

Kapitel P

Vektorräume mit Skalarprodukt

*Unsere Allergrößten, wie Archimedes, Newton, Gauß,
haben stets Theorie und Anwendungen gleichmäßig umfasst.*

Felix Klein (1849–1925)

Inhalt dieses Kapitels P

1 Skalarprodukte

- Skalarprodukte über \mathbb{R} , euklidische Vektorräume
- Skalarprodukte über \mathbb{C} , unitäre Vektorräume
- Erste Anwendungen, von Pythagoras zu Fourier

2 Orthonormalisierung

- Gram–Schmidt–Verfahren und QR–Zerlegung
- Bestapproximation und Methode der kleinsten Quadrate
- Näherungslösung eines überbestimmten Gleichungssystems

3 Orthogonale und unitäre Endomorphismen

- Orthogonale und unitäre Endomorphismen
- Orthogonale und unitäre Gruppen
- Geometrie des dreidimensionalen Raumes

Motivation und Überblick

Wir untersuchen in diesem Kapitel reelle und komplexe Vektorräume mit einem Skalarprodukt. Damit messen wir Winkel, Längen und Abstände, wir betreiben also **Geometrie** im traditionellen, wörtlichen Sinne.

Diese fundamentale Idee kennen Sie bereits aus der Schulgeometrie, im Folgenden ist sie ebenso grundlegend für die **Lineare Algebra und Analytische Geometrie**. Darüber hinaus spielt sie (verallgemeinert) die zentrale Rolle in der Untersuchung von gekrümmten Flächen und allgemein in der Riemannschen Geometrie auf Mannigfaltigkeiten, bis hin zur speziellen und allgemeinen Relativitätstheorie.

Skalarprodukte begegnen Ihnen ebenso in der **Analysis**, zunächst als euklidisches Skalarprodukt bei der Untersuchung der metrischen und topologischen Eigenschaften des euklidischen Raumes \mathbb{R}^n , später dann ebenso in der Fourier–Theorie bei der Zerlegung periodischer Funktionen in harmonische Grundschwingungen, sowie in zahlreichen weiteren Anwendungen.

Motivation und Überblick

Ebenso wie die Analysis nutzt auch die **Numerik** Skalarprodukte, Normen und Metriken zur Messung von Abständen und von Fehlern, zur Kontrolle von Näherungen und zur Definition der Konvergenz.

Auch die **Methode der kleinsten Quadrate** beruht auf der Technik von Skalarprodukten und wird überall eingesetzt, wo fehlerbehaftete Daten verarbeitet werden, insbesondere linearisiert oder ähnlich geglättet.

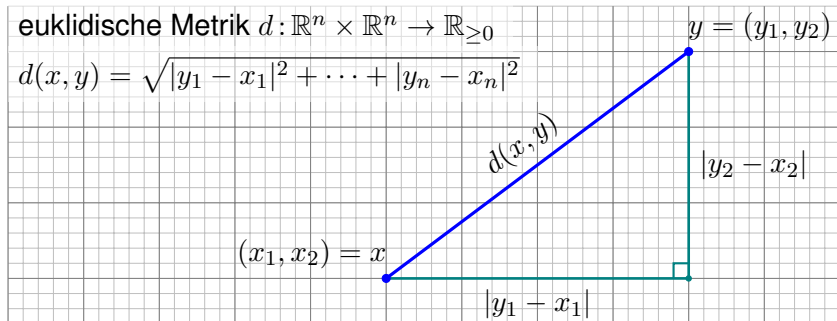
In der **Stochastik** verhält sich die Co/Varianz reeller Zufallsvariablen wie ein Skalarprodukt, und die lineare Regression nutzt die Methode der kleinsten Quadrate zur Untersuchung linearer Zusammenhänge.

In der **Physik** schließlich beruht die Quantenmechanik auf dem Modell, dass die Zustände eines Systems beschrieben werden durch Vektoren in einem Hilbert-Raum, also einem \mathbb{C} -Vektorraum mit Skalarprodukt.

Heisenbergs Unschärferelation entspricht (in geeigneter Übersetzung) der Cauchy-Schwarz-Ungleichung für Skalarprodukte von Vektoren.

Die euklidische Metrik: Satz des Pythagoras im \mathbb{R}^n

Wir betrachten den affinen Raum \mathbb{R}^n über den reellen Zahlen \mathbb{R} .



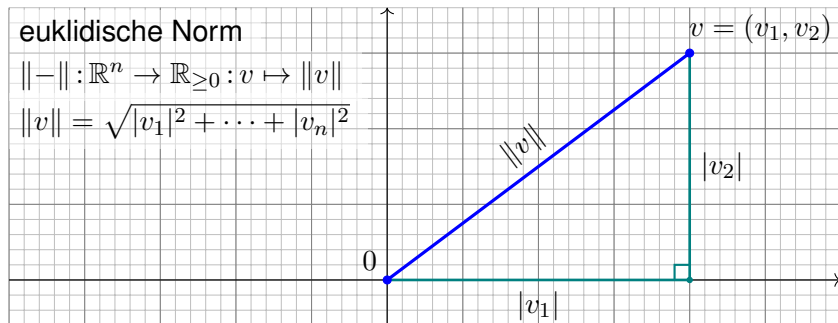
Die **euklidische Metrik** $d(x, y)$ misst den Abstand vom Punkt x zum Punkt y , also die Länge des Vektors $y - x$. Dank Pythagoras gilt:

$$d(x, y) = \sqrt{|y_1 - x_1|^2 + \dots + |y_n - x_n|^2}$$

Die obige Skizze zeigt dies für $n = 2$. Übung: Skizzieren und begründen Sie dies für $n = 3$. Der allgemeine Fall $n \in \mathbb{N}$ gelingt dann per Induktion.

Die euklidische Norm: Satz des Pythagoras im \mathbb{R}^n

Wir betrachten den Vektorraum \mathbb{R}^n über den reellen Zahlen \mathbb{R} .



Die **euklidische Norm** $\|v\|$ misst die Länge des Vektors $v \in \mathbb{R}^n$ als den Abstand vom Ursprung 0 zum Punkt v . Dank Pythagoras gilt:

$$\|v\| = \sqrt{|v_1|^2 + \dots + |v_n|^2}$$

Das Normquadrat $\|v\|^2 = v_1^2 + \dots + v_n^2$ ist eine quadratische Form. Die zugehörige Bilinearform ist das **euklidische Skalarprodukt**.

Das euklidische Skalarprodukt auf dem Raum \mathbb{R}^n

Beispiel P1A: das euklidische Skalarprodukt auf dem Raum \mathbb{R}^n

Auf $V = \mathbb{R}^n$ über \mathbb{R} definieren wir das **euklidische Skalarprodukt**

$$\langle - | - \rangle : V \times V \rightarrow \mathbb{R} : (u, v) \mapsto \langle u | v \rangle := u_1 v_1 + \cdots + u_n v_n.$$

Für alle Vektoren $u, v, w \in V$ und Skalare $\lambda, \mu \in \mathbb{R}$ gilt:

S0: Positivität, $\langle u | u \rangle \geq 0 = \langle 0 | 0 \rangle$

S1: positive Definitheit, $\langle u | u \rangle > 0$ für $u \neq 0$

S2: Symmetrie, $\langle v | u \rangle = \langle u | v \rangle$

S3: Linearität rechts, $\langle u | \lambda v + \mu w \rangle = \lambda \langle u | v \rangle + \mu \langle u | w \rangle$

Aus (S2) und (S3) folgt Linearität auch in der ersten Variablen:

S4: Linearität links, $\langle \lambda u + \mu v | w \rangle = \lambda \langle u | w \rangle + \mu \langle v | w \rangle$

Aus dem euklidischen Skalarprodukt erhalten wir die **euklidische Norm**

$\| - \| : V \rightarrow \mathbb{R}_{\geq 0} : v \mapsto \|v\| = \sqrt{\langle v | v \rangle}$, und aus dieser Norm wiederum die **euklidische Metrik** $d : V \times V \rightarrow \mathbb{R}_{\geq 0} : (x, y) \mapsto d(x, y) = \|y - x\|$.

Das euklidische Skalarprodukt auf dem Raum \mathbb{R}^n

Somit ist das euklidische Skalarprodukt $\langle - | - \rangle$ auf $V = \mathbb{R}^n$ über \mathbb{R} eine **symmetrische Bilinearform** (S2,3) und **positiv definit** (S0,1). Dies sind die wesentlichen Eigenschaften des Skalarprodukts!

Aufgabe: Rechnen Sie die hier gemachten Aussagen nach.

Lösung: Symmetrie (S2) und Bilinearität (S3,4) sind klar.

(S0) Für jeden Vektor $u \in \mathbb{R}^n$ gilt $\langle u | u \rangle = u_1^2 + \dots + u_n^2 \geq 0$.

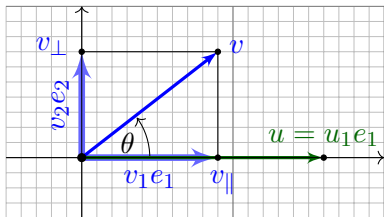
(S1) Im Falle $u \neq 0$ gilt $u_i \neq 0$ für mindestens ein $i \in \{1, \dots, n\}$.

Daraus folgt sofort die strikte Ungleichung $\langle u | u \rangle \geq u_i^2 > 0$.

Bemerkung: Aus (S1) folgt (S0), denn $\langle 0 | 0 \rangle = 0$ dank Linearität (S3). Die Formulierung der Eigenschaften (S0–3) ist daher etwas redundant; ich möchte damit die positive Definitheit (S1) besonders hervorheben.

Gilt in späteren Anwendungen statt positiver Definitheit (S1) nur die schwächere Eigenschaft (S0), so heißt $\langle - | - \rangle$ **positiv semidefinit**. Daher ist es sinnvoll, die Eigenschaft (S1) separat zu formulieren.

Geometrische Interpretation des euklidischen Skalarprodukts



Aus der Schule kennen Sie die geometrische Interpretation:

$$\langle u | v \rangle = \|u\| \cdot \|v\| \cdot \cos \angle(u, v)$$

Diese Formel gilt offensichtlich für $u = u_1 e_1$ und $v = v_1 e_1 + v_2 e_2$ dank $v_1 = \|v\| \cos \theta$, $v_2 = \|v\| \sin \theta$.

😊 Insbesondere folgt daraus die Ungleichung $|\langle u | v \rangle| \leq \|u\| \cdot \|v\|$.

Wir nutzen hierzu unsere geometrische Anschauung der Ebene. Das ist hilfreich zur Vorstellung und willkommen als Motivation.

⚠ Die Anschauung ist in der Mathematik meist eine gute Stütze, doch sie allein ist erfahrungsgemäß kein tragfähiges Fundament.

Wir werden für diese geometrischen Begriffe im Folgenden eine solide und allgemeine Grundlage erarbeiten, die sich nicht auf Anschauung oder Vorwissen beruft. Im Gegenteil werden wir so die geometrische Anschauung und die Anfänge der Trigonometrie begründen.

Geometrische Interpretation des euklidischen Skalarprodukts

Zu der obigen Interpretation $\langle u | v \rangle = \|u\| \cdot \|v\| \cdot \cos \angle(u, v)$ können Sie zu Recht einwenden, dass dies nur in dieser sehr speziellen Lage gilt.

😊 Die gute Nachricht ist jedoch: Wir können zu u und v unsere Basis e_1, e_2, \dots immer so wählen, dass genau diese spezielle Lage entsteht, also $u = u_1 e_1$ und $v = v_1 e_1 + v_2 e_2$ gilt. Das vereinfacht, wie gesehen!

Die geometrische Idee lässt sich auch ohne Koordinaten beschreiben: Bezüglich u zerlegen wir den Vektor $v = v_{\parallel} + v_{\perp}$ in seinen tangentialen Anteil v_{\parallel} parallel zu u und seinen normalen Anteil v_{\perp} senkrecht zu u .

😊 Mit dem Skalarprodukt gelingt dies bequem und explizit wie folgt:

$$v_{\parallel} = u \frac{\langle u | v \rangle}{\langle u | u \rangle} \quad \text{und} \quad v_{\perp} = v - v_{\parallel}, \quad \text{somit} \quad \langle u | v_{\perp} \rangle = 0.$$

Das ist die Grundidee des Gram–Schmidt–Verfahrens (Satz P2A), das aus jeder Basis b_1, \dots, b_n eine Orthonormalbasis konstruiert.

😊 Der obige Spezialfall erweist sich damit als allgemein: Wir wählen eine Orthonormalbasis e_1, e_2, \dots mit $u = u_1 e_1$ und $v = v_1 e_1 + v_2 e_2$.

Geometrische Interpretation des euklidischen Skalarprodukts

Als leuchtendes Beispiel haben wir eingangs den Vektorraum $V = \mathbb{R}^n$ über \mathbb{R} betrachtet. Hierauf haben wir das **euklidische Skalarprodukt**. Es heißt auch das **Standardskalarprodukt** auf dem Raum $V = \mathbb{R}^n$ oder das **kanonische Skalarprodukt**, weil es das Modellbeispiel eines Skalarprodukts ist. (Es gibt darüber hinaus viele weitere.)

Das so gebildete Paar $(\mathbb{R}^n, \langle - | - \rangle)$ nennen wir den **n -dimensionalen euklidischen Vektorraum** oder auch kurz den **euklidischen Raum**: Mit diesen Daten können wir Winkel, Längen und Abstände messen. Er ist das grundlegende Modell für die **euklidische Geometrie** (altgr. γεωμετρία, 'Erdmessung', 'Landmessung').

Speziell in niedriger Dimension $n = 2$ und $n = 3$ haben wir hierzu eine geometrische Anschauung und (mit etwas Glück schon in der Schule) eine mathematisch-physikalische Vorerfahrung. Dies wollen wir nutzen und zu einer tragfähigen Theorie und Rechenmethoden ausbauen.

Geometrische Interpretation des euklidischen Skalarprodukts

Daneben gibt es weitere nützliche Skalarprodukte auf dem Raum \mathbb{R}^n sowie auf anderen \mathbb{R} -Vektorräumen V , auch unendlich-dimensional.

Damit können wir auch in diesen allgemeineren Räumen Winkel, Längen und Abstände messen, also Geometrie betreiben wie im \mathbb{R}^n .

Die Erfahrung zeigt, dass die oben gesammelten Eigenschaften (S1–3) die wesentliche Grundlage für all unsere weiteren Rechnungen sind.

Diese Eigenschaften erheben wir daher nun zur Definition P1B und erklären so, was wir allgemein unter einem Skalarprodukt verstehen.

Das ist ein kühner Schritt der Verallgemeinerung, doch erweist sich als effizient und klärend: Sie schärft den Blick für das Wesentliche.

Abstraktion strukturiert und vereinfacht: Eine allgemeine Tatsache ist oft leichter zu verstehen und zu erklären als ihre zahlreichen Spezialfälle.

Skalarprodukt: allgemeine Definition

Definition P1B: Skalarprodukt über \mathbb{R}

Ein **Skalarprodukt** auf einem \mathbb{R} -Vektorraum V ist eine positiv definite, symmetrische Bilinearform $\langle - | - \rangle : V \times V \rightarrow \mathbb{R}$, erfüllt also (S1–3).

Die zugehörige **Norm** ist dann $\| - \| : V \rightarrow \mathbb{R}_{\geq 0} : v \mapsto \|v\| = \sqrt{\langle v | v \rangle}$.

◆ Beispiel P1A: die Produktsumme auf \mathbb{R}^n

Auf dem \mathbb{R} -Vektorraum \mathbb{R}^n haben wir das euklidische Skalarprodukt

$$\langle u | v \rangle = \sum_{k=1}^n u_k v_k.$$

◆ Beispiel P1Q: das Produktintegral auf $\mathcal{C}([a, b], \mathbb{R})$

Sei $a < b$ in \mathbb{R} und $\mathcal{C}([a, b], \mathbb{R})$ der \mathbb{R} -Vektorraum aller stetigen Funktionen $f, g : [a, b] \rightarrow \mathbb{R}$. Hierauf haben wir das Skalarprodukt

$$\langle f | g \rangle = \int_{t=a}^b f(t)g(t) dt.$$

Die Cauchy–Schwarz–Ungleichung

Wir legen hier nur die drei Eigenschaften (S1–3) zugrunde. In vorigen Beispiel P1A des euklidischen Skalarprodukts waren dies *Folgerungen* aus der explizit gegebenen Formel. In Definition P1B erheben wir diese nun zu den grundlegenden *Forderungen*. Damit können wir arbeiten:

Satz P1c: Cauchy–Schwarz–Ungleichung (CSU)

Aus (S1–3) folgt für alle $u, v \in V$ die **Cauchy–Schwarz–Ungleichung**:

$$|\langle u | v \rangle|^2 \leq \langle u | u \rangle \langle v | v \rangle \quad \text{kurz} \quad |\langle u | v \rangle| \leq \|u\| \cdot \|v\|$$

Gleichheit gilt genau dann, wenn u, v über \mathbb{R} linear abhängig sind.

Kurzbeweis: Dies ist klar für $v = 0$. Sei also $v \neq 0$. Für alle $t \in \mathbb{R}$ gilt:

$$0 \stackrel{(S1)}{\leq} \langle u + tv | u + tv \rangle \stackrel{(S2,3)}{=} \underbrace{\langle u | u \rangle}_{c \in \mathbb{R}_{\geq 0}} + 2 \underbrace{\langle u | v \rangle}_{b \in \mathbb{R}} t + \underbrace{\langle v | v \rangle}_{a \in \mathbb{R}_{> 0}} t^2$$

Demnach ist die Diskriminante nicht positiv: $b^2 - ac \leq 0$.

QED

Die Cauchy–Schwarz–Ungleichung

Ausführlicher Beweis: Für $v = 0$ ist alles klar. Sei also $v \neq 0$.

(1) Wir setzen $z := au - bv$ mit $a = \langle v | v \rangle$ und $b = \langle v | u \rangle$ und rechnen:

$$\begin{aligned} 0 &\stackrel{(S0)}{\leq} \langle z | z \rangle \stackrel{\text{Def}}{=} \langle au - bv | au - bv \rangle \\ &\stackrel{(S3,4)}{=} a^2 \langle u | u \rangle - ab \langle u | v \rangle - ab \langle v | u \rangle + b^2 \langle v | v \rangle \\ &\stackrel{(S2)}{=} \langle v | v \rangle [\langle u | u \rangle \langle v | v \rangle - |\langle u | v \rangle|^2] \end{aligned}$$

Dank (S1) gilt $\langle v | v \rangle > 0$. Wir erhalten so die ersehnte Ungleichung:

$$\langle u | u \rangle \langle v | v \rangle - |\langle u | v \rangle|^2 \geq 0$$

(2) Gleichheit impliziert $0 = \langle z | z \rangle$, also $0 = z$ und somit $u = (b/a)v$.

Umgekehrt folgt aus linearer Abhängigkeit $u = \lambda v$ direkt die Gleichheit

$$\begin{aligned} |\langle u | v \rangle|^2 &\stackrel{\text{Def}}{=} |\langle \lambda v | v \rangle|^2 \stackrel{(S2)}{=} \langle \lambda v | v \rangle \langle v | \lambda v \rangle \\ &\stackrel{(S3)}{=} \langle \lambda v | \lambda v \rangle \langle v | v \rangle \stackrel{\text{Def}}{=} \langle u | u \rangle \langle v | v \rangle. \end{aligned}$$

Damit ist die Cauchy–Schwarz–Ungleichung bewiesen.

QED

Die Cauchy–Schwarz–Ungleichung

Beide Beweise sind elegant, genial-einfach und einfach-genial!
Beide Argumente zaubern trickreich ein Kaninchen aus dem Hut:
Wir betrachten hier eine Linearkombination $z = au - bv$ mit $a, b \in \mathbb{R}$.
Als profitabel erweisen sich die Wahlen $a = \langle v | v \rangle$ und $b = \langle v | u \rangle$.

😊 Der Rest ist Ausrechnen und Ablesen der ersehnten Ungleichung.

Wie kommt man auf diese Koeffizienten? Wie merkt man sie sich?
Anders gefragt: Wie führt man den Beweis, wenn man sich zwar
an den Ansatz $z = au - bv$ erinnert, aber nicht an a und b ?

Die Rechnung bis zur zweiten Zeile ist noch ganz allgemein.
Wenn wir die letzten beiden Summanden auslöschen möchten,
dann gelingt dies mit der Wahl $a = \langle v | v \rangle$ und $b = \langle v | u \rangle$. Voilà!

😊 So lassen sich Ansatz und Strategie leicht verstehen und merken.

Der folgende, dritte Beweis gibt hierzu eine geometrische Motivation,
und erklärt insbesondere, wie man auf den Ansatz $z = au - bv$ kommt.

Die Cauchy–Schwarz–Ungleichung

Dritter Beweis: Für $v = 0$ ist alles klar. Sei also $v \neq 0$.

Wir zerlegen $u = u_{\parallel} + u_{\perp}$ mit $u_{\parallel} = \lambda v$ und $u_{\perp} = u - u_{\parallel}$.

Um Orthogonalität $v \perp u_{\perp}$ zu erreichen, betrachten wir

$\langle v | u_{\perp} \rangle = \langle v | u \rangle - \lambda \langle v | v \rangle$ und setzen $\lambda = \langle v | u \rangle / \langle v | v \rangle$.

Hier nutzen wir (S1), dies garantiert $\langle v | v \rangle > 0$. Damit gilt:

$$\begin{aligned} 0 &\stackrel{(S0)}{\leq} \langle u_{\perp} | u_{\perp} \rangle \stackrel{\text{Def}}{=} \langle u - \lambda v | u - \lambda v \rangle \\ &\stackrel{(S3,4)}{=} \langle u | u \rangle - \lambda \langle u | v \rangle - \lambda \langle v | u \rangle + \lambda^2 \langle v | v \rangle \\ &\stackrel{(S2)}{=} \langle u | u \rangle - |\langle u | v \rangle|^2 \langle v | v \rangle^{-1} \end{aligned}$$

Hieraus lesen wir sofort die Cauchy–Schwarz–Ungleichung ab.

Gleichheit gilt genau dann, wenn $u_{\perp} = 0$, also $u = u_{\parallel} = \lambda v$. QED

Übung: Vergleichen Sie den zweiten und den dritten Beweis.

Beide beruhen im Wesentlichen auf derselben Rechnung!

Der dritte ist geometrisch motiviert und dadurch instruktiv.

Der zweite ist algebraisch optimiert und etwas eleganter.

Die Cauchy–Schwarz–Ungleichung

In unserem ursprünglichen Beispiel des euklidischen Skalarprodukts bedeutet die Cauchy–Schwarz–Ungleichung ganz konkret und explizit:

$$(u_1^2 + \cdots + u_n^2)(v_1^2 + \cdots + v_n^2) \geq (u_1v_1 + \cdots + u_nv_n)^2$$

für alle reellen Zahlen $u_1, \dots, u_n, v_1, \dots, v_n \in \mathbb{R}$.

Übung: Bitte versuchen Sie, diese Ungleichung möglichst „direkt“ zu zeigen, ohne „Umweg“ über die allgemeine Theorie der Skalarprodukte.

Vermutlich werden Sie spüren, dass diese ganz konkret und harmlos erscheinende Ungleichung alles andere als leicht zu beweisen ist.

Nach eigenen Versuchen werden Sie feststellen, dass der oben erklärte allgemeine Begriff des Skalarprodukts und der so entwickelte Beweis gar kein Umweg ist, sondern eine Abkürzung! Ich betone daher erneut:

Abstraktion strukturiert und vereinfacht: Eine allgemeine Tatsache ist oft leichter zu verstehen und zu erklären als ihre zahlreichen Spezialfälle.

Die Cauchy–Schwarz–Ungleichung

Die Cauchy–Schwarz–Ungleichung ist eine der zentralen und universell nützlichen Ungleichungen in der Mathematik und ebenso für die Physik. Sie findet Anwendungen in der Geometrie, der Analysis, der Stochastik und der Quantenmechanik (etwa als Heisenbergs Unschärferelation).

Augustin-Louis Cauchy (1789–1857) veröffentlichte seine berühmte Ungleichung 1821 und in seinem Lehrbuch *Cours d'Analyse Algèbrique*, dem wohl ersten, aus heutiger Sicht strengen Aufbau der Analysis.

Victor Yacovlevich Bunyakovsky (1804–1889) studierte in Paris bei Cauchy und übertrug dessen Ungleichung von Summen auf Integrale, veröffentlicht 1859 in St. Petersburg, allerdings noch ohne Beweis.

Hermann Amandus Schwarz (1843–1921) arbeitete 1885 in Göttingen an Flächen minimaler Krümmung und benötigte dazu auch Chauchys Ungleichung für Integrale. Ihm gelang die allgemeine Formulierung, und der obige Beweis geht im Wesentlichen auf Schwarz zurück.




Eine genial-einfache Idee liefert alle nötigen Informationen.

Die Cauchy–Schwarz–Ungleichung

Im deutsch-sprachigen Raum ist die Bezeichnung *Cauchy–Schwarz–Ungleichung* üblich. In englisch-sprachigen Texten findet man häufig nur *Schwarz's inequality*, in französisch-sprachigen entsprechend *l'inegalité de Schwarz*, die russische Tradition hingegen bevorzugt *Cauchy–Bunyakovsky–Schwarz Ungleichung*. Alle drei Sichtweisen haben ihre Gründe; historische Gerechtigkeit ist meist schwierig.

Die gesamte Analysis beruht zu weiten Teilen auf Ungleichungen, und die Cauchy–Schwarz–Ungleichung ist hier ein erstes wichtiges Ergebnis sowie Vorbild für zahlreiche nachfolgende Ungleichungen.

 J.M. Steele: *The Cauchy–Schwarz master class. An Introduction to the Art of Mathematical Inequalities*. Cambridge University Press 2004.

Übung: Sei V ein Vektorraum mit Skalarprodukt und $u, v \in V$.

- (1) Orthogonalität $u \perp v$ ist äquivalent zu $\|u + tv\| \geq \|u\|$ für alle $t \in \mathbb{R}$.
- (2) Machen Sie eine Skizze und interpretieren Sie dies geometrisch.

Winkel dank Cauchy–Schwarz–Ungleichung

Sei V ein \mathbb{R} -Vektorraum mit einem Skalarprodukt $\langle - | - \rangle : V \times V \rightarrow \mathbb{R}$.

Folgerung aus der CSU: Für je zwei Vektoren $u, v \in V \setminus \{0\}$ gilt

$$\frac{\langle u | v \rangle}{\|u\| \cdot \|v\|} \in [-1, 1].$$

Daher existiert genau eine reelle Zahl $\theta \in [0, \pi]$ mit

$$\cos \theta = \frac{\langle u | v \rangle}{\|u\| \cdot \|v\|}.$$

Wir definieren den **Winkel** zwischen den Vektoren u und v durch

$$\angle(u, v) := \theta = \arccos\left(\frac{\langle u | v \rangle}{\|u\| \cdot \|v\|}\right).$$

So überführen wir unsere geometrische Intuition in eine solide Definition und bequeme Rechnung: Damit können wir Winkel effizient berechnen. Die obige Formel ist vollkommen explizit und bereit zum Einsatz.

Winkel dank Cauchy–Schwarz–Ungleichung

😊 Das ist eine genial-einfache Idee und elegante Definition. Wie sonst sollten wir den Winkel zwischen zwei Vektoren u und v ermitteln?

In der euklidischen Ebene $\mathbb{C} = \mathbb{R}^2$ gelingt dies mit dem Bogenmaß und den trigonometrischen Funktionen \sin und \cos , wie oben erklärt (P105).

Noch informativer ist der **orientierte Winkel** $\alpha \in]-\pi, \pi]$, definiert durch

$$\frac{v}{\|v\|} = e^{i\alpha} \frac{u}{\|u\|} \quad \text{also} \quad \frac{1}{\|v\|} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \cdot \frac{1}{\|u\|} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

😊 Die oben erklärte Gleichung $\langle u | v \rangle = \|u\| \cdot \|v\| \cdot \cos \theta$ genügt, um den **absoluten Winkel** $\theta = |\alpha| \in [0, \pi]$ zu bestimmen. Das **Vorzeichen** entspricht der Orientierung des Winkels und gelingt nur in der Ebene, wo wir Links- und Rechtsdrehung unterscheiden vermöge $\det(u, v) \geq 0$ oder $\det(u, v) < 0$. Schon im Raum \mathbb{R}^3 ist dies nicht mehr möglich.

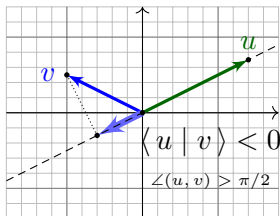
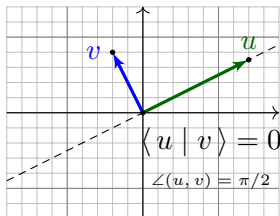
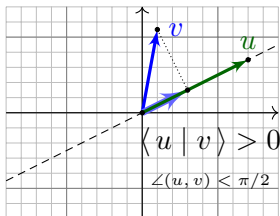
In höherdimensionalen Räumen verlässt uns die Anschauung vollends. Wie wollen Sie den Winkel zwischen zwei Vektoren im \mathbb{R}^{100} messen? So gesehen ist die obige Definition wirklich einfach und elegant.

Winkel dank Cauchy–Schwarz–Ungleichung

Damit gilt weiterhin die gewohnte Gleichung

$$\langle u | v \rangle = \|u\| \cdot \|v\| \cdot \cos \angle(u, v).$$

Wie zuvor ist $\|v\| \cos \angle(u, v)$ die Länge der Projektion von v parallel zu u .



Im wichtigen Spezialfall $\langle u | v \rangle = 0$ sagen wir, die Vektoren u und v sind **orthogonal** oder **stehen senkrecht** zueinander, geschrieben

$$u \perp v \quad :\Leftrightarrow \quad \langle u | v \rangle = 0.$$

Winkel dank Cauchy–Schwarz–Ungleichung

⚠ Den Winkel $\angle(u, v)$ können wir nur für Vektoren $u, v \in V$ definieren, die ungleich Null sind, da wir durch die Normen $\|u\|$ und $\|v\|$ dividieren.

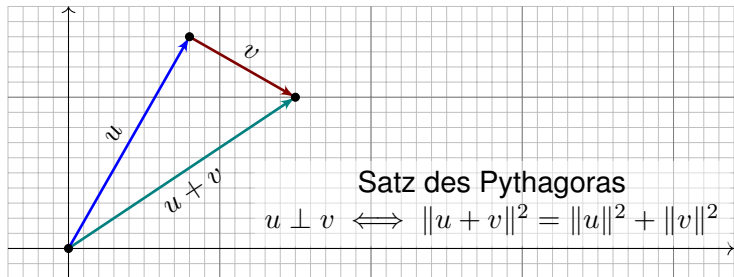
😊 Orthogonalität hingegen können wir für je zwei beliebige Vektoren $u, v \in V$ definieren durch die einfache Bedingung $\langle u | v \rangle = 0$.

Orthogonalität wird im Folgenden eine große Rolle spielen, sowohl für die Entwicklung der Theorie als auch in der Anwendung der Methoden.

Als zwei einfacher doch eindruckliche Beispiele diskutieren wir den Satz des Pythagoras und die Parallelogrammgleichung.

Eingangs haben wir das euklidische Skalarprodukt durch Pythagoras motiviert und hergeleitet. Schön zu sehen und beruhigend zu wissen, dass dieser wichtige Satz bei unserer Abstraktion nicht verloren geht. Im Gegenteil, er erhält nun seine gebührend allgemeine Formulierung.

Satz des Pythagoras



😊 In Worten: In jedem rechtwinkligen Dreieck ist das Normquadrat der Hypotenuse gleich der Summe der Normquadrate der Katheten.

Satz des Pythagoras

Satz P1D: Pythagoras

Sei V ein \mathbb{R} -Vektorraum mit einem Skalarprodukt

$$\langle - | - \rangle : V \times V \rightarrow \mathbb{R} : (u, v) \mapsto \langle u | v \rangle.$$

(1) Für je zwei Vektoren $u, v \in V$ gilt dann die Gleichung

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2 \langle u | v \rangle.$$

Mit dem oben eingeführten Winkel $\angle(u, v)$ schreibt sich dies wie folgt:

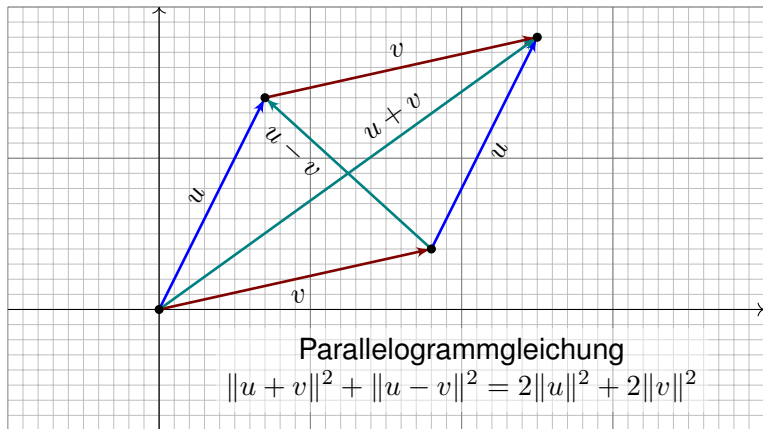
$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2 \|u\| \cdot \|v\| \cdot \cos \angle(u, v)$$

(2) Stehen u und v senkrecht zueinander, so entfällt der letzte Term:

$$u \perp v \iff \|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Übung: Alles steht explizit da. Rechnen Sie es sorgsam nach!

Die Parallelogrammgleichung



😊 In Worten: In jedem Parallelogramm ist die Summe der Quadrate der vier Seiten gleich der Summe der Quadrate der beiden Diagonalen.

Die Parallelogrammgleichung

Satz P1E: Parallelogrammgleichung

Sei V ein \mathbb{R} -Vektorraum mit einem Skalarprodukt

$$\langle - | - \rangle : V \times V \rightarrow \mathbb{R} : (u, v) \mapsto \langle u | v \rangle.$$

(1) Für je zwei Vektoren $u, v \in V$ gilt dann die Gleichung

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

(2) Allein aus der Norm $\|-\|$ lässt sich das Skalarprodukt $\langle - | - \rangle$ rekonstruieren dank der folgenden Polarisationsformel:

$$\langle u | v \rangle = \frac{1}{4} \left[\|u + v\|^2 - \|u - v\|^2 \right]$$

😊 Insbesondere folgt damit $u \perp v \Leftrightarrow \|u + v\| = \|u - v\|$.
Jeder gute Handwerker kennt diesen Test zur Rechtwinkligkeit.

Norm und Dreiecksungleichung

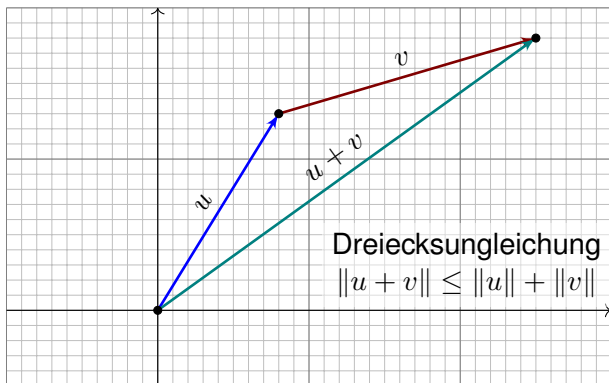
Sei V ein \mathbb{R} -Vektorraum mit einem Skalarprodukt

$$\langle - | - \rangle : V \times V \rightarrow \mathbb{R} : (u, v) \mapsto \langle u | v \rangle.$$

Die zugehörige **Norm** misst die Länge von Vektoren:

$$\|-\| : V \rightarrow \mathbb{R}_{\geq 0} : v \mapsto \|v\| = \sqrt{\langle v | v \rangle}$$

Was sind ihre wesentlichen Eigenschaften?



Satz P1F: Eigenschaften der Norm

Für alle Vektoren $u, v \in V$ und Skalare $\lambda \in \mathbb{R}$ gilt:

- N0: Positivität, $\|u\| \geq 0 = \|0\|$
 N1: positive Definitheit, $\|u\| > 0$ für $u \neq 0$
 N2: absolute Homogenität, $\|\lambda u\| = |\lambda| \cdot \|u\|$
 N3: Dreiecksungleichung, $\|u + v\| \leq \|u\| + \|v\|$

Wir benötigen: Die Wurzelfunktion $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} : x \mapsto \sqrt{x}$ erfüllt $\sqrt{0} = 0$, ist streng monoton, $x < y \Rightarrow \sqrt{x} < \sqrt{y}$, und multiplikativ, $\sqrt{xy} \leq \sqrt{x}\sqrt{y}$.

Beweis: Aus (S0,1,2) folgt (N0,1,2), und (N3) folgt dank CSU:

$$\begin{aligned} \|u + v\|^2 &\stackrel{\text{Def}}{=} \langle u + v \mid u + v \rangle \stackrel{\text{Bin}}{=} \langle u \mid u \rangle + \langle u \mid v \rangle + \langle v \mid u \rangle + \langle v \mid v \rangle \\ &\stackrel{\text{CSU}}{\leq} \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \stackrel{\text{Bin}}{=} (\|u\| + \|v\|)^2 \end{aligned}$$

Daraus folgt $\|u + v\| \leq \|u\| + \|v\|$, wie behauptet.

QED

Weitere Beispiele: Taxinorm und Maximumsnorm

Eine Abbildung $\|-\| : V \rightarrow \mathbb{R}_{\geq 0}$ mit den obigen Eigenschaften (N1–3) nennen wir eine **Norm** auf dem Vektorraum V , siehe Definition P1L.

Beispiel: Auf dem Vektorraum \mathbb{R}^n nutzen wir je nach Bedarf vor allem

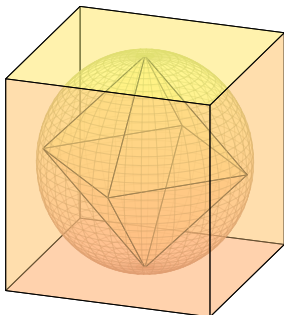
die Taxinorm	$\ x\ _1 := x_1 + x_2 + \cdots + x_n ,$
die Maximumsnorm	$\ x\ _\infty := \max\{ x_1 , x_2 , \dots, x_n \},$
die euklidische Norm	$\ x\ _2 := \sqrt{ x_1 ^2 + x_2 ^2 + \cdots + x_n ^2}.$

bung: (1) Weisen Sie nach, dass dies tatsachlich Normen sind.
 (2) Taxinorm und Maximumsnorm kommen jedoch nicht von einem Skalarprodukt, denn sie erfullen nicht die Parallelogrammgleichung P1E.

Statt der Schreibweise $\|u\|$ fur Normen ist abkurzend auch $|u|$ ublich. Erstere dient zur Betonung und zur Unterscheidung vom Betrag.

Ein Vektor $u \in V$ heit **normiert**, falls $\|u\| = 1$ gilt. Wir konnen jeden Vektor $v \in V \setminus \{0\}$ normieren zu $u = \|v\|^{-1}v$ mit $\|u\| = \|v\|^{-1}\|v\| = 1$.

Weitere Beispiele: Taxinorm und Maximumsnorm



Die euklidische Norm $\|-\|_2 : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} : v \mapsto \sqrt{\langle v | v \rangle}$ definiert den **abgeschlossenen / offenen Einheitsball** und die **Einheitssphere**:

$$\mathbb{D}^n := \{ x \in \mathbb{R}^n \mid \|x\|_2 \leq 1 \} = \{ x \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 \leq 1 \},$$

$$\mathbb{B}^n := \{ x \in \mathbb{R}^n \mid \|x\|_2 < 1 \} = \{ x \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 < 1 \},$$

$$\mathbb{S}^{n-1} := \{ x \in \mathbb{R}^n \mid \|x\|_2 = 1 \} = \{ x \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 = 1 \}.$$

Fur die Maximumsnorm erhalten wir hier stattdessen den **Wurfel**.
Fur die Taxinorm erhalten wir dual hierzu das **Kreuzpolytop**.

Metrik und Dreiecksungleichung

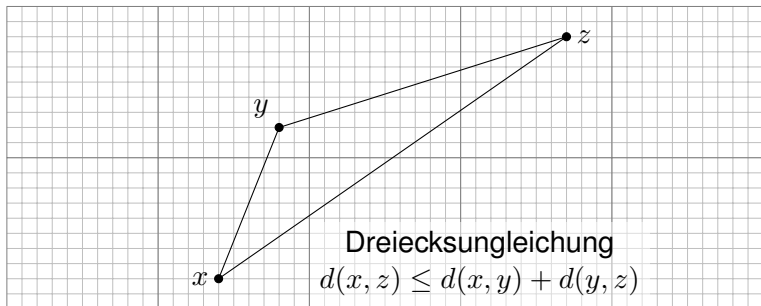
Sei V ein \mathbb{R} -Vektorraum mit einer Norm

$$\|-\| : V \rightarrow \mathbb{R}_{\geq 0} : v \mapsto \|v\| = \sqrt{\langle v | v \rangle}$$

Die zugehörige **Metrik** misst den Abstand von Punkten:

$$d : V \times V \rightarrow \mathbb{R}_{\geq 0} : (x, y) \mapsto d(x, y) = \|y - x\|$$

Was sind ihre wesentlichen Eigenschaften?



Die Dreiecksungleichung besagt anschaulich: Der Weg von x nach z wird nicht kürzer, wenn wir einen Umweg über y machen.

Satz P1G: Eigenschaften der Metrik

Die zur Norm $\|-\| : V \rightarrow \mathbb{R}_{\geq 0}$ gehorige Metrik

$$d : V \times V \rightarrow \mathbb{R}_{\geq 0} : (x, y) \mapsto d(x, y) := \|y - x\|$$

erfreut sich folgender Eigenschaften fur alle $x, y, z \in V$:

M0: Positivitat, $d(x, y) \geq 0 = d(x, x)$

M1: positive Definitheit, $d(x, y) > 0$ fur $x \neq y$

M2: Symmetrie, $d(x, y) = d(y, x)$

M3: Dreiecksungleichung, $d(x, z) \leq d(x, y) + d(y, z)$

Aufgabe: Rechnen Sie dies zur bung nach!**Losung:** Aus (N0,1,2,3) folgt (M0,1,2,3), hier fur (M3) ausfuhrlich:

$$d(x, z) \stackrel{\text{Def}}{=} \|z - x\| \stackrel{\text{Vek}}{=} \|z - y + y - x\| \stackrel{(N3)}{\leq} \|z - y\| + \|y - x\| \stackrel{\text{Def}}{=} d(x, y) + d(y, z)$$

Weitere Beispiele für Metriken

Ein **metrischer Raum** (X, d) ist eine Menge X mit einer Metrik d , also einer Abbildung $d: X \times X \rightarrow \mathbb{R}_{\geq 0}$, die (M0–3) erfüllt.

Mit einer Metrik können wir immerhin noch Abstände messen. Sie ist auf jeder beliebigen Menge definierbar und hat im Allgemeinen keinen Bezug mehr zu Vektorräumen oder Längen und Winkeln von Vektoren.

Beispiel: Auf jeder beliebigen Menge X lässt sich eine besonders einfache Metrik definieren, die nur die Werte 0 und 1 annimmt: Diese **diskrete Metrik** auf X ist gegeben durch

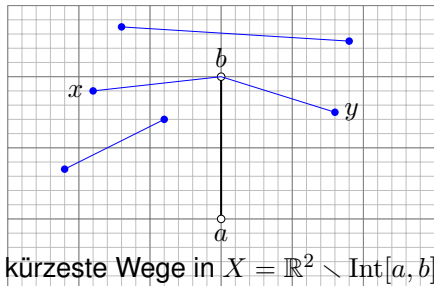
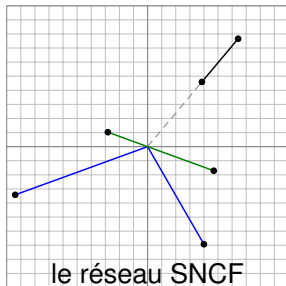
$$d : X \times X \rightarrow \{0, 1\} : (x, y) \mapsto \begin{cases} 0 & \text{falls } x = y, \\ 1 & \text{falls } x \neq y. \end{cases}$$

Übung: (1) Weisen Sie nach, dass dies tatsächlich eine Metrik ist.
(2) Die diskrete Metrik auf dem \mathbb{R}^n kommt nicht von einer Norm.
(3) Dasselbe gilt für folgende Metrik: Die kürzeste Verbindung zwischen zwei französischen Städten x und y führt über Paris, es sei denn beide Städte liegen auf einer gemeinsamen Eisenbahnstrecke nach Paris.

Weitere Beispiele fur Metriken

Beispiel: Die **franzosische Eisenbahnmetrik** auf der Menge \mathbb{R}^n ist

$$d = d_{\text{SNCF}} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} : (x, y) \mapsto \begin{cases} |x - y| & \text{falls } \mathbb{R}x = \mathbb{R}y, \\ |x| + |y| & \text{falls } \mathbb{R}x \neq \mathbb{R}y. \end{cases}$$



Beispiel: Sei $X \subseteq \mathbb{R}^n$ eine Teilmenge des euklidischen Raum oder allgemein ein beliebiger metrischer Raum (X, d) . Wir definieren die **Wegmetrik** $\hat{d} : X \times X \rightarrow [0, \infty]$ fur je zwei Punkte $x, y \in X$ als die infimale Lange aller Wege von x nach y . **ubung:** Fuhren Sie dies aus.

Abstand Punkt-Menge und Menge-Menge

Sei (X, d) ein metrischer Raum. Der Abstand eines Punktes $a \in X$ zu einer Teilmenge $B \subseteq X$ bzw. zwischen zwei Teilmengen $A, B \subseteq X$ ist definiert als das Infimum der punktweisen Abstande:

$$d(a, B) := \inf \{ d(a, b) \mid b \in B \}$$

$$d(A, B) := \inf \{ d(a, b) \mid a \in A, b \in B \}$$

Aufgabe: Zeichnen Sie ein Handballfeld nach folgenden Magaben, siehe Regeln 1:1 bis 1:9 der *International Handball Federation*:

$X = [-20, 20] \times [-10, 10] \subset \mathbb{R}^2$ die Spielflache

$M = \{0\} \times [-10, 10]$ die Mittellinie

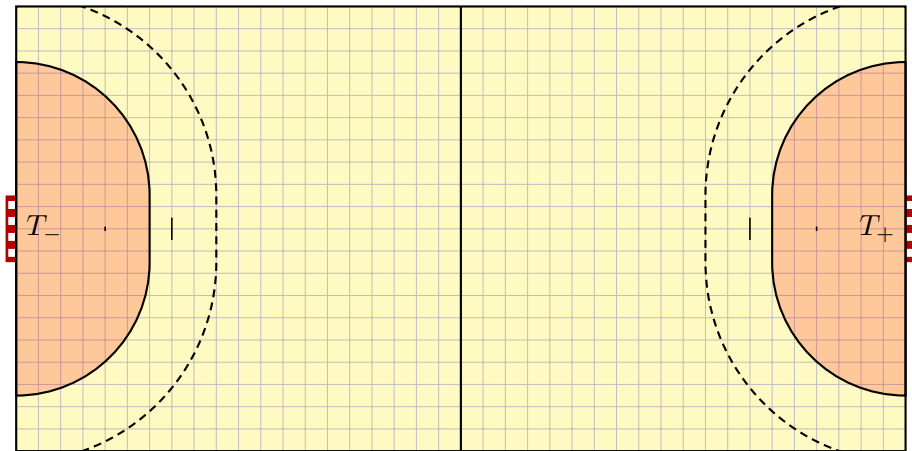
$T = \{-20, 20\} \times [-1.5, 1.5]$ die beiden Torlinien

$S = \{x \in X \mid d(x, T) = 6\}$ Sechs-Meter-Linie („Kreis“)

$N = \{x \in X \mid d(x, T) = 9\}$ Neun-Meter-Linie („Freiwurflinie“)

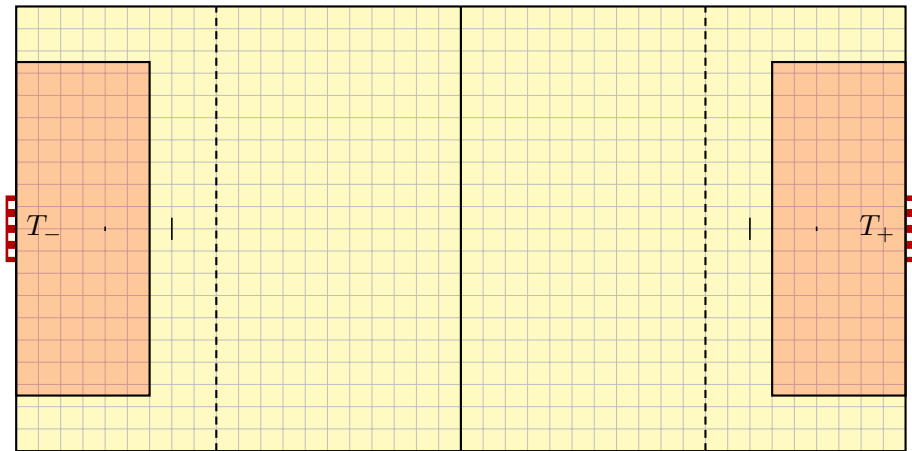
Wie sieht das Feld bezuglich der ublichen, euklidischen Metrik aus?
zum Vergleich mit der Maximumsmetrik? und mit der Taximetrik?

Handballfeld in der euklidischen Metrik



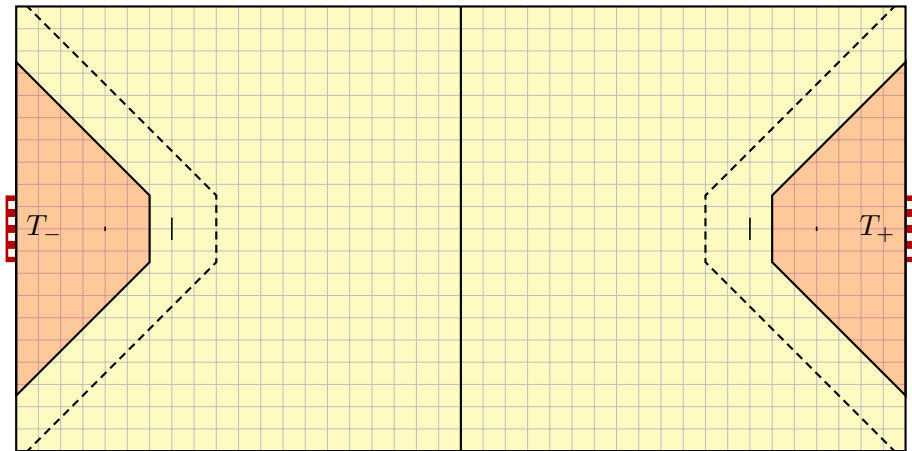
☺ Die Sechs-Meter-Linie, der sogenannte „Kreis“, ist gar keiner!
Sie besteht aus zwei Viertelkreisen und einem Geradenstuck.

Handballfeld in der Maximumsmetrik



😊 In der Maximumsmetrik ist die Sechs-Meter-Linie ein Rechteck.
Fun fact: Im Fußball folgen Tor- und Strafraum dieser Konvention.

Handballfeld in der Taximetrik



☺ In der Taximetrik ist die Sechs-Meter-Linie ein Trapez.
Das sieht ulkig aus und illustriert eindrucklich die Taximetrik .

Anwendungen von Skalarprodukt, Norm und Metrik

Das euklidische Skalarprodukt auf dem \mathbb{R}^n sowie die zugehörige euklidische Norm und die euklidische Metrik sind grundlegend: Sie spielen in vielen Anwendungen eine tragende Rolle.

Daneben gibt es andere Normen und andere Metriken, die je nach Anwendung natürlich auftreten, aber gänzlich andere geometrische Eigenschaften haben. Mein Anliegen in den obigen Ergänzungen ist, durch illustrative Beispiele den nötigen Kontrast zu schaffen. Nur so sehen Sie, wie schön und einfach das euklidische Skalarprodukt ist.

Dies ist das zentrale Beispiel und ein leuchtendes Vorbild für die gesamte Lineare Algebra und viele Anwendungen in der Analysis, der Numerik, der Stochastik, der Physik und vielen weiteren Gebieten.

Anwendungen von Skalarprodukt, Norm und Metrik

Normen und Metriken begegnen Ihnen vor allem in der Analysis und der Numerik, aber ebenso auch in der Geometrie und der Topologie. Sie dienen zur Messung von Abständen, zur Abschätzung von Fehlern, zur Kontrolle von Näherungen und zur Definition der Konvergenz.

Beide Handlungsstränge, Vektorräume und Normen, führen schließlich in der Funktionalanalysis wieder zusammen. Für die Untersuchung von unendlich-dimensionalen Vektorräumen sind Skalarprodukte, Normen, Metriken und Topologien unentbehrliche Hilfsmittel.

Zunächst sind Normen und Metriken hier nur ein amüsanter Ausblick, doch sie sind zugleich wichtig genug, jetzt schon erwähnt zu werden. Wir konzentrieren uns im Folgenden wieder auf das Skalarprodukt.

Anwendungen von Skalarprodukt, Norm und Metrik

😊 Mit dem Skalarprodukt auf V messen wir Winkel und Längen. Damit gewinnen wir für V wichtige geometrische Werkzeuge:

- **Orthogonalität:** $u, v \in V$ stehen senkrecht, wenn $\langle u | v \rangle = 0$.
- **Norm:** Die Länge eines Vektors $v \in V$ ist $\|v\| = \sqrt{\langle v | v \rangle}$.
- **Cauchy–Schwarz–Ungleichung:** Es gilt $|\langle u | v \rangle| \leq \|u\| \cdot \|v\|$.
- **Winkel:** $\langle u | v \rangle = \|u\| \cdot \|v\| \cdot \cos(\theta)$ mit $\theta = \angle(u, v) \in [0, \pi]$.
- **Dreiecksungleichung:** Es gilt $\|u + v\| \leq \|u\| + \|v\|$.
- **Metrik:** Der Abstand zweier Vektoren u, v ist $\|v - u\|$.
- **Konvergenz** $v_n \rightarrow v$ ist definiert durch $\|v_n - v\| \rightarrow 0$.
- **Vollständigkeit:** Jede Cauchy–Folge in V konvergiert in V .
- **Stetigkeit** von Funktionen $f : V \rightarrow W$, linear oder nicht.
- **Differenzierbarkeit**, lineare und höhere Approximation.

Skalarprodukt und Norm auf einem reellen Vektorraum

Zusammenfassend vereinbaren wir den folgenden Sprachgebrauch für Skalarprodukte auf Vektorräumen über \mathbb{R} (und anschließend über \mathbb{C}):

◆ Definition P1L: Skalarprodukt und Norm

Sei V ein Vektorraum über dem Körper \mathbb{R} . Ein **Skalarprodukt** auf V ist eine Abbildung $\langle - | - \rangle : V \times V \rightarrow \mathbb{R}$, die (S1–3) erfüllt.

Das Paar $(V, \langle - | - \rangle)$ heißt dann **\mathbb{R} -Vektorraum mit Skalarprodukt** oder **euklidischer Vektorraum** oder ein reeller **Prä-Hilbert-Raum** und bei metrischer Vollständigkeit auch ein reeller **Hilbert-Raum**.

Eine **Norm** auf V ist eine Abbildung $\|-\| : V \rightarrow \mathbb{R}_{\geq 0}$, die (N1–3) erfüllt.

Das Paar $(V, \|-\|)$ heißt dann **normierter \mathbb{R} -Vektorraum** oder auch **Prä-Banach-Raum** und bei Vollständigkeit reeller **Banach-Raum**.

Für ein **semidefinites Skalarprodukt** fordern wir nur (S0,2,3), für eine **Seminorm** oder **Halbnorm** entsprechend nur (N0,2,3). Beide Abschwächungen kommen in Anwendungen natürlich vor.

Der Betrag komplexer Zahlen

Erinnerung: Zu jeder komplexen Zahl $z = x + iy$ mit $x, y \in \mathbb{R}$ haben wir die konjugierte Zahl $\bar{z} = x - iy$. Ihr Produkt ist somit $z\bar{z} = x^2 + y^2 \geq 0$.

Den Betrag $|\cdot| : \mathbb{C} \rightarrow \mathbb{R}_{\geq 0}$ komplexer Zahlen definieren wir durch

$$|z| := \sqrt{z\bar{z}}, \quad \text{also} \quad |x + iy| := \sqrt{x^2 + y^2} \quad \text{für alle } x, y \in \mathbb{R}.$$

Für alle komplexen Zahlen $u, v \in \mathbb{C}$ gilt dann:

N0: Null und Eins, $|0| = 0$ und $|1| = 1$

N1: positive Definitheit, $|u| > 0$ für $u \neq 0$

N2: Multiplikativität, $|u \cdot v| = |u| \cdot |v|$

N3: Dreiecksungleichung, $|u + v| \leq |u| + |v|$

Aufgabe: (1) Wiederholen bzw. beweisen Sie diese Aussagen.

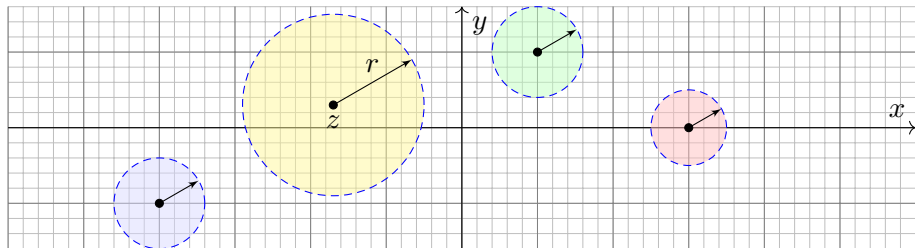
(2) Zeichnen Sie für verschiedene Werte von $z \in \mathbb{C}$ und $r \in \mathbb{R}_{>0}$ die Menge $B(z, r) := \{ u \in \mathbb{C} \mid |u - z| < r \}$ in der komplexen Ebene \mathbb{C} .

Der Betrag komplexer Zahlen

Lösung: (1) Wir haben $\mathbb{C} = \mathbb{R}^2$ und hierauf ist $|\cdot|$ die euklidische Norm. Daraus folgt (N0,1,3) dank P1F. Dank A3B ist die komplexe Konjugation $\bar{\cdot} : \mathbb{C} \rightarrow \mathbb{C} : (x, y) \mapsto (x, -y)$ ein Körperautomorphismus, also gilt:

$$|uv| = \sqrt{uv \cdot \overline{uv}} = \sqrt{u\bar{u} \cdot v\bar{v}} = \sqrt{u\bar{u}} \cdot \sqrt{v\bar{v}} = |u| \cdot |v|$$

(2) Die Menge $B(z, r) := \{ u \in \mathbb{C} \mid |u - z| < r \}$ ist eine offene Kreisscheibe um den Mittelpunkt $z \in \mathbb{C}$ mit Radius $r \in \mathbb{R}_{>0}$.



Das euklidische Skalarprodukt auf dem Raum \mathbb{C}^n

Beispiel P1H: das euklidische Skalarprodukt auf dem Raum \mathbb{C}^n

Auf $V = \mathbb{C}^n$ über \mathbb{C} definieren wir das **euklidische Skalarprodukt**

$$\langle - | - \rangle : V \times V \rightarrow \mathbb{C} : (u, v) \mapsto \langle u | v \rangle := \overline{u_1}v_1 + \cdots + \overline{u_n}v_n.$$

Für alle Vektoren $u, v, w \in V$ und Skalare $\lambda, \mu \in \mathbb{C}$ gilt:

S0: Positivität, $\langle u | u \rangle \geq 0 = \langle 0 | 0 \rangle$

S1: positive Definitheit, $\langle u | u \rangle > 0$ für $u \neq 0$

S2: konjugierte Symmetrie, $\langle v | u \rangle = \overline{\langle u | v \rangle}$

S3: Linearität rechts, $\langle u | \lambda v + \mu w \rangle = \lambda \langle u | v \rangle + \mu \langle u | w \rangle$

Aus (S2) und (S3) folgt konjugierte Linearität in der ersten Variablen:

S4: konjugierte Linearität links, $\langle \lambda u + \mu v | w \rangle = \overline{\lambda} \langle u | w \rangle + \overline{\mu} \langle v | w \rangle$

Zu (S4) sagt man auch **semilinear**, zu (S3,4) daher kurz **sesquilinear** (lat. *sesqui*, 'anderthalb'). Zu (S2) sagt man **hermitesch**, zu Ehren des französischen Mathematikers Charles Hermite (1822–1901).

Das euklidische Skalarprodukt auf dem Raum \mathbb{C}^n

Somit ist das euklidische Skalarprodukt $\langle - | - \rangle$ auf $V = \mathbb{C}^n$ über \mathbb{C} eine **hermitesche Sesquilinearform** (S2,3) und **positiv definit** (S0,1).

Eigenschaft (S2) impliziert $\overline{\langle u | u \rangle} = \langle u | u \rangle$, also ist dieser Wert reell und die Frage nach $\langle u | u \rangle \geq 0$ oder $\langle u | u \rangle > 0$ ist überhaupt sinnvoll.

Aufgabe: Rechnen Sie die hier gemachten Aussagen nach.

Lösung: Hermitizität (S2) und Sesquilinearität (S3,4) sind klar.

(S0) Für jeden Vektor $u \in \mathbb{C}^n$ gilt $\langle u | u \rangle = |u_1|^2 + \dots + |u_n|^2 \geq 0$.

(S1) Im Falle $u \neq 0$ gilt $u_i \neq 0$ für mindestens ein $i \in \{1, \dots, n\}$. Daraus folgt sofort die strikte Ungleichung $\langle u | u \rangle \geq |u_i|^2 > 0$.

Bemerkung: Auf dem Vektorraum \mathbb{R}^n ist die Bilinearform $(u, v) \mapsto u_1v_1 + \dots + u_nv_n$ positiv definit, auf \mathbb{C}^n jedoch nicht.

Zur Korrektur müssen wir eine der beiden Variablen konjugieren. Ich plädiere für die erste, dadurch werden einige Formeln schöner.

 Manche Autoren wählen die zweite; das ist eine Geschmacksfrage. Beide Konventionen sind durch Konjugation ineinander umzurechnen.

Skalarprodukt und Cauchy–Schwarz–Ungleichung

Definition P1i: Skalarprodukt über \mathbb{C}

Ein **Skalarprodukt** auf einem \mathbb{C} -Vektorraum V ist eine positiv definite, hermitesche Sesquilinearform $\langle - | - \rangle : V \times V \rightarrow \mathbb{C}$, erfüllt also (S1–3).

Die zugehörige **Norm** ist dann $\| - \| : V \rightarrow \mathbb{R}_{\geq 0} : v \mapsto \|v\| = \sqrt{\langle v | v \rangle}$.

Bemerkung: Die reelle und die komplexe Definition stimmen überein, wenn wir die beiden Grundkörper \mathbb{R} und \mathbb{C} ausstatten mit der Identität $- : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x$ bzw. der Konjugation $- : \mathbb{C} \rightarrow \mathbb{C} : (x, y) \mapsto (x, -y)$.

Satz P1j: Cauchy–Schwarz–Ungleichung (CSU)

Aus (S1–3) folgt für alle $u, v \in V$ die **Cauchy–Schwarz–Ungleichung**:

$$|\langle u | v \rangle|^2 \leq \langle u | u \rangle \langle v | v \rangle \quad \text{kurz} \quad |\langle u | v \rangle| \leq \|u\| \cdot \|v\|$$

Gleichheit gilt genau dann, wenn u, v über \mathbb{C} linear abhängig sind.

Aufgabe: Beweisen Sie dies, indem Sie den reellen Beweis anpassen.

Skalarprodukt und Cauchy–Schwarz–Ungleichung

Lösung: Die Aussage ist klar für $v = 0$. Im Folgenden sei also $v \neq 0$.

Wir setzen $z := au - bv$ mit $a = \langle v | v \rangle$ und $b = \langle v | u \rangle$ und rechnen:

$$\begin{aligned} 0 &\stackrel{(S0)}{\leq} \langle z | z \rangle \stackrel{\text{Def}}{=} \langle au - bv | au - bv \rangle \\ &\stackrel{(S3,4)}{=} |a|^2 \langle u | u \rangle - \bar{a}b \langle u | v \rangle - a\bar{b} \langle v | u \rangle + |b|^2 \langle v | v \rangle \\ &\stackrel{(S2)}{=} \langle v | v \rangle [\langle u | u \rangle \langle v | v \rangle - |\langle u | v \rangle|^2] \end{aligned}$$

Dank (S1) gilt $\langle v | v \rangle > 0$. Wir erhalten so die ersehnte Ungleichung:

$$\langle u | u \rangle \langle v | v \rangle - |\langle u | v \rangle|^2 \geq 0$$

Gleichheit impliziert $0 = \langle z | z \rangle$, also $0 = z$ und somit $u = (b/a)v$.

Umgekehrt folgt aus linearer Abhängigkeit $u = \lambda v$ direkt die Gleichheit

$$\begin{aligned} |\langle u | v \rangle|^2 &\stackrel{\text{Def}}{=} |\langle \lambda v | v \rangle|^2 \stackrel{(S2)}{=} \langle \lambda v | v \rangle \langle v | \lambda v \rangle \\ &\stackrel{(S3)}{=} \langle \lambda v | \lambda v \rangle \langle v | v \rangle \stackrel{\text{Def}}{=} \langle u | u \rangle \langle v | v \rangle. \end{aligned}$$

Damit ist die Cauchy–Schwarz–Ungleichung bewiesen.

QED

Eigenschaften des reellen und komplexen Skalarprodukts

Aufgabe: Übertragen Sie soweit möglich die Eigenschaften des reellen auf das komplexe Skalarprodukt: Norm? Metrik? Winkel? Orthogonalität? Pythagoras? Parallelogrammgleichung? Polarisationsformel?

Lösung: (1) Die Eigenschaften der zugehörigen Norm (P1F) und der daraus abgeleiteten Metrik (P1G) gelten wörtlich genauso über \mathbb{C} .

(2) Im Reellen ist der Winkel $\theta = \angle(u, v)$ zwischen u und v definiert durch $\langle u | v \rangle = \|u\| \cdot \|v\| \cdot \cos \theta$. Im Komplexen ist $\langle u | v \rangle$ i.A. nicht reell. Notgedrungen definieren wir θ daher durch $\operatorname{Re}\langle u | v \rangle = \|u\| \cdot \|v\| \cdot \cos \theta$.

(3) Orthogonalität vermöge $u \perp v : \Leftrightarrow \langle u | v \rangle = 0$ ist weiterhin sinnvoll. Vorsicht: Hierzu ist $\angle(u, v) = \pi/2$ notwendig, aber nicht hinreichend.

(4) Der Satz des Pythagoras (P1D) gilt weiterhin, nun in der Form $\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2 \operatorname{Re}\langle u | v \rangle = \|u\|^2 + \|v\|^2 + 2 \|u\| \cdot \|v\| \cdot \cos \theta$. Aus $u \perp v$ folgt $\|u + v\|^2 = \|u\|^2 + \|v\|^2$, die Umkehrung gilt nicht mehr.

(5) Die Parallelogrammgleichung (P1E) gilt wörtlich genauso wie zuvor. Die Polarisationsformel muss um den Imaginärteil ergänzt werden. Der folgende Satz führt beide Fälle noch einmal explizit aus.

Satz P1κ: Polarisationsformel, reell und komplex

(1) Über \mathbb{R} lässt sich aus der Norm $\|-\|$ das Skalarprodukt $\langle - | - \rangle$ rekonstruieren dank der folgenden, reellen Polarisationsformel:

$$\langle u | v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2}{4}$$

(2) Über \mathbb{C} lässt sich aus der Norm $\|-\|$ das Skalarprodukt $\langle - | - \rangle$ rekonstruieren dank der folgenden, komplexen Polarisationsformel:

$$\langle u | v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2}{4} + i \frac{\|u - iv\|^2 - \|u + iv\|^2}{4}.$$

Übung: Alles steht explizit da. Rechnen Sie es sorgsam nach!

Insbesondere gilt $\angle(u, v) = \pi/2 \Leftrightarrow \|u + v\| = \|u - v\|$ und vollständig $u \perp v \Leftrightarrow \|u + v\| = \|u - v\| \wedge \|u - iv\| = \|u + iv\|$.

😊 Komplexe Handwerker benötigen zwei Tests zur Rechtwinkligkeit.

Skalarprodukt und Norm über $\mathbb{K} = \mathbb{R}, \mathbb{C}$

Zusammenfassend vereinbaren wir den folgenden Sprachgebrauch:

Definition P1L: Skalarprodukt und Norm über $\mathbb{K} = \mathbb{R}, \mathbb{C}$

Sei V ein Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Ein **Skalarprodukt** auf V ist eine Abbildung $\langle - | - \rangle : V \times V \rightarrow \mathbb{K}$, die (S1–3) erfüllt.

Das Paar $(V, \langle - | - \rangle)$ heißt dann **\mathbb{K} -Vektorraum mit Skalarprodukt** oder **euklidischer \mathbb{R} -Vektorraum** bzw. **unitärer \mathbb{C} -Vektorraum**,

In der Analysis heißt $(V, \langle - | - \rangle)$ auch **Prä-Hilbert-Raum** über \mathbb{K} und bei metrischer Vollständigkeit schließlich **Hilbert-Raum** über \mathbb{K} .

Eine **Norm** auf V ist eine Abbildung $\| - \| : V \rightarrow \mathbb{R}_{\geq 0}$, die (N1–3) erfüllt.

Das Paar $(V, \| - \|)$ heißt dann **normierter \mathbb{K} -Vektorraum** oder auch **Prä-Banach-Raum** und bei Vollständigkeit **Banach-Raum** über \mathbb{K} .

Für ein **semidefinites Skalarprodukt** fordern wir nur (S0,2,3).

Für eine **Seminorm** oder **Halbnorm** fordern wir nur (N0,2,3).

Beide Abschwächungen kommen in Anwendungen natürlich vor.

Orthogonale und orthonormale Familien

Definition P1M: Orthonormalbasis

Sei V ein \mathbb{K} -Vektorraum mit Skalarprodukt $\langle - | - \rangle$ und Norm $\|-\|$.

Eine Familie $(u_i)_{i \in I}$ in V heißt **orthogonal**, falls $\langle u_i | u_j \rangle = 0$ für alle $i \neq j$ in I gilt, und **orthonormal**, falls zudem $\|u_i\| = 1$ für alle $i \in I$ gilt:

$$\langle u_i | u_j \rangle = \begin{cases} 0 & \text{falls } i \neq j, \\ 1 & \text{falls } i = j. \end{cases}$$

Ist $(u_i)_{i \in I}$ zudem eine Basis von V , so nennen wir dies eine **Orthogonalbasis** bzw. eine **Orthonormalbasis** (kurz ONB).

Daraus können wir das Skalarprodukt rekonstruieren gemäß

$$\left\langle \sum_{i \in I} a_i u_i \mid \sum_{j \in I} b_j u_j \right\rangle \stackrel{\text{ONB}}{=} \sum_{k \in I} \overline{a_k} b_k.$$

😊 Jede orthogonale Familie in $V \setminus \{0\}$ ist linear unabhängig (P10).
Zur Basis fehlt dann nur noch, dass $(u_i)_{i \in I}$ den Vektorraum V erzeugt.

Satz des Pythagoras

Satz P1N: Pythagoras

Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}, \mathbb{C}$ mit Skalarprodukt $\langle - | - \rangle$.

(1) Sind $u_1, \dots, u_n \in V$ orthogonal, also $\langle u_k | u_\ell \rangle = 0$ für $k \neq \ell$, so gilt

$$\|u_1 + \dots + u_n\|^2 = \|u_1\|^2 + \dots + \|u_n\|^2.$$

(2) Sind $e_1, \dots, e_n \in V$ orthonormal und $c_1, \dots, c_n \in \mathbb{K}$, so gilt demnach

$$\|c_1 e_1 + \dots + c_n e_n\|^2 = |c_1|^2 + \dots + |c_n|^2.$$

Beweis: (1) Das Normquadrat erhalten wir aus dem Skalarprodukt:

$$\|\sum_k u_k\|^2 = \langle \sum_k u_k | \sum_\ell u_\ell \rangle = \sum_k \sum_\ell \langle u_k | u_\ell \rangle = \sum_k \|u_k\|^2$$

(2) Für $u_k = c_k e_k$ mit $\|e_k\| = 1$ gilt $\|u_k\|^2 = |c_k|^2 \|e_k\|^2 = |c_k|^2$. ◻

😊 Der klassische Satz des Pythagoras ist der erste interessante Fall, nämlich die Ebene $V = \mathbb{R}^2$ über $\mathbb{K} = \mathbb{R}$ mit euklidischem Skalarprodukt.

Satz P10: Fourier–Koeffizienten

Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}, \mathbb{C}$ mit Skalarprodukt $\langle - | - \rangle$.

Sei $(u_k)_{k \in I}$ eine orthogonale Familie von Vektoren $u_k \in V \setminus \{0\}$.

(1) Für jede Linearkombination $v = \sum_{\ell} c_{\ell} u_{\ell}$ über \mathbb{K} gilt dann

$$c_k = \frac{\langle u_k | v \rangle}{\langle u_k | u_k \rangle}.$$

(2) Insbesondere ist die Familie $(u_k)_{k \in I}$ linear unabhängig.

Beweis: (1) Die Koeffizientenformel folgt direkt aus der Orthogonalität:

$$\langle u_k | v \rangle = \langle u_k | \sum_{\ell} c_{\ell} u_{\ell} \rangle = \sum_{\ell} c_{\ell} \langle u_k | u_{\ell} \rangle = c_k \langle u_k | u_k \rangle.$$

(2) Gilt $v = 0$, so folgt $c_k = 0$ für alle $k \in I$.

◻

😊 Das Skalarprodukt filtert den gewünschten Koeffizienten heraus!
In allen Rechnungen ist das überaus praktisch und hilfreich.

Kanonisches Beispiel: das euklidische Skalarprodukt

Wir betrachten weiterhin die beiden Grundkörper $\mathbb{K} = \mathbb{R}, \mathbb{C}$ gemeinsam. Dazu statten wir den Körper \mathbb{R} aus mit der Identität $\bar{\cdot} : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x$ und den Körper \mathbb{C} mit der Konjugation $\bar{\cdot} : \mathbb{C} \rightarrow \mathbb{C} : (x, y) \mapsto (x, -y)$.

Beispiel P1P: die Produktsumme auf \mathbb{K}^n

(1) Auf dem \mathbb{K} -Vektorraum \mathbb{K}^n haben wir das euklidische Skalarprodukt

$$\langle u | v \rangle = \sum_{k=1}^n \bar{u}_k v_k.$$

Somit gilt $\langle u | v \rangle = u^T v$ über \mathbb{R} und entsprechend $\langle u | v \rangle = \bar{u}^T v$ über \mathbb{C} . Die Standardbasis (e_1, \dots, e_n) wird hierdurch zur Orthonormalbasis.

Zu jeder Matrix $A = (a_{ij})_{ij}$ in $\mathbb{C}^{m \times n}$ ist $A^\dagger = A^H = \bar{A}^T = (\bar{a}_{ij})_{ji}$ in $\mathbb{C}^{n \times m}$ die **transponiert-konjugierte Matrix** (G2K). Wir erhalten die Abbildung

$$\dagger : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{n \times m} : A \mapsto A^\dagger.$$

Dies nennt man auch die **hermitesch transponierte Matrix** oder auch die **adjungierte Matrix** (nicht zu verwechseln mit der Adjunkten L2S).

Kanonisches Beispiel: das euklidische Skalarprodukt

Beispiel P1P: die Produktsumme auf $\mathbb{K}^{(I)}$

(2) Sei I eine beliebige Menge, egal ob endlich oder unendlich.

Auf dem \mathbb{K} -Vektorraum $\mathbb{K}^{(I)}$ haben wir das euklidische Skalarprodukt

$$\langle u \mid v \rangle_{\ell^2} := \sum_{k \in I} \overline{u_k} v_k.$$

Die Standardbasis $(e_i)_{i \in I}$ wird hierdurch zur Orthonormalbasis.

Die Vervollständigung $\ell^2(I, \mathbb{K})$ untersuchen Sie in der Analysis.

(3) Jeder \mathbb{K} -Vektorraum V besitzt eine Basis $(b_i)_{i \in I}$ für eine geeignete Indexmenge I (dank Satz J2B) und erlaubt somit das Skalarprodukt

$$\langle \sum_{i \in I} u_i b_i \mid \sum_{j \in I} v_j b_j \rangle := \sum_{k \in I} \overline{u_k} v_k.$$

Die gewählte Basis $(b_i)_{i \in I}$ wird hierdurch zur Orthonormalbasis.

Bemerkung: Die Summe ist jeweils endlich, denn nach Voraussetzung sind die beiden Träger $\text{supp}(u)$ und $\text{supp}(v)$ in I endlich, also hat auch $w : k \mapsto \overline{u_k} v_k$ endlichen Träger, genauer $\text{supp}(w) = \text{supp}(u) \cap \text{supp}(v)$.

Kanonisches Beispiel: das euklidische Skalarprodukt

Aufgabe: Rechnen Sie die hier gemachten Aussagen sorgsam nach.

Lösung: Hermitizität (S2) ist klar, denn die Konjugation $\bar{\cdot} : \mathbb{K} \rightarrow \mathbb{K}$ ist ein Körperautomorphismus, also insbesondere verträglich mit der Addition:

$$\overline{\langle u | v \rangle} = \overline{\sum_{k \in I} \bar{u}_k v_k} = \sum_{k \in I} \overline{\bar{u}_k v_k} = \sum_{k \in I} \bar{v}_k u_k = \langle v | u \rangle$$

Linearität (S3) ist ebenfalls klar, denn die Summe ist linear in v :

$$\begin{aligned} \langle u | \lambda v + \mu w \rangle &= \sum_{k \in I} \bar{u}_k (\lambda v_k + \mu w_k) \\ &= \sum_{k \in I} \lambda (\bar{u}_k v_k) + \mu (\bar{u}_k w_k) \\ &= \lambda \sum_{k \in I} (\bar{u}_k v_k) + \mu \sum_{k \in I} (\bar{u}_k w_k) \\ &= \lambda \langle u | v \rangle + \mu \langle u | w \rangle \end{aligned}$$

(S0) Für jeden Vektor $u \in \mathbb{K}^{(I)}$ gilt $\langle u | u \rangle = \sum_{i \in I} |u_i|^2 \geq 0$.

(S1) Im Falle $u \neq 0$ gilt $u_i \neq 0$ für mindestens ein $i \in I$.

Daraus folgt sofort die strikte Ungleichung $\langle u | u \rangle \geq |u_i|^2 > 0$.

Analytisches Beispiel: das Produktintegral

Beispiel P1Q: das Produktintegral auf $\mathcal{C}([a, b], \mathbb{K})$

Sei $a < b$ in \mathbb{R} und $V = \mathcal{C}([a, b], \mathbb{K})$ der \mathbb{K} -Vektorraum aller stetigen Funktionen $f, g: [a, b] \rightarrow \mathbb{K}$. Hierauf haben wir das Skalarprodukt

$$\langle f | g \rangle_{L^2} = \int_{t=a}^b \overline{f(t)} g(t) dt.$$

Beweis: Die Eigenschaften (S2,3) sind klar, denn das Integral ist linear.

Auch Positivität (S0) ist klar, denn es gilt $\langle f | f \rangle = \int_{t=a}^b |f(t)|^2 dt \geq 0$.

Für positive Definitheit (S1) müssen wir jedoch genauer hinschauen!

Sei $f \neq 0$, also $f(t_0) \neq 0$ für ein $t_0 \in [a, b]$, und somit $\delta := |f(t_0)|^2 > 0$.

Wir können $t_0 \in]a, b[$ annehmen, die Randfälle $t_0 \in \{a, b\}$ sind analog.

Dank der Stetigkeit von f existiert ein hinreichend kleines $\varepsilon \in \mathbb{R}_{>0}$,

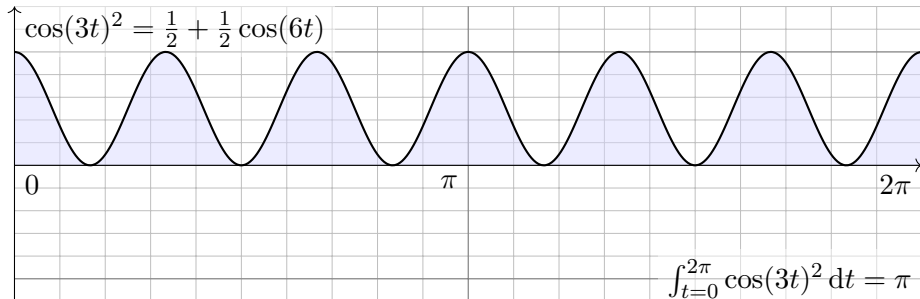
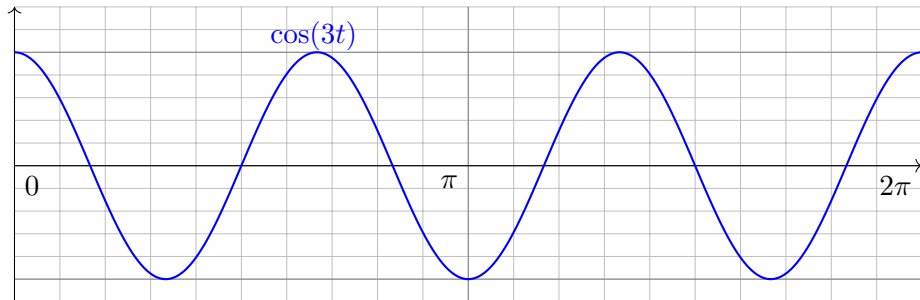
sodass $|f(t)|^2 \geq \delta/2$ für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon] \subseteq [a, b]$ gilt.

Daraus folgt $\langle f | f \rangle = \int_{t=a}^b |f(t)|^2 dt \geq \varepsilon \delta > 0$.

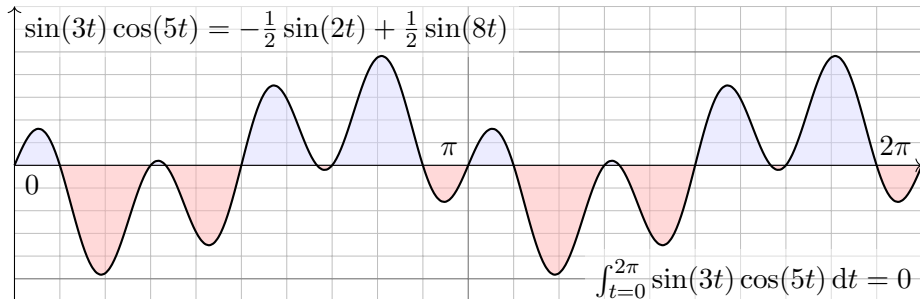
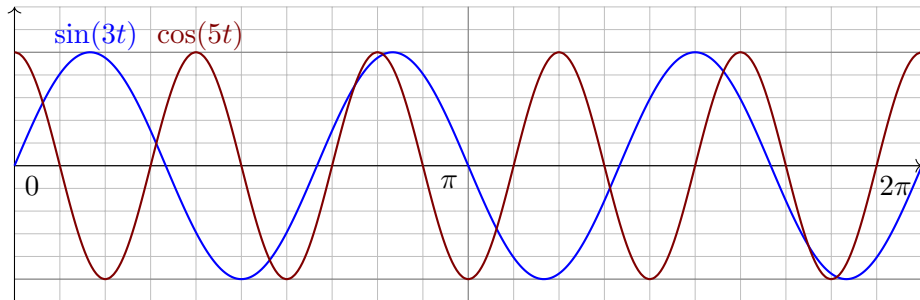
QED

😊 Die Vervollständigung $L^2([a, b], \mathbb{K})$ dieses Raumes untersuchen Sie in der Analysis und nutzen dies insbesondere für die Fourier-Theorie.

Beispiel: Orthogonalität trigonometrischer Funktionen



Beispiel: Orthogonalität trigonometrischer Funktionen



Beispiel: Orthogonalität trigonometrischer Funktionen

Aufgabe: Zur Periode $T > 0$ ist die Grundfrequenz $\omega = 2\pi/T$.

(1) Integrieren Sie für $k, \ell \in \mathbb{N}$ folgende Funktionen über $[0, T]$:

$$\cos(k\omega t) \cos(\ell\omega t), \quad \sin(k\omega t) \sin(\ell\omega t), \quad \sin(k\omega t) \cos(\ell\omega t).$$

Wir erinnern hierzu an die stets nützlichen Additionstheoreme

$$\cos(\alpha) \cos(\beta) = \frac{1}{2} [\cos(\alpha - \beta) + \cos(\alpha + \beta)],$$

$$\sin(\alpha) \sin(\beta) = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)],$$

$$\sin(\alpha) \cos(\beta) = \frac{1}{2} [\sin(\alpha - \beta) + \sin(\alpha + \beta)].$$

(0) Leiten Sie diese Additionstheoreme her aus der Euler-Formel

$$e^{i\alpha} = \cos \alpha + i \sin \alpha \quad \text{und dem Exponentialgesetz } e^{i\alpha+i\beta} = e^{i\alpha} e^{i\beta}.$$

Lösung: (1) Zur Berechnung nutzen wir die Grundintegrale

$$\int_{t=0}^T \sin(n\omega t) dt = 0 \quad \text{für alle } n \in \mathbb{Z},$$

$$\int_{t=0}^T \cos(n\omega t) dt = \begin{cases} 0 & \text{für } n \neq 0, \\ T & \text{für } n = 0. \end{cases}$$

Beispiel: Orthogonalität trigonometrischer Funktionen

Die Familie $1, \cos(\omega t), \sin(\omega t), \cos(2\omega t), \sin(2\omega t), \cos(3\omega t), \sin(3\omega t), \dots$ ist orthogonal bezüglich des Skalarprodukts $\langle f | g \rangle = \int_{t=0}^T f(t)g(t) dt$:

$$\int_{t=0}^T \cos(k\omega t) \cos(l\omega t) dt = \frac{1}{2} \int_{t=0}^T \cos((k-l)\omega t) + \cos((k+l)\omega t) dt$$

$$= \begin{cases} 0 & \text{falls } k \neq l, \\ T/2 & \text{falls } k = l \geq 1, \\ T & \text{falls } k = l = 0. \end{cases}$$

$$\int_{t=0}^T \sin(k\omega t) \sin(l\omega t) dt = \frac{1}{2} \int_{t=0}^T \cos((k-l)\omega t) - \cos((k+l)\omega t) dt$$

$$= \begin{cases} 0 & \text{falls } k \neq l, \\ T/2 & \text{falls } k = l \geq 1, \\ 0 & \text{falls } k = l = 0. \end{cases}$$

$$\int_{t=0}^T \sin(k\omega t) \cos(l\omega t) dt = \frac{1}{2} \int_{t=0}^T \sin((k-l)\omega t) + \sin((k+l)\omega t) dt = 0$$

😊 Das ist schön. Alles wird noch schöner und übersichtlicher für die komplexe Funktion $e^{ik\omega t} = \cos(k\omega t) + i \sin(k\omega t)$, siehe nächste Aufgabe.

Beispiel: die trigonometrische Orthonormalbasis

Für Funktionen $f, g: \mathbb{R} \rightarrow \mathbb{C}$ mit Periode T nutzen wir das Skalarprodukt

$$\langle f | g \rangle := \frac{1}{T} \int_{t=0}^T \overline{f(t)} g(t) dt.$$

Sei $\omega = 2\pi/T$. Als **Basisfunktion** $e_k: \mathbb{R} \rightarrow \mathbb{C}$ mit $k \in \mathbb{Z}$ definieren wir

$$e_k(t) := e^{ik\omega t} = \cos(k\omega t) + i \sin(k\omega t).$$

Ihre Linearkombinationen nennen wir **Fourier-Polynome**:

$$f(t) = \sum_{k=-n}^n \hat{f}(k) e^{ik\omega t}, \quad g(t) = \sum_{\ell=-n}^n \hat{g}(\ell) e^{i\ell\omega t} \quad \text{mit} \quad \hat{f}(k), \hat{g}(\ell) \in \mathbb{C}.$$

Aufgabe: Wie bestimmt die Funktion $f: \mathbb{R} \rightarrow \mathbb{C}$ ihr Spektrum $\hat{f}: \mathbb{Z} \rightarrow \mathbb{C}$?

Wir nutzen Orthonormalität: Berechnen Sie hierzu die Skalarprodukte

(0) $\langle 1 | e_n \rangle$, (1) $\langle e_k | e_\ell \rangle$, (2) $\langle e_k | g \rangle$, (3) $\langle f | g \rangle$, (4) $\langle f | f \rangle$.

(5) Entwickeln Sie $f(t) = \sin^2 t$ und $g(t) = \cos^3 t$ in Fourier-Polynome.

(6) Berechnen Sie daraus $\frac{1}{2\pi} \int_{t=0}^{2\pi} \cos^4 t dt$ und $\frac{1}{2\pi} \int_{t=0}^{2\pi} \cos^6 t dt$.

Beispiel: die trigonometrische Orthonormalbasis

Lösung: (0) Wir berechnen $\langle 1 | e_n \rangle$. Für $n = 0$ ist es besonders leicht:

$$\langle 1 | e_0 \rangle \stackrel{\text{Def}}{=} \frac{1}{T} \int_{t=0}^T 1 \cdot e^{i0\omega t} dt = \frac{1}{T} \int_{t=0}^T 1 dt = 1.$$

Für $n \in \mathbb{Z} \setminus \{0\}$ nutzen wir die vorige Aufgabe oder den HDI:

$$\langle 1 | e_n \rangle \stackrel{\text{Def}}{=} \frac{1}{T} \int_{t=0}^T 1 \cdot e^{in\omega t} dt \stackrel{\text{HDI}}{=} \frac{1}{T} \left[\frac{1}{in\omega} e^{in\omega t} \right]_{t=0}^T = 0.$$

(1) **Orthonormalität** — Wir berechnen die gesuchten Skalarprodukte:

$$\begin{aligned} \langle e_k | e_\ell \rangle &\stackrel{\text{Def}}{=} \frac{1}{T} \int_{t=0}^T \overline{e_k(t)} e_\ell(t) dt \stackrel{\text{Def}}{=} \frac{1}{T} \int_{t=0}^T e^{-ik\omega t} e^{i\ell\omega t} dt \\ &\stackrel{\text{Exp}}{=} \frac{1}{T} \int_{t=0}^T e^{i(\ell-k)\omega t} dt \stackrel{(0)}{=} \begin{cases} 1 & \text{für } k = \ell, \\ 0 & \text{für } k \neq \ell. \end{cases} \end{aligned}$$

😊 Die Basis $(e_k)_{k \in \mathbb{Z}}$ ist orthonormal bezüglich des Skalarprodukts! Das ist analog zur Geometrie des euklidischen Raumes \mathbb{R}^n bzw. \mathbb{C}^n .

😊 Im Komplexen ist alles halb so schwer und doppelt so schön!

Beispiel: die trigonometrische Orthonormalbasis

(2) **Fourier** — Dank Linearität und Orthonormalität erhalten wir:

$$\langle e_k | g \rangle \stackrel{\text{Def}}{=} \left\langle e_k \left| \sum_{\ell=-n}^n \widehat{g}(\ell) e_\ell \right. \right\rangle \stackrel{\text{Lin}}{=} \sum_{\ell=-n}^n \widehat{g}(\ell) \langle e_k | e_\ell \rangle \stackrel{(1)}{=} \widehat{g}(k)$$

😊 Das Skalarprodukt filtert den gewünschten Koeffizienten heraus!

(3) **Parseval** — Dank Bilinearität und Orthonormalität erhalten wir:

$$\begin{aligned} \langle f | g \rangle &\stackrel{\text{Def}}{=} \left\langle \sum_{k=-n}^n \widehat{f}(k) e_k \left| \sum_{\ell=-n}^n \widehat{g}(\ell) e_\ell \right. \right\rangle \stackrel{\text{Lin}}{=} \sum_{k=-n}^n \overline{\widehat{f}(k)} \left\langle e_k \left| \sum_{\ell=-n}^n \widehat{g}(\ell) e_\ell \right. \right\rangle \\ &\stackrel{\text{Lin}}{=} \sum_{k=-n}^n \sum_{\ell=-n}^n \overline{\widehat{f}(k)} \widehat{g}(\ell) \langle e_k | e_\ell \rangle \stackrel{(1)}{=} \sum_{k=-n}^n \overline{\widehat{f}(k)} \widehat{g}(k). \end{aligned}$$

😊 Diese Rechnung gilt allgemein für Orthonormalbasen.

(4) **Energiegleichung** — Für das Normquadrat gilt Pythagoras (P1N):

$$\langle f | f \rangle \stackrel{(3)}{=} \sum_{k=-n}^n |\widehat{f}(k)|^2$$

😊 Das Normquadrat ist die Summe der Koeffizientenquadrate.

Beispiel: die trigonometrische Orthonormalbasis

(5) Wir entwickeln f und g dank der Euler-Formel $e^{it} = \cos t + i \sin t$:

$$f(t) = \sin(t)^2 = \left(\frac{e^{it} - e^{-it}}{2i} \right)^2 = -\frac{1}{4} e^{2it} + \frac{1}{2} - \frac{1}{4} e^{-2it} = \frac{1}{2} - \frac{1}{2} \cos(2t)$$

$$\begin{aligned} g(t) = \cos(t)^3 &= \left(\frac{e^{it} + e^{-it}}{2} \right)^3 = \frac{1}{8} e^{3it} + \frac{3}{8} e^{it} + \frac{3}{8} e^{-it} + \frac{1}{8} e^{-3it} \\ &= \frac{3}{4} \cos(t) + \frac{1}{4} \cos(3t) \end{aligned}$$

😊 Dank Orthonormalität lesen wir die Fourier-Koeffizienten ab (2).

(6) Wir nutzen die Energiegleichung (4) und Fourier-Koeffizienten (5):

$$\frac{1}{2\pi} \int_{t=0}^{2\pi} \sin^4 t \, dt \stackrel{\text{Def}}{=} \langle f | f \rangle \stackrel{(4)}{=} \sum_{k=-n}^n |\hat{f}(k)|^2 \stackrel{(5)}{=} \frac{3}{8}$$

$$\frac{1}{2\pi} \int_{t=0}^{2\pi} \cos^6 t \, dt \stackrel{\text{Def}}{=} \langle g | g \rangle \stackrel{(4)}{=} \sum_{k=-n}^n |\hat{g}(k)|^2 \stackrel{(5)}{=} \frac{5}{16}$$

😊 Die Energiegleichung gilt allgemein für Fourier-Reihen!

Die trigonometrische Orthonormalbasis

Satz P1R: trigonometrische Orthonormalbasis

Die Menge aller Funktionen $f: \mathbb{R} \rightarrow \mathbb{C}$ ist ein \mathbb{C} -Vektorraum. Hierin ist die Teilmenge aller T -periodischen Funktionen ein Untervektorraum. Als Basisfunktion $e_k: \mathbb{R} \rightarrow \mathbb{C}$ mit $k \in \mathbb{Z}$ und $\omega = 2\pi/T$ definieren wir

$$e_k(t) := e^{ik\omega t} = \cos(k\omega t) + i \sin(k\omega t).$$

Diese erzeugen den Unterraum $V = \{ \sum_{k=-n}^n c_k e^{ik\omega t} \mid n \in \mathbb{N}, c_k \in \mathbb{C} \}$ der Fourier-Polynome. Hierauf haben wir das Skalarprodukt

$$V \times V \rightarrow \mathbb{C} : (f, g) \mapsto \langle f \mid g \rangle := \frac{1}{T} \int_{t=0}^T \overline{f(t)} g(t) dt.$$

Damit gelten die Orthonormalitätsrelationen

$$\langle e_k \mid e_\ell \rangle = \begin{cases} 0 & \text{für } k \neq \ell: & \text{paarweise Orthogonalität,} \\ 1 & \text{für } k = \ell: & \text{Normierung auf Länge 1.} \end{cases}$$

Die trigonometrische Orthonormalbasis

Korollar P1s: Fourier–Koeffizienten durch Skalarprodukt

(1) Wir betrachten ein trigonometrisches Polynom:

$$f(t) = \sum_{\ell=-n}^n c_{\ell} e^{i\ell\omega t} = \frac{a_0}{2} + \sum_{\ell=1}^n a_{\ell} \cos(\ell\omega t) + b_{\ell} \sin(\ell\omega t)$$

Die Funktion f bestimmt die Koeffizienten durch Fourier–Integrale:

$$c_k = \langle e_k \mid f \rangle = \frac{1}{T} \int_{t=0}^T e^{-ik\omega t} f(t) dt,$$

bzw.

$$a_k = \langle 2 \cos(k\omega t) \mid f \rangle = \frac{2}{T} \int_{t=0}^T \cos(k\omega t) f(t) dt,$$

$$b_k = \langle 2 \sin(k\omega t) \mid f \rangle = \frac{2}{T} \int_{t=0}^T \sin(k\omega t) f(t) dt.$$

😊 Die Formeln für die Koeffizienten c_k sind besonders schön, da die Funktionen $e_k(t) = e^{ik\omega t}$ orthonormal sind. Hingegen sind $\cos(k\omega t)$ und $\sin(k\omega t)$ zwar orthogonal, aber mit L^2 –Norm $\sqrt{2}/2$ statt Normierung 1.

Die trigonometrische Orthonormalbasis

- 😊 Das Skalarprodukt beschert uns Struktur, Klarheit und Übersicht. Die Orthonormalität der Basis $(e_k)_{k \in \mathbb{Z}}$ vereinfacht die Rechnung.
- 😊 Das Fourier-Integral filtert den gewünschten Koeffizienten heraus! Diese Gleichungen nutzen wir ebenso für Fourier-Reihen ($n = \infty$).

Korollar P1s: Jede Funktion bestimmt ihre Koeffizienten.

(2) Die Funktionen $f, g: \mathbb{R} \rightarrow \mathbb{C}$ seien gegeben als Fourier-Polynome

$$f(t) = \sum_{k=-n}^n \hat{f}(k) e^{ik\omega t} \quad \text{und} \quad g(t) = \sum_{k=-n}^n \hat{g}(k) e^{ik\omega t}.$$

Aus $\hat{f}(k) = \hat{g}(k)$ für alle $k = -n, \dots, n$ folgt offensichtlich $f = g$.

Umgekehrt folgt aus $f = g$ auch $\hat{f} = \hat{g}$, dank der Fourier-Integrale:

$$\hat{f}(k) = \frac{1}{T} \int_{t=0}^T e^{-ik\omega t} f(t) dt = \frac{1}{T} \int_{t=0}^T e^{-ik\omega t} g(t) dt = \hat{g}(k)$$

- 😊 Für Letzteres genügt bereits Gleichheit $f = g$ fast überall.

Die trigonometrische Orthonormalbasis

Korollar P1s: Norm und Skalarprodukt

(3) Koeffizienten $\hat{f}(k), \hat{g}(k) \in \mathbb{C}$ definieren Fourier–Polynome

$$f(t) = \sum_{k=-n}^n \hat{f}(k) e^{ik\omega t} \quad \text{und} \quad g(t) = \sum_{k=-n}^n \hat{g}(k) e^{ik\omega t}.$$

Für ihre Norm und ihr Skalarprodukt gilt nach Pythagoras (P1N)

$$\frac{1}{T} \int_{t=0}^T |f(t)|^2 dt = \sum_{k=-n}^n |\hat{f}(k)|^2, \quad \text{kurz} \quad \|f\|_{L^2} = \|\hat{f}\|_{\ell^2},$$

$$\frac{1}{T} \int_{t=0}^T \overline{f(t)} g(t) dt = \sum_{k=-n}^n \overline{\hat{f}(k)} \hat{g}(k), \quad \text{kurz} \quad \langle f | g \rangle_{L^2} = \langle \hat{f} | \hat{g} \rangle_{\ell^2}.$$

😊 Diese Isometrie ist eine zentrale Eigenschaft der Fourier–Theorie. Für Fourier–Polynome folgt dies direkt aus der Orthonormalität der Basis $(e_k)_{k \in \mathbb{Z}}$. Erfreulicherweise gilt dies nach Vervollständigung sogar allgemein für alle quadrat-integrierbaren Funktionen!

Orthonormalisierung nach Gram–Schmidt

Wie können wir aus einer Basis eine Orthonormalbasis konstruieren?

Satz P2A: Laplace 1816, Gram 1883, Schmidt 1907

Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}, \mathbb{C}$ mit Skalarprodukt $\langle - | - \rangle$.

Sei $b_1, \dots, b_n \in V$ eine Basis des Unterraums $U_n = \langle b_1, \dots, b_n \rangle_{\mathbb{K}}$.

(1) Daraus erhalten wir rekursiv die Orthogonalbasis u_1, \dots, u_n durch

$$u_n := b_n - \sum_{k=1}^{n-1} u_k \lambda_k \quad \text{mit} \quad \lambda_k = \frac{\langle u_k | b_n \rangle}{\langle u_k | u_k \rangle}.$$

(2) Optional können wir u_n ersetzen durch $u'_n = u_n \mu_n$ mit $\mu_n \in \mathbb{K}^\times$.

(3) Normiert zu $e_k := u_k / \|u_k\|$ erhalten wir eine Orthonormalbasis:

$$\langle u_k | u_\ell \rangle = \begin{cases} 0 & \text{für } k \neq \ell, \\ \|u_k\|^2 > 0 & \text{für } k = \ell, \end{cases} \quad \text{und} \quad \langle e_k | e_\ell \rangle = \begin{cases} 0 & \text{für } k \neq \ell, \\ 1 & \text{für } k = \ell. \end{cases}$$

Dasselbe Verfahren gelingt für jede abzählbare Basis $(b_k)_{k \in \mathbb{N}}$ von V .

Orthonormalisierung nach Gram–Schmidt

Aufgabe: Alles steht explizit da. Rechnen Sie es sorgsam nach!

Lösung: (1) Wir führen Induktion über n . Für $n = 1$ ist die Aussage klar.

Sei nun $n \geq 2$. Im Unterraum $U_{n-1} \leq V$ haben wir die gegebene Basis (b_1, \dots, b_{n-1}) bereits zur Orthogonalbasis (u_1, \dots, u_{n-1}) transformiert.

Im Unterraum U_n erhalten wir aus der Basis $(b_1, \dots, b_{n-1}, b_n)$ zunächst $(u_1, \dots, u_{n-1}, b_n)$ und dann $(u_1, \dots, u_{n-1}, u_n)$ mit $u_n := b_n - \sum_{k=1}^{n-1} u_k \lambda_k$.

Für alle $j = 1, \dots, n-1$ gilt $\langle u_j | u_n \rangle = \langle u_j | b_n \rangle - \langle u_j | u_j \rangle \lambda_j \stackrel{!}{=} 0$.

Dies verschwindet genau für $\lambda_j = \langle u_j | b_n \rangle / \langle u_j | u_j \rangle$. Voilà!

(2) Basiseigenschaft und Orthogonalität bleiben nach Skalierung von u_n zu $u'_n = u_n \mu_n$ mit $\mu_n \in \mathbb{K}^\times$. Das verschafft uns zusätzlichen Spielraum.

(3) Da (u_1, \dots, u_n) eine Basis von U_n ist, gilt insbesondere $u_n \neq 0$. Normierung zu $e_n := u_n / \|u_n\|$ liefert also eine Orthonormalbasis.

Dasselbe Verfahren gelingt für jede abzählbare Basis $(b_k)_{k \in \mathbb{N}}$:

Aus b_0, b_1, b_2, \dots konstruieren wir u_0, u_1, u_2, \dots und e_0, e_1, e_2, \dots QED

Orthonormalisierung nach Gram–Schmidt

Aufgabe: Vergleichen Sie das Gram–Schmidt–Verfahren P2A mit dem Diagonalisierungsverfahren O2A für symmetrische Bilinearformen. Was ist dabei gleich? Wo liegt der entscheidende Unterschied?

Lösung: Das Gram–Schmidt–Verfahren P2A verläuft wörtlich genau so wie das allgemeine Verfahren O2A für jede symmetrische Bilinearform!

😊 Dies wird stark vereinfacht durch die Garantie $\langle u_n | u_n \rangle > 0$, daher ist nun keine Fallunterscheidung zu $\langle u_n | u_n \rangle = 0$ mehr nötig.

😊 Zudem können wir in \mathbb{R} die Quadratwurzel aus $\langle u_n | u_n \rangle$ ziehen, sodass wir im letzten Schritt zu $e_n = u_n / \|u_n\|$ normieren können.

Aufgabe: Formulieren Sie eine Verallgemeinerung dieses Verfahrens, die als Eingabe ein beliebiges Erzeugendensystem b_1, \dots, b_m von U erlaubt und daraus wie zuvor eine Orthonormalbasis konstruiert.

Lösung: Wir wenden das Gram–Schmidt–Verfahren P2A an; im Falle $u_n = 0$ löschen wir b_n aus dem Erzeugendensystem.

😊 Der folgende Algorithmus präzisiert die Buchführung der Indizes.

Orthonormalisierung nach Gram–Schmidt

😊 Aus jedem Erzeugendensystem können wir eine ONB konstruieren:

Algo P2A: Orthonormalisierung nach Gram–Schmidt

Eingabe: ein Erzeugendensystem $\mathcal{B} = (b_1, \dots, b_m)$ des Unterraums U

Ausgabe: eine Orthonormalbasis $\mathcal{E} = (e_1, \dots, e_n)$ des Unterraums U

```

1:  $n \leftarrow 0$ 
2: for  $\ell$  from 1 to  $m$  do
3:    $n \leftarrow n + 1$ ;  $b'_n \leftarrow b_\ell$ ;  $u_n \leftarrow b'_n - \sum_{k=1}^{n-1} u_k \langle u_k | b'_n \rangle / \langle u_k | u_k \rangle$ 
4:   if  $u_n = 0$  then  $n \leftarrow n - 1$  else  $e_n \leftarrow u_n / \|u_n\|$ 
5: return  $(e_1, \dots, e_n)$ 

```

Dieser Algorithmus berechnet aus \mathcal{B} drei interessante Basen von U :

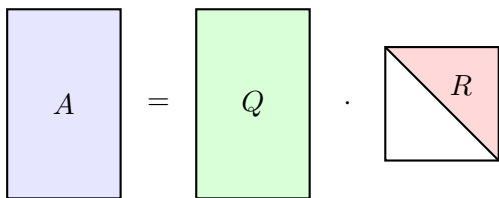
(1) die Teilfamilie $\mathcal{B}' = (b'_1, \dots, b'_n)$ wie im Basisauswahlsatz J2B,

(2) hieraus abgeleitet die Orthogonalbasis $\mathcal{U} = (u_1, \dots, u_n)$

(3) und daraus die Orthonormalbasis $\mathcal{E} = (e_1, \dots, e_n)$.

Übung: Sei $V = \mathbb{K}^d$ mit dem Standardskalarprodukt. Zählen Sie die Operationen in \mathbb{K} . Vergleichen Sie (1) mit dem Gauß–Algorithmus B2c.

Die QR-Zerlegung



Satz P2B: Existenz und Eindeutigkeit der QR-Zerlegung

Gegeben sei eine Matrix $A \in \mathbb{K}^{m \times n}$ mit linear unabhängigen Spalten, also $\text{rang } A = n \leq m$. Dann existiert genau eine Zerlegung der Form

$$A = QR$$

wobei $Q \in \mathbb{K}^{m \times n}$ orthonormale Spalten hat, also $Q^\dagger Q = I_n$ erfüllt, und $R \in \mathbb{K}^{n \times n}$ eine obere Dreiecksmatrix ist, also $r_{i,j} = 0$ für $i > j$, und zudem positive Diagonaleinträge hat, also $r_{i,i} > 0$ für alle i .

Insbesondere gilt $\det R > 0$ und $R \in \text{GL}_n \mathbb{K}$.

Aufgabe: Folgern Sie dies aus dem Gram-Schmidt-Verfahren P2A.

Die QR-Zerlegung

Lösung: Wir beweisen die **Existenz** von (Q, R) durch Konstruktion.

Dank Gram-Schmidt konstruieren wir aus den Spalten $a_1, \dots, a_n \in \mathbb{K}^m$ der Matrix A eine orthonormale Familie $q_1, \dots, q_n \in \mathbb{K}^m$ wie folgt:

$$\begin{aligned} b_1 &:= a_1, & q_1 &:= b_1 / \|b_1\|, \\ b_2 &:= a_2 - q_1 \langle q_1 | a_2 \rangle, & q_2 &:= b_2 / \|b_2\|, \\ &\vdots & &\vdots \\ b_n &:= a_n - \sum_{k=1}^{n-1} q_k \langle q_k | a_n \rangle, & q_n &:= b_n / \|b_n\|. \end{aligned}$$

Umgekehrt gelesen erhalten wir daraus $A = QR$, denn

$$\begin{aligned} a_1 &= q_1 \|b_1\|, \\ a_2 &= q_2 \|b_2\| + q_1 \langle q_1 | a_2 \rangle, \\ &\vdots \\ a_n &= q_n \|b_n\| + \sum_{k=1}^{n-1} q_k \langle q_k | a_n \rangle. \end{aligned}$$

Dabei ist $Q = (q_1, \dots, q_n) \in \mathbb{K}^{m \times n}$ orthonormal und $R \in \text{GL}_n \mathbb{K}$ mit $r_{i,i} = \|b_i\| > 0$ sowie $r_{i,j} = 0$ für $i > j$ und $r_{i,j} = \langle q_i | a_j \rangle$ für $i < j$.

Die QR-Zerlegung

Die **Eindeutigkeit** von (Q, R) folgt ebenso per Induktion über n . Angenommen, es gilt $A = QR$, in den Spalten von A und Q also

$$a_1 = q_1 r_{1,1}$$

$$a_2 = q_1 r_{1,2} + q_2 r_{2,2}$$

$$\vdots$$

$$a_n = q_1 r_{1,n} + \cdots + q_{n-1} r_{n-1,n} + q_n r_{n,n}$$

Hierbei seien $q_1, q_2, \dots, q_n \in \mathbb{K}^m$ orthonormal und $r_{i,i} > 0$ für alle i . Dann sind (Q, R) die Daten aus dem Gram–Schmidt–Verfahren:

Für $k < n$ folgt $\langle q_k | a_n \rangle = \langle q_k | q_n \rangle r_{n,n} + \sum_{\ell=1}^{n-1} \langle q_k | q_\ell \rangle r_{\ell,n} = r_{k,n}$.

Wir setzen $b_n := a_n - \sum_{k=1}^{n-1} q_k \langle q_k | a_n \rangle$ und erhalten $b_n = q_n r_{n,n}$.

Wegen $\|q_n\| = 1$ und $r_{n,n} > 0$ gilt $r_{n,n} = \|b_n\|$ und $q_n = b_n / \|b_n\|$. QED

😊 Alternativ kann man auch die Eindeutigkeit zuerst beweisen... und entdeckt auf diesem Wege erneut das Gram–Schmidt–Verfahren!

Die QR-Zerlegung

Die QR-Zerlegung ist in der Numerik ein wichtiges Werkzeug, etwa zur Berechnung einer Orthonormalbasis (wie oben) oder zur Behandlung von linearen Ausgleichsproblemen.

Die Zerlegung $A = QR$ ist eindeutig, also eine Abbildung $A \mapsto (Q, R)$. Sie kann jedoch mit verschiedenen Algorithmen berechnet werden!

Die bekanntesten sind neben dem obigen Gram-Schmidt-Verfahren vor allem Givens-Rotationen und Householder-Spiegelungen.

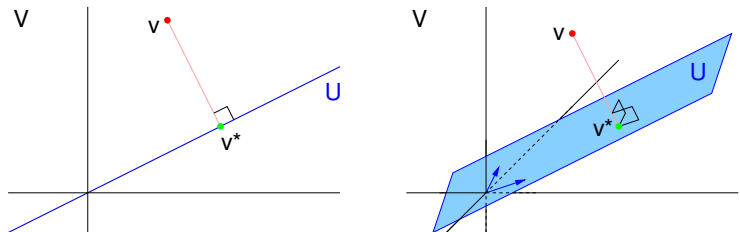
Ersteres wird üblicherweise in der Linearen Algebra benutzt, da es geometrisch anschaulich ist und leicht zu erklären.

In der oben angegebenen Standardform ist es leider numerisch instabil: Es kann Eingabefehler und Rundungsfehler ungünstig verstärken.

In der Numerik werden Sie daher weitere Verfahren zur QR-Zerlegung kennenlernen und ihre Eigenschaften noch genauer analysieren.

😊 Die Grundlage hierzu ist der Existenz- und Eindeutigkeitsatz P2B; er erklärt, was wir erreichen wollen, und dass es eindeutig möglich ist.

Bestapproximation durch Orthogonalprojektion



Sei V ein \mathbb{K} -Vektorraum mit Skalarprodukt $\langle - | - \rangle$ und Norm $\| - \|$.
Hierin sei $U \leq V$ ein endlich-dim. Untervektorraum, $\dim U = n < \infty$.

Approximationsproblem: Zu einem gegebenen Vektor $v \in V$ suchen wir den / einen Vektor $v^* \in U$, der v am nächsten liegt.

Lösung: Wir projizieren v orthogonal auf U , siehe Skizze und Satz P2c.

😊 Schon der Fall $V = \mathbb{R}^2$ und $n = 1$ (und ebenso $V = \mathbb{R}^3$ und $n = 1, 2$) ist interessant und Ihnen vielleicht noch gut aus der Schule vertraut.

Die folgenden Argumente gelten jedoch ganz allgemein: Der Raum V darf beliebig groß sein, zum Beispiel ein Funktionenraum. Lediglich der Unterraum U muss zunächst noch endlich-dimensional bleiben.

Bestapproximation durch Orthogonalprojektion

Satz P2C: Gauß 1795, Bessel 1818

Sei V ein \mathbb{K} -Vektorraum mit Skalarprodukt $\langle - | - \rangle$ und Norm $\|-\|$.
Hierin sei $U \leq V$ ein Unterraum mit Orthonormalbasis e_1, \dots, e_n .

(0) Zu jedem Vektor $v \in V$ existiert genau eine **Bestapproximation** in U , also ein Vektor $v^* \in U$ mit $\|v - v^*\| < \|v - u\|$ für alle $u \in U \setminus \{v^*\}$.

(1) Diese ist gegeben durch die **Orthogonalprojektion** von v auf U :

$$v^* = \sum_{k=1}^n v_k e_k \quad \text{mit Fourier-Koeffizienten} \quad v_k = \langle e_k | v \rangle.$$

(2) Zudem ist v^* der einzige Vektor in U , für den $(v - v^*) \perp U$ gilt.

(3) Dank Pythagoras P1N gilt daher die **Bessel-Gleichung**:

$$\underbrace{|v_1|^2 + \dots + |v_n|^2}_{\text{Längenquadrat von } v^*} + \underbrace{\|v - v^*\|^2}_{\text{Approximationsfehler}} = \underbrace{\|v\|^2}_{\text{Längenquadrat von } v}$$

(4) Hieraus folgt vergrößernd die **Bessel-Ungleichung**:

$$|v_1|^2 + \dots + |v_n|^2 = \|v^*\|^2 \leq \|v\|^2$$

Bestapproximation durch Orthogonalprojektion

Zusatz: Für die Aussagen (1,2,3) genügt es zu fordern, dass $\langle - | - \rangle$ auf V eine hermitesche Sesquilinearform ist. Dank der vorausgesetzten Orthonormalbasis ist sie dann positiv definit auf dem Unterraum $U \leq V$. Für (4) benötigen wir zudem, dass $\langle - | - \rangle$ auf V positiv semidefinit ist.

Beweis: (1) Wir betrachten den genannten Vektor

$$v^* := \sum_{k=1}^n v_k e_k \quad \text{mit den Koeffizienten} \quad v_k = \langle e_k | v \rangle.$$

Für jeden Vektor $u \in U$ gilt $u = \sum_{k=1}^n u_k e_k$ mit Koeffizienten $u_k \in \mathbb{K}$. Den Abstand zwischen v und u berechnen wir dank Skalarprodukt:

$$\begin{aligned} \|v - u\|^2 &= \langle v - u | v - u \rangle = \langle v | v \rangle - \langle v | u \rangle - \langle u | v \rangle + \langle u | u \rangle \\ &= \langle v | v \rangle - \langle v | \sum_k u_k e_k \rangle - \langle \sum_k u_k e_k | v \rangle + \langle \sum_k u_k e_k | \sum_\ell u_\ell e_\ell \rangle \\ &= \langle v | v \rangle - \sum_k u_k \langle v | e_k \rangle - \sum_k \overline{u_k} \langle e_k | v \rangle + \sum_k \sum_\ell \overline{u_k} u_\ell \langle e_k | e_\ell \rangle \\ &= \langle v | v \rangle - \sum_k u_k \overline{v_k} - \sum_k \overline{u_k} v_k + \sum_k |u_k|^2 \\ &= \|v\|^2 - 2 \operatorname{Re} \langle u | v^* \rangle + \|u\|^2 \end{aligned}$$

Bestapproximation durch Orthogonalprojektion

Speziell für $u = v^*$ folgt $\|v - v^*\|^2 = \|v\|^2 - \|v^*\|^2$. Im Vergleich gilt:

$$\begin{aligned} \|v - u\|^2 - \|v - v^*\|^2 &= \sum_k |u_k|^2 - \sum_k u_k \bar{v}_k - \sum_k \bar{u}_k v_k + \sum_k |v_k|^2 \\ &= \sum_k (\bar{u}_k - \bar{v}_k)(u_k - v_k) = \sum_k |u_k - v_k|^2 \geq 0 \end{aligned}$$

Gleichheit gilt hierbei nur für $u = v^*$, andernfalls $\|v - u\| > \|v - v^*\|$.

(2) Die behauptete Orthogonalität rechnen wir ebenso direkt nach.

Für alle $u, u^* \in U$, also $u = \sum_k u_k e_k$ und $u^* = \sum_\ell u_\ell^* e_\ell$, gilt:

$$\langle u | v - u^* \rangle = \left\langle \sum_k u_k e_k \mid v - \sum_\ell u_\ell^* e_\ell \right\rangle = \sum_k \bar{u}_k [\langle e_k | v \rangle - u_k^*] \stackrel{!}{=} 0$$

Demnach gilt $u \perp (v - u^*)$ für alle $u \in U$ genau dann, wenn die Koeffizienten $u_k^* = \langle e_k | v \rangle$ gewählt werden, also $u^* = v^*$ gilt.

Aus der expliziten Formel (1) und der Orthogonalität (2) folgt sofort die Bessel-Gleichung (3). Dank positiver Semidefinitheit von $\langle - | - \rangle$ auf V gilt $\|v - v^*\| \geq 0$ und somit die Bessel-Ungleichung (4). **QED**

Näherungslösung eines überbestimmten Gleichungssystems

Problemstellung: Zu lösen sei das lineare Gleichungssystem

$$Ax = b$$

mit $A \in \mathbb{K}^{m \times n}$ von Rang $\text{rang } A = n \leq m$ und rechter Seite $b \in \mathbb{K}^m$.

Seien $a_1, \dots, a_n \in \mathbb{K}^m$ die Spalten von A und $U = \langle a_1, \dots, a_n \rangle \leq \mathbb{K}^m$ der Spaltenraum von A . Genau dann ist $Ax = b$ lösbar, wenn $b \in U$ gilt.

Allgemein suchen wir eine Näherungslösung $x^* \in \mathbb{K}^n$. Der Fehlervektor

$$v = Ax^* - b$$

soll dabei möglichst klein sein, wir wollen also die Norm $\|v\|$ minimieren. Nach Satz P2c ist dies äquivalent zu $v \perp U$, und somit zu $v \perp a_i$ für alle $i = 1, \dots, n$, kurz $A^\dagger v = 0$. Ausgeschrieben bedeutet das:

$$A^\dagger Ax^* = A^\dagger b$$

😊 Statt unserer ursprünglichen, eventuell überbestimmten Gleichung $Ax = b$ lösen wir diese Umformung; letztere ist eindeutig lösbar.

Näherungslösung eines überbestimmten Gleichungssystems

Lemma P2D

Für jede Matrix $A \in \mathbb{K}^{m \times n}$ ist $A^\dagger A \in \mathbb{K}^{n \times n}$ symmetrisch (über \mathbb{R}) bzw. hermitesch (über \mathbb{C}), erfüllt $\ker(A^\dagger A) = \ker A$ und $\text{rang}(A^\dagger A) = \text{rang } A$.

Beweis: Zunächst gilt $(A^\dagger A)^\dagger = A^\dagger (A^\dagger)^\dagger = A^\dagger A$.

Wir zeigen $\ker(A^\dagger A) = \ker A$. Die Inklusion „ \supseteq “ ist klar. Wir zeigen „ \subseteq “: Sei $v \in \ker(A^\dagger A)$, also $A^\dagger A v = 0$. Dann gilt $0 = v^\dagger (A^\dagger A v) = (A v)^\dagger (A v)$, also $A v = 0$ und somit $v \in \ker A$.

Daraus folgt $\text{rang}(A^\dagger A) = \text{rang } A$ dank Dimensionsformel. □

Bemerkung: Wir nutzen hier wesentlich $\mathbb{K} = \mathbb{R}, \mathbb{C}$ und die positive Definitheit des euklidischen Skalarprodukts $\langle u | v \rangle = u^\dagger v$.

Für beliebige Körper gilt diese Aussage nicht! Als einfaches Gegenbeispiel betrachte man die Matrix $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{F}_2^2$ mit $A^\top A = 0$.

Näherungslösung eines überbestimmten Gleichungssystems

Satz P2E: Näherungslösung einer überbestimmten Gleichung

Zu lösen ist das lineare Gleichungssystem

$$Ax = b$$

mit $A \in \mathbb{K}^{m \times n}$ von Rang $\text{rang } A = n \leq m$ und rechter Seite $b \in \mathbb{K}^m$.

(0) Dann P2C existiert genau eine Bestapproximation $x^* \in \mathbb{K}^n$ mit minimaler Fehlernorm $\|Ax^* - b\|$, gegeben durch $A^\dagger Ax^* = A^\dagger b$.

(1) Dank des vorigen Lemmas ist $A^\dagger A \in \mathbb{K}^{n \times n}$ invertierbar, also gilt:

$$x^* = (A^\dagger A)^{-1} A^\dagger b$$

(2) Ist $A = QR$ die QR-Zerlegung von A gemäß Satz P2B, so folgt:

$$Rx^* = Q^\dagger b$$

Diese Gleichung kann durch Rückwärtseinsetzen leicht gelöst werden.

Näherungslösung eines überbestimmten Gleichungssystems

Beweis: Die Aussagen (0) und (1) haben wir zuvor schon hergeleitet.

(2) Aus $A^\dagger Ax^* = A^\dagger b$ und $A = QR$ folgt $R^\dagger Q^\dagger QRx^* = R^\dagger Q^\dagger b$,
dank $Q^\dagger Q = E_n$ also $Rx^* = Q^\dagger b$. QED

😊 Wir nennen $A^+ = (A^\dagger A)^{-1} A^\dagger$ daher die **Pseudoinverse** zu A .
Zur Gleichung $Ax = b$ berechnet sie die Näherungslösung $x = A^+ b$.

Ist die Matrix A invertierbar, so auch A^\dagger , und dann gilt $A^+ = A^{-1}$.
In diesem Falle ist unsere Näherungslösung die exakte Lösung.

Aus der QR-Zerlegung $A = QR$ folgt $A^+ = R^{-1} Q^\dagger$.
Das ist besonders elegant und einfach zu rechnen.

😊 Die QR-Zerlegung ist ein Standardverfahren der (numerischen) Linearen Algebra und in jeder guten Softwarebibliothek professionell implementiert. Darauf können und sollten wir zurückgreifen.

Daher ist es sinnvoll, ein gegebenes Problem, wie oben $Ax = b$,
auf ein solches, bewährtes Standardverfahren zurückzuführen.
Sobald dies gelingt, ist unser Problem gelöst.

Die Methode der kleinsten Quadrate

Die hier gezeigte Technik heißt **Methode der kleinsten Quadrate**. Wir minimieren dabei die Summe der Fehlerquadrate.

Sie ist ein wunderbar-geniales Universalwerkzeug, beginnend mit der Bestapproximation P2C nach Gauß–Bessel, über die Näherungslösung P2E einer überbestimmten Gleichung, hin zu zahlreichen Varianten und Anwendungen der Ausgleichsrechnung etwa in der Numerik, der Statistik oder dem maschinellen Lernen.

Überbestimmte Gleichungssysteme $Ax = b$ treten typischerweise wie folgt auf: Wir interessieren uns für eine (kleine) Anzahl von Kenngrößen x_1, \dots, x_n . Zu ihrer Bestimmung haben wir eine (große) Anzahl von Messungen b_1, \dots, b_m . Bei jedem Messvorgang treten unvermeidliche Messfehler auf, daher ist es im Prinzip vorteilhaft, möglichst viele Messungen zu machen. Andererseits können wir dann nicht mehr erwarten, hierzu eine exakte Lösung x zu finden.

Die Methode der kleinsten Quadrate

Die Methode der kleinsten Quadrate wurde unabhängig von Gauß und Legendre entwickelt (Legendre veröffentlichte sie 1805, Gauß 1809). Beide nutzten dieses Verfahren, um astronomische Umlaufbahnen anhand von Beobachtungsdaten möglichst genau zu bestimmen.

Am Neujahrstag 1801 entdeckte der Astronom Giuseppe Piazzi den Zwergplaneten Ceres. Vierzig Tage lang konnte er seine Bahn verfolgen, doch dann verschwand Ceres hinter der Sonne. Viele Astronomen versuchten erfolglos, anhand von Piazzis Beobachtungsdaten die Bahn zu berechnen und zu verfolgen, doch Ceres blieb verschwunden.

Dem 24-jährigen Gauß gelang der Durchbruch, die Bahn wesentlich genauer zu berechnen, indem er zur Ausgleichsrechnung die Methode der kleinsten Quadrate nutzte. Ausgehend von Gauß' Vorhersage konnte der Astronom Franz Xaver von Zach nach einem Jahr, am 7. und 31. Dezember 1801 Ceres tatsächlich wiederfinden.

Die Methode der kleinsten Quadrate

Ceres befand sich etwa 7° von der zuvor vermuteten Stelle, also mehr als 13 Vollmondbreiten. Das illustriert die schwierige Datenlage und die durchschlagende Verbesserung durch Mathematik und neue Verfahren. Dieser sensationelle Erfolg machten Gauß und seine Methode berühmt.

Übrigens nutzte Gauß im Zuge seiner astronomischen Berechnungen ganz systematisch noch ein weiteres wichtiges Werkzeug: die Lösung linearer Gleichungssysteme durch sukzessive Elimination der Variablen. Diese Rechenmethode war zuvor schon benutzt worden, doch Gauß führte sie virtuos zu höchster Blüte. Seine erste Veröffentlichung zu diesem Thema stammt aus dem Jahr 1810. Die Mathematik allgemein, und speziell die Lineare Algebra, entspringt und nützt direkt praktischen Anwendungen. Sie als abstrakt zu beschimpfen zeugt von Ignoranz.

Sowohl das Ausgleichs- als auch das Eliminationsverfahren wurden in der Folgezeit in der Geodäsie zur Landvermessung eingesetzt. Daher ist der zweite Namensgeber des Gauß–Jordan–Verfahrens nicht etwa der Mathematiker Camille Jordan, sondern der Geodät Wilhelm Jordan.

Die Methode der kleinsten Quadrate

Dieser historische Rückblick ist in sich schon eine faszinierende Geschichte. Zudem sehen wir daran erneut sehr eindrücklich, dass die scheinbar so abstrakten Methoden der Linearen Algebra auf ganz konkrete Fragestellungen und Bedürfnisse antworten.

Daher das eingangs zitierte Motto dieses Kapitels:

*Unsere Allergrößten, wie Archimedes, Newton, Gauß,
haben stets Theorie und Anwendungen gleichmäßig umfasst.*

Felix Klein (1849–1925)

Mathematik ist immer beides: sowohl abstrakte Theorie als auch konkrete Anwendung; sie sind keine Gegensätze, sie ergänzen sich, die eine kann nur mit der anderen dauerhaft erfolgreich sein.

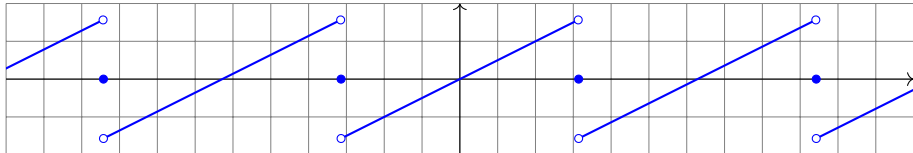
Am besten, Sie beherrschen beides!

Anwendung: Fourier–Entwicklung der Sägezahnfunktion

Sei V der \mathbb{C} –Vektorraum aller 2π –periodischen, stückweise stetigen Funktionen $f, g: \mathbb{R} \rightarrow \mathbb{C}$ mit dem (semidefiniten) Skalarprodukt

$$\langle f | g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{f(x)} g(t) dx.$$

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ ungerade und 2π –periodisch mit $f(x) = x/2$ für $|x| < \pi$. Der erste Teil der Übung ist wie immer, die Funktion zu skizzieren:



Aufgabe: Berechnen Sie die Bestapproximation $f_n \in U_n$ im Unterraum $U_n = \langle e_k \mid -n \leq k \leq n \rangle \subseteq V$ der Fourier–Polynome vom Grad $\leq n$.

😊 Das scheint zunächst ein verblüffend schwieriges Problem zu sein, doch dank Satz P2c wird alles leicht! Integration aus der Schule genügt.

Anwendung: Fourier-Entwicklung der Sägezahnfunktion

Lösung: Der Fourier-Koeffizient c_0 ist der Mittelwert über eine Periode:

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{x}{2} dx = 0 \quad (\text{ungerader Integrand})$$

Für $k \neq 0$ nutzen wir partielle Integration:

$$\begin{aligned} c_k &\stackrel{\text{Def}}{=} \frac{1}{4\pi} \int_{-\pi}^{\pi} e^{-ikx} x dx && \stackrel{\text{part}}{=} \frac{1}{4\pi} \left(\left[\frac{i}{k} e^{-ikx} x \right]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} \frac{i}{k} e^{-ikx} dx \right) \\ &= \frac{i}{4\pi k} \left[e^{-i\pi k} \pi - e^{i\pi k} \pi \right] = (-1)^k \frac{i}{2k} \end{aligned}$$

Damit haben wir zu $f \in V$ die Bestapproximation $f_n \in U_n$ gefunden:

$$\begin{aligned} f_n(x) &= \sum_{k=1}^n (-1)^k i \frac{e^{ikx} - e^{-ikx}}{2k} = \sum_{k=1}^n (-1)^{k+1} \frac{\sin(kx)}{k} \\ &= \sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \frac{\sin 4x}{4} + \dots + (-1)^{n+1} \frac{\sin nx}{n} \end{aligned}$$

Anwendung: Fourier-Entwicklung der Sägezahnfunktion

Umrechnung der Koeffizienten für die Sinus-Cosinus-Reihe:

$$a_k = c_k + c_{-k} = 0, \quad b_k = i(c_k - c_{-k}) = (-1)^{k+1} \frac{1}{k}.$$

Zum Vergleich nochmal direkt die Integrale für a_k, b_k mit $k \geq 1$:

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{x}{2} \cos(kx) \, dx \quad (\text{ungerader Integrand}) \\ &= \frac{1}{2\pi} \left(\left[x \frac{\sin(kx)}{k} \right]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} \frac{\sin(kx)}{k} \, dx \right) = 0 \end{aligned}$$

$$\begin{aligned} b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{x}{2} \sin(kx) \, dx \\ &= \frac{1}{2\pi} \left(\left[x \frac{-\cos(kx)}{k} \right]_{-\pi}^{\pi} + \int_{-\pi}^{\pi} \frac{\cos(kx)}{k} \, dx \right) = (-1)^{k+1} \frac{1}{k} \end{aligned}$$

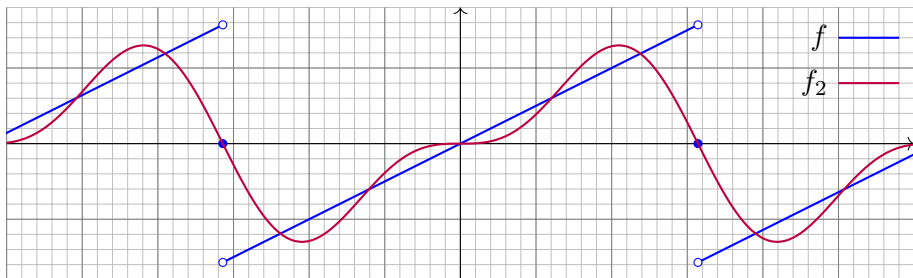
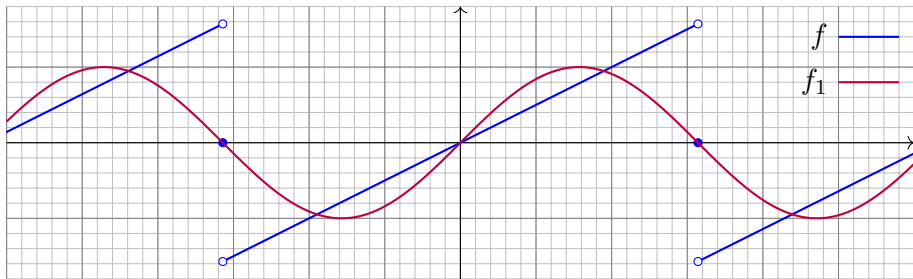
Zur Berechnung von a_k, b_k sind zwei reelle Integrale nötig, für c_k nur ein komplexes; die Wahl des Rechenwegs ist meist Geschmackssache. Die Umrechnung zwischen a_k, b_k und c_k gelingt jedenfalls leicht.

Anwendung: Fourier–Entwicklung der Sägezahnfunktion

- 😊 Die Fourier–Koeffizienten a_k, b_k, c_k sind hier leicht zu berechnen. Da f reell ist, gilt $a_k, b_k \in \mathbb{R}$ und $c_{-k} = \overline{c_k}$. Da f ungerade, gilt $a_k = 0$.
- 😊 Die folgenden Graphiken zeigen hierzu die **Fourier–Polynome** f_n . Wir wollen verstehen, ob und in welchem Sinne f_n gegen f konvergiert.
- 😊 Für jeden Punkt $x \in \mathbb{R}$ gilt augenscheinlich $f_n(x) \rightarrow f(x)$ für $n \rightarrow \infty$: In den Punkten $x = 0$ und $x = \pi$ ist dies klar, ansonsten keineswegs!
- 😞 Die Koeffizienten klingen nur langsam ab ($\sim 1/k$), das heißt auch hohe Frequenzen tragen noch deutlich bei: Die Fourier–Reihe ist „rau“.
- ⚠ Wir sehen recht eindringlich das sogenannte **Gibbs–Phänomen**: Die Funktionen f_n überschwingen in Sprungstellen um ca. 9%.
- 😞 Es gilt daher **keine gleichmäßige Konvergenz** $f_n \rightarrow f$ auf \mathbb{R} : Ein kleiner ε –Schlauch um f enthält nicht alle f_n für $n \geq n_0$.
- 😊 Auf jedem Intervall $I = [-\pi + \delta, \pi - \delta]$ **abseits der Sprungstellen** konvergiert f_n gleichmäßig gegen f : Zu jedem $\varepsilon > 0$ liegen schließlich alle f_n im ε –Schlauch um f auf I . Auch das ist bemerkenswert!

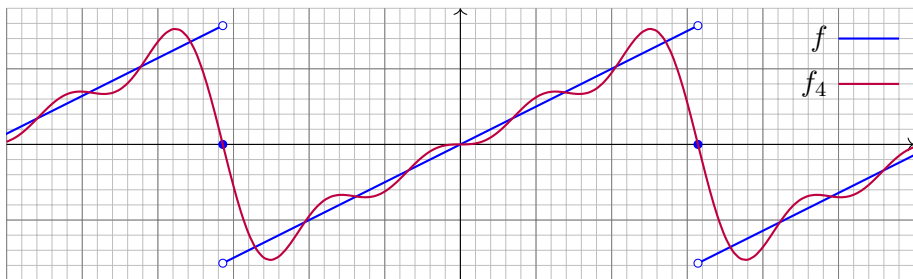
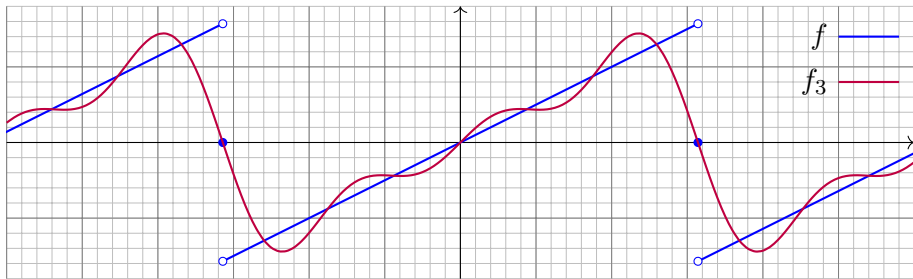
Anwendung: Fourier-Entwicklung der Sägezahnfunktion

Die ersten Fourier-Polynome f_1, f_2 ähneln f zunächst nur grob:



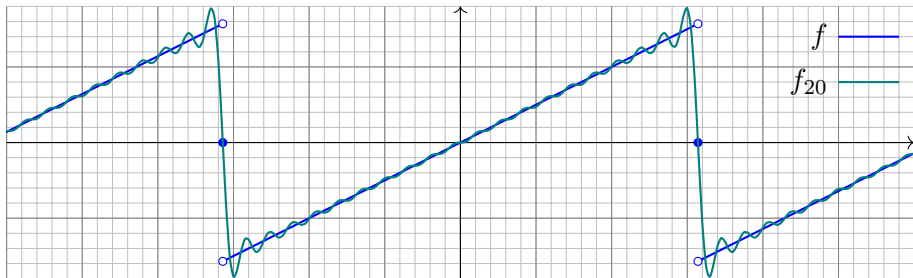
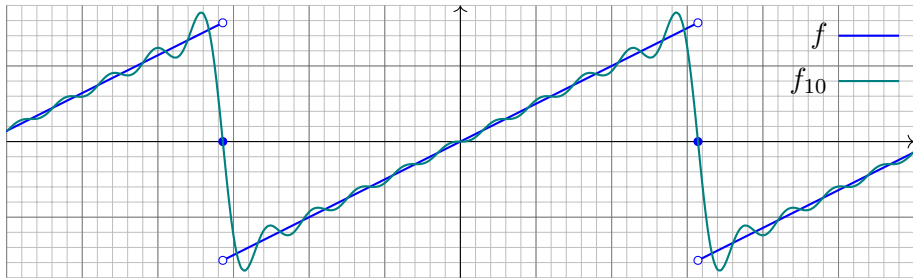
Anwendung: Fourier–Entwicklung der Sägezahnfunktion

Die nächsten Fourier–Polynome f_3, f_4 ähneln f schon etwas mehr:



Anwendung: Fourier–Entwicklung der Sägezahnfunktion

Die Fourier–Polynome überschwingen bis zu 9% der Sprunghöhe:



Was bedeutet Konvergenz der Fourier-Reihe?

Wir haben die Sägezahnfunktion $f : \mathbb{R} \rightarrow \mathbb{R}$ durch Fourier-Polynome $f_1, f_2, f_3, \dots : \mathbb{R} \rightarrow \mathbb{R}$ angenähert: Dies ist die L^2 -Bestapproximation! Die gezeigten Graphiken suggerieren, dass dies erstaunlich gut gelingt. Obwohl f unstetig ist, nähern sich die Funktionen f_n doch recht gut an.

Ich habe dieses Beispiel so weit getrieben, wie es mit den Werkzeugen der Linearen Algebra (und etwas Integration aus der Schule) möglich ist. Die genauere Untersuchung lernen Sie in der Analysis. Als Ausblick will ich die Antworten auf die Konvergenzfrage wenigstens kurz skizzieren.

Es gilt Konvergenz im quadratischen Mittel: Für jede 2π -periodische Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ mit $\|f\| < \infty$ gilt $\|f - f_n\| \searrow 0$, somit $\|f_n\| \nearrow \|f\|$. Bezüglich der L^2 -Norm gilt also: Jede quadrat-integrierbare Funktion f lässt sich beliebig gut durch ihre Fourier-Polynome f_n approximieren.

Die Frage der punktweisen Konvergenz $|f(x) - f_n(x)| \rightarrow 0$ oder der gleichmäßigen Konvergenz $\sup_{x \in \mathbb{R}} |f(x) - f_n(x)| \rightarrow 0$ ist kniffliger. Hierzu nenne ich das folgende berühmte Kriterium von Dirichlet, das für viele praktischen Anwendungen genügt.

Konvergenz im quadratischen Mittel

Satz P2F: Konvergenz bezüglich der L^2 -Norm

Sei $f: \mathbb{R} \rightarrow \mathbb{C}$ absolut integrierbar auf $[0, T]$ und T -periodisch.

Die **Fourier-Analyse** zerlegt das Signal f in sein Spektrum $\hat{f}: \mathbb{Z} \rightarrow \mathbb{C}$ gegeben durch die Koeffizienten $\hat{f}(k) := \langle e_k | f \rangle = \frac{1}{T} \int_{t=0}^T e^{-ik\omega t} f(t) dt$.

Dabei gilt die **Parseval-Gleichung**, auch **Energiegleichung** genannt:

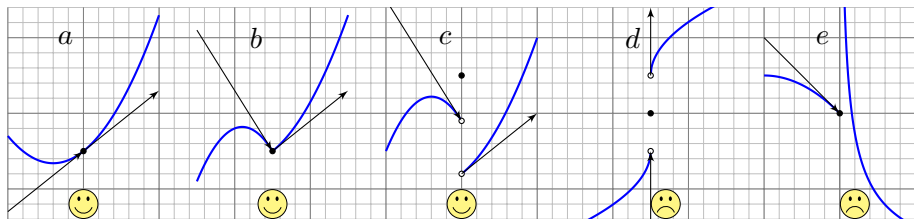
$$\|f\|_{L^2} = \|\hat{f}\|_{\ell^2} \quad \text{also} \quad \frac{1}{T} \int_{t=0}^T |f(t)|^2 dt = \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2$$

Ist f quadrat-integrierbar, also $\|f\| < \infty$, so konvergieren die Fourier-Polynome f_n bezüglich der L^2 -Norm, also $\|f - f_n\| \searrow 0$ für $n \rightarrow \infty$.

Beispiel: Die Sägezahnfunktion liefert die bemerkenswerte Gleichung

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Punktweise Konvergenz von Fourier-Reihen



Links-/rechtsseitiger Grenzwert und Ableitungen von f im Punkt x :

$$f(x-) := \lim_{\xi \nearrow x} f(\xi), \quad f(x+) := \lim_{\xi \searrow x} f(\xi),$$

$$f'(x-) := \lim_{\xi \nearrow x} \frac{f(\xi) - f(x-)}{\xi - x}, \quad f'(x+) := \lim_{\xi \searrow x} \frac{f(\xi) - f(x+)}{\xi - x}.$$

Die **Dirichlet-Bedingung** fordert, dass alle vier Grenzwerte existieren. Wir nennen f **sprungnormiert**, falls $f(x) = \frac{1}{2} [f(x+) + f(x-)]$ gilt.

Stetigkeit im Punkt $x \in \mathbb{R}$ ist äquivalent zu $f(x+) = f(x-) = f(x)$. Im Falle $f(x+) \neq f(x-)$ hat f in x eine **Sprungstelle** (siehe Skizze).

Differenzierbarkeit im Punkt x impliziert Stetigkeit und ist äquivalent zu Dirichlet mit $f(x+) = f(x-) = f(x)$ und $f'(x+) = f'(x-) = f'(x)$.

Punktweise Konvergenz von Fourier-Reihen

Satz P2G: Dirichlet-Kriterium für Fourier-Reihen

Sei $f: \mathbb{R} \rightarrow \mathbb{C}$ absolut integrierbar auf $[0, 2\pi]$ und 2π -periodisch.

(1) Angenommen, $f: \mathbb{R} \rightarrow \mathbb{C}$ erfüllt die Dirichlet-Bedingung im Punkt x , d.h. beide Grenzwerte $f(x\pm)$ und beide Ableitungen $f'(x\pm)$ existieren. Dann konvergiert in diesem Punkt x die Fourier-Reihe $f_n(x)$ gemäß

$$f_n(x) = \sum_{k=-n}^n c_k e^{ikx} \rightarrow \frac{1}{2} [f(x+) + f(x-)] \quad \text{für } n \rightarrow \infty.$$

Spezialfälle: (1a) Es gilt $f_n(x) \rightarrow f(x)$ falls f in x sprungnormiert ist, also $f(x) = \frac{1}{2} [f(x+) + f(x-)]$, oder sogar stetig, also $f(x\pm) = f(x)$.

(1b) Ist $f: \mathbb{R} \rightarrow \mathbb{C}$ stückweise stetig differenzierbar und überall stetig bzw. sprungnormiert, dann konvergiert $f_n(x) \rightarrow f(x)$ in jeden Punkt $x \in \mathbb{R}$.

(2) Ist $f: \mathbb{R} \rightarrow \mathbb{C}$ stetig und stückweise stetig differenzierbar mit $|f'| \leq L$, so konvergiert die Fourier-Reihe $f_n \rightarrow f$ sogar gleichmäßig auf ganz \mathbb{R} :

$$|f_n(x) - f(x)| \leq 2L \cdot \ln(n)/n \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

Punktweise Konvergenz von Fourier-Reihen

Beispiel: Unsere Sägezahnfunktion erfüllt die Dirichlet-Bedingung in jedem Punkt $x \in \mathbb{R}$ (Übung!) und ist zudem sprungnormiert. Daher gilt in jedem Punkt $x \in \mathbb{R}$ die Konvergenz $f_n(x) \rightarrow f(x)$ für $n \rightarrow \infty$.

😊 Damit haben wir die Funktion f in ihre Fourier-Reihe entwickelt! Das ist kein Fourier-Polynom mehr, sondern eine unendliche Reihe:

$$\begin{aligned}
 f(x) &= \sum_{k=1}^{\infty} (-1)^k i \frac{e^{ikx} - e^{-ikx}}{2k} = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\sin(kx)}{k} \\
 &= \sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \frac{\sin 4x}{4} + \frac{\sin 5x}{5} \mp \dots
 \end{aligned}$$

😊 Daraus können wir für erstaunliche Reihen den Grenzwert ablesen. Die Auswertung im Punkt $x = \pi/2$ bzw. $x = \pi/3$ zum Beispiel ergibt:

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \frac{1}{13} - \frac{1}{15} + \dots = \frac{\pi}{4} = 0.7853981633\dots$$

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{5} + \frac{1}{7} - \frac{1}{8} + \frac{1}{10} - \frac{1}{11} + \dots = \frac{\pi}{3\sqrt{3}} = 0.6045997880\dots$$