Universität des Saarlandes
Fakultät für Mathematik und Informatik
Fachrichtung Mathematik

Analysis II

Skript

Sommersemester 2024

(Version vom 27. September 2024)

Prof. Dr. Roland Speicher
mit der Unterstützung von
Dr. Johannes Hoffmann

Inhaltsverzeichnis

1. Eigenschaften von $\mathcal{C}[0,1]$: sup-Norm und Vollständigkeit

In diesem Kapitel betrachten wir die Menge

$$\mathcal{C}[0,1] := \{f : [0,1] \to \mathbb{C} \mid f \text{ stetig}\}$$

aller stetigen Funktionen, die das abgeschlossene *Einheitsintervall* $[0,1] \subseteq \mathbb{R}$ in den Körper der komplexen Zahlen $\mathbb{C}$ abbilden.

Für beliebige Funktionen $[0,1] \to \mathbb{C}$ betrachten wir zunächst unterschiedliche Konvergenzarten:

Definition 1.1. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n : [0,1] \to \mathbb{C}$.

(a) Die Folge $(f_n)_{n\in\mathbb{N}}$ *konvergiert punktweise* gegen eine Funktion $f : [0,1] \to \mathbb{C}$, falls

$$\forall x \in [0,1]\ \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N : \quad |f_n(x) - f(x)| < \varepsilon.$$

(b) Die Folge $(f_n)_{n\in\mathbb{N}}$ *konvergiert gleichmäßig* gegen eine Funktion $f : [0,1] \to \mathbb{C}$, falls

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\ \forall x \in [0,1] : \quad |f_n(x) - f(x)| < \varepsilon.$$

Der Unterschied ist, dass im Falle von punktweiser Konvergenz der Index N für jedes x neu gewählt werden kann, während im Falle von gleichmäßiger Konvergenz ein Index N für alle x funktionieren muss.

Definition 1.2. Für $f \in \mathcal{C}[0,1]$ definieren wir die *Supremumsnorm* oder *sup-Norm*

$$\|f\|_\infty := \sup_{x\in[0,1]} |f(x)| = \sup\{|f(x)| : x \in [0,1]\}.$$

Laut dem Satz von Minimum und Maximum (vgl. Satz 10.6 aus Analysis I) ist

$$|f| : [0,1] \to \mathbb{R}, \quad x \mapsto |f(x)|$$

beschränkt, also folgt $\|f\|_\infty = \||f|\|_\infty < \infty$ für alle $f \in \mathcal{C}[0,1]$.

Satz 1.3. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n \in \mathcal{C}[0,1]$ und zusätzlich $f \in \mathcal{C}[0,1]$. Dann sind äquivalent:

(1) Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen f.

(2) Es gilt $\|f_n - f\|_\infty \xrightarrow{n\to\infty} 0$.

Beweis. Es gilt

$$\begin{aligned} \|f_n - f\|_\infty \xrightarrow{n\to\infty} 0 \quad &\Leftrightarrow \quad \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N : \|f_n - f\|_\infty < \varepsilon \\ &\Leftrightarrow \quad \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N : \sup_{x\in[0,1]} |f_n(x) - f(x)| < \varepsilon \\ &\Leftrightarrow \quad \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\ \forall x \in [0,1] : |f_n(x) - f(x)| < \varepsilon. \end{aligned}$$ □

Somit ist obige Definition von gleichmäßiger Konvergenz konsistent mit Definition 17.1 aus Analysis I. Dort haben wir in Satz 17.3 auch schon gesehen:

Satz 1.4. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n \in \mathcal{C}[0,1]$, die gleichmäßig gegen eine Funktion $f : [0,1] \to \mathbb{C}$ konvergiert. Dann ist auch $f \in \mathcal{C}[0,1]$.

Am Beispiel $f_n(x) := x^n$ kann man sehen, dass Stetigkeit im Grenzwert verloren gehen kann, wenn man statt gleichmäßiger Konvergenz nur punktweise Konvergenz fordert.

Bemerkung 1.5. Die Menge $\mathcal{C}[0,1]$ ist ein komplexer Vektorraum mit Addition

$$\mathcal{C}[0,1] \times \mathcal{C}[0,1] \to \mathcal{C}[0,1], \quad (f,g) \mapsto f+g,$$

wobei $(f+g)(x) := f(x) + g(x)$, und Skalarmultiplikation

$$\mathbb{C} \times \mathcal{C}[0,1] \to \mathcal{C}[0,1], \quad (\lambda, f) \mapsto \lambda \cdot f,$$

wobei $(\lambda \cdot f)(x) := \lambda \cdot f(x)$.

Der Nullvektor für die Addition ist dabei natürlich die Funktion $f = 0$, die konstant Null ist, d.h. $f(x) = 0$ für alle $x \in [0,1]$.

Proposition 1.6. Die Supremumsnorm $\|\cdot\|_\infty$ ist eine Norm auf $\mathcal{C}[0,1]$, d.h.:

(i) $\forall f \in \mathcal{C}[0,1]: \quad \|f\|_\infty \geq 0$.

(ii) $\forall f \in \mathcal{C}[0,1]: \quad \|f\|_\infty = 0 \Leftrightarrow f = 0$.

(iii) $\forall \lambda \in \mathbb{C}\ \forall f \in \mathcal{C}[0,1]: \quad \|\lambda f\|_\infty = |\lambda| \cdot \|f\|_\infty$.

(iv) $\forall f, g \in \mathcal{C}[0,1]: \quad \|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. (Dreiecksungleichung)

Man nennt $\mathcal{C}[0,1]$ auch einen *normierten Vektorraum.*

Beweis. (i) Klar.

(ii) Sei $f \in \mathcal{C}[0,1]$. Für $f = 0$ gilt sofort $\|f\|_\infty = 0$. Ist andererseits $\|f\|_\infty = 0$, so gilt für alle $x_0 \in [0,1]$

$$0 \leq |f(x_0)| \leq \sup_{x\in[0,1]} |f(x)| = \|f\|_\infty = 0,$$

also $f(x_0) = 0$ für alle $x_0 \in [0,1]$ und somit $f = 0$.

(iii) Für alle $\lambda \in \mathbb{C}$ und alle $f \in \mathcal{C}[0,1]$ gilt

$$\|\lambda f\|_\infty = \sup_{x\in[0,1]} |\lambda f(x)| = \sup_{x\in[0,1]} \big(|\lambda| \cdot |f(x)|\big) = |\lambda| \cdot \sup_{x\in[0,1]} |f(x)| = |\lambda| \cdot \|f\|_\infty.$$

(iv) Für alle $f, g \in \mathcal{C}[0,1]$ gilt

$$\begin{aligned}\|f+g\|_\infty = \sup_{x\in[0,1]} |f(x)+g(x)| &\leq \sup_{x\in[0,1]} \big(|f(x)| + |g(x)|\big) \\ &\leq \sup_{x\in[0,1]} |f(x)| + \sup_{x\in[0,1]} |g(x)| = \|f\|_\infty + \|g\|_\infty. \quad \square\end{aligned}$$

Daher können wir gleichmäßige Konvergenz in $\mathcal{C}[0,1]$ ähnlich zur Konvergenz in $\mathbb{R}$ oder in $\mathbb{C}$ behandeln.

Definition 1.7. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{C}[0,1]$.

(i) Die Folge $(f_n)_{n\in\mathbb{N}}$ *konvergiert* gegen einen *Grenzwert* (oder auch eine *Grenzfunktion*) $f \in \mathcal{C}[0,1]$, falls

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n \geq N : \quad \|f - f_n\|_\infty < \varepsilon.$$

Wir schreiben dafür wieder $f_n \xrightarrow{n\to\infty} f$.

(ii) Die Folge $(f_n)_{n\in\mathbb{N}}$ ist eine *Cauchy-Folge*, falls

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n, m \geq N : \quad \|f_n - f_m\|_\infty < \varepsilon.$$

Lemma 1.8. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{C}[0,1]$.

(i) Die Folge $(f_n)_{n\in\mathbb{N}}$ hat höchstens einen Grenzwert.

(ii) Wenn $(f_n)_{n\in\mathbb{N}}$ konvergiert, so ist $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge.

Beweis. (i) Seien $f, g \in \mathcal{C}[0,1]$ zwei Grenzwerte von $(f_n)_{n\in\mathbb{N}}$. Dann gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für $n \geq N$ gilt

$$\begin{aligned} \|f - g\|_\infty = \|(f - f_n) + (f_n - g)\|_\infty &\leq \|f - f_n\|_\infty + \|f_n - g\|_\infty \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

also $\|f - g\|_\infty = 0$. Mit den Normeigenschaften folgt dann $f - g = 0$, also $f = g$.

(ii) Sei $f \in \mathcal{C}[0,1]$ mit $f_n \xrightarrow{n\to\infty} f$ und $\varepsilon > 0$. Dann gibt es ein $N \in \mathbb{N}$, so dass $\|f - f_n\|_\infty < \frac{\varepsilon}{2}$ für alle $n \geq N$ gilt. Für alle $n, m \geq N$ gilt also

$$\begin{aligned} \|f_n - f_m\|_\infty = \|(f_n - f) + (f - f_m)\|_\infty &\leq \|f_n - f\|_\infty + \|f - f_m\|_\infty \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$ □

Die Analogie zwischen $\mathcal{C}[0,1]$ mit der Supremumsnorm $\|\cdot\|_\infty$ auf der einen und $\mathbb{R}$ mit dem euklidischen Betrag $|\cdot|$ auf der anderen Seite reicht sogar noch weiter:

Satz 1.9. Der normierte Raum $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ ist *vollständig*: Jede Cauchy-Folge in $\mathcal{C}[0,1]$ konvergiert gegen ein Element von $\mathcal{C}[0,1]$.

Beweis. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{C}[0,1]$, also

$$\forall \varepsilon > 0 \; \exists N(\varepsilon) \in \mathbb{N} \; \forall n, m \geq N(\varepsilon) : \quad \sup_{x\in[0,1]} |f_n(x) - f_m(x)| = \|f_n - f_m\|_\infty < \varepsilon.$$

Für beliebiges $x \in [0,1]$ gilt somit

$$\forall n, m \geq N(\varepsilon) : \quad |f_n(x) - f_m(x)| < \varepsilon,$$

also ist $\big(f_n(x)\big)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{C}$. Laut Korollar 12.12 aus Analysis I ist $\mathbb{C}$ vollständig, also existiert der Grenzwert $f(x) := \lim_{n\to\infty} f_n(x) \in \mathbb{C}$ für alle $x \in [0,1]$. Somit haben wir eine Funktion $f : [0,1] \to \mathbb{C}$ (als punktweisen Limes der f_n) konstruiert.

Es reicht nun zu zeigen, dass $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f konvergiert, denn dann folgt $f \in \mathcal{C}[0,1]$ aus Satz 1.4: Sei $\varepsilon > 0$. Für alle $x \in [0,1]$ und alle $n, m \geq N\big(\frac{\varepsilon}{2}\big) =: N$ gilt

$$|f_n(x) - f_m(x)| < \frac{\varepsilon}{2},$$

also

$$|f_n(x) - f(x)| = \left| f_n(x) - \lim_{m\to\infty} f_m(x) \right| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leq \frac{\varepsilon}{2} < \varepsilon.$$

Somit gilt

$$\forall n \geq N \ \forall x \in [0,1]: \quad |f_n(x) - f(x)| < \varepsilon,$$

also konvergiert $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f. $\square$

Bemerkung 1.10. Man mache sich die geänderte Betrachtungsweise bewusst: Komplizierte konkrete Objekte (die stetigen Funktionen) werden ersetzt durch “Punkte” in einem abstrakten Raum ($\mathcal{C}[0,1]$). Die meisten Eigenschaften werden unterdrückt und nur die wesentlichen behalten: So reduziert sich etwa der genaue Verlauf von f im Vergleich zu g auf den “Abstand” $\|f - g\|_\infty$.

2. Metrische Räume

Definition 2.1. (a) Sei X eine Menge. Eine Funktion $d : X \times X \to [0, \infty)$ heißt *Metrik* auf X, falls die folgenden Eigenschaften erfüllt sind:

(i) $\forall x, y \in X : \quad d(x, y) = 0 \Leftrightarrow x = y$.

(ii) $\forall x, y \in X : \quad d(x, y) = d(y, x)$.

(iii) $\forall x, y, z \in X : \quad d(x, z) \leq d(x, y) + d(y, z)$. (Dreieicksungleichung)

(b) Ein *metrischer Raum* (X, d) ist eine Menge X mit einer Metrik d.[1]

Beispiel 2.2. (a) Für $X \in \{\mathbb{R}, \mathbb{C}, \mathbb{Z}, \mathbb{Q}\}$ kennen wir bereits den *Euklidischen Abstand*

$$d_{|\cdot|}(x, y) := |x - y|.$$

(b) Für beliebiges X ist die *diskrete Metrik* gegeben durch

$$d(x, y) := \begin{cases} 1, & x \neq y, \\ 0, & x = y. \end{cases}$$

(c) Aus dem vorherigen Kapitel haben wir schließlich $X = \mathcal{C}[0, 1]$ mit

$$d_\infty(f, g) = \|f - g\|_\infty.$$

Definition 2.3. Sei (X, d) ein metrischer Raum und $(x_n)_{n \in \mathbb{N}}$ eine Folge in X.

(i) Die Folge $(x_n)_{n \in \mathbb{N}}$ heißt *konvergent* gegen ein $x \in X$, falls

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n \geq N : \quad d(x_n, x) < \varepsilon.$$

Wir schreiben dann $x_n \xrightarrow{n \to \infty} x$.

(ii) Die Folge $(x_n)_{n \in \mathbb{N}}$ heißt *Cauchy-Folge*, falls

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n, m \geq N : \quad d(x_n, x_m) < \varepsilon.$$

Lemma 2.4. Sei (X, d) ein metrischer Raum.

(i) Eine Folge $(x_n)_{n \in \mathbb{N}}$ in X kann höchstens gegen ein $x \in X$ konvergieren.

(ii) Jede konvergente Folge in X ist auch eine Cauchy-Folge.

Beweis. Vergleiche auch den Beweis von Lemma 1.8: Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in X.

(i) Seien $x, y \in X$ mit $x_n \xrightarrow{n \to \infty} x$ und $x_n \xrightarrow{n \to \infty} y$. Zu $\varepsilon > 0$ existiert dann ein hinreichend großes $n \in \mathbb{N}$ mit

$$0 \leq d(x, y) \leq d(x, x_n) + d(x_n, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

also $d(x, y) = 0$. Da d eine Metrik auf X ist, folgt $x = y$.

[1] Wenn die Metrik d aus dem Kontext klar ist, schreibt man häufig nur X für (X, d) – so wie wir meistens $\mathbb{R}$ statt $(\mathbb{R}, +, \cdot)$ für den Körper der reellen Zahlen schreiben.

(ii) Sei $x \in X$ mit $x_n \xrightarrow{n\to\infty} x$ und $\varepsilon > 0$. Dann gibt es ein $N \in \mathbb{N}$, so dass $d(x_n, x) < \frac{\varepsilon}{2}$ für alle $n \geq N$ gilt. Für alle $n, m \geq N$ gilt also

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \square$$

Statt $x_n \xrightarrow{n\to\infty} x$ schreiben wir auch $\lim_{n\to\infty} x_n = x$ und nennen x den *Grenzwert* der Folge $(x_n)_{n\in\mathbb{N}}$.

Definition 2.5. Ein metrischer Raum (X, d) heißt *vollständig*, falls jede Cauchy-Folge $(x_n)_{n\in\mathbb{N}}$ in X gegen ein $x \in X$ konvergiert.

Beispiel 2.6. Nach Analysis I sind $(\mathbb{R}, |\cdot|)$ und $(\mathbb{C}, |\cdot|)$ vollständig, $(\mathbb{Q}, |\cdot|)$ hingegen nicht. Laut Satz 1.9 ist auch $(\mathcal{C}[0,1], d_\infty)$ vollständig.

Satz 2.7 (Banachscher Fixpunktsatz, Satz von der kontrahierenden Abbildung)**.** Sei (X, d) ein vollständiger metrischer Raum und $T : X \to X$ eine Abbildung mit der Eigenschaft

$$\exists c \in [0,1)\ \forall x, y \in X : \quad d\big(T(x), T(y)\big) \leq c \cdot d(x, y).$$ [2]

Dann existiert genau ein $z \in X$ mit $T(z) = z$.[3] Außerdem konvergiert die Folge $\big(T^n(x)\big)_{n\in\mathbb{N}}$ für jedes $x \in X$ gegen z.[4]

Beweis. Wir gliedern den Beweis in mehrere Teilaussagen:

(i) Behauptung: Für alle $x, y \in X$ und alle $n \in \mathbb{N}$ gilt

$$d\big(T^n(x), T^n(y)\big) \leq c^n \cdot d(x, y)$$

Dies folgt per vollständiger Induktion nach n: Für $n = 1$ hat man

$$d\big(T(x), T(y)\big) \leq c \cdot d(x, y)$$

für alle $x, y \in X$ nach Voraussetzung, im Induktionsschritt erhält man

$$\begin{aligned} d\big(T^{n+1}(x), T^{n+1}(y)\big) &= d\Big(T\big(T^n(x)\big), T\big(T^n(y)\big)\Big) \\ &\leq c \cdot d\big(T^n(x), T^n(y)\big) \\ &\leq c \cdot c^n \cdot d(x, y) \\ &= c^{n+1} \cdot d(x, y), \end{aligned}$$

wobei die erste Ungleichung der Fall $n = 1$ ist und die zweite Ungleichung nach Induktionsvoraussetzung gilt.

(ii) Behauptung: Für alle $x \in X$ ist $\big(T^n(x)\big)_{n\in\mathbb{N}}$ eine Cauchy-Folge.
Sei $x \in X$ beliebig. Für alle $n, k \in \mathbb{N}$ gilt

$$d\big(T^n(x), T^{n+k}(x)\big) = d\Big(T^n(x), T^n\big(T^k(x)\big)\Big) \overset{\text{(i)}}{\leq} c^n \cdot d\big(x, T^k(x)\big).$$

[2]So eine Abbildung heißt *kontrahierend* oder *Kontraktion*.
[3]So ein z heißt *Fixpunkt* von T, die Gleichung $T(z) = z$ heißt *Fixpunktgleichung*.
[4]Hierbei ist $T^n := T \circ T \circ \ldots T$ die n-fache Komposition von T mit sich selbst.

Weiterhin gilt für alle $k \in \mathbb{N}$ mit der Dreiecksungleichung und der geometrischen Reihe

$$\begin{aligned} d\big(x, T^k(x)\big) &\leq d\big(x, T(x)\big) + d\big(T(x), T^2(x)\big) + \ldots + d\big(T^{k-1}(x), T^k(x)\big) \\ &= \sum_{j=0}^{k-1} d\big(T^j(x), T^{j+1}(x)\big) \\ &\overset{\text{(i)}}{\leq} \sum_{j=0}^{k-1} c^j \cdot d\big(x, T(x)\big) \\ &= d\big(x, T(x)\big) \cdot \sum_{j=0}^{k-1} c^j \\ &\leq d\big(x, T(x)\big) \cdot \sum_{j=0}^{\infty} c^j \\ &= \frac{d\big(x, T(x)\big)}{1-c}. \end{aligned}$$

Zusammenfassend folgt aus

$$d\big(T^n(x), T^{n+k}(x)\big) \leq c^n \cdot d\big(x, T^k(x)\big) \leq \underbrace{c^n}_{\to 0} \cdot \underbrace{\frac{d\big(x, T(x)\big)}{1-c}}_{\text{konstant}} \xrightarrow{n\to\infty} 0,$$

dass $\big(T^n(x)\big)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist.

Da X vollständig ist, konvergiert $\big(T^n(x)\big)_{n\in\mathbb{N}}$ gegen ein $F(x) \in X$. Dies liefert uns eine Abbildung $F : X \to X,\ x \mapsto F(x)$.

(iii) Behauptung: Für alle $x, y \in X$ gilt $F(x) = F(y)$.

Wegen

$$d\big(T^n(x), T^n(y)\big) \overset{\text{(i)}}{\leq} c^n \cdot d(x, y) \xrightarrow{n\to\infty} 0$$

folgt

$$\begin{aligned} 0 &\leq d\big(F(x), F(y)\big) \\ &\leq d\big(F(x), T^n(x)\big) + d\big(T^n(x), T^n(y)\big) + d\big(T^n(y), F(y)\big) \xrightarrow{n\to\infty} 0, \end{aligned}$$

also $d\big(F(x), F(y)\big) = 0$ und damit $F(x) = F(y)$.

Somit ist $z := F(x)$ unabhängig von $x \in X$ und es gilt $T^n(x) \xrightarrow{n\to\infty} z$ für jedes $x \in X$.

(iv) Behauptung: Dieses z ist ein Fixpunkt von T, also $T(z) = z$.

Es gilt

$$\begin{aligned} d\big(z, T(z)\big) &\leq d\big(z, T^n(z)\big) + d\big(T^n(z), T(z)\big) \\ &= d\big(z, T^n(z)\big) + d\Big(T\big(T^{n-1}(z)\big), T(z)\Big) \\ &\overset{\text{(i)}}{\leq} \underbrace{d\big(z, T^n(z)\big)}_{\to 0} + c \cdot \underbrace{d\big(T^{n-1}(z), z\big)}_{\to 0} \xrightarrow{n\to\infty} 0, \end{aligned}$$

also $d\big(z, T(z)\big) = 0$ und damit $z = T(z)$.

(v) Behauptung: Die Lösung der Fixpunktgleichung ist eindeutig.

Sei $z' \in X$ eine weitere Lösung, also $T(z') = z'$. Dann gilt

$$z' = T^n(z') \xrightarrow{n\to\infty} z.$$

Da laut Lemma 2.4 Grenzwerte eindeutig sind, muss $z' = z$ gelten. $\square$

Beispiel 2.8. (1) Der metrische Raum $([0,3], d)$ mit $d = d_{|\cdot|}$ ist vollständig. Die Abbildung

$$T : [0,3] \to [0,3], \quad x \mapsto \sqrt{1+x}$$

ist streng monoton wachsend, also

$$T\big([0,3]\big) \subseteq \big[T(0), T(3)\big] = [1,2] \subseteq [0,3].$$

Weiterhin ist T kontrahierend: Für alle $x, y \in [0,3]$ gilt $\sqrt{1+x} + \sqrt{1+y} \geq 2$ und daher

$$d\big(T(x), T(y)\big) = \left|\sqrt{1+x} - \sqrt{1+y}\right| = \frac{|x-y|}{\sqrt{1+x}+\sqrt{1+y}} \leq \frac{1}{2} d(x,y).$$

Laut Satz 2.7 hat die Gleichung $z = T(z) = \sqrt{1+z}$ eine eindeutig bestimmte Lösung in $[0,3]$, die man als Grenzwert von $(x_n)_{n\in\mathbb{N}}$ mit $x_1 \in [0,3]$ beliebig und $x_{n+1} := \sqrt{1+x_n}$ für alle $n \in \mathbb{N}$ berechnen kann. Man erhält den *Goldenen Schnitt*

$$z = \frac{1}{2} + \frac{1}{2}\sqrt{5} = 1.61803\ldots.$$

(2) Seien $a \in \mathbb{R}$ und $c \in (0,1)$ gegeben. Wir betrachten $X := \mathcal{C}[0,c]$ und

$$d : X \times X \to [0,\infty), \quad (f,g) \mapsto \sup_{x\in X} |f(x) - g(x)|.$$

Analog zu Satz 1.9 ist auch (X,d) vollständig. Die Abbildung $T : X \to X$, $f \mapsto T(f)$, definiert durch

$$T(f) : [0,c] \to \mathbb{R}, \quad x \mapsto T(f)(x) := a + \int_0^x f(t)\,\mathrm{d}t,$$

ist kontrahierend: Seien $f, g \in X$, für alle $x \in [0,c]$ gilt dann

$$\begin{aligned}
|T(f)(x) - T(g)(x)| &= \left| a + \int_0^x f(t)\,\mathrm{d}t - a - \int_0^x g(t)\,\mathrm{d}t \right| \\
&= \left| \int_0^x \big(f(t) - g(t)\big)\,\mathrm{d}t \right| \\
&\leq \int_0^x |f(t) - g(t)|\,\mathrm{d}t \\
&\leq \int_0^x d(f,g)\,\mathrm{d}t \\
&= x \cdot d(f,g) \\
&\leq c \cdot d(f,g),
\end{aligned}$$

also auch $d\big(T(f),T(g)\big) \leq c \cdot d(f,g)$. Laut Satz 2.7 hat die Gleichung $T(f) = f$ eine eindeutige Lösung f, d.h.

$$f(x) = a + \int_0^x f(t)\,\mathrm{d}t \quad \text{für alle} \quad x \in [0,c].$$

Laut Hauptsatz ist dies äquivalent zu $f' = f$ und $f(0) = a$.

Somit hat die Gleichung $f' = f$ mit $f(0) = a$ auf $[0,c]$ eine eindeutige Lösung. Aus der Analysis I, Satz 15.12 wissen wir schon, dass diese durch $f(x) = a \cdot \exp(x)$ gegeben ist. (Dort hatten wir die Eindeutigkeit der Lösung mit dem Mittelwertsatz gezeigt.)

(3) Für $c \in (0,2)$ ist $X := \{f \in \mathcal{C}[0,c] \mid f \geq 1\}$ mit der Metrik aus dem vorherigen Beispiel ebenfalls vollständig.[5] Setze

$$T(f)(x) := 1 + \int_0^x \sqrt{f(t)}\,\mathrm{d}t,$$

dann ist $T : X \to X$ kontrahierend: Seien $f,g \in X$, für alle $x \in [0,c]$ gilt dann

$$\begin{aligned}
|T(f)(x) - T(g)(x)| &\leq \int_0^x \left|\sqrt{f(t)} - \sqrt{g(t)}\right| \mathrm{d}t \\
&= \int_0^x \frac{|f(t) - g(t)|}{\sqrt{f(t)} + \sqrt{g(t)}}\,\mathrm{d}t \\
&\leq \int_0^x \frac{|f(t) - g(t)|}{2}\,\mathrm{d}t \\
&\leq \frac{1}{2} \cdot d(f,g) \cdot x \\
&\leq \frac{c}{2} \cdot d(f,g),
\end{aligned}$$

also $d\big(T(f),T(g)\big) \leq \frac{c}{2} d(f,g)$ mit $\frac{c}{2} < 1$. Laut Satz 2.7 hat die Gleichung $T(f) = f$ eine eindeutige Lösung f, d.h.

$$f(x) = 1 + \int_0^x \sqrt{f(t)}\,\mathrm{d}t.$$

Dies ist äquivalent zu $f'(x) = \sqrt{f(x)}$ für alle x und $f(0) = 1$. Auf solche *Differenzialgleichungen* kommen wir in späteren Kapiteln noch zurück.

In vielen Fällen werden Metriken von Normen induziert. Dafür betrachten wir im Folgenden $\mathbb{K}$-Vektorräume, wobei $\mathbb{K}$ für $\mathbb{R}$ oder $\mathbb{C}$ steht.

Definition 2.9. (a) Sei V ein $\mathbb{K}$-Vektorraum. Eine *Norm* auf V ist eine Abbildung $\|\cdot\| : V \to [0,\infty)$ mit den folgenden Eigenschaften:

(i) $\forall x \in V: \quad \|x\| = 0 \Leftrightarrow x = 0.$

(ii) $\forall \lambda \in \mathbb{K}\ \forall x \in V: \quad \|\lambda x\| = |\lambda| \cdot \|x\|.$

(iii) $\forall x,y \in V: \quad \|x+y\| \leq \|x\| + \|y\|.$ (Dreiecksungleichung)

[5] Hierbei steht $f \geq 1$ abkürzend für $f(x) \geq 1$ für alle $x \in [0,c]$.

(b) Ein *normierter Raum* $(V, \|\cdot\|)$ ist ein $\mathbb{K}$-Vektorraum V zusammen mit einer Norm $\|\cdot\|$ auf V.

Bemerkung 2.10. Ist $(V, \|\cdot\|)$ ein normierter Raum, so induziert die Norm $\|\cdot\|$ eine Metrik d auf V durch

$$d(x, y) := \|x - y\| \quad \text{für alle} \quad x, y \in V.$$

Diese ist *translationsinvariant*:

$$\forall x, y, z \in V: \quad d(x + z, y + z) = d(x, y).$$

Bezüglich dieser Metrik können wir dann auch in normierten Räumen von konvergenten Folgen, Cauchy-Folgen und Vollständigkeit sprechen. Einen vollständigen normierten Raum nennen wir auch einen *Banachraum*.[6]

Beispiel 2.11. Die Euklidische Abstandsmetrik auf $\mathbb{C}$ und $\mathbb{R}$ wird von der Euklidischen Norm induziert. Die d_∞-Metrik auf $\mathcal{C}[0, 1]$ wird von der Supremumsnorm induziert.

In einigen Fällen werden Normen selbst durch eine stärkere geometrische Struktur auf dem Vektoraum induziert:

Definition 2.12. (a) Ein Skalarprodukt auf einem $\mathbb{K}$-Vektorraum V ist eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ mit den folgenden Eigenschaften:

(i) Für alle $\lambda, \mu \in \mathbb{K}$ und alle $x, x_1, x_2, y, y_1, y_2 \in V$ gilt [7]

$$\langle \lambda x_1 + \mu x_2, y \rangle = \lambda \langle x_1, y \rangle + \mu \langle x_2, y \rangle$$

und

$$\langle x, \lambda y_1 + \mu y_2 \rangle = \overline{\lambda} \langle x, y_1 \rangle + \overline{\mu} \langle x, y_2 \rangle.$$

(ii) Für alle $x, y \in V$ gilt $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

(iii) Für alle $x \in V$ gilt $\langle x, x \rangle \geq 0$ und

$$\langle x, x \rangle = 0 \quad \Leftrightarrow \quad x = 0.$$

(b) Ein *Prähilbertraum* $(V, \langle \cdot, \cdot \rangle)$ ist ein $\mathbb{K}$-Vektorraum V mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$.

Beispiel 2.13. (a) Auf $\mathbb{C}^n$ und $\mathbb{R}^n$ hat man das *Standardskalarprodukt*

$$\langle x, y \rangle := \sum_{j=1}^{n} x_j \overline{y_j}.$$

(b) Auf $\mathcal{C}[0, 1]$ ist ein Skalarprodukt gegeben durch

$$\langle f, g \rangle := \int_0^1 f(t) \overline{g(t)} \, \mathrm{d}t.$$

[6] Nicht zu verwechseln mit einem *Bananachraum*, der Antwort auf folgende klassische mathematische Scherzfrage: "Was ist gelb, krumm, normiert und vollständig?"

[7] Achtung Physiker: In der Physik wird das Skalarprodukt, gemäß der bra-ket-Denkweise, normalerweise linear in der zweiten Komponente und anti-linear in der ersten angenommen. Das ändert natürlich nichts an der allgemeinen Theorie, man muss nur bei den Formulierungen und Rechnungen aufpassen.

Satz 2.14 (Cauchy-Schwarz)**.** Sei $(V, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum über $\mathbb{K}$.

(a) Für alle $x, y \in V$ gilt die *Cauchy-Schwarz-Ungleichung*

$$|\langle x,y\rangle|^2 \leq \langle x,x\rangle \cdot \langle y,y\rangle$$

mit Gleichheit genau dann, wenn x und y linear abhängig sind.

(b) Durch $\|x\| := \sqrt{\langle x,x\rangle}$ wird eine Norm auf V induziert.

Beweis. (a) Seien $x, y \in V$. Für alle $\lambda \in \mathbb{K}$ gilt

$$0 \leq \langle x+\lambda y, x+\lambda y\rangle = \langle x,x\rangle + \lambda\langle y,x\rangle + \overline{\lambda}\langle x,y\rangle + |\lambda|^2\langle y,y\rangle.$$

- Falls $\langle y,y\rangle = 0$, so folgt $y = 0$ und die Behauptung ist erfüllt (denn dann sind x und y linear abhängig und beide Seiten der Ungleichung sind gleich 0).
- Falls $\langle y,y\rangle \neq 0$, so betrachten wir

$$\lambda := -\frac{\langle x,y\rangle}{\langle y,y\rangle}.$$

Damit gilt

$$\begin{aligned}
0 &\leq \langle x+\lambda y, x+\lambda y\rangle \\
&= \langle x,x\rangle + \lambda\langle y,x\rangle + \overline{\lambda}\langle x,y\rangle + |\lambda|^2\langle y,y\rangle \\
&= \langle x,x\rangle - \frac{\langle x,y\rangle}{\langle y,y\rangle}\langle y,x\rangle - \frac{\overline{\langle x,y\rangle}}{\overline{\langle y,y\rangle}}\langle x,y\rangle + \frac{|\langle x,y\rangle|^2}{|\langle y,y\rangle|^2}\langle y,y\rangle \\
&= \langle x,x\rangle - \frac{\langle x,y\rangle}{\langle y,y\rangle}\overline{\langle x,y\rangle} - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} + \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} \\
&= \langle x,x\rangle - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle},
\end{aligned}$$

also

$$|\langle x,y\rangle|^2 \leq \langle x,x\rangle \cdot \langle y,y\rangle.$$

Gleichheit gilt nun genau dann, wenn

$$0 = \langle x,x\rangle - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} = \langle x+\lambda y, x+\lambda y\rangle,$$

also wenn $x + \lambda y = 0$. Dies ist aber (im Falle $y \neq 0$) dazu äquivalent, dass x und y linear abhängig sind.

(b) Wir weisen die Normeigenschaften aus Definition 2.9 nach:

(i) Sei $x \in V$, dann gilt

$$\|x\| = 0 \quad\Leftrightarrow\quad \sqrt{\langle x,x\rangle} = 0 \quad\Leftrightarrow\quad \langle x,x\rangle = 0 \quad\Leftrightarrow\quad x = 0.$$

(ii) Seien $\lambda \in \mathbb{K}$ und $x \in V$, dann gilt

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\lambda\overline{\lambda}\langle x,x\rangle} = \sqrt{|\lambda|^2\langle x,x\rangle} = |\lambda|\sqrt{\langle x,x\rangle} = |\lambda|\cdot\|x\|.$$

(iii) Seien $x, y \in V$, dann gilt

$$\begin{aligned}\|x+y\|^2 = \langle x+y, x+y\rangle &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \|x\|^2 + \langle x,y\rangle + \overline{\langle x,y\rangle} + \|y\|^2 \\ &= \|x\|^2 + 2\operatorname{Re}(\langle x,y\rangle) + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ &\overset{\text{(a)}}{\leq} \|x\|^2 + 2\sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle} + \|y\|^2 \\ &= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2,\end{aligned}$$

also auch $\|x+y\| \leq \|x\| + \|y\|$. □

Beispiel 2.15. Auf $V = \mathbb{K}^n$ sind für $x = (x_1, \ldots, x_n) \in V$ durch folgende Ausdrücke Normen gegeben:

- $\|x\|_1 := \sum_{j=1}^{n} |x_j|$ (1-Norm, Betragssummennorm)
- $\|x\|_2 := \sqrt{\sum_{j=1}^{n} |x_j|^2}$ (2-Norm, Euklidische Norm)
- $\|x\|_\infty := \max_{j=1,\ldots,n} |x_j|$ (∞-Norm, Maximumsnorm)

Die induzierten Metriken werden üblicherweise mit d_1, d_2 und d_∞ bezeichnet. Für $\|\cdot\|_1$ und $\|\cdot\|_\infty$ kann man die Normeigenschaften leicht nachrechnen. Weiterhin zeigt man leicht, dass das Standardskalarprodukt

$$\langle x, y\rangle_2 := \sum_{j=1}^{n} x_j \overline{y_j}$$

die Euklidische Norm induziert (vgl. Beispiel 2.13).

Definition 2.16. Ist (X, d) ein metrischer Raum, so setzen wir

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

für alle $x \in X$ und $r \in (0, \infty)$. Wir nennen $\mathrm{B}(x, r)$ auch die *Kugel* in (X, d) mit Mittelpunkt x und Radius r.

Beispiel 2.17. Die Kugeln im $\mathbb{R}^2$ um 0 mit Radius 1 sind in Abbildung 1 dargestellt. Für alle $x \in \mathbb{R}^2$ gilt

$$2d_\infty(x, 0) \geq d_1(x, 0) \geq d_2(x, 0) \geq d_\infty(x, 0),$$

dies folgt aus

$$2\max\{|x_1|, |x_2|\} \geq |x_1| + |x_2| \geq \sqrt{|x_1|^2 + |x_2|^2} \geq \max\{|x_1|, |x_2|\}.$$

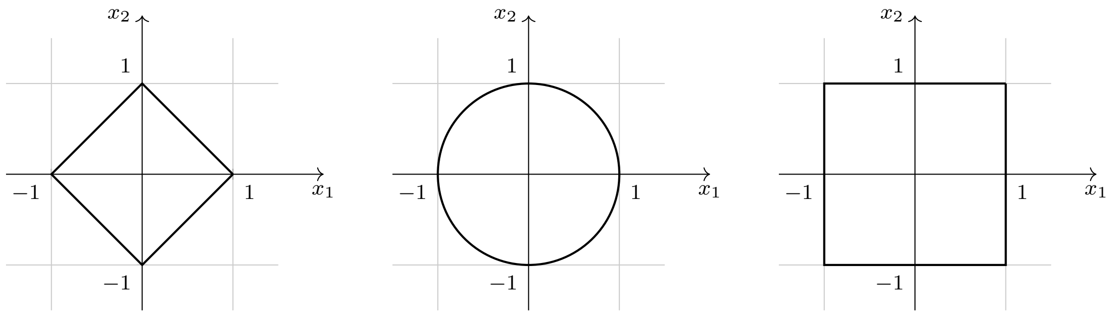

Abbildung 1: Graphische Darstellung der Einheitskugeln im $\mathbb{R}^2$ bezüglich unterschiedlicher Metriken: d_1, d_2 und d_∞ (von links nach rechts).

Bemerkung 2.18. (i) Die übliche Metrik auf $\mathbb{K}^n$ ist die Euklidische Metrik d_2, induziert von der Euklidischen Norm $\|\cdot\|_2$.

(ii) Eine Folge $(x^{(k)})_{k\in\mathbb{N}}$ in $\mathbb{K}^n$ konvergiert genau dann bezüglich d_2, wenn sie komponentenweise konvergiert, also wenn für jedes $j = 1, \dots, n$ die Folgen $(x_j^{(k)})_{k\in\mathbb{N}}$ in $\mathbb{K}$ bezüglich dem Euklidischen Abstand $|\cdot|$ konvergieren.

(iii) $(\mathbb{K}^n, d_2)$ ist ein vollständiger metrischer Raum.

Bemerkung 2.19. (a) Auf $\mathcal{C}[0,1]$ ist durch

$$\langle f, g\rangle := \int_0^1 f(t)\overline{g(t)}\,\mathrm{d}t$$

ein Skalarprodukt gegeben (vgl. Beispiel 2.13). Laut Satz 2.14 gilt somit für alle $f, g \in \mathcal{C}[0,1]$

$$\left|\int_0^1 f(t)\overline{g(t)}\,\mathrm{d}t\right| \le \left(\int_0^1 |f(t)|^2\,\mathrm{d}t\right)^{\frac{1}{2}} \left(\int_0^1 |g(t)|^2\,\mathrm{d}t\right)^{\frac{1}{2}}$$

und durch

$$\|f\|_2 := \left(\int_0^1 |f(t)|^2\,\mathrm{d}t\right)^{\frac{1}{2}}$$

wird eine Norm $\|\cdot\|_2$ auf $\mathcal{C}[0,1]$ definiert. Diese stimmt offenbar nicht mit der sup-Norm aus Abschnitt 1 überein.

(b) Für alle $f \in \mathcal{C}[0,1]$ gilt

$$\|f\|_2^2 = \int_0^1 |f(t)|^2\,\mathrm{d}t \le \|f\|_\infty^2 \int_0^1 \mathrm{d}t = \|f\|_\infty^2,$$

also

$$\|f\|_2 \le \|f\|_\infty.$$

Aber es gibt kein $c \in (0, \infty)$ mit

$$\|f\|_\infty \le c \cdot \|f\|_2 \quad \text{für alle} \quad f \in \mathcal{C}[0,1].$$ [8]

[8] Man sagt daher auch, dass diese beiden Normen nicht *äquivalent* sind, mehr dazu in der Übung.

Sei dazu

$$f_n : [0,1] \to \mathbb{C}, \quad t \mapsto \begin{cases} \sqrt{1 - n \cdot t}, & \text{falls } 0 \le t \le \frac{1}{n}, \\ 0, & \text{falls } \frac{1}{n} \le t \le 1. \end{cases}$$

Für alle $n \in \mathbb{N}$ gilt dann einerseits $\|f_n\|_\infty = 1$ und andererseits

$$\|f_n\|_2^2 = \int_0^{\frac{1}{n}} (1 - n \cdot t)\,\mathrm{d}t = \left[t - \frac{1}{2}nt^2\right]_0^{\frac{1}{n}} = \frac{1}{2n}.$$

Also konvergiert $(f_n)_{n\in\mathbb{N}}$ bezüglich $\|\cdot\|_2$ gegen 0, aber nicht bezüglich $\|\cdot\|_\infty$. Somit ist $\|\cdot\|_2$ *schwächer* als $\|\cdot\|_\infty$.

(c) Laut Satz 1.9 ist $\mathcal{C}[0,1]$ bezüglich $\|\cdot\|_\infty$ vollständig. Bezüglich $\|\cdot\|_2$ ist $\mathcal{C}[0,1]$ hingegen nicht vollständig: Für

$$f_n : [0,1] \to \mathbb{R}, \quad t \mapsto \begin{cases} 0, & \text{falls } 0 \le t \le \frac{1}{2} - \frac{1}{n}, \\ n \cdot t + 1 - \frac{n}{2}, & \text{falls } \frac{1}{2} - \frac{1}{n} \le t \le \frac{1}{2}, \\ 1, & \text{falls } \frac{1}{2} \le t \le 1, \end{cases}$$

gilt im Falle $n \le m$

$$\|f_n - f_m\|_2^2 = \int_0^1 |f_n(t) - f_m(t)|^2\,\mathrm{d}t = \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} |f_n(t) - f_m(t)|^2\,\mathrm{d}t \le \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} 1\,\mathrm{d}t = \frac{1}{n},$$

also ist $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $(\mathcal{C}[0,1], \|\cdot\|_2)$. Allerdings konvergiert $(f_n)_{n\in\mathbb{N}}$ gegen die unstetige Funktion

$$f : [0,1] \to \mathbb{R}, \quad t \mapsto \begin{cases} 0, & \text{falls } 0 \le t < \frac{1}{2}, \\ 1, & \text{falls } \frac{1}{2} \le t \le 1, \end{cases}$$

also $f \notin \mathcal{C}[0,1]$, vergleiche auch Abbildung 2.

Man muss $\mathcal{C}[0,1]$ also bezüglich $\|\cdot\|_2$ "vervollständigen", ähnlich wie $\mathbb{Q}$ zu $\mathbb{R}$ vervollständigt werden musste.

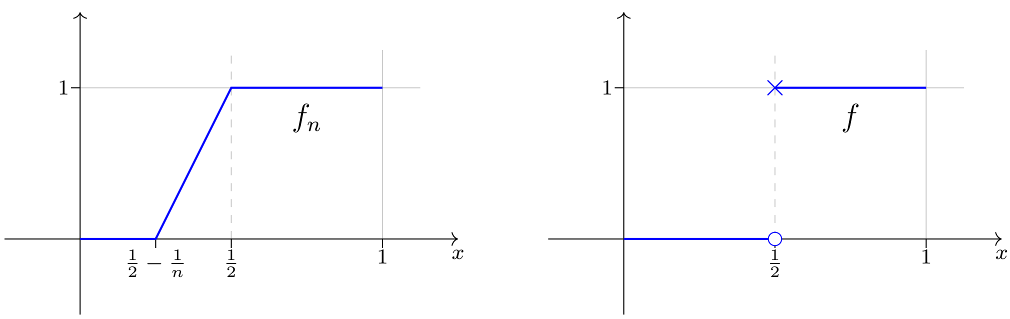

Abbildung 2: Die Funktionen f_n und f, die zeigen, dass $(\mathcal{C}[0,1], \|\cdot\|_2)$ nicht vollständig ist.

Definition 2.20. Ein Prähilbertraum, der bezüglich seiner induzierten Norm vollständig ist, heißt *Hilbertraum*.

Beispiel 2.21. (i) $(\mathbb{K}^n, \langle\cdot,\cdot\rangle_2)$ ist ein Hilbertraum.

(ii) $(\mathcal{C}[0,1], \langle\cdot,\cdot\rangle_2)$ ist *kein* Hilbertraum.

3. Topologie metrischer Räume

Definition 3.1. Sei (X, d) ein metrischer Raum.

(1) Eine Teilmenge $U \subseteq X$ heißt *offen*, falls

$$\forall x \in U \ \exists \varepsilon > 0 : \quad \mathrm{B}(x, \varepsilon) \subseteq U.$$

(2) Eine Teilmenge $A \subset X$ heißt *abgeschlossen*, falls $X \setminus A$ offen ist.

Bemerkung 3.2. In jedem metrischen Raum (X, d) sind sowohl $\emptyset$ als auch X offen und damit (wegen $X \setminus X = \emptyset$ und $X \setminus \emptyset = X$) auch abgeschlossen. "Offen" und "abgeschlossen" schließen sich nicht gegenseitig aus, sie sind auch keine Negationen voneinander.[9]

Beispiel 3.3. (1) Sei (X, d) ein diskreter metrischer Raum, also

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

Dann ist jede Teilmenge von X offen und abgeschlossen.

(2) In jedem metrischen Raum (X, d) ist für alle $x \in X$ und alle $\varepsilon > 0$ die Kugel $\mathrm{B}(x, \varepsilon)$ offen.

Satz 3.4. Sei (X, d) ein metrischer Raum. Dann gilt:

(a) $\emptyset$ und X sind offen.

(b) Jeder endliche Durchschnitt offener Mengen ist wieder offen: Sind $U_1, \dots, U_n$ offene Teilmengen von X, so ist auch $\bigcap_{i=1}^n U_i = U_1 \cap \ldots \cap U_n$ offen.[10]

(c) Jede beliebige Vereinigung offener Mengen ist wieder offen: Für jede Familie $(U_i)_{i \in I}$ von offenen Teilmengen von X ist auch $\bigcup_{i \in I} U_i$ offen.[11]

Beweis. (a) Klar (für $\emptyset$ ist nichts zu zeigen und für X ist die Bedingung $\mathrm{B}(x, \varepsilon) \subseteq X$ trivialerweise immer erfüllt.).

(b) Sei $x \in \bigcap_{i=1}^n U_i$, dann liegt x in jeder offenen Menge U_i. Somit existiert für jedes $i = 1, \dots, n$ ein $\varepsilon_i > 0$ mit $\mathrm{B}(x, \varepsilon_i) \subseteq U_i$. Setze $\varepsilon := \min\{\varepsilon_1, \dots, \varepsilon_n\} > 0$. Für alle $i = 1, \dots, n$ gilt also $\varepsilon \leq \varepsilon_i$ und somit

$$\mathrm{B}(x, \varepsilon) \subseteq \mathrm{B}(x, \varepsilon_i) \subseteq U_i,$$

also auch $\mathrm{B}(x, \varepsilon) \subseteq \bigcap_{i=1}^n U_i$. Somit ist $\bigcap_{i=1}^n U_i$ offen.

(c) Sei $x \in \bigcup_{i \in I} U_i$, dann existiert ein $j \in I$ mit $x \in U_j$. Da U_j offen ist, existiert ein $\varepsilon > 0$ mit $\mathrm{B}(x, \varepsilon) \subseteq U_j \subseteq \bigcup_{i \in I} U_i$. Somit ist $\bigcup_{i \in I} U_i$ offen. □

[9] Mengen, die sowohl offen als auch abgeschlossen sind, heißen manchmal auch *abgeschloffen*.

[10] Die Einschränkung auf endliche Durchschnitte ist hier wichtig: So ist etwa $(1 - \frac{1}{n}, 1 + \frac{1}{n})$ für jedes $n \in \mathbb{N}$ eine offene Menge in $(\mathbb{R}, |\cdot|)$, aber $\bigcap_{n \in \mathbb{N}} (1 - \frac{1}{n}, 1 + \frac{1}{n}) = \{1\}$ ist nicht offen.

[11] So wie eine Folge in X eine Abbildung $\mathbb{N} \to X$ ist, ist eine *Familie* eine Abbildung $I \to X$, wobei I eine beliebige Indexmenge sein kann.

Bemerkung 3.5. (1) Da Kugeln offen sind, ist jede Vereinigung von Kugeln offen. Umgekehrt ist jede offene Menge auch eine Vereinigung von Kugeln.

(2) Eine Menge $\mathcal{T}$ von Teilmengen einer Menge X, die die Bedingungen aus Satz 3.4 erfüllt, heißt *Topologie* auf X, $(X, \mathcal{T})$ ist dann ein *topologischer Raum*. Es gibt topologische Räume, die nicht von Metriken induziert sind.

(3) Analog zu Satz 3.4 gilt für die abgeschlossenen Mengen in X:

  (a) $\emptyset$ und X sind abgeschlossen.

  (b) Jede endliche Vereinigung von abgeschlossenen Mengen ist abgeschlossen.

  (c) Jeder beliebige Durchschnitt abgeschlossener Mengen ist abgeschlossen.

  Dies folgt direkt aus $X \setminus \bigcap A_i = \bigcup (X \setminus A_i)$ und $X \setminus \bigcup A_i = \bigcap (X \setminus A_i)$.

Satz 3.6. Sei A eine Teilmenge eines metrischen Raumes (X, d). Dann sind äquivalent:

(1) A ist abgeschlossen.

(2) Für jede Folge $(x_n)_{n \in \mathbb{N}}$ in A, die einen Grenzwert $x \in X$ besitzt, gilt $x \in A$.

Beweis. (1)⇒(2): Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in A mit $x_n \xrightarrow{n \to \infty} x \in X$. Wäre nun $x \notin A$, so liegt x stattdessen in der offenen Menge $X \setminus A$. Also gibt es ein $\varepsilon > 0$ mit $\mathrm{B}(x, \varepsilon) \subseteq X \setminus A$. Insbesondere gilt nun $\mathrm{B}(x, \varepsilon) \cap A = \emptyset$, also kann $(x_n)_{n \in \mathbb{N}}$ nicht gegen x konvergieren.

(2)⇒(1): Wir führen den Beweis per Kontraposition, die Aussage ist also: Wenn $A \subseteq X$ nicht abgeschlossen ist, gibt es eine Folge $(x_n)_{n \in \mathbb{N}}$ in A, die gegen ein $x \in X \setminus A$ konvergiert.

Sei also $A \subseteq X$, dann ist $X \setminus A$ nicht offen. Somit gibt es ein $x \in X \setminus A$, so dass für jedes $\varepsilon > 0$ die Kugel $\mathrm{B}(x, \varepsilon)$ Elemente von A enthält. Sei $x_n \in \mathrm{B}(x, \frac{1}{n}) \cap A$ ein solches Element für $\varepsilon = \frac{1}{n}$. Dann gilt $x_n \xrightarrow{n \to \infty} x$, da $d(x, x_n) = \frac{1}{n} \xrightarrow{n \to \infty} 0$ und $x_n \in A$, aber $x \notin A$. □

Bemerkung 3.7. (a) $\mathbb{Q}$ ist als Teilmenge von $\mathbb{R}$ nicht abgeschlossen, als Teilmenge von sich selbst aber natürlich schon. Abgeschlossenheit (und Offenheit) ist also immer relativ zu einem gegebenen metrischen Raum.

(b) Seien $a, b \in \mathbb{R}$ mit $a < b$ und betrachte $(\mathbb{R}, |\cdot|)$.

- $[a, b]$ ist abgeschlossen und nicht offen.
- (a, b) ist offen und nicht abgeschlossen.
- $(a, b]$ und $[a, b)$ sind weder offen noch abgeschlossen.

Bemerkung 3.8. Zwei Metriken d_1 und d_2 auf X heißen *äquivalent*, falls es Konstanten $c_1, c_2 > 0$ gibt mit

$$\forall x, y \in X: \qquad d_1(x, y) \le c_1 d_2(x, y) \quad \text{und} \quad d_2(x, y) \le c_2 d_1(x, y).$$

Wenn d_1 und d_2 äquivalent sind, so gilt laut dem Quetschlemma für jede Folge $(x_n)_{n \in \mathbb{N}}$ in X und jedes $x \in X$

$$d_1(x, x_n) \xrightarrow{n \to \infty} 0 \quad \Leftrightarrow \quad d_2(x, x_n) \xrightarrow{n \to \infty} 0,$$

also konvergiert die Folge $(x_n)_{n\in\mathbb{N}}$ genau dann bezüglich d_1, wenn sie für d_2 konvergiert. Insbesondere haben (X, d_1) und (X, d_2) dieselben offenen und dieselben abgeschlossenen Mengen.

Definition 3.9. Sei Y eine Teilmenge eines metrischen Raumes X. Der *Abschluss* $\overline{Y}$ von Y in X ist der Durchschnitt aller abgeschlossenen Mengen von X, die Y enthalten:

$$\overline{Y} := \bigcap_{\substack{A\subseteq X \text{ abg.} \\ Y\subseteq A}} A.$$

Somit ist $\overline{Y}$ die kleinste abgeschlossene Teilmenge von X, die Y enthält.

Beispiel 3.10. Für $X = \mathbb{R}$ gilt $\overline{\mathbb{Q}} = \mathbb{R}$. Seien $a, b \in \mathbb{R}$ mit $a < b$, dann gilt

$$\overline{(a, b)} = \overline{(a, b]} = \overline{[a, b)} = [a, b].$$

Lemma 3.11. Sei Y eine Teilmenge eines metrischen Raumes (X, d). Dann ist $\overline{Y}$ die Menge aller Grenzwerte von konvergenten Folgen in Y.

Beweis. Sei H die Menge aller Grenzwerte von konvergenten Folgen in Y. Dann gilt $Y \subseteq H$ (denn jedes $y \in Y$ ist Grenzwert der konstanten Folge, die nur den Wert y annimmt) und $H \subseteq \overline{Y}$ (da $\overline{Y}$ abgeschlossen ist und jede Folge in Y auch eine Folge in $\overline{Y}$ ist). Es reicht also zu zeigen, dass H abgeschlossen ist.

Sei $(h_n)_{n\in\mathbb{N}}$ eine Folge in H mit $h_n \xrightarrow{n\to\infty} z \in X$. Da h_n nach Definition von H der Grenzwert einer Folge in Y ist, gibt es ein $y_n \in Y$ mit $d(h_n, y_n) < \frac{1}{n}$. Somit gilt

$$d(z, y_n) \leq d(z, h_n) + d(h_n, y_n) < \underbrace{d(z, h_n)}_{\to 0} + \underbrace{\frac{1}{n}}_{\to 0} \xrightarrow{n\to\infty} 0.$$

Somit ist z auch der Grenzwert der Folge $(y_n)_{n\in\mathbb{N}}$ mit $y_n \in Y$ für alle $n \in \mathbb{N}$. Per Definition ist also $z \in H$. $\square$

Satz 3.12. Seien (X, d) und (Y, d') metrische Räume und $f : X \to Y$ eine Abbildung. Dann sind folgende Eigenschaften äquivalent:

(1) Für jede Folge $(x_n)_{n\in\mathbb{N}}$ in X und jedes $x_0 \in X$ gilt

$$x_n \xrightarrow{n\to\infty} x_0 \quad \Rightarrow \quad f(x_n) \xrightarrow{n\to\infty} f(x_0).$$

(2) Für alle $x_0 \in X$ gilt

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in X : \quad d(x, x_0) < \delta \Rightarrow d'\big(f(x), f(x_0)\big) < \varepsilon.$$ [12]

(3) Für jede offene Menge $U \subseteq Y$ ist $f^{-1}(U) \subseteq X$ offen. [13]

(4) Für jede abgeschlossene Menge $A \subseteq Y$ ist $f^{-1}(A) \subseteq X$ abgeschlossen.

[12] In Kugeln ausgedrückt: $\forall x_0 \in X\ \forall \varepsilon > 0\ \exists \delta > 0 : \quad f\big(\mathrm{B}(x_0, \delta)\big) \subseteq \mathrm{B}\big(f(x_0), \varepsilon\big)$.

[13] Hier ist $f^{-1}(U) := \{x \in X \mid f(x) \in U\}$ das Urbild von U unter f. Nicht mit dem Konzept der Umkehrabbildung verwechseln, die hier überhaupt nicht existieren muss!

Beweis. (1)⇒(4): Sei $A \subseteq Y$ abgeschlossen und $(x_n)_{n\in\mathbb{N}}$ eine Folge in $f^{-1}(A)$ mit $x_n \xrightarrow{n\to\infty} x_0 \in X$. Dann ist $\big(f(x_n)\big)_{n\in\mathbb{N}}$ eine Folge in der abgeschlossenen Menge A, für die laut (1) $f(x_n) \xrightarrow{n\to\infty} f(x_0)$ gilt. Also muss auch $f(x_0) \in A$ gelten Somit folgt $x_0 \in f^{-1}(A)$ und $f^{-1}(A)$ abgeschlossen.

(4)⇒(3): Sei $U \subseteq Y$ offen, dann ist $Y \setminus U$ abgeschlossen. Laut (4) ist dann

$$X \setminus f^{-1}(U) = f^{-1}(Y \setminus U)$$

als Urbild einer abgeschlossenen Menge selbst wieder eine abgeschlossene Menge, also ist $f^{-1}(U) \subseteq X$ offen.

(3)⇒(2): Sei $x \in X$ und $\varepsilon > 0$. Dann ist $\mathrm{B}\big(f(x), \varepsilon\big) \subseteq Y$ eine offene Menge, also ist laut (3) auch

$$U := f^{-1}\Big(\mathrm{B}\big(f(x), \varepsilon\big)\Big) \subseteq X$$

eine offene Menge. Wegen $x \in U$ gibt es also ein $\delta > 0$, so dass $\mathrm{B}(x, \delta) \subseteq U$. Folglich gilt

$$f\big(\mathrm{B}(x, \delta)\big) \subseteq f(U) \subseteq \mathrm{B}\big(f(x), \varepsilon\big).$$

(2)⇒(1): Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $x_n \xrightarrow{n\to\infty} x_0 \in X$ und $\varepsilon > 0$. Dann existiert laut (2) ein $\delta > 0$, so dass

$$\forall x \in X : \quad d(x, x_0) < \delta \Rightarrow d'\big(f(x), f(x_0)\big) < \varepsilon.$$

Da $(x_n)_{n\in\mathbb{N}}$ gegen x_0 konvergiert, gibt es ein $N \in \mathbb{N}$ mit $d(x_n, x_0) < \delta$ für alle $n \geq N$. Zusammengenommen gilt also $d'\big(f(x_n), f(x_0)\big) < \varepsilon$ für alle $n \geq N$. Folglich konvergiert $\big(f(x_n)\big)_{n\in\mathbb{N}}$ in Y gegen $f(x_0)$. □

Definition 3.13. Eine Abbildung $f : (X, d) \to (Y, d')$ heißt *stetig*, falls eine (und damit alle) der äquivalenten Bedingungen aus Satz 3.12 erfüllt sind.

Definition 3.14. Sei $f : (X, d) \to (Y, d')$ eine Abbildung zwischen metrischen Räumen.

(a) Die Abbildung f heißt *stetig* im Punkt $x_0 \in X$, falls für alle Folgen $(x_n)_{n\in\mathbb{N}}$ in X

$$x_n \xrightarrow{n\to\infty} x_0 \quad \Rightarrow \quad f(x_n) \xrightarrow{n\to\infty} f(x_0)$$

gilt. Dies ist äquivalent zu

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in X : \quad d(x, x_0) < \delta \Rightarrow d'\big(f(x), f(x_0)\big) < \varepsilon.$$

Stetigkeit von f ist äquivalent zu Stetigkeit in allen Punkten, d.h. zu

$$\forall x_0 \in X \; \forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in X : \quad d(x, x_0) < \delta \Rightarrow d'\big(f(x), f(x_0)\big) < \varepsilon.$$

(b) Die Abbildung f heißt *gleichmäßig stetig*, falls

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x_0, x \in X : \quad d(x, x_0) < \delta \Rightarrow d'\big(f(x), f(x_0)\big) < \varepsilon.$$

4. Kompaktheit in metrischen Räumen

Definition 4.1. Sei A Teilmenge eines metrischen Raumes X.

(1) Eine *offene Überdeckung* von A ist eine Familie $(U_i)_{i\in I}$ von offenen Teilmengen von X mit $A \subseteq \bigcup_{i\in I} U_i$.

(2) Die Menge A heißt *kompakt*, falls jede offene Überdeckung von A eine endliche Teilüberdeckung besitzt: Für jede offene Überdeckung $(U_i)_{i\in I}$ existieren endlich viele $i_1,\ldots,i_n \in I$ mit $A \subseteq \bigcup_{j=1}^n U_{i_j}$.

(3) Die Menge A heißt *folgenkompakt*, falls jede Folge in A eine Teilfolge besitzt, die gegen einen Punkt in A konvergiert.

Im Folgenden wollen wir zeigen, dass in metrischen Räumen die Konzepte "kompakt" und "folgenkompakt" äquivalent sind.

Beispiel 4.2. (1) Die Menge $A = (0,1]$ in $\mathbb{R}$ ist

- nicht folgenkompakt, da $\frac{1}{n} \in A$ für alle $n \in \mathbb{N}$, aber $\frac{1}{n} \xrightarrow{n\to\infty} 0 \notin A$ (also kann auch keine Teilfolge gegen einen Punkt in A konvergieren), und
- nicht kompakt: Für alle $n \in \mathbb{N}$ ist $U_n := (\frac{1}{n}, 2)$ eine offene Menge in $\mathbb{R}$. Wegen

$$\bigcup_{n\in\mathbb{N}} U_n = (0,2) \supseteq (0,1] = A$$

ist $(U_n)_{n\in\mathbb{N}}$ eine offene Überdeckung von A, aber diese hat keine endliche Teilüberdeckung: Für $n_1,\ldots,n_k \in \mathbb{N}$ setze $m := \max\{n_1,\ldots,n_k\}$, dann gilt $\frac{1}{2m} \in A$, aber

$$\frac{1}{2m} \notin U_m = \bigcup_{j=1}^{k} U_{n_k}.$$

(2) Die Menge $A = [0,1]$ in $\mathbb{R}$ ist

- folgenkompakt, da jede Folge in A laut Bolzano-Weierstraß eine in A konvergente Teilfolge besitzt, und
- kompakt. Als Beispiel betrachten wir die offene Überdeckung $(U_n)_{n\in\mathbb{N}_0}$ mit $U_n := (\frac{1}{n}, 2)$ für alle $n \in \mathbb{N}$ und $U_0 := \mathrm{B}(0,\varepsilon)$ für ein beliebiges $\varepsilon > 0$. Nach Archimedes gibt es ein $m \in \mathbb{N}$ mit $\frac{1}{m} < \varepsilon$, also ist $\{U_0,\ldots,U_m\}$ eine endliche Teilüberdeckung.

(3) Die Menge $A = \mathbb{R}$ als Teilmenge von sich selbst ist

- nicht folgenkompakt, denn $(n)_{n\in\mathbb{N}}$ hat keine konvergente Teilfolge, und
- nicht kompakt, denn $(U_n)_{n\in\mathbb{N}}$ mit $U_n := (-n,n)$ ist eine offene Überdeckung von $\mathbb{R}$, die keine endliche Teilüberdeckung besitzt (da eine endliche Vereinigung beschränkter Intervalle immer noch beschränkt ist).

(4) Sei $X = \mathcal{C}[0,1]$ mit der von der sup-Norm induzierten Metrik und

$$A := \overline{\mathrm{B}(0,1)} = \{f \in \mathcal{C}[0,1] : \|f\|_\infty \leq 1\}.$$

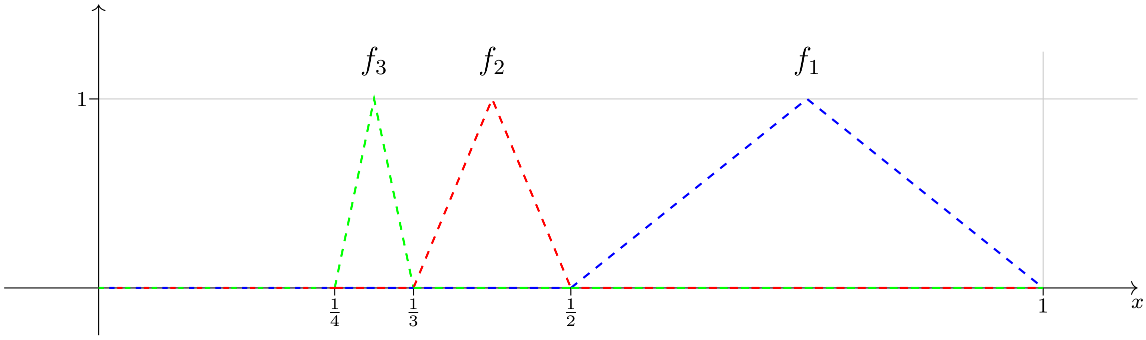

Abbildung 3: Funktionen f_n, die zeigen, dass $\overline{\mathrm{B}(0,1)}$ in $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ nicht kompakt ist.

Betrachte die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ mit $f_n : [0,1] \to \mathbb{R}$ wie in Abbildung 3. Dann gilt $\|f_n\|_\infty = 1$ für alle $n \in \mathbb{N}$, also $f_n \in A$. Weiterhin gilt $\|f_n - f_m\|_\infty = 1$ für alle $n \neq m$. Somit hat $(f_n)_{n\in\mathbb{N}}$ keine konvergente Teilfolge, also ist A nicht folgenkompakt. Weiterhin ist A auch nicht kompakt: Betrachte dazu die offene Überdeckung $(U_f)_{f\in\mathcal{C}[0,1]}$ mit

$$U_f := \mathrm{B}\left(f, \frac{1}{4}\right) = \left\{g \in \mathcal{C}[0,1] : \|g - f\|_\infty < \frac{1}{4}\right\}.$$

Angenommen, es gäbe eine endliche Teilüberdeckung

$$\{U_{h_1}, \ldots, U_{h_n}\}.$$

Wegen $\|f_n - f_m\|_\infty = 1$ für $n \neq m$ kann jedes U_{h_i} höchstens ein Element der Folge $(f_n)_{n\in\mathbb{N}}$ enthalten. Somit kann $\{U_{h_1}, \ldots, U_{h_n}\}$ nicht alle f_n überdecken und insbesondere auch nicht A. Folglich kann es keine endliche Teilüberdeckung von A geben.

Der Grund dafür ist hier also, dass man A nicht mit endlich vielen Kugeln vom Radius $\frac{1}{4}$ überdecken kann.

Definition 4.3. Eine Teilmenge A eines metrischen Raumes X heißt *total beschränkt* (oder *präkompakt*), falls es für jedes $\varepsilon > 0$ endlich viele $x_1, \ldots, x_n \in X$ gibt mit

$$A \subseteq \bigcup_{i=1}^{n} \mathrm{B}(x_i, \varepsilon).$$

Beispiel 4.4. (1) Aus dem letzten Beispiel wissen wir also, dass $\overline{\mathrm{B}(0,1)}$ in $\mathcal{C}[0,1]$ nicht total beschränkt ist.

(2) Das Intervall $(0,1]$ in $\mathbb{R}$ ist total beschränkt (für jedes $\varepsilon > 0$ kann man es mit endlich vielen offenen Intervallen der Länge ε überdecken). Der Grund für die Nichtkompaktheit ist hier, dass $(0,1]$ nicht vollständig ist.

Wir werden sehen, dass Kompaktheit einer Teilmenge A äquivalent ist zu Vollständigkeit (“A hat keine Löcher”) und totaler Beschränktheit (A ist “klein”).

Satz 4.5 (Borel-Lebesgue)**.** Sei A eine Teilmenge eines metrischen Raumes (X, d). Dann sind äquivalent:

(1) A ist kompakt.

(2) A ist folgenkompakt.

(3) A ist vollständig und total beschränkt.

Bemerkung 4.6. Ist (X, d) ein metrischer Raum und $A \subseteq X$, so ist auch (A, d) ein metrischer Raum und die Fragen nach Kompaktheit, Folgenkompaktheit, Vollständigkeit und totaler Beschränktheit von A haben die gleichen Antworten in (A, d) wie in (X, d).[14] Somit genügt es, im folgenden Beweis direkt metrische Räume zu betrachten statt beliebige Teilmengen metrischer Räume, also den Fall $A = X$.

Beweis von Satz 4.5. (1)⇒(2): Sei X kompakt und $(x_n)_{n\in\mathbb{N}}$ eine Folge in X. Angenommen, $(x_n)_{n\in\mathbb{N}}$ hat keine konvergente Teilfolge. Wir können annehmen, dass $x_n \neq x_m$ für alle $n \neq m$. [15] Dann existiert für jedes $n \in \mathbb{N}$ ein $\varepsilon_n > 0$, so dass $x_m \notin U_n := \mathrm{B}(x_n, \varepsilon_n)$ für alle $m \neq n$, da andernfalls eine Teilfolge existieren würde, die gegen x_n konvergiert.

Die Menge $A := \{x_n \mid n \in \mathbb{N}\}$ ist laut Satz 3.6 abgeschlossen. [16] Also ist $U_0 := X \setminus A$ offen. Somit ist

$$\{U_0\} \cup \{U_n \mid n \in \mathbb{N}\}$$

eine offene Überdeckung von X. Jede endliche Vereinigung der U_i enthält nur endlich viele Folgenglieder, kann also keine Überdeckung von X sein. Dies ist ein Widerspruch zur Kompaktheit von X, laut der eine endliche Teilüberdeckung existieren müsste.

Also muss $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge besitzen. Damit ist X folgenkompakt.

(2)⇒(3): Sei X folgenkompakt.

(i) Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in X und $\varepsilon > 0$, dann existiert ein $N_1 \in \mathbb{N}$, so dass

$$\forall n, m \geq N_1 : \quad d(x_n, x_m) < \frac{\varepsilon}{2}.$$

Da X folgenkompakt ist, existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$, die gegen ein $x \in X$ konvergiert. Dann existiert ein $N_2 \in \mathbb{N}$, so dass

$$\forall k \geq N_2 : \quad d(x_{n_k}, x) < \frac{\varepsilon}{2}.$$

Setze $N := \max\{N_1, N_2\}$. Für alle $m \geq N$ ist $n_m \geq m \geq N$, also gilt

$$\forall m \geq N : \quad d(x_m, x) \leq \underbrace{d(x_m, x_{n_m})}_{<\frac{\varepsilon}{2}} + \underbrace{f(x_{n_m}, x)}_{<\frac{\varepsilon}{2}} < \varepsilon.$$

Somit konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen $x \in X$, also ist X vollständig.

[14] Vorsicht: Für Offenheit etwa gilt dies nicht; zwar ist $A \subseteq A$ immer offen, aber $A \subseteq X$ nicht unbedingt.

[15] Taucht ein Wert unendlich oft in der Folge auf, so gibt dies direkt eine konvergente Teilfolge. Taucht jeder Wert nur endlich oft in der Folge auf, so können wir zu der Teilfolge übergehen, wo wir jede Wiederholung entfernen. Die Frage nach einer konvergenten Teilfolge wird davon nicht berührt.

[16] Jede konvergente Folge in A muss irgendwann konstant werden, da dies ansonsten eine konvergente Teilfolge der ursprünglichen Folge (x_n) liefern würde.

(ii) Angenommen, X ist nicht total beschränkt, dann gibt es ein $\varepsilon > 0$, so dass X nicht mit endlich vielen ε-Kugeln überdeckt werden kann. Wähle nun $x_1 \in X$ und

$$x_{n+1} \in X \setminus \bigcup_{j=1}^{n} \mathrm{B}(x_j, \varepsilon) \quad \text{für alle} \quad n \in \mathbb{N}.$$

Laut Annahme ist $\bigcup_{j=1}^{n} \mathrm{B}(x_j, \varepsilon)$ eine echte Teilmenge von X, also kann man immer ein x_{n+1} wählen. Insbesondere gilt $d(x_n, x_m) \geq \varepsilon$ für alle $n \neq m$. Folglich kann $(x_n)_{n\in\mathbb{N}}$ keine konvergente Teilfolge besitzen im Widerspruch zur Folgenkompaktheit.

Somit muss X total beschränkt sein.

(3)⇒(2): Sei X vollständig und total beschränkt. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X. Da X total beschränkt ist, gilt

$$\forall N \in \mathbb{N}\ \exists y_1^{(N)}, \ldots, y_{r_N}^{(N)} \in X : \quad X \subseteq \bigcup_{j=1}^{r_N} \mathrm{B}\left(y_j^{(N)}, \frac{1}{2N}\right).$$

Für $N = 1$ gilt

$$X \subseteq \bigcup_{j=1}^{r_1} \mathrm{B}\left(y_j^{(1)}, \frac{1}{2}\right).$$

Nun muss es (mindestens) ein $\ell_1 \in \{1, \ldots, r_1\}$ geben, so dass $\mathrm{B}\left(y_{\ell_1}^{(1)}, \frac{1}{2}\right)$ unendliche viele Folgenglieder von $(x_n)_{n\in\mathbb{N}}$ enthält. Sei also $(x_n^{(1)})_{n\in\mathbb{N}}$ die Teilfolge von $(x_n)_{n\in\mathbb{N}}$, deren Folgenglieder alle in $\mathrm{B}\left(y_{\ell_1}^{(1)}, \frac{1}{2}\right)$ liegen.

Für $N = 2$ gilt

$$X \subseteq \bigcup_{j=1}^{r_2} \mathrm{B}\left(y_j^{(2)}, \frac{1}{4}\right).$$

Nun muss es (mindestens) ein $\ell_2 \in \{1, \ldots, r_2\}$ geben, so dass $\mathrm{B}\left(y_{\ell_2}^{(2)}, \frac{1}{4}\right)$ unendliche viele Folgenglieder von $(x_n^{(1)})_{n\in\mathbb{N}}$ enthält, wir können also wieder eine Teilfolge $(x_n^{(2)})_{n\in\mathbb{N}}$ von $(x_n^{(1)})_{n\in\mathbb{N}}$ wählen, die vollständig in $\mathrm{B}\left(y_{\ell_2}^{(2)}, \frac{1}{4}\right)$ liegt.

Durch Iteration erhalten wir also Teilfolgen $(x_n^{(k)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$, so dass

- $(x_n^{(k+1)})_{n\in\mathbb{N}}$ eine Teilfolge von $(x_n^{(k)})_{n\in\mathbb{N}}$ ist für alle $k \in \mathbb{N}$, und
- $(x_n^{(k)})_{n\in\mathbb{N}}$ enthalten ist in $\mathrm{B}\left(y_{\ell_k}^{(k)}, \frac{1}{2k}\right)$.

Definiere nun $(z_n)_{n\in\mathbb{N}}$ als "Diagonalfolge" mittels $z_n := x_n^{(n)}$ für alle $n \in \mathbb{N}$, dann ist $(z_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(x_n)_{n\in\mathbb{N}}$.

Sei nun $\varepsilon > 0$. Laut Archimedes gibt es ein $N \in \mathbb{N}$ mit $N > \frac{1}{\varepsilon}$. Seien $n, m \geq N$ beliebig und oBdA $n \leq m$. Dann gilt $z_m, z_n \in \mathrm{B}\left(y_{\ell_n}^{(n)}, \frac{1}{2n}\right)$ und somit

$$d(z_m, z_n) \leq \underbrace{d(z_m, y_{\ell_n}^{(n)})}_{<\frac{1}{2n}} + \underbrace{d(y_{\ell_n}^{(n)}, z_n)}_{<\frac{1}{2n}} < \frac{1}{n} < \varepsilon,$$

also ist $(z_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge. Da X vollständig ist, konvergiert $(z_n)_{n\in\mathbb{N}}$ in X. Somit hat die ursprüngliche Folge $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge, also ist X folgenkompakt.

(2)⇒(1): Sei X folgenkompakt und $(U_i)_{i\in I}$ eine offene Überdeckung von X. Wir zeigen zunächst:

$$\exists r > 0\ \forall x \in X\ \exists i \in I : \quad \mathrm{B}(x, r) \subseteq U_i.$$

Dazu nehmen wir das Gegenteil an:

$$\forall r > 0\ \exists x \in X\ \forall i \in I : \quad \mathrm{B}(x, r) \not\subseteq U_i,$$

insbesondere

$$\forall n \in \mathbb{N}\ \exists x_n \in X\ \forall i \in I : \quad \mathrm{B}\left(x_n, \frac{1}{n}\right) \not\subseteq U_i.$$

Da X folgenkompakt ist, existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$, die gegen ein $x \in X$ konvergiert. Da $(U_i)_{i\in I}$ eine Überdeckung von X ist, existiert ein $\ell \in I$ mit $x \in U_\ell$. Da U_ℓ offen ist, existiert ein $\varepsilon > 0$ mit $\mathrm{B}(x, \varepsilon) \subseteq U_\ell$. Da $(x_{n_k})_{k\in\mathbb{N}}$ gegen x konvergiert, gibt es ein $N \in \mathbb{N}$, so dass

$$d(x_{n_k}, x) < \frac{\varepsilon}{2} \quad \text{für alle} \quad k \geq N.$$

Laut Archimedes können wir N so wählen, dass $N > \frac{2}{\varepsilon}$. Nun ist $n_N \geq N > \frac{2}{\varepsilon}$, also $\frac{1}{n_N} < \frac{\varepsilon}{2}$ und somit

$$\mathrm{B}\left(x_{n_N}, \frac{1}{n_N}\right) \subseteq \mathrm{B}\left(x_{n_N}, \frac{\varepsilon}{2}\right) \subseteq \mathrm{B}(x, \varepsilon) \subseteq U_\ell,$$

denn für alle $y \in \mathrm{B}(x_{n_N}, \frac{\varepsilon}{2})$ gilt

$$d(y, x) \leq \underbrace{d(y, x_{n_N})}_{<\frac{\varepsilon}{2}} + \underbrace{d(x_{n_N}, x)}_{<\frac{\varepsilon}{2}} < \varepsilon$$

und somit $y \in \mathrm{B}(x, \varepsilon)$. Nun ist $\mathrm{B}(x_{n_N}, \frac{1}{n_N}) \subseteq U_\ell$ ein Widerspruch, da diese Kugel laut Annahme in keinem U_i enthalten sein darf.

Also gilt wie gewünscht das Zwischenresultat

$$\exists r > 0\ \forall x \in X\ \exists i \in I : \quad \mathrm{B}(x, r) \subseteq U_i.$$

Wegen (2)⇒(3) ist X total beschränkt, also existieren zu diesem r endlich viele $y_1, \ldots, y_k \in X$ mit

$$X \subseteq \bigcup_{j=1}^{k} \mathrm{B}(y_j, r).$$

Zu jedem y_j existiert nun ein $i_j \in I$ mit $\mathrm{B}(y_j, r) \subseteq U_{i_j}$, also

$$X \subseteq \bigcup_{j=1}^{k} \mathrm{B}(y_j, r) \subseteq \bigcup_{j=1}^{k} U_{i_j}.$$

Dann ist $U_{i_1}, \ldots, U_{i_k}$ eine endliche Teilüberdeckung von X. Somit ist X kompakt. $\square$

Satz 4.7. Sei $f : X \to Y$ eine stetige Abbildung zwischen metrischen Räumen. Dann ist für jede kompakte Teilmenge K von X auch das Bild $f(K)$ kompakt.

Beweis. Sei $(U_i)_{i\in I}$ eine offene Überdeckung von $f(K)$, also

$$f(K) \subseteq \bigcup_{i\in I} U_i.$$

Laut Satz 3.12 sind alle $f^{-1}(U_i) \subseteq X$ offen, wegen

$$K \subseteq f^{-1}\big(f(K)\big) \subseteq f^{-1}\left(\bigcup_{i\in I} U_i\right) = \bigcup_{i\in I} f^{-1}(U_i)$$

ist also $\big(f^{-1}(U_i)\big)_{i\in I}$ eine offene Überdeckung von K. Da K kompakt ist, existiert eine offene Teilüberdeckung $f^{-1}(U_{i_1}), \dots, f^{-1}(U_{i_k})$ von K. Dann ist aber auch $U_{i_1}, \dots, U_{i_k}$ eine endliche Teilüberdeckung von $f(K)$. $\square$

Satz 4.8. Seien K und Y metrische Räume, wobei K kompakt. Dann ist jede stetige Abbildung $f : K \to Y$ sogar gleichmäßig stetig.

Beweis. Angenommen, f wäre nicht gleichmäßig stetig, also (mit $\delta = \frac{1}{n}$)

$$\exists \varepsilon > 0 \ \forall n \in \mathbb{N} \ \exists x_n, y_n \in K : \quad d(x_n, y_n) < \frac{1}{n} \wedge d\big(f(x_n), f(y_n)\big) \geq \varepsilon.$$

Als kompakte Menge ist K auch folgenkompakt, also existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$, die gegen $x \in K$ konvergiert. Wegen

$$0 \leq d(y_{n_k}, x) \leq \underbrace{d(y_{n_k}, x_{n_k})}_{<\frac{1}{n_k}\to 0} + \underbrace{d(x_{n_k}, x)}_{\to 0} \xrightarrow{k\to\infty} 0$$

konvergiert auch $(y_{n_k})_{k\in\mathbb{N}}$ gegen x. Da f stetig ist, gilt

$$f(x_{n_k}) \xrightarrow{k\to\infty} f(x) \xleftarrow{k\to\infty} f(y_{n_k}),$$

also $d\big(f(x_{n_k}, y_{n_k})\big) \xrightarrow{k\to\infty} 0$ im Widerspruch zu $d\big(f(x_n), f(y_n)\big) \geq \varepsilon$ für alle $n \in \mathbb{N}$.

Somit muss f gleichmäßig stetig sein. $\square$

Satz 4.9. Sei K eine kompakte Teilmenge eines metrischen Raumes. Dann nimmt jede stetige Funktion $f : K \to \mathbb{R}$ ihr Supremum und Infimum auf K an.

Beweis. Sei $\alpha := \inf_{x\in K} f(x)$, wobei $\alpha = -\infty$ zunächst möglich ist. Dann existiert eine Folge $(x_n)_{n\in\mathbb{N}}$ in K, so dass

$$f(x_n) \xrightarrow{n\to\infty} \alpha.$$

Da K kompakt ist, existiert eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit

$$x_{n_k} \xrightarrow{k\to\infty} x \in K.$$

Da f stetig ist, folgt

$$f(x_{n_k}) \xrightarrow{k\to\infty} f(x).$$

Da $\big(f(x_{n_k})\big)_{k\in\mathbb{N}}$ als Teilfolge von $\big(f(x_n)\big)_{n\in\mathbb{N}}$ aber auch gegen α konvergiert, muss aufgrund der Eindeutigkeit von Grenzwerten $\alpha = f(x) \in \mathbb{R}$ gelten.

Analog zeigt man die Aussage für das Supremum. $\square$

5. Die Sätze von Heine-Borel und Arzelà-Ascoli

Satz 5.1. Jede kompakte Teilmenge K eines metrischen Raumes X ist abgeschlossen und beschränkt.[17]

Beweis. Wir zeigen zunächst, dass K abgeschlossen ist: Sei dazu $(x_n)_{n\in\mathbb{N}}$ eine Folge in K mit $x_n \xrightarrow{n\to\infty} x \in X$. Da K kompakt ist, gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$, die gegen ein $y \in K$ konvergiert. Da nun notwendigerweise $x = y \in K$ gilt, ist K laut Satz 3.6 abgeschlossen.

Zur Beschränktheit sei $x \in K$ beliebig. Dann ist $\big(\mathrm{B}(x,n)\big)_{n\in\mathbb{N}}$ eine offene Überdeckung von K (für jedes $y \in K$ gibt es nach Archimedes ein $n \in \mathbb{N}$ mit $n > d(x,y)$, also gilt $y \in \mathrm{B}(x,n)$). Da K kompakt ist, gibt es nun eine endliche Teilüberdeckung $\mathrm{B}(x,n_1),\dots,\mathrm{B}(x,n_k)$. Dann gilt aber

$$K \subseteq \bigcup_{i=1}^{k} \mathrm{B}(x,n_i) = \mathrm{B}\big(x, \max\{n_1,\dots,n_k\}\big),$$

also ist K beschränkt. □

Bemerkung 5.2. Die Umkehrung der obigen Aussage gilt nicht: Abgeschlossenheit und Beschränktheit reicht im Allgemeinen nicht für Kompaktheit. Zwar gilt dies für $\mathbb{R}^n$ laut dem folgenden Satz von Heine-Borel, aber nicht für $X = \mathcal{C}[0,1]$ laut dem Satz von Arzelà-Ascoli (Satz 5.8) (dort erhalten wir aber immerhin eine zusätzliche Bedingung, mit der die kompakten Teilmengen charakterisiert werden können).

Zu beachten ist auch, dass $\mathbb{R}^n$ ein endlich-dimensionaler $\mathbb{R}$-Vektorraum ist, während $\mathcal{C}[0,1]$ unendlich-dimensional ist. Die Theorie unendlich-dimensionaler Vektorräume ist wesentlich komplizierter (und damit interessanter) als die Theorie endlich-dimensionaler Vektorräume.

Satz 5.3 (Heine-Borel)**.** Eine Teilmenge K von $\mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis. Sei $K \subseteq \mathbb{R}^n$ abgeschlossen und beschränkt, wir müssen zeigen, dass K kompakt ist. Sei dazu $(x_k)_{k\in\mathbb{N}}$ eine Folge in K, wobei $x_k = (x_k^{(1)},\dots,x_k^{(n)})$ mit $x_k^{(i)} \in \mathbb{R}$. Mit K ist auch $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ beschränkt, also ist für jedes $i = 1,\dots,n$ auch die Folge $(x_k^{(i)})_{k\in\mathbb{N}}$ in $\mathbb{R}$ beschränkt. Insbesondere ist $(x_k^{(1)})_{k\in\mathbb{N}}$ beschränkt. Laut Bolzano-Weierstraß gibt es eine konvergente Teilfolge $(x_{k_\ell}^{(1)})_{\ell\in\mathbb{N}}$. Nun ist $(x_{k_\ell}^{(2)})_{\ell\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, hat also wiederum laut Bolzano-Weierstraß eine konvergente Teilfolge $(x_{k_{\ell_p}}^{(2)})_{p\in\mathbb{N}}$. Wichtig ist, dass $(x_{k_{\ell_p}}^{(1)})_{p\in\mathbb{N}}$ eine Teilfolge von $(x_{k_\ell}^{(1)})_{\ell\in\mathbb{N}}$ ist und somit weiterhin konvergiert.

Wenn man dies nun bis n iteriert, erhält man eine Teilfolge $(x_{m_r})_{r\in\mathbb{N}}$ von $(x_k)_{k\in\mathbb{N}}$, so dass alle Komponenten $(x_{m_r}^{(i)})_{r\in\mathbb{N}}$ in $\mathbb{R}$ konvergieren, also konvergiert auch $(x_{m_r})_{r\in\mathbb{N}}$ in $\mathbb{R}^n$. Da K abgeschlossen ist, liegt der Grenzwert in K, also ist K (folgen)kompakt. □

Beispiel 5.4. (1) Die abgeschlossene Einheitskugel

$$\overline{\mathrm{B}(0,1)} = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

im $\mathbb{R}^n$ ist beschränkt, nach Heine-Borel also kompakt.

[17]Beschränkt heißt hier, dass es ein $x \in X$ und ein $r \in (0,\infty)$ gibt mit $K \subseteq \mathrm{B}(x,r)$.

(2) In Beispiel 4.2 haben wir gesehen, dass die abgeschlossene Einheitskugel

$$\overline{\mathrm{B}(0,1)} = \{f \in \mathcal{C}[0,1] : \|f\| \leq 1\}$$

in $\mathcal{C}[0,1]$ nicht kompakt ist. Diese ist zwar beschränkt, aber nicht total beschränkt. Es stellt sich die Frage, wie wir in $\mathcal{C}[0,1]$ Kompaktheit (beziehungsweise totale Beschränktheit) besser charakterisieren können.

Bemerkung 5.5. Betrachte eine Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{C}[0,1]$, die gegen $f \in \mathcal{C}[0,1]$ konvergiert, und ein $\varepsilon > 0$.

- Aufgrund der Konvergenz existiert ein $N \in \mathbb{N}$, so dass

$$\forall n \geq N \ \forall x \in [0,1] : \quad |f_n(x) - f(x)| < \frac{\varepsilon}{3}.$$

- Aufgrund der Stetigkeit von f existiert ein $\delta > 0$, so dass [18]

$$\forall x, y \in [0,1] : \quad |x-y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{3}.$$

Zusammengenommen gilt also für alle $x, y \in [0,1]$

$$|f_n(x) - f_n(y)| \leq \underbrace{|f_n(x) - f(x)|}_{<\frac{\varepsilon}{3} \text{ für } n \geq N} + \underbrace{|f(x) - f(y)|}_{<\frac{\varepsilon}{3} \text{ für } |x-y|<\delta} + \underbrace{|f_n(y) - f(y)|}_{<\frac{\varepsilon}{3} \text{ für } n \geq N} < \varepsilon.$$

Da auch alle f_i stetig sind, existieren $\delta_i > 0$, so dass

$$\forall x, y \in [0,1] : \quad |x-y| < \delta_i \Rightarrow |f_i(x) - f_i(y)| < \varepsilon.$$

Setzt man nun $\tilde{\delta} := \min\{\delta, \delta_1, \ldots, \delta_{N-1}\}$, so gilt

$$\forall m \in \mathbb{N} \ \forall x, y \in [0,1] : \quad |x-y| < \tilde{\delta} \Rightarrow |f_m(x) - f_m(y)| < \varepsilon.$$

Das ist die Stetigkeitsbedingung für die ganze Folge $(f_n)_{n\in\mathbb{N}}$ gleichzeitig mit dem gleichen "Grad" $\tilde{\delta}$.

Definition 5.6. Eine Teilmenge $A \subseteq \mathcal{C}[0,1]$ heißt *gleichgradig stetig*, falls

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall f \in A \ \forall x, y \in [0,1] : \quad |x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$ [19]

Bemerkung 5.7. Wenn A gleichgradig stetig ist, dann auch $\overline{A}$.

Satz 5.8 (Arzelà-Ascoli)**.** Sei K eine Teilmenge von $\mathcal{C}[0,1]$. Dann sind äquivalent:

(1) K ist kompakt.

(2) K ist abgeschlossen, beschränkt und gleichgradig stetig.

[18] Genau genommen benutzen wir hier die gleichmäßige Stetigkeit, aber wir wissen ja, dass stetige Funktionen auf kompakten Mengen wie dem Intervall $[0,1]$ gleichmäßig stetig sind.

[19] Entscheidend ist hier natürlich, dass ein δ für alle $f \in A$ funktionieren muss: "one δ to rule them all".

Beweis. (1)⇒(2): Sei K kompakt, dann ist K laut Satz 5.1 abgeschlossen und beschränkt. Angenommen, K wäre nicht gleichgradig stetig, dann existiert ein $\varepsilon > 0$, so dass für alle $n \in \mathbb{N}$ (mit $\delta = \frac{1}{n}$) gilt:

$$\exists f_n \in K \; \exists x_n, y_n \in [0,1] : \quad |x_n - y_n| < \frac{1}{n} \wedge |f_n(x_n) - f_n(y_n)| \geq \varepsilon.$$

Insbesondere ist keine Teilfolge von $(f_n)_{n\in\mathbb{N}}$ gleichgradig stetig. Da K kompakt ist, existiert eine konvergente Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$. Laut Bemerkung 5.5 ist $(f_{n_k})_{k\in\mathbb{N}}$ dann aber gleichgradig stetig, Widerspruch.

Somit muss K gleichgradig stetig sein.

(2)⇒(1): Sei K abgeschlossen, beschränkt und gleichgradig stetig. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in K, dann müssen wir zeigen, dass eine konvergente Teilfolge existiert.

Wir zeigen zunächst, dass es eine Teilfolge gibt, die punktweise an rationalen Argumenten konvergiert. Sei dazu $x_1, x_2, x_3, \dots$ eine Abzählung von $[0,1] \cap \mathbb{Q}$.

Da K beschränkt ist, ist $\big(f_n(x_1)\big)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, die laut Bolzano-Weierstraß eine konvergente Teilfolge besitzt, d.h. es gibt eine Teilfolge $(f_n^{(1)})_{n\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$, so dass

$$y_1 := \lim_{n\to\infty} f_n^{(1)}(x_1) \in \mathbb{R}$$

existiert. Da K beschränkt ist, ist $\big(f_n^{(1)}(x_2)\big)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, die laut Bolzano-Weierstraß eine konvergente Teilfolge besitzt, d.h. es gibt eine Teilfolge $(f_n^{(2)})_{n\in\mathbb{N}}$ von $(f_n^{(1)})_{n\in\mathbb{N}}$, so dass

$$y_2 := \lim_{n\to\infty} f_n^{(2)}(x_2) \in \mathbb{R}$$

existiert, wobei auch $\lim_{n\to\infty} f_n^{(2)}(x_1) = y_1$ gilt. Durch Iteration erhalten wir also Teilfolgen $(f_n^{(k)})_{n\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$, so dass

- $(f_n^{(k+1)})_{n\in\mathbb{N}}$ eine Teilfolge von $(f_n^{(k)})_{n\in\mathbb{N}}$ ist für alle $k \in \mathbb{N}$, und
- $\lim_{n\to\infty} f_n^{(k)}(x_i) = y_i$ für alle $i = 1, \dots, k$.

Da wir unsere Folge in jedem Schritt weiter ausdünnen, können wir als unsere gesuchte Teilfolge, welche an allen rationalen Punkten konvergiert, nicht einfach die Folge nehmen, die für $k \to \infty$ übrigbleibt – eventuell bleibt ja gar nichts übrig. Wir müssen uns etwas geschickter anstellen und wie beim Beweis von dem Satz von Borel-Lebesgue greift hier das Diagonalfolgen-Argument. Definiere also $(g_n)_{n\in\mathbb{N}}$ als "Diagonalfolge" mittels $g_n := f_n^{(n)}$. Abgesehen von endlich vielen Anfangsgliedern ist $(g_n)_{n\in\mathbb{N}}$ für jedes $k \in \mathbb{N}$ eine Teilfolge von $(f_n^{(k)})_{n\in\mathbb{N}}$, also gilt

$$\lim_{n\to\infty} g_n(x_k) = y_k \quad \text{für alle} \quad k \in \mathbb{N}.$$

Die Folge $(g_n)_{n\in\mathbb{N}}$ ist nun unser Kandidat für eine konvergente Teilfolge. Wir wissen, dass

- $(g_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(f_n)_{n\in\mathbb{N}}$ ist, und dass
- $\big(g_n(x)\big)_{n\in\mathbb{N}}$ konvergiert für alle $x \in [0,1] \cap \mathbb{Q}$.

Wir müssen zeigen, dass

- $\big(g_n(x)\big)_{n\in\mathbb{N}}$ für alle $x \in [0,1]$ konvergiert, und dass
- $(g_n)_{n\in\mathbb{N}}$ gleichmäßig konvergiert.

Betrachte also $x \in [0,1]$ und sei $\varepsilon > 0$. Da K gleichgradig stetig ist, gibt es ein $\delta > 0$, so dass

$$\forall n \in \mathbb{N}\ \forall i \in \mathbb{N}: \quad |x - x_i| < \delta \Rightarrow |g_n(x) - g_n(x_i)| < \frac{\varepsilon}{3}.$$

Sei also $i \in \mathbb{N}$ so, dass $|x - x_i| < \delta$ (x ist Grenzwert einer rationalen Folge in $[0,1]$, also muss es so ein i geben). Die Folge $\big(g_n(x_i)\big)_{n\in\mathbb{N}}$ konvergiert, ist also eine Cauchy-Folge. Somit existiert ein $N_i \in \mathbb{N}$, so dass

$$\forall m, n \geq N_i: \quad |g_n(x_i) - g_m(x_i)| < \frac{\varepsilon}{3}.$$

Für alle $n, m \geq N_i$ gilt dann aber auch

$$|g_n(x) - g_m(x)| \leq \underbrace{|g_n(x) - g_n(x_i)|}_{<\frac{\varepsilon}{3}} + \underbrace{|g_n(x_i) - g_m(x_i)|}_{<\frac{\varepsilon}{3}} + \underbrace{|g_m(x_i) - g_m(x)|}_{<\frac{\varepsilon}{3}} < \varepsilon,$$

also ist $\big(g_n(x)\big)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$ und somit konvergent.

Um gleichmäßige Konvergenz zu erhalten, muss obige Abschätzung von x unabhängig sein. Dies gilt im Augenblick nicht, da N_i von der Wahl von x_i abhängt. Falls wir hingegen nur endlich viele x_i benutzen müssen, dann haben wir auch nur endlich viele N_i und können dann N als Maximum dieser N_i wählen. Somit brauchen wir endlich viele $x_i \in \mathbb{Q}$, so dass für jedes $x \in [0,1]$ mindestens ein x_i existiert mit $|x - x_i| < \delta$ für ein (durch die gleichgradige Stetigkeit vorgegebenes) δ.

Wähle also $r \in \mathbb{Q}$ mit $0 < r < \min\{\delta, 1\}$ und als x_i die rationalen Zahlen

$$r, 2r, 3r, \ldots, \left\lfloor \frac{1}{r} \right\rfloor \cdot r.$$

Dann haben wir die gleichmäßige Abschätzung

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall x \in [0,1]\ \forall n, m \geq N: \quad |g_n(x) - g_m(x)| < \varepsilon,$$

also konvergiert $(g_n)_{n\in\mathbb{N}}$ gleichmäßig. □

Beispiel 5.9. Betrachte eine Folge $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{C}[0,1]$ mit $\|f_n\|_\infty \leq 1$ und der zusätzlichen Eigenschaft, dass alle f_n auf $(0,1)$ differenzierbar sind mit $\|f_n'\|_\infty \leq c$ für ein $c \in [0,\infty)$.[20] Dann ist $K := \{f_n \mid n \in \mathbb{N}\}$ gleichgradig stetig, denn für alle $x, y \in [0,1]$ existiert laut Mittelwertsatz ein ξ zwischen x und y mit

$$|f_n(x) - f_n(y)| = |f_n'(\xi)| \cdot |x - y| \leq \|f_n'\|_\infty \cdot |x - y| \leq c \cdot |x - y|.$$

Somit ist $\overline{K}$ gleichgradig stetig (laut Bemerkung 5.7), abgeschlossen und beschränkt, laut Arzelà-Ascoli also kompakt. Folglich enthält $(f_n)_{n\in\mathbb{N}}$ eine gleichmäßig konvergente Teilfolge.

[20] Die Beschränkung der Ableitung schließt Beispiele wie die immer schmaler (und damit steiler) werdenden Zackenfunktionen aus Beispiel 4.2 aus.

6. Parametrisierte Kurven im $\mathbb{R}^n$

Konvention. In diesem Kapitel seien stets $a, b \in \mathbb{R}$ mit $a < b$, so dass $[a,b] \subseteq \mathbb{R}$ ein kompaktes Intervall ist. Wir betrachten $\mathbb{R}^n$ mit der Euklidischen Norm $\|\cdot\|_2$ und schreiben dafür abkürzend nur $\|\cdot\|$.

Definition 6.1. Eine *(parametrisierte) Kurve* in $\mathbb{R}^n$ ist eine Abbildung $\gamma : [a,b] \to \mathbb{R}^n$.

Bemerkung 6.2. (1) Sei $\gamma : [a,b] \to \mathbb{R}^n$ eine parametrisierte Kurve. Dann ist

$$\gamma(t) = \big(\gamma_1(t), \dots, \gamma_n(t)\big) \in \mathbb{R}^n \quad \text{für alle} \quad t \in [a,b]$$

für geeignete Funktionen $\gamma_i : [a,b] \to \mathbb{R}$. Weiterhin ist γ genau dann stetig, wenn alle Komponentenfunktionen γ_i stetig sind.

(2) Eine parametrisierte Kurve ist eine Kurve als geometrisches Objekt mit einer Parametrisierung, also einer Vorschrift, wie die Kurve durchlaufen wird.

Beispiel 6.3. Für $n = 2$:

(i) Sei $r > 0$ und $\gamma : [a,b] \to \mathbb{R}^2,\ t \mapsto \big(r\cos(t), r\sin(t)\big)$. Dann durchläuft γ den Kreis mit Radius r gegen den Uhrzeigersinn von Winkel a bis Winkel b. Insbesondere erhält man für $[a,b] = [0, 2\pi]$ und $[a,b] = [0, 4\pi]$ das gleiche geometrische Objekt (den Kreis mit Radius r), allerdings mit unterschiedlichen Parametrisierung: Im ersten Fall wird es einmal, im zweiten Fall zweimal durchlaufen.

(ii) Für $c, d \in (0, \infty)$ durchläuft $\gamma : [0, 2\pi] \to \mathbb{R}^2$ mit $\gamma(t) = \big(c\cos(t), d\sin(t)\big)$ die Ellipse mit Halbachsen c und d.

(iii) Für $\gamma : [a,b] \to \mathbb{R}^2$ mit $\gamma(t) = (t, t^2)$ erhält man als geometrisches Objekt den Graphen der Funktion $f : [a,b] \to \mathbb{R},\ x \mapsto x^2$ (Parabel).

Definition 6.4. Eine Abbildung $\gamma : [a,b] \to \mathbb{R}^n$ heißt *differenzierbar* im Punkt $t_0 \in [a,b]$ mit Ableitung $\dot\gamma(t_0)$, falls für jede Folge $(t_k)_{k\in\mathbb{N}}$ in $[a,b] \setminus \{t_0\}$ mit Grenzwert t_0

$$\dot\gamma(t_0) = \lim_{k\to\infty} \frac{\gamma(t_k) - \gamma(t_0)}{t_k - t_0}.$$

Hierbei heißt $\dot\gamma(t_0)$ *Tangential-* oder *Geschwindigkeitsvektor* zu γ im Punkt t_0. Wie üblich heißt γ differenzierbar, wenn γ in jedem $t_0 \in [a,b]$ differenzierbar ist. Dann ist $\dot\gamma : [a,b] \to \mathbb{R}^n$ wieder eine Abbildung. Weiterhin heißt γ *stetig differenzierbar*, wenn γ differenzierbar ist und $\dot\gamma$ stetig ist.

Bemerkung 6.5. (1) Eine Abbildung $\gamma(t) = \big(\gamma_1(t), \dots, \gamma_n(t)\big)$ ist genau dann in t_0 differenzierbar, wenn $\gamma_1, \dots, \gamma_n$ jeweils in t_0 differenzierbar sind. In diesem Fall gilt

$$\dot\gamma(t_0) = \big(\gamma_1'(t_0), \dots, \gamma_n'(t_0)\big).$$

(2) Jede differenzierbare Abbildung $\gamma : [a,b] \to \mathbb{R}^n$ ist stetig.

Das Ziel ist es nun, die Länge einer Kurve γ durch lineare Approximation zu bestimmen.

Erinnerung. Eine *Unterteilung* des Intervalls $[a, b]$ ist ein Tupel $T = (x_0, \ldots, x_n)$ mit $x_0, \ldots, x_n \in \mathbb{R}$, so dass

$$a = x_0 < x_1 < \ldots < x_n = b.$$

Hierbei heißen die $x_0, \ldots, x_n$ auch *Stützstellen* von T.

Ist $S = (y_0, \ldots, y_m)$ eine weitere Unterteilung von $[a, b]$ mit m Abschnitten, so schreiben wir $S \cup T$ für die eindeutige Unterteilung von $[a, b]$, die $x_0, \ldots, x_n, y_0, \ldots, y_m$ als Stützstellen hat (nach Aussortieren von Duplikaten).

Wir schreiben $S \subseteq T$, wenn jede Stützstelle von S auch eine Stützstelle von T ist. In diesem Fall heißt T *feiner* als S.

Notation 6.6. Sei $\gamma : [a, b] \to \mathbb{R}^n$ eine stetige Kurve und $P = (t_0, \ldots, t_k)$ eine Unterteilung von $[a, b]$.[21] Wir setzen

$$L(P) := \sum_{i=1}^{k} \|\gamma(t_i) - \gamma(t_{i-1})\|$$

(die Länge des von P beschriebenen "Sehnenpolygons"). Ist P' feiner als P, so gilt nach Dreiecksungleichung $L(P') \geq L(P)$.

Definition 6.7. Eine stetige Kurve $\gamma : [a, b] \to \mathbb{R}^n$ heißt *rektifizierbar*, falls

$$L(\gamma) := \sup \{L(P) \mid P \text{ Unterteilung von } [a, b]\}$$

in $\mathbb{R}$ existiert. In diesem Fall heißt $L(\gamma)$ *Länge* von γ.

Bemerkung 6.8. Nicht jede stetige Kurve ist rektifizierbar. Beispielsweise gibt es eine stetige Kurve $\gamma : [0, 1] \to \mathbb{R}^2$ mit $\gamma([0, 1]) = [0, 1] \times [0, 1]$, für die daher $L(\gamma) = \infty$ gilt (*Peano-Kurve*).

Satz 6.9. Jede stetig differenzierbare Kurve $\gamma : [a, b] \to \mathbb{R}^n$ ist rektifizierbar mit

$$L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| \, \mathrm{d}t.$$[22]

Beispiel 6.10. Sei

$$\gamma : [0, \theta] \to \mathbb{R}^2, \quad t \mapsto \big(\cos(t), \sin(t)\big).$$

Für alle $t \in [0, \theta]$ gilt $\dot{\gamma}(t) = \big(-\sin(t), \cos(t)\big)$, also

$$\|\dot{\gamma}(t)\| = \sqrt{\sin^2(t) + \cos^2(t)} = 1.$$

Daraus folgt

$$L(\gamma) = \int_0^\theta \|\dot{\gamma}(t)\| \, \mathrm{d}t = \int_0^\theta 1 \, \mathrm{d}t = \theta.$$

Der Winkel θ im Bogenmaß ist also eigentlich die Länge des Kreisbogens von 0 bis θ.

Wir zeigen zunächst einige Hilfsaussagen:

Lemma 6.11. Sei $\gamma : [a, b] \to \mathbb{R}^n$ eine stetig differenzierbare Kurve. Dann gilt:

[21] Zur Erinnerung: Das bedeutet, dass $a = t_0 < t_1 < \ldots < t_{k-1} < t_k = b$.

[22] Anschaulich formuliert: Der zurückgelegte Weg ist das Integral über die Geschwindigkeit.

(a) Die Abbildung $\dot\gamma$ ist gleichmäßig stetig, also

$$\forall \alpha > 0\ \exists \delta > 0\ \forall \theta_1, \theta_2 \in [a,b]: \quad |\theta_1 - \theta_2| < \delta \Rightarrow \|\dot\gamma(\theta_1) - \dot\gamma(\theta_2)\| < \alpha.$$

(b) Für alle $\alpha > 0$ existiert ein $\delta > 0$, so dass für alle $s, t, \theta \in [a,b]$ mit $a \le s \le \theta \le t \le b$ und $|s-t| < \delta$

$$\left\| \dot\gamma(\theta) - \frac{\gamma(t) - \gamma(s)}{t-s} \right\| < \alpha.$$

(c) Für alle $\alpha > 0$ existiert ein $\delta > 0$, so dass für alle $s, t, \theta \in [a,b]$ mit $a \le s \le \theta \le t \le b$ und $|s-t| < \delta$

$$\left| \|\dot\gamma(\theta)\| - \left\| \frac{\gamma(t) - \gamma(s)}{t-s} \right\| \right| < \alpha.$$

(d) Für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für jede Unterteilung $P = (t_0, \ldots, t_k)$ von $[a,b]$ mit Feinheit höchstens δ (also $|t_i - t_{i-1}| < \delta$ für alle i) gilt:

$$\left| \int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t - L(P) \right| < \varepsilon.$$

Beweis. (a) Folgt aus Satz 4.8, da $\dot\gamma$ stetig und $[a,b]$ kompakt.

(b) Sei $\alpha > 0$ und $\gamma(t) = \big(\gamma_1(t), \ldots, \gamma_n(t)\big)$ für alle $t \in [a,b]$. Laut (a) gibt ein $\delta > 0$, so dass für alle $\theta_1, \theta_2 \in [a,b]$ mit $|\theta_1 - \theta_2|$ schon

$$\|\dot\gamma(\theta_1) - \dot\gamma(\theta_2)\| < \sqrt{\frac{\alpha^2}{n}}$$

gilt, also auch

$$|\gamma_k'(\theta_1) - \gamma_k'(\theta_2)| \le \|\dot\gamma(\theta_1) - \dot\gamma(\theta_2)\| < \sqrt{\frac{\alpha^2}{n}} \quad \text{für alle} \quad k = 1, \ldots, n.$$

Seien nun $s, t, \theta \in [a,b]$ mit $a \le s \le \theta \le t \le b$ und $|s-t| < \delta$. Für alle $k = 1, \ldots, n$ existiert laut Mittelwertsatz ein $\theta_k \in [s,t]$ mit

$$\frac{\gamma_k(t) - \gamma_k(s)}{t-s} = \gamma_k'(\theta_k).$$

Wegen $\theta, \theta_k \in [s,t]$ und $|s-t| < \delta$ folgt $|\theta - \theta_k| < \delta$, also

$$\begin{aligned}
\left\| \dot\gamma(\theta) - \frac{\gamma(t) - \gamma(s)}{t-s} \right\|^2 &= \sum_{k=1}^n \left| \gamma_k'(\theta) - \frac{\gamma_k(t) - \gamma_k(s)}{t-s} \right|^2 \\
&= \sum_{k=1}^n \underbrace{|\gamma_k'(\theta) - \gamma_k'(\theta_k)|^2}_{<\sqrt{\frac{\alpha^2}{n}}} \\
&< \sum_{k=1}^n \frac{\alpha^2}{n} = \alpha^2.
\end{aligned}$$

Mit der Monotonie der Wurzelfunktion folgt die Aussage.

(c) Folgt aus (b) mit der umgekehrten Dreiecksungleichung: $\|x-y\| \geq |\|x\|-\|y\||$ für alle $x, y \in \mathbb{R}^n$.

(d) Sei $\varepsilon > 0$. Zu $\alpha := \frac{\varepsilon}{2(b-a)}$ wähle $\delta = \min(\delta_a, \delta_c) > 0$ mit δ_a wie in (a) und δ_c wie in (c) (so dass für dieses δ sowohl (a) als auch (c) gelten) und beliebige $\theta_i \in [t_{i-1}, t_i]$. Einerseits gilt

$$
\begin{aligned}
&\left| \int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t - \sum_{i=1}^k \|\dot\gamma(\theta_i)\| \cdot (t_i - t_{i-1}) \right| \\
&= \left| \sum_{i=1}^k \int_{t_{i-1}}^{t_i} \|\dot\gamma(t)\| \,\mathrm{d}t - \sum_{i=1}^k \int_{t_{i-1}}^{t_i} \|\dot\gamma(\theta_i)\| \,\mathrm{d}t \right| \\
&= \left| \sum_{i=1}^k \int_{t_{i-1}}^{t_i} \|\dot\gamma(t)\| - \|\dot\gamma(\theta_i)\| \,\mathrm{d}t \right| \\
&\leq \sum_{i=1}^k \int_{t_{i-1}}^{t_i} |\|\dot\gamma(t)\| - \|\dot\gamma(\theta_i)\|| \,\mathrm{d}t \\
&\overset{\text{(a)}}{\leq} \sum_{i=1}^k \int_{t_{i-1}}^{t_i} \alpha \,\mathrm{d}t = \sum_{i=1}^k \alpha (t_i - t_{i-1}) = \alpha(b-a)
\end{aligned}
$$

und andererseits

$$
\begin{aligned}
&\left| \sum_{i=1}^k \|\dot\gamma(\theta_i)\| \cdot (t_i - t_{i-1}) - \sum_{i=1}^k \|\gamma(t_i) - \gamma(t_{i-1})\| \right| \\
&= \left| \sum_{i=1}^k \left(\|\dot\gamma(\theta_i)\| - \frac{\|\gamma(t_i) - \gamma(t_{i-1})\|}{t_i - t_{i-1}} \right) \cdot (t_i - t_{i-1}) \right| \\
&\leq \sum_{i=1}^k \left| \|\dot\gamma(\theta_i)\| - \frac{\|\gamma(t_i) - \gamma(t_{i-1})\|}{t_i - t_{i-1}} \right| \cdot (t_i - t_{i-1}) \\
&\overset{\text{(c)}}{\leq} \sum_{i=1}^k \alpha \cdot (t_i - t_{i-1}) = \alpha(b-a),
\end{aligned}
$$

also

$$
\begin{aligned}
\left| \int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t - L(P) \right| &= \left| \int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t - \sum_{i=1}^k \|\gamma(t_i) - \gamma(t_{i-1})\| \right| \\
&\leq \left| \int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t - \sum_{i=1}^k \|\dot\gamma(\theta_i)\| \cdot (t_i - t_{i-1}) \right| \\
&\quad + \left| \sum_{i=1}^k \|\dot\gamma(\theta_i)\| \cdot (t_i - t_{i-1}) - \sum_{i=1}^k \|\gamma(t_i) - \gamma(t_{i-1})\| \right| \\
&\leq \alpha(b-a) + \alpha(b-a) = 2\alpha(b-a) < \varepsilon.
\end{aligned}
$$

□

Beweis von Satz 6.9. Sei $\gamma : [a,b] \to \mathbb{R}^n$ stetig differenzierbar. Aus Lemma 6.11 folgt:

- Für alle Unterteilungen P von $[a,b]$ gilt

$$L(P) \leq \int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t,$$

denn für jedes $\varepsilon > 0$ wähle ein $\delta > 0$ wie in Teil (d). Dann können wir eine Unterteilung $P' \supseteq P$ mit Feinheit höchstens δ wählen, so dass

$$L(P) \leq L(P') \leq \int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t + \varepsilon.$$

Da dies für alle $\varepsilon > 0$ gilt, folgt die Aussage.

- Die Zahl $\int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t$ kann beliebig genau durch Zahlen der Form $L(P)$ approximiert werden.

Zusammengenommen gilt also

$$\int_a^b \|\dot\gamma(t)\| \,\mathrm{d}t = \sup\{L(P) \mid P \text{ Unterteilung von } [a,b]\} = L(\gamma). \qquad \square$$

Bemerkung 6.12. (1) Sei $\gamma : [a,b] \to \mathbb{R}^n$ eine stetige Kurve. Sind $a', b' \in \mathbb{R}$ mit $a' < b'$ und $\sigma : [a',b'] \to [a,b]$ eine stetige und monotone Bijektion, so ist

$$\tilde\gamma : [a',b'] \to \mathbb{R}^n, \quad t \mapsto \tilde\gamma(t) := \gamma\big(\sigma(t)\big)$$

eine *Umparametrisierung* von γ mit $L(\gamma) = L(\tilde\gamma)$ (da die Sehnenpolygone von $\tilde\gamma$ die selben sind wie die von γ).

(2) Eine stetig differenzierbare Kurve $\gamma : [a,b] \to \mathbb{R}^n$ heißt *regulär*, falls $\|\dot\gamma(t)\| \neq 0$ für alle $t \in [a,b]$. Falls γ regulär ist, ist

$$s : [a,b] \to [0, L(\gamma)], \quad t \mapsto s(t) := \int_a^t \|\dot\gamma(u)\| \,\mathrm{d}u$$

stetig differenzierbar mit $s'(t) = \|\dot\gamma(t)\| > 0$, also ist s monoton und bijektiv. Somit existiert eine differenzierbare Umkehrfunktion $\varphi : [0, L(\gamma)] \to [a,b]$ von s. Für

$$\tilde\gamma : [0, L(\gamma)] \to \mathbb{R}^n, \quad t \mapsto \tilde\gamma(t) := \gamma\big(\varphi(t)\big)$$

gilt

$$\dot{\tilde\gamma}(t) = \dot\gamma\big(\varphi(t)\big) \cdot \varphi'(t) = \dot\gamma\big(\varphi(t)\big) \cdot \frac{1}{s'\big(\varphi(t)\big)} = \dot\gamma\big(\varphi(t)\big) \cdot \frac{1}{\big\|\dot\gamma\big(\varphi(t)\big)\big\|},$$

also $\big\|\dot{\tilde\gamma}(t)\big\| = 1$ für alle t. Man nennt $\tilde\gamma$ die *natürliche Parametrisierung* für γ.

7. Differenzierbare Funktionen in mehreren Variablen

Konvention. (1) Im Folgenden benutzen wir auf $\mathbb{R}^n$ (und auf $\mathbb{R}^m$) eine fixierte Norm. Da wir wissen, dass dort alle Normen äquivalent sind, spielt die spezifische Wahl von $\|\cdot\|$ keine Rolle für unsere Aussagen.

(2) Im Folgenden bezeichnen wir mit $e_i \in \mathbb{R}^n$ stets den i-ten Standardbasisvektor des $\mathbb{R}^n$, also den Vektor, der in der i-ten Zeile den Wert 1 annimmt und sonst überall den Wert 0. Somit ist e_i die i-te Spalte der Einheitsmatrix vom Format $n \times n$.

Bemerkung 7.1. Seien $n, m \in \mathbb{N}$ fest. Wir betrachten Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$. Diese sind stetig, falls

$$\forall x \in \mathbb{R}^n \ \forall \varepsilon > 0 \ \exists \delta > 0 \ \forall y \in \mathbb{R}^n : \quad \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \varepsilon$$

oder äquivalenterweise falls für jede Folge $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ und jedes $x \in \mathbb{R}^n$ aus der Konvergenz $x_k \xrightarrow{k\to\infty} x$ in $\mathbb{R}^n$ auch die Konvergenz $f(x_k) \xrightarrow{k\to\infty} f(x)$ in $\mathbb{R}^m$ folgt, siehe auch Satz 3.12.

Falls wir f koordinatenweise schreiben als $f = (f_1, \dots, f_m)$ mit $f_i : \mathbb{R}^n \to \mathbb{R}$, so ist f genau dann stetig, wenn alle f_i stetig sind (analog zur Situation für Kurven, Bemerkung 6.2).

Allerdings ist Stetigkeit nicht äquivalent zu "Stetigkeit in jeder Variablen"!

Beispiel 7.2. Betrachte

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2}, & \text{falls } (x, y) \neq (0, 0), \\ 0, & \text{falls } (x, y) = (0, 0). \end{cases}$$

Dann ist für jedes feste $x_0 \in \mathbb{R}$ die Funktion

$$g : \mathbb{R} \to \mathbb{R}, \quad y \mapsto f(x_0, y) = \begin{cases} \frac{2x_0 y}{x_0^2+y^2}, & \text{falls } x_0 \neq 0, \\ 0, & \text{falls } x_0 = 0 \end{cases}$$

stetig. Analog ist auch für jedes feste $y_0 \in \mathbb{R}$ die Funktion

$$h : \mathbb{R} \to \mathbb{R}, \quad x \mapsto f(x, y_0) = \begin{cases} \frac{2xy_0}{x^2+y_0^2}, & \text{falls } y_0 \neq 0, \\ 0, & \text{falls } y_0 = 0 \end{cases}$$

stetig. Allerdings ist f in $(0, 0)$ nicht stetig, denn für die Folge $(\frac{1}{k}, \frac{1}{k})_{k\in\mathbb{N}}$ gilt

$$\left(\frac{1}{k}, \frac{1}{k}\right) \xrightarrow{k\to\infty} (0, 0),$$

aber

$$f\left(\frac{1}{k}, \frac{1}{k}\right) = \frac{2 \cdot \frac{1}{k} \cdot \frac{1}{k}}{(\frac{1}{k})^2 + (\frac{1}{k})^2} = 1 \xrightarrow{k\to\infty} 1 \neq 0 = f(0, 0).$$

Motivation 7.3. Welche Art von Objekt ist die Ableitung einer Funktion $f : \mathbb{R}^n \to \mathbb{R}^m$?

Im Spezialfall $n = m = 1$ mit $f : \mathbb{R} \to \mathbb{R}$ haben wir

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

oder äquivalent

$$f(x+h) - f(x) = f'(x) \cdot h + r(h),$$

wobei der Restterm $r(h)$ klein ist im Sinne von $\lim_{h\to 0} \frac{r(h)}{h} = 0$. Anders ausgedrückt approximieren wir $f(x+h) - f(x)$ durch eine lineare Funktion

$$\mathbb{R} \to \mathbb{R}, \quad h \mapsto f'(x) \cdot h.$$

Im allgemeinen Fall $f : \mathbb{R}^n \to \mathbb{R}^m$ wollen wir ebenso

$$f(x+h) - f(x) = A(h) + r(h),$$

wobei $r(h)$ für $h \to 0$ "klein" wird, durch eine lineare Funktion

$$\mathbb{R}^n \to \mathbb{R}^m, \quad h \mapsto A(h)$$

approximieren. Wir werden dann $f'(x) := A$ setzen und f' mit Df bezeichnen.

Definition 7.4. Eine *lineare Abbildung* (oder *Operator*) zwischen zwei $\mathbb{K}$-Vektorräumen X und Y ist eine Abbildung $A : X \to Y$ mit

$$A(x_1 + x_2) = A(x_1) + A(x_2) \quad \text{für alle} \quad x_1, x_2 \in X \qquad \text{(Additivität)}$$

und

$$A(\alpha x) = \alpha A(x) \quad \text{für alle} \quad x \in X \quad \text{und alle} \quad \alpha \in \mathbb{K}. \qquad \text{(Homogenität)}$$

Bemerkung 7.5. (1) Ist A linear, so schreibt man oft Ax für $A(x)$.

(2) Lineare Operatoren auf endlich-dimensionalen Vektorräumen wie $\mathbb{R}^n$ sind Gegenstand der Linearen Algebra.

(3) Lineare Operatoren auf unendlich-dimensionalen Vektorräumen wie $\mathcal{C}[0,1]$ sind Gegenstand der Funktionalanalysis.

(4) Ist A linear, so folgt $A(0) = 0$.

(5) Sei $X = \mathbb{R}^n$ mit Basis $b_1, \dots, b_n$ und $Y = \mathbb{R}^m$ mit Basis $c_1, \dots, c_m$, dann ist

$$Ab_i = \sum_{j=1}^{m} a_{ji} c_j \quad \text{für alle} \quad i = 1, \dots, n$$

für geeignete $a_{ji} \in \mathbb{K}$ und A kann mit der Matrix $(a_{ji})_{\substack{j=1,\dots,m \\ i=1,\dots,n}}$ identifiziert werden. Dann ist Ax gerade die Multiplikation der Matrix A mit x als Spaltenvektor.

Definition 7.6. (i) Mit $L(\mathbb{R}^n, \mathbb{R}^m)$ bezeichnen wir die Menge aller linearen Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$.

(ii) Für $A_1, A_2 \in L(\mathbb{R}^n, \mathbb{R}^m)$ und $\alpha_1, \alpha_2 \in \mathbb{R}$ definieren wir $\alpha_1 A_1 + \alpha_2 A_2 \in L(\mathbb{R}^n, \mathbb{R}^m)$ durch

$$(\alpha_1 A_1 + \alpha_2 A_2)(x) := \alpha_1 A_1(x) + \alpha_2 A_2(x) \quad \text{für alle} \quad x \in \mathbb{R}^n.$$

(iii) Für $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ und $B \in L(\mathbb{R}^m, \mathbb{R}^p)$ definieren wir $BA \in L(\mathbb{R}^n, \mathbb{R}^p)$ durch

$$(BA)(x) := B\big(A(x)\big) \quad \text{für alle} \quad x \in \mathbb{R}^n.$$

Bemerkung 7.7. (1) Unter der Identifikation mit Matrizen (für eine feste Wahl der Basen) entspricht obige Definition der Addition und Multiplikation von Matrizen.

(2) $L(\mathbb{R}^n, \mathbb{R}^m)$ wird so zu einem (endlich-dimensionalen) $\mathbb{R}$-Vektorraum. Durch

$$\|A\| := \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} = \sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|Ax\| < \infty$$

wird eine Norm auf $L(\mathbb{R}^n, \mathbb{R}^m)$ gegeben. Damit wird $L(\mathbb{R}^n, \mathbb{R}^m)$ zu einem Banachraum.

(3) Aus der Definition folgt $\|Ax\| \leq \|A\| \cdot \|x\|$ für alle $x \in \mathbb{R}^n$. Ist also $(x_k)_{k \in \mathbb{N}}$ eine Folge in $\mathbb{R}^n$, die gegen ein $x \in \mathbb{R}^n$ konvergiert, so gilt

$$\|Ax_k - Ax\| = \|A(x_k - x)\| \leq \|A\| \cdot \|x_k - x\| \xrightarrow{k \to \infty} 0,$$

also ist jede lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ stetig. Sie ist sogar gleichmäßig stetig: Für $\varepsilon > 0$ setze $\delta := \frac{\varepsilon}{\|A\|}$.[23] Für alle $x, y \in \mathbb{R}^n$ mit $\|x - y\| < \delta$ gilt dann

$$\|Ax - Ay\| = \|A(x - y)\| \leq \|A\| \cdot \|x - y\| < \|A\| \cdot \frac{\varepsilon}{\|A\|} = \varepsilon.$$

(4) Man kann obiges auch auf lineare Abbildungen $A : X \to Y$ zwischen beliebigen normierten Vektorräumen verallgemeinern. Allerdings gilt dann $\|A\| < \infty$ nicht mehr unbedingt für alle $A \in L(X, Y)$: Lineare Abbildungen zwischen unendlich-dimensionalen Vektorräumen sind nicht automatisch stetig!

Definition 7.8. Sei $U \subseteq \mathbb{R}^n$ offen. Eine Abbildung $f : U \to \mathbb{R}^m$ heißt *(total) differenzierbar* in $x \in U$, falls es eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ und ein $\varepsilon > 0$ gibt, so dass

$$f(x + \xi) = f(x) + L(\xi) + \varphi(\xi)$$

für alle $\xi \in \mathbb{R}^n$ mit $\|\xi\| < \varepsilon$, wobei $\varphi : \mathrm{B}(0, \varepsilon) \to \mathbb{R}^m$ eine Funktion ist mit

$$\lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \frac{\varphi(\xi)}{\|\xi\|} = 0.$$

Wir bezeichnen L mit $Df(x)$ und nennen es die *(totale) Ableitung* von f im Punkt x.

[23] Für $\|A\| = 0$ geht das natürlich nicht, aber in diesem Fall ist $A = 0$ als konstante Abbildung sowieso gleichmäßig stetig.

Bemerkung 7.9. (1) Falls f in x differenzierbar ist, so ist f in x auch stetig, denn ist $(\xi_k)_{k\in\mathbb{N}}$ eine Nullfolge in $\mathrm{B}(0,\varepsilon)$, so gilt mit der Stetigkeit von L

$$f(x+\xi_k) = f(x) + L(\xi_k) + \underbrace{\frac{\varphi(\xi_k)}{\|\xi_k\|}}_{\to 0} \cdot \underbrace{\|\xi_k\|}_{\to 0} \xrightarrow{k\to\infty} f(x) + \underbrace{L(0)}_{=0} + 0 = f(x).$$

(2) Sei $f = A$ linear, dann gilt

$$A(x+\xi) = Ax + A\xi + 0,$$

mit $L = A$ und $\varphi = 0$ ist A also in jedem $x \in \mathbb{R}^n$ differenzierbar mit $DA(x) = A$, also $DA = A$.

(3) Die lineare Abbildung L in Definition 7.8 ist eindeutig bestimmt (falls sie existiert), also ist $Df(x)$ wohldefiniert.

(4) Falls f in x differenzierbar ist, so ist f in alle Richtungen *partiell differenzierbar*: Für festes $\xi \in \mathbb{R}^n$ und $\gamma(t) := \gamma_\xi(t) = x + t\xi$ mit $t \in \mathbb{R}$ und $|t|$ hinreichend klein ist $f \circ \gamma$ in $t = 0$ differenzierbar, denn aus

$$f\big(\gamma(t)\big) = f\big(x + t\xi\big) = f(x) + L(t\xi) + \varphi(t\xi) = f\big(\gamma(0)\big) + tL(\xi) + \varphi(t\xi)$$

folgt

$$\frac{f\big(\gamma(t)\big) - f\big(\gamma(0)\big)}{t} = L(\xi) + \underbrace{\frac{\varphi(t\xi)}{t}}_{\to 0} \xrightarrow{t\to 0} L(\xi),$$

also

$$\frac{\mathrm{d}}{\mathrm{d}t}(f\circ\gamma)(t)|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} f(x+t\xi)|_{t=0} = \lim_{t\to 0} \frac{f\big(\gamma(t)\big) - f\big(\gamma(0)\big)}{t} = L(\xi) = L\xi,$$

wobei $L = Df(x)$.

Definition 7.10. Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$.

(1) Die *Richtungsableitung* $D_\xi f(x)$ von f in Richtung $\xi \in \mathbb{R}^n$ an der Stelle $x \in U$ ist die Ableitung der Kurve $\gamma(t) = f(x+t\xi)$ in $t = 0$.

(2) Wir schreiben

$$\frac{\partial f}{\partial x_i} := D_i f := D_{e_i} f$$

für die Ableitung in Richtung des i-ten Standardbasisvektors e_i.

(3) Ist $\frac{\partial f}{\partial x_i}$ in Richtung e_j differenzierbar, so schreiben wir

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = D_j D_i f$$

und analog für höhere Ableitungen.

Bemerkung 7.11. (1) Bezüglich der Standardbasen $(e_i)_{i=1}^n$ und $(e_i)_{i=1}^m$ von $\mathbb{R}^n$ und $\mathbb{R}^m$ ist $Df(x)$ also als Matrix der partiellen Ableitungen gegeben:

$$Df(x) = \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{i=1,\dots,m\\ j=1,\dots,n}} \in \mathbb{R}^{m\times n} \quad \text{für} \quad f = (f_1, \dots, f_m) \quad \text{mit} \quad f_i : U \to \mathbb{R}.$$

Für

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} xy \\ x+y \end{pmatrix}$$

mit $f_1(x,y) = xy$ und $f_2(x,y) = x+y$ erhält man beispielsweise

$$Df(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} y & x \\ 1 & 1 \end{pmatrix}.$$

(2) Die Existenz der totalen Ableitung ist stärker als die Existenz aller partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}$: Für das f aus Beispiel 7.2 existieren $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ überall (insbesondere auch im Punkt $(0,0)$!), aber da f in $(0,0)$ nicht stetig ist, kann f dort auch nicht differenzierbar sein.

Satz 7.12. Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ (hier ist also $m = 1$) eine Funktion, für die die partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ für alle $i = 1, \dots, n$ in U existieren und stetig sind. Dann ist f in U differenzierbar.

Beweis. Sei $x \in U$ und wähle $\varepsilon > 0$ mit $\mathrm{B}(x, \varepsilon) \subseteq U$. Weiterhin sei $\xi \in \mathbb{R}^n$ mit $\|\xi\| < \varepsilon$. Wir wollen zeigen, dass $f(x+\xi) = f(x) + L\xi + \varphi(\xi)$ für geeignete L und φ gilt.

Schreibe $\xi = \sum_{i=1}^n \xi_i e_i$ in der Standardbasis mit $\xi_i \in \mathbb{R}$ sowie

$$z^{(k)} := x + \sum_{i=1}^k \xi_i e_i \quad \text{für alle} \quad k = 0, \dots, n,$$

dann gilt $z^{(0)} = x$, $z^{(n)} = x + \xi$ und $z^{(k)} - z^{(k-1)} = \xi_k e_k$ für alle $k = 1, \dots, n$, also unterscheiden sich $z^{(k)}$ und $z^{(k-1)}$ nur in Komponente k um den Wert ξ_k. Betrachtet man also die Funktion

$$f_k : [0, \xi_k] \to \mathbb{R}, \quad t \mapsto f(z^{(k-1)} + te_k),$$

so ist diese nach Voraussetzung differenzierbar. Laut Mittelwertsatz gibt es ein $\theta_k \in [0,1]$ mit

$$f_k(\xi_k) - f_k(0) = f_k'(\theta_k \xi_k) \cdot (\xi_k - 0) = f_k'(\theta_k \xi_k) \cdot \xi_k,$$

also

$$f(z^{(k)}) = f_k(\xi_k) = f_k(0) + \xi_k \cdot f_k'(\theta_k \xi_k) = f(z^{(k-1)}) + \xi_k \cdot \frac{\partial f}{\partial x_k}(y^{(k)})$$

für $y^{(k)} := z^{(k-1)} + \theta_k \xi_k e_k$. Somit gilt

$$f(x+\xi) = f(z^{(n)}) = f(z^{(n-1)}) + \xi_n \cdot \frac{\partial f}{\partial x_n}(y^{(n)}) = \ldots = \underbrace{f(z^{(0)})}_{f(x)} + \sum_{k=1}^n \xi_k \cdot \frac{\partial f}{\partial x_k}(y^{(k)}).$$

Setzt man

$$A = \left(\frac{\partial f}{\partial x_i}(x)\right)_{i=1}^n = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right) \in \mathbb{R}^{1\times n},$$

so gilt

$$f(x+\xi) = f(x) + A\xi + \underbrace{\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(y^{(i)}) - \frac{\partial f}{\partial x_i}(x) \right) \xi_i}_{=:\varphi(\xi)}$$

Nun ist A offensichtlich linear und für $\xi \to 0$ gilt $\xi_i \to 0$ und somit

$$y^{(i)} = z^{(i)} + \theta_i \xi_i e_i = x + \xi_1 e_1 + \ldots + \xi_{i-1} e_{i-1} + \theta_i \xi_i e_i \to x \quad \text{für alle} \quad i = 1, \ldots, n.$$

Da die partiellen Ableitungen als stetig vorausgesetzt wurden, folgt

$$\frac{\partial f}{\partial x_i}(y^{(i)}) \to \frac{\partial f}{\partial x_i}(x) \quad \text{für alle} \quad i = 1, \ldots, n,$$

also

$$\frac{|\varphi(\xi)|}{\|\xi\|} = \frac{\left| \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(y^{(i)}) - \frac{\partial f}{\partial x_i}(x) \right) \xi_i \right|}{\|\xi\|} \leq \sum_{i=1}^{n} \underbrace{\left| \frac{\partial f}{\partial x_i}(y^{(i)}) - \frac{\partial f}{\partial x_i}(x) \right|}_{\to 0} \underbrace{\frac{|\xi_i|}{\|\xi\|}}_{\leq 1} \to 0$$

und somit $\frac{\varphi(\xi)}{\|\xi\|} \to 0$ für $\xi \to 0$ mit $\xi \neq 0$. □

Korollar 7.13. Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ stetig partiell differenzierbar. Dann ist f stetig.

Beweis. Wegen Satz 7.12 folgt aus stetiger partieller Differenzierbarkeit schon totale Differenzierbarkeit, laut Bemerkung 7.9 folgt daraus Stetigkeit. □

Definition 7.14. Sei $U \subseteq \mathbb{R}^n$ offen. Eine Abbildung $f : U \to \mathbb{R}$ heißt *k-mal stetig partiell differenzierbar*, falls alle partiellen Ableitungen der Form

$$D_{i_1} \ldots D_{i_k} f \quad \text{für} \quad i_1, \ldots, i_k \in \{1, \ldots, n\}$$

existieren und stetig sind.

Satz 7.15 (Schwarz)**.** Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Dann gilt

$$D_i D_j f = D_j D_i f \quad \text{für alle} \quad i, j = 1, \ldots, n.$$

Beispiel 7.16. Für $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = x\mathrm{e}^{xy}$ gilt

$$\frac{\partial f}{\partial x} = \mathrm{e}^{xy} + xy\mathrm{e}^{xy} \quad \text{und} \quad \frac{\partial f}{\partial y} = x^2 \mathrm{e}^{xy}.$$

Die gemischten zweiten partiellen Ableitungen sind

$$\frac{\partial^2 f}{\partial y \, \partial x} = \frac{\partial}{\partial y} \frac{\partial f}{\partial x} = \frac{\partial}{\partial y} (\mathrm{e}^{xy} + xy\mathrm{e}^{xy}) = x\mathrm{e}^{xy} + x\mathrm{e}^{xy} + x^2 y\mathrm{e}^{xy}$$

und

$$\frac{\partial^2 f}{\partial x \, \partial y} = \frac{\partial}{\partial x} \frac{\partial f}{\partial y} = \frac{\partial}{\partial x} x^2 \mathrm{e}^{xy} = 2x\mathrm{e}^{xy} + x^2 y\mathrm{e}^{xy},$$

also

$$\frac{\partial^2 f}{\partial y \, \partial x} = \frac{\partial^2 f}{\partial x \, \partial y}.$$

Beweis von Satz 7.15. Sei $x \in U$. Wir betrachten den Ausdruck

$$\varphi(s,t) := \big(f(x+te_i+se_j) - f(x+te_i)\big) - \big(f(x+se_j) - f(x)\big) \quad \text{für} \quad s,t>0.$$

Setze $g(t) := f(x+te_i+se_j) - f(x+te_i)$, dann liefert der Mittelwertsatz angewendet auf g ein $\theta \in (0,t)$ mit

$$\varphi(s,t) = g(t) - g(0) = t \cdot g'(\theta).$$

Weiterhin gilt

$$g'(\theta) = \underbrace{\frac{\partial f}{\partial x_i}(x+\theta e_i + se_j)}_{=:h(s)} - \frac{\partial f}{\partial x_i}(x+\theta e_i) = h(s) - h(0) = s \cdot h'(\tilde{\theta})$$

für ein $\tilde{\theta} \in (0,s)$ laut Mittelwertsatz angewendet auf h. Wegen

$$h'(\tilde{\theta}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x+\theta e_i + \tilde{\theta} e_j)$$

gilt nun

$$\varphi(s,t) = t \cdot g'(\theta) = t \cdot s \cdot h'(\tilde{\theta}) = t \cdot s \cdot \frac{\partial^2 f}{\partial x_j \partial x_i}(x+\theta e_i + \tilde{\theta} e_j).$$

Vollzieht man die gleiche Rechnung mit vertauschten Rollen von s und t, so erhält man

$$\varphi(s,t) = s \cdot t \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}(x+\eta e_j + \tilde{\eta} e_i)$$

für $\eta \in (0,s)$ und $\tilde{\eta} \in (0,t)$, also

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x+\theta e_i + \tilde{\theta} e_j) = \frac{\varphi(s,t)}{st} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x+\eta e_j + \tilde{\eta} e_i).$$

Für $s,t \to 0$ gilt $\theta, \tilde{\theta}, \eta, \tilde{\eta} \to 0$, also

$$x+\theta e_i + \tilde{\theta} e_j \xrightarrow{s,t\to 0} x \quad \text{und} \quad x+\eta e_j + \tilde{\eta} e_i \xrightarrow{s,t\to 0} x.$$

Da die zweiten partiellen Ableitungen nach Voraussetzung stetig sind, folgt

$$\begin{aligned}\frac{\partial^2 f}{\partial x_j \partial x_i}(x) &= \lim_{s,t\to 0} \frac{\partial^2 f}{\partial x_j \partial x_i}(x+\theta e_i + \tilde{\theta} e_j) \\ &= \lim_{s,t\to 0} \frac{\partial^2 f}{\partial x_i \partial x_j}(x+\eta e_j + \tilde{\eta} e_i) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x).\end{aligned}$$

□

Bemerkung 7.17. Sei $U \subseteq \mathbb{R}^n$ offen. Eine Funktion $f : U \to \mathbb{R}^m$ ist genau dann stetig beziehungsweise differenzierbar, wenn alle Komponentenfunktionen $f_i : U \to \mathbb{R}$ stetig beziehungsweise differenzierbar sind (für $f = (f_1, \ldots, f_m)$). Daher gelten die Aussagen Satz 7.12, Korollar 7.13 und Satz 7.15 auch für Funktionen $f : U \to \mathbb{R}^m$.

Definition 7.18. Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$.

(1) Falls f differenzierbar ist, bezeichnen wir

$$Df(x) = \left(\frac{\partial f_i}{\partial x_j}(x)\right)_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$$

als *Funktional-*, *Jacobi-* oder *Differentialmatrix*.

(2) Im Fall $m = 1$, also $f : U \to \mathbb{R}$, ist die Funktionalmatrix eine $(1 \times n)$-Matrix. Sie heißt dann *Gradient* von f und wird mit $\operatorname{grad} f$ bezeichnet:

$$\operatorname{grad} f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right).$$

Bemerkung 7.19. (1) Für $f : U \to \mathbb{R}$ definiert $\operatorname{grad} f$ ein *Vektorfeld*: Jedes $x \in U \subseteq \mathbb{R}^n$ wird wieder auf einen Vektor in $\mathbb{R}^n$ abgebildet.

(2) Für Richtungsableitungen gilt in dem Fall

$$D_\xi f = Df \cdot \xi = \underbrace{\operatorname{grad} f}_{\in \mathbb{R}^{1\times n}} \cdot \underbrace{\xi}_{\in \mathbb{R}^{n\times 1}} = \sum_{i=1}^{n} (\operatorname{grad} f)_i \cdot \xi_i = \langle \operatorname{grad} f, \xi \rangle \in \mathbb{R}.$$

Insbesondere gilt mit Cauchy-Schwarz

$$\|D_\xi f\| = |\langle \operatorname{grad} f, \xi\rangle| \le \|\operatorname{grad} f\| \cdot \|\xi\|$$

mit Gleichheit (also maximaler Ableitung bei fester Länge von ξ) genau dann, wenn ξ in Richtung von $\operatorname{grad} f$ zeigt.

Satz 7.20 (Kettenregel)**.** Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen. Wir betrachten Abbildungen $g : U \to V$ und $f : V \to \mathbb{R}^k$. Sei g in $x \in U$ differenzierbar und f in $g(x) \in V$ differenzierbar. Dann ist $f \circ g : U \to \mathbb{R}^k$ in x differenzierbar und es gilt

$$D(f \circ g)(x) = Df\big(g(x)\big) \cdot Dg(x).$$[24]

Beweis. Da g in x differenzierbar ist, gilt

$$g(x + \xi) = g(x) + Dg(x) \cdot \xi + \varphi(\xi)$$

für $\|\xi\|$ hinreichend klein und $\frac{\varphi(\xi)}{\|\xi\|} \xrightarrow{\xi \to 0} 0$. Da f in $g(x)$ differenzierbar ist, gilt

$$\begin{aligned}
&(f \circ g)(x + \xi)\\
&= f\big(g(x + \xi)\big)\\
&= f\big(g(x) + Dg(x) \cdot \xi + \varphi(\xi)\big)\\
&= f\big(g(x)\big) + Df\big(g(x)\big) \cdot \big(Dg(x) \cdot \xi + \varphi(\xi)\big) + \psi\big(Dg(x) \cdot \xi + \varphi(\xi)\big)\\
&= f\big(g(x)\big) + Df\big(g(x)\big) \cdot Dg(x) \cdot \xi + Df\big(g(x)\big) \cdot \varphi(\xi) + \psi\big(Dg(x) \cdot \xi + \varphi(\xi)\big)
\end{aligned}$$

[24] Auf der rechten Seite steht hier die Komposition von linearen Abbildungen, welche wir mit der Multiplikation ihrer darstellenden Matrizen, also der Multiplikation der Jacobi-Matrizen, identifzieren können.

für $\frac{\psi(\xi)}{\|\xi\|} \xrightarrow{\xi\to 0} 0$. Für die Restterme gilt einerseits

$$\frac{Df\big(g(x)\big)\cdot\varphi(\xi)}{\|\xi\|} = Df\big(g(x)\big)\Bigg(\underbrace{\frac{\varphi(x)}{\|\xi\|}}_{\to 0}\Bigg) \xrightarrow{\xi\to 0} 0,$$

da $Df\big(g(x)\big)$ als Abbildung linear und stetig ist. Andererseits muss

$$\frac{\psi\big(Dg(x)\cdot\xi+\varphi(\xi)\big)}{\|\xi\|} = \frac{\psi\big(Dg(x)\cdot\xi+\varphi(\xi)\big)}{\|Dg(x)\cdot\xi+\varphi(\xi)\|}\cdot\frac{\|Dg(x)\cdot\xi+\varphi(\xi)\|}{\|\xi\|}$$

für $\xi\to 0$ auch gegen Null gehen: Wegen $Dg(x)\cdot\xi+\varphi(\xi) \xrightarrow{\xi\to 0} 0$ gilt

$$\frac{\psi\big(Dg(x)\cdot\xi+\varphi(\xi)\big)}{\|Dg(x)\cdot\xi+\varphi(\xi)\|} \xrightarrow{\xi\to 0} 0,$$

es reicht also zu zeigen, dass $\frac{\|Dg(x)\cdot\xi+\varphi(\xi)\|}{\|\xi\|}$ beschränkt ist: Es gilt

$$\|Dg(x)\cdot\xi+\varphi(\xi)\| \le \|Dg(x)\cdot\xi\| + \|\varphi(\xi)\| \le \|Dg(x)\|\cdot\|\xi\| + \|\varphi(\xi)\|,$$

also

$$\frac{\|Dg(x)\cdot\xi+\varphi(\xi)\|}{\|\xi\|} \le \|Dg(x)\| + \frac{\|\varphi(\xi)\|}{\|\xi\|}.$$

Hierbei ist $\|Dg(x)\|$ konstant und es gilt $\frac{\|\varphi(\xi)\|}{\|\xi\|} \xrightarrow{\xi\to 0} 0$, also ist die rechte Seite und damit auch die linke Seite beschränkt.

Somit ist also $f\circ g$ in x differenzierbar mit $D(f\circ g)(x) = Df\big(g(x)\big)\cdot Dg(x)$. □

Beispiel 7.21. (1) Seien $\varphi:\mathbb{R}\to\mathbb{R}^n$ und $f:\mathbb{R}^n\to\mathbb{R}$ differenzierbar, dann gilt

$$Df = \left(\frac{\partial f}{\partial x_1},\dots,\frac{\partial f}{\partial x_n}\right),$$

also $Df(x)\in\mathbb{R}^{1\times n}$ für alle $x\in\mathbb{R}^n$, und

$$D\varphi = \begin{pmatrix}\varphi_1'\\ \vdots\\ \varphi_n'\end{pmatrix} = \begin{pmatrix}\frac{\partial\varphi_1}{\partial t}\\ \vdots\\ \frac{\partial\varphi_n}{\partial t}\end{pmatrix},$$

also $D\varphi(t)\in\mathbb{R}^{n\times 1}$ für alle $t\in\mathbb{R}$. Somit gilt

$$(f\circ\varphi)'(t) = Df\big(\varphi(t)\big)\cdot D\varphi(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}\big(\varphi(t)\big)\cdot\frac{\partial\varphi_i}{\partial t}(t)\in\mathbb{R} \quad \text{für alle} \quad t\in\mathbb{R}.$$

Häufig schreibt man in dieser Situation kurz

$$f = f(x_1,\dots,x_n) \quad \text{mit} \quad x_i := \varphi_i(t),$$

dann gilt

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}\cdot\frac{\partial x_i}{\partial t}.$$

(2) Für $f(x,y) = x^2 + y^2$ und $\varphi(t) = \big(\sin(t), \cos(t)\big)$ gilt einerseits

$$(f \circ \varphi)(t) = f\big(\sin(t), \cos(t)\big) = \sin^2(t) + \cos^2(t) = 1,$$

also $(f \circ \varphi)' = 0$. In der anderen Schreibweise (mit $x = \sin(t)$ und $y = \cos(t)$) gilt ebenfalls

$$\begin{aligned} f'(t) = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial t} &= 2x \cdot \cos(t) + 2y \cdot \big(-\sin(t)\big) \\ &= 2\sin(t) \cdot \cos(t) + 2\cos(t) \cdot \big(-\sin(t)\big) \\ &= 0. \end{aligned}$$

(3) In der allgemeineren Situation mit

$$f : \mathbb{R}^m \to \mathbb{R}^k, \quad g : \mathbb{R}^n \to \mathbb{R}^m \quad \text{und} \quad h = f \circ g : \mathbb{R}^n \to \mathbb{R}^k$$

schreiben wir

$$f = f(y_1, \ldots, y_m) \quad \text{und} \quad g = g(x_1, \ldots, x_n) = \big(y_1(x), \ldots, y_m(x)\big).$$

Für

$$h(x_1, \ldots, x_n) = f\big(y_1(x), \ldots, y_m(x)\big)$$

gilt dann

$$\frac{\partial h_i}{\partial x_j} = \sum_{s=1}^{m} \frac{\partial f_i}{\partial y_s} \cdot \frac{\partial y_s}{\partial x_j}.$$

8. Taylorformel und lokale Extrema

Konvention. Im Folgenden bezeichne $\langle\cdot,\cdot\rangle$ stets das Standardskalarprodukt auf $\mathbb{R}^n$ und $\|\cdot\|$ die davon induzierte Euklidische Norm $\|\cdot\|_2$.

Erinnerung 8.1. Sei $f:\mathbb{R}\to\mathbb{R}$ $(k+1)$-mal stetig differenzierbar. Dann gilt die Taylorformel

$$\begin{aligned} f(x+t) &= f(x)+f'(x)\cdot t+\frac{1}{2}f''(x)\cdot t^2+\ldots+\frac{1}{n!}f^{(n)}(x)\cdot t^n+R_{n+1}(t) \\ &= \sum_{k=0}^{n}\frac{f^{(k)}(x)}{k!}\cdot t^k+R_{n+1}(t), \end{aligned}$$

wobei für das Restglied $R_{n+1}(t)$ gilt:

$$R_{n+1}(t)=\frac{1}{n!}\int_0^t (t-s)^n f^{(n+1)}(x+s)\,\mathrm{d}s \qquad \text{(Integraldarstellung)}$$

oder alternativ

$$R_{n+1}(t)=\frac{f^{(n+1)}(\theta)}{(n+1)!}\cdot t^{n+1}\quad \text{für ein}\quad \theta\in[x,x+t]. \qquad \text{(Lagrange-Form)}$$

Definition 8.2. Ein *Multiindex* ist ein Tupel $\alpha=(\alpha_1,\ldots,\alpha_n)\in\mathbb{N}_0^n$. Wir setzen

$$|\alpha|:=\alpha_1+\ldots+\alpha_n\in\mathbb{N}_0 \quad\text{und}\quad \alpha!:=\alpha_1!\cdot\ldots\cdot\alpha_n!\in\mathbb{N}.$$

Für $\xi=(\xi_1,\ldots,\xi_n)\in\mathbb{R}^n$ schreiben wir

$$\xi^\alpha:=\xi_1^{\alpha_1}\cdot\ldots\cdot\xi_n^{\alpha_n}.$$

Ist $U\subseteq\mathbb{R}^n$ offen und $f:U\to\mathbb{R}$ $|\alpha|$-mal stetig differenzierbar, so setzen wir

$$D^\alpha f(x):=D_1^{\alpha_1}D_2^{\alpha_2}\ldots D_n^{\alpha_n}f(x).$$

Satz 8.3. Sei $U\subseteq\mathbb{R}^n$ offen und $f:U\to\mathbb{R}$ k-mal stetig differenzierbar. Weiter seien $x\in U$ und $\xi\in\mathbb{R}^n$ so, dass $x+t\xi\in U$ für alle $t\in[0,1]$. Dann ist die Funktion

$$g:[0,1]\to\mathbb{R},\quad t\mapsto g(t):=f(x+t\xi)$$

k-mal stetig differenzierbar und es gilt

$$g^{(k)}(t)=\frac{\mathrm{d}^k g}{\mathrm{d}t^k}(t)=\sum_{\substack{\alpha\in\mathbb{N}_0^n\\|\alpha|=k}}\frac{k!}{\alpha!}D^\alpha f(x+t\xi)\xi^\alpha.$$

Beweis. Für $k=1$ gilt: $g(t)=f\big(\varphi(t)\big)$ mit $\varphi(t)=x+t\xi$, also nach Kettenregel

$$g'(t)=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\big(\varphi(t)\big)\cdot\underbrace{\frac{\partial\varphi_i}{\partial t}}_{=\xi_i}=\sum_{i=1}^{n}D_i f(x+t\xi)\cdot\xi_i.$$

Für höhere k erhalten wir

$$g''(t) = \sum_{i=1}^{n} \underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\big(D_i f(x+t\xi)\big)}_{=\sum_{j=1}^{n} D_j D_i f(x+t\xi)\cdot\xi_j} \cdot \xi_i = \sum_{i,j=1}^{n} D_j D_i f(x+t\xi)\cdot\xi_j\cdot\xi_i$$

und durch Iteration

$$g^{(k)}(t) = \sum_{i_1,\ldots,i_k=1}^{n} D_{i_1}\ldots D_{i_k} f(x+t\xi)\cdot\xi_{i_1}\cdot\ldots\cdot\xi_{i_k}.$$

Nach dem Satz von Schwarz 7.15 vertauschen die partiellen Ableitungen: Jeder Term $D_{i_1}\ldots D_{i_k}f$ ist von der Form $D^\alpha f = D_1^{\alpha_1}\ldots D_n^{\alpha_n} f$ für ein $\alpha = (\alpha_1,\ldots,\alpha_n)\in\mathbb{N}_0^n$. Es gibt genau $\frac{k!}{\alpha!}$ Tupel $(i_1,\ldots,i_k)$, die zu dem gleichen α führen, [25] also gilt

$$g^{(k)}(t) = \sum_{i_1,\ldots,i_k=1}^{n} D_{i_1}\ldots D_{i_k} f(x+t\xi)\cdot\xi_{i_1}\cdot\ldots\cdot\xi_{i_k} = \sum_{\substack{\alpha\in\mathbb{N}_0^n\\|\alpha|=k}} \frac{k!}{\alpha!} D^\alpha f(x+t\xi)\xi^\alpha. \qquad \square$$

Für $n=k=2$ hat man beispielsweise

$$\begin{aligned}
g''(t) &= \sum_{i,j=1}^{2} D_i D_j f(x+t\xi)\cdot\xi_i\cdot\xi_j \\
&= \underbrace{D_1^2 f(x+t\xi)\cdot\xi_1^2}_{(i,j)=(1,1)} + \underbrace{D_1D_2 f(x+t\xi)\cdot\xi_1\cdot\xi_2}_{(i,j)=(1,2)} + \underbrace{D_2D_1 f(x+t\xi)\cdot\xi_2\cdot\xi_1}_{(i,j)=(2,1)} \\
&\quad + \underbrace{D_2^2 f(x+t\xi)\cdot\xi_2^2}_{(i,j)=(2,2)} \\
&= \underbrace{D_1^2 f(x+t\xi)\cdot\xi_1^2}_{\alpha=(2,0)} + \underbrace{2D_1D_2 f(x+t\xi)\cdot\xi_1\cdot\xi_2}_{\alpha=(1,1)} + \underbrace{D_2^2 f(x+t\xi)\cdot\xi_2^2}_{\alpha=(0,2)} \\
&= \sum_{\substack{\alpha\in\mathbb{N}_0^2\\|\alpha|=2}} \frac{2!}{\alpha_1!\alpha_2!} D_1^{\alpha_1} D_2^{\alpha_2} f(x+t\xi)\cdot\xi_1^{\alpha_1}\cdot\xi_2^{\alpha_2}.
\end{aligned}$$

Satz 8.4 (Taylorformel)**.** Sei $U\subseteq\mathbb{R}^n$ offen und $f: U\to\mathbb{R}$ $(k+1)$-mal stetig differenzierbar. Weiter seien $x\in U$ und $\xi\in\mathbb{R}^n$ so, dass $x+t\xi\in U$ für alle $t\in[0,1]$. Dann existiert ein $\theta\in[0,1]$, so dass

$$f(x+\xi) = f(x) + \sum_{\substack{\alpha\in\mathbb{N}_0^n\\1\le|\alpha|\le k}} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha + \sum_{\substack{\alpha\in\mathbb{N}_0^n\\|\alpha|=k+1}} \frac{D^\alpha f(x+\theta\xi)}{\alpha!}\xi^\alpha.$$

Beweis. Die Funktion

$$g:[0,1]\to\mathbb{R},\quad t\mapsto g(t):=f(x+t\xi)$$

[25] Klassische Kombinatorik: Um $\alpha=(\alpha_1,\ldots,\alpha_n)$ zu erhalten, ziehen wir aus einer Urne mit α_1-vielen Kugeln mit der Zahl 1, α_2-vielen Kugel mit der Zahl 2, … und α_n-vielen Kugeln mit der Zahl n. Es gibt $k!$ Reihenfolgen, in denen wir alle $k=|\alpha|$ Kugeln aus der Urne ziehen können. Da die α_i-vielen Kugeln mit der Zahl i nicht unterscheidbar sind, müssen wir für jedes i noch durch die Anzahl der Permutationen dieser α_i-vielen Kugeln teilen, also durch $\alpha_i!$.

ist $(k+1)$-mal stetig differenzierbar. Mit der eindimensionalen Taylorformel existiert ein $\theta \in [0,1]$, so dass

$$f(x+\xi) = g(1) = g(0+1) = \sum_{m=0}^{k} \frac{g^{(m)}(0)}{m!} + \frac{g^{(k+1)}(\theta)}{(k+1)!}.$$

Die Aussage folgt durch Einsetzen der Formeln für die Ableitungen von g aus Satz 8.3. □

Definition 8.5. Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$.

(1) Die Funktion f hat in $x_0 \in U$ ein *lokales Maximum* (bzw. *lokales Minimum*), falls es ein $\varepsilon > 0$ mit $\mathrm{B}(x_0, \varepsilon) \subseteq U$ gibt, so dass $f(y) \leq f(x_0)$ (bzw. $f(y) \geq f(x_0)$) für alle $y \in \mathrm{B}(x_0, \varepsilon)$ gilt.

(2) Die Funktion f hat in $x_0 \in U$ ein *isoliertes Maximum* (bzw. *isoliertes Minimum*), falls es ein $\varepsilon > 0$ mit $\mathrm{B}(x_0, \varepsilon) \subseteq U$ gibt, so dass $f(y) < f(x_0)$ (bzw. $f(y) > f(x_0)$) für alle $y \in \mathrm{B}(x_0, \varepsilon)$ mit $y \neq x_0$ gilt.

Satz 8.6. Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ differenzierbar. Ist $x_0 \in U$ ein lokales Extremum von f, so gilt $\operatorname{grad} f(x_0) = 0$.[26]

Beweis. Sei $\xi \in \mathbb{R}^n$ und $\varepsilon > 0$ so, dass $x_0 + t\xi \in U$ für alle $t \in (-\varepsilon, \varepsilon)$. Dann gilt $D_\xi f(x_0) = g'(0)$ für

$$g : (-\varepsilon, \varepsilon) \to \mathbb{R}, \quad t \mapsto g(t) := f(x_0 + t\xi).$$

Nun hat g in x_0 ein lokales Extremum. Laut Satz 15.2 aus Analysis I folgt $g'(0) = 0$. □

Bemerkung 8.7. Gilt $f'(x_0) = 0$ im Fall $n = 1$, so entscheidet $f''(x_0)$ darüber, ob in x_0 ein Maximum ($f''(x_0) < 0$) oder ein Minimum ($f''(x_0) > 0$) vorliegt. Im allgemeinen Fall gilt

$$\begin{aligned} f(x+\xi) &= f(x) + \underbrace{\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \cdot \xi_i}_{=\langle \operatorname{grad} f(x), \xi\rangle} + \underbrace{\sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha| = 2}} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha}_{=\frac{1}{2}\sum_{i,j=1}^{n} D_i D_j f(x) \cdot \xi_i \xi_j} + \dots \\ &= f(x) + \langle \operatorname{grad} f(x), \xi\rangle + \frac{1}{2}\langle A\xi, \xi\rangle + \dots \end{aligned}$$

für

$$A := \big(D_i D_j f(x_0)\big)_{i,j=1}^{n}.$$

Definition 8.8. Sei $U \subseteq \mathbb{R}^n$ offen, $x \in U$ und $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann heißt

$$(\operatorname{Hess} f)(x) := \big(D_i D_j f(x)\big)_{i,j=1}^{n} \in \mathbb{R}^{n \times n}$$

die *Hesse-Matrix* von f im Punkt x.

[26] Anders ausgedrückt: Im Punkt x_0 verschwinden alle Richtungsableitungen, also $D_\xi f(x_0) = 0$ für alle $\xi \in \mathbb{R}^n$.

Satz 8.9. Sei $U \subseteq \mathbb{R}^n$ offen, $x \in U$ und $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. Für alle $\xi \in \mathbb{R}^n$ mit $\|\xi\|$ hinreichend klein gilt dann

$$f(x+\xi) = f(x) + \langle \operatorname{grad} f(x), \xi \rangle + \frac{1}{2}\langle (\operatorname{Hess} f)(x) \cdot \xi, \xi \rangle + \varphi(\xi),$$

wobei für die Funktion φ gilt

$$\lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \frac{|\varphi(\xi)|}{\|\xi\|^2} = 0.$$

Beweis. Nach der Taylorformel 8.4 für $k = 2$ gibt es ein $\theta \in [0, 1]$ mit

$$f(x+\xi) = f(x) + \langle \operatorname{grad} f(x), \xi \rangle + \frac{1}{2}\langle (\operatorname{Hess} f)(x + \theta\xi) \cdot \xi, \xi \rangle,$$

somit gilt die Behauptung, falls

$$\varphi(\xi) := \frac{1}{2}\Big\langle \big((\operatorname{Hess} f)(x) - (\operatorname{Hess} f)(x+\theta\xi)\big) \cdot \xi, \xi \Big\rangle$$

die Zusatzbedingung erfüllt. Mit Cauchy-Schwarz gilt

$$\begin{aligned} |\varphi(x)| &= \frac{1}{2}\Big|\Big\langle \big((\operatorname{Hess} f)(x) - (\operatorname{Hess} f)(x+\theta\xi)\big) \cdot \xi, \xi \Big\rangle\Big| \\ &\leq \frac{1}{2} \cdot \Big\|\big((\operatorname{Hess} f)(x) - (\operatorname{Hess} f)(x+\theta\xi)\big) \cdot \xi\Big\| \cdot \|\xi\| \\ &\leq \frac{1}{2} \cdot \Big\|\big((\operatorname{Hess} f)(x) - (\operatorname{Hess} f)(x+\theta\xi)\big)\Big\| \cdot \|\xi\| \cdot \|\xi\|. \end{aligned}$$

Da f zweimal stetig differenzierbar ist, gilt

$$\Big\|\big((\operatorname{Hess} f)(x) - (\operatorname{Hess} f)(x+\theta\xi)\big)\Big\| \xrightarrow{\xi \to 0} 0,$$

also

$$\frac{|\varphi(\xi)|}{\|\xi\|^2} \xrightarrow{\xi \to 0} 0. \qquad \square$$

Bemerkung 8.10. Wegen $D_i D_j f = D_j D_i f$ ist $\operatorname{Hess} f$ eine symmetrische Matrix, kann also diagonalisiert werden und hat n reelle Eigenwerte $\lambda_1, \ldots, \lambda_n$ (mit Multiplizitäten gezählt).

- Wir nennen $\operatorname{Hess} f$ *positiv definit*, falls

  $$\forall \xi \in \mathbb{R}^n \setminus \{0\} : \quad \langle (\operatorname{Hess} f)\xi, \xi \rangle > 0.$$

  Dies ist äquivalent dazu, dass alle Eigenwerte von $\operatorname{Hess} f$ positiv sind: $\lambda_i > 0$ für alle $i = 1, \ldots, n$.

- Wir nennen $\operatorname{Hess} f$ *negativ definit*, falls

  $$\forall \xi \in \mathbb{R}^n \setminus \{0\} : \quad \langle (\operatorname{Hess} f)\xi, \xi \rangle < 0.$$

  Dies ist äquivalent dazu, dass alle Eigenwerte von $\operatorname{Hess} f$ negativ sind: $\lambda_i < 0$ für alle $i = 1, \ldots, n$.

- Wir nennen $\operatorname{Hess} f$ *indefinit*, falls

$$\exists \xi, \eta \in \mathbb{R}^n : \quad \langle (\operatorname{Hess} f)\xi, \xi\rangle > 0 \quad \text{und} \quad \langle (\operatorname{Hess} f)\eta, \eta\rangle < 0.$$

  Dies ist äquivalent dazu, dass $\operatorname{Hess} f$ sowohl positive als auch negative Eigenwerte hat: Es gibt $i, j \in \{1, \dots, n\}$ mit $\lambda_i > 0$ und $\lambda_j < 0$.[27]

Satz 8.11. Sei $U \subseteq \mathbb{R}^n$ offen, $x \in U$ und $f : U \to \mathbb{R}$ zweimal stetig differenzierbar mit $\operatorname{grad} f(x) = 0$.

(i) Falls $(\operatorname{Hess} f)(x)$ positiv definit ist, hat f ein isoliertes Minimum in x.

(ii) Falls $(\operatorname{Hess} f)(x)$ negativ definit ist, hat f ein isoliertes Maximum in x.

(iii) Falls $(\operatorname{Hess} f)(x)$ indefinit ist, hat f kein isoliertes Extremum in x.

(Falls $(\operatorname{Hess} f)(x)$ Null als Eigenwert hat, ist keine allgemeine Aussage möglich.)

Beweis. Sei $A := (\operatorname{Hess} f)(x)$.

(i) Da $\{\eta \in \mathbb{R}^n : \|\eta\| = 1\}$ kompakt und $\mathbb{R}^n \to \mathbb{R}^n$, $\eta \mapsto \langle A\eta, \eta\rangle$ stetig ist, existiert

$$M := \min\{\langle A\eta, \eta\rangle : \|\eta\| = 1\} > 0.$$

Für alle $\xi \in \mathbb{R}^n$ mit $\xi \neq 0$ und $\|\xi\|$ hinreichend klein gilt

$$\langle A\xi, \xi\rangle = \|\xi\|^2 \cdot \left\langle A\frac{\xi}{\|\xi\|}, \frac{\xi}{\|\xi\|}\right\rangle \geq M \cdot \|\xi\|^2,$$

mit Satz 8.9 also

$$f(x+\xi) = f(x) + \underbrace{\langle \operatorname{grad} f(x), \xi\rangle}_{=0} + \frac{1}{2}\underbrace{\langle A\xi, \xi\rangle}_{\geq M\cdot\|\xi\|^2} + \varphi(\xi) \geq f(x) + \frac{1}{2}M\|\xi\|^2 + \varphi(\xi),$$

wobei $\frac{|\varphi(\xi)|}{\|\xi\|^2} \xrightarrow{\xi\to 0} 0$. Somit gibt es $\varepsilon > 0$, so dass

$$|\varphi(\xi)| < \frac{1}{2}M\|\xi\|^2 \quad \text{für} \quad \|\xi\| < \varepsilon$$

und somit

$$\varphi(\xi) > -\frac{1}{2}M\|\xi\|^2 \quad \text{für} \quad \|\xi\| < \varepsilon.$$

Folglich gilt

$$f(x+\xi) \geq f(x) + \frac{1}{2}M\|\xi\|^2 + \varphi(\xi) > f(x) + \frac{1}{2}M\|\xi\|^2 - \frac{1}{2}M\|\xi\|^2 = f(x),$$

also hat f in x ein isoliertes Minimum.

(ii) Ist $(\operatorname{Hess} f)(x)$ negativ definit, so ist $\big(\operatorname{Hess}(-f)\big)(x) = -(\operatorname{Hess} f)(x)$ positiv definit. Also hat $-f$ ein isoliertes Minimum in x und somit f ein isoliertes Maximum in x.

[27] Dies deckt nur die Fälle ab, in denen $\operatorname{Hess} f$ nie Null als Eigenwert hat. Falls Null als Eigenwert auftritt, spricht man manchmal von positiv oder negativ semidefinit. Genau wie der Fall $f''(x_0) = 0$ im Eindimensionalen ist das für die Bestimmung von Extrema aber nicht sonderlich hilfreich.

(iii) Seien $\xi, \eta \in \mathbb{R}^n$ mit $\langle (\operatorname{Hess} f)\xi, \xi\rangle > 0$ und $\langle (\operatorname{Hess} f)\eta, \eta\rangle < 0$. Weiterhin sei $\varepsilon > 0$, so dass $x + t\xi, x + t\eta \in U$ für alle $t \in (-\varepsilon, \varepsilon)$. Betrachte

$$g : (-\varepsilon, \varepsilon) \to \mathbb{R}, \quad t \mapsto g(t) := f(x + t\xi).$$

Dann gilt $g'(0) = D_\xi f(x) = 0$ und $g''(0) = \langle A\xi, \xi\rangle > 0$, also hat g in 0 ein isoliertes Minimum und somit gibt es $\varepsilon_1 > 0$ so dass

$$f(x + t\xi) = g(t) > g(0) = f(x) \quad \text{für alle} \quad t \in (0, \varepsilon_1).$$

Analog erhält man für $h(t) := f(x + t\eta)$ ein $\varepsilon_2 > 0$ mit

$$f(x + t\eta) = h(t) < h(0) = f(x) \quad \text{für alle} \quad t \in (0, \varepsilon_2).$$

Zusammengenommen gilt (mit $\varepsilon_3 := \min(\varepsilon_1, \varepsilon_2)$)

$$f(x + t\xi) > f(x) > f(x + t\eta) \quad \text{für alle} \quad t \in (0, \varepsilon_3).$$

Somit kann f in x kein isoliertes Extremum haben. □

Beispiel 8.12. Die Taylorformel erlaubt die Approximation von f durch eine quadratische Funktion; deren Verhalten bestimmt über die Extrema von f. Die folgenden Funktionen haben stets $(0, 0)$ als Nullstelle ihres Gradienten.

(i) Für $f(x, y) = c + x^2 + y^2$ ist
$$\operatorname{Hess} f = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$
positiv definit, also hat f ein isoliertes Minimum bei $(0, 0)$.

(ii) Für $f(x, y) = c - x^2 - y^2$ ist

$$\operatorname{Hess} f = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

negativ definit, also hat f ein isoliertes Maximum bei $(0, 0)$.

(iii) Für $f(x, y) = c + x^2 - y^2$ ist

$$\operatorname{Hess} f = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

indefinit, also hat f kein isoliertes Extremum bei $(0, 0)$ (hier liegt eine sogenannte *Sattelfläche* vor).

(iv) Für $f(x, y) = x^2 - 2xy + y^2$ hat

$$\operatorname{Hess} f = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

die Eigenwerte 0 und 4, also ist mit Satz 8.11 keine Aussage möglich. Wegen $f(x, y) = (x - y)^2$ liegt offensichtlich ein globales Minimum bei $(0, 0)$ vor, das aber nicht isoliert ist, da $f(0, 0) = f(t, t)$ für alle $t \in \mathbb{R}$.

9. Satz von der lokalen Umkehrbarkeit

Motivation 9.1. Eine stetig differenzierbare Funktion $f : \mathbb{R} \to \mathbb{R}$ ist lokal invertierbar bei $x_0 \in \mathbb{R}$, falls $f'(x_0) \neq 0$, denn dann ist $f : U \to W$ bijektiv für eine Umgebung U von x_0 und eine Umgebung W von $f(x_0)$.[28]

In der Sprache der Abbildungen bedeutet $f'(x_0) \neq 0$, dass

$$Df(x_0) : \mathbb{R} \to \mathbb{R}, \quad h \mapsto f'(x_0) \cdot h$$

invertierbar ist.

Somit erwarten wir, dass $f : \mathbb{R}^n \to \mathbb{R}^n$ bei x_0 lokal invertierbar ist, falls die lineare Approximation $Df(x_0)$ invertierbar (also bijektiv) ist. Wichtig: Eine lineare Abbildung $\mathbb{R}^p \to \mathbb{R}^q$ kann nur invertierbar sein, wenn $p = q$.

Satz 9.2 (Lokale Umkehrbarkeit)**.** Sei $A \subseteq \mathbb{R}^n$ offen und $f : A \to \mathbb{R}^n$ stetig partiell differenzierbar. Weiterhin sei $x_0 \in A$ so, dass $Df(x_0)$ invertierbar ist. Dann gibt es offene Mengen $U, W \subseteq \mathbb{R}^n$ mit $x_0 \in U \subseteq A$, $f(x_0) \in W$ und $f(U) = W$, so dass $f : U \to W$ ein Inverses $f^{-1} : W \to U$ hat. Außerdem ist f^{-1} stetig partiell differenzierbar mit

$$Df^{-1}(y) = \big(Df(x)\big)^{-1} \quad \text{für alle} \quad y \in W, \quad \text{wobei} \quad x = f^{-1}(y).$$

Bemerkung 9.3. Bezüglich der Standardbasis gilt

$$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

und $Df(x_0)$ ist genau dann invertierbar, wenn $\det\big(Df(x_0)\big) \neq 0$. Wir nennen $\det(Df)$ auch *Jacobi-Determinante*.

Beweis von Satz 9.2. Wir führen den Beweis in mehreren Schritten:

(1) Durch Verschieben mit x_0 und $f(x_0)$ können wir annehmen, dass sowohl $x_0 = 0$ als auch $f(x_0) = 0$ (sonst betrachten wir $f(x + x_0) - f(x_0)$).

(2) Setze $T := Df(x_0)$ und $\tilde{f} := T^{-1} \circ f$. Da T und somit auch T^{-1} linear ist, gilt $D(T^{-1})(x) = T^{-1}$ für alle $x \in \mathbb{R}^n$, mit der Kettenregel folgt

$$D\tilde{f}(x_0) = \underbrace{D(T^{-1})\big(f(x_0)\big)}_{=T^{-1}} \cdot \underbrace{Df(x_0)}_{=T} = \mathrm{id}\,.$$

Aus $f = T \circ \tilde{f}$ folgt $f^{-1} = \tilde{f}^{-1} \circ T^{-1}$, also reicht es, den Satz für $\tilde{f}$ zu zeigen.

(3) Somit können wir annehmen, dass $x_0 = 0$, $f(x_0) = 0$ und $Df(0) = \mathrm{id}$. Wir wollen $y = f(x)$ lokal nach x auflösen und dazu den Banachschen Fixpunktsatz verwenden. Die Fixpunktgleichung erhalten wir aus

$$y = f(x) \quad \Leftrightarrow \quad 0 = y - f(x) \quad \Leftrightarrow \quad x = x + y - f(x) = x - f(x) + y.$$

[28] Eine *Umgebung* von x ist eine offene Menge U mit $x \in U$.

Für $g(x) := x - f(x)$ gilt $g(0) = 0$ und $Dg(0) = 0$. Weiterhin ist

$$\|g(x)\| = \big\|\big(g_1(x), \dots, g_n(x)\big)\big\| \leq \sum_{i=1}^{n} |g_i(x)|.$$

Mit der Taylorformel 8.4 für $k = 0$ gibt es für jedes $i = 1, \dots, n$ ein $\theta_i \in [0,1]$ mit

$$g_i(x) = g_i(0 + x) = \underbrace{g_i(0)}_{=0} + \langle \operatorname{grad} g_i(\theta_i x), x\rangle = \langle \operatorname{grad} g_i(\theta_i x), x\rangle.$$

Da

$$0 = Dg(0) = \begin{pmatrix} \operatorname{grad} g_1(0) \\ \vdots \\ \operatorname{grad} g_n(0) \end{pmatrix}$$

und Dg nach Voraussetzung stetig ist, gibt es ein $r > 0$, so dass

$$\forall x \in \mathbb{R}^n \ \forall i \in \{1, \dots, n\} : \quad \|x\| \leq r \ \Rightarrow \ \|\operatorname{grad} g_i(x)\| < \frac{1}{2n}.$$

Sei $x \in \mathbb{R}^n$ mit $\|x\| \leq r$, dann gilt mit Cauchy-Schwarz

$$\begin{aligned} \|g(x)\| \leq \sum_{i=1}^{n} |g_i(x)| = \sum_{i=1}^{n} |\langle \operatorname{grad} g_i(\theta_i x), x\rangle| &\leq \sum_{i=1}^{n} \underbrace{\|\operatorname{grad} g_i(\theta_i x)\|}_{<\frac{1}{2n}} \cdot \underbrace{\|x\|}_{\leq r} \\ &< \sum_{i=1}^{n} \frac{r}{2n} = \frac{r}{2}, \end{aligned}$$

also gilt

$$g\left(\overline{\mathrm{B}(0,r)}\right) \subseteq \mathrm{B}\left(0, \frac{r}{2}\right).$$

Sei nun $y \in \mathrm{B}(0, \frac{r}{2})$ und betrachte

$$g_y : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto x - f(x) + y = g(x) + y.$$

Für $x \in \overline{\mathrm{B}(0,r)}$ gilt

$$\|g_y(x)\| = \|g(x) + y\| \leq \underbrace{\|g(x)\|}_{<\frac{r}{2}} + \underbrace{\|y\|}_{<\frac{r}{2}} < r,$$

also

$$g_y\left(\overline{\mathrm{B}(0,r)}\right) \subseteq \mathrm{B}(0,r).$$

Für alle $x_1, x_2 \in \overline{\mathrm{B}(0,r)}$ gilt

$$\|g_y(x_1) - g_y(x_2)\| = \|g(x_1) - g(x_2)\| \leq \sum_{i=1}^{n} |g_i(x_1) - g_i(x_2)|.$$

Wieder mit der Taylorformel gibt es ein $t_i \in [0,1]$ mit

$$g_i(x_2) = g_i\big(x_1 + (x_2 - x_1)\big) = g_i(x_1) + \big\langle \operatorname{grad} g_i\big(x_1 + t_i(x_2 - x_1)\big), x_2 - x_1\big\rangle.$$

Hierbei ist $a_i := x_1 + t_i(x_2 - x_1)$ ein Punkt auf der Strecke zwischen x_1 und x_2, also insbesondere $a_i \in \overline{\mathrm{B}(0,r)}$. Mit Cauchy-Schwarz gilt

$$\begin{aligned}
|g_i(x_1) - g_i(x_2)| &= |g_i(x_1) - g_i(x_1) - \langle \operatorname{grad} g_i(a_i), x_2 - x_1\rangle| \\
&= |\langle \operatorname{grad} g_i(a_i), x_2 - x_1\rangle| \\
&\leq \|\operatorname{grad} g_i(a_i)\| \cdot \|x_2 - x_1\| \\
&< \frac{1}{2n} \cdot \|x_2 - x_1\|,
\end{aligned}$$

da $\|a_i\| \leq r$. Zusammengenommen folgt

$$\|g_y(x_1) - g_y(x_2)\| \leq \sum_{i=1}^{n} |g_i(x_1) - g_i(x_2)| < \sum_{i=1}^{n} \frac{1}{2n} \cdot \|x_2 - x_1\| = \frac{1}{2}\|x_2 - x_1\|,$$

also ist g_y eine Kontraktion auf dem vollständigen metrischen Raum $\overline{\mathrm{B}(0,r)}$. Mit dem Banachschen Fixpunktsatz 2.7 gibt es einen eindeutigen Fixpunkt $x \in \overline{\mathrm{B}(0,r)}$ mit

$$x = g_y(x) = x - f(x) + y,$$

also $y = f(x)$. Setze $W := \mathrm{B}(0, \frac{r}{2})$ und $U := f^{-1}(W)$, dann ist folglich $f : U \to W$ bijektiv, also existiert die Umkehrabbildung $f^{-1} : W \to U$.

(4) Für die Stetigkeit von f^{-1} seien $x_1, x_2 \in \overline{\mathrm{B}(0,r)}$, dann gilt

$$\begin{aligned}
\|x_1 - x_2\| &= \big\|\big(g(x_1) + f(x_1)\big) - \big(g(x_2) + f(x_2)\big)\big\| \\
&= \big\|\big(f(x_1) - f(x_2)\big) + \big(g(x_1) - g(x_2)\big)\big\| \\
&\leq \|f(x_1) - f(x_2)\| + \|g(x_1) - g(x_2)\| \\
&< \|f(x_1) - f(x_2)\| + \frac{1}{2}\|x_1 - x_2\|,
\end{aligned}$$

also $\|x_1 - x_2\| \leq 2\|f(x_1) - f(x_2)\|$.
Für $y_1, y_2 \in W = \mathrm{B}(0, \frac{r}{2})$ sind $x_i := f^{-1}(y_i) \in \overline{\mathrm{B}(0,r)}$, also

$$\big\|f^{-1}(y_1) - f^{-1}(y_2)\big\| = \|x_1 - x_2\| \leq 2\|f(x_1) - f(x_2)\| = 2\|y_1 - y_2\|.$$

Aus dieser Ungleichung folgt die Stetigkeit von f^{-1}.

(5) Nach Voraussetzung ist Df stetig mit $(\mathrm{id} - Df)(0) = \mathrm{id} - Df(0) = 0$, für alle x aus einer Umgebung V von 0 gilt also

$$\|\mathrm{id} - Df(x)\| \leq \frac{1}{2}.$$

Ist nun $x \in V$ und $\xi \in \mathbb{R}^n$ mit $D_\xi f(x) = 0$, so gilt

$$\begin{aligned}
\|\xi\| = \|\mathrm{id}(\xi) - D_\xi f(x)\| &= \|\mathrm{id}(\xi) - Df(x) \cdot \xi\| \\
&= \big\|\big(\mathrm{id} - Df(x)\big) \cdot \xi\big\| \\
&\leq \|\mathrm{id} - Df(x)\| \cdot \|\xi\| \leq \frac{1}{2} \cdot \|\xi\|,
\end{aligned}$$

also $\|\xi\| = 0$ und damit $\xi = 0$. Folglich ist $Df(x)$ für alle $x \in V$ invertierbar.[29] OBdA können wir im Folgenden annehmen, dass r klein genug gewählt ist, dass $\overline{\mathrm{B}(0,r)} \subseteq V$ gilt, also dass $Df(x)$ für alle $x \in \overline{\mathrm{B}(0,r)}$ invertierbar ist.

[29] Genau genommen folgt zunächst die Injektivität, die für quadratische Matrizen beziehungsweise lineare Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^n$ aber äquivalent zur Bijektivität ist.

(6) Für die Differenzierbarkeit von f^{-1} sei $y \in W$ und $x := f^{-1}(y)$. Da f in x total differenzierbar ist, existiert ein $\tilde{\varepsilon} > 0$ mit

$$f(x+\eta) = f(x) + Df(x)\cdot\eta + \varphi(\eta) \quad \text{für alle} \quad \|\eta\| < \tilde{\varepsilon},$$

wobei $\frac{\varphi(\eta)}{\|\eta\|} \xrightarrow{\eta\to 0} 0$. Sei $\varepsilon := \frac{\tilde{\varepsilon}}{2}$. Für alle $\xi \in \mathbb{R}^n$ mit $\|\xi\| < \varepsilon$ setze

$$\eta := f^{-1}(y+\xi) - f^{-1}(y) = f^{-1}(y+\xi) - x,$$

dann gilt mit der Abschätzung aus (4)

$$\begin{aligned}\|\eta\| = \big\|f^{-1}(y+\xi) - f^{-1}(y)\big\| \le 2\big\|f\big(f^{-1}(y+\xi)\big) - f\big(f^{-1}(y)\big)\big\| &= 2\|y+\xi-y\| \\ &= 2\|\xi\|,\end{aligned}$$

also insbesondere $\|\eta\| \le 2\|\xi\| < 2\varepsilon = \tilde{\varepsilon}$ und somit

$$\begin{aligned}f(x) + Df(x)\cdot\eta + \varphi(\eta) = f(x+\eta) = f\big(x + f^{-1}(y+\xi) - x\big) &= f\big(f^{-1}(y+\xi)\big) \\ &= y+\xi \\ &= f(x)+\xi.\end{aligned}$$

Da $Df(x)$ laut (5) invertierbar ist, gilt $\eta = Df(x)^{-1}\big(\xi - \varphi(\eta)\big)$. Für alle $\|\xi\| < \varepsilon$ gilt also

$$f^{-1}(y+\xi) = f^{-1}(y) + \eta = f^{-1}(y) + Df(x)^{-1}\cdot\xi \underbrace{-Df(x)^{-1}\cdot\varphi(\eta)}_{=:\psi(\xi)}$$

mit

$$\begin{aligned}\frac{\|\psi(\xi)\|}{\|\xi\|} = \frac{\|Df(x)^{-1}\varphi(\eta)\|}{\|\xi\|} &\le \frac{\|Df(x)^{-1}\|\cdot\|\varphi(\eta)\|}{\|\xi\|} \\ &= \underbrace{\big\|Df(x)^{-1}\big\|}_{\text{konstant}}\cdot\underbrace{\frac{\|\eta\|}{\|\xi\|}}_{\le 2}\cdot\underbrace{\frac{\|\varphi(\eta)\|}{\|\eta\|}}_{\to 0} \xrightarrow{\xi\to 0} 0,\end{aligned}$$

denn für $\xi \to 0$ gilt auch $\eta \to 0$.[30] Somit ist f^{-1} in y total differenzierbar mit Ableitung $Df(x)^{-1}$.

Mit der Cramerschen Regel sind die Einträge von $Df(x)^{-1}$ gegeben durch Quotienten aus Unterdeterminanten von $Df(x)$ und der Determinante von $Df(x)$, hängen also stetig von den stetigen partiellen Ableitungen von f ab und sind somit ebenfalls stetig. Also ist f^{-1} in y auch stetig partiell differenzierbar. □

Die Formel für Df^{-1} folgt auch aus der Kettenregel: Sei $y \in W$ und $x := f^{-1}(y)$, dann gilt

$$\mathrm{id} = D\,\mathrm{id}(x) = D(f^{-1}\circ f)(x) = Df^{-1}\big(f(x)\big)\cdot Df(x),$$

also

$$Df^{-1}(y) = Df^{-1}\big(f(x)\big) = \big(Df(x)\big)^{-1}.$$

Diese Rechnung beweist allerdings nicht die Differenzierbarkeit von f^{-1}, sie dient eher als Erinnerungshilfe für die Form von Df^{-1}.

[30] Ist $\xi \ne 0$ und damit $y+\xi \ne y$, so folgt mit der Injektivität von f^{-1} schon $\eta = f^{-1}(y+\xi) - f^{-1}(y) \ne 0$.

Beispiel 9.4. Identifizieren wir $\mathbb{C}$ mit $\mathbb{R}^2$, so ist

$$p : (0, \infty) \times \mathbb{R} \to \mathbb{R}^2, \quad (r, \varphi) \mapsto \big(r\cos(\varphi), r\sin(\varphi)\big)$$

die Abbildung, die Polarkoordinaten ($z = r\mathrm{e}^{\mathrm{i}\varphi}$) in kartesische Koordinaten ($z = x + \mathrm{i}y$) verwandelt. Mit

$$Dp(r, \varphi) = \begin{pmatrix} \frac{\partial p_1}{\partial r} & \frac{\partial p_1}{\partial \varphi} \\ \frac{\partial p_2}{\partial r} & \frac{\partial p_2}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix}$$

folgt

$$\det\big(Dp(r, \varphi)\big) = r\big(\cos^2(\varphi) + \sin^2(\varphi)\big) = r \neq 0,$$

also ist p lokal invertierbar. Somit kann jede komplexe Zahl $z = x + \mathrm{i}y$ auch wieder in Polarkoordinaten $z = r\mathrm{e}^{\mathrm{i}\varphi}$ ausgedrückt werden (außer $z = 0$, da Null nicht im Bild von p liegt).

10. Implizite Funktionen und Lagrangesche Multiplikatoren

Satz 10.1 (Implizite Funktionen)**.** Sei $U \subseteq \mathbb{R}^p \times \mathbb{R}^q$ offen und $z_0 = (x_0, y_0) \in U$. Weiterhin sei $F : U \to \mathbb{R}^q$ stetig partiell differenzierbar mit $F(z_0) = F(x_0, y_0) = 0$ und die Matrix

$$\frac{\partial F}{\partial y}(x_0, y_0) = \left(\frac{\partial F_i}{\partial y_j}(x_0, y_0)\right)_{i,j=1,\dots,q} \in \mathbb{R}^{q\times q}$$

invertierbar. Dann gibt es $\varepsilon, \delta > 0$ mit $\mathrm{B}(x_0, \delta) \times \mathrm{B}(y_0, \varepsilon) \subseteq U$ und genau eine stetige Abbildung $g : \mathrm{B}(x_0, \delta) \to \mathrm{B}(y_0, \varepsilon)$, so dass

$$F\big(x, g(x)\big) = 0 \quad \text{für alle} \quad x \in \mathrm{B}(x_0, \delta).$$

Für jedes $x \in \mathrm{B}(x_0, \delta)$ ist $g(x) = y$ sogar die einzige Lösung der Gleichung $F(x, y) = 0$.

Beispiel 10.2. Sei $p = q = 1$ und $F(x, y) = x^2 + y^2 - 1$. Dann ist $F(x, y) = 0$ eine implizite Beschreibung für alle Punkte auf dem Einheitskreis. Es ist $\frac{\partial F}{\partial y} = 2y$ eine 1×1-Matrix, die genau dann invertierbar ist, wenn $y \neq 0$.

- Für $z = (0, 1)$ erhalten wir $y = f(x) = \sqrt{1 - x^2}$ auf $(-1, 1)$.
- Für $z = (0, -1)$ erhalten wir $y = f(x) = -\sqrt{1 - x^2}$ auf $(-1, 1)$.
- Für $z = (1, 0)$ und $z = (-1, 0)$ gibt es keine eindeutige lokale Lösung.

Beweis von Satz 10.1. Setze

$$f : U \to \mathbb{R}^p \times \mathbb{R}^q, \quad (x, y) \mapsto \big(x, F(x, y)\big).$$

Dann gilt

$$Df(x, y) = \begin{pmatrix} I_p & 0 \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{pmatrix},$$

also ist Df genau dann invertierbar, wenn $\frac{\partial F}{\partial y}$ invertierbar ist. Mit $n = p + q$ folgt aus Satz 9.2, dass f^{-1} in einer Umgebung von

$$f(z_0) = f(x_0, y_0) = \big(x_0, F(x_0, y_0)\big) = (x_0, 0)$$

existiert, dort gilt $f^{-1}(x, 0) = (x, y)$. Mit $g(x) := y$ folgt

$$\Big(x, F\big(x, g(x)\big)\Big) = f\big(x, g(x)\big) = (x, 0),$$

also $F\big(x, g(x)\big) = 0$. Somit ist g die gesuchte Funktion. Die restlichen Eigenschaften von g folgen aus den Eigenschaften von f mit Satz 9.2. $\square$

Motivation 10.3. Als nächsten Schritt wollen wir Funktionen $h : \mathbb{R}^n \to \mathbb{R}$ maximieren, aber unter der Nebenbedingung $f(x) = 0$ für eine andere Funktion $f : \mathbb{R}^n \to \mathbb{R}$. Beispielsweise interessiert man sich für die größtmögliche Fläche $h(x, y) = x \cdot y$ eines Rechtecks mit Seitenlängen x und y, wobei der Umfang $2(x + y) = 1$ vorgegeben ist, also $f(x, y) = 2(x + y) - 1$.

Die Idee dahinter ist es, die Maximierung unter Nebenbedingung auf ein entsprechendes Problem ohne Nebenbedingung zurückzuführen. Dazu führt man eine zusätzliche (skalare) Variable λ ein, die die Nebenbedingung repräsentiert. Setzt man $F(x,\lambda) = h(x) - \lambda f(x)$, so gilt

$$\operatorname{grad} F(x,\lambda) = \big(\operatorname{grad} h(x) - \lambda \operatorname{grad} f(x), -f(x)\big),$$

also ist genau dann $\operatorname{grad} F(x,\lambda) = 0$, wenn $f(x) = 0$ und $\operatorname{grad} h(x) - \lambda \operatorname{grad} f(x) = 0$. Eine Lösung des neuen Problems erfüllt damit automatisch die Nebenbedingung und sollte dann auch eine Lösung des ursprünglichen Problems sein.

Satz 10.4. Seien $U \subseteq \mathbb{R}^n$ offen, $f, h : U \to \mathbb{R}$ zwei stetig partiell differenzierbare Funktionen und $M := \{x \in U \mid f(x) = 0\}$. Weiterhin sei $a \in M$ so, dass $\operatorname{grad} f(a) \neq 0$ und h in a auf M ein lokales Maximum oder Minimum hat.[31] Dann existiert ein *Lagrange-Multiplikator* $\lambda \in \mathbb{R}$, so dass

$$\operatorname{grad} h(a) = \lambda \cdot \operatorname{grad} f(a).$$

Beweis. Sei $a = (a_1, \ldots, a_n)$. Nach Voraussetzung gilt $\operatorname{grad} f(a) \neq 0$, sei oBdA $\frac{\partial f}{\partial x_1}(a) \neq 0$ (sonst nummerieren wir die Variablen um). Laut Satz 10.1 existieren $\delta, \varepsilon > 0$ und

$$g : \mathrm{B}\big((a_2, \ldots, a_n), \delta\big) \to \mathrm{B}(a_1, \varepsilon),$$

so dass wir $x \in M$ in der Nähe von a parametrisieren können durch

$$x = \big(g(x_2, \ldots, x_n), x_2, \ldots, x_n\big),$$

also

$$f\big(g(x_2, \ldots, x_n), x_2, \ldots, x_n\big) = 0. \tag{1}$$

Sei

$$H(x_2, \ldots, x_n) := h\big(g(x_2, \ldots, x_n), x_2, \ldots, x_n\big).$$

Da h in a auf M ein Extremum hat, muss H in $\tilde{a} := (a_2, \ldots, a_n)$ ein lokales Extremum haben, also gilt

$$\frac{\partial H}{\partial x_i}(\tilde{a}) = 0 \quad \text{für alle} \quad i = 2, \ldots, n.$$

Mit $\tilde{g}(\tilde{x}) := (g(\tilde{x}), \tilde{x}) = x$ für $\tilde{x} := (x_2, \ldots, x_n)$ gilt für alle $i = 2, \ldots, n$

$$\begin{aligned}
0 = \frac{\partial H}{\partial x_i}(\tilde{a}) = \frac{\partial (h \circ \tilde{g})}{\partial x_i}(\tilde{a}) &= \sum_{j=1}^{n} \frac{\partial h}{\partial x_j}\underbrace{\big(\tilde{g}(\tilde{a})\big)}_{=a} \cdot \frac{\partial \tilde{g}_j}{\partial x_i}(\tilde{a}) \\
&= \frac{\partial h}{\partial x_1}(a) \cdot \frac{\partial \tilde{g}_1}{\partial x_i}(\tilde{a}) + \sum_{j=2}^{n} \frac{\partial h}{\partial x_j}(a) \cdot \frac{\partial \tilde{g}_j}{\partial x_i}(\tilde{a}) \\
&= \frac{\partial h}{\partial x_1}(a) \cdot \frac{\partial g}{\partial x_i}(\tilde{a}) + \sum_{j=2}^{n} \frac{\partial h}{\partial x_j}(a) \cdot \underbrace{\frac{\partial x_j}{\partial x_i}(\tilde{a})}_{=0 \text{ für } j \neq i} \\
&= \frac{\partial h}{\partial x_1}(a) \cdot \frac{\partial g}{\partial x_i}(\tilde{a}) + \frac{\partial h}{\partial x_i}(a).
\end{aligned}$$

[31] Im Sinne von: Es gibt ein $\varepsilon > 0$, so dass für alle $x \in M$ mit $d(x,a) < \varepsilon$ entweder $h(a) \geq h(x)$ (Maximum) oder $h(a) \leq h(x)$ (Minimum) gilt.

Aus (1) erhalten wir analog

$$0 = \frac{\partial(f \circ \tilde{g})}{\partial x_i}(\tilde{a}) = \frac{\partial f}{\partial x_1}(a) \cdot \frac{\partial g}{\partial x_i}(\tilde{a}) + \frac{\partial f}{\partial x_i}(a),$$

also

$$\frac{\partial g}{\partial x_i}(\tilde{a}) = -\left(\frac{\partial f}{\partial x_1}(a)\right)^{-1} \cdot \frac{\partial f}{\partial x_i}(a).$$

Zusammen hat man also für alle $i = 2, \ldots, n$

$$\frac{\partial h}{\partial x_i}(a) = \underbrace{\frac{\partial h}{\partial x_1}(a) \cdot \left(\frac{\partial f}{\partial x_1}(a)\right)^{-1}}_{=:\lambda} \cdot \frac{\partial f}{\partial x_i}(a).$$

Da die Gleichung trivialerweise auch für $i = 1$ gilt, folgt die Aussage. $\square$

Beispiel 10.5. Für $h(x, y) = xy$ und $f(x, y) = 2(x + y) - 1$ folgt aus dem Ansatz

$$(y, x) = \operatorname{grad} h = \lambda \operatorname{grad} f = \lambda \cdot (2, 2)$$

direkt $x = 2\lambda = y$. Weiterhin gilt

$$0 = f(x, y) = f(2\lambda, 2\lambda) = 2(2\lambda + 2\lambda) - 1 = 8\lambda - 1,$$

also $\lambda = \frac{1}{8}$. Somit ist $(x, y) = (\frac{1}{4}, \frac{1}{4})$ mit $h(\frac{1}{4}, \frac{1}{4}) = \frac{1}{16}$ der einzige Kandidat für das Maximum von h auf $M = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}$. Aus praktischen Gründen wollen wir uns sicherheitshalber auf nicht-negative Kantenlängen beschränken: Die Menge

$$\tilde{M} := \{(x, y) \in M \mid x \geq 0, y \geq 0\} = M \cap \big([0, \infty) \times [0, \infty)\big)$$

ist kompakt und $h : \tilde{M} \to \mathbb{R}$ ist stetig, also gibt es ein Maximum von h auf $\tilde{M}$. An den äußeren Punkten von $\tilde{M}$ (also für $x = 0$ oder $y = 0$) gilt $h(x, y) = 0$, dort kann es also nicht liegen. Folglich muss das Maximum im Punkt $(\frac{1}{4}, \frac{1}{4})$ angenommen werden.

11. Differentiation von parameterabhängigen Integralen

Satz 11.1. Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \times [a,b] \to \mathbb{R}$ stetig und

$$F : U \to \mathbb{R}, \quad x \mapsto \int_a^b f(x,t)\,\mathrm{d}t.$$

Dann gilt:

(a) F ist stetig.

(b) Falls f nach $x_1, \dots, x_n$ stetig partiell differenzierbar ist, so ist F stetig partiell differenzierbar und es gilt

$$\frac{\partial F}{\partial x_j}(x) = \int_a^b \frac{\partial f}{\partial x_j}(x,t)\,\mathrm{d}t \quad \text{für alle} \quad j = 1, \dots, n.$$

Beweis. (a) Sei $x_0 \in U$ und $r > 0$ mit $\overline{\mathrm{B}(x_0,r)} \subseteq U$. Dann ist $K := \overline{\mathrm{B}(x_0,r)} \times [a,b]$ kompakt und $f : K \to \mathbb{R}$ stetig, nach Satz 4.8 sogar gleichmäßig stetig. Sei $\varepsilon > 0$ und setze $\tilde{\varepsilon} := \frac{\varepsilon}{2(b-a)} > 0$, dann gibt es ein $\delta > 0$, so dass

$$\forall x, y \in \overline{\mathrm{B}(x_0,r)}\ \forall s,t \in [a,b] : \quad \|(x,t)-(y,s)\| < \delta \Rightarrow |f(x,t) - f(y,s)| < \tilde{\varepsilon}.$$

Für alle $x \in \overline{\mathrm{B}(x_0,r)}$ mit $\|(x,t)-(x_0,t)\| = \|x - x_0\| < \delta$ und alle $t \in [a,b]$ folgt also

$$\begin{aligned}
|F(x) - F(x_0)| &= \left| \int_a^b f(x,t)\,\mathrm{d}t - \int_a^b f(x_0,t)\,\mathrm{d}t \right| \\
&= \left| \int_a^b \big(f(x,t) - f(x_0,t)\big)\,\mathrm{d}t \right| \\
&\le \int_a^b \underbrace{|f(x,t) - f(x_0,t)|}_{<\tilde{\varepsilon}}\,\mathrm{d}t \\
&\le \int_a^b \tilde{\varepsilon}\,\mathrm{d}t = (b-a)\cdot\tilde{\varepsilon} = \frac{\varepsilon}{2} < \varepsilon.
\end{aligned}$$

Somit ist F stetig in x_0 und damit auf ganz U.

(b) Sei $j \in \{1, \dots, n\}$, $x_0 \in U$ und $r > 0$ mit $\overline{\mathrm{B}(x_0,r)} \subseteq U$. Dann ist $K := \overline{\mathrm{B}(x_0,r)} \times [a,b]$ kompakt und $\frac{\partial f}{\partial x_j} : K \to \mathbb{R}$ stetig, nach Satz 4.8 sogar gleichmäßig stetig. Sei $\varepsilon > 0$ und setze $\tilde{\varepsilon} := \frac{\varepsilon}{2(b-a)} > 0$, dann gibt es ein $\delta > 0$, so dass

$$\forall x, y \in \overline{\mathrm{B}(x_0,r)}\ \forall s,t \in [a,b] : \quad \|(x,t)-(y,s)\| < \delta \Rightarrow \left| \frac{\partial f}{\partial x_j}(x,t) - \frac{\partial f}{\partial x_j}(y,s) \right| < \tilde{\varepsilon}.$$

Sei $(h_n)_{n\in\mathbb{N}}$ eine Nullfolge in $\mathbb{R}$ mit $0 < |h_n| < \delta$. Für festes $t \in [a,b]$ gibt es laut dem Mittelwertsatz ein $\theta_n \in [0,1]$ mit

$$\frac{f(x_0 + h_n e_j, t) - f(x_0,t)}{h_n} = \frac{\partial f}{\partial x_j}(x_0 + \theta_n h_n e_j, t).$$

Wegen

$$\|(x_0+\theta_n h_n e_j,t)-(x_0,t)\| = \|\theta_n h_n e_j\| = |\theta_n h_n|\cdot\|e_j\| = |\theta_n|\cdot|h_n| \le |h_n| < \delta$$

gilt

$$\left|\frac{\partial f}{\partial x_j}(x_0+\theta_n h_n e_j,t)-\frac{\partial f}{\partial x_j}(x_0,t)\right| < \tilde{\varepsilon},$$

also folgt

$$\begin{aligned}
&\left|\frac{F(x_0+h_n e_j)-F(x_0)}{h_n}-\int_a^b \frac{\partial f}{\partial x_j}(x_0,t)\,\mathrm{d}t\right| \\
&= \left|\frac{\int_a^b f(x_0+h_n e_j,t)\,\mathrm{d}t-\int_a^b f(x_0,t)\,\mathrm{d}t}{h_n}-\int_a^b \frac{\partial f}{\partial x_j}(x_0,t)\,\mathrm{d}t\right| \\
&= \left|\int_a^b \frac{f(x_0+h_n e_j,t)-f(x_0,t)}{h_n}\,\mathrm{d}t-\int_a^b \frac{\partial f}{\partial x_j}(x_0,t)\,\mathrm{d}t\right| \\
&= \left|\int_a^b \frac{\partial f}{\partial x_j}(x_0+\theta_n h_n e_j,t)\,\mathrm{d}t-\int_a^b \frac{\partial f}{\partial x_j}(x_0,t)\,\mathrm{d}t\right| \\
&= \left|\int_a^b \left(\frac{\partial f}{\partial x_j}(x_0+\theta_n h_n e_j,t)-\frac{\partial f}{\partial x_j}(x_0,t)\right)\mathrm{d}t\right| \\
&\le \int_a^b \underbrace{\left|\frac{\partial f}{\partial x_j}(x_0+\theta_n h_n e_j,t)-\frac{\partial f}{\partial x_j}(x_0,t)\right|}_{<\tilde{\varepsilon}}\mathrm{d}t \\
&\le \int_a^b \tilde{\varepsilon}\,\mathrm{d}t = (b-a)\cdot\tilde{\varepsilon} = \frac{\varepsilon}{2} < \varepsilon,
\end{aligned}$$

also

$$\frac{\partial F}{\partial x_j}(x_0) = \lim_{n\to\infty}\frac{F(x_0+h_n e_j)-F(x_0)}{h_n} = \int_a^b \frac{\partial f}{\partial x_j}(x_0,t)\,\mathrm{d}t.$$

Da $\frac{\partial f}{\partial x_j}$ nach Voraussetzung stetig ist, folgt die Stetigkeit von $\frac{\partial F}{\partial x_j}$ mit Teil (a). □

Beispiel 11.2. Statt mit partieller Integration kann man $\int_0^a t^2\cos(t)\,\mathrm{d}t$ auch wie folgt berechnen: Für

$$F(x) := \int_0^a \cos(tx)\,\mathrm{d}t$$

gilt einerseits

$$F'(x) = \int_0^a \frac{\partial}{\partial x}\cos(tx)\,\mathrm{d}t = -\int_0^a t\sin(tx)\,\mathrm{d}t$$

und

$$F''(x) = -\int_0^a t\frac{\partial}{\partial x}\sin(tx)\,\mathrm{d}t = -\int_0^a t^2\cos(tx)\,\mathrm{d}t,$$

also

$$\int_0^a t^2\cos(t)\,\mathrm{d}t = -F''(1).$$

Andererseits gilt

$$F(x) = \int_0^a \cos(tx)\,\mathrm{d}t = \left[\frac{\sin(tx)}{x}\right]_{t=0}^{t=a} = \frac{\sin(ax)}{x},$$

also

$$F'(x) = \frac{a\cos(ax)x - \sin(ax)}{x^2} = \frac{a\cos(ax)}{x} - \frac{\sin(ax)}{x^2}$$

und

$$\begin{aligned} F''(x) &= \frac{-a^2\sin(ax)x - a\cos(ax)}{x^2} - \frac{a\cos(ax)x^2 - 2x\sin(ax)}{x^4} \\ &= \frac{-a^2\sin(ax)}{x} - \frac{a\cos(ax)}{x^2} - \frac{a\cos(ax)}{x^2} + \frac{2\sin(ax)}{x^3} \\ &= \frac{-a^2\sin(ax)}{x} - \frac{2a\cos(ax)}{x^2} + \frac{2\sin(ax)}{x^3}. \end{aligned}$$

Somit folgt

$$\begin{aligned} \int_0^a t^2\cos(t)\,\mathrm{d}t = -F''(1) &= -\big(-a^2\sin(a) - 2a\cos(a) + 2\sin(a)\big) \\ &= (a^2-2)\sin(a) + 2a\cos(a). \end{aligned}$$

Motivation 11.3. Sei $U \subseteq \mathbb{R}^n$ offen und $v : U \to \mathbb{R}^n$ ein stetig partiell differenzierbares Vektorfeld. Können wir charakterisieren, wann v der Gradient eines Skalarfeldes $f : U \to \mathbb{R}$ ist?[32]

Falls $v = \operatorname{grad} f$, also $v_i = \frac{\partial f}{\partial x_i}$ für alle i, gilt notwendigerweise

$$\frac{\partial v_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial v_j}{\partial x_i}$$

nach dem Satz von Schwarz.[33]

Im Allgemeinen ist dies nicht hinreichend: Im $\mathbb{R}^2$ mit $U := \mathbb{R}^2 \setminus \{0\}$ und

$$v(x,y) = \frac{1}{x^2+y^2}(-y,x)$$

gilt zwar $\frac{\partial v_1}{\partial y} = \frac{\partial v_2}{\partial x}$, aber v kann nicht als $\operatorname{grad} f$ geschrieben werden.

Das Problem ist, dass U ein Loch hat.[34] Falls U keine Löcher hat, so ist die notwendige Bedingung auch hinreichend. Wir beschränken uns hier auf konvexe Mengen U.

Definition 11.4. Sei V ein reeller Vektorraum. Eine Teilmenge $M \subseteq V$ heißt *konvex*, falls

$$\forall x,y \in M\ \forall t \in [0,1]: \quad x + t(y-x) \in M.$$

Satz 11.5. Sei $U \subseteq \mathbb{R}^n$ offen und konvex. Weiterhin sei $v : U \to \mathbb{R}^n$ ein stetig partiell differenzierbares Vektorfeld. Dann sind äquivalent:

(1) Es gibt eine stetig partiell differenzierbare Funktion $f : U \to \mathbb{R}$ mit $v = \operatorname{grad} f$.[35]

(2) Es gilt $\dfrac{\partial v_i}{\partial x_j} = \dfrac{\partial v_j}{\partial x_i}$ für alle $i,j = 1,\dots,n$.

---

[32] In der Elektrostatik betrachtet man etwa das Potential f eines elektrischen Feldes v.

[33] Laut Voraussetzung ist $v = \operatorname{grad} f$ stetig differenzierbar, also ist f zweimal stetig partiell differenzierbar.

[34] Formal würde man sagen, dass U nicht einfach zusammenhängend ist.

[35] Man nennt v dann auch ein *konservatives* Vektorfeld mit *(Skalar-)Potential* f.

Beweis. (1)⇒(2): Folgt aus den Rechnungen in Motivation 11.3.

(2)⇒(1): Durch Verschieben von f können wir oBdA annehmen, dass $0 \in U$. Da U konvex ist, gilt dann $xt \in U$ für alle $x \in U$ und alle $t \in [0,1]$. Setze

$$\varphi : U \times [0,1] \to \mathbb{R}, \quad (x,t) \mapsto \langle v(tx), x\rangle = \sum_{i=1}^{n} v_i(tx)x_i$$

und

$$f : U \to \mathbb{R}, \quad x \mapsto \int_0^1 \varphi(x,t)\,\mathrm{d}t = \int_0^1 \sum_{i=1}^{n} v_i(tx)x_i\,\mathrm{d}t.$$[36]

Einerseits ist φ stetig partiell differenzierbar mit

$$\begin{aligned}
\frac{\partial\varphi}{\partial x_j}(x,t) = \frac{\partial}{\partial x_j}\sum_{i=1}^{n} v_i(tx)x_i &= \sum_{i=1}^{n} \frac{\partial}{\partial x_j} v_i(tx)x_i \\
&= \sum_{i=1}^{n}\left(\frac{\partial v_i(tx)}{\partial x_j}\cdot x_i + v_i(tx)\cdot\frac{\partial x_i}{\partial x_j}\right) \\
&= \left(\sum_{i=1}^{n}\frac{\partial v_i}{\partial x_j}(tx)\cdot t\cdot x_i\right) + v_j(tx) \\
&= \left(\sum_{i=1}^{n}\frac{\partial v_j}{\partial x_i}(tx)\cdot t\cdot x_i\right) + v_j(tx),
\end{aligned}$$

andererseits gilt

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big(t\cdot v_j(tx)\big) &= \frac{\mathrm{d}t}{\mathrm{d}t}\cdot v_j(tx) + t\cdot\frac{\mathrm{d}v_j(tx)}{\mathrm{d}t} \\
&= v_j(tx) + t\cdot\sum_{i=1}^{n}\frac{\partial v_j}{\partial x_i}(tx)\cdot\frac{\mathrm{d}(tx_i)}{\mathrm{d}t} \\
&= v_j(tx) + t\cdot\sum_{i=1}^{n}\frac{\partial v_j}{\partial x_i}(tx)\cdot x_i,
\end{aligned}$$

also

$$\frac{\partial\varphi}{\partial x_j}(x,t) = \frac{\mathrm{d}}{\mathrm{d}t}\big(t\cdot v_j(tx)\big).$$

Mit Satz 11.1 folgt

$$\begin{aligned}
\frac{\partial f}{\partial x_j}(x) = \frac{\partial}{\partial x_j}\int_0^1 \varphi(x,t)\,\mathrm{d}t = \int_0^1 \frac{\partial\varphi}{\partial x_j}(x,t)\,\mathrm{d}t &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\big(t\cdot v_j(tx)\big)\,\mathrm{d}t \\
&= \big[t\cdot v_j(tx)\big]_{t=0}^{t=1} \\
&= v_j(x),
\end{aligned}$$

also $\operatorname{grad} f = v$. □

[36] Ist $\gamma : [0,1] \to U,\ t \mapsto tx$ die Parametrisierung der direkten Verbindung von 0 zu x, so schreibt man auch $f(x) = \left\langle \int_\gamma v(y),\, \mathrm{d}y \right\rangle$.

Bemerkung 11.6. In physikalischen Anwendungen ist besonders der Fall $n = 3$ relevant. Ist $v : U \to \mathbb{R}^3$ ein differenzierbares Vektorfeld, so definiert man die *Rotation* von v als

$$\operatorname{rot} v := \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right).$$

Bedingung (2) aus Satz 11.5 ist dann äquivalent zu $\operatorname{rot} v = 0$.

12. Gewöhnliche Differentialgleichungen und elementare Lösungsmethoden

Motivation 12.1. Eine Differentialgleichung (DGL) ist eine Gleichung für eine Funktion, in der Ableitungen der Funktion vorkommen. Klassische Beispiele aus der Physik sind:

(1) Radioaktiver Zerfall: Ist $y(x)$ die Masse eines radioaktiven Stoffes zum Zeitpunkt x, so ist die Zerfallsrate proportional zur verbleibenden Masse. Dies führt zur Differentialgleichung

$$y'(x) = -k \cdot y(x)$$

für eine stoffspezifische Zerfallskonstante $k > 0$, wobei das Minus auf der rechten Seite dem Umstand geschuldet ist, dass die Masse durch den Zerfall abnimmt. Die Lösung der Gleichung ist

$$y(x) = c \cdot \mathrm{e}^{-kx}.$$

Die Konstante c ist nicht durch die Differentialgleichung bestimmt: $c = y(0)$ ist die Masse zum Zeitpunkt $x = 0$.

(2) Federpendel: Ist $y(x)$ zum Zeitpunkt x die Auslenkung einer Masse m, die an einer Feder mit Federkonstante D aufgehängt ist, so gilt nach dem Hookeschen Gesetz

$$F = -D \cdot y(x)$$

für die Rückstellkraft F der Feder. Andererseits gilt in der Newtonschen Mechanik ("Kraft ist Masse mal Beschleunigung")

$$F = m \cdot y''(x).$$

Für die Bewegung erhält man also die Differentialgleichung

$$m \cdot y''(x) = -D \cdot y(x)$$

oder

$$y''(x) + \frac{D}{m} \cdot y(x) = 0$$

beziehungsweise mit $\omega := \sqrt{\frac{D}{m}}$

$$y''(x) + \omega^2 \cdot y(x) = 0.$$

Lösungen davon sind $y(x) = \sin(\omega x)$ oder $y(x) = \cos(\omega x)$ oder allgemeiner

$$y(x) = A \cdot \sin(\omega x) + B \cdot \cos(\omega x).$$

Hierbei sind die Konstanten A und B wieder nicht durch die Differentialgleichung bestimmt, sondern entsprechen zusätzlich vorgegebenen Anfangsbedingungen zur Zeit $x = 0$: Es gilt $y(0) = B$ und aus

$$y'(x) = A\omega \cos(\omega x) - B\omega \sin(\omega x)$$

folgt $y'(0) = A\omega$, also $A = \frac{y'(0)}{\omega}$. Anschaulich bestimmen außer den physikalischen Größen D und m nur die Anfangsposition $y(0)$ und die Anfangsgeschwindigkeit $y'(0)$ die zeitliche Entwicklung des Systems.

(3) Obige Beispiele sind *gewöhnliche* Differentialgleichungen[37] für Funktionen von einer Variablen. Es gibt auch *partielle* Differentialgleichungen[38] für Funktionen von mehreren Variablen, diese werden wir hier aber nicht behandeln. Beispiele dafür sind die Wärmeleitungsgleichung

$$\frac{\partial}{\partial t} y(x,t) = a \frac{\partial^2}{\partial x^2} y(x,t),$$

die Wellengleichung

$$\frac{\partial^2}{\partial t^2} y(x,t) = c^2 \frac{\partial^2}{\partial x^2} y(x,t)$$

und die Navier-Stokes-Gleichung, die die Strömung von Flüssigkeiten beschreibt.

Definition 12.2. Sei $G \subseteq \mathbb{R}^2$ und $f : G \to \mathbb{R},\ (x,y) \mapsto f(x,y)$ stetig. Dann heißt

$$y' = f(x,y)$$

eine (gewöhnliche) *Differentialgleichung erster Ordnung*. Ist $I \subseteq \mathbb{R}$ ein offenes Intervall und $y : I \to \mathbb{R}$ eine differenzierbare Funktion mit

$$\big\{ \big(x, y(x)\big) : x \in I \big\} \subseteq G$$

und

$$y'(x) = f\big(x, y(x)\big) \quad \text{für alle} \quad x \in I,$$

dann heißt y *Lösung* der obigen Differentialgleichung auf I.

Definition 12.3. Seien $I, J \subseteq \mathbb{R}$ offene Intervalle und $f : I \to \mathbb{R}$ sowie $g : J \to \mathbb{R}$ stetig. Weiterhin sei $g(y) \neq 0$ für alle $y \in J$. Die Differentialgleichung

$$y' = f(x)g(y)$$

auf $G = I \times J$ heißt *Differentialgleichung mit getrennten Variablen.*

Bemerkung 12.4. Die Lösung dieser Differentialgleichung ergibt sich durch folgende informelle Rechnung:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y' = f(x)g(y) \quad \Rightarrow \quad \frac{\mathrm{d}y}{g(y)} = f(x)\,\mathrm{d}x \quad \Rightarrow \quad \int \frac{1}{g(y)}\,\mathrm{d}y = \int f(x)\,\mathrm{d}x.$$

Nun berechnet man Stammfunktionen auf beiden Seiten und löst nach y auf.

Satz 12.5. Seien $I, J \subseteq \mathbb{R}$ offene Intervalle und $f : I \to \mathbb{R}$ sowie $g : J \to \mathbb{R}$ stetig. Weiterhin sei $g(y) \neq 0$ für alle $y \in J$. Betrachte die Differentialgleichung

$$y' = f(x)g(y)$$

und sei $(x_0, y_0) \in I \times J$. Definiere

$$F : I \to \mathbb{R}, \quad x \mapsto F(x) := \int_{x_0}^{x} f(t)\,\mathrm{d}t$$

[37] Kurz GDGL oder ODE für *ordinary differential equation.*
[38] Kurz PDGL oder PDE für *partial differential equation.*

und

$$G : J \to \mathbb{R}, \quad y \mapsto G(y) := \int_{y_0}^{y} \frac{1}{g(t)} \,\mathrm{d}t.$$

Sei $I' \subseteq I$ ein Intervall mit $x_0 \in I'$ und $F(I') \subseteq G(J)$. Dann gibt es genau eine Lösung $y : I' \to \mathbb{R}$ der obigen Differentialgleichung mit Anfangsbedingung $y(x_0) = y_0$. Diese Lösung erfüllt

$$G\big(y(x)\big) = F(x) \quad \text{für alle} \quad x \in I'.$$

Die Funktion G ist auf $G(J)$ invertierbar, also gilt

$$y(x) = G^{-1}\big(F(x)\big) \quad \text{für alle} \quad x \in I'.$$

Beweis. Für

$$G : J \to \mathbb{R}, \quad y \mapsto G(y) := \int_{y_0}^{y} \frac{1}{g(t)} \,\mathrm{d}t.$$

gilt mit dem Hauptsatz der Differential- und Integralrechnung

$$G'(y) = \frac{1}{g(y)} \neq 0 \quad \text{für alle} \quad y \in J.$$

Da $\frac{1}{g}$ auf J stetig ist, folgt mit dem Zwischenwertsatz

$$\forall y \in J : G'(y) > 0 \quad \text{oder} \quad \forall y \in J : G'(y) < 0,$$

also ist G streng monoton. Da G stetig ist, ist $G(J)$ laut Zwischenwertsatz ein Intervall und somit $G : J \to G(J)$ bijektiv, also existiert die Umkehrabbildung $G^{-1} : G(J) \to J$. Mit G ist auch G^{-1} stetig differenzierbar, es gilt

$$(G^{-1})'\big(G(y)\big) = \frac{1}{G'(y)} = g(y) \quad \text{für alle} \quad y \in J.$$

Setze

$$y : I' \to \mathbb{R}, \quad x \mapsto y(x) := G^{-1}\big(F(x)\big),$$

dann gilt

$$y(x_0) = G^{-1}\big(F(x_0)\big) = G^{-1}(0) = y_0$$

und

$$\begin{aligned} y'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(G^{-1}\big(F(x)\big)\Big) = (G^{-1})'\big(F(x)\big) \cdot F'(x) &= g\Big(G^{-1}\big(F(x)\big)\Big) \cdot f(x) \\ &= g\big(y(x)\big) \cdot f(x). \end{aligned}$$

Somit haben wir die Existenz einer Lösung gezeigt.

Die Eindeutigkeit folgt daraus, dass wir die Rechnung aus Bemerkung 12.4 rigoros machen: Sei $y(x)$ Lösung der gegebenen Differentialgleichung mit $y(x_0) = y_0$. Da $y(x)$ im Definitionsbereich von g liegen muss, gilt dann $y(I') \subseteq J$ und auf I' gilt

$$(G \circ y)'(x) = G'\big(y(x)\big) \cdot y'(x) = \frac{1}{g\big(y(x)\big)} \cdot f(x) g\big(y(x)\big) = f(x) = F'(x),$$

also $G \circ y = F + c$ auf I' für eine Konstante c. Aber für $x = x_0$ gilt

$$0 = G(y_0) = G\big(y(x_0)\big) = F(x_0) + c = c,$$

also $c = 0$ und damit $G(y(x)) = F(x)$ für alle $x \in I'$. Somit muss $y(x) = G^{-1}(F(x))$ für alle $x \in I'$ gelten. $\square$

Beispiel 12.6. Auf $I \times J = (0, \infty) \times (0, \infty)$ betrachten wir die Differentialgleichung

$$y' = -\frac{y}{x} \quad \text{mit Anfangsbedingung} \quad y(1) = c > 0.$$

Mit

$$f : (0, \infty) \to \mathbb{R},\ x \mapsto -\frac{1}{x} \quad \text{und} \quad g : (0, \infty) \to \mathbb{R},\ y \mapsto y$$

erhält man

$$F(x) = \int_1^x f(t)\,\mathrm{d}t = -\int_1^x \frac{1}{t}\,\mathrm{d}t = -\ln(x)$$

und

$$G(y) = \int_c^y \frac{1}{g(t)}\,\mathrm{d}t = \int_c^y \frac{1}{t}\,\mathrm{d}t = \big[\ln(t)\big]_c^y = \ln(y) - \ln(c) = \ln\left(\frac{y}{c}\right).$$

Da $G(J) = \mathbb{R}$ können wir $I' = I = (0, \infty)$ wählen. Es gibt also eine auf $(0, \infty)$ definierte eindeutige Lösung $y : (0, \infty) \to \mathbb{R}$ der Differentialgleichung $y' = -\frac{y}{x}$ mit Anfangsbedingung $y(1) = c$. Diese können wir konkret bestimmen: Aus

$$\ln\left(\frac{y(x)}{c}\right) = G\big(y(x)\big) = F(x) = -\ln(x) = \ln\left(\frac{1}{x}\right)$$

folgt $\frac{y(x)}{c} = \frac{1}{x}$, also $y(x) = \frac{c}{x}$. Die Probe $y(1) = \frac{c}{1} = c$ und

$$y'(x) = -\frac{c}{x^2} = -\frac{c}{x} \cdot \frac{1}{x} = -y(x) \cdot \frac{1}{x}$$

bestätigt die Korrektheit der Lösung.

Beispiel 12.7. Auf $I \times J = \mathbb{R} \times (0, \infty)$ betrachten wir die Differentialgleichung

$$y' = y^2 \quad \text{mit Anfangsbedingung} \quad y(0) = c > 0.$$

Mit

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto 1 \quad \text{und} \quad g : (0, \infty) \to \mathbb{R},\ y \mapsto y^2$$

erhält man

$$F(x) = \int_0^x 1\,\mathrm{d}t = x \quad \text{und} \quad G(y) = \int_c^y \frac{1}{t^2}\,\mathrm{d}t = \left[-\frac{1}{t}\right]_c^y = \frac{1}{c} - \frac{1}{y}$$

Da $G((0, \infty)) = (-\infty, \frac{1}{c})$ können wir $I' = (-\infty, \frac{1}{c})$ wählen. Die eindeutige Lösung y auf I' ergibt sich dann gemäß

$$x = F(x) = G\big(y(x)\big) = \frac{1}{c} - \frac{1}{y(x)},$$

also

$$\frac{1}{y(x)} = \frac{1}{c} - x = \frac{1 - cx}{c} \quad \text{und damit} \quad y(x) = \frac{c}{1 - cx} \quad \text{für alle} \quad x \in I'.$$

Die Probe $y(0) = \frac{c}{1-0} = c$ und

$$y'(x) = \frac{-c \cdot (-c)}{(1 - cx)^2} = \frac{c^2}{(1 - cx)^2} = y(x)^2$$

bestätigt die Korrektheit der Lösung.

Die Lösung explodiert für $x \to \frac{1}{c}$, also kann man sie nicht auf ein größeres Intervall ausdehnen. Dieses Verhalten der Lösung ist aus der Differentialgleichung nicht ersichtlich, ist aber typisch für nicht-lineare Differentialgleichungen, die nicht-linear von y abhängen. Für lineare Differentialgleichungen ist die Situation besser!

Definition 12.8. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $a, b : I \to \mathbb{R}$ stetig. Dann heißt

$$y' = a(x)y + b(x)$$

eine *lineare Differentialgleichung erster Ordnung.*[39] Obige Differentialgleichung heißt *homogen*, falls $b = 0$.

Eine homogene lineare Differentialgleichung erster Ordnung ist eine Differentialgleichung mit getrennten Variablen, wir können ihre Lösung also direkt berechnen:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = a(x)y \quad \Rightarrow \quad \frac{\mathrm{d}y}{y} = a(x)\,\mathrm{d}x \quad \Rightarrow \quad \ln(y) + c = \int \frac{1}{y}\,\mathrm{d}y = \int a(x)\,\mathrm{d}x,$$

also

$$y = \exp\left(\int a(x)\,\mathrm{d}x + c\right).$$

Satz 12.9. Sei $I \subseteq \mathbb{R}$ offenes Intervall und $a : I \to \mathbb{R}$ stetig. Weiterhin seien $x_0 \in I$ und $y_0 \in \mathbb{R}$. Dann gibt es genau eine Lösung $y : I \to \mathbb{R}$ der homogenen linearen Differentialgleichung

$$y' = a(x)y \quad \text{mit Anfangsbedingung} \quad y(x_0) = y_0.$$

Diese ist gegeben durch

$$y(x) = y_0 \cdot \exp\left(\int_{x_0}^{x} a(t)\,\mathrm{d}t\right).$$

Beweis. Wegen $y(x_0) = y_0 \cdot \exp(0)) = y_0$ und

$$y'(x) = \underbrace{y_0 \cdot \exp\left(\int_{x_0}^{x} a(t)\,\mathrm{d}t\right)}_{=y(x)} \cdot \underbrace{\frac{\mathrm{d}}{\mathrm{d}x}\int_{x_0}^{x} a(t)\,\mathrm{d}t}_{=a(x)} = a(x)y(x)$$

ist $y(x)$ eine Lösung. Für die Eindeutigkeit können wir Satz 12.5 nicht direkt anwenden, da $g(y) = y \neq 0$ nicht für alle $y \in J$ garantiert werden kann. Deshalb argumentieren wir wie folgt: Sei $y(x)$ eine Lösung von

$$y' = a(x)y(x) \quad \text{mit Anfangsbedingung} \quad y(x_0) = y_0.$$

Wir betrachten

$$z(x) := y(x) \cdot \exp\left(-\int_{x_0}^{x} a(t)\,\mathrm{d}t\right),$$

für alle $x \in I$ gilt dann

$$\begin{aligned} z'(x) &= y'(x) \cdot \exp\left(-\int_{x_0}^{x} a(t)\,\mathrm{d}t\right) + y(x) \cdot \big(-a(x)\big) \cdot \exp\left(-\int_{x_0}^{x} a(t)\,\mathrm{d}t\right) \\ &= \exp\left(-\int_{x_0}^{x} a(t)\,\mathrm{d}t\right) \cdot \underbrace{\big(y'(x) - a(x)y(x)\big)}_{=0} = 0. \end{aligned}$$

Folglich ist z konstant, also gibt es ein $c \in \mathbb{R}$ mit $z(x) = c$ für alle $x \in I$, insbesondere

$$z(x) = z(x_0) = y(x_0) \cdot \exp(0) = y_0,$$

also

$$y(x) = \underbrace{z(x)}_{=y_0} \cdot \exp\left(\int_{x_0}^{x} a(t)\,\mathrm{d}t\right) = y_0 \cdot \exp\left(\int_{x_0}^{x} a(t)\,\mathrm{d}t\right). \qquad \square$$

[39] Wichtig ist, dass Linearität nur in y und nicht in x gefordert wird.

Bemerkung 12.10. Die Lösung existiert hier immer auf ganz I und nicht nur auf einem Teilintervall I'. Im Gegensatz zu nicht-linearen Differentialgleichungen wie in Beispiel 12.7 können lineare Differentialgleichungen nicht zu einer Explosion in endlicher Zeit führen!

Beispiel 12.11. (1) Für $k \in \mathbb{R}$ hat $y' = ky$ mit $I = \mathbb{R}$ und $y(0) = c \in \mathbb{R}$ die Lösung

$$y(x) = c \cdot \exp\left(\int_0^x k \,\mathrm{d}t\right) = c \cdot \mathrm{e}^{kx}.$$

(2) Die Lösung der Differentialgleichung $y' = -\frac{y}{x}$ mit $I = (0, \infty)$ und $y(1) = c$ aus Beispiel 12.6 lässt sich nun erneut berechnen als

$$y(x) = c \cdot \exp\left(\int_1^x \left(-\frac{1}{t}\right) \,\mathrm{d}t\right) = c \cdot \exp\Big(\big[-\ln(t)\big]_1^x\Big) = c \cdot \exp\big(-\ln(x)\big) = c \cdot \frac{1}{x}.$$

Motivation 12.12. Wir betrachten jetzt die allgemeine (inhomogene) lineare Differentialgleichung erster Ordnung

$$y' = a(x)y + b(x) \quad \text{mit Anfangsbedingung} \quad y(x_0) = y_0$$

auf I mit $x_0 \in I$ und $y_0 \in \mathbb{R}$. Sei $A(x) := \int_{x_0}^x a(t)\,\mathrm{d}t$, dann ist $c \cdot \exp(A(x))$ die Lösung der zugehörigen homogenen linearen Differentialgleichung $y' = a(x)y$. Wir machen nun den Ansatz

$$y(x) = c(x) \cdot \exp\big(A(x)\big)$$

für die Lösung der inhomogenen Differentialgleichung, wobei die Funktion $c(x)$ zu bestimmen ist.[40] Dann gilt $A'(x) = a(x)$ und

$$y'(x) = c'(x) \cdot \exp\big(A(x)\big) + c(x) \cdot a(x) \cdot \exp\big(A(x)\big),$$

also ist $y' = a(x)y + b(x)$ äquivalent zu

$$a(x) \cdot c(x) \cdot \exp\big(A(x)\big) + b(x) = y'(x) = c'(x) \cdot \exp\big(A(x)\big) + c(x) \cdot a(x) \cdot \exp\big(A(x)\big),$$

und somit zu $c'(x) = b(x)\mathrm{e}^{-A(x)}$. Durch Integrieren erhält man

$$c(x) = \int_{x_0}^x b(s)\mathrm{e}^{-A(s)}\,\mathrm{d}s + c(x_0),$$

wobei

$$c(x_0) = y(x_0)\mathrm{e}^{-A(x_0)} = y(x_0)\mathrm{e}^0 = y_0$$

und somit

$$\begin{aligned} y(x) = c(x)\mathrm{e}^{A(x)} &= \left(y_0 + \int_{x_0}^x b(s)\mathrm{e}^{-A(s)}\,\mathrm{d}s\right)\mathrm{e}^{A(x)} \\ &= y_0 \cdot \mathrm{e}^{A(x)} + \mathrm{e}^{A(x)} \cdot \int_{x_0}^x b(s)\mathrm{e}^{-A(s)}\,\mathrm{d}s. \end{aligned}$$

Somit haben wir den folgenden Satz bewiesen.

[40] Dieses Vorgehen nennt man auch *Variation der Konstanten*, da die Konstante c in der homogenen Lösung zu einer Funktion $c(x)$ umgewandelt wird.

Satz 12.13. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $x_0 \in I$, $y_0 \in \mathbb{R}$ und $a, b : I \to \mathbb{R}$ stetig. Die allgemeine lineare Differentialgleichung

$$y' = a(x)y + b(x) \quad \text{mit Anfangsbedingung} \quad y(x_0) = y_0$$

hat auf I die eindeutig bestimmte Lösung

$$y(x) = y_0 \cdot \mathrm{e}^{A(x)} + \mathrm{e}^{A(x)} \cdot \int_{x_0}^{x} b(s)\mathrm{e}^{-A(s)}\,\mathrm{d}s, \quad \text{wobei} \quad A(x) := \int_{x_0}^{x} a(t)\,\mathrm{d}t.$$

Beispiel 12.14. Auf $I = \mathbb{R}$ betrachten wir

$$y' = 3x^2 y + x^5 \quad \text{mit Anfangsbedingung} \quad y(0) = 1.$$

Mit $a(x) = 3x^2$ und $b(x) = x^5$ ist

$$A(x) = \int_0^x 3t^2\,\mathrm{d}t = x^3$$

und wegen $(\mathrm{e}^{-s^3})' = -3s^2\mathrm{e}^{-s^3}$ gilt

$$\begin{aligned}
\int_0^x b(s)\mathrm{e}^{-A(s)}\,\mathrm{d}s = \int_0^x s^5\mathrm{e}^{-s^3}\,\mathrm{d}s &= -\frac{1}{3}\int_0^x s^3 \cdot \left(\mathrm{e}^{-s^3}\right)'\,\mathrm{d}s \\
&= -\frac{1}{3}\left(\left[s^3\mathrm{e}^{-s^3}\right]_0^x - \int_0^x 3s^2\mathrm{e}^{-s^3}\,\mathrm{d}s\right) \\
&= -\frac{1}{3}\left(x^3\mathrm{e}^{-x^3} + \int_0^x \left(\mathrm{e}^{-s^3}\right)'\,\mathrm{d}s\right) \\
&= -\frac{1}{3}\left(x^3\mathrm{e}^{-x^3} + \left[\mathrm{e}^{-s^3}\right]_0^x\right) \\
&= -\frac{1}{3}\left(x^3\mathrm{e}^{-x^3} + \mathrm{e}^{-x^3} - 1\right) \\
&= \frac{1}{3} - \frac{1}{3}(x^3+1)\mathrm{e}^{-x^3},
\end{aligned}$$

also

$$y(x) = \mathrm{e}^{x^3} + \mathrm{e}^{x^3}\left(\frac{1}{3} - \frac{1}{3}(x^3+1)\mathrm{e}^{-x^3}\right) = \frac{4}{3}\mathrm{e}^{x^3} - \frac{1}{3}(x^3+1).$$

13. Existenz- und Eindeutigkeitssatz

Definition 13.1. Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ und

$$f : G \to \mathbb{R}^n, \quad (x, y) \mapsto f(x, y) \qquad (\text{wobei } x \in \mathbb{R} \text{ und } y \in \mathbb{R}^n)$$

stetig. Dann heißt

$$y' = f(x, y)$$

ein *System von n Differentialgleichungen erster Ordnung*. Ist $I \subseteq \mathbb{R}$ ein Intervall und $y : I \to \mathbb{R}^n$ differenzierbar mit

- $\{(x, y(x)); x \in I\} \subseteq G$ und
- $y'(x) = f(x, y(x))$ für alle $x \in I$,

dann heißt y *Lösung* des obigen Differentialgleichungssystems auf I.

Bemerkung 13.2. In Komponenten $y = (y_1, \dots, y_n)$ und $f = (f_1, \dots, f_n)$ ausgedrückt bedeutet $y' = f(x, y)$ gerade

$$\begin{aligned} y_1' &= f_1(x, y_1, \dots, y_n), \\ y_2' &= f_2(x, y_1, \dots, y_n), \\ &\vdots \\ y_n' &= f_n(x, y_1, \dots, y_n). \end{aligned}$$

Definition 13.3. Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ stetig. Dann heißt

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$$

eine *Differentialgleichung n-ter Ordnung*. Ist $I \subseteq \mathbb{R}$ ein Intervall und $y : I \to \mathbb{R}$ n-mal differenzierbar mit

- $\left\{(x, y(x), y'(x), \dots, y^{(n-1)(x)}) : x \in I\right\} \subseteq G$ und
- $y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x))$ für alle $x \in I$,

dann heißt y *Lösung* der obigen Differentialgleichung auf I.

Bemerkung 13.4. Eine Differentialgleichung n-ter Ordnung wie in Definition 13.3 kann immer auf ein System von n Differentialgleichungen erster Ordnung wie in Definition 13.1 umgeschrieben werden. Sei dazu für eine Funktion z eine Differentialgleichung n-ter Ordnung

$$z^{(n)} = f(x, z, z', \dots, z^{(n-1)})$$

gegeben. Dann ersetzen wir $z, z', \dots, z^{(n-1)}$ durch neue Variablen

$$y_1 := z, \quad y_2 := z', \quad y_3 := z'', \quad \dots, \quad y_n := z^{(n-1)},$$

die nun dem folgenden speziellen System von n Differentialgleichungen erster Ordnung genügen:

$$\begin{aligned} y_1' &= z' = y_2, \\ y_2' &= z'' = y_3, \\ &\vdots \\ y_{n-1}' &= z^{(n-1)} = y_n, \\ y_n' &= z^{(n)} = f(x, y_1, y_2, \dots, y_n). \end{aligned}$$

Aus den Lösungen des System erhält man dann die Lösung der Differentialgleichung n-ter Ordnung.

Es reicht also die allgemeine Theorie für Systeme von Differentialgleichungen erster Ordnung zu entwickeln und die Resultate dann auf Differentialgleichungen höherer Ordnung umzuschreiben.

Beispiel 13.5. Betrachte die Differentialgleichung zweiter Ordnung

$$z'' = -\omega^2 z \quad \text{mit} \quad \omega > 0$$

des Federpendels aus Motivation 12.1. Als System erhält man

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= -\omega^2 \cdot y_1, \end{aligned}$$

also $y' = f(x, y)$ mit

$$f\left(x, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} y_2 \\ -\omega^2 y_1 \end{pmatrix}.$$

Bemerkung 13.6.

(1) Wir betrachten also im Folgenden Systeme von Differentialgleichungen erster Ordnung wie in Definition 13.1 und versuchen Aussagen zu Existenz und Eindeutigkeit der Lösungen davon zu machen. Wie im Fall $n = 1$ erwarten wir, dass wir neben der Differentialgleichung selbst noch einen Anfangswert $y_0 = y(x_0)$ für ein $x_0 \in I$ vorgeben müssen, um die Eindeutigkeit der Lösung zu erhalten.

(2) Die grobe Idee zu den Existenz- und Eindeutigkeitsaussagen ist es, die Differentialgleichung als Fixpunktgleichung unter einer Abbildung T aufzufassen und den Banachschen Fixpunktsatz zu verwenden. Für die Differentialgleichung $y' = f(x, y)$ direkt ist allerdings nicht klar, wie T gewählt werden sollte. Allerdings können wir die Differentialgleichung mit Hilfe des Hauptsatzes der Differential- und Integralrechnung in eine Integralgleichung umschreiben, für die die Wahl von T dann offensichtlich ist.

Lemma 13.7. Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$, $(x_0, y_0) \in G$ und $f : G \to \mathbb{R}^n$ stetig. Weiterhin sei $I \subseteq \mathbb{R}$ ein Intervall mit $x_0 \in I$ und $y : I \to \mathbb{R}^n$ stetig mit $\{(x, y(x)) : x \in I\} \subseteq G$. Dann sind äquivalent:

(i) Die Funktion y ist eine Lösung des Systems von Differentialgleichungen $y' = f(x, y)$ auf I mit Anfangsbedingung $y(x_0) = y_0$.

(ii) Die Funktion y erfüllt für alle $x \in I$ die Integralgleichung

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big) \, \mathrm{d}t.$$

Beweis. (ii)⇒(i): Sei y eine Lösung der Integralgleichung auf I, dann gilt

$$y(x_0) = y_0 + \int_{x_0}^{x_0} f\big(t, y(t)\big) \, \mathrm{d}t = y_0$$

und nach dem Hauptsatz

$$y'(x) = f\big(x, y(x)\big)$$

(denn $t \mapsto f(t, y(t))$ ist nach Voraussetzung stetig).

(i)⇒(ii): Sei y eine Lösung der Differentialgleichung und $y(x_0) = y_0$. Dann ist y' stetig und laut Hauptsatz gilt

$$\int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t = \int_{x_0}^{x} y'(t)\,\mathrm{d}t = y(x) - y(x_0) = y(x) - y_0,$$

also

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t.$$

□

Bemerkung 13.8. Somit suchen wir nun eine Lösung der Integralgleichung

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t$$

und in dieser Form haben wir eine Fixpunktgleichung, für die T offensichtlich ist:

$$T(y)(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t.$$

Wir brauchen also nur noch einen Banachraum X und die Kontraktionseigenschaft von T, um den Banachschen Fixpunktsatz anwenden zu können. Der Raum $X = \mathcal{C}(I, \mathbb{R}^n)$ ist offensichtlich, für die Kontraktionseigenschaft brauchen wir allerdings stärkere Eigenschaften von f als Stetigkeit.

Wir hoffen also, $\|T(y) - T(\tilde{y})\|_\infty \leq c\|y - \tilde{y}\|_\infty$ mit $c < 1$ abschätzen zu können. Hierzu sollte

$$\begin{aligned}
\|T(y) - T(\tilde{y})\|_\infty &= \sup_x \|T(y)(x) - T(\tilde{y})(x)\| \\
&\leq \sup_x \int_{x_0}^{x} \big\|f\big(t, y(t)\big) - f\big(t, \tilde{y}(t)\big)\big\|\,\mathrm{d}t \\
&\overset{!}{\leq} \sup_x \int_{x_0}^{x} L\|y(t) - \tilde{y}(t)\|\,\mathrm{d}t \\
&\leq \sup_x \int_{x_0}^{x} L\|y - \tilde{y}\|_\infty\,\mathrm{d}t
\end{aligned}$$

gelten. Eine Bedingung, die obige Ungleichungskette vervollständigt, heißt auch *Lipschitz-Bedingung.*

Definition 13.9. Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n,\ (x, y) \mapsto f(x, y)$.

(a) Die Funktion f genügt auf G einer *Lipschitz-Bedingung* bezüglich y mit *Lipschitz-Konstante* $L \geq 0$, falls

$$\forall (x, y), (x, \tilde{y}) \in G : \quad \|f(x, y) - f(x, \tilde{y})\| \leq L\|y - \tilde{y}\|.$$

(b) Die Funktion f genügt auf G *lokal einer Lipschitz-Bedingung*, falls jeder Punkt $(a, b) \in G$ eine Umgebung $U \subseteq \mathbb{R} \times \mathbb{R}^n$ besitzt, so dass f auf $U \cap G$ einer Lipschitz-Bedingung genügt.

Beispiel 13.10. (a) Die Funktion $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ (x, y) \mapsto y^2$ genügt lokal auf $\mathbb{R} \times \mathbb{R}$ einer Lipschitz-Bedingung: Zu $c > 0$ gilt für alle $(x, y), (x, \tilde{y}) \in \mathbb{R} \times [-c, c]$

$$|f(x, y) - f(x, \tilde{y})| = \left|y^2 - \tilde{y}^2\right| = \underbrace{|y + \tilde{y}|}_{\leq 2c} \cdot |y - \tilde{y}| \leq 2c \cdot |y - \tilde{y}|$$

mit der (lokalen) Lipschitz-Konstante $2c$. Allerdings finden wir keine Lipschitz-Konstante für ganz $\mathbb{R} \times \mathbb{R}$: Für $x \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt

$$|f(x, n + 1) - f(x, n)| = (n + 1)^2 - n^2 = 2n + 1 \xrightarrow{n\to\infty} \infty,$$

also gibt es keine Konstante $L \geq 0$ mit

$$|f(x, n + 1) - f(x, n)| \leq L = L|(n + 1) - n|$$

für alle $n \in \mathbb{N}$.

(b) Die Funktion $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ (x, y) \mapsto \sqrt{|y|}$ ist stetig, genügt aber nicht lokal auf $\mathbb{R} \times \mathbb{R}$ einer Lipschitz-Bedingung. Das Problem liegt in der Nähe von $y = 0$: Falls

$$\forall y, \tilde{y} \in [-\varepsilon, \varepsilon] : \quad |f(x, y) - f(x, \tilde{y})| \leq L|y - \tilde{y}|$$

gelten würde, dann wäre

$$\frac{1}{\sqrt{n}} = \left|f\left(x, \frac{1}{n}\right) - f(x, 0)\right| \leq L\left|\frac{1}{n} - 0\right| = L \cdot \frac{1}{n}.$$

für hinreichend großes n, also $\sqrt{n} \leq L$ im Widerspruch zu $\sqrt{n} \xrightarrow{n\to\infty} \infty$.

Lemma 13.11. Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \to \mathbb{R}^n,\ (x, y) \mapsto f(x, y)$ bezüglich $y_1, \ldots, y_n$ stetig partiell differenzierbar. Dann genügt f auf G lokal einer Lipschitz-Bedingung.

Beweis. Übungsaufgabe. □

Satz 13.12 (Eindeutigkeitssatz)**.** Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n,\ (x, y) \mapsto f(x, y)$ eine stetige Funktion, die auf G lokal einer Lipschitz-Bedingung genügt. Seien $y, \tilde{y} : I \to \mathbb{R}^n$ zwei Lösungen des Differentialgleichungssystems $y' = f(x, y)$ auf einem Intervall I. Wenn es ein $x_0 \in I$ gibt mit $y(x_0) = \tilde{y}(x_0)$, dann ist $y(x) = \tilde{y}(x)$ für alle $x \in I$.

Beweis. Wir zeigen zunächst die lokale Eindeutigkeit und weiten diese dann aus.

(1) Sei $y_0 := y(x_0) = \tilde{y}(x_0)$. Da f lokal einer Lipschitz-Bedingung genügt, gibt es eine Umgebung U von (x_0, y_0) und $L > 0$, so dass

$$\forall (x, y), (x, \tilde{y}) \in U : \quad \|f(x, y) - f(x, \tilde{y})\| \leq L|y - \tilde{y}|.$$

Da $(x_0, y_0) \in U$ und $y, \tilde{y} : I \to \mathbb{R}^n$ stetig gibt es ein $\delta > 0$, so dass

$$\forall x \in [x_0 - \delta, x_0 + \delta] \subseteq I : \quad \big(x, y(x)\big), \big(x, \tilde{y}(x)\big) \in U.$$

Sei $\varepsilon := \left\{\delta, \frac{1}{2L}\right\}$ und betrachte $I_0 := [x_0 - \varepsilon, x_0 + \varepsilon] \subseteq I$. Wir zeigen, dass y und $\tilde{y}$ auf I_0 übereinstimmen. Setze dazu

$$M := \sup\left\{\|y(x) - \tilde{y}(x)\| : x \in I_0\right\}$$

(da y und $\tilde{y}$ auf I_0 beschränkt sind, gilt $M < \infty$). Gemäß Lemma 13.7 sind y und $\tilde{y}$ auch Lösungen der zu $y' = f(x, y)$ assoziierten Integralgleichung, also gilt für alle $x \in I_0$ wegen $y(x_0) = \tilde{y}(x_0)$

$$\begin{aligned}
\|y(x) - \tilde{y}(x)\| &= \left\| y(x_0) + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t - \left(\tilde{y}(x_0) + \int_{x_0}^{x} f\big(t, \tilde{y}(t)\big)\,\mathrm{d}t\right)\right\| \\
&= \left\| \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t - \int_{x_0}^{x} f\big(t, \tilde{y}(t)\big)\,\mathrm{d}t \right\| \\
&= \left\| \int_{x_0}^{x} \Big(f\big(t, y(t)\big) - f\big(t, \tilde{y}(t)\big)\Big)\,\mathrm{d}t \right\| \\
&\le \int_{x_0}^{x} \|f\big(\underbrace{t, y(t)}_{\in U}\big) - f\big(\underbrace{t, \tilde{y}(t)}_{\in U}\big)\|\,\mathrm{d}t \\
&\le \int_{x_0}^{x} L \cdot \|y(t) - \tilde{y}(t)\|\,\mathrm{d}t \\
&\le \int_{x_0}^{x} L \cdot M\,\mathrm{d}t = (x - x_0) \cdot L \cdot M \le \varepsilon \cdot L \cdot M \le \frac{1}{2L} \cdot L \cdot M = \frac{M}{2},
\end{aligned}$$

also $\|y(x) - \tilde{y}(x)\| \le \frac{M}{2}$ für alle $x \in I_0$ und damit auch

$$M = \sup_{x \in I_0} \|y(x) - \tilde{y}(x)\| \le \frac{M}{2}.$$

Daraus folgt $M = 0$ und somit $y(x) = \tilde{y}(x)$ für alle $x \in I_0$.

(2) Sei $J \subseteq I$ das größte Intervall mit $x_0 \in J$, auf dem y und $\tilde{y}$ übereinstimmen. Sei x_1 ein Randpunkt von J. Da y und $\tilde{y}$ stetig sind, gilt $y(x_1) = \tilde{y}(x_1)$. Falls x_1 nicht im Rand von I liegt, können wir J lokal um x_1 gemäß Schritt (1) vergrößern, was im Widerspruch zu der Maximalität von J steht. Somit muss $J = I$ sein. □

Satz 13.13 (Existenzsatz von Picard-Lindelöf)**.** Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \to \mathbb{R}^n$ eine stetige Funktion, die auf G lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu jedem $(x_0, y_0) \in G$ ein $\varepsilon > 0$ und eine Lösung $y : [x_0 - \varepsilon, x_0 + \varepsilon] \to \mathbb{R}^n$ des Differentialgleichungssystems $y' = f(x, y)$ mit $y(x_0) = y_0$.

Beweis. Nach Lemma 13.7 genügt es, eine stetige Funktion $y : [x_0 - \varepsilon, x_0 + \varepsilon] \to \mathbb{R}^n$ zu konstruieren mit

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t \quad \text{für alle} \quad x \in [x_0 - \varepsilon, x_0 + \varepsilon].$$

Da G offen ist, gibt es $\gamma, \delta > 0$ mit

$$V := \{(x, y) \in \mathbb{R} \times \mathbb{R}^n : |x - x_0| \le \delta \text{ und } \|y - y_0\| \le \gamma\} \subseteq G,$$

so dass f auf V einer Lipschitz-Bedingung mit Lipschitz-Konstanten $L > 0$ genügt.[41] Da f stetig und V kompakt ist, gibt es ein $M > 0$ mit $\|f(x, y)\| \le M$ für alle $(x, y) \in V$. Setze

$$\varepsilon := \min\left\{\delta, \frac{\gamma}{M}, \frac{1}{2L}\right\} \quad \text{und} \quad I := [x_0 - \varepsilon, x_0 + \varepsilon].$$

[41] Da f auf G lokal einer Lipschitz-Bedingung genügt, gibt es eine Umgebung $\tilde{V}$ von (x_0, y_0), auf der die Lipschitz-Bedingung erfüllt ist. Nun muss man nur noch γ und δ solange verkleinern, bis $V \subseteq \tilde{V}$.

Wir betrachten $\mathcal{C}(I,\mathbb{R}^n) := \{y : I \to \mathbb{R}^n \text{ stetig}\}$ mit der Supremumsnorm

$$\|y\|_I := \sup_{x\in I}\|y(x)\| \quad \text{für alle} \quad y \in \mathcal{C}(I,\mathbb{R}^n).$$

Dann ist $(\mathcal{C}(I,\mathbb{R}^n), \|\cdot\|_I)$ vollständig, also ein Banachraum. Sei $c : I \to \mathbb{R}^n,\ x \mapsto y_0$ und

$$A := \{y \in \mathcal{C}(I,\mathbb{R}^n) : \|y - c\|_I \le \gamma\},$$

dann ist A abgeschlossen in $\mathcal{C}(I,\mathbb{R}^n)$, also ein vollständiger metrischer Raum. Für $y \in A$ definiere $Ty : I \to \mathbb{R}^n$ durch

$$(Ty)(x) := y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t \quad \text{für alle} \quad x \in I.$$ [42]

Die Abbildung Ty ist differenzierbar, also auch stetig und somit gilt $Ty \in \mathcal{C}(I,\mathbb{R}^n)$. Für alle $x \in I$ gilt

$$\begin{aligned}
\|Ty(x) - y_0\| &= \left\| y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t - y_0 \right\| \\
&= \left\| \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t \right\| \\
&\le \int_{x_0}^{x} \big\| f\big(t, y(t)\big) \big\|\,\mathrm{d}t \le \int_{x_0}^{x} M\,\mathrm{d}t = (x - x_0)M \le \varepsilon M \le \gamma,
\end{aligned}$$

also auch $\|Ty - c\|_I \le \gamma$ und damit $Ty \in A$, also $T(A) \subseteq A$. Es bleibt zu zeigen, dass $T : A \to A$ eine Kontraktion ist: Seien $y, \tilde{y} \in A$. Für alle $x \in I$ gilt

$$\begin{aligned}
\|(Ty)(x) - (T\tilde{y})(x)\| &= \left\| y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t - \left(y_0 + \int_{x_0}^{x} f\big(t, \tilde{y}(t)\big)\,\mathrm{d}t \right) \right\| \\
&= \left\| \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t - \int_{x_0}^{x} f\big(t, \tilde{y}(t)\big)\,\mathrm{d}t \right\| \\
&= \left\| \int_{x_0}^{x} \Big(f\big(t, y(t)\big) - f\big(t, \tilde{y}(t)\big) \Big)\,\mathrm{d}t \right\| \\
&\le \int_{x_0}^{x} \big\| f\big(t, y(t)\big) - f\big(t, \tilde{y}(t)\big) \big\|\,\mathrm{d}t \\
&\le \int_{x_0}^{x} L \cdot \|y(t) - \tilde{y}(t)\|\,\mathrm{d}t \\
&\le \int_{x_0}^{x} L \cdot \|y - \tilde{y}\|_I\,\mathrm{d}t \\
&\le |x - x_0| L \cdot \|y - \tilde{y}\|_I \\
&\le \varepsilon L \cdot \|y - \tilde{y}\|_I \\
&\le \frac{1}{2}\|y - \tilde{y}\|_I,
\end{aligned}$$

also $\|Ty - T\tilde{y}\|_I \le \frac{1}{2}\|y - \tilde{y}\|_I$ und somit ist $T : A \to A$ eine Kontraktion. Nach dem Banachschen Fixpunktsatz 2.7 gibt es ein $y \in A$ mit $T(y) = y$, also

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t \quad \text{für alle} \quad x \in I.$$

Somit ist y Lösung des Differentialgleichungssystems $y' = f(x, y)$ auf I mit $y(x_0) = y_0$. $\square$

[42] Für alle $t \in [x_0, x] \subseteq I$ gilt $\|y(t) - y_0\| \le \|y - c\|_I \le \gamma$, also $(t, y(t)) \in V \subseteq G$. Somit sind $f(t, y(t))$ und folglich auch T wohldefiniert.

Bemerkung 13.14. Picard-Lindelöf garantiert nur eine lokale Lösung in der Nähe von x_0. Wir wollen diese natürlich so weit wie möglich ausdehnen. In Beispiel 12.7 haben wir gesehen, dass bei nicht-linearen Differentialgleichungen keine globale Lösung auf ganz I existieren muss:

$$y' = y^2 \quad \text{mit Anfangswert} \quad y(0) = c \quad \text{auf} \quad G = \mathbb{R} \times (0, \infty)$$

hat die Lösung $y(x) = \frac{c}{1-cx}$. Diese ist auf $J := (-\infty, \frac{1}{c})$ definiert, kann aber nicht fortgesetzt werden, da sie für $x \nearrow \frac{1}{c}$ explodiert. Eine solche maximale Lösung existiert immer.

Korollar 13.15. Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \to \mathbb{R}^n$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu jedem $(x_0, y_0) \in G$ zum Anfangswertproblem

$$y' = f(x, y) \quad \text{mit} \quad y(x_0) = y_0 \tag{2}$$

eine eindeutige maximale Lösung $y : J \to \mathbb{R}^n$, wobei J ein Intervall mit $x_0 \in J$ ist, genauer:

(a) Die Funktion y löst (2).

(b) Ist I ein Intervall mit $x_0 \in I$ und $z : I \to \mathbb{R}^n$ eine weitere Lösung von (2), so gilt $I \subseteq J$ und $z(x) = y(x)$ für alle $x \in I$.

Beweis. Sei

$$M := \{I \subseteq \mathbb{R} \text{ Intervall} : x_0 \in I, (2) \text{ hat Lösung } y_I : I \to \mathbb{R}^n \text{ auf } I\}.$$

Nach Satz 13.13 ist $M \neq \emptyset$. Setze $J := \bigcup_{I \in M} I$, dann ist J ein Intervall mit $x_0 \in J$. Sind $I_1, I_2 \in M$, so gilt laut Satz 13.12 $y_{I_1}(x) = y_{I_2}(x)$ für alle $x \in I_1 \cap I_2$. Somit ist

$$y : J \to \mathbb{R}^n, \quad x \mapsto y_I(x) \quad \text{falls} \quad x \in I \in M$$

wohldefiniert und erfüllt beide gewünschten Eigenschaften nach Konstruktion, ist also die eindeutige maximale Lösung. □

Bemerkung 13.16. Die Iteration gemäß dem Banachschen Fixpunktsatz im Beweis von Picard-Lindelöf ergibt eine Folge, die gegen die Lösung konvergiert. Für $x \in I$ setze $y_0(x) := y_0$ und

$$y_k(x) := y_0 + \int_{x_0}^{x} f\big(t, y_{k-1}(t)\big)\, \mathrm{d}t \quad \text{für alle} \quad k \in \mathbb{N}.$$

Die Lipschitz-Bedingung garantiert dann die gleichmäßige Konvergenz dieser Folge gegen die eindeutige Lösung. Ohne Lipschitz-Bedingung kann man diese Folge nicht kontrollieren. Man kann dann aber andere Approximationen der Lösung benutzen und eventuell Existenz einer Lösung ohne Eindeutigkeit zeigen. Dahinter stehen dann Kompaktheitsargumente, insbesondere Arzelà-Ascoli.

Wir betrachten den Fall $n = 1$.

Satz 13.17 (Peano)**.** Sei $f : [x_0, b] \times [y_0 - R, y_0 + R] \to \mathbb{R}$ eine stetige Funktion. Dann hat die Differentialgleichung

$$y' = f(x, y) \quad \text{mit Anfangswert} \quad y(x_0) = y_0$$

eine Lösung auf $[x_0, x_0 + h]$, wobei $h := \min \left\{ b - x_0, \frac{R}{\|f\|_\infty} \right\}$.

Beweisskizze. Wir formulieren die Differentialgleichung wieder als Fixpunktproblem in Integralform um für $T : \mathcal{C}[x_0, x_0 + h] \to \mathcal{C}[x_0, x_0 + h]$ mit

$$(Ty)(x) := y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\, \mathrm{d}t.$$

Für jedes $k \in \mathbb{N}$ definieren wir nun $y_k \in \mathcal{C}[x_0, x_0 + h]$ durch[43]

$$y_k(x) := \begin{cases} y_0, & \text{falls } x_0 \leq x \leq x_0 + \frac{1}{k}, \\ y_0 + \displaystyle\int_{x_0}^{x-\frac{1}{k}} f\big(t, y_k(t)\big)\, \mathrm{d}t, & \text{falls } x_0 + \frac{1}{k} \leq x \leq x_0 + h. \end{cases}$$

Dann gilt:

(a) Die Folge $(y_k)_{k\in\mathbb{N}}$ hat eine gleichmäßig konvergente Teilfolge.

(b) Der Grenzwert dieser Teilfolge erfüllt die Integralgleichung und damit auch die Differentialgleichung.

Hierbei folgt (b) durch Grenzwertbildung in der Definition von y_k. Um (a) zu sehen, benutzen wir Arzelà-Ascoli (5.8); wir müssen also sehen, dass $(y_k)_{k\in\mathbb{N}}$ beschränkt und gleichgradig stetig ist.

(i) Zur Beschränktheit: Für alle $x \in [x_0, x_0 + h]$ gilt

$$|y_k(x) - y_0| \leq \int_{x_0}^{x-\frac{1}{k}} \big|f\big(t, y_k(t)\big)\big|\, \mathrm{d}t \leq \int_{x_0}^{x-\frac{1}{k}} \|f\|_\infty \, \mathrm{d}t \leq h \cdot \|f\|_\infty \leq R,$$

also $\|y_k\| \leq |y_0| + R$ für alle $k \in \mathbb{N}$.[44]

(ii) Zur gleichgradigen Stetigkeit: Sei $\varepsilon > 0$ und setze $\delta := \frac{\varepsilon}{\|f\|_\infty}$ (unabhängig von k). Für alle $x_1, x_2 \in [x_0, x_0 + h]$ mit $\|x_2 - x_1\| < \delta$ gilt dann

$$|y_k(x_2) - y_k(x_1)| = \left| \int_{x_1-\frac{1}{k}}^{x_2-\frac{1}{k}} f\big(t, y_k(t)\big)\, \mathrm{d}t \right| \leq \|f\|_\infty \cdot |x_2 - x_1| < \|f\|_\infty \cdot \delta = \varepsilon.$$

Mit Arzelà-Ascoli gibt es also eine gleichmäßig konvergente Teilfolge $(y_{k_\ell})_{\ell\in\mathbb{N}}$ von $(y_k)_{k\in\mathbb{N}}$ mit Grenzwert

$$y := \lim_{\ell\to\infty} y_{k_\ell} \in \mathcal{C}[x_0, x_0 + h]. \qquad \square$$

Beispiel 13.18. Betrachte

$$y' = y^{\frac{2}{3}} \quad \text{mit Anfangswert} \quad y(0) = 0 \quad \text{auf} \quad [0, b].$$

Die Funktion $f(x, y) = y^{\frac{2}{3}}$ ist nicht lokal Lipschitz (Problem bei $y = 0$), die Voraussetzungen für Peano sind aber erfüllt, also gibt es eine Lösung. Die Eindeutigkeit ist aber verlorengegangen, es gibt sogar unendlich viele Lösungen: Für jedes $c > 0$ ist eine Lösung gegeben durch

$$y(x) := \begin{cases} 0, & \text{falls } x \leq c, \\ \dfrac{(x-c)^3}{27}, & \text{falls } x \geq c. \end{cases}$$

[43] Dies ist keine rekursive Definition in den Folgengliedern! Für jedes k wird y_k auf Intervallen der Länge $\frac{1}{k}$ durch die Werte auf den vorhergehenden Intervallen definiert.

[44] Die Rechnung rechtfertigt auch, dass wir $y_k(t)$ überhaupt in f einsetzen können und somit y_k wohldefiniert ist.

Wie in Bemerkung 13.4 angekündigt schreiben wir unsere Ergebnisse nun von Differentialgleichungssystemen erster Ordnung auf Differentialgleichungen höherer Ordnung um.

Satz 13.19. Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \to \mathbb{R}$ eine stetige Funktion, die auf G lokal einer Lipschitz-Bedingung genügt. Betrachte die Differentialgleichung

$$y^{(n)} = f(x, y', \dots, y^{(n-1)}). \tag{3}$$

Dann gilt:

(a) Eindeutigkeit: Sind $y, \tilde{y} : I \to \mathbb{R}$ Lösungen von (3) auf einem Intervall I mit

$$y(x_0) = \tilde{y}(x_0), \quad y'(x_0) = \tilde{y}'(x_0), \quad \dots, \quad y^{(n-1)}(x_0) = \tilde{y}^{(n-1)}(x_0)$$

für ein $x_0 \in I$, dann gilt $y(x) = \tilde{y}(x)$ für alle $x \in I$.

(b) Existenz: Zu jedem $(x_0, y_0, y_1, \dots, y_{n-1}) \in G$ gibt es ein $\varepsilon > 0$ und eine Lösung $y : [x_0 - \varepsilon, x_0 + \varepsilon] \to \mathbb{R}$ von (3) mit

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1}.$$

(c) Maximale Lösung: Zu jedem $(x_0, y_0, y_1, \dots, y_{n-1}) \in G$ existiert eine eindeutige maximale Lösung $y : J \to \mathbb{R}$ von (3) mit

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1}.$$

Beweis. Folgt aus Satz 13.12, Satz 13.13 und Korollar 13.15. Entscheidend ist, dass sich die lokale Lipschitz-Bedingung für $f : G \to \mathbb{R}$ auf

$$F : G \to \mathbb{R}^n, \qquad \begin{pmatrix} x \\ y_1 \\ \vdots \\ y_n \end{pmatrix} \mapsto \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_n \\ f(x, y_1, \dots, y_n) \end{pmatrix}$$

für das Differentialgleichungssystem überträgt. □

Beispiel 13.20. Um eine eindeutige Lösung der Differentialgleichung

$$y'' = -\omega^2 y \quad \text{mit} \quad \omega > 0$$

zu erhalten, müssen wir (für $x_0 = 0$) die Anfangswerte $y(0)$ und $y'(0)$ vorschreiben. Für $A, B \in \mathbb{R}$ sind

$$y(x) = A\cos(\omega x) + B\sin(\omega x)$$

Lösungen mit $y(0) = A$ und

$$y'(0) = -A\omega\sin(\omega \cdot 0) + B\omega\cos(\omega \cdot 0) = B\omega.$$

Somit hat das Anfangswertproblem

$$y'' = -\omega^2 y \quad \text{mit Anfangswerten} \quad y(0) = y_0 \quad \text{und} \quad y'(0) = y_1$$

die Lösung

$$y(x) = y_0\cos(\omega x) + \frac{y_1}{\omega}\sin(\omega x).$$

Diese ist eindeutig und maximal auf $I = \mathbb{R}$.

14. Lineare Differentialgleichungssysteme

Konvention. In diesem Kapitel sei stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Wir identifizieren $\mathbb{C}^n$ mit $\mathbb{R}^{2n}$ und fassen Funktionen $I \to \mathbb{C}^n$ mit $I \subseteq \mathbb{R}$ als Funktionen $I \to \mathbb{R}^{2n}$ auf. Ableitungen und Integrale solcher Funktionen sind komponentenweise zu verstehen.

Definition 14.1. Sei $I \subseteq \mathbb{R}$ ein Intervall und $A : I \to \mathbb{K}^{n\times n}$ sowie $b : I \to \mathbb{K}^n$ stetig, also

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

für stetige Funktionen $a_{ij} : I \to \mathbb{K}$ und $b_k : I \to \mathbb{K}$ mit $i, j, k = 1, \dots, n$. Dann heißt

$$y' = A(x)y$$

ein *homogenes lineares Differentialgleichungssystem* und

$$y' = A(x)y + b(x)$$

ein *inhomogenes lineares Differentialgleichungssystem* (falls $b \neq 0$).

Bemerkung 14.2. In Komponenten $y = (y_1, \dots, y_n)$ geschrieben lautet das inhomogene System

$$\begin{aligned} y_1' &= a_{11}(x)y_1 + \ldots + a_{1n}y_n + b_1(x), \\ y_2' &= a_{21}(x)y_1 + \ldots + a_{2n}y_n + b_2(x), \\ &\vdots \\ y_n' &= a_{n1}(x)y_1 + \ldots + a_{nn}y_n + b_n(x). \end{aligned}$$

Satz 14.3. Sei $I \subseteq \mathbb{R}$ ein Intervall und $A : I \to \mathbb{K}^{n\times n}$ sowie $b : I \to \mathbb{K}^n$ stetig. Dann gibt es zu jedem $(x_0, y_0) \in I \times \mathbb{K}^n$ genau eine Lösung $y : I \to \mathbb{K}^n$ von

$$y' = A(x)y + b(x) \quad \text{mit Anfangswert} \quad y(x_0) = y_0.$$

Beweis. Sei $G := I \times \mathbb{K}^n$ und

$$f : G \to \mathbb{K}^n, \quad (x, y) \mapsto A(x)y + b(x).$$

Sei $J \subseteq I$ ein kompaktes Teilintervall mit $x_0 \in J$ und $L := \sup_{x\in J} \|A(x)\| < \infty$. Für alle $x \in J$ und alle $y, \tilde{y} \in \mathbb{K}^n$ gilt dann

$$\|f(x, y) - f(x, \tilde{y})\| = \|A(x)(y - \tilde{y})\| \leq \|A(x)\| \cdot \|y - \tilde{y}\| \leq L \cdot \|y - \tilde{y}\|,$$

also genügt f auf $J \times \mathbb{K}^n$ einer (globalen) Lipschitz-Bedingung mit Lipschitz-Konstante L. Die Eindeutigkeit der Lösung folgt dann aus Satz 13.12. Die Existenz einer Lösung auf ganz J wird in der Übung gezeigt.

Da das Anfangswertproblem auf jedem kompakten Teilintervall $J \subseteq I$ mit $x_0 \in J$ eine eindeutige Lösung hat, folgt mit Korollar 13.15, dass die maximale Lösung auf ganz I definiert ist. $\square$

Satz 14.4. Sei $I \subseteq \mathbb{R}$ ein Intervall und $A : I \to \mathbb{K}^{n\times n}$ stetig. Weiterhin sei

$$L_H := \{y : I \to \mathbb{K}^n : y \text{ ist Lösung von } y' = A(x)y\}\,.$$

Dann gilt:

(a) Die Menge L_H ist ein $\mathbb{K}$-Vektorraum.

(b) Für k Elemente $y_1, \ldots, y_k \in L_H$ sind äquivalent:
   (i) Die Funktionen $y_1, \ldots, y_k : I \to \mathbb{K}^n$ sind linear unabhängig über $\mathbb{K}$.
   (ii) Es gibt ein $x_0 \in I$, so dass $y_1(x_0), \ldots, y_k(x_0) \in \mathbb{K}^n$ linear unabhängig über $\mathbb{K}$ sind.
   (iii) Für alle $x_0 \in I$ sind $y_1(x_0), \ldots, y_k(x_0) \in \mathbb{K}^n$ linear unabhängig über $\mathbb{K}$.

(c) Es gilt $\dim(L_H) = n$.

Beweis. (a) Die konstante Nullfunktion ist sicherlich eine Lösung von $y' = A(x)y$, also gilt $0 \in L_H$. Für alle $y, z \in L_H$ und alle $\lambda, \mu \in \mathbb{K}$ gilt

$$(\lambda y + \mu z)' = \lambda y' + \mu z' = \lambda A(x)y + \mu A(x)z = A(x)(\lambda y + \mu z),$$

also $\lambda y + \mu z \in L_H$. Folglich ist L_H ein Unterraum von $\{f : I \to \mathbb{K}^{n\times n}\}$ und damit auch selbst ein $\mathbb{K}$-Vektorraum.

(b) (iii)$\Rightarrow$(ii): Klar.

(ii)$\Rightarrow$(i): Sei $\lambda_1 y_1 + \ldots + \lambda_k y_k = 0$ für $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$. Dann gilt insbesondere auch

$$\lambda_1 y_1(x_0) + \ldots + \lambda_k y_k(x_0) = 0.$$

Laut (ii) sind $y_1(x_0), \ldots, y_k(x_0)$ linear unabhängig, also gilt $\lambda_1 = \ldots = \lambda_k = 0$. Somit sind $y_1, \ldots, y_k$ linear unabhängig.

(i)$\Rightarrow$(iii): Sei $x_0 \in I$. Sind $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ mit

$$\lambda_1 y_1(x_0) + \ldots + \lambda_k y_k(x_0) = 0,$$

so sind $y = 0$ und $\tilde{y} = \lambda_1 y_1 + \ldots + \lambda_k y_k$ zwei Lösungen von $y' = A(x)y$, die an der Stelle x_0 übereinstimmen. Laut Satz 14.3 müssen diese beiden Lösungen dann überall übereinstimmen, also

$$\lambda_1 y_1 + \ldots + \lambda_k y_k = 0.$$

Da $y_1, \ldots, y_k$ laut (i) linear unabhängig sind, folgt $\lambda_1 = \ldots = \lambda_k = 0$. Somit sind $y_1(x_0), \ldots, y_k(x_0)$ linear unabhängig.

(c) Wenn es $n + 1$ linear unabhängige Lösungen $y_1, \ldots, y_{n+1} \in L_H$ gäbe, dann wären laut (b) für jedes $x_0 \in I$ auch die $n + 1$ Vektoren $y_1(x_0), \ldots, y_{n+1}(x_0) \in \mathbb{K}^n$ linear unabhängig im Widerspruch zu $\dim(\mathbb{K}^n) = n$. Also muss $\dim(L_H) \leq n$ gelten.

Sei andererseits $x_0 \in I$ und betrachte die Standardbasis $e_1, \ldots, e_n$ des $\mathbb{K}^n$. Laut Satz 14.3 gibt es dann für jedes $k = 1, \ldots, n$ ein $y_k \in L_H$ mit $y_k(x_0) = e_k$. Dann sind $y_1(x_0), \ldots, y_n(x_0) \in \mathbb{K}^n$ linear unabhängig, laut (b) müssen also auch $y_1, \ldots, y_n$ linear unabhängig sein. Somit folgt $\dim(L_H) \geq n$ und zusammengenommen $\dim(L_H) = n$. □

Definition 14.5. Sei $I \subseteq \mathbb{R}$ ein Intervall und $A : I \to \mathbb{K}^{n\times n}$ stetig. Ein *Lösungsfundamentalsystem* von $y' = A(x)y$ ist eine Basis $(y_1, \dots, y_n)$ des Lösungsraumes L_H.

Bemerkung 14.6. Schreibt man $y_1, \dots, y_n \in L_H$ als Spaltenvektoren

$$y_k = \begin{pmatrix} y_{1k} \\ \vdots \\ y_{nk} \end{pmatrix} \quad \text{und} \quad Y := \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} = \begin{pmatrix} y_{11} & \dots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{n1} & \dots & y_{nn} \end{pmatrix},$$

dann folgt aus Satz 14.4, dass $(y_1, \dots, y_n)$ genau dann ein Lösungsfundamentalsystem sind, wenn $\det Y(x_0) \neq 0$ für ein $x_0 \in I$ gilt (oder, äquivalenterweise, für alle $x_0 \in I$). In diesem Fall hat jede Lösung $y \in L_H$ die Form $y = \lambda_1 y_1 + \ldots + \lambda_n y_n$ für geeignete $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. In Matrixschreibweise gilt also

$$y = Yc \quad \text{für} \quad c = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Man nennt Y dann auch eine *Fundamentalmatrix* von $y' = A(x)y$.

Satz 14.7. Sei $I \subseteq \mathbb{R}$ ein Intervall und $A : I \to \mathbb{K}^{n\times n}$ sowie $b : I \to \mathbb{K}^n$ stetig. Betrachte

$$L_I := \{y : I \to \mathbb{K}^n \mid \forall x \in I : y'(x) = A(x)y(x) + b(x)\}$$

und

$$L_H := \{y : I \to \mathbb{K}^n \mid \forall x \in I : y'(x) = A(x)y(x)\}.$$

Für beliebiges $z \in L_I$ gilt dann $L_I = z + L_H$.[45]

Beweis. Sei zunächst $y \in L_I$. Dann gilt

$$(y - z)' = y' - z' = Ay + b - (Az + b) = A(y - z),$$

also $y - z \in L_H$ und $y = z + (y - z) \in z + L_H$. Dies zeigt $L_I \subseteq z + L_H$.

Sei nun $y \in L_H$. Dann gilt

$$(z + y)' = z' + y' = Az + b + Ay = A(z + y) + b,$$

also $z + y \in L_H$. Dies zeigt $z + L_H \subseteq L_I$. □

Bemerkung 14.8. Formal sieht ein lineares System für allgemeines n genauso aus wie eine lineare Gleichung erster Ordnung (also ein System mit $n = 1$), wobei Zahlen in $\mathbb{K}$ durch Vektoren und Matrizen ersetzt wurden. In Motivation 12.12 haben wir aus der Lösung der homogenen Gleichung durch Variation der Konstanten eine Lösung der inhomogenen Gleichung gefunden. Das funktioniert auch für allgemeines n: Sei $y' = A(x)y + b(x)$ gegeben. Angenommen, wir haben eine Fundamentalmatrix $Y(x)$ des homogenen Systems $y' = A(x)y$, also

$$Y'(x) = A(x)Y(x)$$ [46]

[45] Wie bei herkömmlichen linearen Gleichungssystemen gilt also: Die allgemeine Lösung der inhomogenen Gleichung ist die Summe einer speziellen Lösung der inhomogenen Gleichung und der allgemeinen Lösung der homogenen Gleichung.

[46] Diese kompakte Notation besagt, dass jede Spalte von Y die homogene Differentialgleichung erfüllt.

und jede Lösung des homogenen Systems hat die Form $y(x) = Y(x)c$ für ein $c \in \mathbb{K}^n$. Für das inhomogene System "variieren" wir nun c, wir machen also den Ansatz

$$y(x) = Y(x)c(x)$$

für die Lösung des inhomogenen Systems und rechnen wie im Fall $n = 1$:

$$\begin{aligned} A(x)y(x) + b(x) \overset{!}{=} y'(x) = \big(Y(x)c(x)\big)' &= Y'(x)c(x) + Y(x)c'(x) \\ &= A(x)\underbrace{Y(x)c(x)}_{=y(x)} + Y(x)c'(x) \\ &= A(x)y(x) + Y(x)c'(x). \end{aligned}$$

Somit sollte $b(x) = Y(x)c'(x)$ gelten. Da $Y(x)$ als Fundamentalmatrix invertierbar ist, folgt $c'(x) = Y^{-1}(x)b(x)$, integrieren liefert

$$c(x) = \int_{x_0}^{x} Y^{-1}(t)b(t)\,\mathrm{d}t + c(x_0).$$

Einsetzen und Anpassen der Anfangsbedingung liefert den folgenden Satz.

Satz 14.9 (Variation der Konstanten)**.** Sei $I \subseteq \mathbb{R}$ ein Intervall und $A : I \to \mathbb{K}^{n\times n}$ sowie $b : I \to \mathbb{K}^n$ stetig. Sei $Y(x)$ eine Fundamentalmatrix des homogenen Systems $y' = A(x)y$. Für gegebene $x_0 \in I$ und $c \in \mathbb{K}^n$ ist die eindeutige Lösung des Anfangswertproblems

$$y' = A(x)y + b(x) \quad \text{mit} \quad y(x_0) = c$$

gegeben durch

$$y(x) = Y(x)u(x), \quad \text{wobei} \quad u(x) := \int_{x_0}^{x} Y(t)^{-1}b(t)\,\mathrm{d}t + Y(x_0)^{-1}c.$$

Beweis. (Nochmal durch direktes Nachrechnen:) Es gilt

$$y(x_0) = Y(x_0)u(x_0) = Y(x_0)Y(x_0)^{-1}c = c$$

und

$$\begin{aligned} y'(x) = \big(Y(x)u(x)\big)' = Y'(x)u(x) + Y(x)u'(x) &= A(x)\underbrace{Y(x)u(x)}_{=y(x)} + Y(x)Y(x)^{-1}b(x) \\ &= A(x)y(x) + b(x). \end{aligned}$$ □

Beispiel 14.10. Betrachte die Differentialgleichung zweiter Ordnung

$$z'' = -z + \sin(x).$$[47]

Wir schreiben dies um in ein System

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -y_1 + \sin(x) \end{aligned}$$

[47]Dies ist eine Schwingungsgleichung wie in Motivation 12.1 (2) und Beispiel 13.20 mit $\omega = 1$ und zusätzlicher Anregung $\sin(x)$ von außen.

beziehungsweise $y' = A(x)y + b(x)$ mit

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad A(x) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{und} \quad b(x) = \begin{pmatrix} 0 \\ \sin(x) \end{pmatrix}.$$

Der Bezug zur ursprünglichen Differentialgleichung ist gegeben durch $z = y_1$ und $z' = y_2$. Zwei Lösungen des homogenen Systems

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -y_1 \end{aligned}$$

beziehungsweise $y' = A(x)y$ sind gegeben durch

$$y(x) = \begin{pmatrix} \cos(x) \\ -\sin(x) \end{pmatrix} \quad \text{und} \quad \tilde{y}(x) = \begin{pmatrix} \sin(x) \\ \cos(x) \end{pmatrix}.$$

Setze

$$Y(x) = \begin{pmatrix} y(x) & \tilde{y}(x) \end{pmatrix} = \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix},$$

dann gilt

$$\det Y(x) = \cos^2(x) + \sin^2(x) = 1 \neq 0 \quad \text{für alle} \quad x \in \mathbb{R},$$

also sind y und $\tilde{y}$ linear unabhängig und Y ist eine Fundamentalmatrix des homogenen Systems. Man erhält leicht

$$Y^{-1}(x) = \begin{pmatrix} \cos(x) & -\sin(x) \\ \sin(x) & \cos(x) \end{pmatrix}.$$

Für $x_0 = 0$ und $c = (0, 0)$ ist die Lösung des inhomogenen Systems mit $y(0) = (0, 0)$ gegeben durch $y(x) = Y(x)u(x)$, wobei

$$\begin{aligned} u(x) = \int_0^x \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} 0 \\ \sin(t) \end{pmatrix} \mathrm{d}t &= \int_0^x \begin{pmatrix} -\sin^2(t) \\ \cos(t)\sin(t) \end{pmatrix} \mathrm{d}t \\ &= \begin{pmatrix} \left[\frac{1}{2}\sin(t)\cos(t) - \frac{t}{2}\right]_0^x \\ \left[\frac{1}{2}\sin^2(t)\right]_0^x \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} \sin(x)\cos(x) - x \\ \sin^2(x) \end{pmatrix}, \end{aligned}$$

also[48]

$$\begin{aligned} y(x) &= \frac{1}{2}\begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix} \begin{pmatrix} \sin(x)\cos(x) - x \\ \sin^2(x) \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} \sin(x)\cos^2(x) - x\cos(x) + \sin(x)\sin^2(x) \\ -\sin^2(x)\cos(x) + x\sin(x) + \cos(x)\sin^2(x) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \sin(x) - x\cos(x) \\ x\sin(x) \end{pmatrix}. \end{aligned}$$

Somit ist die Lösung der Differentialgleichung zweiter Ordnung

$$z'' = -z + \sin(x) \quad \text{mit Anfangswerten} \quad z(0) = 0 \quad \text{und} \quad z'(0) = 0$$

gegeben durch

$$z(x) = y_1(x) = \frac{1}{2}\big(\sin(x) - x\cos(x)\big).$$

[48] Da das hier gelöste System aus einer Gleichung höherer Ordnung konstruiert wurde, müsste man an dieser Stelle y_2 nicht ausrechnen, da per Konstruktion $y_2 = y_1'$ gilt.

Bemerkung 14.11. (1) Wir haben die Lösung des inhomogenen Systems also auf die Lösung des homogenen Systems reduziert, allerdings brauchen wir für das homogene System die allgemeine Lösung, oder äquivalenterweise n linear unabhängige Lösungen. Für allgemeine lineare Systeme können wir dazu wenig sagen. Hat das System allerdings *konstante Koeffizienten*, das heißt, $A \in \mathbb{K}^{n\times n}$ hängt nicht von x ab, dann gibt es eine vollständige Theorie dazu. Da A dann auf ganz $\mathbb{R}$ definiert ist, können wir $I = \mathbb{R}$ wählen, und gemäß Satz 14.3 existiert dann eine (durch die Anfangsbedingungen eindeutig bestimmte) Lösung auf ganz $\mathbb{R}$.

(2) Es ist praktischer, die Lösungstheorie über $\mathbb{C}$ zu betrachten, selbst wenn wir nur an Differentialgleichungen und deren Lösungen über $\mathbb{R}$ interessiert sind. Betrachte beispielsweise die Differentialgleichung $z'' = rz$ für $r \in \mathbb{R}\setminus\{0\}$ und das zugehörige System

$$\begin{aligned} y_1' &= y_2,\\ y_2' &= ry_1. \end{aligned}$$

In $\mathbb{R}$ müssen wir unterscheiden, ob $r > 0$ oder $r < 0$:

- Für $r = +1$ hat $z'' = z$ zwei linear unabhängige Lösungen $z(x) = \mathrm{e}^x$ und $z(x) = \mathrm{e}^{-x}$.
- Für $r = -1$ hat $z'' = z$ zwei linear unabhängige Lösungen $z(x) = \sin(x)$ und $z(x) = \cos(x)$.

Über $\mathbb{C}$ haben wir immer die unabhängigen Lösungen

$$z(x) = \mathrm{e}^{\sqrt{r}x} \quad \text{und} \quad z(x) = \mathrm{e}^{-\sqrt{r}x},$$ [49]

für $r = -1$ hat man also auch $z(x) = \mathrm{e}^{\mathrm{i}x}$ und $z(x) = \mathrm{e}^{-\mathrm{i}x}$. Wollen wir reellwertige Lösungen, so erhalten wir diese aus geeigneten Linearkombinationen, etwa

$$\cos(x) = \frac{1}{2}(\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x}) \quad \text{und} \quad \sin(x) = \frac{1}{2\mathrm{i}}(\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x}).$$

Wir betrachten deshalb im Folgenden lineare Systeme mit konstanten Koeffizienten über $\mathbb{C}$.

Satz 14.12. Sei $A \in \mathbb{C}^{n\times n}$ und $v \in \mathbb{C}^n$ ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$. Dann ist

$$y : \mathbb{R} \to \mathbb{C}^n, \quad x \mapsto \mathrm{e}^{\lambda x} v$$

eine Lösung des linearen Differentialgleichungssystems $y' = Ay$.

Beweis. Es gilt $y'(x) = \lambda v \mathrm{e}^{\lambda x} = Av\mathrm{e}^{\lambda x} = Ay(x)$ für alle $x \in \mathbb{R}$. □

Korollar 14.13. Besitzt $A \in \mathbb{C}^{n\times n}$ eine Basis $v_1, \dots, v_n \in \mathbb{C}^n$ von n Eigenvektoren zu Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$, so bilden die Funktionen

$$y_k : \mathbb{R} \to \mathbb{C}^n, \quad x \mapsto \mathrm{e}^{\lambda_k x} v_k$$

mit $k = 1, \dots, n$ ein Fundamentalsystem von $y' = Ay$.

[49] Für $r < 0$ setzen wir $\sqrt{r} := \mathrm{i}\sqrt{-r}$.

Beweis. Laut Satz 14.12 sind alle y_k Lösungen von $y' = Ay$. Da $y_1(0) = v_1, \dots, y_n(0) = v_n$ als Basis linear unabhängig sind, sind laut Satz 14.4 auch $y_1, \dots, y_k$ linear unabhängig. Also sind $y_1, \dots, y_n$ ein Fundamentalsystem von $y' = Ay$. □

Beispiel 14.14. Die Matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$$

hat Eigenwerte $\lambda_1 = \lambda_2 = 1$ und $\lambda_3 = -1$ mit Eigenvektorbasis

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Somit ist die allgemeine komplexwertige Lösung von $y' = Ay$ von der Form

$$y(x) = c_1 \mathrm{e}^x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \mathrm{e}^x \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + c_3 \mathrm{e}^{-x} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} c_1 \mathrm{e}^x \\ c_2 \mathrm{e}^x - c_3 \mathrm{e}^{-x} \\ c_2 \mathrm{e}^x + c_3 \mathrm{e}^{-x} \end{pmatrix}$$

für $c_1, c_2, c_3 \in \mathbb{C}$. Da A nur reelle Eigenwerte hat, erhält man die gleiche allgemeine reellwertige Lösung, nur mit $c_1, c_2, c_3 \in \mathbb{R}$.

Bemerkung 14.15. (1) Die wesentliche Idee bei obigem ist, das System $y' = Ay$ auf ein einfacheres System $z' = Bz$ zu transformieren, dieses zu lösen und die Lösung zurückzutransformieren. In der Situation von Korollar 14.13 setzen wir $S := (v_1, \dots, v_n) \in \mathbb{C}^{n\times n}$ und $z := S^{-1}y$ (oder äquivalenterweise $y = Sz$).[50] Mit $B := S^{-1}AS$ gilt dann

$$z' = S^{-1}y' = S^{-1}Ay = S^{-1}ASz = Bz.$$

Die Matrix B hat eine einfache Struktur, da

$$Be_k = S^{-1}ASe_k = S^{-1}Av_k = S^{-1}\lambda_k v_k = \lambda_k S^{-1} v_k = \lambda_k e_k,$$

also

$$B = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Das System $z' = Bz$ lautet also

$$\begin{aligned} z_1' &= \lambda_1 z_1, \\ z_2' &= \lambda_2 z_2, \\ &\vdots \\ z_n' &= \lambda_n z_n, \end{aligned} \qquad \text{wobei} \quad z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \quad \text{mit} \quad z_k : \mathbb{R} \to \mathbb{C},$$

[50] Die Matrix S ist invertierbar, da $v_1, \dots, v_n$ linear unabhängig sind.

und hat die n linear unabhängigen Lösungen

$$\mathrm{e}^{\lambda_1 x} e_1 = \mathrm{e}^{\lambda_1 x} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathrm{e}^{\lambda_2 x} e_2 = \mathrm{e}^{\lambda_2 x} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathrm{e}^{\lambda_n x} e_n = \mathrm{e}^{\lambda_n x} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Durch Anwendung von S wird dies auf

$$\mathrm{e}^{\lambda_1 x} S e_1 = \mathrm{e}^{\lambda_1 x} v_1, \quad \mathrm{e}^{\lambda_2 x} S e_2 = \mathrm{e}^{\lambda_2 x} v_2, \quad \ldots, \quad \mathrm{e}^{\lambda_n x} S e_n = \mathrm{e}^{\lambda_n x} v_n$$

transformiert und man erhält die Lösungen aus Korollar 14.13.

(2) Große Klassen von Matrizen (wie etwa symmetrische Matrizen) können diagonalisiert werden, aber nicht alle. Man kann Matrizen aber immer auf folgende *Jordan-Normalform* bringen: Sei $A \in \mathbb{C}^{n \times n}$. Dann gibt es eine invertierbare Matrix $S \in \mathbb{C}^{n \times n}$, so dass

$$B := S^{-1} A S = \begin{pmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & J_r \end{pmatrix},$$

wobei jedes J_k ein *Jordan-Kästchen* von der folgenden Form ist:

$$J = J(\lambda, m) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \in \mathbb{C}^{m \times m}$$

mit $\lambda \in \mathbb{C}$ und $m \in \mathbb{N}$.[51] Da die verschiedenen J_k entkoppeln, müssen wir nur noch ein Fundamentalsystem für die Jordan-Kästchen finden. Sei also ein solches $J = J(\lambda, m) \in \mathbb{C}^{m \times m}$ gegeben und betrachte das zugehörige System

$$\begin{aligned} z_1' &= \lambda z_1 + z_2, \\ z_2' &= \lambda z_2 + z_3, \\ &\vdots \\ z_{m-1}' &= \lambda z_{m-1} + z_m, \\ z_m' &= \lambda z_m. \end{aligned}$$

Dies können wir, beginnend bei der letzten Gleichung, rekursiv lösen: Aus $z_m' = \lambda z_m$ folgt $z_m(x) = c_m \mathrm{e}^{\lambda x}$, aus

$$z_{m-1}' = \lambda z_{m-1} + z_m = \lambda z_{m-1} + c_m \mathrm{e}^{\lambda x}$$

[51] Für $\lambda \in \mathbb{C}$ hat man etwa $J(\lambda, 1) = \begin{pmatrix} \lambda \end{pmatrix} \in \mathbb{C}^{1 \times 1}$, $J(\lambda, 2) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ und $J(\lambda, 3) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$.

folgt (Variation der Konstanten oder direktes Nachrechnen)

$$z_{m-1}(x) = c_{m-1}\mathrm{e}^{\lambda x} + c_m x \mathrm{e}^{\lambda x},$$

aus

$$z'_{m-2} = \lambda z_{m-2} + z_{m-1} = \lambda z_{m-2} + c_{m-1}\mathrm{e}^{\lambda x} + c_m x \mathrm{e}^{\lambda x}$$

folgt

$$z_{m-2}(x) = c_{m-2}\mathrm{e}^{\lambda x} + c_{m-1} x \mathrm{e}^{\lambda x} + c_m \frac{x^2}{2}\mathrm{e}^{\lambda x}.$$

Satz 14.16. Seien $\lambda \in \mathbb{C}$ und $m \in \mathbb{N}$ und betrachte ein Jordan-Kästchen

$$J = J(\lambda, m) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \in \mathbb{C}^{m \times m}.$$

Dann ist eine Fundamentalmatrix des homogenen Differentialgleichungssystems $z' = Jz$ gegeben durch

$$z(x) = \mathrm{e}^{\lambda x} \begin{pmatrix} 1 & x & \frac{x^2}{2} & \dots & \frac{x^{m-2}}{(m-2)!} & \frac{x^{m-1}}{(m-1)!} \\ 0 & 1 & x & \dots & \frac{x^{m-3}}{(m-3)!} & \frac{x^{m-2}}{(m-2)!} \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & x & \frac{x^2}{2} \\ 0 & 0 & 0 & \dots & 1 & x \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Für $m = 4$ hat man beispielsweise

$$z(x) = \mathrm{e}^{\lambda x} \begin{pmatrix} 1 & x & \frac{x^2}{2} & \frac{x^3}{6} \\ 0 & 1 & x & \frac{x^2}{2} \\ 0 & 0 & 1 & x \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Beispiel 14.17. Betrachte das Differentialgleichungssystem

$$\begin{array}{rcrrr} y_1' & = & y_1 & -3y_2 & -2y_3, \\ y_2' & = & -y_1 & +y_2 & -y_3, \\ y_3' & = & 2y_1 & +4y_2 & +5y_3, \end{array}$$

also $y' = Ay$ mit

$$A = \begin{pmatrix} 1 & -3 & -2 \\ -1 & 1 & -1 \\ 2 & 4 & 5 \end{pmatrix}.$$

Dann kann A mit

$$S := \begin{pmatrix} -1 & 1 & -1 \\ -1 & 0 & 0 \\ 2 & 0 & 1 \end{pmatrix} \quad \text{und} \quad S^{-1} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$$

auf Jordan-Form

$$B = S^{-1}AS = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

gebracht werden.[52] Das transformierte System für $z = S^{-1}y$ lautet also $z' = Bz$, also

$$\begin{aligned} z_1' &= 2z_1 + z_2, \\ z_2' &= 2z_2, \\ z_3' &= 3z_3. \end{aligned}$$

Drei linear unabhängige Lösungen für z sind gegeben durch

$$\begin{pmatrix} 0 \\ 0 \\ \mathrm{e}^{3x} \end{pmatrix}, \quad \begin{pmatrix} \mathrm{e}^{2x} \\ 0 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} x\mathrm{e}^{2x} \\ \mathrm{e}^{2x} \\ 0 \end{pmatrix}.$$

Rücktransformation mit S ergibt für y die drei linear unabhängigen Lösungen

$$S\begin{pmatrix} 0 \\ 0 \\ \mathrm{e}^{3x} \end{pmatrix} = \mathrm{e}^{3x}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad S\begin{pmatrix} \mathrm{e}^{2x} \\ 0 \\ 0 \end{pmatrix} = \mathrm{e}^{2x}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$$

und

$$S\begin{pmatrix} x\mathrm{e}^{2x} \\ \mathrm{e}^{2x} \\ 0 \end{pmatrix} = x\mathrm{e}^{2x}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} + \mathrm{e}^{2x}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{2x} - x\mathrm{e}^{2x} \\ -x\mathrm{e}^{2x} \\ 2x\mathrm{e}^{2x} \end{pmatrix} =: \tilde{y}.$$

Probe für die dritte Lösung $\tilde{y}$: Es gilt

$$\tilde{y}' = \begin{pmatrix} (\mathrm{e}^{2x} - x\mathrm{e}^{2x})' \\ (-x\mathrm{e}^{2x})' \\ (2x\mathrm{e}^{2x})' \end{pmatrix} = \begin{pmatrix} 2\mathrm{e}^{2x} - \mathrm{e}^{2x} - 2x\mathrm{e}^{2x} \\ -\mathrm{e}^{2x} - 2x\mathrm{e}^{2x} \\ 2\mathrm{e}^{2x} + 4x\mathrm{e}^{2x} \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{2x} - 2x\mathrm{e}^{2x} \\ -\mathrm{e}^{2x} - 2x\mathrm{e}^{2x} \\ 2\mathrm{e}^{2x} + 4x\mathrm{e}^{2x} \end{pmatrix}$$

und

$$\begin{aligned} A\tilde{y} = \begin{pmatrix} 1 & -3 & -2 \\ -1 & 1 & -1 \\ 2 & 4 & 5 \end{pmatrix}\begin{pmatrix} \mathrm{e}^{2x} - x\mathrm{e}^{2x} \\ -x\mathrm{e}^{2x} \\ 2x\mathrm{e}^{2x} \end{pmatrix} &= \begin{pmatrix} \mathrm{e}^{2x} - x\mathrm{e}^{2x} - 3(-x\mathrm{e}^{2x}) - 2\cdot 2x\mathrm{e}^{2x} \\ -(\mathrm{e}^{2x} - x\mathrm{e}^{2x}) - x\mathrm{e}^{2x} - 2x\mathrm{e}^{2x} \\ 2(\mathrm{e}^{2x} - x\mathrm{e}^{2x}) + 4(-x\mathrm{e}^{2x}) + 5\cdot 2x\mathrm{e}^{2x} \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{e}^{2x} - x\mathrm{e}^{2x} + 3x\mathrm{e}^{2x} - 4x\mathrm{e}^{2x} \\ -\mathrm{e}^{2x} + x\mathrm{e}^{2x} - x\mathrm{e}^{2x} - 2x\mathrm{e}^{2x} \\ 2\mathrm{e}^{2x} - 2x\mathrm{e}^{2x} - 4x\mathrm{e}^{2x} + 10x\mathrm{e}^{2x} \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{e}^{2x} - 2x\mathrm{e}^{2x} \\ -\mathrm{e}^{2x} - 2x\mathrm{e}^{2x} \\ 2\mathrm{e}^{2x} + 4x\mathrm{e}^{2x} \end{pmatrix} = \tilde{y}'. \end{aligned}$$

[52] Wie man S findet, lernt man in der Linearen Algebra.

15. Lineare Differentialgleichungen höherer Ordnung

Definition 15.1. Sei $I \subseteq \mathbb{R}$ ein Intervall und $a_0, \ldots, a_{n-1}, b : I \to \mathbb{K}$ stetig. Dann heißt

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \ldots + a_1(x)y' + a_0(x)y = 0 \tag{4}$$

eine *homogene lineare Differentialgleichung n-ter Ordnung* und

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \ldots + a_1(x)y' + a_0(x)y = b(x) \tag{5}$$

eine *inhomogene lineare Differentialgleichung n-ter Ordnung* (falls $b \neq 0$).

Satz 15.2. Sei $I \subseteq \mathbb{R}$ ein Intervall und $a_0, \ldots, a_{n-1}, b : I \to \mathbb{K}$ stetig.

(a) Sei
$$L_H := \{y : I \to \mathbb{K} \mid y \text{ löst } (4)\}.$$
Dann ist L_H ein n-dimensionaler $\mathbb{K}$-Vektorraum.

(b) Sei
$$L_I := \{y : I \to \mathbb{K} \mid y \text{ löst } (5)\}.$$
Dann gilt für beliebiges $z \in L_I$ schon $L_I = z + L_H$.

(c) Funktionen $y_1, \ldots, y_n \in L_H$ sind genau dann linear unabhängig, wenn für ein $x \in I$ (oder äquivalenterweise: für alle $x \in I$) gilt

$$W(x) := \det \begin{pmatrix} y_1(x) & y_2(x) & \ldots & y_n(x) \\ y_1'(x) & y_2'(x) & \ldots & y_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \ldots & y_n^{(n-1)}(x) \end{pmatrix} \neq 0.$$

Definition 15.3. Der Ausdruck $W(x)$ heißt *Wronski-Determinante* von $y_1, \ldots, y_n$.

Beweis von Satz 15.2. Wie in Bemerkung 13.4 schreiben wir (4) als System mit $z_k = y^{(k)}$, also

$$\begin{aligned} z_0' &= z_1, \\ z_1' &= z_2, \\ &\vdots \\ z_{n-2}' &= z_{n-1}, \\ z_{n-1}' &= -a_0(x)z_0 - a_1(x)z_1 - \ldots - a_{n-1}(x)z_{n-1}, \end{aligned} \tag{6}$$

und setze
$$\mathcal{L}_H := \{z : I \to \mathbb{K}^n \mid z = (z_0, \ldots, z_{n-1}) \text{ löst } (6)\}.$$
Dann ist $L_H \to \mathcal{L}_H,\ y \mapsto (y, y', y'', \ldots, y^{(n-1)})$ ein Vektorraumisomorphismus; (a) und (c) folgen dann aus Satz 14.4; (b) folgt aus Satz 14.7. $\square$

Beispiel 15.4. Betrachte die inhomogene lineare Differentialgleichung

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 1 \quad \text{auf} \quad I = (0, \infty).$$

Zwei Lösungen der homogenen Differentialgleichung

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0$$

sind $y_1(x) = x$ und $y_2(x) = \sqrt{x}$, wie eine Probe leicht zeigt: Mit $y_1'(x) = 1$ und $y_1''(x) = 0$ gilt

$$y_1''(x) - \frac{1}{2x}y_1'(x) + \frac{1}{2x^2}y_1(x) = 0 - \frac{1}{2x} \cdot 1 + \frac{1}{2x^2} \cdot x = -\frac{1}{2x} + \frac{1}{2x} = 0,$$

aus $y_2'(x) = \frac{1}{2}x^{-\frac{1}{2}}$ und $y_2''(x) = -\frac{1}{4}x^{-\frac{3}{2}}$ folgt

$$y_2''(x) - \frac{1}{2x}y_2'(x) + \frac{1}{2x^2}y_2(x) = -\frac{1}{4}x^{-\frac{3}{2}} - \frac{1}{4}x^{-\frac{3}{2}} + \frac{1}{2}x^{-\frac{3}{2}} = 0.$$

Diese Lösungen sind linear unabhängig, da

$$W(x) = \det \begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix} = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{pmatrix} = \frac{1}{2}\sqrt{x} - \sqrt{x} = -\frac{1}{2}\sqrt{x} \neq 0$$

für alle $x \in I$. Die allgemeine Lösung der homogenen Differentialgleichung ist also von der Form

$$y_H(x) = \alpha x + \beta\sqrt{x} \quad \text{mit} \quad \alpha, \beta \in \mathbb{R}.$$

Eine spezielle Lösung der inhomogenen Differentialgleichung ist $y_I(x) = \frac{2}{3}x^2$, denn mit $y_I'(x) = \frac{4}{3}x$ und $y_I''(x) = \frac{4}{3}$ gilt

$$y_I''(x) - \frac{1}{2x}y_I'(x) + \frac{1}{2x^2}y_I(x) = \frac{4}{3} - \frac{1}{2x}\frac{4}{3}x + \frac{1}{2x^2}\frac{2}{3}x^2 = \frac{4}{3} - \frac{2}{3} + \frac{1}{3} = 1.$$

Die allgemeine Lösung der inhomogenen Differentialgleichung ist also

$$y(x) = y_H(x) + y_I(x) = \alpha x + \beta\sqrt{x} + \frac{2}{3}x^2 \quad \text{mit} \quad \alpha, \beta \in \mathbb{R}.$$

Bemerkung 15.5. Wie bei (oder: als Spezialfall von) Systemen von linearen Differentialgleichungen erster Ordnung können wir aus der allgemeinen Lösung der homogenen Differentialgleichung durch Variation der Konstanten die allgemeine Lösung der inhomogenen Differentialgleichung bestimmen. Die allgemeine Lösung der homogenen Differentialgleichung können wir allerdings wieder nur im Fall konstanter Koeffizienten einfach bestimmen.

Definition 15.6. Sei $I \subseteq \mathbb{R}$ ein Intervall und $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$. Dann heißt

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1y' + a_0y = 0 \tag{7}$$

eine *homogene lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten.*

Satz 15.7. Sei $I \subseteq \mathbb{R}$ ein Intervall und $a_0, a_1, \dots, a_{n-1} \in \mathbb{C}$.

(a) Sei
$$L_H := \{y : I \to \mathbb{C} \mid y \text{ löst } (7)\}.$$
Dann ist L_H ein n-dimensionaler $\mathbb{C}$-Vektorraum.

(b) Jede Funktion $y \in L_H$ ist beliebig oft differenzierbar.

Beweis. (a) Dies ist ein Spezialfall von Satz 15.2.

(b) Sei $y \in L_H$. Als Lösung von (7) ist y n-mal differenzierbar. Wegen
$$y^{(n)} = -a_{n-1}y^{(n-1)} - \dots - a_0 y$$
ist $y^{(n)}$ als Linearkombination differenzierbarer Funktionen selbst differenzierbar. Also ist y $(n+1)$-mal differenzierbar. Mittels Iteration folgt, dass y beliebig oft differenzierbar ist. □

Bemerkung 15.8. Im Prinzip wissen wir, wie wir (7) lösen können: Wir betrachten das System
$$\begin{aligned} z_0' &= z_1, \\ z_1' &= z_2, \\ &\vdots \\ z_{n-2}' &= z_{n-1}, \\ z_{n-1}' &= -(a_0 z_0 + a_1 z_1 + \dots + a_{n-1} z_{n-1}), \end{aligned}$$
also $z' = Az$ für
$$A := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$
Die Lösung hängt dann von den Eigenwerten und der Jordan-Normalform von A gemäß Bemerkung 14.15 und Satz 14.16. Wir können diese Daten hier aber auch direkter berechnen, ohne mit A zu argumentieren. Wir werden später auf diesen Zusammenhang mit A zurückkommen.

Notation 15.9. Sei $\mathbb{C}[x]$ die Menge aller Polynome in x mit Koeffizienten in $\mathbb{C}$. Für ein Intervall $I \subseteq \mathbb{R}$ sei
$$\mathcal{C}^\infty(I) := \{f : I \to \mathbb{C} \mid f \text{ ist beliebig oft differenzierbar}\}.$$
Wir setzen
$$D : \mathcal{C}^\infty(I) \to \mathcal{C}^\infty(I), \quad f \mapsto f'.$$
Für $p(x) = \sum_{j=0}^n a_j x^j \in \mathbb{C}[x]$ sei
$$p(D) := \sum_{j=0}^n a_j D^j, \quad \text{wobei} \quad D^0 = \mathrm{id}, D^1 = D, D^2 = D \circ D, \dots$$

Für $f \in \mathcal{C}^\infty(I)$ gilt also

$$p(D)f = a_0 f + a_1 f' + a_2 f'' + \ldots + a_{n-1} f^{(n-1)} + a_n f^{(n)}.$$

Dann ist

$$\mathbb{C}[x] \to \{F : \mathcal{C}^\infty(I) \to \mathcal{C}^\infty(I) \mid F \text{ linear}\}, \quad p \mapsto p(D)$$

ein Algebrenhomomorphismus: Für alle $p, q \in \mathbb{C}[x]$ gilt

$$(p+q)(D) = p(D) + q(D) \quad \text{und} \quad (p \cdot q)(D) = p(D) \circ q(D),$$

was man durch Nachrechnen verifizieren sollte.

Lemma 15.10. Für alle $p \in \mathbb{C}[x]$ und alle $\lambda \in \mathbb{C}$ gilt

$$p(D)\mathrm{e}^{\lambda x} = p(\lambda)\mathrm{e}^{\lambda x}.$$

Beweis. Sei $p(x) = \sum_{j=0}^n a_j x^j \in \mathbb{C}[x]$. Dann gilt $D^0 \mathrm{e}^{\lambda x} = 1 \cdot \mathrm{e}^{\lambda x}$,

$$D\mathrm{e}^{\lambda x} = \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{\lambda x} = \lambda \mathrm{e}^{\lambda x},$$

$$D^2\mathrm{e}^{\lambda x} = DD\mathrm{e}^{\lambda x} = D(\lambda \mathrm{e}^{\lambda x}) = \frac{\mathrm{d}}{\mathrm{d}x}(\lambda \mathrm{e}^{\lambda x}) = \lambda^2 \mathrm{e}^{\lambda x}$$

und induktiv $D^j \mathrm{e}^{\lambda x} = \lambda^j \mathrm{e}^{\lambda x}$, also

$$p(D)\mathrm{e}^{\lambda x} = \sum_{j=0}^n a_j D^j \mathrm{e}^{\lambda x} = \sum_{j=0}^n a_j \lambda^j \mathrm{e}^{\lambda x} = p(\lambda)\mathrm{e}^{\lambda x}. \qquad \square$$

Satz 15.11. Sei $p(x) = \sum_{j=0}^n a_j x^j \in \mathbb{C}[x]$ mit $a_n = 1$, also

$$p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0.$$

Das Polynom p habe n paarweise verschiedene Nullstellen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Dann bilden $y_k(x) = \mathrm{e}^{\lambda_k x}$ mit $k = 1, \ldots, n$ ein Lösungsfundamentalsystem der Differentialgleichung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0.$$

Beweis. Laut Lemma 15.10 gilt

$$p(D)y_k = p(D)\mathrm{e}^{\lambda_k x} = \underbrace{p(\lambda_k)}_{=0}\mathrm{e}^{\lambda_k x} = 0,$$

also ist jedes y_k eine Lösung der Differentialgleichung. Wegen $y_k^{(j)} = \lambda_k^j \mathrm{e}^{\lambda_k x}$ ist dann die Wronski-Determinante von $y_1, \ldots, y_n$ an der Stelle $x = 0$ gegeben durch

$$W(0) = \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{pmatrix}.$$

Man zeigt (z.B. in der linearen Algebra), dass diese sogenannte *Vandermonde-Determinante* den Wert

$$W(0) = \prod_{1 \le i < j \le n} (\lambda_j - \lambda_i)$$

hat. Da nach Voraussetzung $\lambda_1, \ldots, \lambda_n$ paarweise verschieden sind, gilt also $\lambda_i \neq \lambda_j$ für alle $i \neq j$ und damit $W(0) \neq 0$. Somit sind nach Satz 15.2 $y_1, \ldots, y_n$ linear unabhängig und bilden ein Lösungsfundamentalsystem der Differentialgleichung. $\square$

Beispiel 15.12. Die Differentialgleichung

$$y''' - y'' + y' - y = 0$$

korrespondiert zum Polynom

$$p(x) = x^3 - x^2 + x - 1 = (x^2 + 1)(x - 1) = (x - \mathrm{i})(x + \mathrm{i})(x - 1)$$

mit Nullstellen $\lambda_1 = \mathrm{i}$, $\lambda_2 = -\mathrm{i}$ und $\lambda_3 = 1$. Über $\mathbb{C}$ haben wir somit das Fundamentalsystem

$$y_1(x) = \mathrm{e}^{\mathrm{i}x}, \quad y_2(x) = \mathrm{e}^{-\mathrm{i}x} \quad \text{und} \quad y_3(x) = \mathrm{e}^{x}.$$

Über $\mathbb{R}$ nehmen wir reelle Linearkombinationen von y_1 und y_2 und erhalten

$$\frac{1}{2}\big(y_1(x) + y_2(x)\big) = \cos(x) \quad \text{sowie} \quad \frac{1}{2\mathrm{i}}\big(y_1(x) - y_2(x)\big) = \sin(x),$$

ein Fundamentalsystem über $\mathbb{R}$ ist also gegeben durch $\cos(x)$, $\sin(x)$ und e^x.

Bemerkung 15.13. Nach dem Fundamentalsatz der Algebra faktorisiert jedes komplexe normierte Polynom in Linearfaktoren, allerdings können Nullstellen öfters auftreten, also

$$p(x) = (x - \lambda_1)^{k_1}(x - \lambda_2)^{k_2} \cdot \ldots \cdot (x - \lambda_r)^{k_r}$$

für paarweise verschiedene $\lambda_j \in \mathbb{C}$, wobei $k_j \in \mathbb{N}$ mit $\sum_{j=1}^r k_j = \deg(p)$.[53] Da wir nur Polynome betrachten, bei denen der Vorfaktor von x^n gleich 1 ist, haben wir $a = 1$. Wir haben dann also Lösungen für

$$0 = p(D)y = \big((D - \lambda_1)^{k_1}(D - \lambda_2)^{k_2} \cdot \ldots \cdot (D - \lambda_r)^{k_r}\big)y,$$

falls $(D - \lambda_j)^{k_j} y = 0$, also insbesondere falls $(D - \lambda_j)y = 0$, was der Lösung $\mathrm{e}^{\lambda_j x}$ entspricht. Allerdings brauchen wir nicht nur eine, sondern k_j linear unabhängige Lösungen von $(D - \lambda_j)^{k_j} y = 0$.

Lemma 15.14. Sei $I \subseteq \mathbb{R}$ ein Intervall, $\lambda \in \mathbb{C}$, $k \in \mathbb{N}_0$ und $f \in \mathcal{C}^\infty(I)$. Dann gilt

$$(D - \lambda)^k\big(f(x)\mathrm{e}^{\lambda x}\big) = f^{(k)}(x)\mathrm{e}^{\lambda x}.$$

Beweis. Wir zeigen die Aussage mit vollständiger Induktion. Für $k = 0$ ist nichts zu zeigen. Wenn die Aussage für k gilt, folgt

$$\begin{aligned}
(D - \lambda)^{k+1}\big(f(x)\mathrm{e}^{\lambda x}\big) &= (D - \lambda)(D - \lambda)^k\big(f(x)\mathrm{e}^{\lambda x}\big) \\
&= (D - \lambda)\big(f^{(k)}(x)\mathrm{e}^{\lambda x}\big) \\
&= D\big(f^{(k)}(x)\mathrm{e}^{\lambda x}\big) - \lambda f^{(k)}(x)\mathrm{e}^{\lambda x} \\
&= f^{(k+1)}(x)\mathrm{e}^{\lambda x} + f^{(k)}(x)\lambda\mathrm{e}^{\lambda x} - \lambda f^{(k)}(x)\mathrm{e}^{\lambda x} \\
&= f^{(k+1)}(x)\mathrm{e}^{\lambda x}.
\end{aligned}$$

□

Somit haben wir für jedes λ_j die Lösungen

$$\mathrm{e}^{\lambda_j x}, \quad x\mathrm{e}^{\lambda_j x}, \quad x^2\mathrm{e}^{\lambda_j x}, \quad \ldots \quad x^{k_j - 1}\mathrm{e}^{\lambda_j x}.$$

Wir müssen nur noch zeigen, dass diese linear unabhängig sind.

[53] Um Verwechslung mit dem Gradienten vorzubeugen, schreiben wir den Grad eines Polynoms p als $\deg(p)$ vom englischen *degree*.

Lemma 15.15. Seien $p \in \mathbb{C}[x]$ und $\lambda \in \mathbb{C}$ mit $p(\lambda) \neq 0$. Ist $g \in \mathbb{C}[x]$ mit $\deg(g) = k \in \mathbb{N}_0$, dann gilt

$$p(D)\big(g(x)\mathrm{e}^{\lambda x}\big) = h(x)\mathrm{e}^{\lambda x},$$

wobei $h \in \mathbb{C}[x]$ mit $\deg(h) = k$.

Beweis. Die Taylorreihe von p um den Punkt λ ist

$$p(x) = \sum_{j=0}^{k} c_j (x-\lambda)^j$$

für geeignete $c_j \in \mathbb{C}$. Insbesondere gilt $c_0 = p(\lambda) \neq 0$ und laut Lemma 15.14

$$p(D)\big(g(x)\mathrm{e}^{\lambda x}\big) = \sum_{j=0}^{k} c_j (D-\lambda)^j \big(g(x)\mathrm{e}^{\lambda x}\big) = \underbrace{\sum_{j=0}^{k} c_j g^{(j)}(x)}_{=:h(x)} \cdot \mathrm{e}^{\lambda x}$$

Wegen $c_0 \neq 0$ gilt $\deg(h) = \deg(g)$. $\square$

Satz 15.16. Sei $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{C}[x]$ mit Linearfaktorisierung $p(x) = (x-\lambda_1)^{k_1} \cdot \ldots \cdot (x-\lambda_r)^{k_r}$. Dann besitzt die Differentialgleichung

$$p(D)y = y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + y_0 y = 0$$

das Lösungsfundamentalsystem

$$y_{jm}(x) = x^m \mathrm{e}^{\lambda_j x} \quad \text{wobei} \quad 1 \leq j \leq r \quad \text{und} \quad 0 \leq m \leq k_j - 1.$$

Beweis. (1) Zunächst ist jedes y_{jm} eine Lösung der Differentialgleichung, da

$$p(D)y_{jm} = \frac{p(D)}{(D-\lambda_j)^{k_j}} \cdot (D-\lambda_j)^{k_j}\big(x^m \mathrm{e}^{\lambda_j x}\big) = \frac{p(D)}{(D-\lambda_j)^{k_j}} \mathrm{e}^{\lambda_j x} \underbrace{\frac{\mathrm{d}^{k_j} x^m}{\mathrm{d}x^{k_j}}}_{=0} = 0$$

(2) Weiterhin sind die y_{jm} linear unabhängig. Dafür reicht es, folgendes zu zeigen: Sind $g_j \in \mathbb{C}[x]$ mit $\deg(g_j) < k_j$ so, dass $\sum_{j=1}^{r} g_j(x)\mathrm{e}^{\lambda_j x} = 0$, so folgt $g_j = 0$ für alle $j = 1, \ldots, r$. Dies beweisen wir durch Induktion nach $r \in \mathbb{N}$:

Im Falle $r = 1$ folgt aus $g(x)\mathrm{e}^{\lambda x} = 0$ sofort $g(x) = 0$, da die Exponentialfunktion nur echt positive Werte annehmen kann. Angenommen, die Aussage gilt für r und sei $\sum_{j=1}^{r+1} g_j(x)\mathrm{e}^{\lambda_j x} = 0$ für $g_j \in \mathbb{C}[x]$ mit $\deg(g_j) < k_j$. Dann folgt

$$\begin{aligned}
0 &= (D-\lambda_{r+1})^{k_{r+1}} \sum_{j=1}^{r+1} g_j(x)\mathrm{e}^{\lambda_j x} \\
&= \sum_{j=1}^{r} (D-\lambda_{r+1})^{k_{r+1}} g_j(x)\mathrm{e}^{\lambda_j x} + (D-\lambda_{r+1})^{k_{r+1}} g_{r+1}(x)\mathrm{e}^{\lambda_{r+1} x} \\
&\overset{15.14}{=} \sum_{j=1}^{r} (D-\lambda_{r+1})^{k_{r+1}} g_j(x)\mathrm{e}^{\lambda_j x} + \underbrace{\big(g_{r+1}(x)\big)^{(k_{r+1})}}_{=0 \text{ da } \deg(g_{r+1})<k_{j+1}} \mathrm{e}^{\lambda_{r+1} x} \\
&= \sum_{j=1}^{r} (D-\lambda_{r+1})^{k_{r+1}} g_j(x)\mathrm{e}^{\lambda_j x} \\
&= \sum_{j=1}^{r} h_j(x)\mathrm{e}^{\lambda_j x}
\end{aligned}$$

laut Lemma 15.15 für geeignete $h_j \in \mathbb{C}[x]$ mit $\deg(h_j) = \deg(g_j)$. Nach Induktionsvoraussetzung folgt $h_j = 0$ für alle $j = 1, \dots, r$. Da die g_j den gleichen Grad wie die h_j haben müssen, folgt $g_j = 0$ für alle $j = 1, \dots, r$. Also haben wir noch

$$0 = \sum_{j=1}^{r+1} g_j(x)\mathrm{e}^{\lambda_j x} = g_{r+1}\mathrm{e}^{\lambda_{r+1}x}$$

übrig, wie im Fall $r = 1$ folgt schließlich $g_{r+1} = 0$. Insgesamt gilt also $g_j = 0$ für alle $j = 1, \dots, r+1$. $\square$

Beispiel 15.17. (1) Die Differentialgleichung $y'' - 2y' + y = 0$ entspricht dem Polynom $p(x) = x^2 - 2x + 1 = (x-1)^2$. Somit bilden $y_1(x) = \mathrm{e}^x$ und $y_2(x) = x\mathrm{e}^x$ ein Lösungsfundamentalsystem.

(2) Die Differentialgleichung $y'''' + 8y'' + 16y = 0$ entspricht dem Polynom

$$p(x) = x^4 + 8x^2 + 16 = (x^2+4)^2 = \big((x-2\mathrm{i})(x+2\mathrm{i})\big)^2,$$

also $\lambda_1 = 2\mathrm{i}$ und $\lambda_2 = -2\mathrm{i}$ mit Vielfachheiten $k_1 = 2$ und $k_2 = 2$. Über $\mathbb{C}$ erhalten wir das Lösungsfundamentalsystem

$$y_{10}(x) = \mathrm{e}^{2\mathrm{i}x}, \quad y_{11}(x) = x\mathrm{e}^{2\mathrm{i}x}, \quad y_{20}(x) = \mathrm{e}^{-2\mathrm{i}x} \quad \text{und} \quad y_{21}(x) = x\mathrm{e}^{-2\mathrm{i}x}.$$

Über $\mathbb{R}$ können wir durch Linearkombinationen folgendes Fundamentalsystem bilden:

$$\begin{aligned}
\cos(2x) &= \frac{1}{2}\big(y_{10}(x) + y_{20}(x)\big),\\
\sin(2x) &= \frac{1}{2\mathrm{i}}\big(y_{10}(x) - y_{20}(x)\big),\\
x\cos(2x) &= \frac{1}{2}\big(y_{11}(x) + y_{21}(x)\big),\\
x\sin(2x) &= \frac{1}{2\mathrm{i}}\big(y_{11}(x) - y_{21}(x)\big).
\end{aligned}$$

Für diese Linearkombinationen brauchen wir, dass die nicht-reellen Nullstellen in komplex-konjugierten Paaren (also $\lambda_1 = \overline{\lambda_2}$) mit gleicher Multiplizität (also $k_1 = k_2$) auftreten. Dies gilt aber für alle Polynome mit reellen Koeffizienten!

Bemerkung 15.18. Sei $p(x) \in \mathbb{R}[x]$ und $\lambda \in \mathbb{C} \setminus \mathbb{R}$ eine k-fache Nullstelle von p. Dann ist auch $\overline{\lambda}$ eine k-fache Nullstelle von p und es gilt

$$\overline{y_{\lambda j}(x)} = \overline{x^j\mathrm{e}^{\lambda x}} = x^j\overline{\mathrm{e}^{\lambda x}} = x^j\mathrm{e}^{\overline{\lambda x}} = x^j\mathrm{e}^{\overline{\lambda}x} = y_{\overline{\lambda}j}(x)$$

für alle $j = 0, \dots, k-1$. Ist $\lambda = a + ib$ mit $a, b \in \mathbb{R}$ und $b \neq 0$, so erhält man aus den komplexen Lösungen $y_{\lambda j}(x)$ und $y_{\overline{\lambda}j}(x)$ von $p(D)y = 0$ also immer die (linear unabhängigen) reellen Lösungen

$$\frac{1}{2}\big(y_{\lambda j}(x) + y_{\overline{\lambda}j}(x)\big) = \frac{1}{2}\Big(y_{\lambda j}(x) + \overline{y_{\lambda j}(x)}\Big) = \mathrm{Re}(y_{\lambda j}(x)) = x^j\mathrm{e}^{ax}\cos(bx)$$

und

$$\frac{1}{2i}\big(y_{\lambda j}(x) - y_{\overline{\lambda}j}(x)\big) = \frac{1}{2i}\Big(y_{\lambda j}(x) - \overline{y_{\lambda j}(x)}\Big) = \mathrm{Im}(y_{\lambda j}(x)) = x^j\mathrm{e}^{ax}\sin(bx).$$

Bemerkung 15.19. Wir kommen nun auf den Zusammenhang zwischen linearen Differentialgleichungen höherer Ordnung und linearen Differentialgleichungssystemen zurück, der in Bemerkung 15.8 schon angedeutet wurde.

(1) Im Fall konstanter Koeffizienten haben wir jeweils die folgenden allgemeinen Aussagen:

	DGL n-ter Ordnung $p(D)y = 0$	System von n DGLs 1. Ordnung $z' = Az$
gegeben durch:	$p(x) \in \mathbb{C}[x]$ Polynom vom Grad n	$A \in \mathbb{C}^{n \times n}$ $n \times n$-Matrix
Lösungen bestimmt durch:	Nullstellen $\lambda_1, \dots, \lambda_r$ von p mit Multiplizitäten $k_1, \dots, k_r$	Eigenwerte $\mu_1, \dots, \mu_q$ von A mit Größen $\ell_1, \dots, \ell_q$ der Jordan-Kästchen J_j
Form der Lösungen:	$x^m e^{\lambda x}$ mit $0 \leq m \leq k_j - 1$ für jedes $\lambda = \lambda_j$	$e^{\mu x} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x^m/m! \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ mit $0 \leq m \leq \ell_j - 1$ für jedes Jordan-Kästchen J_j und $\mu = \mu_j$
Unterschied:	die $\lambda_1, \dots, \lambda_r$ sind paarweise verschieden	die $\mu_1, \dots, \mu_q$ können gleich sein, aber zu verschiedenen J_j gehören

(2) Wir betrachten nun eine Differentialgleichung n-ter Ordnung $p(D)y = 0$, wobei $p(x) = \sum_{j=0}^{n} a_j x^j \in \mathbb{C}[x]$ mit $a_n = 1$. Das zugehörige System ist dann gegeben durch $z' = A_p z$, wobei

$$A_p = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

die sogenannte *Begleitmatrix* von p ist. Der Vergleich unserer beiden Lösungstheorien suggeriert dann:

(i) Die Nullstellen von p sind die Eigenwerte von A_p.

(ii) Die Jordan-Form von A_p hat die spezielle Eigenschaft, dass zu jedem Eigenwert von A_p genau ein Jordan-Kästchen existiert.

Wir wollen im Folgenden diese beiden Tatsachen überprüfen.

(3) Die Eigenwerte von A_p sind durch die Nullstellen des charakteristischen Polynoms

$$q(\lambda) = \det(\lambda I_n - A_p)$$

gegeben. Im Fall $n = 4$ erhält man durch Entwicklung nach der letzten Zeile

$$
\begin{aligned}
q(\lambda) &= \det \begin{pmatrix} \lambda & -1 & 0 & 0 \\ 0 & \lambda & -1 & 0 \\ 0 & 0 & \lambda & -1 \\ a_0 & a_1 & a_2 & \lambda + a_3 \end{pmatrix} \\
&= -a_0 \cdot \underbrace{\det \begin{pmatrix} -1 & 0 & 0 \\ \lambda & -1 & 0 \\ 0 & \lambda & -1 \end{pmatrix}}_{=-1} + a_1 \cdot \underbrace{\det \begin{pmatrix} \lambda & 0 & 0 \\ 0 & -1 & 0 \\ 0 & \lambda & -1 \end{pmatrix}}_{=\lambda} \\
&\quad - a_2 \cdot \underbrace{\det \begin{pmatrix} \lambda & -1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{=-\lambda^2} + (\lambda + a_3) \cdot \underbrace{\det \begin{pmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ 0 & 0 & \lambda \end{pmatrix}}_{=\lambda^3} \\
&= a_0 + a_1\lambda + a_2\lambda^2 + a_3\lambda^3 + \lambda^4 = p(\lambda).
\end{aligned}
$$

Für allgemeine n zeigt man die Aussage per Induktion.

(4) Für (ii) betrachten wir zunächst den Fall $n = 2$, also $y'' + a_1 y' + y_0 y = 0$ und $p(x) = x^2 + a_1 x + a_0$, also

$$
A_p = \begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix}.
$$

Für eine 2×2-Matrix gibt es allgemein genau die folgenden drei Möglichkeiten für die Jordan-Form:

Fall 1: $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \operatorname{diag}(\lambda_1, \lambda_2)$ mit $\lambda_1 \neq \lambda_2$.

Fall 2: $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \operatorname{diag}(\lambda, \lambda)$.

Fall 3: $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = J(\lambda, 2)$.

Wir müssen zeigen, dass Fall 2 für A_p nicht eintreten kann! Es gilt

$$
0 = p(\lambda) = \det(\lambda I_2 - A_p) = \lambda^2 + a_1\lambda + a_0
$$

genau dann, wenn $\lambda = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_0}$. Fall 2 kann nur auftreten bei einem Eigenwert λ mit Multiplizität 2, also bei $\frac{a_1^2}{4} - a_0 = 0$; dann gilt

$$
A_p = \begin{pmatrix} 0 & 1 \\ -\frac{a_1^2}{4} & -a_1 \end{pmatrix}.
$$

Für $a_1 = 0$ hat man direkt die Jordan-Form

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

also Fall 3 mit $\lambda = 0$. Für $a_1 \neq 0$ sind die Transformationsmatrizen gegeben durch

$$
S = \begin{pmatrix} \frac{a_1}{2} & 1 \\ -\frac{a_1^2}{4} & 0 \end{pmatrix} \quad \text{und} \quad S^{-1} = \begin{pmatrix} 0 & -\frac{4}{a_1^2} \\ 1 & \frac{2}{a_1} \end{pmatrix},
$$

also gilt

$$S^{-1}A_pS = \begin{pmatrix} -\frac{a_1}{2} & 1 \\ 0 & -\frac{a_1}{2} \end{pmatrix},$$

wir sind also im Fall 3 mit $\lambda = -\frac{a_1}{2}$: Ein Jordan-Kästchen der Größe 2. Somit kann Fall 2 also nicht auftreten.

(5) Für größere n ist eine solche Fallunterscheidung nicht praktikabel. Mit etwas mehr linearer Algebra kann man über das Minimalpolynom $m(x)$ von A_p argumentieren.[54] Die Theorie besagt, dass p ein Vielfaches von m ist und alle Linearfaktoren von p in m vorkommen müssen. Ist also

$$p(x) = (x-\lambda_1)^{k_1} \cdot \ldots \cdot (x-\lambda_r)^{k_r}$$

die Faktorisierung von p, so gilt

$$m(x) = (x-\lambda_1)^{t_1} \cdot \ldots \cdot (x-\lambda_r)^{t_r} \quad \text{wobei} \quad 1 \leq t_j \leq k_j.$$

Weiterhin ist t_j gerade die Größe des größten Jordan-Kästchens zu λ_j in der Jordan-Normalform von A_p. In unserer Situation reicht es also zu zeigen, dass $p(x) = m(x)$, denn dann gilt $k_j = t_j$ für alle j und somit gibt es in der Jordan-Normalform von A_p genau ein Jordan-Kästchen pro Eigenwert.

Wir zeigen $p(x) = m(x)$ wieder für den Fall $n = 4$: Mit

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 \end{pmatrix}$$

gilt $A^0 = I_4$,

$$A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 \\ a_0a_3 & a_1a_3 - a_0 & a_2a_3 - a_1 & a_3^2 - a_2 \end{pmatrix}$$

und

$$A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 \\ a_0a_3 & a_1a_3 - a_0 & a_2a_3 - a_1 & a_3^2 - a_2 \\ a_0a_2 - a_0a_3^2 & a_0a_3 + a_1a_2 - a_1a_3^2 & a_1a_3 + a_2^2 - a_2a_3^2 - a_0 & 2a_2a_3 - a_3^3 - a_1 \end{pmatrix}.$$

Wir konzentrieren uns jeweils nur auf die erste Zeile: Es gilt $e_1^T A^j = e_{j+1}^T$ für alle $j = 0, \ldots, 3$. Ist also $g(x) = b_3x^3 + b_2x^2 + b_1x + b_0 \in \mathbb{C}[x]$ ein Polynom mit $g(A) = 0$, so gilt

$$0 = e_1^T g(A) = e_1^T \sum_{j=0}^{3} b_j A^j = \sum_{j=0}^{3} b_j e_1^T A^j = \sum_{j=0}^{3} b_j e_{j+1}^T = \begin{pmatrix} b_0 & b_1 & b_2 & b_3 \end{pmatrix}.$$

Folglich muss schon $g = 0$ sein. Es gibt also außer dem Nullpolynom kein Polynom vom Grad höchstens 3, welches p als Nullstelle hat. Somit muss $m(x) = p(x)$ sein.

[54] Das *Minimalpolynom* von A ist das (eindeutig bestimmte) normierte Polynom kleinsten Grades, das A als "Nullstelle" hat (und nicht das Nullpolynom ist).

A. Aussagen aus der Linearen Algebra

Bemerkung. Die Aussagen in diesem Anhang sind nur für den Körper der reellen Zahlen $\mathbb{R}$ formuliert, sie gelten sinngemäß aber über jedem Körper und insbesondere auch über den komplexen Zahlen $\mathbb{C}$.

Matrizen. Eine (reelle) $m \times n$-Matrix ist eine Tabelle

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

mit m Zeilen, n Spalten und $m \cdot n$ Einträgen $a_{i,j} \in \mathbb{R}$. Wir schreiben $\mathbb{R}^{m\times n}$ für die Menge aller $m \times n$-Matrizen.

Sind $A, B \in \mathbb{R}^{m\times n}$ und $\lambda \in \mathbb{R}$, so definieren wir

$$A + B := (a_{ij} + b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \quad \text{und} \quad \lambda \cdot A := (\lambda \cdot a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}.$$

Mit diesen Operationen wird $\mathbb{R}^{m\times n}$ zu einem $\mathbb{R}$-Vektorraum.

Matrixprodukt. Sind $A \in \mathbb{R}^{m\times n}$ und $B \in \mathbb{R}^{k\times m}$, so definieren wir

$$BA := B \cdot A := \left(\sum_{\ell=1}^{m} a_{i\ell} b_{\ell j} \right)_{\substack{i=1,\dots,k \\ j=1,\dots,n}} \in \mathbb{R}^{k\times n}.$$

Wichtig: Das Produkt BA ist nur definiert, wenn die Anzahl der Spalten von B mit der Anzahl der Zeilen von A übereinstimmt. Selbst wenn sowohl AB als auch BA existieren, gilt im Allgemeinen $AB \neq BA$.

Lineare Abbildungen. Eine Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ heißt linear, falls

$$L(x_1 + x_2) = L(x_1) + L(x_2) \quad \text{für alle} \quad x_1, x_2 \in \mathbb{R}^n$$

und

$$L(\lambda x) = \lambda L(x) \quad \text{für alle} \quad x \in \mathbb{R}^n \quad \text{und alle} \quad \lambda \in \mathbb{R}.$$

Wir schreiben $L(\mathbb{R}^n, \mathbb{R}^m)$ für die Menge aller linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$. Sind $L, L_1, L_2 \in L(\mathbb{R}^n, \mathbb{R}^m)$ und $\lambda \in \mathbb{R}$, so definieren wir $L_1 + L_2, \lambda L \in L(\mathbb{R}^n, \mathbb{R}^m)$ mittels

$$(L_1 + L_2)(x) := L_1(x) + L_2(x) \quad \text{und} \quad (\lambda L)(x) := \lambda \cdot L(x) \quad \text{für alle} \quad x \in \mathbb{R}^n.$$

Mit diesen Operationen wird $L(\mathbb{R}^n, \mathbb{R}^m)$ zu einem $\mathbb{R}$-Vektorraum.

Identifikation. Jede Matrix $A \in \mathbb{R}^{m\times n}$ induziert eine lineare Abbildung

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto L_A(x) := Ax.$$

Ist umgekehrt $L \in L(\mathbb{R}^n, \mathbb{R}^m)$, so setzt man

$$A_L := \big(L(e_1) \quad \dots \quad L(e_n)\big) \in \mathbb{R}^{m\times n},$$

wobei $e_i \in \mathbb{R}^n$ der i-te Standardbasisvektor von $\mathbb{R}^n$ ist, also der Vektor, der in der i-ten Komponente den Wert 1 und sonst überall den Wert 0 hat. Die Matrix A_L heißt die darstellende Matrix von L bezüglich der Standardbasen.

Man zeigt leicht, dass diese beiden Konstruktionen zueinander invers sind. Somit gibt es eine Bijektion zwischen $\mathbb{R}^{m\times n}$ und $L(\mathbb{R}^n, \mathbb{R}^m)$, die sogar mit den Operationen verträglich ist (genauer: ein $\mathbb{R}$-Vektorraum-Isomorphismus), so dass wir häufig eine lineare Abbildung mit ihrer darstellenden Matrix identifizieren. Insbesondere entspricht die Komposition $K \circ L$ von $L \in L(\mathbb{R}^n, \mathbb{R}^m)$ und $K \in L(\mathbb{R}^m, \mathbb{R}^k)$ gerade der Matrix $A_K A_L$.

Inverse Matrizen. Die darstellende Matrix der Identitätsabbildung $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto x$ ist die Einheitsmatrix

$$I := I_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & \dots e_n \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt invertierbar, wenn eine Matrix $B \in \mathbb{R}^{n\times n}$ mit $AB = BA = I_n$ existiert. In diesem Fall schreiben wir $B = A^{-1}$; die Matrix A^{-1} heißt Inverse von A und ist eindeutig bestimmt. Man kann leicht zeigen, dass A genau dann invertierbar ist, wenn die lineare Abbildung $L_A : \mathbb{R}^n \to \mathbb{R}^n$ bijektiv ist.

Determinante. In der Linearen Algebra konstruiert man eine Abbildung $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$. Diese sogenannte Determinante von A ist eine Polynomfunktion in den Einträgen von A. Eine Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann invertierbar, wenn $\det(A) \neq 0$. Weiterhin gibt es immer eine Matrix $\tilde{A} \in \mathbb{R}^{n\times n}$ mit

$$A\tilde{A} = I_n \cdot \det(A) = \tilde{A}A,$$

falls also $\det(A) \neq 0$ gilt, ist $A^{-1} := \frac{1}{\det(A)} \cdot \tilde{A}$.

Für 2×2-Matrizen ist etwa

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

und es gilt

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \text{falls} \quad ad - bc \neq 0.$$

Sind $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$ und $D \in \mathbb{R}^{m\times m}$, so gilt

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det(A)\det(D) = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix}.$$

Eigenwerte und Eigenvektoren. Sei $A \in \mathbb{R}^{n\times n}$. Ein Vektor $x \in \mathbb{R}^n \setminus \{0\}$ heißt Eigenvektor zum Eigenwert $\lambda \in \mathbb{R}$, falls

$$Ax = \lambda x.$$

Die Eigenwerte von A sind genau die (möglicherweise komplexen) Nullstellen des charakteristischen Polynoms $\chi_A(\lambda) := \det(\lambda I_n - A)$, insbesondere gibt es (mit Vielfachheiten gezählt) genau n (möglicherweise komplexe) Eigenwerte von A. Das Produkt aller Eigenwerte von A ist gerade $\det(A)$.

Sind $\mathbb{R}^{n\times n}$ und $B \in \mathbb{R}^{m\times m}$, so ist die Menge der Eigenwerte von

$$\operatorname{diag}(A, B) := \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}$$

gerade die Vereinigung der Mengen der Eigenwerte von A und von B.

Ist zum Beispiel

$$\operatorname{diag}(\lambda_1, \ldots, \lambda_n) := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \in \mathbb{R}^{n\times n}$$

eine Diagonalmatrix, dann ist e_j ein Eigenvektor zum Eigenwert λ_j.

Eigenraum. Sei λ ein Eigenwert von A. Der Eigenraum von A zu λ ist die Menge aller Eigenvektoren von A zu λ inklusive dem Nullvektor. Er ist gegeben als Lösung des homogenen Gleichungssystems $(\lambda I_n - A)x = 0$.

Untervektorraum. Ein Untervektorraum von $\mathbb{R}^n$ ist eine nichtleere Teilmenge $U \subset \mathbb{R}^n$, so dass für alle $x, y \in U$ und alle $\lambda \in \mathbb{R}$ auch $x + y \in U$ und $\lambda x \in U$ gilt.

Erzeugendensystem. Für jede endliche Menge $\{x_1, \ldots, x_k\} \subseteq \mathbb{R}^n$ ist

$$U := \left\{ \sum_{j=1}^{k} \lambda_j x_j \mid \lambda_j \in \mathbb{R} \right\}$$

der von $\{x_1, \ldots, x_n\}$ erzeugte Untervektorraum von $\mathbb{R}^n$. Man nennt $\{x_1, \ldots, x_n\}$ ein (endliches) Erzeugendensystem von U.

Lineare Unabhängigkeit. Eine Menge $\{x_1, \ldots, x_k\} \subseteq \mathbb{R}^n$ heißt linear unabhängig, falls für alle $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ gilt:

$$\sum_{j=1}^{k} \lambda_j x_j = 0 \quad \Rightarrow \quad \lambda_1 = \ldots = \lambda_k = 0.$$

Basis und Dimension. Eine Basis eines Untervektorraums U von $\mathbb{R}^n$ ist ein linear unabhängiges Erzeugendensystem von U. Je zwei Basen von U haben die selbe Kardinalität. Die Anzahl der Elemente einer Basis von U heißt Dimension von U, geschrieben $\dim(U)$. Es gilt $\dim(\mathbb{R}^n) = n$.

Diagonalisierbarkeit. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt diagonalisierbar, wenn es eine invertierbare Matrix $S \in \mathbb{R}^{n\times n}$ gibt, so dass $S^{-1}AS$ eine Diagonalmatrix ist. Diagonalisierbarkeit von A ist äquivalent dazu, dass es eine Basis von $\mathbb{R}^n$ gibt, die nur aus Eigenvektoren von A besteht.

Dramatis personae

Index