Kapitel C

Distanzlehre: Metrische Räume

Alles messen, was messbar ist –
und messbar machen, was noch nicht messbar ist.
Galileo Galilei (1564–1642)

Inhalt dieses Kapitels C

Mathematik: isoliert oder vernetzt?

Dieses Kapitel ist ein Rückblick auf die ersten Semester mit einigen Präzisierungen. Ich erstrebe damit eine gemeinsame Grundlage und einen reichhaltigen Beispielfundus für die Topologie.

Ich importiere hier freizügig Themen aus der Linearen Algebra und der Analysis, die sie ja alle gehört haben (sofern Sie Mathematik studieren), manchmal auch aus der Stochastik oder der Numerik oder anschauliche Beispiele aus der Physik, der Informatik oder weiteren Anwendungen.

Der Kenntnisstand der Teilnehmer ist naturgemäß sehr unterschiedlich. Wenn Sie dabei schöne Themen wiedererkennen, dürfen Sie sich freuen und stolz sein auf ihre mathematische Allgemeinbildung. Wenn nicht, dann motiviert Sie dieser Ausblick vielleicht, mehr wissen zu wollen.

Ich will und muss mich hier kurz fassen — aber doch nicht zu kurz!
Ich versuche dabei, die Mathematik vernetzt darzustellen und hilfreiche Querverbindungen zu zeigen. Natürlich kann ich dies oft nur skizzieren und auf Ihre Neugier vertrauen. Manche lieben es, andere hassen es.

Mathematik: isoliert oder vernetzt?

Das Studium verlockt zu zwei Extremen:

- Es besteht die Gefahr, dass Sie von allem ein Wenig gehört haben, aber nichts recht beherrschen: Überblickswissen ohne Detailwissen.
- Es besteht auch die umgekehrte Gefahr, dass Sie sich allzu früh auf ein schmales Thema spezialisieren und kaum mehr kennenlernen.

Ich denke, die Wahrheit liegt in der goldenen Mitte zwischen den beiden: möglichst breiter Überblick und zudem ausreichend tiefes Detailwissen.

Mein tiefempfundener Ratschlag an Sie als hoffnungsfrohe Studierende: Finden Sie mindestens ein Thema, besser zwei oder drei oder mehr, die Sie wirklich begeistern und vertiefen Sie sich darin.

Besonders bereichernd und erleuchtend in der Mathematik sind die Querverbindungen! Dies gilt auch und besonders zur Physik, zur Informatik oder zu den Wirtschaftswissenschaften (Spieltheorie).

Die Mathematik ist keine sterile Ansammlung isolierter Fakten, sondern ein lebendiges Ganzes mit vielfältige Vernetzung.

Das euklidische Skalarprodukt

Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Auf dem $\mathbb{K}$–Vektorraum $V = \mathbb{K}^n$ definieren wir das **euklidische Skalarprodukt** $\langle - \mid - \rangle : V \times V \to \mathbb{K}$ durch

$$\langle (u_1, \dots, u_n) \mid (v_1, \dots, v_n) \rangle := \overline{u_1} v_1 + \dots + \overline{u_n} v_n.$$

Im reellen Fall $\mathbb{K} = \mathbb{R}$ gilt $\bar{x} = x$. Diese Paarung erfreut sich folgender Eigenschaften für alle Vektoren $u, v, w \in V$ und Skalare $\lambda, \mu \in \mathbb{K}$:

S0: $\langle u \mid u \rangle \in \mathbb{R}_{\geq 0} \subset \mathbb{K}$ (Positivität)

S1: $\langle u \mid u \rangle > 0$ für $u \neq 0$ (Definitheit)

S2: $\langle u \mid v \rangle = \overline{\langle v \mid u \rangle}$ (Symmetrie)

S3: $\langle u \mid \lambda v + \mu w \rangle = \lambda \langle u \mid v \rangle + \mu \langle u \mid w \rangle$ (Linearität)

Aus (S2,3) folgt Semi/Linearität in der ersten Variablen:

$$\langle \lambda u + \mu v \mid w \rangle = \overline{\lambda} \langle u \mid w \rangle + \overline{\mu} \langle v \mid w \rangle$$

Im reellen Fall $\mathbb{K} = \mathbb{R}$ ist $\langle - \mid - \rangle$ eine **Bilinearform**, im komplexen Fall $\mathbb{K} = \mathbb{C}$ eine **Sesquilinearform**.

Die Cauchy–Schwarz–Ungleichung

Satz C1A: Cauchy–Schwarz–Ungleichung (CSU)

Aus (S0–3) folgt für alle $u, v \in V$ die **Cauchy–Schwarz–Ungleichung**:

$$|\langle u \mid v \rangle|^2 \leq \langle u \mid u \rangle \langle v \mid v \rangle \qquad \text{bzw.} \qquad |\langle u \mid v \rangle| \leq |u| \cdot |v|$$

für die **Norm** $|u| := \sqrt{\langle u \mid u \rangle}$. Gleichheit gilt hierbei genau dann, wenn u, v über $\mathbb{K}$ linear abhängig sind. Die so definierte Norm erfreut sich folgender Eigenschaften für alle Vektoren $u, v \in V$ und Skalare $\lambda \in \mathbb{K}$:

N0: $|u| \geq 0 = |0|$ (Positivität)

N1: $|u| > 0$ für $u \neq 0$ (Definitheit)

N2: $|\lambda u| = |\lambda| \cdot |u|$ (Homogenität)

N3: $|u + v| \leq |u| + |v|$ (Dreiecksungleichung)

Statt der Schreibweise $|u|$ für Normen ist alternativ auch $\|u\|$ üblich, etwa zur besseren Unterscheidung in $|\langle u \mid v \rangle| \leq \|u\| \cdot \|v\|$ und $\|\lambda u\| = |\lambda| \cdot \|u\|$.

Übung: Beweisen Sie die Cauchy–Schwarz–Ungleichung und damit N3!

Anwendungen von Skalarprodukt und Norm

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

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

Vektorräume mit Skalarprodukt und mit Norm

☺ Wir erheben die grundlegenden Eigenschaften nun zur Definition:

Definition C1B: Skalarprodukt und Norm

Sei V ein Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Ein **Skalarprodukt** auf V ist eine Abbildung $\langle - \mid - \rangle : V \times V \to \mathbb{K}$, die (S0–3) erfüllt.

Das Paar $(V, \langle - \mid - \rangle)$ heißt dann **$\mathbb{K}$–Vektorraum mit Skalarprodukt**, oder auch **Prä–Hilbert–Raum** und bei Vollständigkeit **Hilbert–Raum**.

Eine **Norm** auf V ist eine Abbildung $|-| : V \to \mathbb{R}_{\ge 0}$, die (N0–3) erfüllt.
Jedes Skalarprodukt definiert eine zugehörige Norm $|u| := \sqrt{\langle u \mid u \rangle}$.

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

Übung: In vielen Anwendungen tritt auch der Wert $|u| = \infty$ natürlich auf; dies nennen wir eine **Pseudonorm**. Eine **Halbnorm** erfüllt nur (N0,2,3). Die Teilmengen $V_{<\infty} = \{ v \in V \mid |v| < \infty \}$ und $V_0 = \{ v \in V \mid |v| = 0 \}$ in V sind Untervektorräume, $V_0 < V_{<\infty} < V$, dank Dreiecksungleichung (N3). Auf dem Quotientenvektorraum $V_{<\infty}/V_0$ induziert $|-|$ eine echte Norm.

Vektorräume mit Skalarprodukt: ℓ^2

Beispiel / Übung C1C: der Hilbert–Raum ℓ^2

Weiterhin sei $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Für jede Menge Ω ist der **Funktionenraum**

$$\mathbb{K}^\Omega = \{ f : \Omega \to \mathbb{K} \}$$

ein $\mathbb{K}$–Vektorraum mit punktweiser Addition und Skalarmultiplikation. Die **Trägermenge** von f ist $\operatorname{supp}(f) := \{ x \in \Omega \mid f(x) \neq 0 \}$. Die Teilmenge

$$\mathbb{K}^{(\Omega)} := \{ f : \Omega \to \mathbb{K} \mid \operatorname{supp}(f) \text{ endlich } \}$$

ist ein Untervektorraum in $\mathbb{K}^\Omega$. Hierauf haben wir das **Skalarprodukt**

$$\langle - \mid - \rangle \;:\; \mathbb{K}^{(\Omega)} \times \mathbb{K}^{(\Omega)} \to \mathbb{K} \;:\; (f,g) \mapsto \langle f \mid g \rangle := \textstyle\sum_{x \in \Omega} \overline{f(x)}\, g(x).$$

Die kanonische Basis $(e_a)_{a \in \Omega}$ von $\mathbb{K}^{(\Omega)}$ ist orthonormal bzgl. $\langle - \mid - \rangle$. Vervollständigt erhalten wir aus $\mathbb{K}^{(\Omega)}$ den **Hilbert–Raum**

$$\ell^2 = \ell^2(\Omega) = \ell^2(\Omega, \mathbb{K}) := \{ f : \Omega \to \mathbb{K} \mid \textstyle\sum_{x \in \Omega} |f(x)|^2 < \infty \}.$$

Es gilt $\mathbb{K}^{(\Omega)} < \ell^2 < \mathbb{K}^\Omega$, und das obige Skalarprodukt setzt sich auf ℓ^2 fort.

Vektorräume mit Skalarprodukt: L^2

Beispiel / Übung C1D: der Hilbert–Raum L^2

Sei $a < b$ in $\mathbb{R}$. Wir betrachten wir den Vektorraum $\mathscr{C} = \mathscr{C}([a,b], \mathbb{K})$ aller stetigen Funktionen $f, g : [a,b] \to \mathbb{K}$. Hierauf haben wir das Skalarprodukt

$$\langle f, g \rangle = \frac{1}{\tau} \int_{x=a}^{b} \overline{f(x)}\, g(x)\, \mathrm{d}x \quad \text{mit} \quad \tau = b - a.$$

Die Funktionen $e_k : [a,b] \to \mathbb{C} : x \mapsto \mathrm{e}^{\mathrm{i}k\omega x}$ mit $\omega = 2\pi/\tau$ und $k \in \mathbb{Z}$ sind orthonormal und bilden eine Basis der **trigonometrischen Polynome**:

$$T = T([a,b], \mathbb{C}) := \left\{ f = \sum_{k=-n}^{n} c_k e_k \;\middle|\; n \in \mathbb{N},\ c_{-n}, \dots, c_n \in \mathbb{C} \right\}$$

Reell betrachten wir $f(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(k\omega x) + b_k \sin(k\omega x)$ mit $n \in \mathbb{N}$ und $a_0, a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$. Beide Basen lassen sich ineinander umrechnen dank der Euler–Formel $\mathrm{e}^{\mathrm{i}k\omega x} = \cos(k\omega x) + \mathrm{i} \sin(k\omega x)$.

🙂 Die komplexe Formulierung ist einfacher und schöner!

Vektorräume mit Skalarprodukt: L^2

Beispiel / Übung C1D: der Hilbert–Raum L^2, fortgesetzt

Vervollständigt erhalten wir aus $T < \mathscr{C}$ den Raum

$$\mathscr{L}^2([a,b],\mathbb{K}) := \left\{ f : [a,b] \to \mathbb{K} \text{ messbar} \;\middle|\; \int_{x=a}^{b} |f(x)|^2 \,\mathrm{d}x < \infty \right\}.$$

Hierbei gilt $T < \mathscr{C} < \mathscr{L}^2 < \mathbb{K}^{[a,b]}$, und das obige Skalarprodukt setzt sich auf $\mathscr{L}^2$ fort, allerdings ist es hier nur noch semidefinit:

Aus $\int |f(x)|^2 \,\mathrm{d}x = 0$ folgt $f(x) = 0$ nicht unbedingt für alle x, sondern nur für fast alle $x \in [a,b]$. Das bedeutet, es existiert eine Teilmenge $N \subset [a,b]$ vom Lebesgue–Maß $\mathrm{vol}_1(N) = 0$, sodass $f(x) = 0$ für alle $x \in [a,b] \smallsetminus N$.

Diese Nullfunktionen bilden den Untervektorraum

$$\mathscr{N} := \left\{ f : [a,b] \to \mathbb{K} \;\middle|\; f(x) = 0 \text{ für fast alle } x \in [a,b] \right\}.$$

🙂 Nullfunktionen wollen und können und werden wir vernachlässigen. Sie liefern bei der Integration keinen Beitrag, sind somit masselos.

Vektorräume mit Skalarprodukt: L^2

Beispiel / Übung C1D: der Hilbert–Raum L^2, fortgesetzt

Der relevante Vektorraum ist demnach der Quotientenvektorraum

$$L^2([a,b],\mathbb{K}) := \mathscr{L}^2([a,b],\mathbb{K}) / \mathscr{N}.$$

Hierauf ist das obige Skalarprodukt $(f,g) \mapsto \langle f \mid g \rangle$ wohldefiniert und nun auch tatsächlich positiv-definit. Zudem ist $L^2([a,b],\mathbb{K})$ vollständig, wir erhalten also schließlich einen Hilbert–Raum.

Die Räume ℓ^2 und L^2 sehen verschieden aus, sind aber isometrisch: Es gilt die bemerkenswerte Fourier–Isometrie $\ell^2(\mathbb{Z},\mathbb{C}) \cong L^2([a,b],\mathbb{C})$!

$$\mathbb{C}^{(\mathbb{Z})} \cong T([a,b],\mathbb{C}) \;:\; e_k \leftrightarrow e_k$$
$$\ell^2(\mathbb{Z},\mathbb{C}) \cong L^2([a,b],\mathbb{C}) \;:\; e_k \leftrightarrow e_k$$

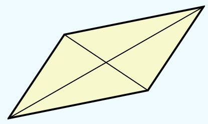
🙂 Das ist die abstrakte Grundlage der L^2–Theorie für Fourier–Reihen. Die Fourier–Theorie hat noch viel mehr Schönes zu bieten: Integrieren und Differenzieren, punktweise und gleichmäßige Konvergenz, explizite Fourier–Entwicklungen und praktische Fehlerabschätzungen, etc.

Vom Skalarprodukt zur Norm und zurück

Lemma C1I: Parallelogrammgleichung und Polarisation

Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Für jedes Skalarprodukt $\langle - \mid - \rangle$ und zugehörige Norm $|u| = \sqrt{\langle u \mid u \rangle}$ gilt die **Parallelogrammgleichung**

$$|u+v|^2 + |u-v|^2 = 2|u|^2 + 2|v|^2 \quad \text{für alle } u,v \in V.$$



Über $\mathbb{C}$ rekonstruieren wir das Skalarprodukt dank **Polarisationsformel**

$$\langle u \mid v \rangle = \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^{-k}|u+\mathrm{i}^k v|^2 = \frac{|u+v|^2 - |u-v|^2}{4} + \mathrm{i}\frac{|u-\mathrm{i}v|^2 - |u+\mathrm{i}v|^2}{4}.$$

Über $\mathbb{R}$ gilt die Polarisationsformel entsprechend ohne den Imaginärteil.

Satz C1J: Jordan–von Neumann 1935

Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Erfüllt eine beliebige Norm $|-|$ auf V die Parallelogrammgleichung, so definiert die Polarisationsformel ein Skalarprodukt $\langle - \mid - \rangle$ auf V, das einzige Skalarprodukt mit Norm $|-|$.

Vektorräume mit Norm: ℓ^p

Beispiel / Übung C1K: der Banach–Raum ℓ^p

Sei Ω eine Menge. Auf dem Funktionenraum $\mathbb{K}^\Omega = \{ f : \Omega \to \mathbb{K} \}$ definieren wir die **Supremumsnorm**, auch **ℓ^∞–Norm** genannt, durch

$$\|f\|_\infty := |f|_\Omega = \sup\big\{ |f(x)| \;\big|\; x \in \Omega \big\}.$$

Die beschränkten Funktionen bilden den Untervektorraum

$$\ell^\infty(\Omega, \mathbb{K}) := \big\{ f : \Omega \to \mathbb{K} \;\big|\; \|f\|_\infty < \infty \big\}.$$

Entsprechend definieren wir die **ℓ^p–Norm** für $1 \le p < \infty$ durch

$$\|f\|_p := \big(\textstyle\sum_{x\in\Omega} |f(x)|^p\big)^{1/p}.$$

Die p–summierbaren Funktionen bilden den Untervektorraum

$$\ell^p(\Omega, \mathbb{K}) := \big\{ f : \Omega \to \mathbb{K} \;\big|\; \|f\|_p < \infty \big\}.$$

Hierauf ist $\|-\|_p$ tatsächlich eine Norm dank **Minkowski–Ungleichung**. Besonders einfach und prominent sind die drei Spezialfälle $p = 1, 2, \infty$.

Beispiel / Übung C1K: der Banach–Raum ℓ^p, fortgesetzt

Der Raum $L^p(\Omega, \mathscr{A}, \mu)$ ist vollständig, also ein Banach–Raum.

Für je zwei konjugierte Exponenten $1 \le p, q \le \infty$ mit $1/p + 1/q = 1$ erhalten wir das **Produkt** $\cdot : \ell^p \times \ell^q \to \ell^1$ und daraus die **Paarung**

$$\langle - \mid - \rangle \;:\; \ell^p \times \ell^q \to \mathbb{K} \;:\; \langle f \mid g \rangle = \sum_{x \in \Omega} \overline{f(x)}\, g(x).$$

Das Produkt $\overline{f} g$ ist absolut summierbar dank der **Hölder–Ungleichung**:

$$\|f \cdot g\|_{\ell^1} \le \|f\|_{\ell^p} \cdot \|g\|_{\ell^q}$$

Im Spezialfall $p = q = 2$ ist dies die Cauchy–Schwarz–Ungleichung, und die Paarung ist das oben diskutierte Skalarprodukt auf ℓ^2.

Für alle $1 < p < \infty$ ist dies eine duale Paarung, das heißt $(\ell^p)' \cong \ell^q$.

Für $\sharp\Omega \ge 2$ und $p \ne 2$ erfüllt $\|-\|_{\ell^p}$ nicht die Parallelogrammgleichung:

$$\|e_a + e_b\|_{\ell^p}^2 + \|e_a - e_b\|_{\ell^p}^2 = 2^{2/p} + 2^{2/p} \quad \ne \quad 2\|e_a\|_{\ell^p}^2 + 2\|e_b\|_{\ell^p}^2 = 2 + 2$$

Speziell für $\Omega = \{1, \dots, n\}$ erhalten wir die ℓ^p–Normen auf $\mathbb{K}^n$.
Ist Ω unendlich, so erhalten wir für verschiedene $p \in [1, \infty]$ verschiedenen Vektorräume $\ell^p(\Omega, \mathbb{K})$. Sie enthalten alle den Untervektorraum $\mathbb{K}^{(\Omega)}$ der Funktionen mit endlichem Träger.

Auch die Exponenten $0 < p < 1$ und sogar $p = 0$ liefern interessante Beispiele für „Normen“, allerdings nicht mit allen Eigenschaften.

Vektorräume mit Norm: L^p

Beispiel / Übung C1K: der Banach–Raum L^p

Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, etwa $\Omega \subset \mathbb{R}^n$ offen mit Lebesgue–Maß μ oder Ω diskret mit Zählmaß $\mu(A) = \sharp A$. Für $1 \le p < \infty$ definieren wir die **L^p–Halbnorm** und jede messbare Funktion $f : \Omega \to \mathbb{K}$ durch

$$\|f\|_p := \left(\int_{x \in \Omega} |f(x)|^p \, \mathrm{d}x \right)^{1/p}.$$

Die p–integrierbaren Funktionen bilden den Untervektorraum

$$\mathscr{L}^p(\Omega, \mathscr{A}, \mu) := \left\{ f : \Omega \to \mathbb{K} \;\middle|\; \|f\|_p < \infty \right\}.$$

Der relevante Vektorraum ist der Quotientenvektorraum

$$L^p(\Omega, \mathscr{A}, \mu) := \mathscr{L}^p(\Omega, \mathscr{A}, \mu) / \mathscr{N}.$$

Hierauf ist $\|-\|_p$ tatsächlich eine Norm dank **Minkowski–Ungleichung**.

⚠ Wir unterscheiden Funktionen f und Äquivalenzklassen $\bar{f} = f + \mathscr{N}$. Eine solche „$L^p$–Funktion“ $\bar{f} \in L^p$ hat keine punktweisen Werte „$\bar{f}(x)$“.

Vektorräume mit Norm: L^p

Beispiel / Übung C1K: der Banach–Raum L^p, fortgesetzt

Die **L^∞–Halbnorm** definieren wir durch das **essentielle Supremum**:

$$\begin{aligned} |f|_\Omega &= \sup |f(x)| &&:= \inf\left\{ a \in \mathbb{R} \;\middle|\; \quad \{ x \mid |f(x)| > a \} \ = \emptyset \right\}, \\ \|f\|_\infty &= \operatorname{ess\,sup} |f(x)| &&:= \inf\left\{ a \in \mathbb{R} \;\middle|\; \mu(\{ x \mid |f(x)| > a \}) = 0 \right\}. \end{aligned}$$

Die essentiell beschränkten Funktionen bilden den Untervektorraum

$$\mathscr{L}^\infty(\Omega, \mathscr{A}, \mu) = \left\{ f : \Omega \to \mathbb{K} \;\middle|\; \|f\|_\infty < \infty \right\}.$$

Der relevante Vektorraum ist der Quotientenvektorraum

$$L^\infty(\Omega, \mathscr{A}, \mu) := \mathscr{L}^\infty(\Omega, \mathscr{A}, \mu) / \mathscr{N}.$$

Hierauf ist $\|-\|_\infty$ tatsächlich eine Norm, genannt die **L^∞–Norm**.

Für das Zählmaß gilt $L^p(\Omega, \mathfrak{P}(\Omega), \sharp) = \mathscr{L}^p(\Omega, \mathfrak{P}(\Omega), \sharp) = \ell^p(\Omega)$.
Für $1 < p < q < \infty$ gilt $\ell^1 \subset \ell^p \subset \ell^q \subset \ell^\infty$, im Allgemeinen strikt.
Für $\mathrm{vol}(\Omega) < \infty$ gilt $L^1 \supset L^p \supset L^q \supset L^\infty$, im Allgemeinen strikt.

Vektorräume mit Norm: L^p

Beispiel / Übung C1K: der Banach–Raum L^p, fortgesetzt

Der Raum $L^p(\Omega, \mathscr{A}, \mu)$ ist vollständig, also ein Banach–Raum.

Für je zwei konjugierte Exponenten $1 \le p, q \le \infty$ mit $1/p + 1/q = 1$ erhalten wir das **Produkt** $\cdot : L^p \times L^q \to L^1$ und daraus die **Paarung**

$$\langle - \mid - \rangle \;:\; L^p \times L^q \to \mathbb{K} \;:\; \langle f \mid g \rangle := \int_{x \in \Omega} \overline{f(x)}\, g(x)\, \mathrm{d}x.$$

Das Produkt $\overline{f}g$ ist absolut integrierbar dank der **Hölder–Ungleichung**:

$$\|f \cdot g\|_{L^1} \le \|f\|_{L^p} \cdot \|g\|_{L^q}$$

Im Spezialfall $p = q = 2$ ist dies die **Cauchy–Schwarz–Ungleichung**, und die Paarung ist das oben diskutierte Skalarprodukt auf L^2.

Für alle $1 < p < \infty$ ist dies eine **duale Paarung**, das heißt $(L^p)' \cong L^q$.

☺ Die L^p–Räume sind zentral für Analysis, Numerik und Stochastik: Erwartungswert, Co/Varianz, höhere Momente, Grenzwertsätze, ...

Vektorräume mit Norm: L^p

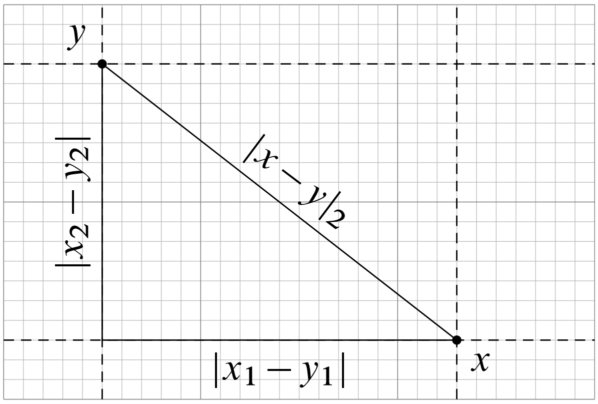
Die L^p–Räume spielen eine zentrale Rolle in der **Analysis**, insb. der Maß- und Integrationstheorie. Der Buchstabe L erinnert an Henri Lebesgue (1875–1941) und seinen grundlegenden Integralbegriff.

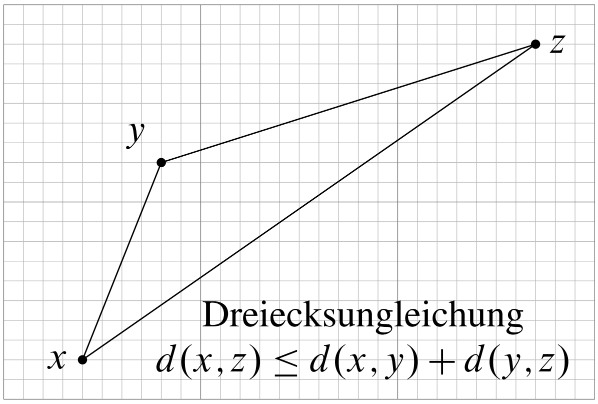
Die L^p–Norm sind ebenso wichtig in der **Numerik** und **Optimierung**: sie wertet Abweichungen / bestraft Fehler mit dem Exponenten p. Besonders prominent sind auch hier die Fälle $p = 1, 2, \infty$: L^1-, quadratische, gleichmäßige Approximation.

Die L^p–Räume finden sich auch ganz natürlich in der **Stochastik**: Dort ist $(\Omega, \mathscr{A}, \mu)$ ein Wahrscheinlichkeitsraum, also $\mu(\Omega) = 1$, und messbare Funktionen $f : \Omega \to \mathbb{R}$ modellieren Zufallsvariablen. Das Integral $m_1 = \int_{x \in \Omega} f(x)\, \mathrm{d}x$ ist der Erwartungswert, dazu ist offensichtlich absolute Integrierbarkeit $f \in \mathscr{L}^1$ die Minimalforderung.

Quadratische Integrierbarkeit $f \in \mathscr{L}^2$ sichert die Existenz der Varianz $m_2 = \int_{x \in \Omega} |f(x) - m_1|^2\, \mathrm{d}x$, das Skalarprodukt misst die Covarianz. Entsprechend nutzt man die höheren Momente $m_p = \int_{x \in \Omega} |f(x) - m_1|^p\, \mathrm{d}x$, zum Beispiel im zentralen Grenzwertsatz mit expliziten Fehlerschranken.

Metrische Räume: die euklidische Metrik





Satz C2A: die euklidische Metrik

Auf $X = \mathbb{R}^n$ definieren wir die **euklidische Metrik** $d : X \times X \to \mathbb{R}$ durch

$$d(x,y) := |x-y|_2 = \sqrt{\langle x-y \mid x-y \rangle} = \sqrt{|x_1-y_1|^2+\dots+|x_n-y_n|^2}.$$

Diese Abbildung erfreut sich folgender Eigenschaften für alle $x,y,z \in X$:

M0: $d(x,y) \ge 0 = d(x,x)$ (Positivität)

M1: $d(x,y) > 0$ für $x \ne y$ (Definitheit)

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

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

Metrische Räume: die euklidische Metrik

Beweis: Klar sind (M0–2) einzig die Dreiecksungleichung (M3) ist nicht trivial. Sie folgt aus der Cauchy–Schwarz–Ungleichung (C1A) dank Dreiecksungleichung (N3): $|x-z|_2 = |x-y+y-z|_2 \le |x-y|_2+|y-z|_2$.

🙂 Dies formuliert anschauliche Eigenschaften des Abstands: (M0) Der Abstand von x nach x ist Null, der von x nach y ist Null oder positiv. (M1) Der Abstand zwischen $x \ne y$ ist positiv. Umgekehrt: Ist der Abstand von x nach y gleich Null, so ist dies nur für $x = y$ möglich. Anders gesagt: Aus $d(x,y) = 0$ folgt $x = y$. (M2) Der Abstand von x nach y ist gleich dem Abstand von y nach x. (M3) Die Dreiecksungleichung besagt, dass in einem Dreieck die Summe der Längen von zwei Seiten nicht kleiner sein kann als die Länge der dritten Seite, siehe Abbildung C:2. Anschaulich gesagt: Der Weg von x nach z wird nicht kürzer, möglicherweise länger, wenn wir einen Umweg über y machen.

Metrische Räume: allgemeine Definition

Übung: Beweisen Sie diese Aussagen für die euklidischen Metrik:

M0: $d(x,y) \ge 0 = d(x,x)$ (Positivität)

M1: $d(x,y) > 0$ für $x \ne y$ (Definitheit)

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

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

🙂 Wir erheben diese grundlegenden Eigenschaften nun zur Definition:

Definition C2B: Metrik

Eine **Metrik** auf einer Menge X ist eine Abbildung $d : X \times X \to [0,\infty]$, die obige Bedingungen (M0–3) erfüllt. Das Paar (X,d) nennen wir dann einen **metrischen Raum** mit Grundmenge X und Metrik d.

Für eine **Halbmetrik** d fordern wir nur (M0,2,3), nicht unbedingt (M1): Allgemeiner als eine Metrik darf eine Halbmetrik zwei verschiedenen Punkten $x \ne y$ in X den Abstand $d(x,y) = 0$ beimessen.

Metrische Räume: allgemeine Definition

Manche Autoren erlauben für Metriken nicht den Wert ∞. Gilt $d(x,y) < \infty$ für alle $x,y \in X$, so nennen wir d zur Betonung eine **endliche Metrik**. Gelegentlich ist es aber bequem, den Begriff der Metrik etwas zu erweitern und auch den Wert $d(x,y) = \infty$ zuzulassen. Hierdurch wird keines der folgenden Argumente komplizierter, aber manches einfacher! Für die nötigen Rechnungen in $[0,\infty]$ vereinbaren wir $0 \le a \le \infty$ sowie $a + \infty = \infty + a = \infty$ für alle $a \in [0,\infty]$; dies entspricht den üblichen Konvention der erweiterten Zahlengeraden.

Durch Weglassen der Forderung (M1) schwächen wir den Begriff der Metrik zum Begriff der Halbmetrik ab. Halbmetriken sind meist weniger wirkungsvoll, kommen aber in manchen Anwendungen natürlich vor und sind daher ein sinnvoller Hilfsbegriff. Auch wenn man ihn schließlich vermeiden will, so ist es meist bequem ihn zunächst zuzulassen.

Euklidische und diskrete Metrik

Beispiel: Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Jedes Skalarprodukt $\langle - \mid - \rangle$ auf V definiert durch $|u| = \sqrt{\langle u \mid u \rangle}$ eine zugehörige Norm.

Jede Halb/Norm $|-| : V \to [0, \infty]$ definiert eine zugehörige Halb/Metrik $d : V \times V \to [0, \infty] : (x, y) \mapsto |x - y|$, denn aus (N0–3) folgt sofort (M0–3):

$$d(x,z) = |x-z| = |x-y+y-z| \le |x-y| + |y-z| = d(x,y) + d(y,z)$$

Prominenteste Beispiele sind $\mathbb{K}^n$ sowie $\ell^p(\Omega)$ und $L^p(\Omega, \mathscr{A}, \mu)$.

Beispiel / Übung C2C: die diskrete Metrik

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

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

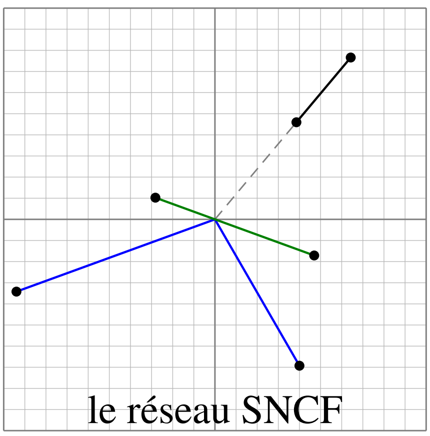
Die für jede Metrik geforderten Eigenschaften (M0–3) prüft man leicht.

Eisenbahnmetrik und Wegmetrik

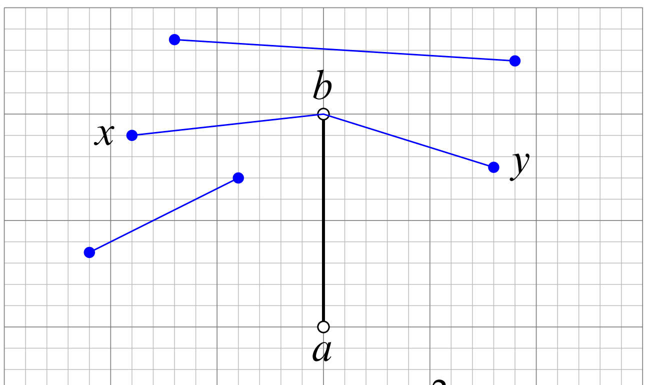
Beispiel C2D: die französische Eisenbahnmetrik

Die **französische Eisenbahnmetrik** auf der Menge $\mathbb{R}^n$ ist

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



le réseau SNCF



kürzeste Wege in $X = \mathbb{R}^2 \setminus \mathrm{Int}[a,b]$

Beispiel: Sei $X \subset \mathbb{R}^n$ ein euklidischer Teilraum oder allgemein ein metrischer Raum (X, d). Wir definieren die **Wegmetrik** $\bar{d} : X \times X \to [0, \infty]$ für je zwei Punkte $x, y \in X$ als die infimale Länge aller Wege von x nach y.

Neue Räume aus alten: Teil und Summe

Beispiel / Übung C2E: Teilraum

Sei (X,d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Dann ist die Einschränkung $d_A := d|_{A\times A} : A\times A \to [0,\infty]$ eine Metrik auf A, genannt die **Teilraummetrik** von A in (X,d). Hierdurch wird (A,d_A) selbst zu einem metrischen Raum, und wir nennen (A,d_A) einen **Teilraum** von (X,d).

Soweit nicht anders vereinbart, statten wir jede Teilmenge $A \subset X$ mit dieser Metrik aus und machen sie so zu einem Teilraum (A,d_A). Insbesondere erbt jede Teilmenge $A \subset \mathbb{R}^n$ die euklidische Metrik.

Beispiel / Übung C2F: Summenraum

Sei $(X_i,d_i)_{i\in I}$ eine Familie metrischer Räume (X_i,d_i) mit $i \in I$. Auf ihrer disjunkten Vereinigung $X = \bigsqcup_{i\in I} X_i$ definieren wir $d : X\times X \to [0,\infty]$ durch

$$d(x,y) = \begin{cases} d_i(x,y) & \text{falls } x,y \in X_i, \\ \infty & \text{falls } x \in X_i \text{ und } y \in X_j \text{ mit } i \neq j. \end{cases}$$

Dies ist eine Metrik auf X, die wir die **Summenmetrik** nennen.

Neue Räume aus alten: Produkt und Potenz

Beispiel / Übung C2G: Produktraum

Sei $(X_i,d_i)_{i\in I}$ eine Familie metrischer Räume (X_i,d_i) mit $i \in I$.
Auf ihrem Produkt $X = \prod_{i\in I} X_i$ definieren wir die **Supremumsmetrik**

$$d \;:\; X\times X \to [0,\infty] \;:\; d(x,y) = \sup\big\{\, d_i(x_i,y_i) \;\big|\; i \in I \,\big\}.$$

Wir nennen $(X,d) = \prod_{i\in I}(X_i,d_i)$ den **Produktraum** von $(X_i,d_i)_{i\in I}$.
Allgemein haben wir auf $X = \prod_{i\in I} X_i$ die **ℓ^p–Metrik** für $1 \le p \le \infty$.
Ist I endlich und sind alle Metriken d_i endlich, so ist auch d endlich.

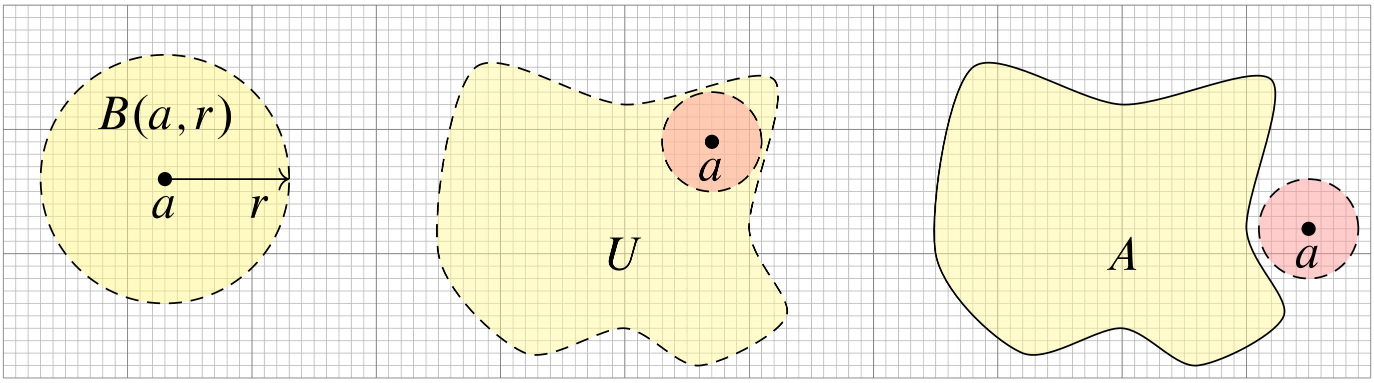
Beispiel / Übung C2H: Abbildungsraum

Sei X eine Menge und (Y,d_Y) ein metrischer Raum. Wir betrachten die Menge $Y^X = \mathrm{Abb}(X,Y)$ aller Abbildungen $f,g : X \to Y$. Hierauf definieren wir die **Supremumsmetrik** $d : Y^X \times Y^X \to [0,\infty]$ wie zuvor durch

$$d(f,g) = \big|d_Y(f,g)\big|_X = \sup\big\{\, d_Y(f(x),g(x)) \;\big|\; x \in X \,\big\}.$$

Allgemein haben wir auf Y^X ebenso die **ℓ^p–Metrik** für $1 \le p \le \infty$.

Offene und abgeschlossene Mengen



Definition C2L: metrische Topologie

Sei (X,d) ein metrischer Raum. Zum Punkt $a \in X$ und Radius $r \in [0,\infty]$ ist

$$B(a,r) = B_X(a,r) = B_d(a,r) = B_{(X,d)}(a,r) := \{\, x \in X \mid d(a,x) < r \,\},$$
$$\bar{B}(a,r) = \bar{B}_X(a,r) = \bar{B}_d(a,r) = \bar{B}_{(X,d)}(a,r) := \{\, x \in X \mid d(a,x) \le r \,\},$$
$$S(a,r) = S_X(a,r) = S_d(a,r) = S_{(X,d)}(a,r) := \{\, x \in X \mid d(a,x) = r \,\}$$

der **offene Ball** / der **abgeschlossene Ball** / die **Sphäre**.

Offene und abgeschlossene Mengen

Definition C2L: metrische Topologie, fortgesetzt

Wir nennen $U \subset X$ **Umgebung** von a, falls $B(a,\varepsilon) \subset U$ für ein $\varepsilon \in \mathbb{R}_{>0}$:

$$\exists \varepsilon \in \mathbb{R}_{>0}\ \forall x \in X : \left[\, d(a,x) < \varepsilon \Rightarrow x \in U \,\right]$$

Das System aller Umgebungen von a in (X,d) bezeichnen wir mit

$$\mathcal{U}_a = \mathcal{U}_a(X,d) := \{\, U \subset X \mid \exists \varepsilon \in \mathbb{R}_{>0} : B(a,\varepsilon) \subset U \,\}.$$

Eine Menge $U \subset X$ ist **offen**, wenn sie Umgebung jedes ihrer Punkte ist:

$$\forall a \in U\ \exists \varepsilon \in \mathbb{R}_{>0} : B(a,\varepsilon) \subset U$$
$$\Leftrightarrow \quad \forall a \in U\ \exists \varepsilon \in \mathbb{R}_{>0}\ \forall x \in X : \left[\, d(a,x) < \varepsilon \Rightarrow x \in U \,\right]$$

Das System aller offenen Mengen im Raum (X,d) ist die **Topologie**

$$\mathcal{T} = \mathcal{T}(X,d) := \{\, U \subset X \mid U \text{ offen in } (X,d) \,\}$$
$$= \{\, U \subset X \mid \forall a \in U\ \exists \varepsilon \in \mathbb{R}_{>0} : B(a,\varepsilon) \subset U \,\}.$$

Eine Menge $A \subset X$ heißt **abgeschlossen**, wenn $X \smallsetminus A$ offen ist.

Metrische Topologie: erste Eigenschaften

Proposition C2M: Eigenschaften der metrischen Topologie

Sei (X,d) ein metrischer Raum (oder allgemeiner d eine Halbmetrik).

Jeder offene Ball $B(a,r)$ ist offen in der Topologie von (X,d),
jeder abgeschlossene Ball $\bar{B}(a,r)$ ist abgeschlossen.

Die offenen Mengen in (X,d) erfreuen sich folgender Eigenschaften:

O1: Die leere Menge $\emptyset$ und der Raum X sind offen:

$$\emptyset, X \in \mathcal{T}$$

O2: Sind $U_1,\dots,U_n$ offen, dann auch $U_1 \cap \dots \cap U_n$:

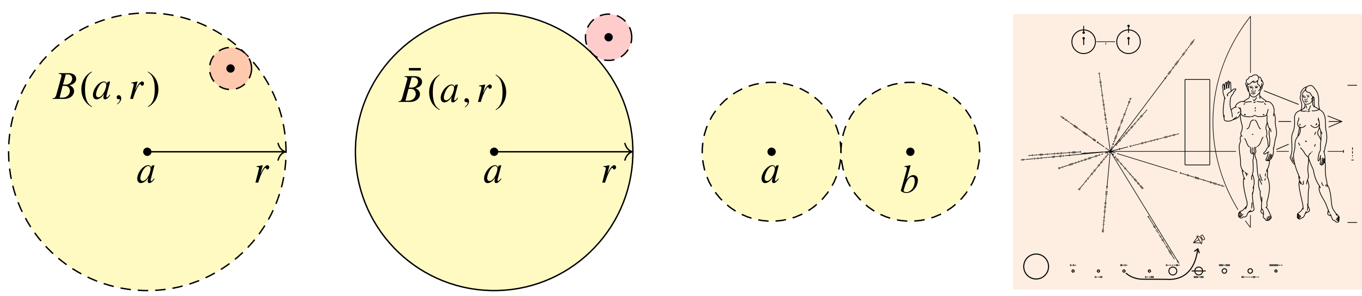
$$\forall n \in \mathbb{N} \;\; \forall U_1,\dots,U_n \in \mathcal{T} \;:\; U_1 \cap \dots \cap U_n \in \mathcal{T}$$

O3: Sind alle U_i offen für $i \in I$, dann auch $\bigcup_{i\in I} U_i$:

$$\forall (I \to \mathcal{T} : i \mapsto U_i) \;:\; \bigcup_{i\in I} U_i \in \mathcal{T}$$

Die abgeschlossenen Mengen sind komplementär zu den offenen,
somit stabil unter endlichen Vereinigungen und beliebigen Schnitten.

Metrische Topologie: erste Eigenschaften

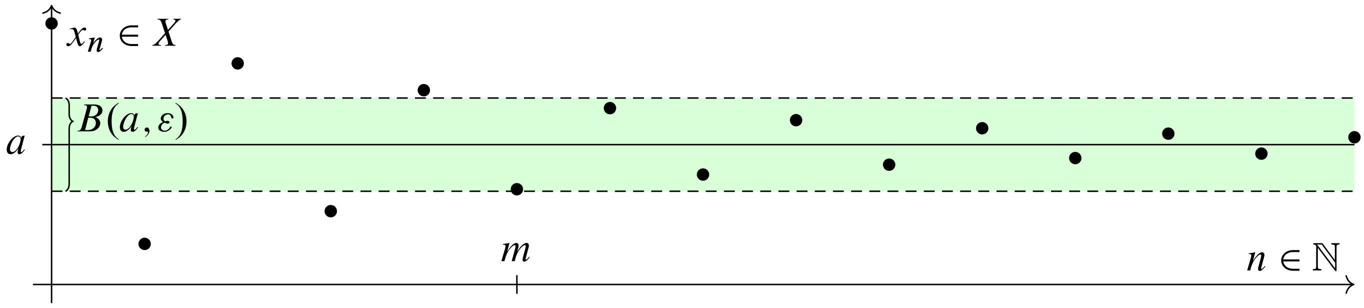


Proposition C2O: die Hausdorff–Eigenschaft

In jedem metrischen Raum (X,d) gilt die Hausdorff–Eigenschaft (T_2):
Zu je zwei Punkten $a \neq b$ in X existieren disjunkte offene Umgebungen.

Beweis: Es gilt $r := d(a,b) > 0$ dank (M1). Die Bälle $U = B(a,r/2)$ und $V = B(b,r/2)$ sind offen (C2M), erfüllen $a \in U$ und $b \in V$ und sind disjunkt: Wäre $x \in U \cap V$, so führten Symmetrie (M2) und Dreiecksungleichung (M3) zum Widerspruch $r = d(a,b) \le d(a,x) + d(x,b) < r/2 + r/2 = r$. QED

Konvergenz einer Folge in einem metrischen Raum



Definition C3A: Konvergenz einer Folge

Gegeben seien eine Folge $(x_n)_{n\in\mathbb{N}}$ und ein Punkt a in einem metrischen Raum (X,d). Die Folge $(x_n)_{n\in\mathbb{N}}$ **konvergiert** gegen a, falls $d(x_n,a) \to 0$:

$$(x_n)_{n\in\mathbb{N}} \to a \text{ in } (X,d) \quad :\Leftrightarrow \quad d(x_n,a) \to 0 \text{ für } n \to \infty$$

$$\Leftrightarrow \quad \forall \varepsilon \in \mathbb{R}_{>0}\ \exists m \in \mathbb{N}\ \forall n \in \mathbb{N}_{\geq m} : d(x_n,a) < \varepsilon$$

$$\Leftrightarrow \quad \forall \varepsilon \in \mathbb{R}_{>0}\ \exists m \in \mathbb{N}\ \forall n \in \mathbb{N}_{\geq m} : x_n \in B(a,\varepsilon)$$

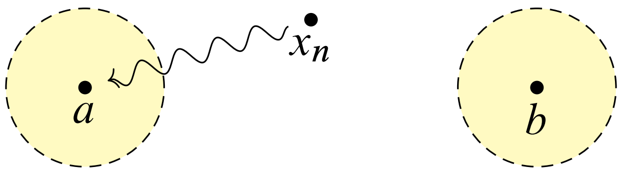
$\Leftrightarrow$ Jede Umgebung $U \in \mathscr{U}_a$ enthält schließlich alle Folgenglieder x_n.

Die letzte Formulierung verwendet die Metrik nicht explizit, sondern nur offene Mengen: Konvergenz ist somit eine topologische Eigenschaft!

Konvergenz einer Folge in einem metrischen Raum

Konvergenz in (X,d) ist eine **zweistellige Relation** zwischen Folgen $(x_n)_{n\in\mathbb{N}}$ und Punkten a. Wir schreiben hierfür kurz $x_n \to a$ oder $x_n \xrightarrow{d} a$. Zur Betonung der Indizes $n \in \mathbb{N}$ schreiben wir „$x_n \to a$ für $n \to \infty$“.

Die bequeme Schreibweise $\lim x_n = a$ ist problematisch, solange die Konvergenz nicht geklärt ist: eventuell existiert kein Grenzwert „$\lim x_n$“! Kann eine Folge in (X,d) womöglich mehrere Grenzwerte haben?

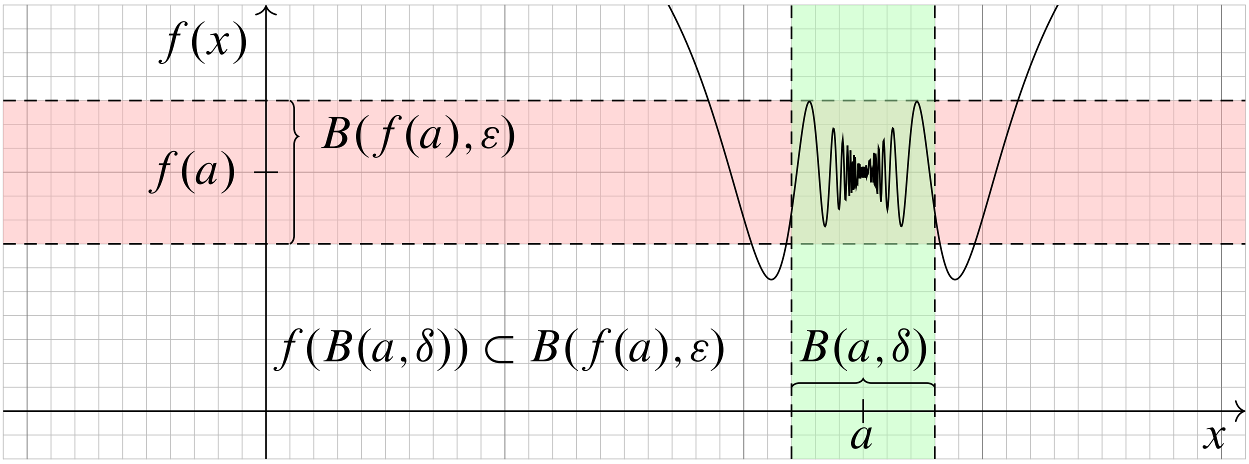


Satz C3B: Eindeutigkeit des Grenzwertes

Sei (X,d) ein metrischer Raum. Jede Folge $(x_n)_{n\in\mathbb{N}}$ in (X,d) hat höchstens einen Grenzwert: Aus $x_n \to a$ und $x_n \to b$ folgt $a = b$.

Beweis: Sei $x_n \to a$ und $a \neq b$. Dank Hausdorff–Eigenschaft existieren $U \in \mathscr{U}_a$ und $V \in \mathscr{U}_b$ mit $U \cap V = \emptyset$. Wegen $x_n \to a$ existiert ein Index $m \in \mathbb{N}$ mit $x_n \in U$ für alle $n \geq m$, folglich gilt $x_n \notin V$, somit $x_n \not\to b$. QED

Stetigkeit von Abbildungen



🙂 Für eine reelle Funktion $f:\mathbb{R}\to\mathbb{R}$ umschreibt man Stetigkeit gemeinhin so, dass der Graph von f „keine Sprünge macht“, oder sich „ohne abzusetzen in einem Zug zeichnen lässt“.

☹ Das mag der Intuition nützen, ist aber keine tragfähige Definition! Anschauung kann die Definition motivieren, sie aber nicht ersetzen.

Probieren Sie es: Ist die Funktion $h:\mathbb{Q}\to\mathbb{R}: x\mapsto \operatorname{sign}(x^2-2)$ stetig?

Stetigkeit von Abbildungen

Definition C3E: Stetigkeit von Abbildungen

Eine Abbildung $f:X\to Y$ heißt **stetig im Punkt** $a\in X$, wenn zu jedem $\varepsilon\in\mathbb{R}_{>0}$ ein $\delta\in\mathbb{R}_{>0}$ existiert, sodass $f(B_X(a,\delta))\subset B_Y(f(a),\varepsilon)$ gilt:

$$\forall\varepsilon\in\mathbb{R}_{>0}\ \exists\delta\in\mathbb{R}_{>0}\ \forall x\in X\ :\ \big[d_X(a,x)<\delta\ \Rightarrow\ d_Y(f(a),f(x))<\varepsilon\big]$$

Gilt dies in jedem Punkt $a\in X$, so nennen wir $f:X\to Y$ **stetig (auf X)**:

$$\forall a\in X\ \forall\varepsilon\in\mathbb{R}_{>0}\ \exists\delta\in\mathbb{R}_{>0}\ \forall x\in X\ :\ \big[d_X(a,x)<\delta\ \Rightarrow d_Y(f(a),f(x))<\varepsilon\big]$$

Strenger nennen wir $f:X\to Y$ **gleichmäßig stetig (auf X)**, wenn gilt:

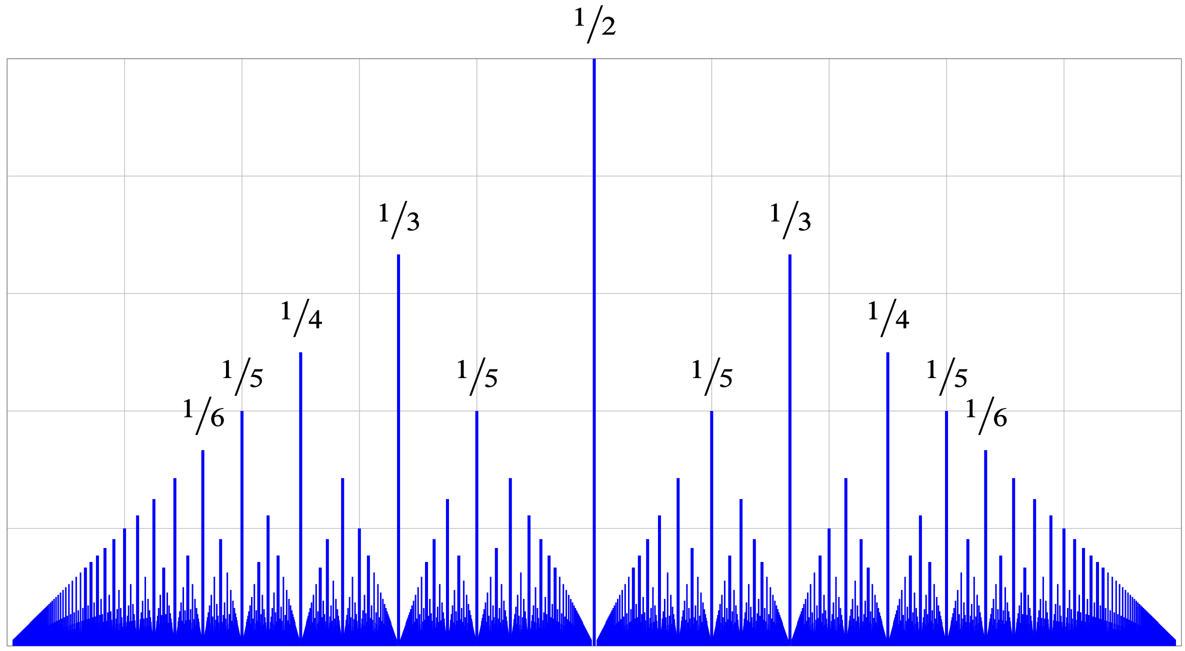
$$\forall\varepsilon\in\mathbb{R}_{>0}\ \exists\delta\in\mathbb{R}_{>0}\ \forall a\in X\ \forall x\in X\ :\ \big[d_X(a,x)<\delta\ \Rightarrow\ d_Y(f(a),f(x))<\varepsilon\big]$$

Mit $\mathscr{C}(X,Y)=\mathscr{C}(X,d_X;Y,d_Y)$ bezeichnen wir die Menge aller stetigen Abbildungen $f:X\to Y$. Ist die Abbildung f bijektiv sowie f und f^{-1} stetig, so nennen wir f einen **Homöomorphismus** (A1A).

🙂 Diese Begriffe und zahlreiche Beispiele kennen Sie aus der Analysis!

Mahnende Beispiele für un/stetige Funktionen

Beispiel: Die **Dirichlet–Funktion** $\mathbf{I}_\mathbb{Q} : \mathbb{R} \to \mathbb{R}$ ist nirgends stetig.
Das Produkt $f : \mathbb{R} \to \mathbb{R} : x \mapsto x\,\mathbf{I}_\mathbb{Q}(x)$ ist stetig in 0, aber nirgends sonst.



Beispiel: Die **kleine Dirichlet–Funktion** $g : \mathbb{R} \to \mathbb{R}$ mit $g(p/q) = 1/q$ für jeden gekürzten Bruch $p/q \in \mathbb{Q}$ mit $q \in \mathbb{Z}_{\ge 1}$ und $g(x) = 0$ für $x \in \mathbb{R} \smallsetminus \mathbb{Q}$ ist unstetig in jedem rationalen Punkt, aber stetig in jedem irrationalen.

Stetigkeit ist eine topologische Eigenschaft.

Satz C3G: Charakterisierungen der Stetigkeit

Für jede Abbildung $f : X \to Y$ metrischer Räume sind äquivalent:
(1) f ist folgenstetig: Aus $x_n \to a$ in X folgt $f(x_n) \to f(a)$ in Y.
(2) f ist stetig auf X (im Sinne obiger ε–δ–Definition C3E).
(3) Für jede offene Menge V in Y ist das Urbild $f^{-1}(V)$ offen in X:

$$\forall V \in \mathfrak{T}_Y \;:\; f^{-1}(V) \in \mathfrak{T}_X$$

Lokal um jeden Punkt $a \in X$ sind äquivalent:
(4) f ist folgenstetig in a: Aus $x_n \to a$ in X folgt $f(x_n) \to f(a)$ in Y.
(5) f ist stetig im Punkt a (im Sinne obiger ε–δ–Definition C3E).
(6) Für jede Umgebung V von $f(a)$ ist $f^{-1}(V)$ eine Umgebung von a:

$$\forall V \in \mathscr{U}_{f(a)} \;:\; f^{-1}(V) \in \mathscr{U}_a$$

Die Formulierungen (3,6) verwenden die Metrik nicht explizit, sondern nur offene Mengen: Stetigkeit ist somit eine topologische Eigenschaft!

Konvergenz und Stetigkeit sind somit topologische Eigenschaften. In der Analysis geht es über die **qualitativen** Fragen der Konvergenz und Stetigkeit hinaus auch **quantitativ** darum, wie schnell eine Folge konvergiert oder wie stark $f(x)$ mit x variiert. Im euklidischen Raum $\mathbb{R}^n$ oder in jedem normierten $\mathbb{K}$–Vektorraum ist das wichtigste Instrument die Differenzierbarkeit zur linearen Approximation, oder noch höhere Ableitungen zur polynomialen Approximation (Taylor–Entwicklung). Zur feineren Untersuchung dienen ebenso die Lipschitz–Stetigkeit (§C2f) oder die Hölder–Stetigkeit (§C2g). Speziell in der Numerik wird eine geeignete Metrik explizit für praktische Fehlerschranken genutzt.

Topologisch äquivalente Metriken

In vielen Fällen kommt es uns nicht auf die Metrik selbst an, sondern nur auf Konvergenz und Stetigkeit, also insbesondere auf offene Mengen!

Definition C3I: topologischer Vergleich von Metriken

Zwei Metriken $d,e : X \times X \to [0,\infty]$ heißen **(topologisch) äquivalent**, wenn sie dieselben offenen Mengen definieren, kurz $\mathcal{T}(X,d) = \mathcal{T}(X,e)$.

Wir nennen d **feiner** als e, oder e **gröber** als d, falls $\mathcal{T}(X,d) \supset \mathcal{T}(X,e)$. Der Kürze halber bedeutet „feiner" hierbei „mindestens so fein wie".

Äquivalent, jeder ε–Ball bezüglich e enthält einen δ–Ball bezüglich d:

$$\forall a \in X \;\; \forall \varepsilon \in \mathbb{R}_{>0} \;\; \exists \delta \in \mathbb{R}_{>0} \; : \; B_{(X,d)}(a,\delta) \subset B_{(X,e)}(a,\varepsilon).$$

Äquivalenz bedeutet also, d ist feiner als e, und e ist feiner als d. Wir nennen d **echt feiner** als e, falls $\mathcal{T}(X,d) \supsetneq \mathcal{T}(X,e)$ gilt.

Topologisch äquivalente Metriken

Proposition C3J: topologischer Vergleich von Metriken

Für je zwei Metriken d,e auf X sind gleichbedeutend:

1. Die Metrik d ist feiner als die Metrik e (im Sinne der Definition C3I).
2. Jede Umgebung bezüglich e ist auch eine Umgebung bezüglich d.
3. Jede offene Menge bezüglich e ist auch offen bezüglich d.
4. Die Identität $\mathrm{id}_X : (X,d) \to (X,e) : x \mapsto x$ ist stetig.
5. Gilt $x_n \to a$ bezüglich d, dann gilt auch $x_n \to a$ bezüglich e.
6. Ist $f : X \to Y$ stetig bezüglich e, dann ist f stetig bezüglich d.
7. Ist $f : Y \to X$ stetig bezüglich d, dann ist f stetig bezüglich e.

Seien $|-|$ und $\|-\|$ Normen auf einem $\mathbb{K}$–Vektorraum V.

8. Genau dann ist $|-|$ feiner als $\|-\|$, falls $\|x\| \le L\,|x|$ für alle $x \in V$ und eine geeignete Konstante $L \in \mathbb{R}_{>0}$ gilt.
9. Topologische Äquivalenz bedeutet $\ell\,|x| \le \|x\| \le L\,|x|$ für alle $x \in V$ und geeignete Konstanten $\ell, L \in \mathbb{R}_{>0}$.

Beispiel / Übung C3K: Vergleich von L^p–Normen

Zur Illustration einige klassische Beispiele: Sei $1 \le p < q \le \infty$.

1. Auf $\mathbb{R}^n$ sind die ℓ^p–Norm und die ℓ^q–Norm äquivalent:
$$|x|_\infty \le |x|_p \le |x|_1 \le n \cdot |x|_\infty$$
2. Auf $\mathbb{R}^{(\mathbb{N})} \subset \mathbb{R}^{\mathbb{N}}$ ist die ℓ^p–Norm echt feiner als die ℓ^q–Norm:
$$\|f\|_{\ell^q} \le \|f\|_{\ell^p}$$
3. Auf $\mathscr{C}([0,1],\mathbb{R})$ ist die L^q–Norm echt feiner als die L^p–Norm:
$$\|f\|_{L^p} \le \|f\|_{L^q}$$
4. Auf $\mathscr{C}_c(\mathbb{R},\mathbb{R})$ sind die L^p– und die L^q–Norm unvergleichbar, das heißt, keine ist feiner als die andere.

Beispiel / Übung C3L: gestauchte & gestutzte Metrik

Ist $d : X \times X \to [0,\infty]$ eine Metrik, so auch die **gestauchte Metrik**

$$d' \ : \ X \times X \to [0,1] \ : \ (x,y) \mapsto \frac{d(x,y)}{1+d(x,y)}.$$

Hieraus lässt sich die ursprüngliche Metrik d rekonstruieren gemäß

$$d \ : \ X \times X \to [0,\infty] \ : \ (x,y) \mapsto \frac{d'(x,y)}{1-d'(x,y)}.$$

Hierbei entspricht $d(x,y) = \infty$ dem Wert $d'(x,y) = 1$ und umgekehrt.

Entsprechend erhalten wir aus d die **gestutzte Metrik**

$$d^* \ : \ X \times X \to [0,1] \ : \ (x,y) \mapsto \min\{\, d(x,y), 1\,\}.$$

Die drei Metriken d, d' und d^* sind topologisch äquivalent. Insbesondere können wir von d zu einer endlichen Metrik d' oder d^* übergehen.

Beispiel / Übung C3L: gestauchte & gestutzte Metrik, fortgesetzt

Die hierbei verwendete Skalierung $f : [0,\infty] \to [0,\infty]$ mit $f(t) = t/(1+t)$ bzw. $f(t) = \min\{t, 1\}$ erfreut sich folgender drei Eigenschaften:

(1) f ist positiv definit: Es gilt $f(0) = 0$ und $f(t) > 0$ für $t > 0$.
(2) f ist monoton wachsend: Für alle $s \le t$ in $[0,\infty]$ gilt $f(s) \le f(t)$.
(3) f ist subadditiv: Für alle $s,t \in [0,\infty]$ gilt $f(s+t) \le f(s) + f(t)$.

In diesem Fall ist mit d auch $d' = f \circ d$ eine Metrik und feiner als d.

Ist f zudem stetig in 0, dann sind d und d' topologisch äquivalent: Mit d ist auch d^α für $0 < \alpha \le 1$ eine Metrik und topologisch äquivalent.

Die Funktion $f = \mathbf{I}_{]0,\infty]} : [0,\infty] \to \{0,1\}$ erfüllt (1–3). Für jede Metrik d ist $d' = f \circ d$ die diskrete Metrik, und i.A. nicht topologisch äquivalent.

Skalierung von Metriken

Skalierung von Metriken

Abstand zwischen Mengen

Definition C3M: Abstand zwischen Punkten und Mengen

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

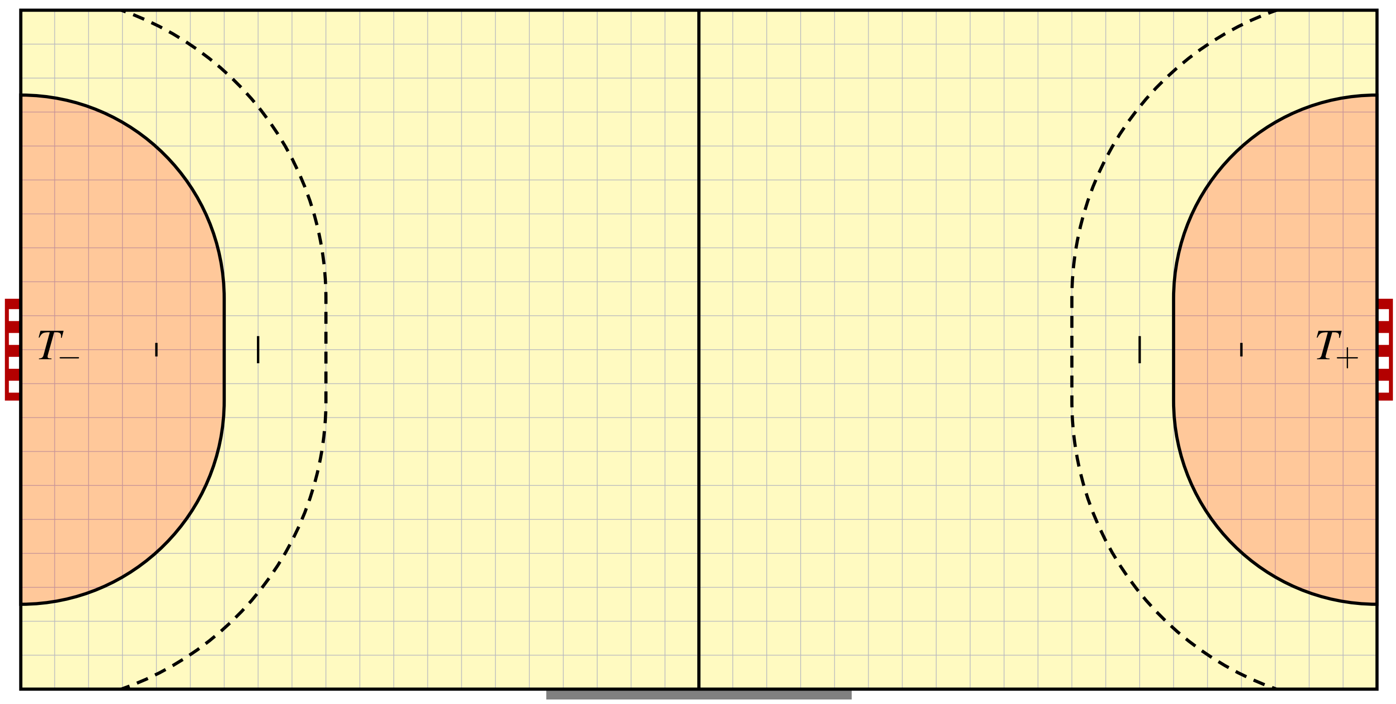
$$d(a,B) := \inf\{\, d(a,b) \,\big|\, b \in B \,\}$$
$$d(A,B) := \inf\{\, d(a,b) \,\big|\, a \in A,\ b \in B \,\}$$

Übung: Zeichnen Sie ein Handballfeld nach folgenden Maßgaben:

$X = [-20,20] \times [-10,10] \subset \mathbb{R}^2$	Spielfeld
$M = \{0\} \times [-10,10]$	Mittellinie
$T = \{-20,20\} \times [-1.5,1.5]$	die beiden Tore
$S = \{\, x \in X \mid d(x,T) = 6 \,\}$	Sechs-Meter-Linie („Kreis“)
$N = \{\, x \in X \mid d(x,T) = 9 \,\}$	Neun-Meter-Linie („Freiwurflinie“)

Wie sieht das Feld bezüglich der euklidischen Metrik aus?
und der Maximumsmetrik (ℓ^∞)? und der Taximetrik (ℓ^1)?

Handballfeld in euklidischer Metrik



Abstand und Abschluss

Proposition C3N: Abstand und Abschluss

Sei (X,d) ein metrischer Raum und $A, B, A_i \subset X$ Teilmengen.

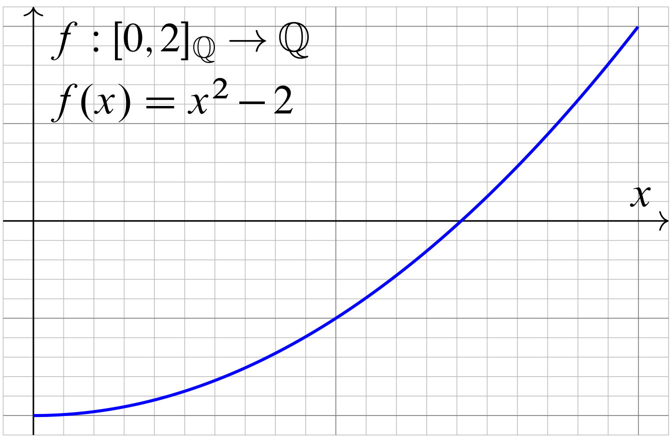
1. Die Funktion $f : X \to [0,\infty] : x \mapsto d(x,A)$ erfüllt $|f(x) - f(y)| \le d(x,y)$. Gleiches gilt für $g(x) = \inf_{i \in I} d(x, A_i)$ und $h(x) = \sup_{i \in I} d(x, A_i)$.
2. Für $x \in A$ gilt $d(A,x) = 0$. Genau dann ist A abgeschlossen, wenn auch die Umkehrung gilt: Für alle $x \in X$ mit $d(A,x) = 0$ folgt $x \in A$.
3. Seien $A, B \subset X$ nicht-leer, abgeschlossen und disjunkt. Dann ist die zugehörige **Urysohn–Funktion**

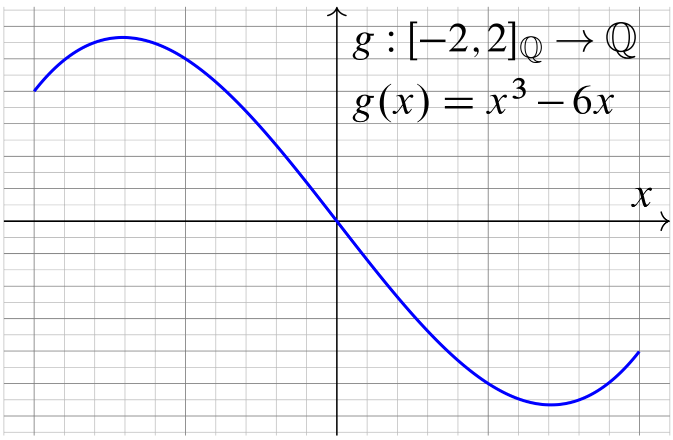
$$f \;:\; X \to [0,1] \;:\; x \mapsto \frac{d(x,A)}{d(x,A) + d(x,B)}$$

stetig und erfüllt $f^{-1}(\{0\}) = A$ und $f^{-1}(\{1\}) = B$.

☺ Mit stetigen Funktionen $f : X \to \mathbb{R}$ können wir somit alle Punkte $a \neq b$ in X trennen und sogar alle disjunkten abgeschlossenen Mengen A, B.

Zwischenwerte, Minimum und Maximum





Satz C3R: Zwischenwertsatz / Zusammenhang

Jede stetige Funktion $f : [a,b] \to \mathbb{R}$ hat die Zwischenwerteigenschaft: Zu jedem $y \in \mathbb{R}$ mit $f(a) \le y \le f(b)$ existiert $x \in [a,b]$ mit $f(x) = y$.

Satz C3S: Minimum und Maximum / Kompaktheit

Jede stetige Funktion $f : [a,b] \to \mathbb{R}$ nimmt Minimum und Maximum an: Es existieren $x_0, x_1 \in [a,b]$ sodass $f(x_0) \le f(x) \le f(x_1)$ für alle $x \in [a,b]$.

Konvergente Folgen und Cauchy–Folgen

Eine Folge $(x_n)_{n\in\mathbb{N}}$ in (X,d) **konvergiert** gegen einen Punkt $a \in X$, wenn der Abstand $d(x_n,a)$ schließlich beliebig klein wird. Ausführlich:

$$(x_n)_{n\in\mathbb{N}} \to a \text{ in } (X,d) \quad :\Leftrightarrow \quad d(x_n,a) \to 0 \text{ für } n \to \infty$$
$$:\Leftrightarrow \quad \forall \varepsilon \in \mathbb{R}_{>0}\ \exists m \in \mathbb{N}\ \forall n \in \mathbb{N}_{\geq m} : d(x_n,a) < \varepsilon$$

Wir nennen dann die Folge $(x_n)_{n\in\mathbb{N}}$ **konvergent** und a ihren **Grenzwert**.

Konvergenz in (X,d) ist eine **zweistellige Relation** zwischen Folgen $(x_n)_{n\in\mathbb{N}}$ und Punkten a in X. Wir schreiben hierfür kurz $x_n \to a$.

Oft noch praktischer wäre ein Kriterium für die Folge $(x_n)_{n\in\mathbb{N}}$ allein.

Definition C4A: Cauchy–Folge

Wir nennen $(x_n)_{n\in\mathbb{N}}$ **Cauchy–Folge** in (X,d), wenn der Durchmesser $\delta_n := \sup_{p,q\geq n} d(x_p,x_q)$ eine Nullfolge ist. In Quantorenschreibweise:

$$\forall \varepsilon \in \mathbb{R}_{>0}\ \exists n \in \mathbb{N}\ \forall p,q \geq n : d(x_p,x_q) \leq \varepsilon$$

Jede konvergente Folge ist eine Cauchy–Folge, aber nicht umgekehrt.

Vollständige metrische Räume

Lemma C4B: Konvergenz impliziert Cauchy.

Jede konvergente Folge $(x_n)_{n\in\mathbb{N}} \to a$ ist eine Cauchy–Folge.

Beweis: Sei $\varepsilon > 0$. Dank Konvergenz $x_n \to a$ existiert $n \in \mathbb{N}$, sodass $\mathrm{d}(x_p,a) < \varepsilon/2$ für alle $p \geq n$. Für alle $p,q \geq n$ folgt somit $d(x_p,x_q) \leq d(x_p,a) + d(a,x_q) < \varepsilon$. Also ist $(x_n)_{n\in\mathbb{N}}$ Cauchy–Folge. QED

Definition C4C: vollständiger metrischer Raum

Ein metrischer Raum (X,d) heißt **vollständig**, wenn jede Cauchy–Folge $(x_n)_{n\in\mathbb{N}}$ in (X,d) konvergiert, also ein Punkt $a \in X$ existiert mit $x_n \to a$.

Beispiel: Im Raum $X =]0,1]$ mit euklidischer Metrik ist $x_n = 2^{-n}$ eine Cauchy–Folge, aber nicht konvergent. (Im Raum $[0,1]$ gilt $x_n \to 0$.)

Beispiel: Im Raum $\mathbb{Q}$ mit euklidischer Metrik ist $x_n = \sum_{k=0}^{n} 1/k!$ eine Cauchy–Folge, hat aber in $\mathbb{Q}$ keinen Grenzwert. (In $\mathbb{R}$ hingegen gilt $x_n \to \mathrm{e} = 2.71828\ldots$, aber dieser Grenzwert liegt nicht in $\mathbb{Q}$.)

Vollständige metrische Räume

Satz C4D: Vollständigkeit der reellen Zahlen

Der Raum $\mathbb{R}$ ist vollständig bezüglich der Metrik $d(x,y) = |x-y|$.

Ebenso sind $\mathbb{R}^m$ und $\mathbb{C}^m$ vollständig bezüglich der p–Metrik, $p \in [1,\infty]$.

Beweis: (1) Die definierende Eigenschaft des geordneten Körpers $(\mathbb{R},+,\cdot,<)$ der reellen Zahlen ist die Supremums-Vollständigkeit (B2C).

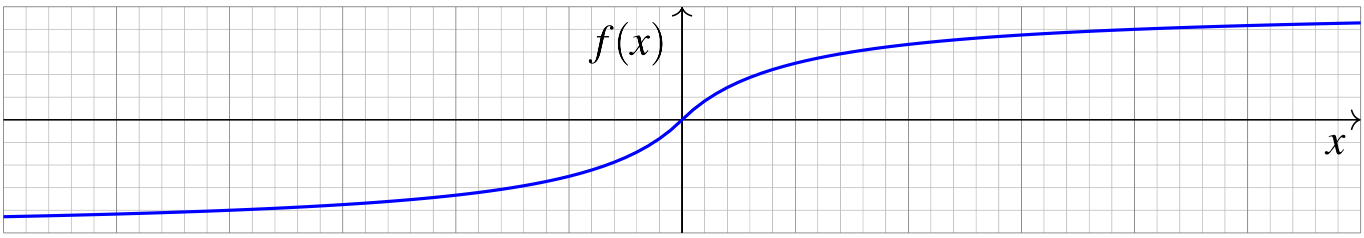
Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchy–Folge in $\mathbb{R}$. Sie ist insbesondere beschränkt, also liegen $a = \liminf x_n$ und $b = \limsup x_n$ in $\mathbb{R}$. Dabei gilt $a \le b$.

Zu $\varepsilon \in \mathbb{R}_{>0}$ sei $n \in \mathbb{N}$ sodass $|x_p - x_q| \le \varepsilon$ für alle $p,q \ge n$. Insbesondere gilt $x_n - \varepsilon \le x_p \le x_n + \varepsilon$ für alle $p \ge n$, also $x_n - \varepsilon \le a \le b \le x_n + \varepsilon$, somit $b - a \le 2\varepsilon$. Da dies für alle $\varepsilon \in \mathbb{R}_{>0}$ gilt, folgt $a = b$. Also gilt $x_n \to a$.

(2) Erinnerung: Für alle $x \in \mathbb{R}^m$ gilt $|x|_\infty \le |x|_p \le |x|_1 \le m \cdot |x|_\infty$.

Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchy–Folge in $\mathbb{R}^m$. Für die k–te Koordinate ist dann $(x_{n,k})_{n\in\mathbb{N}}$ eine Cauchy–Folge in $\mathbb{R}$. Dank Vollständigkeit (1) existiert ein Punkt $a_k \in \mathbb{R}$ mit $x_{n,k} \to a_k$. Daraus folgt $x_n \to a$ in $\mathbb{R}^m$. QED

Metrische Vollständigkeit ist keine topologische Eigenschaft!

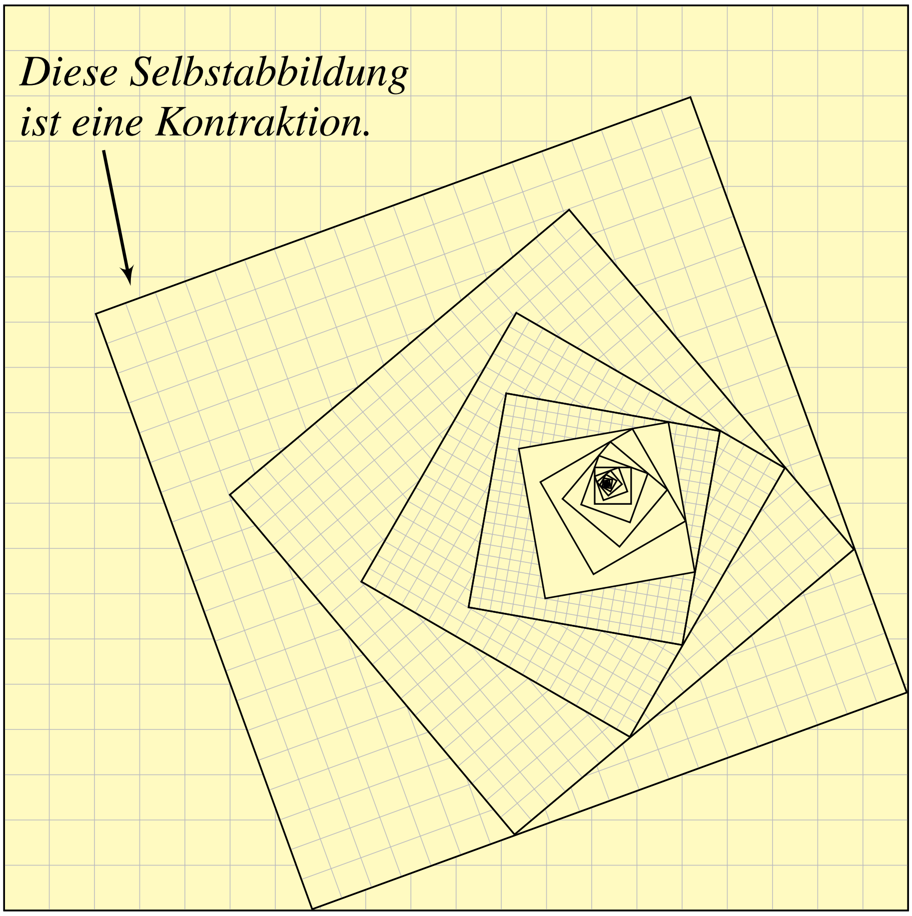


(1) Metrische Vollständigkeit bleibt unter Homöomorphismen im Allgemeinen nicht erhalten: Wir haben $(f,g) : \mathbb{R} \cong \left]-1,+1\right[$ mit $f(x) = x/(1+|x|)$ und $g(x) = x/(1-|x|)$ aus A1B. Bezüglich der euklidischen Metrik ist $\mathbb{R}$ vollständig, $\left]-1,+1\right[$ jedoch nicht.

(2) Cauchy–Folgen bleiben nicht erhalten unter Homöomorphismen: Die Folge $y_n = n/(1+n) \nearrow 1$ in $\left]-1,+1\right[$ ist eine Cauchy–Folge bezüglich der euklidischen Metrik, die Folge $x_n = g(y_n) = n \nearrow +\infty$ in $\mathbb{R}$ jedoch nicht.

(3) Vollständigkeit bleibt nicht erhalten für äquivalente Metriken: Der Raum $\mathbb{R}$ mit der euklidischen Metrik $e(x,y) = |x-y|$ ist vollständig nicht jedoch mit der hierzu äquivalenten Metrik $d(x,y) = |f(x) - f(y)|$.

Illustration zu Banachs Fixpunktsatz



Nicht jede Kontraktion $f : X \to X$ des Raums (X,d) hat einen Fixpunkt!

Ein Gegenbeispiel auf $X = \mathbb{R}^n \setminus \{0\}$ ist $f : X \to X : x \mapsto x/2$.

Wir benötigen zudem, dass der Raum (X,d) vollständig ist!

Banachs Fixpunktsatz ist ein universelles Werkzeug.

Aufgabe: (1) Illustrieren Sie Iterationen und Banachs Fixpunktsatz.
(2) Sie kennen bereits spektakuläre Anwendungen: Nennen Sie einige!
(3) Wiederholung: Formulieren und beweisen Sie Banachs Fixpunktsatz.

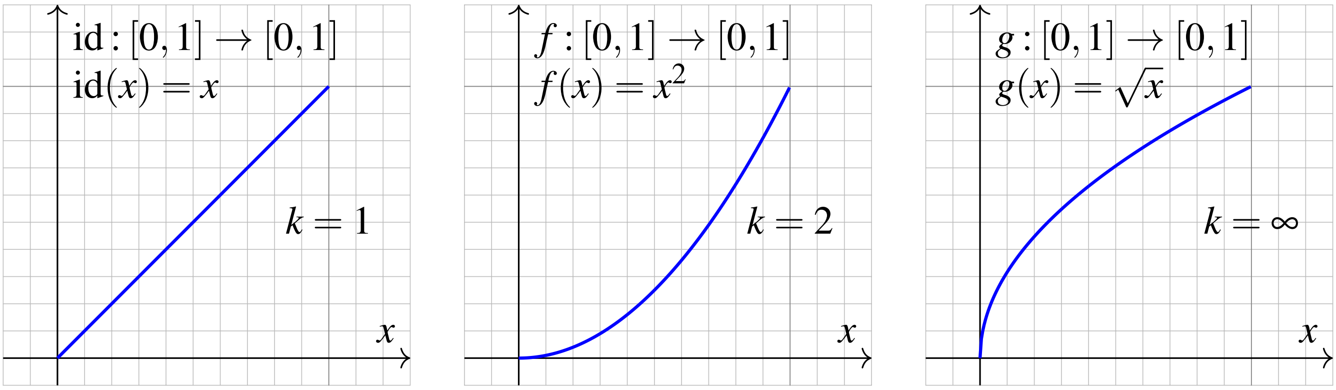
Lösung: (1) Bildhaftes Beispiel: Wenn Sie von Ihrer geographischen Umgebung X eine Landkarte im Maßstab $k = 1 : n$ (mit $n > 1$) vor sich auf den Tisch legen, dann definiert die Zuordnung jedes realen Punktes zu seinem Bildpunkt eine k–kontraktive Abbildung $f : X \to X$. Genau ein Punkt der Karte liegt auf dem geographischen Punkt, den er bezeichnet.

(2) Mit dem Iterationsverfahren können Sie viele Gleichungen numerisch lösen wie $x = \cos(x)$ für $x \in [0,1]$. Das **Newton–Verfahren** baut darauf auf und verbessert ganz wesentlich die Konvergenzgeschwindigkeit.

Weitere Anwendungen sind der **Satz von Picard–Lindelöf** zur Lösung von Differentialgleichungen $y' = f(x,y)$ und der **lokale Umkehrsatz** zur Konstruktion lokaler Diffeomorphismen $(f,g) : \mathbb{R}^n \supseteq U \cong V \subseteq \mathbb{R}^n$.

Das ist schon spektakulär – und doch nur die ersten Anwendungen im Grundstudium, als Spitze des Eisbergs! Wir fügen noch weitere hinzu.

Dehnungsschranken und Kontraktionen



Gegeben seien metrische Räume (X,d_X) und (Y,d_Y) und $k \in \mathbb{R}_{\geq 0}$. Eine Abbildung $f : X \to Y$ nennen wir **k–lipschitz–stetig**, falls gilt:

$$d_Y(f(a), f(b)) \leq k\, d_X(a,b) \qquad \text{für alle } a,b \in X$$

Im Falle $0 \leq k < 1$ nennen wir f **kontraktiv** oder eine **k–Kontraktion**.

Beispiel: Für $f : \mathbb{R}^n \to \mathbb{R}^m : x \mapsto Ax + b$ mit $A \in \mathbb{R}^{m\times n}$ und $b \in \mathbb{R}^n$ gilt $|f(x) - f(y)| = |A(x-y)| \leq \|A\| \cdot |x-y|$ bezüglich der Operatornorm.

Beispiel: Sei $X \subseteq \mathbb{R}^n$ konvex und $f : X \to \mathbb{R}^m$ diff'bar mit $\|f'\| \leq k$. Für alle $x,y \in X$ existiert $z \in [x,y]$, sodass $f(x) - f(y) = f'(z)(x-y)$. Demnach gilt $|f(x) - f(y)| \leq \|f'(z)\| \cdot |x-y| \leq k|x-y|$.

Dehnungsschranken und Kontraktionen

Wir nennen eine solche Funktion f auch **dehnungsbeschränkt**. Der **metrische Differenzenquotient** ist hier beschränkt gemäß

$$\frac{d_Y(f(a), f(b))}{d_X(a,b)} \leq k \qquad \text{für alle } a \neq b \text{ in } X.$$

Für $f : (X,d_X) \to (Y,d_Y)$ definieren wir daher die **Lipschitz–Norm**

$$\|f\| = \|f\|_{\mathrm{Lip}} = \mathrm{Lip}(f) := \sup\left\{ \frac{d_Y(f(a), f(b))}{d_X(a,b)} \;\middle|\; a \neq b \text{ in } X \right\}.$$

Die folgende Formulierung vereinheitlicht Ausnahmen und Sonderfälle:

$$\|f\| := \inf\left\{ k \in \mathbb{R}_{\geq 0} \;\middle|\; \forall a,b \in X \,:\, d_Y(f(a), f(b)) \leq k\, d_X(a,b) \right\}.$$

Dies entspricht der **Operatornorm** von linearen Abbildungen normierter Vektorräume. Genau dann gilt $\|f\| < \infty$, wenn f lipschitz–stetig ist.

Genau dann gilt $\|f\| = 0$, wenn f konstant ist. Speziell für die Identität $\mathrm{id}_X : X \to X$ gilt $\|\mathrm{id}_X\| = 1$, im Sonderfall $X = \{x\}$ jedoch nur $\|f\| = 0$.

Für die Komposition von Abbildungen gilt $\|g \circ f\| \leq \|g\| \cdot \|f\|$.

Banachs Fixpunktsatz

Satz C4Q: Fixpunktsatz von Banach

Sei (X,d) ein metrischer Raum, nicht-leer und vollständig. Hierauf sei $f : X \to X$ eine k–Kontraktion: Wir haben eine Kontraktionskonstante $k \in [0,1[$, und für alle $x,y \in X$ gilt $d(f(x), f(y)) \le k\, d(x,y)$. Dann folgt:

(1) Zur Abbildung f existiert genau ein Fixpunkt $a \in X$, mit $f(a) = a$.

(2) Jede Iteration mit $x_0 \in X$ und $x_{n+1} = f(x_n)$ konvergiert gegen a.

(3) Für alle $n \in \mathbb{N}_{\ge 1}$ gelten dabei die beiden Fehlerschranken

$$d(a,x_n) \;\le\; \underbrace{\frac{k}{1-k} d(x_n,x_{n-1})}_{\text{a posteriori}} \;\le\; \underbrace{\frac{k^n}{1-k} d(x_1,x_0)}_{\text{a priori}} \;\searrow\; 0.$$

Zusatz: Allgemeiner genügt $d(f^n(x), f^n(y)) \le k_n\, d(x,y)$ für alle $x,y \in X$ und alle $n \in \mathbb{N}$ sowie $\sum_{n=0}^{\infty} k_n < \infty$. Damit gilt die feinere Fehlerschranke

$$d(a,x_n) \;\le\; d(x_1,x_0) \sum_{j=n}^{\infty} k_j \;\searrow\; 0.$$

Banachs Fixpunktsatz

Die Aussage (1) garantiert Existenz und Eindeutigkeit des Fixpunktes. Die Konstruktion (2) liefert zudem eine extrem praktische Approximation.

Gemäß (3) ist dabei die Konvergenz $x_n \to a$ mindestens so schnell wie die Konvergenz der geometrischen Folge $k^n \searrow 0$: Dies nutzen wir bei iterativen Berechnungen als *a priori* Abschätzung des Zeitaufwandes.

Dieser wunderbare Satz geht auf Stefan Banach (1892–1945) zurück, der nützliche Zusatz stammt von Johannes Weissinger (1913–1995). Diese Verallgemeinerung ist dabei nicht schwieriger zu beweisen.

Zum Beispiel genügt es, dass eine gewisse Iteration f^m kontraktiv ist, also $d(f^m(x), f^m(y)) \le k\, d(x,y)$ für eine Konstante $k \in [0,1[$ erfüllt.

Ebenso kommt es vor, dass höhere Iterationen f^n stärker kontrahieren, und die Reihe $\sum k_n$ viel kleiner ausfällt als die geometrische $\sum k^n$.

Für die qualitative Konvergenzaussage ist dies zunächst unwesentlich, doch für die praktische Fehlerabschätzung ist es ungemein hilfreich.

Banachs Fixpunktsatz

Beweis: *Eindeutigkeit:* Für je zwei Fixpunkte $a = f(a)$ und $b = f(b)$ gilt $d(a,b) = d(f(a), f(b)) \le k\,d(a,b)$ mit $k < 1$, also $d(a,b) = 0$, somit $a = b$.

Die *Existenz* beweisen wir durch die iterative Folge $(x_n)_{n\in\mathbb{N}}$:
Wir wählen $x_0 \in X \ne \emptyset$ und setzen $x_{n+1} = f(x_n)$ für alle $n \in \mathbb{N}$.

Per Induktion gilt $d(x_{n+1}, x_n) \le k^n d(x_1, x_0)$: Für $n = 0$ ist dies trivial, für $n \ge 1$ gilt $d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \le k\,d(x_n, x_{n-1}) \le k^n d(x_1, x_0)$.

Dank Dreiecksungleichung erhalten wir für alle $n \le p < q$:

$$\begin{aligned} d(x_q, x_p) &\le d(x_q, x_{q-1}) + \dots + d(x_{p+2}, x_{p+1}) + d(x_{p+1}, x_p) \\ &\le (k^{q-p-1} + \dots + k + 1)\, d(x_{p+1}, x_p) = \frac{1 - k^{q-p}}{1-k} d(x_{p+1}, x_p) \\ &\le \frac{k^p - k^q}{1-k} d(x_1, x_0) \le \frac{k^n}{1-k} d(x_1, x_0) \searrow 0 \quad \text{für } n \to \infty. \end{aligned}$$

Demnach ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchy–Folge im metrischen Raum (X, d). Da (X, d) vollständig ist, existiert ein Grenzwert $a \in X$ mit $x_n \to a$.

Da f kontraktiv und somit stetig ist, folgt aus $x_{n+1} = f(x_n)$ für alle $n \in \mathbb{N}$ per Grenzübergang $a = \lim x_{n+1} = \lim f(x_n) = f(\lim x_n) = f(a)$. $\boxed{\text{QED}}$

Banachs Fixpunktsatz

Obige Ungleichung für $n = p$ und $q \to \infty$ ergibt die Fehlerabschätzung

$$d(a, x_n) \le \underbrace{\frac{k}{1-k} d(x_n, x_{n-1})}_{\text{a posteriori}} \le \underbrace{\frac{k^n}{1-k} d(x_1, x_0)}_{\text{a priori}} \searrow 0.$$

Den Zusatz beweisen wir genauso: Wegen $\sum_{n=0}^{\infty} k_n < \infty$ gilt $k_n \to 0$. Insbesondere existiert $m \in \mathbb{N}$ mit $k_m \le 1/2$, das heißt f^m ist kontraktiv. Sind $a = f(a)$ und $b = f(b)$ Fixpunkte von f, so auch von f^m, also folgt $a = b$.

Für jede iterative Folge mit $x_0 \in X$ und $x_{n+1} = f(x_n)$ für $n \in \mathbb{N}$ erhalten wir für $n \le p < q$ wie oben $d(x_q, x_p) \le d(x_1, x_0) \sum_{j=n}^{\infty} k_j$. Somit ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchy–Folge in (X, d), also existiert ein Grenzwert $a \in X$ mit $x_n \to a$.

Für $n = p$ und $q \to \infty$ erhalten wir die feinere Fehlerabschätzung

$$d(a, x_n) \le d(x_1, x_0) \sum_{j=n}^{\infty} k_j \searrow 0.$$

Wie immer bei Konvergenz gilt auch hier: Ende gut, alles gut. $\boxed{\text{QED}}$

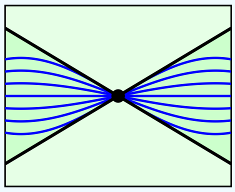
Lösung des AWP durch iterierte Integration

Wir erklären eine allgemeine Lösungsmethode, um 1890 entwickelt von Picard und Lindelöf, für das Anfangswertproblem

$$\boxed{y'(t) = f(t, y(t)), \quad y(t_0) = y_0.}$$

Der folgende Satz garantiert (1) die Existenz und Eindeutigkeit einer Lösung und zudem (2) ein Konstruktionsverfahren zur Approximation dieser Lösung mit (3) expliziter Fehlerschranke. Was will man mehr?

Geometrische Voraussetzungen für die Picard–Lindelöf–Iteration:



Sei $I = [t_0 - a, t_0 + b] \subset \mathbb{R}$ ein Intervall um t_0 mit $a, b \le T$.
Sei $K = \bar{B}(y_0, r) \subset \mathbb{K}^n$ der Ball um y_0 mit Radius $r > 0$.
Sei $f : I \times K \to \mathbb{K}^n$ stetig, somit beschränkt, also $|f| \le M$.
Hierbei gelte $T \cdot M \le r$, notfalls verkleinern wir T und I.
Zudem erfülle f für alle $t \in I$ und $u, v \in K$ die **Lipschitz–Bedingung**

$$\boxed{|f(t,u) - f(t,v)| \le L\,|u - v|.}$$

Ist etwa $f(t,y)$ stetig differenzierbar nach y, so genügt $|\partial f / \partial y| \le L$.

Lösung des AWP durch iterierte Integration

Satz C4R: Picard–Lindelöf–Iteration

Unter den oben erklärten geometrischen Voraussetzungen gilt:

(1) Das Anfangswertproblem hat genau eine Lösung $y : I \to K$.

(2) Die Lösung y ist die Grenzfunktion der Picard–Lindelöf–Iteration

$$u_0 = y_0, \quad u_{n+1}(t) = y_0 + \int_{\tau=t_0}^{t} f(\tau, u_n(\tau))\,\mathrm{d}\tau \quad \text{für } n = 0, 1, 2, 3, \dots.$$

(3) Für alle $n \in \mathbb{N}$ und $t \in I$ gilt die gleichmäßige Fehlerschranke

$$\big|y(t) - u_n(t)\big| \;\le\; \max_I \big|u_1 - u_0\big|\,\mathrm{e}^{LT} \cdot \frac{(LT)^n}{n!} \;\to\; 0.$$

(4) Die Lösung y hängt stetig vom Anfangswert y_0 ab: Ist $\tilde{y} : \tilde{I} \to K$ eine Lösung zum Startwert $\tilde{y}(t_0) = \tilde{y}_0$ auf einem Intervall $\tilde{I}$ mit $t_0 \in \tilde{I}$, so laufen die Lösungen y und $\tilde{y}$ höchstens exponentiell auseinander:

$$|y(t) - \tilde{y}(t)| \;\le\; |y_0 - \tilde{y}_0| \cdot \mathrm{e}^{L|t - t_0|} \quad \text{für alle } t \in I \cap \tilde{I}$$

Lösung des AWP durch iterierte Integration

Aufgabe: Lösen Sie durch Picard–Lindelöf–Iteration das AWP

$$\boxed{y'(t) = y(t), \quad y(0) = 1.}$$

Lösung: Sukzessive Approximation gemäß Picard–Lindelöf:

$$\begin{aligned}
u_0(t) &= 1 \\
u_1(t) &= 1 + \int_{\tau=0}^t u_0(\tau)\,\mathrm{d}\tau = 1 + t \\
u_2(t) &= 1 + \int_{\tau=0}^t u_1(\tau)\,\mathrm{d}\tau = 1 + t + \tfrac{1}{2}t^2 \\
u_3(t) &= 1 + \int_{\tau=0}^t u_2(\tau)\,\mathrm{d}\tau = 1 + t + \tfrac{1}{2}t^2 + \tfrac{1}{3!}t^3 \\
u_4(t) &= 1 + \int_{\tau=0}^t u_3(\tau)\,\mathrm{d}\tau = 1 + t + \tfrac{1}{2}t^2 + \tfrac{1}{3!}t^3 + \tfrac{1}{4!}t^4
\end{aligned}$$

Per Induktion und Grenzübergang $n \to \infty$ finden wir die Lösung:

$$u_n(t) = \sum_{k=0}^{n} \frac{t^k}{k!} \quad \to \quad y(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} = \exp(t)$$

☺ Probe! Dieselbe Lösung finden wir alternativ durch Separation.

Lösung des AWP durch iterierte Integration

Aufgabe: Lösen Sie durch Picard–Lindelöf–Iteration das AWP

$$\boxed{y'(t) = -t\,y(t), \quad y(0) = 1.}$$

Lösung: Sukzessive Approximation gemäß Picard–Lindelöf:

$$\begin{aligned}
u_0(t) &= 1 \\
u_1(t) &= 1 - \int_{\tau=0}^t \tau\,u_0(\tau)\,\mathrm{d}\tau = 1 - \tfrac{1}{2}t^2 \\
u_2(t) &= 1 - \int_{\tau=0}^t \tau\,u_1(\tau)\,\mathrm{d}\tau = 1 - \tfrac{1}{2}t^2 + \tfrac{1}{8}t^4 \\
u_3(t) &= 1 - \int_{\tau=0}^t \tau\,u_2(\tau)\,\mathrm{d}\tau = 1 - \tfrac{1}{2}t^2 + \tfrac{1}{8}t^4 - \tfrac{1}{48}t^6 \\
u_4(t) &= 1 - \int_{\tau=0}^t \tau\,u_3(\tau)\,\mathrm{d}\tau = 1 - \tfrac{1}{2}t^2 + \tfrac{1}{8}t^4 - \tfrac{1}{48}t^6 + \tfrac{1}{384}t^8
\end{aligned}$$

Per Induktion und Grenzübergang $n \to \infty$ finden wir die Lösung:

$$u_n(t) = \sum_{k=0}^{n} \frac{(-1)^k t^{2k}}{2^k k!} \quad \to \quad y(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{2^k k!} = \exp(-t^2/2)$$

☺ Probe! Dieselbe Lösung finden wir alternativ durch Separation.

Lösung des AWP durch iterierte Integration

Aufgabe: (A) Sei $y : I \to K$ stetig. Unsere **Differentialgleichung**

$$y'(t) = f(t, y(t)) \quad \text{für alle } t \in I \text{ und } y(t_0) = y_0$$

ist äquivalent zu folgender (für uns vorteilhaften) **Integralgleichung**:

$$y(t) = y_0 + \int_{\tau=t_0}^{t} f(\tau, y(\tau))\,\mathrm{d}\tau \quad \text{für alle } t \in I$$

Somit sind Lösungen $y : I \to K$ genau die **Fixpunkte** der Abbildung

$$\boxed{\Psi : C(I,K) \to C(I,\mathbb{K}^n), \quad (\Psi u)(t) = y_0 + \int_{\tau=t_0}^{t} f(\tau, u(\tau))\,\mathrm{d}\tau.}$$

(B) Tatsächlich gilt $\Psi : C(I,K) \to C(I,K)$: Wir können Ψ also **iterieren**!

(C) Für je zwei Funktionen $u, v \in C(I,K)$ gilt die **Fehlerschranke**

$$\left|\Psi^n u(t) - \Psi^n v(t)\right| \le \max_{[t_0,t]}\left|u-v\right| \cdot \frac{L^n |t-t_0|^n}{n!} \le \max_{I}\left|u-v\right| \cdot \frac{(LT)^n}{n!}.$$

Für $n \to \infty$ gilt $(LT)^n/n! \to 0$. Demnach ist Ψ^n schließlich **kontraktiv**:
Zu $0 < \alpha < 1$ existiert $n \in \mathbb{N}$ sodass $\left|\Psi^n u - \Psi^n v\right|_I \le \alpha \left|u-v\right|_I$ gilt.

Lösung des AWP durch iterierte Integration

Lösung: (A) Dank HDI. Die Integralgleichung besagt $y = \Psi y$.

(B) Für $u \in C(I,K)$ zeigen wir $\Psi u \in C(I,K)$ durch Nachrechnen:

$$\begin{aligned}\left|\Psi u(t) - y_0\right| &= \left|\int_{t_0}^{t} f(\tau,u(\tau))\,\mathrm{d}\tau\right| \le \left|\int_{t_0}^{t} \left|f(\tau,u(\tau))\right|\mathrm{d}\tau\right| \le \left|\int_{t_0}^{t} M\,\mathrm{d}\tau\right| \\ &= |t-t_0| M \qquad\qquad \le TM \qquad\qquad\qquad \le r\end{aligned}$$

(C) Für $n = 0$ ist die Aussage trivial. Für $n \ge 1$ folgt sie induktiv:

$$\begin{aligned}\left|\Psi^n u(t) - \Psi^n v(t)\right| &= \left|\int_{t_0}^{t} f(\tau, \Psi^{n-1}u(\tau)) - f(\tau, \Psi^{n-1}v(\tau))\,\mathrm{d}\tau\right| \\ &\le \left|\int_{t_0}^{t} L\left|\Psi^{n-1}u(\tau) - \Psi^{n-1}v(\tau)\right|\mathrm{d}\tau\right| \\ &\le \left|\int_{t_0}^{t} \max_{[t_0,t]}\left|u-v\right| \cdot \frac{L^n|\tau-t_0|^{n-1}}{(n-1)!}\,\mathrm{d}\tau\right| \\ &= \max_{[t_0,t]}\left|u-v\right| \cdot \frac{L^n|t-t_0|^n}{n!} \quad \le \quad \max_{I}\left|u-v\right| \cdot \frac{(LT)^n}{n!}\end{aligned}$$

Lösung des AWP durch iterierte Integration

Diese Rechnungen beweisen den Satz C4R von Picard–Lindelöf:

Der Raum $C(I,K)$ ist vollständig bezüglich der Supremumsnorm. Daher können wir auf Ψ den **Fixpunktsatz von Banach** anwenden:

(1) Die Abbildung $\Psi : C(I,K) \to C(I,K)$ hat genau einen Fixpunkt y.
☺ Dank (A) hat unser AWP somit genau eine Lösung $y : I \to K$.

(2) Wir erhalten die Lösung y durch sukzessive Approximation.
☺ Für jede Funktion $u_0 \in C(I,K)$ konvergiert $u_n = \Psi^n u_0$ gegen y.

Für den Fixpunkt $v_0 = y$ liefert die Iteration $v_n = \Psi^n v_0 = y$.
Dank der Fehlerschranke (D) erhalten wir also:

$$\big|y(t) - u_n(t)\big| \;\le\; \max_{[t_0,t]} \big|y - u\big| \cdot \frac{L^n |t-t_0|^n}{n!} \;\le\; \max_I \big|y-u\big| \cdot \frac{(LT)^n}{n!}.$$

☺ Die Konvergenz $u_n \to y$ der Approximationen u_n gegen die Lösung y ist demnach mindestens so schnell wie die Konvergenz $(LT)^n/n! \to 0$.

☺ Die in (3) angegebene Verschärfung nutzt nur u_0, u_1:
Auf der rechten Seite stehen also nur bekannte Daten.

Lösung des AWP durch iterierte Integration

(3) Wir formulieren die Fehlerschranke nur mit u_0 und u_1:

$$\big|u_n(t) - u_{n+1}(t)\big| = \big|\Psi^n u_0(t) - \Psi^n u_1(t)\big| \le \max_{[t_0,t]} \big|u_0 - u_1\big| \cdot \frac{L^n |t-t_0|^n}{n!}$$

$$\big|u_n(t) - u_{n+p}(t)\big| \le |u_n(t) - u_{n+1}(t)| + \dots + |u_{n+p-1}(t) - u_{n+p}(t)|$$

$$\le \sum_{k=n}^{\infty} \max_{[t_0,t]} \big|u_0 - u_1\big| \cdot \frac{L^k |t-t_0|^k}{k!} \;\le\; \max_{[t_0,t]} \big|u_1 - u_0\big| \, \mathrm{e}^{L|t-t_0|} \cdot \frac{L^n |t-t_0|^n}{n!}$$

Für $p \to \infty$ erhalten wir die gewünschte Ungleichung für $|u_n(t) - y(t)|$.

(4) Ist $\tilde{y} : \tilde{I} \to K$ eine Lösung zum Startwert $\tilde{y}(t_0) = \tilde{y}_0$, so gilt dank (A) $\tilde{y}(t) = \tilde{y}_0 + \int_{t_0}^t f(\tau, \tilde{y}(\tau))\,\mathrm{d}\tau$. Wir beginnen die Picard–Lindelöf–Iteration mit der Funktion $u_0 = \tilde{y}$ und erhalten im ersten Schritt die Verschiebung

$$u_1(t) = y_0 + \int_{t_0}^t f(\tau, \tilde{y}(\tau))\,\mathrm{d}\tau = y_0 - \tilde{y}_0 + u_0(t), \quad \text{dank (3) also}$$

$$\big|y(t) - \tilde{y}(t)\big| \le \big|y_0 - \tilde{y}_0\big| \sum_{k=0}^{\infty} \frac{L^k |t-t_0|^k}{k!} = |y_0 - \tilde{y}_0| \, \mathrm{e}^{L|t-t_0|}.$$