an eigenvector of the $n \times n$ matrix $\mathbf{A}$ if

$$\mathbf{A}v = \lambda v \quad \text{and} \quad v \neq 0.$$

- The vector $\mathbf{A}u$ is some vector of $\mathbb{R}^n$ obtained by applying a linear transformation g, whose matrix equals $\mathbf{A}$, to u.
- Hence, $v \neq 0$ is an eigenvector of $\mathbf{A}$ 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 $\mathbf{A}$ 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 $\mathbf{A}$.
- 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^o about the origin in $\mathbb{R}^2$.

Rotation Group

Definition 266 (*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 267 (*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 268

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 3×3 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 $SO(3)$.

Euler's Rotation Theorem

Lemma 269 (*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:

$$\mathbf{R} := \mathbf{R}_n \cdot \ldots \cdot \mathbf{R}_2 \cdot \mathbf{R}_1$$

We will now show that there exists a vector v such that $\mathbf{R}v = v$. We see that the vector v must be an eigenvector of the matrix $\mathbf{R}$ with eigenvalue $\lambda = 1$. Since this requires $(\mathbf{R} - \mathbf{I})v = 0$, we know that $\det(\mathbf{R} - \mathbf{I}) = 0$ is required.

Euler's Rotation Theorem

Proof of Lem. 269 (cont'd) : We use

$$\det(-(\mathbf{R}-\mathbf{I})) = -\det(\mathbf{R}-\mathbf{I}) \qquad \text{and} \qquad \det(\mathbf{R}^{-1}) = 1$$

and obtain

$$\begin{aligned}\det(\mathbf{R}-\mathbf{I}) &= \det\left((\mathbf{R}-\mathbf{I})^t\right) = \det\left(\mathbf{R}^t-\mathbf{I}\right) = \det\left(\mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{R}\right)\\ &= \det\left(\mathbf{R}^{-1}(\mathbf{I}-\mathbf{R})\right) = \det\left(\mathbf{R}^{-1}\right)\det(-(\mathbf{R}-\mathbf{I})) = -\det(\mathbf{R}-\mathbf{I}).\end{aligned}$$

Thus, $\det(\mathbf{R}-\mathbf{I}) = 0$. Hence, there is at least one non-zero vector v such that $\mathbf{R}v = v$. 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. □

Theorem 270 (*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.

Quaternions and Rotation

Lemma 271

Let $\mathcal{Q}$ be a quaternion that is not zero and $\mathcal{P}$ be a pure quaternion. Then $\mathcal{P}' := \mathcal{Q}\mathcal{P}\mathcal{Q}^{-1}$ is a pure quaternion, too.

- This quaternion operation maps the set of all pure quaternions onto itself.
- This set forms a 3-dimensional sub-space of the space of all quaternions.

Theorem 272

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 2φ. Now consider the pure quaternions $\mathcal{P} := (0, p)$ and $\mathcal{P}' := (0, p')$. We have

$$\mathcal{P}' = \mathcal{Q}\mathcal{P}\mathcal{Q}^{-1} \qquad \text{for } \mathcal{Q} := (\cos\varphi, u\sin\varphi).$$

Lemma 273

Consider the setting of Theorem 272 and let $s := \cos\varphi$, $v := u\sin\varphi$. 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 (through the origin) in $\mathbb{R}^3$ corresponds to a unit quaternion.
- Conversely, every unit quaternion represents a rotation about an axis in $\mathbb{R}^3$.

Theorem 274

There is a one-to-one correspondence between unit quaternions and rotations about axes (through the origin) in $\mathbb{R}^3$.

Lemma 275

The inverse quaternion models the opposite rotation.

Proof: We have

$$\mathcal{Q}^{-1}(\mathcal{Q}\mathcal{P}\mathcal{Q}^{-1})\mathcal{Q} = \mathcal{P}. \qquad \square$$

- Geometric interpretation of this fact: Since $\mathcal{Q}^{-1} = (s, -u)$ for a unit quaternion $\mathcal{Q} := (s, u)$, the inverse of $\mathcal{Q}$ 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

Lemma 276

If $\mathcal{Q}$ is a unit quaternion then $\mathcal{Q}$ and $-\mathcal{Q}$ represent the same rotation.

Sketch of Proof : A rotation about the axis u by the angle 2φ equals a rotation about the (inversely oriented) axis $-u$ by the angle -2φ. □

Lemma 277

The square $\mathcal{Q}^2$ of a unit quaternion $\mathcal{Q}$ is a rotation by twice the angle about the same axis.

Lemma 278

The orthogonal matrix that corresponds to a rotation by the unit quaternion $\mathcal{Q} = (s, (a, b, c))$ is given by

$$\begin{pmatrix} s^2 + a^2 - b^2 - c^2 & 2ab - 2sc & 2ac + 2sb \\ 2ab + 2sc & s^2 - a^2 + b^2 - c^2 & 2bc - 2sa \\ 2ac - 2sb & 2bc + 2sa & s^2 - a^2 - b^2 + c^2 \end{pmatrix}.$$

Quaternions and Rotation: SLERP

- Suppose that we are given two unit quaternions $\mathcal{Q}_0, \mathcal{Q}_1$ 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 $\mathcal{Q}_0 := (s_0, (a_0, b_0, c_0))$ and $\mathcal{Q}_1 := (s_1, (a_1, b_1, c_1))$:

Theorem 279 (*Shoemake 1985*)

Consider two unit quaternions $\mathcal{Q}_0 := (s_0, (a_0, b_0, c_0))$ and $\mathcal{Q}_1 := (s_1, (a_1, b_1, c_1))$. Let Θ such that

$$\cos\Theta = s_0 \cdot s_1 + a_0 \cdot a_1 + b_0 \cdot b_1 + c_0 \cdot c_1.$$

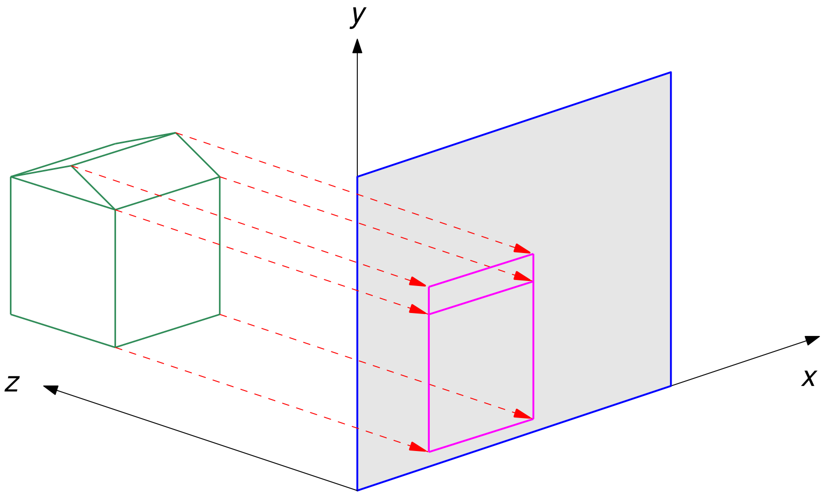
Then, for $t \in [0, 1]$,

$$Slerp(\mathcal{Q}_0, \mathcal{Q}_1, t) := \frac{1}{\sin\Theta}\big(\sin((1-t)\Theta)\mathcal{Q}_0 + \sin(t\Theta)\mathcal{Q}_1\big)$$

corresponds to the interpolated quaternion at time $t \in [0, 1]$. The Slerp interpolation function achieves constant angular velocity.

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

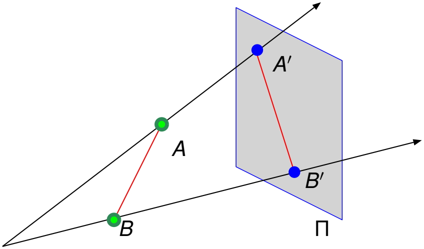
- Plan from Mesopotamia, $\approx$ 2000 BCE.
- Early Greeks: *Agatharchus* ($\approx$ 500 BCE), *Apollonius* of Perga ($\approx$ 262 BCE till $\approx$ 190 BCE) studied projections of quadrics.
- Romans: *Vitruvius* wrote *De Architectura*, published specifications of plan and elevation drawings, and 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.

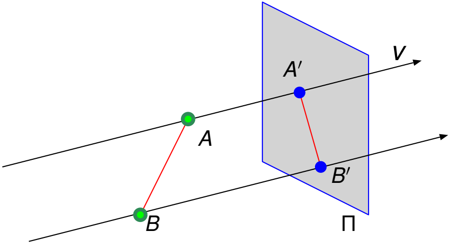
  "A painting is the intersection of a visual pyramid at a given distance, with a fixed center and a definite position of light, represented by art with lines and colors on a given surface."

Geometric Projections

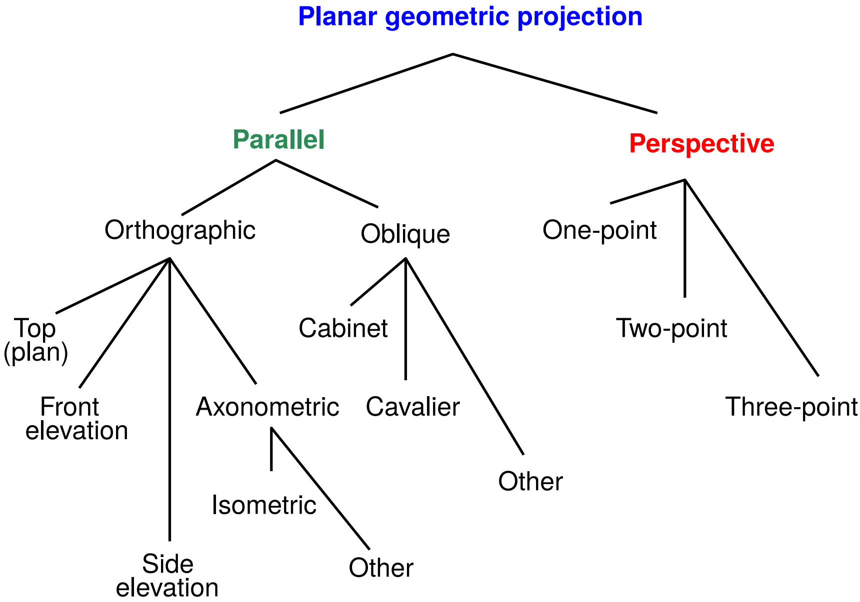
- *Projectors*: Rays emanating from the center of projection and passing through points of the object.
- *Projection*: Intersection of projectors with *projection plane* Π.
- Non-geometric projections used in cartography. E.g., Mercator projection.

- *Perspective:*
  - Center of projection is at a finite distance from Π.
  - *Perspective foreshortening*.
- *Parallel:*
  - Center of projection is at ∞.
  - Defined by a direction v.



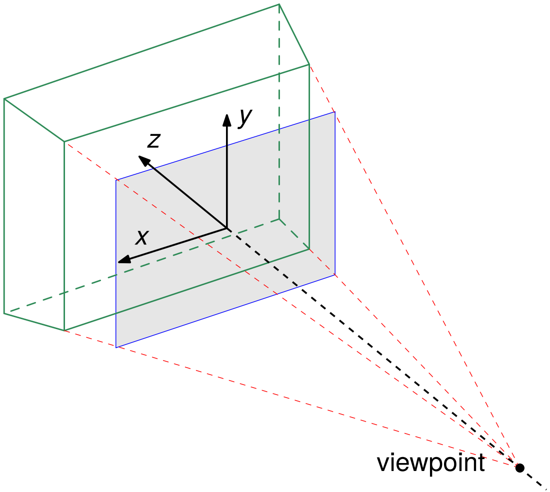


Geometric Projections: Different Types



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 Π is the xy-plane.

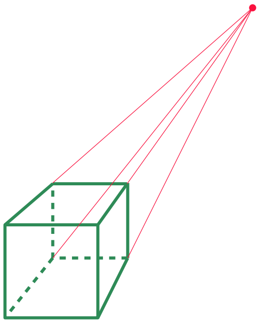


Perspective Projection

- 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 mutually parallel lines that are not parallel to the image plane converge.
- Since buildings tend to have one to three sets of parallel lines, we get one-point perspective, two-point perspective, or three-point perspective.

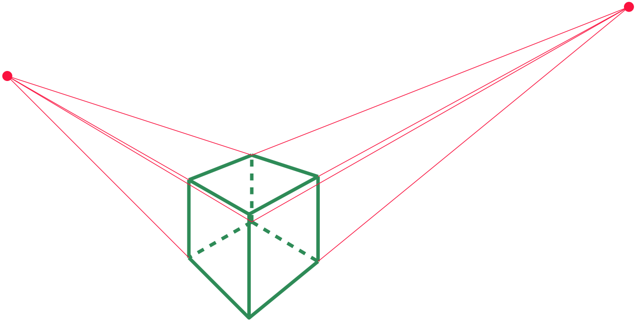
One Vanishing Point

- Π parallel to two principal axes of the cube: one vanishing point.



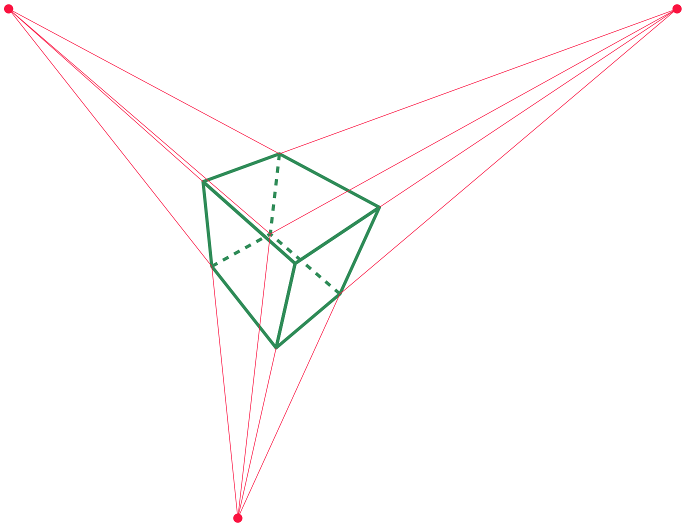
Two Vanishing Points

- Π is parallel to only one principal axis of the cube: two vanishing points.



Three Vanishing Points

- Π is not parallel to any principal axis of the cube: three vanishing points.



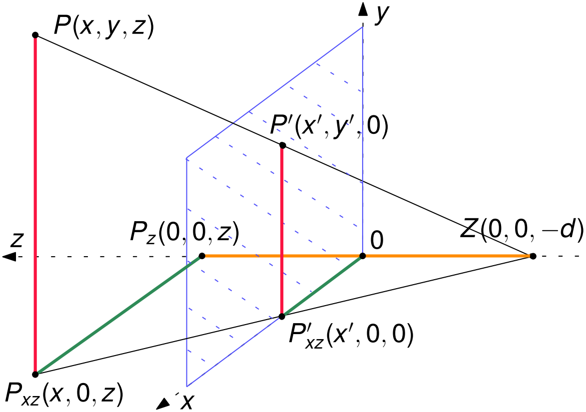
Mathematics of Perspective Projection

- 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,

$$y' = \frac{d \cdot y}{z + d}.$$



Matrix of a Perspective Projection

- Let $\mathbf{P} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{d} & 1 \end{pmatrix}$.
- We get

$$\mathbf{P} \cdot p = \begin{pmatrix} p_x \\ p_y \\ 0 \\ \frac{p_z+d}{d} \end{pmatrix} \equiv \begin{pmatrix} \frac{d \cdot p_x}{p_z+d} \\ \frac{d \cdot p_y}{p_z+d} \\ 0 \\ 1 \end{pmatrix} =: \begin{pmatrix} p'_x \\ p'_y \\ 0 \\ 1 \end{pmatrix}.$$

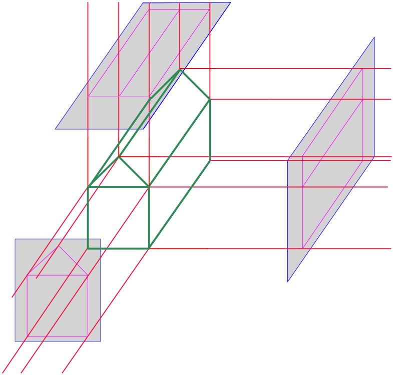
- Apply transformation of coordinate system if the projection plane differs from $z = 0$, or if the eye point is not at $(0, 0, -d)$.

Parallel Projection: Orthographic

- *Orthographic:* Projectors are perpendicular to the projection plane.

$$\rightarrow \mathbf{P}_{xy} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- *Front, top, side views:* Projectors parallel to one of the principal axes.



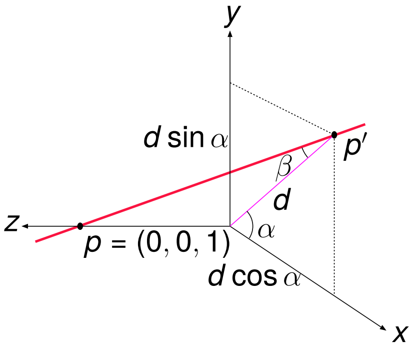
Parallel Projection: Oblique

- *Oblique:* Projectors not perpendicular to the projection plane.
- With $d := \cot\beta$ we get

$$x' = x + z \cdot d\cos\alpha,$$
$$y' = y + z \cdot d\sin\alpha,$$
$$z' = 0.$$

- Thus,

$$\mathbf{P} := \begin{pmatrix} 1 & 0 & d\cos\alpha & 0 \\ 0 & 1 & d\sin\alpha & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

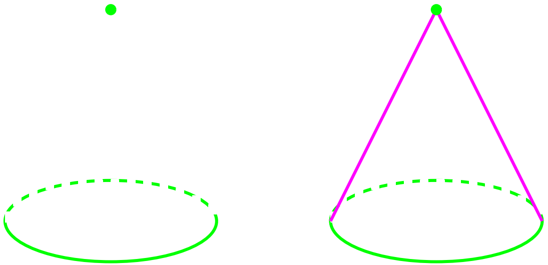


Special Oblique Projections

- *Cavalier projection:*
  - Angle β between projectors and projection plane is 45^o; 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 β between projectors and projection plane is $\mathtt{tan}^{-1} 2 \approx 63.4^o$; 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.



- 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.
- 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

$$\mathbf{N} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}; \quad \mathbf{N} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}; \quad \mathbf{N} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix};$$

  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 **O** is the matrix of an orthogonal projection, and **P** is the matrix of the corresponding perspective projection.
- **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.

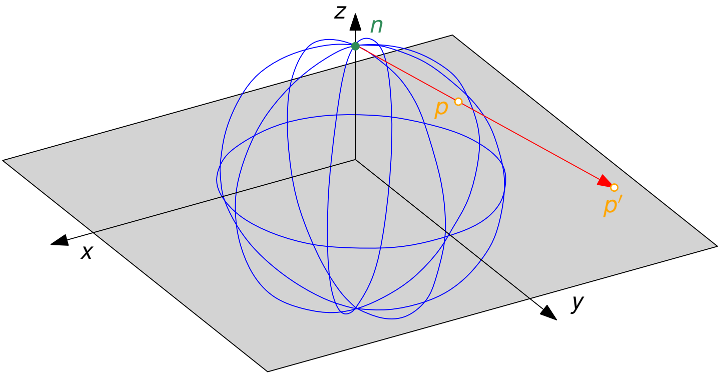
- We can modify **N** such that all z-coordinates are scaled to lie between 0 and 1.

Stereographic Projection

- A *stereographic projection* maps a sphere onto a plane.
- Default setting: Mapping of S^2 onto the xy-plane $z = 0$, with the north pole $n := (0, 0, 1)$ serving as projection point.

Lemma 280

For any point p on S^2 other than n, the line through n and p intersects $z = 0$ in exactly one point p'.

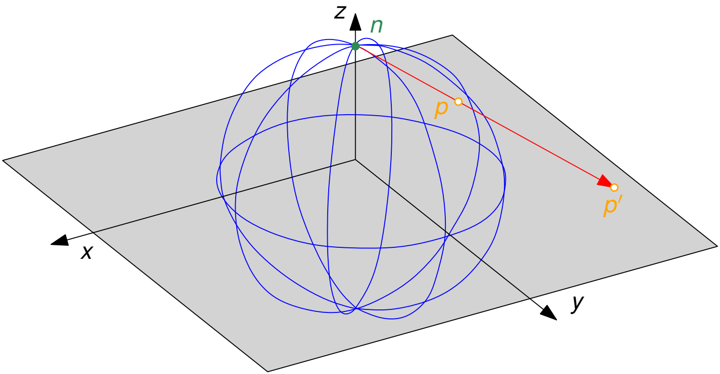


Stereographic Projection

- A *stereographic projection* maps a sphere onto a plane.
- Default setting: Mapping of S^2 onto the xy-plane $z = 0$, with the north pole $n := (0, 0, 1)$ serving as projection point.

Definition 281 (*Stereographic projection*)

The *stereographic projection*, $\mathrm{SP}(p)$, of a point p on $S^2 \setminus \{n\}$ is the point p' uniquely determined according to Lemma 280.



Stereographic Projection

Lemma 282

Let p on $S^2 \setminus \{n\}$ and $p' := \mathrm{SP}(p)$. If $p := (x, y, z)$ and $p' := (a, b)$ then we get the following relations:

$$(a, b) = \left(\frac{x}{1-z}, \frac{y}{1-z}\right),$$

$$(x, y, z) = \left(\frac{2a}{1+a^2+b^2}, \frac{2b}{1+a^2+b^2}, \frac{-1+a^2+b^2}{1+a^2+b^2}\right).$$

Corollary 283

The point p has rational coordinates if and only if p' has rational coordinates.

Lemma 284

The stereographic projection and its inverse are continuous maps.

Sketch of Proof : Both maps are given as fractions of polynomials where the denominator is never zero. □

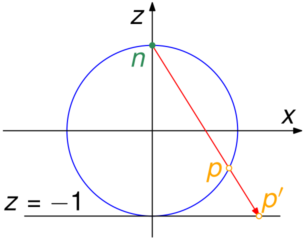
Stereographic Projection

Lemma 285

The stereographic projection establishes a bijection between $S^2 \setminus \{n\}$ and the plane $z = 0$: It maps the south pole to $(0, 0)$, the equator to the unit circle, the southern hemisphere to the region inside the circle, and the northern hemisphere to the region outside the circle.

Lemma 286

The stereographic projection is conformal but neither isometric nor area-preserving.



Other conventions ...

... include a mapping to $z = -1$.

- The intercept theorem implies that this scales the image by a factor of two.

Extended Stereographic Projection

- Recall that the north pole n of S^2 would be mapped to a *point at infinity* of the plane.
- We extend the stereographic projection by defining $\mathrm{SP}(n) := \infty$ and get a bijection between S^2 and $\mathbb{R}^2 \cup \{\infty\}$.
- By identifying $\mathbb{R}^2$ with $\mathbb{C}$, this concept can be generalized to a bijection between S^2 and $\mathbb{C}_\infty$.
- Topologically speaking, since S^2 is compact, the sphere S^2 is homeomorphic to a one-point compactification of the (complex) plane.

Lemma 287

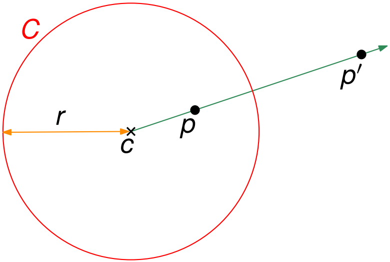
The extended stereographic projection $\mathrm{SP}\colon S^2 \to \mathbb{R}^2 \cup \{\infty\}$, or $\mathrm{SP}\colon S^2 \to \mathbb{C}_\infty$, maps circles to circles and lines.

Inversion

Definition 288 (*Circle inversion*)

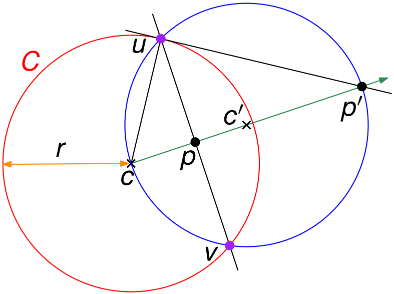
Consider a circle C in the (Euclidean) plane that is centered at c and has radius $r \in \mathbb{R}^+$. The *inversion* of a point p with respect to C, denoted by $\mathrm{Inv}_{c,r}(p)$, is the point p' on the ray from c through p such that $\|c - p\| \cdot \|c - p'\| = r^2$.

- Obviously, $\mathrm{Inv}_{c,r}(c)$ is not defined.
- We turn $\mathrm{Inv}_{c,r}$ into a total function by mapping c to a point at infinity, and vice versa.



Inversion: Construction by Compass and Ruler

- The right triangles $\Delta(c, p', u)$ and $\Delta(c, p, u)$ are similar.
- The center c' of the blue circle is the midpoint of c and p'.



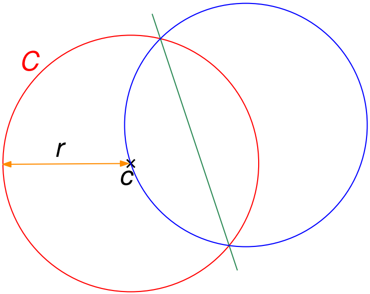
Inversion: Properties

Lemma 289

The inversion $\mathrm{Inv}_{c,r}$ maps a line that does not pass through c to a circle that passes through c. A line that passes through c is mapped onto itself.

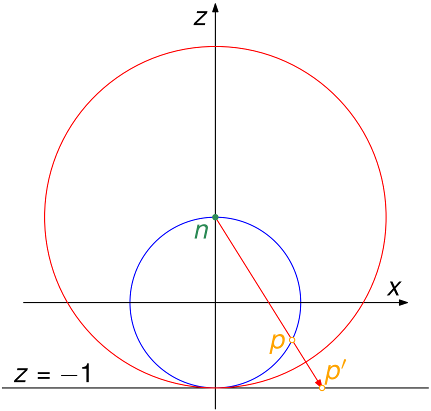
Lemma 290

The inversion $\mathrm{Inv}_{c,r}$ maps a circle that does not pass through c to a circle that does not pass through c.



Lemma 291

The stereographic projection onto the plane $z = -1$ equals an inversion with respect to a sphere centered at n with radius $r := 2$.



Floating-Point Arithmetic and Numerical Mathematics

Floating-Point Arithmetic

- Computers employ floating-point (fp) arithmetic to perform real arithmetic.
- No matter how many bits are used, fp-arithmetic represents a number by a fixed-length binary mantissa and an exponent of fixed size.

Chuck Allison

Floating-point numbers are not real numbers [...]. Real numbers have infinite precision and are therefore continuous and nonlossy; floating-point numbers have limited precision, so they are finite, and they resemble “badly behaved” integers, because they are not evenly spaced throughout their range.

- 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.
- The IEEE 754 standard for fp-arithmetic knows four different rounding modes. The first mode is the default; the others are called directed roundings.

  Round to Nearest
  Round towards 0
  Round towards $+\infty$
  Round towards $-\infty$

Floating-Point Errors

- Hence, there are two sources of error for fp-computations: input error and round-off error.

  Input error: It arises from reading/assigning a value to an fp variable.

  - 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!
  - What do we get if we assign $2^{24} + 1 = 16777217$ to a 32-bit float? We get 16777216!

  Round-off error: It arises from rounding results of fp-computations during an algorithm.

  - E.g., $\sqrt{2}$ cannot be represented exactly since $\sqrt{2}$ is an irrational number.

- 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:

$$1/3 = 0.333333333333333314829616256247390992939472198486328125000$$
$$1/10 = 0.100000000000000005551115123125782702118158340454101562500$$

Machine Precision

- The round-off error is bounded in terms of the *machine precision*, ε, which is the smallest value satisfying

$$|fp(a \circ b) - (a \circ b)| \leq \varepsilon |a \circ b|$$

  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, ε 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 ...
- Random errors tend to cancel on a large scale, and accumulate on small scale.

Floating-Point Arithmetic and Compilers

- Accumulation: Adding 0.001 for 1 000 000 times need not yield exactly 1000.

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,

$$\sum_{i=1}^{10000000} 0.001 = 1000.0000000000009095 \qquad \text{with } \texttt{gcc -O2 -mfpmath=387},$$

$$\sum_{i=1}^{10000000} 0.001 = 999.9999999832650701 \qquad \text{with } \texttt{gcc -O0 -mfpmath=387}.$$

- Newer chips also support the SSE/SSE2 instruction set, and the default option `-mfpmath=sse` avoids this problem for x86-64 compilers.

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 ${}^{56}/_{100} \cdot 10^{-14}$ 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.
- 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 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:

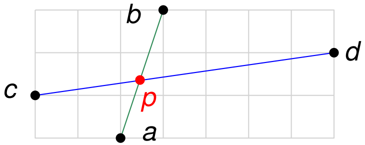
  $$(a =_\varepsilon b) \quad :\Longleftrightarrow \quad (|a - b| \leq \varepsilon),$$

  for some positive value of ε.
- 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: Failure of Basic Mathematical Implications

- Suppose that we are given two line segments $\overline{ab}$ and $\overline{cd}$ such that

$$c_x < a_x < b_x < d_x \qquad a_y < c_y < d_y < b_y.$$



- It is easy to see that the two line segments intersect, without a or b lying on $\overline{cd}$ and without c or d lying on $\overline{ab}$. Furthermore, the line segments cannot overlap. Hence, the two line segments intersect in a point.
- Let $p := \overline{ab} \cap \overline{cd}$. Are the following inequalities guaranteed to be true?

$$a_x < p_x < b_x \qquad a_y < p_y < b_y \qquad c_x < p_x < d_x \qquad c_y < p_y < d_y$$

- Yes in theory, no on an fp-arithmetic!

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 $x_0 := 10$:

$h := 10^{0}$:	$f'(10) \approx 331.0000000$	$h := 10^{-1}$:	$f'(10) \approx 303.0099999$
$h := 10^{-2}$:	$f'(10) \approx 300.3000999$	$h := 10^{-3}$:	$f'(10) \approx 300.0300009$
$h := 10^{-4}$:	$f'(10) \approx 300.0030000$	$h := 10^{-5}$:	$f'(10) \approx 300.0002999$
$h := 10^{-6}$:	$f'(10) \approx 300.0000298$	$h := 10^{-7}$:	$f'(10) \approx 300.0000003$
$h := 10^{-8}$:	$f'(10) \approx 300.0000219$	$h := 10^{-9}$:	$f'(10) \approx 300.0000106$
$h := 10^{-10}$:	$f'(10) \approx 300.0002379$	$h := 10^{-11}$:	$f'(10) \approx 299.9854586$
$h := 10^{-12}$:	$f'(10) \approx 300.1332515$	$h := 10^{-13}$:	$f'(10) \approx 298.9963832$
$h := 10^{-14}$:	$f'(10) \approx 318.3231456$	$h := 10^{-15}$:	$f'(10) \approx 568.4341886$
$h := 10^{-16}$:	$f'(10) \approx 0.000000000$	$h := 10^{-17}$:	$f'(10) \approx 0.000000000$

- 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 significantly greater than the machine precision!

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

$$\begin{aligned} x + 2y &= 3 \\ 0.48x + 0.99y &= 1.47 \end{aligned}$$

  has the exact solution $x = 1, y = 1$, while the system

$$\begin{aligned} x + 2y &= 3 \\ 0.49x + 0.99y &= 1.47 \end{aligned}$$

  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 may dramatically affect the solutions of an equation: *ill-conditioned* or *ill-posed*!

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 2×2 linear system

$$\begin{aligned} 1.2969x + 0.8648y &= 0.8642 \\ 0.2161x + 0.1441y &= 0.1440 \end{aligned} \qquad \text{that is,} \qquad \mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

- The exact solution is $x = 2$ and $y = -2$.
- But we get close-to-zero residuals also for other pairs of x and y:

$$\begin{aligned} x_2 &= 2.001557851 \\ y_2 &= -2.002336236 \end{aligned} \qquad \left\| \mathbf{A}\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\| \approx 10^{-10}$$

$$\begin{aligned} x_1 &= 0.9911 \\ y_1 &= -0.4870 \end{aligned} \qquad \left\| \mathbf{A}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\| \approx 10^{-8}$$

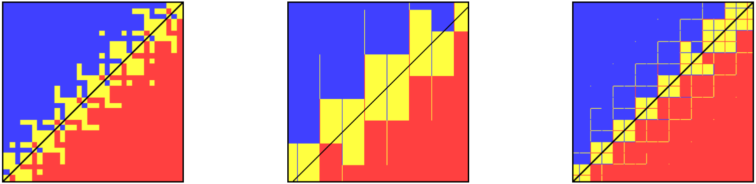
$$\begin{aligned} x_3 &= -0.000004626 \\ y_3 &= 0.999312976 \end{aligned} \qquad \left\| \mathbf{A}\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\| \approx 10^{-9}$$

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 $u := 2^{-53}$:

$$\text{sign}\det\begin{pmatrix} 1 & p_x + x \cdot u & p_y + y \cdot u \\ 1 & q_x & q_y \\ 1 & r_x & r_y \end{pmatrix} \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} 0 \ ?$$

- The resulting 256×256 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!

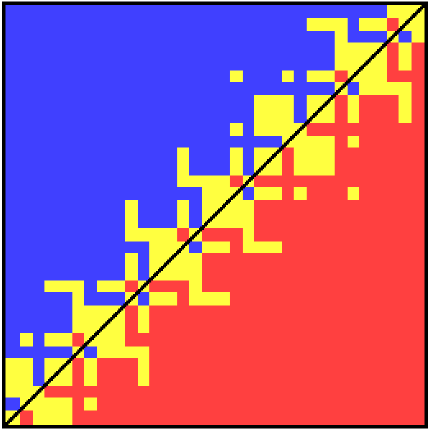


[Image credit: `www.mpi-inf.mpg.de/~kettner/proj/NonRobust/`]

Sample Robustness Problem: Incorrect Orientation Predicate

- [Kettner et alii 2006]: A yellow (red, blue) pixel indicates collinear (negative, positive, resp.) orientation.

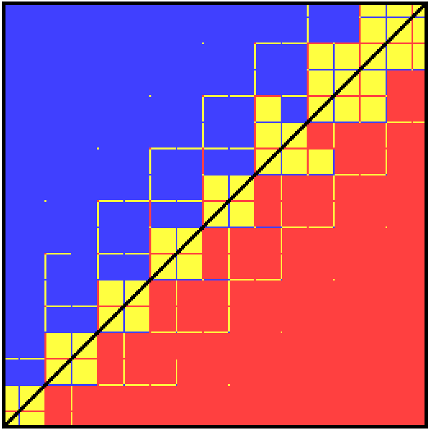
$$p := \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \quad q := \begin{pmatrix} 12 \\ 12 \end{pmatrix} \quad r := \begin{pmatrix} 24 \\ 24 \end{pmatrix}$$



Sample Robustness Problem: Incorrect Orientation Predicate

- [Kettner et alii 2006]: A yellow (red, blue) pixel indicates collinear (negative, positive, resp.) orientation.

$$p := \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \quad q := \begin{pmatrix} 8.8000000000000007 \\ 8.8000000000000007 \end{pmatrix} \quad r := \begin{pmatrix} 12.1 \\ 12.1 \end{pmatrix}$$



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 Overflow Error

- Ariane Flight V88 was the failed maiden flight of the Ariane 5 rocket, vehicle number 501, on 04-June-1996.
- The operating code for the Ariane 4 rocket was reused in the Ariane 5. However, Ariane 5 was faster. . ..
- This triggered a bug in an arithmetic routine inside the rocket's flight computer: The error was in the code that converts a 64-bit floating-point number to a 16-bit signed integer. The faster engines caused the 64-bit numbers to be larger in the Ariane 5 than in the Ariane 4, triggering an overflow condition.
- To make the situation worse, the default IEEE 754 exception-handling policy ("presubstitution") had not been used.
- As a consequence, the overflow resulted in a hardware exception, causing both flight computers to crash: First the backup flight computer crashed, followed 0.05 seconds later by a crash of the primary flight computer.
- As a result of both computers being off, the rocket's primary processor overpowered the rocket's engines, which caused the rocket to disintegrate 40 seconds after launch, and finally self-destruction via its automated flight termination system.
- That failure resulted in a loss of more than €290 million and in a delay of the Ariane program by a year.

Butterfly Effect and Chaos Theory

- In 1961, the mathematician and meteorologist Lorenz noted that very minor changes in the initial conditions (due to numerical rounding) caused repeated runs of his weather model to produce strikingly different results.
- He wanted to rerun a numerical computer model to redo a weather prediction from the middle of the previous run.
- He entered the initial condition 0.506 from the previous result instead of entering the full-precision value 0.506127.
- The result was a completely different weather prediction.
- Lorenz:

  Chaos: When the present determines the future, but the approximate present does not approximately determine the future.
- See `https://upload.wikimedia.org/wikipedia/commons/4/44/Double_pendulum_simultaneous_realisations.ogv` for six slow-motion videos of the same double pendulum (built with Lego). For each recording, the double pendulum was excited in virtually the same manner.

Quote taken from “The Art of Computer Programming” (D.E. Knuth)

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.

Floating-Point Comparisons and Precision Thresholds

- The gap between the theory of the reals and floating-point reality 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.
- All floating-point computations need to be consistent.
- In particular, make sure that different calls of the same function with the “same” input will yield exactly the same output. E.g., when computing 3×3 determinants to determine the orientation of three points p, q, r, the following identities are a must even on a floating-point arithmetic:

$$\begin{aligned} \det(p,q,r) &= \det(q,r,p) = \det(r,p,q) \\ &= -\det(q,p,r) = -\det(p,r,q) = -\det(r,q,p). \end{aligned}$$

- 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.
- Epsilon-based comparisons need to be relative to the absolute values of the numbers to be compared, or the input has to be scaled (by performing shifts!) to fit into the unit square/cube prior to the actual computation.
- Use iterations as back-up for analytical solutions to equations. If at all possible, use methods that bracket the solution sought!

Improving the Reliability of FP-Calculations

- 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:

$$
\begin{aligned}
1 + \frac{1}{2} + \ldots + \frac{1}{1000000} &\approx 14.392726722864989\\
\frac{1}{1000000} + \ldots + \frac{1}{2} + 1 &\approx 14.392726722865723\\
\text{Mathematica:} &\approx 14.392726722865723631 38\ldots
\end{aligned}
$$

Improving the Reliability of FP-Calculations: Quadratic Equations

- 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!
- 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

$$\Delta := -\frac{1}{2}(b + \operatorname{sign}(b)\sqrt{b^2 - 4ac}).$$

Then the roots are obtained more reliably as

$$x_1 := \frac{\Delta}{a} \quad \text{and} \quad x_2 := \frac{c}{\Delta}.$$ (This is a consquence 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

$$x_1 \approx -10000.000000000000000000000000000,$$
$$x_2 \approx -0.0000000000000000000000000000000.$$

- The refined approach yields

$$x_1 \approx -10000.000000000000000000000000000,$$
$$x_2 \approx -0.0000000000001000000000000000030.$$

- According to Mathematica, the true solution is

$$x_1 \approx -9999.999999999999900000000000000,$$
$$x_2 \approx -0.0000000000001000000000000000010.$$

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 $\overline{x} \in \mathbb{R}$ for which $f(\overline{x}) = 0$.
- 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, $\lim_{k\to\infty} x_k = \overline{x}$.
- 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 $\overline{x}$?

Definition 292 (*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 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 $f\colon \mathbb{R} \to \mathbb{R}$, 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 $\overline{x}$ 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:

$$\text{if} \quad \text{sign}(f(\frac{a+b}{2})) \begin{cases} = & 0 & \text{then} & \overline{x} := \frac{a+b}{2}, \text{ stop;} \\ = & \text{sign}(f(a)) & \text{then} & a := \frac{a+b}{2}; \\ = & \text{sign}(f(b)) & \text{then} & b := \frac{a+b}{2}. \end{cases}$$

- 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 $[a, b]$, only one root will be found.
- 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 $\mathrm{sign}(f(c))$, 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 $\overline{x}$ of f, and that x_0 is close to $\overline{x}$.
- The Newton-Raphson method is based on the approximation of a function f by the straight-line tangent at $(x_k, f(x_k))$:

$$y = f(x_k) + f'(x_k)(x - x_k).$$

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 $\overline{x}$ 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.
- 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.
- 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.

Iterative Algorithms for Solving Linear Equations

- Recall that finding the exact solution x of the system of linear equations $\mathbf{A}x = b$ requires $O(n^3)$ time for an $n \times n$ matrix $\mathbf{A}$.
- A direct (and exact) solution turns out to be a waste of time if n goes into the thousands or millions and if $\mathbf{A}$ 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

$$\mathbf{A}\Delta x = \mathbf{A}x - \mathbf{A}x' = b - \mathbf{A}x'.$$

- Interpreting this equation as basis for an iterative formula $x^{(k+1)} = x^{(k)} + \Delta x$ yields

$$\mathbf{A}(x^{(k+1)} - x^{(k)}) = b - \mathbf{A}x^{(k)}.$$

- So far, we would have gained little, as we would still have to solve for $x^{(k+1)}$ ...
- Bold idea: Replace $\mathbf{A}$ on the left-hand side of this equation by an easily invertible matrix $\mathbf{B}$ that is “close to” $\mathbf{A}$.

Iterative Algorithms for Solving Linear Equations

- We get

$$\mathbf{B}(x^{(k+1)} - x^{(k)}) = b - \mathbf{A}x^{(k)},$$

or

$$\mathbf{B}x^{(k+1)} = b - (\mathbf{A} - \mathbf{B})x^{(k)}.$$

- One can formulate conditions under which the solution obtained by this iterative scheme is guaranteed to converge to the exact solution of $\mathbf{A}x = b$.

- 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

$$\mathbf{B}x^{(k+1)} = b - (\mathbf{A} - \mathbf{B})x^{(k)}.$$

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 **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.
- 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 $\mathbf{A}$ is strictly diagonally dominant.
- Tends to converge faster than Jacobi iteration, but is significantly more difficult to parallelize.

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

$$I = \int_a^b f(x)\,dx = F(b) - F(a)\,, \quad \text{where } F'(x) = f(x)\,.$$

- However, there are many functions that cannot be integrated analytically. Thus, methods for approximating the integral through *quadrature rules* of the form

$$\hat{I} = \sum_{i=1}^{n} \omega_i f(x_i)$$

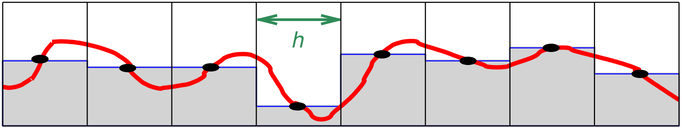
  have been devised, which is essentially a weighted sum of samples of the function f at various points x_i using weights ω_i.
- The many different quadrature rules can be distinguished by their sampling patterns and weights.

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:

$$\begin{aligned} \hat{I} &= h \sum_{i=1}^{n} f(a + (i - \frac{1}{2})h) \\ &= h \left[f(a + \frac{h}{2}) + f(a + \frac{3h}{2}) + \cdots + f(b - \frac{h}{2}) \right] . \end{aligned}$$

- 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 $[a, b]$.

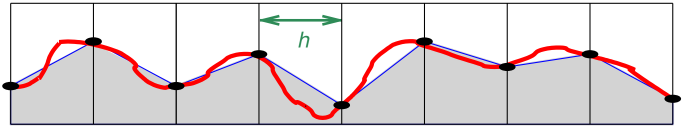


Trapezoidal Rule for Numerical Integration

- 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.

$$\begin{aligned}\hat{I} &= \sum_{i=1}^{n} \frac{h}{2}\left[f(a+(i-1)h)+f(a+ih)\right] \\ &= h\left[\frac{1}{2}f(a)+f(a+h)+f(a+2h)+\cdots+f(b-h)+\frac{1}{2}f(b)\right].\end{aligned}$$

- 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:

$$\hat{I} = h\left[\frac{1}{3}f(a) + \frac{4}{3}f(a+h) + \frac{2}{3}f(a+2h) + \frac{4}{3}f(a+3h) + \frac{2}{3}f(a+4h) + \right.$$

$$\left. \cdots + \frac{4}{3}f(b-h) + \frac{1}{3}f(b)\right] .$$

- Simpson's rule is exact for polynomial functions up to cubics. The error can be bounded by $O(n^{-4})$.
- It converges very quickly if 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 vary rare event in computer graphics!

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 ω_{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 N 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 then 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.)

The End!

I hope that you enjoyed this course, and I wish you all the best for your future studies.

UNIVERSITÄT SALZBURG
Computational Geometry and Applications Lab