Analysis II
Sommersemester 2015

Prof. Dr. Annette Huber-Klawitter

Fassung vom 20. Juli 2015

Dies ist ein Vorlesungsskript und kein Lehrbuch.
Mit Fehlern muss gerechnet werden!

Math. Institut
Eckerstr. 1
79104 Freiburg

0761-888 5495
annette.huber@math.uni-freiburg.de

Einleitung

Im ersten Semester haben wir behandelt:

(i) Die reellen Zahlen und ihre Eigenschaften.

(ii) Grenzwerte von Folgen

(iii) Stetigkeit von Funktionen

(iv) Differenzierbarkeit

(v) Integrierbarkeit

(vi) Reihenentwicklungen

Dieses Semester wiederholen wir große Teile dieses Programms, diesmal für Funktionen

$$f : \mathbb{R}^n \to \mathbb{R}^m.$$

Ein neuer Aspekt ist die Behandlung von Differentialgleichungen.

Die Differentialrechnung in mehreren Variablen ist die Mathemetik zur klassischen Mechanik, in der man Fragen behandelt wie die Beschreibung von Bewegungen von Objekten unter der Wirkung von verschiedenen Kräften und unter zusätzlichen Zwangsbedingungen. Ein Beispiel ist eine Achterbahnfahrt. Man spricht vom Lagrange- oder vom Hamiltonformalismus.

Der Stoff gehört auch zum Rüstzeug aller Ingenieursstudiengänge, aber der Wirtschaftswissenschaften. Dort wird natürlich auf einem niedrigeren formalen Niveau gearbeitet.

Im einzelnen werden wir diese Aspekte ansprechen:

(i) Topologie des $\mathbb{R}^n$

(ii) Differenzierbarkeit in mehreren Variablen

(iii) Anwendungen: Extremwertprobleme, der Satz über implizite Funktionen, Taylor-Reihen

(iv) Der Existenz- und Eindeutigkeitssatz für gewöhnliche Differentialgleichungen

(v) Lineare Differentialgleichungen

Literatur: Es gibt eine Vielzahl von möglichen Lehrbüchern mit Titeln wie Analysis 2 oder Infinitesimalrechnung 2. Ich empfehle dringend, neben dem Skript auch mit wenigstens einem Lehrbuch zu arbeiten. Voraussichtlich wird die Vorlesung sich eng an das Buch von Forster anlehnen.

Kapitel 9

Die Topologie des $\mathbb{R}^n$

Metrische Räume

Definition 9.1. *Ein* metrischer Raum *ist eine Menge X zusammen mit einer* Metrik *d, d.h. einer Abbildung*

$$d : X \times X \to \mathbb{R}$$

so dass gilt

(i) $d(x, y) = 0$ *genau dann, wenn* $x = y$.

(ii) Symmetrie: $d(x, y) = d(y, x)$ *für alle* $x, y \in X$.

(iii) Dreiecksungleichung: Für alle $x, y, z \in X$ *gilt*

$$d(x, z) \leq d(x, y) + d(y, z).$$

Wir nennen $d(x, y)$ den Abstand *von x und y bezüglich der Metrik d.*

Lemma 9.2. *Sei (X, d) metrischer Raum. Dann gilt $d(x, y) \geq 0$ für alle $x, y \in X$.*

Beweis: Es gilt

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y).$$

□

Beispiel. (i) $\mathbb{R}$ mit $d(x, y) = |x - y|$ ist ein metrischer Raum.

(ii) Jede Menge X wird metrischer Raum mit der *trivialen Metrik*

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

(iii) $\mathbb{R}^n$ mit $d(x,y) = \|x-y\|_p$ (vergleiche Kapitel 8). Dieses Beispiel werden wir noch viel ausführlicher behandeln.

(iv) Sei (X,d) metrischer Raum, $Y \subset X$ Teilmenge. Dann ist $(Y, d|_{Y\times Y})$ ein metrische Raum.

Sobald wir einen Abstandsbegriff haben, können wir die Grenzwertdefinitionen der Analysis 1 wiederholen und auch die entsprechenden Sätze wieder beweisen. Wir beginnen damit.

Definition 9.3. *Sei (X,d) ein metrischer Raum.*

(i) Für $x_0 \in X$, $r > 0$ heißt $B(x_0,r) = \{x \in X | d(x,x_0) < r\}$ offene Kugel um x_0 mit Radius r *oder auch* r-Umgebung von x_0.

(ii) Sei $x_0 \in X$. Eine Teilmenge $U \subset X$ heißt Umgebung *von x_0, wenn es ein $\varepsilon > 0$ gibt mit $B(x_0,\varepsilon) \subset U$.*

(iii) Eine Teilmenge $U \subset X$ heißt offen, *wenn sie Umgebung jedes Elementes $x \in U$ ist, d.h. sie enhält eine offene Kugel um jeden ihrer Punkte.*

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

Beispiel. Für $X = \mathbb{R}$ ist $B(x_0,r) = (x_0 - r, x_0 + r)$. Die Mengen $(-1,1)$, $[-1,1)$ und $[-1,1]$ sind Umgebungen von 0. Die Menge $(-1,0]$ ist keine Umgebung von 0. Offene Intervalle (a,b) sind offen im Sinne der obigen Definition, abgeschlossene Intervalle $[a,b]$ sind abgeschlossen.

Satz 9.4. *Sei (X,d) ein metrischer Raum. Dann gilt:*

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

(ii) Sind $U, V \subset X$ offen, so auch $U \cap V$.

(iii) Ist $(U_i)_{i\in I}$ eine Familie von offenen Teilmengen von X, so ist auch $\bigcup_{i\in I} U_i$ offen.

(iv) (Hausdorffsches Trennungsaxiom) Für $x, y \in X$ mit $x \neq y$ gibt es Umgebungen U von x und V von y mit $U \cap V = \emptyset$.

Beweis: Die Aussage für $\emptyset$ ist trivial. Sei $x \in X$ beliebig. Dann enthält X die Kugel $B(x,1)$. Daher ist X offen.

Seien U, V offen, $x \in U \cap V$. Nach Voraussetzung gibt es $\varepsilon, \delta > 0$ mit $B(x,\varepsilon) \subset U$ und $B(x,\delta) \subset V$. Sei $\eta = \min(\varepsilon,\delta)$. Dann ist $B(x,\eta)$ in beiden Kugeln enthalten, also auch in U und in V, also in $U \cap V$.

Schließlich seien U_i für $i \in I$ offen und $x \in \bigcup_{i\in I} U_i$. Dann gibt es $j \in I$ mit $x \in U_i$. Nach Definition gibt es $\varepsilon > 0$ mit $B(x,\varepsilon) \subset U_i$. Dann gilt auch $B(x,\varepsilon) \subset \bigcup_{i\in I} U_i$. Damit ist die Vereinigung offen.

Seien nun $x, y \in X$, $x \neq y$. Sei $r = d(x, y)$. Es gilt $r > 0$. Wir setzen $U = B(x, r/2)$ und $V = B(y, r/2)$. Wir zeigen, dass $U \cap V$ leer ist. Sei also $z \in U \cap V$. Dann gilt

$$d(x, y) \leq d(x, z) + d(z, y) < r/2 + r/2 = r = d(x, y).$$

Dies ist ein Widerspruch. □

Korollar 9.5. *(i) X und $\emptyset$ sind abgeschlossen.*

(ii) Sind A und B abgeschlossene Teilmengen von X, so ist $A \cup B$ abgeschlossen.

(iii) Sei $(A_i)_{i \in I}$ ein System von abgeschlossenen Teilmengen von X. Dann ist $\bigcap_{i \in I} A_i$ abgeschlossen.

Beweis: Diese Aussagen folgen sofort aus den Aussagen für ihre offene Komplemente. □

Beispiel. In $X = \mathbb{R}$ gilt

$$\bigcap_{\varepsilon > 0} (-\varepsilon, 1 + \varepsilon) = [0, 1].$$

Schnitte von unendlich vielen offenen Menge sind im allgemeinen nicht offen.

Grenzwerte

Definition 9.6. *Sei X ein metrischer Raum. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge von Punkten aus X. Die Folge heißt* konvergent gegen $x \in X$

$$x = \lim_{n \to \infty} x_n$$

wenn es zu jeder Umgebung U von x ein $N \in \mathbb{N}$ gibt, so dass $x_n \in U$ für alle $n \geq N$.

Nach Definition enthält jede Umgebung eine offene ε-Umgebung für ein geeignetes ε. Es genügt daher in der Definition nur ε-Umgebungen zu betrachten. Die Folge ist also genau dann konvergent, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $d(x_n, x) < \varepsilon$ für alle $n \geq N$.

Beispiel. Eine Folge von $\mathbb{R}$ ist genau dann konvergent im obigen Sinn, wenn sie konvergent ist im Sinne der Analysis 1.

Lemma 9.7. *Sei (X, d) metrischer Raum, $(x_n)_{n \in \mathbb{N}}$ konvergent. Dann ist der Grenzwert eindeutig.*

Beweis: Seien x, y beide Grenzwerte der Folge. Sie $\varepsilon > 0$. Nach Definition gibt es $N \in \mathbb{N}$, so dass $d(x_n, x) < \varepsilon$ und $d(x_n, y) < \varepsilon$ für alle $n \geq N$. Es folgt

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < 2\varepsilon.$$

Dies gilt für alle $\varepsilon > 0$, also ist $d(x, y) = 0$. Nach den Axiomen eines metrischen Raums ist dann $x = y$. □

Satz 9.8. *Sei X ein metrischer Raum. Eine Teilmenge $A \subset X$ ist genau dann abgeschlossen, wenn für jede konvergente Folge $(x_n)_{n\in\mathbb{N}}$ in A gilt $\lim_{n\to\infty} x_n \in A$.*

Beispiel. Sei $X = \mathbb{R}$, $(-\infty, b]$. Der Satz besagt: Ist $(x_n)_{n\in\mathbb{N}}$ konvergente Folge in $\mathbb{R}$ mit $x_n \leq b$, so gilt $\lim_{n\to\infty} x_n \leq b$.

Beweis: Sei A abgeschlossen, $(x_n)_{n\in\mathbb{N}}$ konvergente Folge in A mit Grenzwert $x \in X$. Angenommen, $x \notin A$. Dann liegt x in der offenen Menge $X \smallsetminus A$. Nach Definition gibt es eine offene Umgebung U von x, die ganz in $X \smallsetminus A$ enthalten ist. Da die Folge gegen x konvergiert, gibt es ein N, so dass $x_N \in U$. Dies ist ein Widerspruch zu $x_N \in A$.

Sei umgekehrt $A \subset X$ ein Teilmenge, die die Grenzwertbedingung erfüllt. Angenommen, A ist nicht abgeschlossen. Dann ist $X \smallsetminus A$ nicht offen. Also gibt es $y \in X \smallsetminus A$, so dass keine ε-Umgebung $B(y, \varepsilon)$ in $X \smallsetminus A$ enthalten ist. Wir betrachten $\varepsilon = 1/n$. Da $B(y, 1/n)$ nicht in $X \smallsetminus A$ enthalten ist, hat $B(y, 1/n)$ einen nicht-leeren Schnitt mit A. Sei $x_n \in A \cap B(y, 1/n)$. Die Folge $(x_n)_{n\in\mathbb{N}}$ liegt in A und konvergiert gegen y. Dies ist ein Widerspruch. □

Definition 9.9. *Sei (X, d) ein metrischer Raum. Eine Folge $(x_n)_{n\in\mathbb{N}}$ heißt* Cauchy-Folge, *wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt mit $d(x_n, x_m) < \varepsilon$ für alle $n, m \geq N$.*

Bemerkung. Jede konvergente Folge ist eine Cauchy-Folge.

Beweis: Wie in der Analysis 1. Sei $(x_n)_{n\in\mathbb{N}}$ konvergent gegen x. Zu $\varepsilon > 0$ gibt es also N, so dass $d(x_n, x) < \varepsilon/2$ für alle $n \geq N$. Es folgt

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

□

Definition 9.10. *Ein metrischer Raum heißt* vollständig, *wenn jede Cauchy-Folge konvergiert.*

Beispiel. $\mathbb{R}$ ist vollständig, $\mathbb{Q}$ nicht.

Definition 9.11. *Seien X, Y metrische Räume und $f : X \to Y$ eine Abbildung. Die Abbildung heißt* stetig in $x \in X$, *falls*

$$\lim_{n\to\infty} f(x_n) = f(x)$$

für jede Folge $(x_n)_{n\in\mathbb{N}}$ in X mit Grenzwert x.

Die Abbildung heißt stetig, *wenn sie in allen Punkten stetig ist.*

Satz 9.12. *Seien X, Y, Z metrische Räume und $f : X \to Y$ und $g : Y \to Z$ Abbildungen. Ist f stetig in x und g stetig in $f(x)$, so ist $g \circ f$ stetig in x.*

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ Folge in X mit Grenzwert x. Nach Voraussetzung ist $\lim_{n\to\infty} f(x_n) = f(x)$. Also ist $(f(x_n))_{n\in\mathbb{N}}$ eine Folge in Y mit Grenzuwert $f(x)$. Nach Voraussetzung ist $\lim_{n\to\infty} g(f(x_n)) = g(f(x))$. Da dies für alle Folgen gilt, ist $g \circ f$ stetig in x. $\square$

Satz 9.13 (ε-δ-Kriterium)**.** *Seien X, Y metrische Räume, $f : X \to Y$ eine Abbildung. Sie ist genau dann stetig in $x \in X$, wenn gilt: für jede $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass $d(f(x), f(x')) < \varepsilon$ für alle $x' \in X$ mit $d(x, x') < \delta$.*

Im Fall $X = \mathbb{R}$ ist dies wörtlich das Kriterium aus der Analysis 1. Auch der Beweis ist derselbe.

Beweis: Sei f stetig in y. Angnommen, das ε-δ-Kriterium ist nicht erfüllt. Dann gibt es ein $\varepsilon > 0$, so dass für alle $\delta > 0$ ein x' die Bedingung verletzt. Wir betrachten dieses ε und $\delta = 1/n$. Sei x_n ein Punkt mit $d(x, x_n) < 1/n$ und $d(f(x_n), f(x)) > \varepsilon$. Die Folge der $(x_n)_{n\in\mathbb{N}}$ konvergiert gegen x, die Folge der $(f(x_n))_{n\in\mathbb{N}}$ jedoch nicht gegen $f(x)$. Dies ist ein Widerspruch zur Stetigkeit.
Sei umgekehrt das ε-δ-Kriterium erfüllt. Wir betrachten eine Folge $(x_n)_{n\in\mathbb{N}}$ mit Grenzwert y. Wir wollen zeigen, dass $\lim_{n\to\infty} f(x_n) = f(y)$. Sei also $\varepsilon > 0$. Sei $\delta > 0$ die zugehörige Zahl aus dem Kriterium. Wegen $\lim_{n\to\infty} x_n = x$ gibt es ein $N \in \mathbb{N}$, so dass $d(x_n, x) < \delta$ für alle $n \geq N$. Nach Wahl von δ folgt dann $d(f(x_n), f(x)) < \varepsilon$ für $n \geq N$. Das heißt, die Folge der Bilder konvergiert. $\square$

Korollar 9.14. *Seien X, Y metrische Räume, $f : X \to Y$ eine Abbildung. Die Abbildung f ist genau dann stetig, wenn $f^{-1}(U)$ offen ist für jede offene Teilmenge $U \subset Y$.*

Beweis: Sei f stetig, $U \subset Y$ offen, $x \in f^{-1}U$, $y = f(x) \in U$. Dann gibt es $\varepsilon > 0$ mit $B(y, \varepsilon) \subset U$. Nach dem ε-δ-Kriterium gibt es $\delta > 0$, so dass alle Elemente von $B(x, \delta)$ nach $B(y, \varepsilon)$ abgebildet werden. D.h. $B(x, \delta) \subset f^{-1}U$. Daher ist $f^{-1}U$ offen.
Sei umgekehrt das Urbild jeder offenen Menge offen. Wir überprüfen Stetigkeit in $x \in X$ nach dem ε-δ-Kriterium. Sei also $\varepsilon > 0$. Die Menge $U = B(f(x), \varepsilon)$ ist offen. Nach Voraussetzung ist dann auch $f^{-1}U$ offen, enthält also eine offene δ-Umgebung von x. Dies ist das gesuchte δ. $\square$

Normierte Räume

Wir diskutieren nun eine große Klasse von metrischen Räumen, darunter auch unser wichtigstes Beispiel $\mathbb{R}^n$.

Definition 9.15. *Sei V ein reeller Vektorraum. Eine* Norm *auf V ist eine Abbildung*

$$|\cdot| : V \to \mathbb{R}$$

mit

(i) (Definitheit)$\|x\| = 0$ genau dann, wenn $x = 0$.

(ii) $\|\lambda x\| = |\lambda|\|x\|$ für alle $\lambda \in \mathbb{R}$, $x \in V$.

(iii) (Dreiecksungleichung) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in V$.

Insbesondere ist $\|x\| = \|-x\|$ und wegen der Dreiecksungleichung

$$\|0\| \leq \|x\| + \|-x\| = 2\|x\| \Rightarrow \|x\| \geq 0.$$

Beispiel. (i) Der Absolutbetrag ist eine Norm auf $\mathbb{R}$ oder $\mathbb{C}$.

(ii) Für jedes $p \geq 1$ ist auf $\mathbb{R}^n$ wird eine Norm definiert durch

$$\|(x_1, \ldots, x_n)\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}.$$

(Beweis Analysis 1, Minkowskische Ungleichung Satz 8.6)

(iii) Für $p = \infty$ setzen wir die Maximumsnorm

$$\|(x_1, \ldots, x_n)\|_\infty = \max |x_i|$$

(iv) Auf $V = C([a, b], \mathbb{R})$ dem Vektorraum der stetigen Funktionen auf $[a, b]$ definieren wir für $p \geq 1$

$$\|f\|_p = \sqrt[p]{\int_a^b |f(x)|^p \mathrm{d}x}.$$

(Beweis Übungsaufgabe).

(v) Sei ℓ^p die Menge $\mathbb{C}$-wertigen Folgen $(z_n)_{n\in\mathbb{N}}$ mit

$$\|(x_n)_{n\in\mathbb{N}}\|_p = \sqrt[p]{\sum_{i=1}^{\infty} |z_i|^p} < \infty.$$

Dies ist ein normierter Vektorraum.

(vi) Auf V der Raum der $\mathbb{C}$-wertigen Riemann-integrierbaren $2\pi i$-periodischen Funktionen definieren wir

$$\|f\|_2 = \int_0^{2\pi i} \|f(x)\|_2 \mathrm{d}x.$$

Dies ist nicht ganz eine Norm, da es periodische Funktionen gibt mit $\|f\|_2 = 0$, aber $f = 0$. Sei also $V_0 = \{f \in V | \|f\|_2\}$. Wir erhalten dann eine Norm auf V/V_0. Der Hauptsatz über Fourierreihen (Analysis 1, Theorem 7.8) kann fomuliert werden als die Aussage, dass es eine Inklusion $V/V_0 \to \ell^2$ gibt. Hierbei wird eine periodische Funktion mit der Folge ihrer Fourier-Koeffizienten identifiziert.

Lemma 9.16. *Jeder normierte Vektorraum ist ein metrischer Raum mit*

$$d(x, y) = \|x - y\|.$$

Alle Definitionen über Grenzwerte etc. lassen sich also jetzt auf den $\mathbb{R}^n$ anwenden. Unser nächstes Ziel ist es zu zeigen, dass $\mathbb{R}^n$ vollständig ist. A priori hängt diese Aussage von der Wahl der Norm ab. Tatsächlich ist dies aber gar nicht der Fall. Wie man in Definition 9.6 sieht, hängt die Konvergenz von Folgen nicht von der Metrik, sondern nur von dem System der offenen Umgebungen ab!

Satz 9.17. *Seien $p, q \in [1, \infty]$. Dann gibt es $C, C' > 0$, so dass für alle $x \in \mathbb{R}^n$ gilt*

$$C'\|x\|_q \leq \|x\|_p \leq C\|x\|_q.$$

Wir sagen, die Normen $\|x\|_p$ und $\|x\|_q$ sind *äquivalent.*

Beweis: Es genügt, $q = \infty$ zu betrachten. Es gilt

$$\|(x_1, \ldots, x_n)\|_p \leq \sqrt[p]{n \max_i |x_i|^p} = \max_i |x_i| = \sqrt[p]{n}\|(x_1, \ldots, x_n)\|_\infty.$$

Also können wir $C = \sqrt[p]{n}$ wählen.
Umgekehrt ist für $i = 1, \ldots, n$

$$|x_i| = \sqrt[p]{|x_i|^p} \leq \|(x_1, \ldots, x_n)\|_p$$

also gilt dies aus für $\max_i |x_i|$. wir können $C' = 1$ wählen. □

Tatsächlich gilt sogar noch ein stärkerer Satz: Zwei *beliebige* Normen auf $\mathbb{R}^n$ sind äquivalent.

Korollar 9.18. *Sei $p \geq 1$. Eine Teilmenge $U \subset \mathbb{R}^n$ ist genau dann offen bezüglich der p-Norm, wenn sie offen bezüglich der Maximumsnorm ist.*

Beweis: Sei U offen bezüglich der p-Norm, $x \in U$. Dann gibt es $\varepsilon > 0$ mit $B(x, \varepsilon) \subset U$, d.h.

$$\|x' - x\|_p < \varepsilon \Rightarrow x' \in U.$$

Wegen der Abschätzung der Normen gibt es ε', so dass

$$\|x' - x\|_\infty < \varepsilon \Rightarrow \|x' - x\|_p < \varepsilon.$$

Also enthält U eine offene ε'-Kugel bezüglich $\|\cdot\|_\infty$. Damit ist U offen bezüglich der Maximumnorm. Das gleiche Argument zeigt auch die Rückrichtung. □

Satz 9.19. *$\mathbb{R}^m$ ist vollständig bezüglich der Norm $\|\cdot\|_p$ für alle $p \geq 1$. Eine Folge $(x_n)_{n\geq 1}$, $x_n = (x_n^1, \dots, x_n^m)$ ist genau dann eine Cauchy-Folge, wenn jede der Folgen $(x_n^i)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Es gilt*

$$\lim_{n\to\infty} x_n = (\lim_{n\to\infty} x_n^1, \dots, \lim_{n\to\infty} x_n^m).$$

Grenzwerte können komponentenweise berechnet werden.

Beweis: Wegen der Äquivalenz der Normen ist eine Folge genau dann auch Cauchy-Folge/konvergent bezüglich der p-Norm, wenn sie eine Cauchy-Folge/konvergent ist bezüglich der Maximumsnorm. Wir betrachten daher nur die Maximumsnorm.

Sei $(x_n)_{n\to\infty}$ eine Cauchy-Folge. Wegen $|x_n^i - x_{n'}^i| \leq \|x_n\|_\infty$ ist die Komponentenfolgen eine Cauchy-Folge. Nach Analysis 1 ist sie konvergent, d.h. $\mathbb{R}$ ist vollständig. Sei y^i der Grenzwert. Sei $y = (y^1, \dots, y^m)$. Wir zeigen nun $y = \lim_{n\to\infty}$. (Dann ist $\mathbb{R}^n$ auch vollständig.) Sei $\varepsilon > 0$. Für jedes i gibt es ein N_i so dass $|x_n^i - y^i| < \varepsilon$ für alle $n \geq N_i$. Sei N das Maximum der N_i. Dann gilt die Abschätzung für alle i, also auch für $\|x_n - y\|_\infty$, wie zu zeigen war. □

Nun wollen wir stetige Abbildungen zwischen $\mathbb{R}^n$s betrachten.

Beispiel. Addition und Multiplikation sind stetige Abbildungen $\mathbb{R}^2 \to \mathbb{R}$. Divison ist eine stetige Abbildung $\mathbb{R} \times \mathbb{R}^* \to \mathbb{R}$.

Beweis: Sei $((x_n, y_n))_{n\in\mathbb{N}}$ eine konvergente Folge, also $\lim_{n\to\infty} x_n = x$, $\lim_{n\to\infty} y_n = y$. Nach den Grenzwertsätzen (Analysis 1) konvergiert dann die Summenfolge $(x_n + y_n)_{n\in\mathbb{N}}$ gegen $x + y$. Also ist $+$ stetig. Das Argument für die Multiplikation und Division ist dasselbe. □

Korollar 9.20. *Sei X ein metrischer Raum. Eine Abbildung $f = (f_1, \dots, f_n) : X \to \mathbb{R}^n$ ist genau dann stetig, wenn jede Komponentenabbildung $f_i : X \to \mathbb{R}$ stetig ist.*

Beweis: Sei $\lim_{n\to\infty} x_n = x$ in X. Wenn f stetig ist, so gilt $\lim_{n\to\infty} f(x_n) = f(x)$. Nach dem Satz ist dies äquivalent dazu, dass jede Komponentenfolge $f_i(x_n)$ gegen $f_i(x)$ konvergiert. Also ist jedes f_i stetig. Ebenso folgt die Umkehrung. □

Korollar 9.21. *Seien $f, g : X \to \mathbb{R}$ stetige Funktionen. Dann sind auch $f + g$ und fg stetig. Gilt $g(x) \neq 0$ für alle x, so ist auch f/g stetig.*

Beweis: Wir schreiben $f + g$ als die Komposition

$$X \to \mathbb{R}^2 \xrightarrow{+} \mathbb{R}$$

wobei die erste Abbildung $x \mapsto (f(x), g(x))$ ist. Wir testen die Stetigkeit der ersten Abbildung mit dem Kriterium aus Korollar 9.20. Sie ist stetig, da f und g

stetig sind. Als Komposition von stetigen Abbildungen ist $f+g$ stetig. Dasselbe Argument funktioniert für die Mulitplikation.

Für die Division beachten wir noch, dass eine Abbildung $f : X \to \mathbb{R}^*$ genau dann stetig ist, wenn die Komposition $X \to \mathbb{R}^* \to \mathbb{R}$ stetig ist. □

Beispiel. Alle Polynomfunktionen

$$f : \mathbb{R}^n \to \mathbb{R}$$

mit

$$f(x_1, \dots, x_n) = \sum_{i_1,\dots,i_n=0}^{N} a_{i_1,\dots,i_n} x_1^{i_1} \dots x_n^{i_n}$$

sind stetig.

Unsere zweite Klasse von normierten Vektorräumen waren Funktionenräume. Konvergenz bezüglich $\|\cdot\|_\infty$ ist gerade die gleichmäßige Konvergenz. Konvergenz bezüglich $\|\cdot\|_2$ ist die Konvergenz im quadratischen Mittel aus Kapitel 7 (Fourierreihen).

Besonders interessant ist die Klasse der linearen Abbildungen, die es immer zwischen normierten Vektorräumen gibt.

Lineare Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$ sind stets stetig, die ist ein Spezialfall einer Polynomfunktion.

Satz 9.22. *Seien V, W normierte Vektorräume, $f : V \to W$ linear. Die Abbildung ist genau dann stetig, wenn es eine Konstante $C > 0$ gibt, so dass*

$$\|f(x)\|_W \leq C\|x\|_V \text{ für alle } x \in V.$$

Beispiel. Die Identität $\mathbb{R}^n \to \mathbb{R}^n$ ist stetig, wenn wir $\mathbb{R}^n$ einmal mit der p-Norm und einmal mit der q-Norm versehen. Dies ist die Aussage von Satz 9.17.

Beweis: Sei f stetig. Wir nutzen die Stetigkeit in 0 aus. Zu $\varepsilon = 1$ gibt es ein $\delta > 0$, so dass

$$\|x\|_V < \delta \Rightarrow \|f(x)\|_W < 1.$$

Wir setzen $C = 2/\delta$. Sei nun $v \in V$ beliebig, $v \neq 0$, $\lambda = (C\|v\|_V)^{-1}$, $z = \lambda v$. Dann gilt

$$\|z\| = |\lambda|\|v\| = \frac{\delta}{2\|v\|}\|v\| = \frac{\delta}{2}.$$

Also folgt $\|f(z)\| < 1$, d.h.

$$\|f(z)\| = \|f(\lambda x)\| = \|\lambda f(x)\| = |\lambda|\|f(x)\| = \frac{1}{C\|x\|}\|f(x)\| < 1.$$

Hieraus folgt die gewünschte Ungleichung.

Sei umgekehrt C eine Konstante wie im Satz. Dann gilt

$$\|f(x) - f(x')\| = \|f(x - x')\| \leq C\|x - x'\|$$

für alle x, x'. Hieraus folgt mit dem ε-δ-Kriterium die Stetikeit. □

Abbildungen zwischen unendlich dimensionalen normierten Räumen sind nicht immer stetig.

Beispiel. Sei $V = C^1([0,1],\mathbb{R})$ der Raum der stetig differenzierbaren Funktionen und $W = C([0,1],\mathbb{R})$ der Raum der stetigen Funktionen. Wir versehen beide mit der Supremumsnorm. Sei $D : V \to W$ die Ableitung, also $Df = f'$. Diese ist nicht stetig.

Beweis: Wir betrachten $f_n = x^n \in V$ und $Df_n = f_n' = nx^{n-1}$. Dann ist $\|f_n\| = 1$ für alle $n \in \mathbb{N}$, aber $\|Df_n\| = n$ divergiert gegen Unendlich. $\square$

Definition 9.23. *Seien V, W normierte Vektorräume, $f : V \to W$ eine stetige lineare Abbildung. Wir setzen*

$$\|f\| = \sup\{\|f(x)\| | x \in V, \|x\| = 1\}.$$

Dieses Supremum existiert, da $\|f(x)\|$ durch die Konstante C des Kriteriums beschränkt ist.

Topologische Räume

Wir führen eine weitere Abstraktionsstufe ein.

Definition 9.24. *Ein* topologischer Raum *ist eine Menge X zusammen mit einer Menge $\mathcal{T}$ von offenen Teilmengen von X, so dass gilt*

(i) $X\mathcal{T}$ und $\emptyset \in \mathcal{T}$.

(ii) Sind $U, V \in \mathcal{T}$, dann auch $U \cap V$.

(iii) Ist $\{U_i\}_{i\in I}$ ein System von Elemente auf $\mathcal{T}$, dann ist auch $\bigcup_{i\in I} U_i$ ein Element von $\mathcal{T}$.

Elemente von $\mathcal{T}$ heißen offen. *Eine Menge $A \subset X$ heißt* abgeschlossen, *wenn $X \smallsetminus A$ offen ist.*
Sei $x \in X$. Eine Umgebung *von x ist eine Teilmenge $V \subset X$, die eine offene Menge U enthält, in der x liegt.*
Ein topologischer Raum heißt hausdorff, *wenn es zu je zwei Punkten $x, y \in X$, $x \neq$ Umgebungen U von x und V von y mit $U \cap V = \emptyset$ gibt.*

Beispiel. Sei (X, d) ein metrischer Raum. Dann ist X ein topologischer Raum mit der Topologie aus Definition 9.3. Dies ist der Inhatl von Satz **??**. Dort wurde auch verifiziert, dass er die Hausdorff-Eigenschaft erfüllt.

Es gibt lange Liste mit gruseligen Beispielen für topologische Räume. Dies gehört nicht wirklich hierher, daher belassen wir es bei den beiden wichtigsten.

Beispiel. (i) Sei X eine Menge. Die *triviale Topologie* hat nur die offenen Mengen X und $\emptyset$.

(ii) Sei X eine Menge. Die *diskrete Topologie* hat alle Teilmengen als offenen Mengen, d.h. $\mathcal{T} = \mathcal{P}(X)$, die Potenzmenge. Tatsächlich ist dies die Topologie, die von der trivialen Metrik induziert wird.

Definition 9.25. *Sei (X, d) ein topologischer Raum, $S \subset X$ eine Teilmenge.*

(i) Der Abschluss $\overline{S}$ *von S ist die kleinste abgeschlossene Teilmenge von X, die S enthält, d.h. der Schnitt aller abgeschlossenen Teilmengen, die S enthalten.*

(ii) Das Innere $\mathring{S}$ *ist die größte offene Teilmenge von X, die in S enthalten ist, d.h. die Vereinigung aller offenen Teilmengen von X, die in S enthalten sind.*

(iii) Der Rand ∂S *ist definiert als $\overline{S} \setminus \mathring{S}$.*

Der Rand ist abgeschlossen, da er als Schnitt der abgeschlossenen Mengen $\overline{S}$ und $X \setminus \mathring{S}$ geschrieben werden kann.

Beispiel. Sei $X = \mathbb{R}$. Der Abschluss von $[a, b)$ ist $[a, b]$, das Innere ist (a, b) und der Rand $\{a, b\}$.

Lemma 9.26. *Sei X ein topologischer Raum, $S \subset X$ eine Teilmenge. Es gilt*

$$\overline{S} = X \setminus (X \setminus S)^\circ.$$

Der Rand von S besteht aus den Punkten $x \in X$, so dass jede Umgebung U von x nicht-leeren Schnitt mit S und mit $X \setminus S$ hat. Es gilt

$$\overline{S} = S \cup \partial S, \mathring{S} = S \setminus \partial S \cap S$$

Beweis: Die Menge $(X \setminus S)^\circ$ ist offen und enthalten in $X \setminus S$. Daher ist ihr Komplement abgeschlossen und enthält S. Dies zeigt $\overline{S} \subset X \setminus (X \setminus S)^\circ$. Umgekehrt ist $\overline{S}$ abgeschlossen, also $X \setminus \overline{S}$ offen und enthalten in $X \setminus S$. Es folgt

$$(X \setminus S)^\circ \supset X \setminus \overline{S} \Rightarrow X \setminus (X \setminus S)^\circ \subset \overline{S}.$$

Wenn es eine Umgebung von s gibt, die ganz in S liegt, so ist $s \in \mathring{S}$.
Wenn es eine Umgebung von s gibt, die ganz in $X \setminus S$ liegt, so ist $s \in X \setminus \overline{S}$.
Wenn es beides nicht gibt, so muss s ein Randpunkt sein. $\square$

Alternativ können wir den Abschluss etc. in Termen von Häufungspunkten charakterisieren.

Definition 9.27. *Sei X ein metrischer Raum, $S \subset X$. Ein Punkt $x \in X$ heißt* Häufungspunkt *von S, wenn es eine konvergente Folge $(x_n)_{n \in \mathbb{N}}$ gibt von Elementen $x_n \in S$ mit Grenzwert x.*

Lemma 9.28. *Sei X ein metrischer Raum, $S \subset X$ eine Teilmenge. Der Abschluss $\overline{S}$ besteht aus allen Häufungspunkten von S.*

Beweis: Nach Satz 9.8 enthält die abgeschlossene Menge als Grenzwerte konvergenter Folgen in S.
Sei umgekehrt $x \in \overline{S}$. Dann liegt keine Umgebung von x ganz in $X \setminus S$. Insbesondere gibt es in jedem $B(x, 1/n)$ ein $x_n \in S$. Diese Folge konvergiert gegen x. $\square$

Definition 9.29. *Seien X, Y topologische Räume. Eine Abbildung $f : X \to Y$ heißt* stetig, *wenn Urbilder offener Mengen offen sind.*

Nach Korollar 9.14 stimmt dies im Fall von metrischen Räumen mit der ursprünglichen Definition überein.

Kapitel 10

Kompaktheit

Der wichtigste Satz über stetige Funktionen war der Zwischenwertsatz: Ist f stetig, so ist das Bild eines abgeschlossenen, beschränkten Intervalls wieder ein abgeschlossenes, beschränktes Intervall. Dies wollen wir verallgemeinern.

Definition 10.1. *Ein topologischer Raum heißt* folgenkompakt*, wenn jede Folge eine konvergente Teilfolge besitzt.*

Lemma 10.2. *Sei $A \subset \mathbb{R}^m$ abgeschlossen und beschränkt. Dann ist A ist folgenkompakt.*

Eine Teilmenge $B \subset X$ in einem metrischen Raum heißt *beschränkt*, wenn $B \subset B(x, r)$ für ein $x \in X$, $r > 0$.

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in A. Wir beginnen mit dem Fall $n = 1$. Die Folge ist nach Voraussetzung beschränkt. Wir wissen aus Analysis 1, dass sie eine konvergente Teilfolge besitzt, mit einem Grenzwert in A. Da A abgeschlossen ist, liegt der Grenzwert bereits in A.

Sei nun $m > 1$. Konvergenz von Folgen hängt nicht von der Norm ab, sondern nur von der induzierten Topologie. Wir arbeiten mit der Maximumsnorm $\|\cdot\|_\infty$ auf $\mathbb{R}^n$. Die ersten Koordinaten bilden eine beschränkt Teilfolge in $\mathbb{R}$. Also gibt es eine Teilfolge, die in $\mathbb{R}$ konvergiert. Wir ersetzen $(x_n)_{n\in\mathbb{N}}$ durch diese Teilfolge. Danach betrachten wir die Folge der zweiten Koordinaten. Sie sind beschränkt, also gibt es eine konvergente Teilfolge. Wieder ersetzen wir die Folge durch die konvergente Teilfolge. Nach m Schritten haben wir eine Teilfolge erreicht, deren Koordinatenfolgen sämtlich konvergieren. Nach Satz 9.19 konvergiert die Folge in $\mathbb{R}^m$. Da A abgeschlossen ist, liegt der Grenzwert bereits in A. $\square$

Wesentlich wichtiger ist eine andere Formulierung der Eigenschaft, die in allen topologischen Räumen funktioniert.

Definition 10.3. *Sei X ein topologischer Raum, $A \subset X$ eine Teilmenge. Eine* Überdeckung von *A ist eine Familie $(U_i)_{i\in I}$ von offenen Teilmengen von X, so dass $A \subset \bigcup_{i\in I} U_i$.*

Beispiel. $X = A = \mathbb{R}$ wird überdeckt durch die offenen Intervalle $(x, x+1)$ für $x \in \mathbb{R}$.

Definition 10.4. *Sei X ein topologischer Raum, $A \subset X$ eine Teilmenge. Die Menge heißt* kompakt, *wenn jede offene Überdeckung $(U_i)_{i\in I}$ von A eine endliche Teilüberdeckung hat, d.h. es gibt eine endliche Teilmenge $F \subset I$, so dass $\bigcup_{i\in F} U_i \supset A$.*

Beispiel. Die Menge $(-1,1)$ ist nicht kompakt, denn die Überdeckung $((-x,x))_{x\in(-1,1)}$ hat keine endliche Teilüberdeckung. Sind $x_1, \dots, x_n$ endliche viele Indizes, so ist

$$\bigcup_{i=1}^{n}(-x_i, x_i) = (-x, x)$$

mit $x = \max\{x_1, \dots, x_n\}$. Die Kompaktheitsbedingung ist verletzt.

Unser Ziel ist es zu zeigen, dass die kompakten Teilmengen des $\mathbb{R}^n$ gerade die beschränkten abgeschlossenen Teilmengen sind. Dies erfordert mehr Vorarbeit. Wenigstens ein positives Beispiel wollen wir sofort sehen.

Beispiel. Sei X ein metrischer Raum, $(x_n)_{n\in\mathbb{N}}$ konvergente Folge mit Grenzwert x. Dann ist die Menge

$$A = \{x, x_1, x_2, \dots\}$$

kompakt.

Beweis: Sei $(U_i)_{i\in I}$ eine offene Überdeckung von A. Es gibt also i_0 mit $x \in U_{i_0}$. Da U_{i_0} offen ist, gibt es eine Kugel $B(x,\varepsilon) \subset U_{i_0}$. Da die Folge konvergiert, gibt es eine natürliche Zahl N mit $x_n \in B(x,\varepsilon)$ für alle $n \geq N$. Insbesondere gilt $x_n \in U_{i_0}$ für $n \geq N$. Nach Voraussetzung gibt es für $n < N$ ein $i_n \in I$ mit $x_n \in U_{i_n}$. Dann ist $J = \{i_0, \dots, i_{n-1}\}$ die gesuchte endliche Indexmenge. □

Satz 10.5 (Bolzano-Weierstraß)**.** *Sei X metrischer Raum, $A \subset X$ kompakt. Dann ist A folgenkompakt.*

Beweis: Angenommen, A ist nicht folgenkompakt. Dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$, die keine konvergente Teilfolge hat. Dann besitzt jeder Punkt $x \in A$ eine Umgebung U_x, die nur endlich viele Folgenglieder enthält. (Andernfalls könnte man eine gegen x konvergente Teilfolge konstruieren). Die Famile der U_x überdeckt A. Da A kompakt ist, wird es von nur endlich vielen $U_{x_1}, \dots, U_{x_N}$ überdeckt. Diese enthalten nur endlich viele Folgenglieder. Dies ist ein Widerspruch. □

Bemerkung. Die Umkehrung ist ebenfalls wahr. Als Korollar erhalten wir dann sofort, dass die beschränkten abgeschlossenen Teilmengen des $\mathbb{R}^m$ kompakt sind. Für einen Beweis der Rückrichtung verweisen wir z.B. auf das Skript von Wang zur Funktionalanalysis. Uns führt es etwas zu weit ab von unseren eigentlichen Zielen, daher führen wir einen direkten Beweis. Dies benötigt noch etwas Vorbereitung.

Satz 10.6. *Sei X metrischer Raum, $A \subset X$ kompakt. Dann ist A abgeschlossen und beschränkt.*

Beweis: Wir zeigen die Beschränktheit. Sei $x \in X$ ein beliebiger, fester Punkt. Dann ist $(B(x,r))_{r>0}$ eine offene Überdeckung von X, also auch von A. Da A kompakt ist, gibt es eine endliche Teilüberdeckung, also eine endliche Menge von Radien $r_1, \dots, r_n$, so dass

$$A \subset \bigcup_{i=1}^{n} B(x, r_i).$$

Sei $r = \max\{r_1, \dots, r_n\}$. Dann ist $A \subset B(x,r)$, also A beschränkt.
Wir zeigen die Abgeschlossenheit, d.h. $X \smallsetminus A$ offen. Sei $x \in X \smallsetminus A$. Für $r > 0$ setzen wir

$$U_r = \{y \in X | d(x,y) > r\}.$$

Das System der U_r ist eine Überdeckung von $X \smallsetminus \{x\}$, also auch eine Überdeckung von A. Da A kompakt ist, reicht eine endliche Teilüberdeckung zu Radien $r_1, \dots, r_n$. Sei $r = \min\{r_1, \dots, r_n\}$. Dann ist $A \subset U_r$ und daher $B(x,r) \cap A = \emptyset$. Dies ist die gesuchte offene Umgebung von x in $X \smallsetminus A$. $\square$

Korollar 10.7. *Jede konvergente Folge in einem metrischen Raum ist beschränkt.*

Beweis: Sie ist kompakt, also beschränkt. $\square$

Satz 10.8. *Sei X metrischer Raum, $K \subset X$ kompakt. Sei $A \subset K$ abgeschlossen. Dann ist auch A kompakt.*

Beweis: Sei $(U_i)_{i \in I}$ eine offene Überdeckung von A. Die Menge $X \smallsetminus A$ ist offen. Nimmt man sie hinzu, so erhält man eine offene Überdeckung von K. Da K kompakt ist, gibt es endliche viele Inidizes $i_1, \dots, i_n \imath I$, so dass

$$K \subset (X \smallsetminus A) \cup U_{i_1} \cup \dots \cup U_{i_n}.$$

Dies ist dann auch eine Überdeckung von A. Dann ist sogar $(U_{i_j})_{j=1,\dots,n}$ eine endliche offene Überdeckung von A. $\square$

Satz 10.9. *Seien X, Y metrische Räume, $f : X \to Y$ stetig. Ist $K \subset X$ kompakt, so ist $f(K) \subset Y$ kompakt.*

Beweis: Sei $(U_i)_{i \in I}$ eine offene Überdeckung in Y. Wir betrachten $V_i = f^{-1}U_i$. Diese Menge ist offen, weil f stetig ist. Das System $(V_i)_{i \in I}$ ist eine offene Überdeckung von K. Da K kompakt ist gibt es eine endliche Menge $F \subset I$, so dass $(V_i)_{i \in F}$ eine Überdeckung von K ist. Dann ist $(U_i)_{i \in F}$ eine Überdeckung von $f(K)$, da $ff^{-1}(U_i) = U_i$. $\square$

Korollar 10.10. *Sei X kompakter metrischer Raum, $f : X \to \mathbb{R}$ stetig. Dann ist f beschränkt und nimmt ihr Minmum und Maximum an, d.h. es gibt $x, y \in X$ mit*

$$f(x) = \sup\{f(x')|x' \in X\}$$
$$f(y) = \inf\{f(y')|y' \in X\}$$

Beweis: Nach dem letzten Satz ist $f(X) \subset \mathbb{R}$ kompakt, also beschränkt und abgeschlossen. Da Supremum und Infimum Häufungspunkte von $f(X)$ sind, liegen sie in $f(X)$. □

Definition 10.11. *Sei X metrischer Raum, $A \subset X$ Teilmenge, $x \in X$. Wir definieren den* Abstand *von x von A als*

$$d(x, A) = \inf\{d(x, x')|x' \in A\}.$$

Sei $B \subset X$ eine weitere Teilmenge, so definieren wir den Abstand *von A und B als*

$$d(A, B) = \inf\{d(x, A)|x \in B\} = \inf\{d(x, y)|x \in B, y \in A\}.$$

Korollar 10.12. *Ist A abgeschlossen, B kompakt mit $A \cap B = \emptyset$, so gilt*

$$d(A, B) > 0.$$

Beweis: Wir zeigen zuerst, dass die Abbildung

$$d(\cdot, A) : X \to \mathbb{R}$$

stetig ist. Sei also $x \in X$, $\varepsilon > 0$. Wir wählen $\delta = \varepsilon$. Für $x' \in X$ mit $d(x, x') < \varepsilon$ folgt dann

$$|d(x, A) - d(x', A)| < \varepsilon$$

da $d(x', A) < d(x, x') + d(x, A)$.
Ist nun B kompakt, so nimmt $d(\cdot, A)$ auf B sein Minimum an, d.h. es gibt $b \in B$ mit $d(b, A) = d(B, A)$. Da A abgeschlossen ist und $b \notin A$, gibt es $\varepsilon > 0$ mit $B(b, \varepsilon) \subset X \setminus A$. Also ist $d(b, A) \geq \varepsilon$. □

Satz 10.13. *Seien X, Y metrische Räume, X kompakt. Dann ist jede stetige Abbildung $f : X \to Y$ gleichmäßig stetig.*

Beweis: Sei $\varepsilon > 0$. Wir suchen $\delta > 0$ so dass

$$d(x, x') < \delta \Rightarrow d(f(x)f(x')) < \varepsilon$$

für alle $x, x' \in X$. Nach Voraussetzung ist f stetig. Also gibt es für jedes $x \in X$ in δ_x mit

$$d(x, x') < \delta_x \Rightarrow d(f(x)f(x')) < \varepsilon/2$$

für alle $x' \in X$. Die $B(x, \delta_x/2)$ bilden eine offene Überdeckung von X. Da X kompakt ist, gibt es endlich viele $x_1, \ldots, x_n$, so dass

$$X \subset B(x_1, \delta_{x_1/2}) \cup \cdots \cup B(x_n, \delta_{x_n}/2).$$

Wir wählen $\delta = 1/2 \min\{\delta_{x_1}, \ldots, \delta_{x_n})$. Seien nun x, x' beliebig mit $d(x, x') < \delta$. Dann gibt es i mit

$$d(x, x_i) < \delta_{x_i}/2.$$

Wegen $d(x, x') < \delta_{x_i}/2$ folgt auch

$$d(x' x_i) < \delta_{x_i}.$$

Aus der Wahl von δ_{x_i} folgt also

$$d(f(x), f(x_i)) < \varepsilon/2, d(f(x'), f(x_i)) < \varepsilon/2$$

und damit die gewünschte Abschätzung. □

Theorem 10.14 (Heine-Borel). *Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis: Die eine Richtung haben wir bereits in beliebigen metrischen Räumen gezeigt. Sei also nun A beschränkt und abgeschlossen. Da abgeschlossenen Teilmengen von kompakten Mengen kompakt sind, genügt es zu zeigen, dass jeder Quader

$$Q = [a_1, b_1] \times [a_2, b_2] \times \ldots [a_n, b_n]$$

kompakt ist. Zur Vereinfachung der Notation betrachten wir nur den Fall $a_i = 0, b_i 1$, also den Einheitswürfel.

Der Beweis wird mit vollständiger Induktion nach n geführt. Wir beginnen mit $Q = [0, 1]$. Gegeben sei eine offene Überdeckung $(U_i)_{i \in I}$ des Einheitsintervalls. Wir betrachten die Menge T aller $t \in [0, 1]$, so dass das Intervall $[0, t]$ von endlich vielen der U_i überdeckt wird. Diese Menge ist nicht leer, denn $0 \in T$ (da es sich um eine Überdeckung handelt, gibt es ein i_0 mit $0 \in U_{i_0}$). Ist $t \in T$, so liegen auch alle $t' < t$ in T (wir benutzen dieselbe endliche Überdeckung). Die Menge ist beschränkt, hat also ein Supremum t_0.

Angenommen, t_0 ist echt kleiner als 1. Nach Voraussetzung gibt es einen Index i mit $t_0 \in U_i$. Da U_i offen ist, gibt es ein Intervall $(t_0 - \varepsilon, t_0 + \varepsilon) \subset U_i$. Das Intervall $(t_0 - \varepsilon, t_0)$ enthält ein Element t' von T. Nach Definition gibt es also eine endliche Menge $i_1, \ldots, i_m$ von Indizes, so dass

$$[0, t'] \subset U_{i_1} \cup \cdots \cup U_{i_m}.$$

Dann gilt

$$[0, t_0 + \varepsilon/2] U_{i_1} \cup \cdots \cup U_{i_m} \cup U,$$

also $t_0 + \varepsilon/2 \in T$. Dies ist ein Widerspruch zur Wahl von t_0. Dies zeigt, dass $t_0 = 1$. Es gibt eine endliche Teilüberdeckung, die ganz $[0, 1]$ überdeckt.

Damit haben wir den Induktionsanfang gezeigt. Wir setzen nun voraus, dass $[0,1]^n$ kompakt ist und wollen auf $[0,1]^{n+1}$ schließen. Zur Abkürzung schreiben wir $[0,1]^n = K$. Sei wieder $(U_i)_{i\in I}$ eine offene Überdeckung von $K \times [0,1]$. Für jedes $t \in [0,1]$ ist $K\times\{t\}$ kompakt. Daher gibt es endlich viele Indizes $i_1,\dots,i_m$, so dass

$$K \times \{t\} \subset U_{i_1} \cup \dots U_{i_m} =: U.$$

Wir wollen zeigen, dass U einen ganzen "Schlauch" um $K \times \{t\}$ enthält, d.h. dass es ein $\varepsilon > 0$ gibt mit

$$K \times (t-\varepsilon, t+\varepsilon) \subset U.$$

Dafür betrachten wir für $x \in K$ die Menge der Punkte

$$A_x = \{x\} \times [0,1] \smallsetminus (U \cap \{x\} \times [0,1].$$

Sie ist abgeschlossen in der kompakten Menge $\{x\} \times [0,1]$, daher nimmt die stetige Funktion

$$d_x : A_x \to \mathbb{R}; \qquad y \mapsto d((x,t),y)$$

ihr Minimum ε_x in einem Punkt y_0 an. Da U offen ist, enthält U eine Kugel um (x,t) und daher ist $d((x,t),y) > 0$. Nun betrachten wir die Funktion

$$K \to \mathbb{R}; \qquad x \mapsto \varepsilon_x$$

Diese ist wieder stetig und nimmt daher ihr Infimum in einem Punkt x_0 an. Insbesondere ist

$$\varepsilon_x \geq \varepsilon_{x_0} > 0$$

für alle x. Dies ist der gesuchte Radius r_t für einen ganzen Schlauch um $K\times\{t\}$. Die Intervalle $(t-r_t, t+r_t)$ für $t \in [0,1]$ bilden eine offene Überdeckung der kompakten Menge $[0,1]$. Daher gibt es endlich viele Punkte $t_1,\dots,t_m$ so dass

$$[0,1] = \bigcup_{j=1}^{m} (t-r_t, t+r_t).$$

Wie wir gesehen haben gibt es für jedes j eine endliche Indexmenge I_j so dass

$$K \times (t-r_t, t+r_t) \subset \bigcup_{i\in I_j} U_i.$$

Daher überdecken die U_i für $i \in \bigcup_{j=1}^m I_j$ ganz $K \times [0,1]$. □

Kapitel 11

Differenzierbarkeit

Wir interessieren uns für Abbildungen

$$f : U \to \mathbb{R}^m$$

mit $U \subset \mathbb{R}^n$ und wollen den Begriff der Differenzierbarkeit verallgemeinern. Geht man vom 1-dimensionalen Fall aus, so gibt es zwei offensichtliche Ansätze.

Beispiel. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x, y) = \sin(x + 2y)$.
Für festes $x \in \mathbb{R}$ erhalten wir eine Funktion $\mathbb{R} \to \mathbb{R}$, nämlich $x \mapsto \sin(x + y)$. Diese ist differenzierbar mit Ableitung $\cos(x + y)$. Wir schreiben

$$\frac{\partial f}{\partial x}(x, y) = \cos(x + 2y)$$

Ebenso erhalten wir

$$\frac{\partial f}{\partial y}(x, y) = 2\cos(x + 2y)$$

Dies sind die *partiellen Ableitungen*.
Andererseits können wir uns von der Idee der besten linearen Approximation leiten lassen. Lineare Funktionen $g : \mathbb{R}^2 \to \mathbb{R}$ haben die Form $g(x, y) = ax+by+c$ für $a, b, c \in \mathbb{R}$. Unser Beispiel ist besonders einfach, denn wir können die lineare Approximation aus der Potenzreihendarstellung ablesen:

$$\begin{aligned}\sin(x+2y) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}(x+2y)^{2k+1} &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \sum_{i=0}^{2k+1} \binom{2k+1}{i} x^i (2y)^{2k+1-i} \\ &= 0 + \frac{1}{0!}(x + 2y) + \dots\end{aligned}$$

Also

$$c = 0 =, a = 1 = \frac{\partial f}{\partial x}(0, 0), b = 2 = \frac{\partial f}{\partial y}(0, 0).$$

Die lineare Abbildung $l(x, y) = x + 2y$ heißt *totale Ableitung* von f.

Totale Differenzierbarkeit

Wir erinnern uns an die lineare Algebra: Eine Abbildung

$$A : \mathbb{R}^n \to \mathbb{R}^m$$

heißt *linear*, falls

$$A(\lambda v + \mu w) = \lambda A(v) + \mu A(w)$$

für alle $v, w \in \mathbb{R}^n$ und $\lambda, \mu \in \mathbb{R}$. Jede solche Abbildung kann als Multiplikation mit einer Matrix mit n Spalten und m Zeilen geschrieben werden.

Definition 11.1. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ eine Abbildung. Sie heißt* (total) differenzierbar *in $x \in U$, falls es eine lineare Abbildung*

$$A : \mathbb{R}^n \to \mathbb{R}^m$$

gibt, so dass in einer Umgebung von x gilt

$$f(x + \xi) = f(x) + A\xi + \phi(\xi)$$

wobei ϕ eine in einer Umgebung von $0 \in \mathbb{R}^n$ definierte Funktion mit Werten in $\mathbb{R}^m$ ist mit

$$\lim_{\xi \to 0} \frac{\phi(\xi)}{\|\xi\|} = 0.$$

Wir schreiben $Df(x)$ für die lineare Abbildung A. Sie heißt totale Ableitung *von f in x.*

Da je zwei Normen äquivalent sind, ist es egal mit welcher Norm wir in dieser Definition arbeiten.

Bemerkung. Für $n = m = 1$ ist dies die Definition aus Analysis 1.

Beispiel. $f(x, y) = 3x + 2y^2$ ist total differenzierbar in 0 mit $Df(0) = 3x$. Wir berechnen

$$\phi(x, y) = f(x, y) - f(0, 0) - Df(0)(x, y) = 2y^2.$$

Wir arbeiten in der 2-Norm. Sei $(x_n, y_n)_{n \in \mathbb{N}}$ eine Nullfolge. Wir müssen zeigen, dass

$$\frac{\phi(x_n, y_n)}{\|(x, y)\|_2} = \frac{2y_n^2}{\sqrt{x_n^2 + y_n^2}}$$

eine Nullfolge ist. Wir schätzen ab

$$\frac{2y_n^2}{\sqrt{x_n^2 + y_n^2}} \leq \frac{2y_n^2}{\sqrt{y_n^2}} = 2|y_n|$$

und dies ist tatsächlich eine Nullfolge.

Lemma 11.2. *Sei $U \subset \mathbb{R}^n$ offen, $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ eine Abbildung. Dann ist f genau dann total differenzierbar in $x \in U$, wenn jede Komponentenfunktion f_i total differenzierbar in x ist.*

Beweis: Alle Rechnungen und Grenzwertübergänge geschehen komponentenweise nach Satz 9.19. □

Satz 11.3. *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ total differenzierbar in $x \in U$. Dann ist f stetig ist x.*

Beweis: Wie in der Analyis 1: Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in U mit Grenzwert x. Wir schreiben $x_n = x + \xi_n$ mit einer Nullfolge x_n. Wir berechnen

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} (f(x) + Df(x)(\xi_n) + \phi(\xi_n)) = f(x) + Df(x)0 + \lim_{n\to\infty} \phi(\xi_n)$$

da Grenzwerte mit Addition vertauschen und lineare Abbildungen stetig sind. Da ξ_n eine Nullfolge ist, gibt es $N \in \mathbb{N}$ mit $\|\xi_n\| < 1$ für $n \geq N$. Für diese n gilt auch

$$\|\phi(\xi_n)\| < \|\phi(\xi_n)\|\|\xi_n\| to 0.$$

□

Satz 11.4 (Kettenregel). *Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen $g : U \to V$, $f : V \to \mathbb{R}^k$ Abbildungen. Ist g total differenzierbar in $x \in U$ und g total differenzierbar in $y = g(x)$, so ist $f \circ g$ total differenzierbar in x mit*

$$D(f \circ g)(x) = Df(y) \circ Dg(x).$$

Im Fall $n = m = k = 1$ ist dies die vertraute Kettenregel aus Analysis 1. Auch der Beweis ist derselbe.

Beweis: Sei $A = Dg(x)$, $B = Dg(y)$ mit $y = f(x)$. Nach Voraussetzung gilt

$$g(x + \xi) = g(x) + A\xi + \phi(\xi)$$
$$f(y + \eta) = f(y) + B\eta + \psi(\eta)$$

mit

$$\lim_{\xi\to 0} \frac{\phi(\xi)}{\|\xi\|} = 0, \lim_{\eta\to 0} \frac{\psi(\eta)}{\|\eta\|} = 0.$$

Wir setzen $\eta = g(x + \xi) - g(x) = A\xi + \phi(\xi)$. Für $\xi \to 0$ strebt η gegen 0. Es folgt

$$\begin{aligned} f(g(x+\xi)) = f(g(x) + \eta) &= f(g(x)) + B(\eta) + \psi(\eta) \\ &= f(g(x)) + BA\xi + B(\phi(\xi)) + \psi(A\xi + \phi(\xi))) \end{aligned}$$

Zu zeigen bleibt

$$\lim_{\xi\to 0} \frac{1}{\|\xi\|}(B(\phi(\xi)) + \psi(A\xi + \phi(\xi))) = 0.$$

Für den ersten Summanden folgt das aus der Linearität von B und der Voraussetzung. Für den zweiten schreiben wir die Vorausetzung um zu

$$\psi(\eta) = \|\eta\|\psi_1(\eta)$$

mit $\lim_{\eta\to 0}\psi_1(\eta) = 0$, also

$$\psi(A\xi + \phi(\xi)) = \|A\xi + \phi(\xi)\|\psi_1(A\xi + \phi(\xi)$$

und daher

$$\|\psi(A\xi + \phi(\xi))\| = \|A\xi\| + \|\phi(\xi)\|\|\psi_1(A\xi + \phi(\xi)\|.$$

Wir schätzen ab

$$\|A\xi\| + \|\phi(\xi)\| \leq \|A\|\|\xi\| + K\|\xi\|$$

für eine geeignete Konstante K und ξ klein genug. Insgesamt also

$$\|\psi(A\xi + \phi(\xi))\| \leq (\|A\| + K)\|\xi\|\|\psi_1(A\xi + \phi(\xi)\|$$

Auch nach Divison durch $\|\xi\|$ strebt die rechte Seite noch gegen 0, da $A\xi+\phi(\xi) \to 0$. Dies war zu zeigen. □

Partielle Ableitungen

Zum praktischen Rechnen ist der bisherige Standpunkt nicht sehr hilfreich. Hier kommen wieder die partiellen Ableitungen ins Spiel.

Definition 11.5. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$. Sei $x_i : U \to \mathbb{R}$ die i-te Koordinatenfunktion, e_i der i-te Standardbasisvektor. Die Abbildung heißt* partiell differenzierbar *in $x \in U$ in der i-ten Koordinate, falls*

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h\to 0}\frac{f(x + he_i) - f(x)}{h}$$

existiert. Sie heißt partiell differenzierbar *in x, wenn die partiellen Ableitungen für $i = 1, \dots, n$ existieren. Sie heißt* stetig partiell differenzierbar*, wenn die partiellen Ableitungen stetig sind.*

Bemerkung. Dies ist ein Spezialfall der totalen Differenzierbarkeit. $\frac{\partial f}{\partial x_i}(x)$ ist die totale Ableitung der Funktion

$$h \mapsto f(x + e_i h)$$

in $h = 0$. Sie ist auf einer offenen Umgebung von x in $\mathbb{R}$ definiert. Wie bei jeder Ableitung kann sie daher komponentenweise berechnet werden: Ist $f = (f_1, \dots, f_m)$, so ist

$$\frac{\partial f}{\partial x_i}(x) = \left(\frac{\partial f_1}{\partial x_i}(x), \dots, \frac{\partial f_m}{\partial x_i}(x)\right)$$

Jede der f_i ist $\mathbb{R}$-wertige. Beim Berechnen der partiellen Ableitungen sind wir dann in der Analysis 1.

Beispiel. Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (x^2, \sin(xy))$. Wir betrachten den (x,y). Es ist

$$\frac{\partial f(x,y)}{\partial x} = (2x, y\cos(xy))$$
$$\frac{\partial f(x,y)}{\partial y} = (0, x\cos(xy))$$

Hier wird die allgemein übliche, etwas unsaubere Notation benutzt. x_i steht einerseits fuer die i-te Koordinate des Punktes x, in dem differnziert werden soll. Gleichzeitig steht es fuer die Koordinatenfunktion auf U, nach der differenziert wird. In der Praxis macht dies keine Probleme.

Lemma 11.6. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ total differenzierbar in $x \in U$. Dann ist f partiell differenzierbar in x, und es gilt*

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

Beispiel. In dem letzten Beispiel ist also die lineare Abbbildung $Df(0)$ gegeben durch die 0-Matrix.

Beweis: Nach Voraussetzung gibt es eine lineare Abbildung A und eine Abbildung ϕ, die auf einer Umgebung von $0 \in \mathbb{R}^n$ definiert ist, so dass

$$f(x + \xi) = f(x) + A\xi + \phi(\xi)$$

mit

$$\lim_{\xi \to 0} \frac{\phi(\xi)}{\|\xi\|} = 0.$$

Wir berechnen die partielle Ableitung, also

$$\begin{aligned} \lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h} &= \lim_{h \to 0} \frac{Ahe_i + \phi(he_i)}{h} \\ &= Ae_i + \lim_{h \to 0} h^{-1}\phi(he_i) \end{aligned}$$

In jeder p-Norm gilt $\|he_i\|_p = h$. Mit $\xi = he_i$ sehen wir, dass der letzte Grenzwert verschwindet. Das Produkt Ae_i ist genau die i-te Spalte der Matrix A. $\square$

Bemerkung. Die Ableitung in x ist also gegeben durch die *Jacobi-Matrix*

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

Wir haben die Kettenregel gesehen, dass die Ableitung einer Komposition die Komposition der Ableitung ist. Also:

$$J_{g \circ f}(x) = J_g(f(x))J_f(x)$$

als Produkt von Matrizen.

Ist speziell $m = 1$, also $f : U \to \mathbb{R}$, so heißt $J_f(x)$ auch *Gradient* $\mathrm{grad}(f)$, es ist der Vektor der partiellen Ableitungen.

Die Umkehrung des Lemmas ist im allgemeinen falsch!

Beispiel. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$(x, y) \mapsto \begin{cases} \frac{xy}{x^2+y^2} & (x, y) \neq 0 \\ 0 & (x, y) = (0, 0) \end{cases}$$

Wir betrachten die partiellen Ableitungen im Urspung. Wegen $f(x, 0) = 0$ und $f(0, y) = 0$ existieren die partiellen Ableitungen, sie sind 0. Aber f ist nicht total differenzierbar, da f nicht stetig ist (betrachte die Nullfolge $(1/n, 1/n)_{n\in\mathbb{N}}$ mit $f(1/n, 1/n) = 1/2$).

Satz 11.7. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ partiell differenzierbar in U, mit stetigen partiellen Ableitungen in x. Dann ist f total differenzierbar in x.*

Beweis: Ohne Einschränkung arbeiten wir in $B(x, \delta) \subset U$. Für $\xi = (\xi_1, \ldots, \xi_n) \in B(0, \delta)$ müssen wir

$$\phi(\xi) = f(x + \xi) - f(x) - J_f(x)\xi$$

untersuchen. Wir gehen entlang der Koordinatenrichtungen von x zu $x + \xi$. Sei also

$$\xi^{(i)} = \sum_{j=1}^{i} \xi_j e_j.$$

Es ist $\xi^{(n)} = \xi$ und $\xi^{(i)} - \xi^{(i-1)} = \xi_i e_i$. Mit variierendem ξ, ändert sich $\xi^{(i)}$ nur in Richtung der i-ten Koordinatenrichtung. Wir wende den Mittelwertsatz der Differentialrechung aus Analyis 1 an auf die i-te partielle Ableitung und die Randpunkte $x + \xi^{(i-1)}$ und $\xi^{(i)} = x + \xi^{(i-1)} + \xi_i e_i$. Es gibt also ein $\theta_i \in [0, 1]$ mit

$$f(x + \xi^{(i)}) - f(x + \xi^{(i-1)}) = \frac{\partial f}{\partial x_i}(x + \xi^{(i-1)} + \theta_i \xi_i e_i)\xi_i.$$

Wir setzen $y^{(i)}$ für den Zwischenpunkt. Zusammen:

$$f(x + \xi) - f(x) = \sum_{i=1}^{n} \frac{\partial f(y^{(i)})}{\partial x_i} \xi_i.$$

und daher

$$\phi(\xi) = \sum_{i=1}^{n} \left[\frac{\partial f(y^{(i)})}{\partial x_i} - \frac{\partial f(x)}{\partial x_i} \right] \xi_i$$

Mit $\xi \to 0$ gilt auch $y^{(i)} \to x$. Da die partiellen Ableitung stetig ist, folgt

$$\lim_{\xi \to 0} \frac{\phi(\xi)}{\|\xi\|} \to 0.$$

□

Korollar 11.8. *Stetig partiell differenzierbare Abbildungen sind stetig.*

Richtungsableitungen

Partielle Ableitungen sind nützlich. Unschön ist die willkürliche Auszeichnung der Koordinatenrichtungen. Geht man von einem pyhsikalischen System aus, so sind dies willkürliche Wahlen. Wir wollen allgemeiner in *jede* Richtung ableiten. Auch nicht-lineare Koordinatenwechsel sollen erlaubt sein. Bei Licht betrachtet, ist es das Wort "Richtung", dass die Schwierigkeit macht.

Definition 11.9. *Ein* Weg *in* $\mathbb{R}^n$ *ist eine stetige Abbildung*

$$\gamma : I \to \mathbb{R}^n$$

wobei I *ein Intervall ist. Der Weg heißt (stückweise) differenzierbar, wenn* γ *(stückweise) differenzierbar ist.*
Sei γ *differenzierbar in* $t \in I$*. Dann heißt*

$$\gamma'(t) = D\gamma(t)$$

Tangentialvektor *von* γ *in* t*. Falls* $\gamma'(t) \neq 0$*, so heißt* $\gamma'(t)/\|\gamma'(t)\|$ *der* Einheitstangentialvektor *von* γ *in* t*.*

Beispiel. Für jedes $x \in \mathbb{R}^n$ und $i = 1, \ldots, n$ ist $t \mapsto x + te_i$ ein stetig differenzierbarer Weg in $\mathbb{R}^n$.

Geometrisch verstehen wir den Tangentenvektor als den Grenzwert der Sekantenvektoren. Physikalisch ist es der Geschwindigkeitsvektor eines Punktes, der sich entlang γ bewegt.

Definition 11.10. *Eine stetig differenzierbare Kurve* $\gamma : I \to \mathbb{R}^n$ *heißt* regulär*, falls* $\gamma'(t) \neq 0$ *für alle* $t \in I$*. Andernfalls heißt sie* singulär*.*

Beispiel. Die Neilsche Parabel

$$t \mapsto (t^2, t^3)$$

ist singulär in $t = 0$, regulär außerhalb.

Definition 11.11. *Sei* $U \subset \mathbb{R}^n$ *offen,* $\gamma : I \to U$ *ein differenzierbarer Weg,* $f : U \to \mathbb{R}^n$ *eine Abbildnung. Wir definieren die* Richtungsableitung *von* f *in* $x_0 = \gamma(t_0)$ *in Richtung* γ *als*

$$\frac{\partial f}{\partial \gamma}(x_0) = \lim_{t \to t_0} \frac{f(\gamma(t)) - f(\gamma(t_0))}{t - t_0}$$

falls der Grenzwert existiert.

Nach Definition ist $\frac{\partial f}{\partial \gamma}(x_0) = (f \circ \gamma)'(t_0)$.

Beispiel. Die partiellen Ableitungen sind die Richtungsableitungen in Richtung der Wege $t \mapsto x + te_i$.

Lemma 11.12. *Sei $U \subset \mathbb{R}^n$ offen, $\gamma : I \to U$ differenzierbarer Weg, $x_0 = \gamma(t_0)$, $f : U \to \mathbb{R}$ total differenzierbar in x_0, Dann ist*

$$\frac{\partial f}{\partial \gamma}(x_0) = \mathrm{grad}(f)(x_0)\gamma'(t_0) = \sum_{i=1}^{n} \frac{\partial f(x_0)}{\partial x_i}\gamma_i'(t_0).$$

Das Produkt in der Mitte ist das Produkt von Matrizen. Oft schreibt man auch $\langle \mathrm{grad}(f)(x_0), \gamma'(t)\rangle$, wie für das Standardskalarprodukt.

Beweis: Kettenregel. □

Wichtig ist, dass die Richtungsableitung nur vom Tangentialvektor abhängt!

Definition 11.13. *Zwei differenzierbare Wege γ_1, γ_2 durch $x_0 \in \mathbb{R}^n$ heißen äquivalent, falls für jede total differenzierbare Funktion f auf einer Umgebung von x_0 gilt*

$$\frac{\partial f}{\partial \gamma_1}(x_0) = \frac{\partial f}{\partial \gamma_2}(x_0).$$

Die Äquivalenzklassen von Wegen durch x_0 heißen Tangentialvektoren. *Der gemeinsame Wert der Ableitung heißt* Ableitung von f in Richtung des Tangentialvektors.

Lemma 11.14. *Die Zuordnung*

$$\gamma \mapsto \gamma'(t)$$

definiert eine Bijektion zwischen Tangentialvektoren und $\mathbb{R}^n$.

Beweis: Seien γ ein Weg. Wir werten aus in der differenzierbaren Funktion $f : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto x_i$. Nach den Formeln für die Richtungsableitung erhalten wir also

$$\frac{\partial f}{\partial \gamma}(x_0) = \mathrm{grad}(f)\gamma'(t_0) = \gamma_i'(t).$$

Daher ist die Abbildung wohldefiniert. Stimmen die Tangentialvektoren $\gamma'(t)$ für zwei Kurven überein, so sind sie äquivalent. Schließlich kann leicht jeder Vektor v als Richtungsvektor der Kurve $t \mapsto x_0 + tv$ geschrieben werden. □

Bemerkung. Punkte und Richtungsvektoren leben in verschiedenen Mengen. Da unsere Puntkte ebenfalls Elemente des $\mathbb{R}^n$ sind, verwechselt man dies leicht. In der Physik ist der Unterschied leichter: Punkte haben Ortskoordinaten, gemessen in einer Längeneinheit. Tangentialvektoren sind Geschwindigkeitskoordinaten, gemessen in Länge/Zeit.

In der Differentialgeometrie definiert man eine *Mannigfaltigkeit* als einen topologischen Raum mit der Hausdorffeigenschaft (und noch ein wenig mehr) und für jeden Punkt $x_0 \in X$ eine Wahl von Koordinatenfunktionen. Die obige Definition von Äquivalenz von Wegen ist unabhängig von der Koordinatenwahl. Unsere Definition von Tangentialvektoren verallgemeinert sich daher sofort auf Mannigfaltigkeiten.

Die Rechenregeln für Richtungsableitungen sind die vertrauten.

Lemma 11.15. *Seien γ ein differenzierbarer Weg in $U \subset \mathbb{R}^n$, $f, g : U \to \mathbb{R}$ differenzierbar, $a, b \in \mathbb{R}$. Dann gilt in $x_0 = \gamma(t_0)$*

(i) *(Linearität)* $\frac{\partial af+bg}{\partial \gamma}(x_0) = a\frac{\partial f}{\partial \gamma}(x_0) + b\frac{\partial g}{\partial \gamma}(x_0)$

(ii) *(Leibniz-Regel)* $\frac{\partial fg}{\partial \gamma}(x_0) = f\frac{\partial g}{\partial \gamma}(x_0) + g\frac{\partial f}{\partial \gamma}(x_0)$.

Beweis: Rechenregeln aus der Analysis 1 für die Ableitungen von Funktion $t \mapsto f(\gamma(t))$. □

Eine Zuordnung mit diesen Eigenschaften heißt *Deriviation.*

Satz 11.16. *Sei $\mathcal{C}$ die Menge der Funktionen, die auf einer Umgebung von x_0 definiert sind und differenzierbar. Sei $D : \mathcal{C} \to \mathbb{R}$ eine Derivation. Dann gibt es einen eindeutigen Tangentialvektor γ, so dass $D = \frac{\partial}{\partial \gamma}(x_0)$.*

Wir finden $\gamma'(t_0)$ indem wir D auf die Koordinatenfunktionen x_i anwenden. Der Rest des Beweises führt hier zu weit.

Höhere Ableitungen

Sei $f : U \to \mathbb{R}$ partiell differenzierbar. Dann erhalten wir Ableitungsfunktionen, die wir wieder differenzieren können.

Definition 11.17. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ partiell differenzierbar. Dann heißt f* zweimal (stetig) partiel differentierbar*, wenn alle partiellen Ableitungen (stetig) partiell differenzierbar sind.*
Allgemein heißt f für $k \in \mathbb{N}$ k-mal (stetig) partiell differenzierbar, wenn all $k-1$-ten partiellen Ableitungen (stetig) partiell differenzierbar sind.

Wir schreiben
$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f$$
und allgemein
$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} f.$$
Dabei kürzen wir im Nenner $\partial x_i \partial x_i$ ab zu ∂x_i^2, also
$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial}{\partial x} f.$$

Beispiel. Sei $f(x,y) = \sin(x+2y)$ mit

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= -\sin(x+2y)\\
\frac{\partial^2 f}{\partial y^2} &= -4\sin(x+2y)\\
\frac{\partial^2 f}{\partial x \partial y} &= -2\sin(x+2y)\\
\frac{\partial^2 f}{\partial y \partial} &= -2\sin(x+2y)
\end{aligned}$$

Satz 11.18. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ zweimal stetig partiell differenzierbar. Dann gilt für alle $x \in U$ und $i, j = 1, \dots, n$*

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(x) = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(x).$$

Beweis: In der Rechnung werden alle Koordinaten bis auf i, j festgehalten, daher ist ohne Einschränkung $n = 2$, $i = 1, j = 2$ und $x = 0$. Wie arbeiten in den Koordinaten (x, y). Ohne Einschränkung ist $U = (-\delta, \delta) \times -(\delta, \delta)$. Für $y \in (-\delta, \delta)$ betrachten wir die Funktion

$$F_y : (-\delta, \delta) \to \mathbb{R}, \qquad x \mapsto f(x,y) - f(x,0).$$

Nach Voraussetzung ist sie differenzierbar. Nach dem Mittelwertsatz der Differentialrechnung angewendet auf das Intervall $[0, x]$ gibt es also $\theta \in [0, x]$ mit

$$F_y(x) - F_y(0) = F_y'(\theta)x.$$

Hierbei ist

$$F_y'(\theta) = \frac{\partial f}{\partial x}(\theta, y) - \frac{\partial f}{\partial x}(\theta, 0).$$

Nach Voraussetzung ist $\frac{\partial f}{\partial x}(\theta, y)$ partiell differenzierbar nach y. Wir wenden wieder den Mittelwertsatz an auf das Intervall $[0, y]$. Es gibt also $\eta \in [0, y]$ mit

$$\frac{\partial f}{\partial x}(\theta, y) - \frac{\partial f}{\partial x}(\theta, 0) = \frac{\partial}{\partial y}\frac{\partial}{\partial x}(\theta, \eta)y.$$

Zusammen haben wir also:

$$f(x,y) - f(x,0) - f(0,y) + f(0,0) = \frac{\partial^2 f}{\partial y \partial x}(\theta, \eta)xy.$$

Dieselbe Argumentation wenden wir in der umgekehrten Reihenfolge an und finden $\eta' \in [0, y]$ und $\theta \in [0, x]$ mit

$$f(x,y) - f(x,0) - f(0,y) + f(0,0) = \frac{\partial^2 f}{\partial x \partial y}(\theta', \eta')xy.$$

Für $xy \neq 0$ folgt also

$$\frac{\partial^2 f}{\partial y \partial x}(\theta, \eta) = \frac{\partial^2 f}{\partial x \partial y}(\theta', \eta').$$

Nun bilden wir den Grenzwert $(x, y) \to 0$. Hieraus folgt $(\theta, \eta) \to 0$ und $(\theta', \eta') \to 0$. Da die zweiten Ableitungen stetig sind, folgt die gesuchte Gleichheit. $\square$

Einige partielle Ableitungen kommen so oft vor, dass es eigene Notationen gibt. Wir haben bereits den Gradienten einer Funktion $f : U \to \mathbb{R}$ mit $U \subset \mathbb{R}^n$ kennengelernt. Wir schreiben auch

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

und damit

$$\nabla f = \mathrm{grad}(f).$$

Ist $v : U \to \mathbb{R}^n$ eine Abbildung ("Vektorfeld"), so ist

$$\mathrm{div}(v) = \langle \nabla, v \rangle = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} v_i : U \to \mathbb{R}$$

die *Divergenz* des Vektorfeldes. Ist speziell $n = 3$, so ist

$$\mathrm{rot}(v) = \nabla \times v = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_2}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right) : U \to \mathbb{R}^3$$

die *Rotation des Vektorfeldes.*

Korollar 11.19. *Sei $U \subset \mathbb{R}^3$, $f : U \to \mathbb{R}$ zweimal stetig partiell diffbar. Dann gilt*

$$\mathrm{rot}(\mathrm{grad}(f)) = 0.$$

Beweis: Z.B. erhalten wir für die erste Komponente

$$\frac{\partial}{\partial x_3} \frac{\partial}{\partial x_1} f - \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_3} f = 0.$$

$\square$

Damit ein Vektorfeld ein Gradient ist, muss also seine Rotation verschwinden. Wieder für $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$ heißt

$$\Delta f = \mathrm{divgrad}(f) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$$

Laplace-Operator. Funktionen mit $\Delta f = 0$ heißen *harmonisch.*

Dieser Operator und seine Verallgemeinerungen spielen eine sehr große Rolle in der Physik, in der komplexen Analysis, in der Differentialgeometrie und auch in der analytischen Zahlentheorie.

Kapitel 12

Anwendungen

Reihenentwicklung

Wir wollen nun auch Funktionen in mehreren Variablen durch Polynome und Potenzreihen approximieren.
Wir betrachten zum Warmwerden Polynome in zwei Variablen, also:

$$a + (bx + cy) + (dx^2 + exy + fy^2) + (gx^3 + hx^2y + ixy^2 + jy^3) + \dots$$

Für jeden Monom $x^n y^m$ benötigen wir also einen Koeffizienten a_{nm} und erhalten die Formel

$$\sum_{n=0}^{N} \sum_{m=0}^{m} a_{nm} x^n y^n.$$

Allgemeiner erhalten wir bei einem Polynom in n Variablen Koeffizienten mit einem n-Tupel von Indizes. Zur Vereinfachung schreiben wir

$$\underline{i} = (i_1, \dots, i_n) \in \mathbb{N}_0^n, \qquad |\underline{i}| = \sum_{j=1}^{n} i_j,$$
$$\underline{i}! = i_1! i_2! \dots i_n!, \qquad x^{\underline{i}} = x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$$

Ein beliebiges Polynom vom Grad N schreibt sich dann als

$$P(\underline{x}) = \sum_{k=1}^{N} \sum_{\underline{i}=k} a_{\underline{i}} x^{\underline{i}}.$$

Aus der Theorie einer Variablen sind wir gewohnt, die Koeffizienten mit den Ableitungen in Verbindung zu bringen. Um den Koeffizienten vor $x^{\underline{i}}$ zu bestimmen wenden wir den Operatior

$$\frac{\partial^{|\underline{i}|}}{\partial x^{\underline{i}}} := \frac{\partial^{|\underline{i}|}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}}$$

an. Dabei erscheint ein Faktor $\underline{i}!$. Daher erwarten wir eine Potenzreihenentwicklung der Form

$$f(x) = \sum_{k=0}^{\infty} \sum_{|\underline{i}|=k} \frac{1}{\underline{i}!} \frac{\partial^k}{\partial x^{\underline{i}}} f(0) x^{\underline{i}}.$$

Theorem 12.1 (Taylor-Formel). *Sei $U = B(x,\delta) \subset \mathbb{R}^n$, $f : U \to \mathbb{R}^n$ eine $(N+1)$-mal stetig partiall differenzierbare Funktion. Sei $\xi \in B(0,\delta)$. Dann existiert ein $\theta \in [0,1]$, so dass*

$$f(x+\xi) = \sum_{|\underline{i}| \leq N} \frac{1}{\underline{i}!} \frac{\partial^{|\underline{i}|} f}{\partial x^{\underline{i}}}(x) \xi^{\underline{i}} + R(x,\theta)$$

mit

$$R(x,\theta) = \sum_{|\underline{i}|=N+1} \frac{\partial^{N+1} f(x+\theta\xi)}{\partial \underline{x}^{\underline{i}}} \xi^{\underline{i}}.$$

Beweis: Wir betrachten die Funktion

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

Sie ist $N+1$-mal stetig differenzierbar. Wir wenden die Taylor-Formel aus Analysis 1 an und berechnen dabei die Ableitungen nach der Kettenregel aus den partiellen Ableitungen (siehe unten). □

Lemma 12.2. *Seien f, x, ξ wie im Theorem, $g : [0,1] \to \mathbb{R}$ mit $g(t) = f(x+t\xi)$. Dann gilt*

$$\frac{\mathrm{d}^k g}{\mathrm{d}t^k}(t) = \sum_{|\underline{i}|=k} \frac{k!}{\underline{i}!} \frac{\partial^k f}{\partial x^{\underline{i}}}(x+t\xi)\xi^{\underline{i}}.$$

Beweis: Wir wenden k-mal die Kettenregel an und erhalten

$$\frac{\mathrm{d}^k g}{\mathrm{d}t^k}(t) = \sum_{j_1=1}^{n} \sum_{j_2}^{n} \cdots \sum_{j_k=1}^{n} \frac{\partial^k f(x+t\xi_i)}{\partial x_{j_1} \ldots \partial x_{j_k}} \xi_{j_1} \ldots \xi_{j_k}$$

Nun fassen wir alle Summanden vom Typ $\underline{i}$ zusammen. Es sind $k!/\underline{i}!$ viele. □

Wirklich interessant ist die Taylor-Formel, da das Restglied klein wird für $\xi \to 0$.

Definition 12.3. *Sei $\phi : B(0,\delta) \to \mathbb{R}$. Wir sagen $\phi = o(\|\xi\|^N)$ (lies: klein o von ....) falls*

$$\lim_{\xi \to} \frac{\phi(\xi)}{\|\xi\|^N} = 0.$$

Korollar 12.4. *Sei $B(x,\delta) \subset \mathbb{R}^n$, $f : B(x,\delta) \to \mathbb{R}$ eine N-mal stetige Funktion. Dann ist*

$$f(x+\xi) = \sum_{|\underline{i}| \leq N} \frac{1}{\underline{i}!} \frac{\partial^{|\underline{i}|} f}{\partial \underline{x}^{\underline{i}}}(x) \xi^{\underline{i}} + o(\|\xi\|^N).$$

Beweis: Nach dem Theorem haben wir den Fehler

$$\phi(\xi) = \sum_{|\underline{i}|=N+1} \left[\frac{\partial^{N+1} f(x+\theta\xi)}{\partial x^{\underline{i}}} - \frac{\partial^{N+1} f(x)}{\partial x^{\underline{i}}} \right] \xi^{\underline{i}}$$

wobei $\theta \in [0,1]$ aber von ξ abhängt. Wegen der Stetigkeitsvoraussetzung geht der Klammerausdruck gegen 0, und $\xi^{uli}/\|\xi\|^{N+1}$ ist beschränkt. Daher folgt die Behauptung. □

Besonders interessant ist der Fall $N = 2$

$$f(x+\xi) = f(x) + \mathrm{grad}(f)\xi + \sum_{i,j=1}^{n} \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j} \xi_i \xi_j + o(\|\xi\|^2).$$

Definition 12.5. *Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$ zweimal partiell differenzierbar. Dann heißt*

$$\mathrm{Hess}(f) = \left(\tfrac{\partial^2 f}{\partial x_i \partial x_j} \right)$$

Hesse Matrix *von f.*

Ist f zweimal stetig partiell differenzierbar, so ist $\mathrm{Hess}(f)$ symmetrisch und

$$f(x+\xi) = f(x) + \mathrm{grad}(f)\xi + \frac{1}{2}\langle \xi, \mathrm{Hess}(f)\xi \rangle + o(\|\xi\|^2).$$

Extremwerte

Die Definition von Extrema ist wie in Analysis 1.

Definition 12.6. *Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$ eine Funktion. Ein Punkt $x \in U$ heißt* lokales Maximum *(bzw.* lokales Minimum*), wenn es eine Umgebung $V \subset$ von x gibt, so dass*

$$f(x') \leq f(x) \text{ für alle } x' \in V$$

(bzw. $f(x') \geq f(x)$). Der Punkt heißt Extremum*, wenn er ein lokales Maximum oder Minimum ist. Ein Maximum (bzw. Minimum) heißt* isoliert*, wenn es eine Umgebung $W \subset V$ gibt, in der*

$$f(x') < f(x) \text{ für alle } x' \in V, x' \neq x.$$

(bzw. $f(x') > f(x)$.)

Wir betrachten einige Prototypen:

Beispiel. (i) $f(x,y) = x^2 + y^2$ hat in 0 ein isoliertes Minimum.

(ii) $f(x,y) = -x^2 - y^2$ hat in 0 ein isoliertes Maximum.

(iii) $f(x,y) = x + y$ hat in 0 weder ein Minimumnoch ein Maximum.

(iv) $f(x, y) = x^2$ hat in 0 ein nicht isoliertes Minimum.

(v) $f(x, y) = x^2 - y^2$ hat in 0 weder ein Minimum noch ein Maximum (Sattelpunkt!) Die Funktion wächst in Richtung x und fällt in Richtung y.

(vi) $f(x, y) = x^3 + y^2$ hat in 0 weder in lokales Minimum noch ein Maximum.

Lemma 12.7. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ partiell differenzierbar. Besitzt f in x ein lokales Extremum, so gilt*

$$\operatorname{grad}(f) = 0.$$

Beweis: Wir betrachten $t \mapsto f(x + te_i)$. Da f ein lokales Extremum hat, gilt dies auch für diese Funktion. Nach Analysis 1 verschwindet die Ableitung. $\square$

Wie wir oben gesehen habe, ist dies nicht hinreichend. Wir müssen die 2. Ableitungen, also die Hesse-Matrix betrachten. Im Beispiel oben sind dies:

- $\operatorname{Hess}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$
- $\operatorname{Hess}(f) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$
- –
- $\operatorname{Hess}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$
- $\operatorname{Hess}(f) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$

Die Beispiele waren besonders einfach gewählt, so dass alle Hesse-Matrizen in Diagonalgestalt sind. Durch geeignete Koordinatenwahl lässt sich das aber stets erreichen!

Definition 12.8. *Sei $A = (a_{ij})$ eine symmetrische reelle $n \times n$-Matrix. Die assoziierte symmetrische Bilinearform*

$$b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad (\xi, \xi') \mapsto \langle \xi, A\xi' \rangle = \sum_{i,j=1}^{n} \xi_i a_{ij} \xi'_j$$

heißt positiv definit *(bzw.* negativ definit*), falls $b(\xi, \xi) > 0$ (bzw. < 0) für alle $\xi \in \mathbb{R}^n$, $\xi \neq 0$. Sie heißt* positiv semidefinit *(bez.* negativ semindefinit*, wenn ≥ 0 (bzw ≤ 0) vorliegt. Sie heißt* indefinit*, wenn keiner Fälle eintritt, d.h. es gibt ξ, ξ' mit $b(\xi, \xi) > 0$ und $b(\xi', \xi' < 0$.*

Aus der linearen Algebra wissen wir, dass $V = \mathbb{R}^n$ eine Orthonormalbasis aus Eigenvektoren von A hat. Seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte. Es gilt

- A positiv definit, genau dann, wenn $\lambda_i > 0$ für alle i.

- A negativ definit, genau dann, wenn $\lambda_i < 0$ für alle i.
- A positiv semidefinit, genau dann wenn $\lambda_i \geq 0$ für alle i
- A negativ semidefinit, genau dann wenn $\lambda_i \leq 0$ üfr alle i
- A indefinit, genau dann wenn es wenigstens einen positiven und einen negativen Eigenwert gibt.

Wir wir also in unsere Beispiel sehen, entscheiden die Definitiheitseigenschaften der Hesse Matrix über die lokalen Extrema.

Satz 12.9. *Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}^n$ zweimal stetig partiell differenzierbar in $x \in U$ mit mit $\operatorname{grad}(f)(x) = 0$. Dann ist*

(i) x ein isoliertes Minium, falls $\operatorname{Hess}(f)(x)$ positiv definit ist;

(ii) x ein isolierters Maximum, falls $\operatorname{Hess}(f)(x)$ negativ definit ist;

(iii) x weder Minimum noch Maxiumum, falls $\operatorname{Hess}(f)(x)$ indefinit ist.

Beweis: Sei ohne Einschränkung $x = 0$, $A = \operatorname{Hess}(f)$. Ist ξ ein Eigenvektor von A zum Eigenwert λ, also $A\xi = \lambda\xi$ und $\langle \xi, A\xi \rangle = \sum_{i=1}^{n} \lambda \xi_i^2$. Entlang des Weges $\gamma(t) = t\xi$ (also in Richtung ξ) gilt

$$\begin{aligned} f(\gamma(t)) &= f(0) + \operatorname{grad}(f)(0)t\xi + \frac{1}{2}\sum_{i=1}^{n} \lambda t^2 \xi_i^2 + o(t^2). \\ &= f(0) + \left(\frac{1}{2}\sum_{i=1}^{n} \lambda \xi_i^2\right) t^2 + o(t^2) \end{aligned}$$

Aus Analysis 1 wissen wir, dass diese Funktion ein isoliertes Minimum/Maximum hat, wenn die 2. te Ableitung positiv/negativ ist. Dies ist der Fall, wenn $\lambda > 0/\lambda < 0$ gilt. Ist A indefinit, so liegt also weder ein Maximum noch ein Minimum vor.

Ist A positiv definit, so liegt in Richtung jedes Eigenvektors ein Minimum vor. Es bleibt zu zeigen, dass f dann wirklich ein isoliertes Minimum hat.

Sei $S = \{\xi \in \mathbb{R}^n | \|\xi\| = 1\}$. Diese Menge ist beschränkt und abgeschlossen, also kompakt. Daher nimmt die stetige Funktion

$$\xi \mapsto b(\xi, \xi) = \langle \xi, A\xi \rangle$$

auf S ein Minimum an. Da A positiv definit ist, ist dieses Minimum positiv. Also gibt es $\alpha > 0$ mit

$$\langle(\xi, A\xi \rangle \geq \alpha \|\xi\|^2.$$

(Klar für $\xi \in S$, $\xi = 0$. Der allgemeine Fall folgt durch Skalieren mit $\|\xi\|$). Nach Voraussetzung ist

$$f(\xi) = f(0) + \frac{1}{2}\langle \xi, A\xi \rangle + \phi(\xi)$$

wobei $\lim_{\xi\to} \phi(\xi)/\|\xi\|^2 = 0$. Sei $\delta > 0$, so dass

$$\phi(\xi) \leq \frac{\alpha}{4}\|\xi\|^2$$

Dann folgt

$$f(\xi) - f(0) \geq \frac{1}{2}\alpha\|\xi\|^2 - \frac{1}{4}\alpha\|\xi\|^2 > 0$$

Das Argument im negativ definiten Fall ist analog.

□

Parameterabhängige Integrale

Sei $U \subset \mathbb{R}^n$, $I = [a,b] \subset \mathbb{R}$ abgeschlossenes Intervall. Gegeben sei

$$f : U \times I \to \mathbb{R} \ .$$

Hieraus definieren wir

$$\phi : U \to \mathbb{R}, \qquad x \mapsto \int_a^b f(x,y)\mathrm{d}y$$

falls das Integral existiert. Wir nennen $\phi(x)$ das *Parameterintegral* zum Parameter x. Wir wollen nun diese Funktion auf Stetigkeit, Differenzierbarkeit und Integrierbarkeit untersuchen.

Satz 12.10. *Sei $U \subset \mathbb{R}^n$ offen, $I = [a,b] \subset \mathbb{R}$ und $f : U \times I \to \mathbb{R}$ stetig. Dann ist*

$$\phi : U \to \mathbb{R}, \qquad x \mapsto \int_a^b f(x,y)\mathrm{d}y$$

wohldefiniert und stetig.

Beweis: Für jedes x ist $f(x,\cdot)$ stetig. Das Integral existiert also nach Analysis 1.
Wir überprüfen Stetigkeit in $x \in U$ mit dem ε-δ-Kriterium. Sei also $x \in U$, $\varepsilon > 0$. Dann gibt es $\delta_0 > 0$ mit $K = \{x' \in \mathbb{R}^n | d(x,x') \leq \delta_0\} \subset U$. Dann ist $K \times I$ kompakt, also f gleichmäßig stetig auf $K \times I$. Zu jedem $\varepsilon > 0$ gibt es ein $0 < \delta < \delta_0$, so dass

$$d(x,x') < \delta \Rightarrow d(f(x',y), f(x,y)) < \frac{\varepsilon}{b-a}$$

Für $x' \in B(x,\delta)$ erhalten wir also die Abschätzung

$$|\phi(x') - \phi(x)| \leq \int_a^b |f(x',y) - f(x,y)|\mathrm{d}y \leq (b-a)\frac{\varepsilon}{b-a}$$

Also ist ϕ stetig.

□

Satz 12.11. *Sei $U \subset \mathbb{R}^n$, $I = [a,b]$, $f : U \times I \to \mathbb{R}$ stetig, und stetig partiell differenzierbar nach x_j für $j = 1, \ldots, n$. Dann ist $\phi : U \to \mathbb{R}$, $\phi(x) = \int_a^b \phi(x,y)\mathrm{d}y$ stetig partiell differenzierbar mit*

$$\frac{\partial \phi}{\partial x_j}(x) = \int_a^b \frac{\partial f}{\partial x_j}(x,y)\mathrm{d}y \ .$$

Beweis: Wir beweisen die Formel für die partielle Ableitung. Nach dem letzen Satz ist sie dann stetig.
Wir wollen zeigen, dass für $x \in U$, $1 \leq j \leq n$ gilt

$$\begin{aligned} 0 = \lim_{h \to 0} & \left| \frac{\phi(x+he_j) - \phi(x)}{h} - \int_a^b \frac{\partial f}{\partial x_j}(x,y)\mathrm{d} < \right| \\ &= \left| \int_a^b \left(\frac{f(x+he_j,y) - f(x,y)}{h} - \frac{\partial f}{\partial x_j}(x,y) \right) \mathrm{d}y \right| \end{aligned}$$

Sei $\varepsilon > 0$. Wir wählen δ_0 und K wie im letzten Beweis. Die Funktion $\frac{\partial f}{\partial x_j}$ ist gleichmäßig stetig auf $K \times I$. Also gibt es $\delta > 0$ so dass

$$d(x,x') < \delta \Rightarrow d\left(\frac{\partial f}{\partial x_j}(x',y), \frac{\partial f}{\partial x_j}(x,y) \right) < \frac{\varepsilon}{b-a}$$

Dafür schreiben wir auch den Differenzialquotienten als Integral. Nach dem Hauptsatz der Differential- und Integralrechnung für $\tilde{f}(t) = f(x+te_j,y)$ und $t = hs$ gilt

$$\begin{aligned} \frac{f(x+he_j,y) - f(x,y)}{h} &= \frac{1}{h} \int_0^h \frac{\mathrm{d}}{\mathrm{d}t} \tilde{f}(t)\mathrm{d}t \\ &= \frac{1}{h} \int_0^1 h \frac{\mathrm{d}\tilde{f}}{\mathrm{d}t}(sh)\mathrm{d}s \\ &= \int_0^1 \frac{\partial f}{\partial x_j}(x+she_j,y)\mathrm{d}s \end{aligned}$$

Wir setzen ein für $h \in [-\delta, \delta]$ und $x' = x + hse_j \in B(x, \varepsilon/(b-a))$

$$\begin{aligned} \left| \int_a^b \left(\frac{f(x+he_j,y) - f(x,y)}{h} - \frac{\partial f}{\partial x_j}(x,y) \right) \mathrm{d}y \right| &= \left| \int_a^b \int_0^1 \left(\frac{\partial f}{\partial x_j}(x+she_j,y) - \frac{\partial f}{\partial x_j}(x,y) \right) \mathrm{d}s\mathrm{d}y \right| \\ &\leq \int_a^b \int_0^1 \left| \frac{\partial f}{\partial x_j}(x+she_j,y) - \frac{\partial f}{\partial x_j}(x,y) \right| \mathrm{d}s\mathrm{d}y \\ &\leq \int_a^b \int_0^1 \frac{\varepsilon}{b-a} \mathrm{d}s\mathrm{d}y = \varepsilon. \end{aligned}$$

$\square$

Im Beweis haben wir bereits ein Doppelintegral benutzt. So wie es bei guten Funktionen nicht auf die Reihenfolge der partiellen Differention ankommt, ist es auch mit der Integration.

Satz 12.12 (Fubini)**.** *Sei $I = [a, b]$, $J = [\alpha, \beta]$ beschränkte abgeschlossene Intervalle. Sei $f : I \times J \to \mathbb{R}$ stetig. Dann gilt*

$$\int_\alpha^\beta \left(\int_a^b f(x,y)\mathrm{d}x \right) \mathrm{d}y = \int_a^b \left(\int_\alpha^\beta f(x,y)\mathrm{d}y \right) \mathrm{d}x.$$

Beweis: Wir halten α, β, a fest und variieren b. Sei

$$\begin{aligned}\phi(x) &= \int_\alpha^\beta \left(\int_a^x f(\xi,y)\mathrm{d}\xi \right) \mathrm{d}y \\ \psi(x) &= \int_a^x \left(\int_\alpha^\beta f(\xi,y)\mathrm{d}y \right) \mathrm{d}\xi\end{aligned}$$

Nach Satz 12.10 sind die inneren Funktionen

$$\begin{aligned}y &\mapsto \int_a^x f(\xi,y)\mathrm{d}\xi \\ \xi &\mapsto \int_\alpha^\beta f(\xi,y)\mathrm{d}y\end{aligned}$$

stetig und daher integrierbar. Daher sind ϕ und ψ wohldefiniert.

Es gilt $\phi(a) = \psi(a) = 0$. Daher genügt es zu zeigen, dass die Ableitungen nach x übereinstimmen.

Im Falle von ψ liefert der Hauptsatz der Differential- und Integralrechnungn

$$\psi'(x) = \int_\alpha^\beta f(\xi,y)\mathrm{d}y$$

Im Falle von ϕ betrachten wir

$$F(x,y) = \int_a^x f(\xi,y).$$

Die Funktion ist stetig partiell differenzierbar nach x mit Ableitung

$$\frac{\partial F}{\partial x} = f.$$

Nach dem Satz über das Ableiten von Parameterintegralen erhalten wir also

$$\begin{aligned}\phi'(x) &= \frac{\mathrm{d}}{\mathrm{d}x} \int_a^b F(x,y)\mathrm{d}y \\ &= \int_a^b \frac{\partial F}{\partial x}(x,y)\mathrm{d}y \\ &= \int_a^b f(x,y)\mathrm{d}y\end{aligned}$$

□

Bemerkung. In allen Sätzen in diesem Abschnitt ging es um das Vertauschen zweier Grenzwerte, z.B. partielle Ableitung/Integration, zwei Integrationen. Dies ist naheliegend, aber nicht immer richtig. Die wesentliche Eigenschaft, die das Vertauschen ermöglicht ist die *gleichmäßige* Stetigkeit. Wir haben sie als Konsequenz aus Kompaktheitsvoraussetzungen erhalten.

In Analysis 1 haben wir bereits ähnliche Beispiele gesehen, z.B. Grenzwerte von Funktionenfolgen und Differentiation, Integration oder die Cauchyproduktformel.

Bemerkung. Der Satz von Fubini erlaubt uns in sehr vielen Situationen mehrdimensionale Integration - einfach eine Variable nach der anderen. Eine allgemeine systematische Theorie ist Gegenstand der *Maßtheorie*, die auch die richtige Grundlage für Wahrscheinlichkeitstheorie ist. Sie ist bei uns Gegenstand der Analyis 3.

Beispiel. Wir berechnen

$$\int_0^a x^2 \cos(x) \mathrm{d}x.$$

Aus Analysis 1 kennen Sie eine Berechnung mittels partieller Integration. Es geht aber auch trickreicher:

$$F(y) = \int_0^a \cos(xy) \mathrm{d}x$$

Es folgt

$$F'(y) = \int_0^a x \sin(xy) \mathrm{d}x, F''(y) = \int_0^a -x^2 \cos(xy) \mathrm{d}x.$$

Wir müssen also $-F''(1)$ ausrechnen. Andererseits:

$$\int_0^a \cos(xy) \mathrm{d}x = \left. \frac{-\sin(xy)}{y} \right|_0^a = \frac{\sin(ay)}{y}$$

Wir erhalten für die zweite Ableitung direkt

$$\begin{aligned}
F'(y) &= \frac{a\cos(ay)y - \sin(ay)}{y^2} = \frac{a\cos(ay)}{y} - \frac{\sin(ay)}{y^2} \\
F''(y) &= \frac{-a^2 \sin(ay)y - a\cos(ay)}{y^2} - \frac{a\cos(ay)y^2 - 2y\sin(ay)}{y^4} \\
&= \frac{-a^2 \sin(ay)}{y} - \frac{2a\cos(ay)}{y^2} + \frac{2\sin(ay)}{y^3} \\
-F''(1) &= -(-a^2 - 2)\sin(a) - 2a\cos(a) = (a^2 - 2)\sin(a) + 2a\cos(a)
\end{aligned}$$

Und damit:

$$\int_0^x x^2 \cos(x) \mathrm{d}x = (x^2 - 2)\sin(x) + 2x\cos(x).$$

Anfänge der Variationsrechnung

Wir betrachten folgende Situation: $I = [a, b]$,

$$L : I \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, (t, q, p) \mapsto L(t, q, p)$$

zweimal stetig partiell differenzierbare Funktion, $c_1, c_2 \in \mathbb{R}^n$. Sei K die Menge $\phi : [a, b] \to \mathbb{R}^n$, die zweimal stetig partiell differenzierbar sind mit $\phi(a) = c_1$, $\phi(b) = c_2$. Wir definieren das *Wirkungsintegral*

$$S(\phi) = \int_a^b L(t, \phi(t), \phi'(t))\mathrm{d}t.$$

Gesucht ist dasjenige ϕ, so dass $S(\phi)$ minimal wird.

Beispiel. Wir betrachten die dreidimensionale Bewegung eines Massepunktes in einem Potential. Die Funktion ϕ beschreibt einen Weg von c_1 nach c_2 in $\mathbb{R}^3$, die Ableitung $\phi'(t)$ ist die jeweilige Geschwindigkeit.
Wir setzen die *Lagrangefunktion*

$$L = T - V$$

wobei T die kinetische und V die potientielle Engergie ist. Die kinetische Energie hängt von der Geschwindigkeit, die potentielle vom Ort ab. Das *Hamiltonsche Prinzip* besagt, dass der tatsächliche ablaufende Vorgang das Wirkungsintegral minimiert.

Satz 12.13 (Eulersche Differentialgleichungen)**.** *Sei $a, b, L, c_1, c_2, K, S : K \to \mathbb{R}^n$ wie oben. Eine notwendige Bedingung für $S(\phi)) = \inf_{\phi' \in S(\phi')}$ ist die Differenzialgleichung*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial p_i} L(t, \phi(t), \phi'(t)) - \frac{\partial L}{\partial q_i}(t, \phi(t), \phi'(t)) = 0$$

für $i = 1, \ldots, n$.

Beweis: Sei ϕ, so dass das Wirkungsintegral minimiert wird. Sei $g : [a, b] \to \mathbb{R}^n$ zweimal stetig differenzierbar mit $g(a) = g(b) = 0$. Für jedes $\lambda \in \mathbb{R}$ ist $\phi + \lambda g \in K$, also

$$S(\phi) \leq S(\phi + \lambda g).$$

Wir definieren

$$F(\lambda) = S(\phi + \lambda g) = \int_a^b L(t, \phi(t) + \lambda g(t), \phi'(t) + \lambda g'(t))\mathrm{d}t.$$

Dies ist eine differenzierbare Funktion $\mathbb{R} \to \mathbb{R}$. Sie hat ein Minimum in $\lambda = 0$, also $\frac{\mathrm{d}F}{\mathrm{d}\lambda}(0) = 0$. Nach dem Satz über die Ableitung von Parameterintegralen folgt

$$\frac{\mathrm{d}F}{\mathrm{d}\lambda}(\lambda) = \int_a^b \frac{\mathrm{d}}{\mathrm{d}\lambda} L(t, \phi(t), \phi(t) + \lambda g(t), \phi'(t) + \lambda g'(t))\mathrm{d}t$$

Wir berechnen die Ableitung unter dem Integral:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} L(t, \phi(t), \phi(t) + \lambda g(t), \phi'(t) + \lambda g'(t))$$
$$= \sum_{i=1}^{n} \frac{\partial L}{\partial q_i}(\dots) g_i(t) + \sum_{i=1}^{n} \frac{\partial L}{\partial p_i}(\dots) g_i'(t)$$

Durch partielle Intergration erhalten wir

$$\int_a^b \frac{\partial L}{\partial p_i}(\dots) g_i'(t)\mathrm{d}t = \frac{\partial L}{\partial p_i}(\dots) g_i(t)|_a^b - \int_a^b g_i(t) \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial p_i}\mathrm{d}t$$

Wegen $g(a) = g(b) = 0$ entfällt der erste Summand! Zusammen erhalten wir in $\lambda = 0$:

$$0 = \int_a^b \sum_{i=1}^{n} \left[\sum_{i=1}^{n} \frac{\partial L}{\partial q_i}(\dots) - \sum_{i=1}^{n} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial p_i}(\dots) \right] \mathrm{d}t.$$

Diese Aussage gilt für *jedes* g. Wählen wir $g_j = 0$ für $j \neq i$, so sehen wir, dass jeder Summand einzeln verschwinden muss. Aus dem untigen Lemma folgt, dass dann jeder einzelne Integrand verschwindet. Das sind genau die Eulerschen Differentialgleichungen. □

Lemma 12.14. *Sei $f : [a, b] \to \mathbb{R}$ stetig. Für jede zweimal stetig differenzierbare Funktion $g : [a, b] \to \mathbb{R}$ mit $g(a) = g(b) = 0$ gelte*

$$\int_a^b f(t)g(t)\mathrm{d}t.$$

Dann ist $f(t) = 0$ für alle t.

Beweis: Angenommen, es ist $f(t_0) \neq 0$ für ein $t \in (a, b)$. (Ohne Einschränkung, da f stetig). Ohne Einschränkung ist $\varepsilon = f(t_0) > 0$. Dann existiert ein $\delta > 0$, so dass $f(t) \geq \varepsilon/2$ für alle $t \in [t_0 - \delta, t_0 + \delta]$. Man konstruiert eine nichtnegative Funktion g mit $g(t_0) = 1$ und $g(t) = 0$ für $t \notin [t_0 - \delta, t_0 + \delta]$. Es folgt

$$\int_a^b f(t)g(t)\mathrm{d}t = \int_{t_0-\delta}^{t_0+\delta} f(t)g(t)\mathrm{d}t \geq \int_{t_0-\delta}^{t_0+\delta} \frac{\varepsilon}{2} g(t)\mathrm{d}t > 0$$

Dies ist ein Widerspruch. □

Beispiel. In unserem physikalischen Beispiel ist $V(x_1, x_2, x_3)$ das Potential, $T = \frac{1}{2}m\sum_{i=1}^n v_i^2$ die kinetische Energie. Dabei ist v_i die Geschwindigkeit, also

$$L(x, v) = \frac{1}{2}m\sum_{i=1}^{2} v_i^2 - V(x_1, x_2, x_3),$$

unabhängig von t. Die partiellen Ableitungen sind

$$\frac{\partial}{\partial x_i} L(x,v) = -\frac{\partial U}{\partial x_i}$$
$$\frac{\partial}{\partial v_i} L(x,v) = m v_i$$

Die Eulersche Differentialgleichung lautet dann

$$\frac{\mathrm{d}}{\mathrm{d}t} m\dot{x}(t) + \frac{\partial U}{\partial x_i}(x(t)) = 0$$

also

$$m\ddot{x}(t) = -\mathrm{grad} U(x(t))$$

Hieraus erhält man Energieerhaltung: Die Energie des Systems ist $T+U$, also

$$E(t) = \frac{1}{2} m \sum \dot{x}_i^2(t) + U(x_1(t), x_2(t), x_3(t))$$

Ableiten ergibt (Kettenregel!)

$$\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}t} &= \frac{1}{2} m \sum 2\dot{x}_i(t)\ddot{x}_i(t) + \sum \frac{\partial U}{\partial x_i}(x(t))\dot{x}_i(t) \\
&= m\langle \dot{x}(t), \ddot{x}(t)\rangle + \langle \mathrm{grad} U, x(t)\rangle \\
&= \langle \dot{x}(t), -\mathrm{grad} U(x(t))\rangle \langle \mathrm{grad} U, x(t)\rangle \\
&= 0
\end{aligned}$$

Kapitel 13

Die Satz über inverse und implizite Funktionen

Beispiel. Sei $F(x,y) = x^2 + y^2 - 1$. Die Nullstellenmenge V dieser Funktion ist die Einheitskreislinie. Diese "sieht lokal aus" wie ein Intervall. Wir suchen daher in einer Umgebung von $(x_0, y_0) \in V$ eine bijektive Abbildung auf ein Intervall. Ein guter Kandidat ist die Projektion auf die x oder y-Koordinate. Damit die Abbildung

$$(x,y) \mapsto x$$

bijektiv werden kann, benötigen wir eine Umkehrabbildung, also ein ϕ mit $(x, \phi(x)) \in V$. Mit anderen Worten, wir wollen V als Graph schreiben. Wir haben die Gleichung

$$x^2 + \phi(x)^2 = 1 \Rightarrow \phi(x) = \pm\sqrt{1 - x^2}$$

Je nach y_0 wählen wir die positive oder negative Wurzel. Damit die Funktion bijektiv bleibt, müssen wir den Wert $x = \pm 1$ vermeiden. Er ist dadurch ausgezeichnet, dass $\frac{\partial F}{\partial x}$ verschwindet.
In den meisten Punkten haben wir zwei Koordinatenwahlen durch x und y. Sie gehen durch

$$y = \pm\sqrt{1 - x^2}, x = \sqrt{1 - y^2}$$

in einander über. Diese Funktionen sind bijektiv und differenzierbar.

Unser Ziel ist es, dieses Beispiel zu verallgemeinern. Der schwierigste Teil ist die Existenz der Funktion ϕ. Hierfür benutzen wir ein Hilfsmittel, dass auch noch viele andere Anwendungen in Theorie und Numerik hat.

Theorem 13.1 (Banachscher Fixpunktsatz)**.** *Sei Sei (X, d) ein vollständiger metrischer Raum, und $F : X \to X$ eine Kontraktion, das heißt es gibt ein $\theta \in [0, 1)$ mit*

$$d(F(x), F(y)) \leq \theta d(x, y)$$

für alle $x, y \in X$. Dann gibt es genau ein $x \in X$ mit $F(x) = x$.

Beweis: Wir beginnen mit der Eindeutigkeit. Seien x, x' Fixpunkte. Dann ist

$$d(x, x') = d(F(x), F(x')) < d(x, x'),$$

Widerspruch.
Für die Existenz wählen wir einen beliebigen Startwert $x_0 \in X$ und betrachten die Folge

$$x_n = F(x_{n-1}).$$

Wir wollen zeigen, dass es sich um eine Cauchy-Folge handelt. Es ist

$$d(x_n, x_{n-1}) = d(F(x_{n-1}, x_{n-2}) \leq \theta d(x_{n-1}, x_{n-2}).$$

Hieraus folgt mit Induktion

$$d(x_n, x_{n-1}) \leq \theta^{n-1} d(x_1, x_0)$$

Hieraus folgt mit der Dreicksungleichung

$$\begin{aligned} d(x_n, x_0) \leq \sum_{i=1}^{n} d(x_i, d_{i-1}) \leq \sum_{i=1}^{n} \theta^{i-1} d(x_1, x_0) \\ \leq \sum_{i=1}^{\infty} \theta^{i-1} d(x_1, x_0) = \frac{1}{1-\theta} d(x_1, x_0) \end{aligned}$$

(geometrische Reihe, da $\theta < 1$). Die gleiche Abschätzung gilt auch, wenn mir x_n als Startwert auffassen, also für $m > n$

$$d(x_m, x_n) \leq \frac{1}{1-\theta} d(x_{n+1}, x_n) \leq \frac{\theta^n}{1-\theta} d(x_1, x_0).$$

Also handelt es sich um eine Cauchyfolge. Da X vollständig ist, konvergiert sie gegen einen Punkt $x \in X$.
Aus der Kontraktionseigenschaft von F folgt insbesondere Stetigkeit (ε-δ-Kriterium mit $\delta = \varepsilon$). Also:

$$F(x) = F(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

□

Bemerkung. Ist $X \subset \mathbb{R}^n$ und F zusätzlich differenzierbar, so sehen wir dass für $v \in \mathbb{R}^n$

$$\left| \frac{\partial F}{\partial v}(x) \right| = \lim_{h\to 0} \left| \frac{F(x+hv) - F(x)}{h} \right| \leq \theta.$$

Wir erinnern an die Operatornorm

$$\|Df(x)\| = \sup_{\|v\|=1} |Df(x)(v)| = \sup_{\|v\|=1} \left| \frac{\partial f}{\partial v}(x) \right|$$

Also erhalten wir

$$\|Df(x)\| \leq \theta$$

als notwendige Bedingung für eine Kontraktion.

Beispiel. Wir untersuchen $F : \mathbb{R} \to \mathbb{R}$ gegeben durch $x \mapsto x^2$. Handelt es sich um eine Kontraktion? Es ist

$$|F(x+h) - F(x)| \leq |F'(x)||h| + o(h)$$

Bis auf einen kleinen Fehler entscheidet also $F'(x) = 2x$ über diese Frage. (Wir werden das gleich sauber ausarbeiten). Auf $[-1/4, 1/4]$ ist $|F'| \leq 1/2$. Auf diesem Bereich handelt es sich um eine Kontraktion. Der Banachsche Fixpunktsatz findet also einen Fixpunkt, z.B. zum Startwert 1/4:

$$\frac{1}{4}, \frac{1}{16}, \dots$$

Wir finden tatsächlich den Fixpunkt 0. Auch mit dem Startpunkt 3/4 würde es gut gehen. Dennoch ist F hier keine Kontraktion, da

$$F(\frac{3}{4}) - F(\frac{1}{2}) = \frac{9}{16} - \frac{1}{4} = \frac{1}{2}$$

Den Fixpunkt 1 finden wir mit dem Verfahren gar nicht!

Um den Fixpunktsatz anwenden zu können, benötigen wir die Umkehrung unseres notwendigen Kriteriums.

Definition 13.2. *Eine Menge $M \subset \mathbb{R}^n$ heißt* konvex*, falls mit je zwei Punkte auch die Streckt zwischen ihnen in M liegt, also*

$$x, y \in M \Rightarrow (1-t)x + ty \in M \text{ für alle } t \in [0,1].$$

Typische Beispiele sind Kugeln bezüglich einer p-Norm.

Satz 13.3. *Sei $U \subset \mathbb{R}^n$ offen, konvex, $X = \overline{U}$, $f : U \to U$ stetig differenzierbar. Sei $\theta < \infty$ mit $|Df(x)| \leq \theta$ auf U. Dann folgt*

$$|f(x) - f(y)| \leq \theta|x - y|$$

für alle $x, y \in U$.

Beweis: Wir betrachten den Weg $\gamma : [0,1] \to U$, $t \mapsto (1-t)x + ty$. Es gilt $\gamma' = (t) = y - x$. Mit dem Hauptsatz für die Funktion $(f \circ \gamma)' = \frac{\partial f}{\partial \gamma}\gamma'(t)$ gilt

$$f(y) - f(x) = \int_0^1 Df(\gamma(t)(y-x)\mathrm{d}t$$

und daher

$$|f(y) - f(x)| \leq \int_0^1 |Df(\gamma(t))(y-x)|\,\mathrm{d}t \leq \theta|y - x|$$

□

Ist $\theta < 1$, so erhalten wir also eine Kontraktion.

Theorem 13.4 (Satz über inverse Funktionen)**.** *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ stetig partiell differenzierbar. Sei $x_0 \in U$ mit $Df(x_0)$ invertierbar als Matrix. Dann gibt es eine offene Umgebung $V \subset U$ von x_0, so dass gilt*

(i) $W = f(V)$ ist offene Umgebung von $y_0 = f(x_0)$,

(ii) $f : V \to W$ ist bijektiv und die Umkehrfunktion ebenfalls stetig partiell differenzierbar.

Bijektive differenzierbare Funktionen mit differenzierbarer Umkehrfunktion heißen *Diffeomorphismen.* Bijektive stetige Funktionen mit stetiger Umkehrfunktion heißen *Homöomorphismen.*

Beweis: Ohne Einschränkung ist $x_0 = 0$, $y_0 = f(x_0) = 0$. Sei $A = Df(0)$. Wir suchen eine Umkehrfunktion von f, wenigstens nahe bei 0, also ein Urbild von y nahe bei 0. Dies formulieren wir als Fixpunktproblem und betrachten $R(x) = f(x) - Ax : U \to \mathbb{R}^n$. Es ist dann

$$f(x) = y \Leftrightarrow Ax + R(x) = y \Leftrightarrow x = A^{-1}(y - R(x)).$$

Sei also für $y \in \mathbb{R}^n$

$$F_y : U \to \mathbb{R}^n, \qquad F_y(x) = A^{-1}(y - R(x)) \ .$$

Wir suchen einen Fixpunkt von F_y, dies ist das Urbild $f^{-1}(y)$. Um den Fixpunktsatz anwenden zu können, müssen wir also erreichen, das F_y eine Kontraktion ist. Sei $C = \|A^{-1}\| > 0$. Es ist

$$DR(x) = Df(x) - A \ .$$

Diese Funktion ist stetig nach Voraussetzung und es gilt $DR(0) = A - A = 0$. Also gibt es $\delta_0 > 0$, so dass der abgeschlossene Ball $B = \overline{B(0, \delta_0)}$ ganz in U liegt und für $x \in B$

$$\|DR(x)\| \leq \frac{1}{2C} \ .$$

Aus dem letzten Satz erhalten wir für $x_1, x_2 \in B$

$$|R(x_1) - R(x_2)| \leq \frac{1}{2C}|x_1 - x_2|.$$

Daraus berechnen wir für F_y, weiter $x \in B$:

$$\begin{aligned} |F_y(x_1) - F_y(x_2)| &= |A^{-1}(y - R(x_1) - A^{-1}(y - R(x_2)| \\ &= |A^{-1}(R(x_1) - R(x_2)| \leq C|R(x_1) - R(x_2)| \leq \frac{1}{2}|x_1 - x_2| \end{aligned}$$

Dies ist die Kontraktionseigenschaft. Allerdings ist nicht klar, ob $F_y(B) \subset B$. Daher schätzen wir ab:

$$\begin{aligned} |F_y(x)| = |A^{-1}(y - R(x))| &\leq \|A^{-1}\|(|y| + |R(x)|) \\ &= C(|y| + |R(x) - R(0)|) \leq C|y| + \frac{1}{2}|x| \end{aligned}$$

Für $0 < \delta \leq \delta_0$, $\varepsilon = \delta/(2C)$ erhalten wir

$$\|x\| \leq \delta, \|y\| \leq \varepsilon \Rightarrow |F_y(x)| < C\varepsilon + \frac{1}{2}\delta = \delta.$$

Damit ist F_y eine Kontraktion of dem abgeschlossenen Ball $\overline{B(0,\delta)}$.

Nach dem Banachschen Fixpunktsatz gibt es einen eindeutigen Fixpunkt x. Dies ist das gesuchte Urbild $f^{-1}(y)$. Er liegt sogar in dem offenen Ball, da nach unserer Abschätzung

$$\|x\| = \|F_y(x)\| < \delta$$

Wir setzen nun $W = B(0,\varepsilon)$, $V = B(0,\delta) \cap f^{-1}(W)$. Dies sind die gesuchten offenen Umgebungen. Wir haben $f : V \to W$. Für jedes $y \in W$ haben wir $f^{-1}(x) \in B(0,\delta)$ konstruiert. Es liegt dann auch in $f^{-1}(W))$, also

$$f^{-1} : W \to V.$$

Damit haben wir gezeigt, dass die Funktion bijektiv ist!

Zu zeigen bleibt die Differenzierbarkeit von $g = f^{-1}$ in 0. Es gilt vorab

$$\|g(y)\| = \|F_y(g(y))\| \leq C\|y\| + \frac{1}{2}\|g(y)\| \Rightarrow \|g(y)\| \leq 2C\|y\|$$

Insbeondere ist g stetig im Nullpunkt mit $g(y) = 0$. Aus der Kettenregel erhalten wir $Dg = A^{-1}$, wenn die Ableitung existiert. Für $y \neq 0$ ist $g(y) \neq 0$, und es gilt (mit der Stetigkeitsüberlegung)

$$\begin{aligned} \frac{\|g(y) - A^{-1}y\|}{\|y\|} = \frac{\|F_y(g(y)) - A^{-1}y\|}{\|y\|} = \frac{\|A^{-1}(R(g(y))\|}{\|y\|} &\leq C\frac{\|R(g(y))\|\|g(y\|}{\|g(y)\|\|y\|} \\ &\leq C\frac{\|R(g(y))\|}{\|g(y)\|}2C \end{aligned}$$

Dies strebt gegen 0 für $y \to 0$, wegen wegen der Stetigkeit von g in 0 und

$$\lim_{x\to 0} \frac{\|g(x)\|}{\|x\|} = \lim_{x\to 0} \frac{\|f(x) - Ax\|}{\|x\|} = 0$$

wegen der Differenzierbarkeit von f in 0.

Damit ist die Umkehrfunktion differenzierbar in 0. Wir können das bisher bewiesene auf jeden Punkt $x_0 \in V$ anwenden. Wegen der Eindeutigkeit der Umkehrunktion erhalten wir dasselbe g und dessen Stetigkeit und Differenzierbarkeit in $y_0 = f(x_0)$. Aus der Kettenregel erhalten wir also

$$Dg = (Df)^{-1} \circ g$$

Die Funktion g ist stetig, ebenso nach Voraussetzung Df. Aus der Kramerschen Regel und der der Stetigkeit von Summe, Produkt und Quotient erhalten wir die Stetigkeit von $(Df)^{-1}$. Damit ist auch Dg stetig. □

Beispiel. Wir erinnern uns an die *Polarkoordinaten*

$$(x, y) = (r \cos\phi, r \sin\phi)$$

auf $\mathbb{R}^2$. Hierbei ist $r \in [0, \infty)$, $\phi \in \mathbb{R}$. Wir berechnen die totale Ableitung:

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -r\sin\phi & r\cos\phi \end{pmatrix}$$

Diese Matrix hat die Determinante $r(\cos^2\phi + \sin^2\phi) = r$. Für $r \neq 0$ ist also der Koordinatenwechsel $(r, \phi) \mapsto (x, y)$ lokal bijektiv und ein Diffeormorphismus. Wir wissen, dass die Abbildung global nicht injektiv ist, wohl aber auf Mengen der Form $\mathbb{R}_{>0} \times (\phi_0, \phi_0 + 2\pi)$.

Wir kehren zurück zu unserem Problem, eine Menge wie $\{x^2 + y^2 = 1\}$ lokal als Graph einer Funktion zu schreiben. Wir betrachten daher die folgende Situation: Sei $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^k$,

$$f : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^k$$

eine Abbildung. Uns interssiert die Menge

$$\{(x, y) \in U \times V | f(x, y) = 0\}$$

Wann ist diese Menge (lokal, also in der Nähe eines Startpunktes (x_0, y_0)) als Graph einer Funktion

$$\phi : U \to V$$

beschreibbar, d.h.

$$f(x, \phi(x)) = 0$$

Wir sagen die Funktion ϕ wird *implizit definiert* durch F. Was sind die Eigenschaften von ϕ? Existenz? Eindeutigkeit? Stetigkeit? Differenzierbarkeit? Aus dem Beispiel sehen wir, dass wir Bedingungen an die Ableitungen in Richtung x stellen wollen.

Für eine Abbildung $(x, y) \mapsto f(x, y)$ führen wir die Notation

$$D_y f = \left(\frac{\partial f}{\partial x_i})\right)_i$$

ein. Hier ist $y = (y_1, \ldots, y_k)$, also erhalten wir eine Matrix aus partiellen Ableitungen.

Theorem 13.5 (Satz über implizite Funktionen)**.** *Sei $U \subset \mathbb{R}^m \times \mathbb{R}^k$ offen, $f : U \to \mathbb{R}^k$ sei stetig partiell differenzierbar. Sei $(x_0, y_0) \in U$ mit $f(x_0, y_0) = 0$ und*

$$D_y f(x_0, y_0) = \left(\frac{\partial f}{\partial x_i}(x_0, y_0)\right)_{i=1}^n$$

invertierbar. Dann gibt es offene Umgebungen W von x_0 und V von y_0 und eine stetig partiell differenzierbare Funktion

$$g : W \to V$$

mit

$$\{(x, y) \in W \times V | f(x, y) = 0\} = \{(x, g(x)) | x \in W\}$$

Die Funktion g hat die Ableitung

$$Dg(x_0) = -(D_y f(x_0, y_0))^{-1} D_x f(x_0, y_0)$$

Beweis: Wir führen die Aussage zurück auf den Satz über inverse Funktionen, indem wir

$$F : U \to \mathbb{R}^m \times \mathbb{R}^n, \qquad (x, y) \mapsto (x, f(x, y))$$

betrachten. Es gilt

$$DF = \begin{pmatrix} E_m & 0 \\ D_x f & D_y f \end{pmatrix} .$$

Diese Matrix ist genau dann invertierbar, wenn $D_y f$ es ist. Dies hatten wir in (x_0, y_0) gerade vorausgesetzt. Nach dem Satz über inverse Funktionen gibt es also Umgebungen V' von $(x_0, y_0$ und W' von $(x_0, f(x_0, y_0)) = (x_0, 0)$, die unter F in Bijektion stehen. Sowohl F als auch $G = F^{-1}$ sind stetig, daher können wir V' in der Form $V_1 \times V_2$ wählen. Für $(x, z) \in W'$ ist also $(x, z) = (x, f(x, y))$ mit eindeutigem $(x, y) \in V_1 \times V_2$, d.h.

$$G(x, z) = G(x, f(x, y)) = G(F(x, y)) = (x, y) .$$

Mit anderen Worten:

$$G(x, z) = (x, g_0(x, z))$$

mit $g_0 : W' \to \mathbb{R}^k$, genauer $g_0 : W' \to V_2$. In W' wählen wir eine kleinere Umgebung von $(x_0, 0)$ von der Form $W_1 \times W_2$. Wir definieren

$$g : W_1 \to V_2, \qquad g(x) = g_0(x, 0).$$

Diese Funktion erfüllt

$$(x, f(x, g(x))) = F(x, g(x)) = F(x, g_0(x, 0)) = FG(x, 0) = (x, 0).$$

Ist umgekehrt $(x, y) \in W_1 \times V_2$ mit $f(x, y) = 0$, so ist

$$(x, y) = GF(x, y) = G(x, f(x, y))) = G(x, 0) = (x, g_0(x, 0)))$$

also $y = g_0(x, 0) = g(x)$ wie behauptet. $\square$

Beispiel. $f(x, y) = x^2 + y^2 - 1$. Wegen $\frac{\partial f}{\partial y} = 2y$ können wir das Theorem anwenden auf die erste Koordinate, solange $y \neq 0$. Wir finden y als Funktion von x. Umgekehrt können wir x als Funktion von y schreiben, wenn $x \neq 0$.

Definition 13.6. *Eine* differenzierbare Mannigfaltikeit *der Dimension k ist ein topologischer Raum M, der das Hausdorff-Trennungsaxiom erfüllt, zusammen mit einer offenen Überdeckung $M = \bigcup_i U_i$ und Homöomorphismen*

$$\phi_i : U_i \to V_i \subset \mathbb{R}^k$$

*(*die Karten*) auf offene Teilmengen $V_i \subset \mathbb{R}^k$, so dass für alle $i, j \in I$ auf $U_i \cap U_j$ die* Kartenwechselabbildungen

$$\phi_{ij} = \phi_j \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$$

stetig differenzierbar sind.

Bemerkung. In der korrekten Definition werden noch weitere Bedingung an die Topologie gestellt, auf die wir hier nicht eingehen wollen. Sie sorgen dafür, dass eine "Teilung der Eins" existiert.

Beispiel. $S^1 = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 = 1\}$ wird zu einer Mannigfaltigkeit mit der Überdeckung durch die Teilmengen

$$U_1 = \{(x,y)|y > 0\}, U_2 = \{(x,y)|y < 0\},$$
$$U_3 = \{(x,y)|x > 0\}, U_4 = \{(x,y)|x < 0\}$$

und den Kartenabbildungen $(x,y) \mapsto x$ auf U_1, U_2 und $(x,y) \mapsto y$ auf U_3, U_4. Die Kartenwechsel haben die Form $x = \pm\sqrt{y^2 - 1}$ bzw. $y = \pm\sqrt{x^2 - 1}$, sind also differenzierbar.

Beispiel. Sei $U \subset \mathbb{R}^{m+k}$ offen, $f : U \to \mathbb{R}^k$ stetig partiell differenzierbar. Sei

$$M = \{x \in U | f(x) = 0\}.$$

Falls die totale Ableitung

$$Df = \left(\frac{\partial f_i}{\partial x_j}\right)$$

in jedem Punkt von M den (maximal möglichen) Rang k hat, so wird M zu einer Mannigfaltikeit der Dimension m mit den Umgebungen, die aus dem Satz über implizite Funktionen gegeben werden. Solche Mannigfaltikeiten heißen *Untermannigfaltigkeiten.*

Im einzelnen: Sei $x \in M$. Nach Voraussetzung gibt es k Koordinaten $x_{j_1}, \dot{,} x_{j_k}$, so dass die quadratische Matrix

$$\left(\frac{\partial f_i}{\partial x_{j_s}}\right)_{i,s=1}^{k}$$

invertierbar ist. Nach Umsortieren der Koordinaten sind wir in der Situation des Satzes über implizite Funktionen. In einer kleinen Umgebung U_x ist die Projektion auf die Koordinaten ungleich $x_{j_1}, \ldots, x_{j_k}$ eine Kartenabbildung. In der Notation des Satzes ist W das Bild in $\mathbb{R}^m$, $U_x = W \times V \cap M$. Die Koordinatenwechselabbildungen können wir explizit in Termen der impliziten Funktion g ausdrücken. Daber sind sie differenzierbar.

Definition 13.7. *Seien M, N Mannigfaltigkeiten der Dimension m und n mit Karten $\{\phi_i : U_i \to \mathbb{R}^m\}_i$, $\{\psi_j : V_j \to \mathbb{R}^n\}_j$ Eine stetige Abbildung*

$$f : M \to N$$

heißt differenzierbar*, wenn es für jedes $x \in M$ Umgebungen U von x und V von $f(x)$ gibt, die ganz in Kartenumgebungen U_i und V_j enthalten sind und so dass $f^{-1}V \subset U$, und so dass die Kompositionen*

$$\psi_j \circ f \circ \phi_i^{-1} : \phi_i(U) \to \psi(V)$$

stetig differenzierabar sind.

Sie heißt Diffeomorphismus*, wenn sie bijektiv ist und die Umkehrabbildungen ebenfalls stetig differenzierbar sind.*

Als Spezialfall können wir die Identität betrachten. Ist sie differenzierbar, so nennen wir die Kartenwahlen *äquivalent*. Es ist üblich, Mannigfaltigkeiten nur bis auf Äquivalenz zu betrachten.

Beispiel. Die Mannigfaltigkeiten $(-\pi/2, \pi/2)$ und $\mathbb{R}$ (jeweils mit der offensichtlichen Karte gegeben durch die Identität) sind diffeomorph via der Abbildung tan.

Die Mannigfaltigkeiten S^1 und $\mathbb{R}$ sind nicht diffeomorph, denn sie sind nicht homöomorph: S^1 ist kompakt, $\mathbb{R}$ nicht.

Bemerkung. Wir erinnern an die Definition der Richtungableitung aus Definition 11.11. Er war so gewählt, dass er nicht nur die Koordinatenrichtungen, sondern beliebige Richtungen erlaubte. In der dort gewählten Form funktioniert er dann auch auf Mannigfaltigkeiten.

Extrema mit Nebenbedingungen

Wir betrachten $h, f : U \to \mathbb{R}$ mit $U \subset \mathbb{R}^n$ offen. Wir suchen nach lokale Extrama von h, aber nur in der Menge

$$\{x \in U | f(x) = 0\}$$

Wir sprechen von einem lokalen Extremum von h unter der *Nebenbedingung* f.

Satz 13.8. *Sei $U \subset \mathbb{R}^n$ offen, $h, f : U \to \mathbb{R}$ stetig partiell differenzierbar, $M = \{x \in U | f(x) = 0\}$. Es sei $a \in M$ mit $\mathrm{grad}(f)(a) \neq 0$ ein lokales Maximum (Minimum) von h unter der Nebenbedingung f. Dann gibt es $\lambda \in \mathbb{R}$ mit*

$$\mathrm{grad} h(a) = \lambda \mathrm{grad} f(a).$$

Die Zahl λ heißt *Lagrangescher Multiplikator*.

Beweis: Da $\mathrm{grad} f(a) \neq 0$, können wir annehmen, dass $\frac{\partial f}{\partial x_n}(a) \neq 0$. Wir schreiben $a = (a', a_n)$. Nach dem Satz über implizite Funktionen gibt es Umgebungen $W \subset \mathbb{R}^{n-1}$ von a' und $V \subset \mathbb{R}$ von a_n mit $W \times V \subset U$, sowie eine stetig partiell differenzierbare Funktion

$$g : W \to V$$

so dass

$$M \cap W \times V = \{x \in W \times V | x_n = g(x_1, \ldots, x_{n-1})\} .$$

Es ist also

$$f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) = 0$$

und aus der Kettenregel folgt wieder für $i = 1, \ldots, n-1$

$$0 = \frac{\partial f}{\partial x_i}(a) + \frac{\partial f}{\partial x_n}(a) \partial x_i(a') .$$

Wir betrachten

$$H : W \to \mathbb{R}, \qquad (x_1, \ldots, x_{n-1}) \mapsto h(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1}).$$

Diese Funktion hat in a' ein lokales Extremum, da a ein Extremum auf M ist. Aus dem notwendigen Kriterium für Extreme erhalten wir

$$0 = \frac{\partial H}{\partial x_i} = \frac{\partial h}{\partial x_i}(a) + \frac{\partial h}{\partial x_n}(a) \frac{\partial g}{\partial x_i}(a')$$

für $i = 1, \ldots, n-1$. Wir setzen

$$\lambda = \frac{\partial h}{\partial x_n}(a) \left(\frac{\partial f}{\partial x_n}(a) \right)^{-1} .$$

Einsetzen ergibt

$$\frac{\partial h}{\partial x_i}(a) = \lambda \frac{\partial f}{\partial x_i}(a)$$

für $i = 1, \ldots, n$ wie behauptet. $\square$

Beispiel. Sei $A = (a_{ij})$ eine symmetrische $n \times n$-Matrix und

$$q : \mathbb{R}^n \to \mathbb{R}, \qquad x \mapsto \langle x, Ax \rangle = \sum_{ij} a_{ij} x_i x_j$$

die zugehörige quadratische Form. Wir suchen ein Extremum unter der Nebenbedingung $\|x\| = 1$. Wir setzen daher

$$f(x) = \langle x, x \rangle - 1 = \sum_i x_i^2 - 1$$

Wir wenden den Satz an. Es ist

$$\mathrm{grad} f(x) = 2x,$$

daher ist grad$f \neq 0$ auf ganz $M = \{x | f(x) = 0\}$. Weiter ist

$$\frac{\partial q}{\partial x_i}(x) = \frac{\partial}{\partial x_i} \sum_{st} a_{st} x_s x_t = 2 \sum_s a_{is} x_s$$

(je ein Summand von $i = s$ und $i = t$). Zusammengefasst also

$$\text{grad} q = 2Ax$$

Die Bedingung lautet also

$$2Ax = \lambda 2x \to Ax = \lambda x.$$

Da die Menge M beschränkt und abgeschlossen ist, existiert das Maximum. Es gibt also ein $x \in \mathbb{R}^n$, dass die Gleichung erfüllt, also einen Eigenvektor. Wir haben damit gezeigt, dass jede symmetrische reelle Matrix einen reellen Eigenwert hat. Wie in der linearen Algebra folgt leicht, dass dann sogar alle Eigenwerte reell sind. Anders als beim dem Verfahren mit dem charakteristischen Polynom kann man aus obigen Argument ein numerisches Verfahren machen, das sehr schnell Eigenwerte und -vektoren bestimmt.

Kapitel 14

Allgemeine Theorie von Differentialgleichungen

Es geht um funktionelle Zusammenhänge zwischen einer Funktion und ihren Ableitungen. Wir beginnen mit ein paar Beispielen. Wir betrachten zunächst Funktionen $f : U \to \mathbb{R}$ mit $U \subset \mathbb{R}$ offen.

Beispiel. (i) $y' = 0$. Wir wissen, dass dann $y = c$ konstant ist. Wir sehen, dass f nicht eindeutig durch die Differentialgleichung festgelegt ist. Die Lösung y wird eindeutig durch die Vorgabe eines *Anfangswertes* $y(x_0)$. Genauer: y ist konstant auf jedem Intervall.

(ii) $y'' = 0$ auf einem Intervall. Alle Lösungen haben die Form $ax + b$ mit $a, b \in \mathbb{R}$. Diesmal werden zwei Anfangswerte gebraucht, also $y(x_0)$ und $y(x_1)$ oder $y(x_0)$ und $y'(x_0)$.

Beispiel. $y' = y$. Wir raten die Lösung $y = \exp$. Ist dies die einzige Lösung? Hier ist eine Methode zur Herleitung:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y \Leftrightarrow \mathrm{d}y = y\mathrm{d}x \Rightarrow \frac{1}{y}\mathrm{d}y = \mathrm{d}x$$

(Darf man das? Ja, $\mathrm{d}y$ ist eine *Differentialform.* Wir diskutieren das jetzt nicht.) Wir integrieren:

$$\int \frac{1}{y}\mathrm{d}y = \int \mathrm{d}x$$

(darf man das? Auf jeden Fall kommt das Richtige heraus) und erhalten

$$\log(y) = x + C \Rightarrow y = \exp(x + C) = C_0 \exp(x).$$

Andere Methode: Angenommen, y ist durch eine konvergente Potenzreihe $y(x) = \sum_{k=0}^{\infty} a_k x^k$ gegeben. Dann erhalten wir

$$y'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty} a_k x^k = y.$$

Koeffizientenvergleich ergibt das Gleichungssystem

$$ka_k = a_{k-1} \qquad k \geq 1.$$

Dieses ist lösbar mit a_0 beliebig, $a_k = a_0/k!$. Dies ergibt

$$y(x) = a_0 \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

Auch wenn wir uns vor allem für Funktionen mit Werten in $\mathbb{R}$ interessieren, ist es sinnvoll, von vornherein mit *Systemen* von Differentialgleichungen zu arbeiten.

Beispiel. $y'' = -y$. Durch Raten oder Potenzreihenansatz erhält man die allgemeine Lösung $y(x) = a \sin x + b \cos x$. Wir gehen statt dessen zu einem *System* von Differentialgleichungen über. Sei $Y_1 = f, Y_2 = y'$. Die Differentialgleichung ist nun äquivalent zu

$$Y_1' = Y_2, Y_2' = -Y_1 \Leftrightarrow Y' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} Y$$

mit $Y = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$.

Definition 14.1. *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen mit Koordinaten (t, x), $f : G \to \mathbb{R}^n$ stetig. Eine stetig differenzierbar Funktion $y : I \to \mathbb{R}$ mit $I \subset \mathbb{R}$ Intervall heißt* Lösung der Differentialgleichung

$$y' = f(\cdot, y)$$

falls $f(t, y(t)) \in U$ für alle $t \in I$ und

$$y'(t) = f(t, y(t)) \qquad \text{für alle } t \in I.$$

Gilt außerdem

$$y(t_0) = c_0$$

für ein vorgegebenes (t_0, c_0), so heißt y Lösung des Anfangswertproblems

Genauer handelt es sich um *gewöhnliche Differentialgleichungen* (oder ODE= ordinary differential equation, da g nur von einer Variabeln abhängt. Allgemeiner gibt es auch *partielle Differentialgleichungen*, in denen Ableitungen in verschiedene Koordinatenrichtungen verknüpft sind. Die Differentialgleichung hat *Ordnung 1*, da nur y und die erste Ableitung vorkommen. Allgemeiner kann man auch Gleichungen mit höheren Ableitungen betrachten. Dies liefert aber nicht neues, wie wir schon im dritten Beispiel gesehen haben.

Bevor wir uns um Lösungsverfahren kümmern, wollen wir Frage nach Existenz und Eindeutigkeit der Lösung behandeln.

Definition 14.2. *Sei $G \subset \mathbb{R}\times\mathbb{R}^n$ offen, $f : G \to \mathbb{R}^n$ eine Abbildung. Wir sagen, f genüge auf G einer* Lipschitz-Bedingung *mit* Lipschitz-Konstante $L \geq 0$*, falls*

$$\|f(t, y_1) - f(t, y_2)\| \leq L\|y_2 - y_1\|$$

Wir sagen, dass G lokal einer Lipschitz-Bedingung genügt*, wenn jeder Punkt (a, b) eine Umgebung hat, auf der eine Lipschitz-Bedingung erfüllt ist.*

Bemerkung. (i) Aus der Lipschitz-Bedingung folgt eine gleichmäßige Stetigkeit als Funktion von x.

(ii) Im Zusammenhang mit dem Satz über inverse Funktionen haben wir in Satz 13.3 gezeigt, dass stetig partiell differenzierbare Funktionen lokal eine Lipschitz-Bedingung erfüllen. Tatsächlich genügt partielle Differenzierbarkeit in Richtung $x_1, \dots, x_n$.

Satz 14.3 (Eindeutigkeit)**.** *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : G \to \mathbb{R}^n$ stetig und genüge lokal einer Lipschitz-Bedingung.*
Sei $I \subset \mathbb{R}$ ein Intervall, $t_0 \in I$, $y_0 \in \mathbb{R}^n$ mit $(t_0, y_0) \in G$. Dann hat das Anfangswertproblem

$$y' = f(\cdot, y), y(t_0) = y_0$$

höchstens eine stetig partiell differenzierbare Lösung.

Beweis: Seien y_1 und y_2 zwei Lösungen. Wir behandeln die Frage zunächst auf einem kleineren Intervall um t_0. Wir wählen $\delta > 0$, $r > 0$, so dass

$$(t_0 - \delta, t_0 + \delta) \times B(y_0, r) \subset G$$

und f auf dieser Menge einer Lipschitz-Bedingung mit Lipschitz-Konstante L erfüllt.
Durch Integration von t_0 nach t erhalten wir:

$$y_i'(t) = f(t, y_i(t)) \Rightarrow y_i(t) - y_i(t_0) = \int_{t_0}^{t} f(s, y_i(s))\mathrm{d}s.$$

Daher erhalten wir durch die Anfangswertbedingung

$$y_2(t) - y_1(t) = \int_{t_0}^{t} \left(f(s, y_2(s)) - f(s, y_1(s))\right) \mathrm{d}s.$$

Hieraus folgt

$$\|y_2(t) - y_1(t)\| \leq \left| \int_{t_0}^{t} L\|y_2(s) - y_1(s)\|\mathrm{d}s \right|$$

Sei

$$M(t) = \sup\{\|y_2(s) - y - 1(s)\| \,|\, |s - t_0| \leq |t - t_0|\}.$$

Dann gilt für $\tau \in (t_0 - \delta, t_0 + \delta)$ mit $|\tau - t_0| \leq |t - t_0|$:

$$\|y_2(\tau) - y_1(\tau)\| \leq L|\tau - t_0|M(\tau) \leq L|t - t_0|M(t)$$

und daher auch

$$M(t) \leq L|t - t_0|M(t).$$

Für t nahe bei t_0 (nämlich $|t - t_0| < 1/2L$) folgt

$$M(t) \leq \frac{1}{2}M(t).$$

Dies impliziert $M(t) = 0$ und daher $y_2(\tau) = y_1(\tau)$ für $|\tau - t_0| < \varepsilon$ mit $\varepsilon = \min(1/2L, \delta)$.

Die beiden Lösungen stimmen also in einer Umgebung von t_0 überein.

Der Rest ist ein Zusammenhangsargument. Wir betrachten die Menge

$$T = \{t \in I | y_1(s) = y_2(s) \text{ für alle } s \in [t_0, t]\} \ .$$

Diese Menge ist offen, da sie mit jedem t auch ein offenes Intervall um t enthält. (Wir wenden den ersten Teil an auf t statt t_0.)

Mit jedem t enthält sie alle anderen Elemente zwischen t_0 und t. Also ist T ein offenes Intervall. Angenommen, $T \subsetneq I$. Sei $t_1 = \sup T$. Da y_1 und y_2 stetig sind und beliebig nahe bei t_1 übereinstimmen, stimmen sie auch in t_1 überein. Damit ist $t_1 \in T$. Mit t_1 enthält T dann auch eine offene Umgebung von t_1. Dies ist ein Widerspruch zu $t_1 = \sup T$. Genauso argumentieren wir fürs Infimum. □

Wenn die Lipschitz-Bedingung nicht erfüllt ist, so sind die Lösungen im allgemeinen nicht eindeutig.

Beispiel. Sei $y' = y^{2/3}$. Die Funktion $y = 0$ ist eine Lösung. Für jedes $a \in \mathbb{R}$ erhalten wir außerdem die Lösung

$$f_a(t) = \frac{1}{27}(t - a)^3$$

Wegen $f_a(0) = 0$ ist der Eindeutigkeitssatz verletzt. Wir überprüfen die Lipschitz-Bedingung. Die Funktion $f(t, x) = x^{2/3}$ ist differenzierbar nach x mit

$$\frac{\partial f}{\partial x}f(t, x) = \frac{2}{3}x^{-1/3}$$

für $x \neq 0$. In keiner Umgebung von $(a, 0)$ ist die Lipschitz-Bedingung erfüllt.

Die Existenz einer Lösung ist im allgemeinen sehr schwierig.

Beispiel. Sei $f(t, y) = y^2$. Wir betrachten das Anfangswertproblem

$$y' = y^2, y(0) = 1.$$

Es hat auf $(-\infty, 1)$ die Lösung $1/(1 - t)$. Wegen

$$\lim_{t \to 1} \frac{1}{1 - t} \to \infty$$

lässt sich die Lösung nicht stetig auf ganz $\mathbb{R}$ fortsetzen.

Wir behandlen daher nur die Frage nach lokaler Existenz von Lösungen.

Theorem 14.4 (Picard-Lindelöf). *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : G \to \mathbb{R}^n$ stetige Abbildung, die lokal einer Lipschitz-Bedingung erfüllt. Dann gibt es zu jedem $(t_0, y_0) \in G$ ein $\varepsilon > 0$ und eine (eindeutige) Lösung*

$$y : (t_0 - \varepsilon, t_0 + \varepsilon) \to \mathbb{R}^n$$

der Differentialgleichung

$$y' = f(x, y)$$

mit der Anfangsbedingung $y(t_0) = y_0$.

Beweis: Wir wollen das Theorem mit Hilfe des Banachschen Fixpunktsatzes beweisen. Als vollständigen metrischen Raum verwenden wir stetige Funktionen auf einem Intervall I mit der Supremumsnorm. (Die Cauchy-Bedingung bezüglich der Supremumsnorm impliziert punkteweise die Cauchy-Bedingung, also existiert eine Grenzfunktion. Konvergenz bezüglich der Supremumsnorm bedeutet gleichmäßige Konvergenz. Nach Analysis 1 ist die Grenzfunktion stetig.)
Nach dem Hauptsatz der Differential- und Integralrechnung genügt die Integralgleichung zu lösen

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\mathrm{d}s \; .$$

Wir benutzen die Abbildung $y \mapsto F(y)$ ($y : I \to \mathbb{R}$ eine Funktion!)

$$(Fy)(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\mathrm{d}s.$$

Ein Fixpunkt von F ist dann die gesuchte Lösung.
Zu (t_0, y_0) wählen wir wieder eine Umgegung $(t_0 - \delta, t_0 + \delta) \times U$, so dass f auf der Umgebung eine Lipschitzbedingung mit Konstante L erfüllt. Sei $0 < \varepsilon \leq \varepsilon$ zunächst beliebig, aber fest. Sei $I = [t_0 - \varepsilon, t_0 + \varepsilon]$. Wir betrachten den Raum $C^{=}(I, \mathbb{R}^n)$ der stetigen Funktionen auf I mit der Supremumsnorm $\|y\|_\infty = \sup_{t \in I} \|y(t)\|$. Wir betrachten die Menge

$$X = \{y : I \to \mathbb{R}^n | \text{ stetig }, \|y - y_0\| \leq \delta\}.$$

Als abgeschlossene Teilmenge eines vollständigen metrischen Raums erhalten wir eine vollständigen metrischen Raum.
Wir zeigen nun für hinreichend kleines ε:

(i) $Fy \in X$ für $y \in X$,

(ii) $\|Fy_2 - Fy_1\|_\infty \leq \frac{1}{2}\|y_2 - y_1\|_\infty$ (Kontraktion).

Mit dem Banachschen Fixpunktsatz finden wir dann einen Fixpunkt; dies ist die gesuchte Lösung.
Als stetige Funktion ist f beschränkt auf $[t_0-\delta, t_0+\delta] \times \overline{B}(y_0, r) \subset [t_0-\delta, t_0+\delta] \times U$. Sei M das Supremum. Für $y \in X$, $y : I \to \mathbb{R}^n$ folgt

$$\|F(x)(t) - y_0\| = \|\int_{t_0}^t f(s, y(s))\mathrm{d}s\| \leq M|t-t_0| \leq M\varepsilon$$

Wir wählen $\varepsilon \leq \delta/M$. Dann ist die erste Behauptung erfüllt.
Weiter erhalten wir aus der Lipschitzbedingung:

$$\begin{aligned}\|F(y_2)(t) - F(y_1)(t)\| = \|&\int_{t_0}^t (f(s, y_2(s)) - f(s, y_1(s)))\mathrm{d}x\| \\ &\leq |t-t_0| \sup_s \|f(s, y_2(s)) - f(s, y_1(s))\| \\ &\leq \varepsilon L \sup_s \|y_2(s) - y_1(s)\|\end{aligned}$$

Wählen wir zusätzlich $\varepsilon \leq 1/2L$, so its auch die zweite Bedingung erfüllt. $\square$

Wir behandeln ein explizites Beispiel, in dem die Beweismethode eine Lösung liefert:

Beispiel. $y' = 2xy$, also $f(x, y) = 2xy$ als Funktion $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Wir suchen eine Lösung mit $y(0) = c$. Die Integralgleichung ist

$$y(x) = c + \int_0^x 2\xi y(\xi)\mathrm{d}\xi.$$

Daher ist das Interationsverfahren

$$y_{n+1}(x) = c + 2\int_0^x \xi y_n(\xi)\mathrm{d}\xi.$$

Mit $y_0 = c$ erhalten wir:

$$\begin{aligned} y_1(x) &= c + 2c\int_0^x \xi\mathrm{d}\xi = c(1+x^2) \\ y_2(x) &= c + 2c\int_0^x \xi(1+\xi^2)\mathrm{d}\xi = c(1+x^2+\frac{x^4}{2} \\ \ldots y_k(x) &= c(1+x^2+\frac{x^4}{2}+\cdots+\frac{x^{2k}}{k!})\end{aligned}$$

Die Folge konvergiert auf ganz $\mathbb{R}$ gegen

$$y = c\sum_{k=0}^\infty \frac{1}{k!}x^{2k} = c\exp(x^2).$$

Wir formulieren noch die Konsequenzen für Differentialgleichungen höherer Ordnung.

Definition 14.5. *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : G \to \mathbb{R}$ stetig. Dann heißt*

$$y^{(n)} = f(t, y, y', \dots, y^{(n-1)})$$

Differentialgleichung der Ordnung *n. Seien $(t_0, y_0, \dots, y_{n-1}) \in G$. Eine Funktion $y : I \to \mathbb{R}$ auf einem Intervall mit $t_0 \in I$ heißt Lösung des* Anfangswertproblems*, wenn sie diese Gleichung erfüllt und $y^{(i)}(t_0) = y_i$.*

Wir können die Differentialgleichung umschreiben auf ein System von Gleichungen erster Ordnung, in dem wir setzen:

$$g_1 = y, g_2 = y', \dots, g_n = y^{(n-1)}$$

und

$$F : G \to \mathbb{R}^n, \qquad F(t, x_1, \dots, x_n) = (x_1, x_2, \dots, x_{n-1}, f(t, x_1, \dots, x_n)).$$

Ist $g = (g_1, \dots, g_n)$ eine Lösung von F, so ist g_1 eine Lösung von f. Umgekehrt erhalten wir aus einer Lösung von f eine Lösung von F.

Korollar 14.6. *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : G \to \mathbb{R}^n$ stetig und erfülle eine Lipschitzbedingung. Sei $(t_0, y_0, \dots, y_{n-1}) \in G$. Dann ist die Lösung des Anfangswertproblems eindeutig auf jedem Intervall. Sie existiert in einer Umgebung von t_0.*

Beweis: Wir wenden den Existenz- und Eindeutigkeitssatz in dieser Situation an. □

Elementare Lösungsverfahren

Trennung der Variablen

Wir betrachten Differentialgleichungen der Form

$$y' = f(x)g(y)$$

wobei $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ beide stetig, I, J Intervalle. Wir setzen voraus, dass $g(y) \neq 0$ auf ganz J. Sie heißt *Differentialgleichung mit getrennten Variablen.*

Beispiel. $y' = y$ ist eine Differentialgleichung mit getrennten Variablen mit $f = 1$, $g(y) = y$.

Satz 14.7. *Sei $y' = f(x)g(y)$ wie oben. Sei $(x_0, y_0) \in I \times J$. Sei*

$$F(x) = \int_x^{x_0} f(t)\mathrm{d}t, \qquad G(y) = \int_{y_0}^{y} \frac{1}{g(s)}\mathrm{d}s$$

für $x \in I$, $y \in J$. Sei $I' \subset I$ ein Intervall mit $x_0 \in I'$ und $F(I') \subset G(J)$. Dann existiert genau eine Lösung

$$\phi : I' \to \mathbb{R}, \qquad \phi(x_0) = y_0.$$

Diese Lösung genügt der Bedingung

$$G(\phi(x)) = F(x) \qquad \text{für alle } x \in I'.$$

Beweis: Wir überprüfen zunächst die Bedingung. Sei ϕ eine Lösung des Anfangswertproblems, also

$$\phi'(x) = f(x)g(\phi(x)) \Rightarrow f(x) = \frac{\phi'(x)}{g(\phi(x))}.$$

Durch Integrieren mit der Substitution $u = \phi(t)$ erhalten wir

$$F(x) = \int_{x_0}^{x} f(t)\mathrm{d}t = \int_{x_0}^{x} \frac{\phi'(t)}{g(\phi(t)}\mathrm{d}t = \int_{y_0}^{\phi(x)} \frac{\mathrm{d}u}{g(u)} = G(\phi(x)).$$

Wir wenden uns der Eindeutigkeit zu. Die Funktion G ist streng monoton, da $G' = 1/g \neq 0$. Sei $H : G(J) \to \mathbb{R}$ die Umkehrfunktion. Aus der Relation folgt also für alle $x \in I'$

$$\phi(x) = H(F(x))\ .$$

Schließlich nehmen wir diese Gleichung als Definition. Da F und H stetig differenzierbar sind (Stammfunktionen stetiger Funktionen sind steig differenzierbar, Umkehrfunktionen stetig differenzierbarer Funktionen auch), ist es ϕ auch. Anwenden von G ergibt wieder

$$G(\phi(x)) = F(x).$$

Hieraus erhalten wir durch Differenzieren

$$G'(\phi(x))\phi'(x) = F'(x) \Rightarrow \frac{\phi'(x)}{g(\phi(x))} = f(x).$$

Damit löst ϕ die Differentialgleichung. Wegen $G(y_0) = 0$

$$\phi(x_0) = H(F(x_0)) = H(0) \Rightarrow \phi(x_0) = y_0.$$

$\square$

Beispiel.

$$y' = y^2, \qquad y(0) = c \in \mathbb{R}.$$

Der Eindeutigkeitssatz garantiert uns die Eindeutigkeit von Lösungen. Wir lösen mit Trennung der Variablen. Was sind I, J, I'? Wir rechnen erstmal drauf los.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y^2 \Rightarrow \frac{\mathrm{d}y}{y^2} = \mathrm{d}x \Rightarrow \int_c^y \frac{\mathrm{d}t}{t^2} = \int_0^x \mathrm{d}x \Rightarrow -\frac{1}{t}\bigg|_c^y = -\frac{1}{y} + \frac{1}{c} = x \Rightarrow y = \frac{c}{1-cx}$$

Offensichtlich gibt es ein Problem mit dieser Rechnung, wenn $c = 0$. In diesem Fall sehen wir aber die offensichtliche Lösunge $y = 0$ und die Formel stimmt sogar.
Sei also jetzt $c \neq 0$. Die Lösung wird singulär in $x = c^{-1}$. Für $c > 0$ erhalten wir also eine Lösung auf $(-\infty, c^{-1})$, für $c < 0$ auf (c^{-1}, ∞), da 0 in I' liegen soll.
Beim Integrieren dürfen wir $y = 0$ nicht überschreiten. Ist $c > 0$, so ist also das maximale $J = \mathbb{R}_{>0}$. Wir erhalten dasselbe I' aus dem Intervall

$$G(J) = (-\frac{1}{y} + \frac{1}{c})(\mathbb{R}^{>0}) = (-\infty, \frac{1}{c}) = I'$$

Beispiel. Wir betrachten die *lineare homogene Differentialgleichung*

$$y' = a(x)y, \qquad y(x_0) = c$$

mit $a : I \to \mathbb{R}$ stetig, $x_0 \in I$. Wir lösen mit Trennung der Variablen

$$\frac{\mathrm{d}y}{\mathrm{d}x} = a(x)y \Rightarrow \frac{\mathrm{d}y}{y} = a(x)\mathrm{d}x \Rightarrow \int_c^y \frac{\mathrm{d}t}{t} = \int_{x_0}^x a(s)\mathrm{d}s$$
$$\Rightarrow \log(y) = \int_{x_0}^x a(s)\mathrm{d}s + \log(c)$$

Hieraus erhalten wir die Lösung

$$y = c\exp\left(\int_{x_0}^x a(s)\mathrm{d}s\right).$$

Sie ist wohldefiniert auf ganz I.

Variation der Konstanten

Wir erhalten die Lösung der *inhomogenen linearen Differentialgleichung*

$$y' = a(x)y + b(x)$$

$(a, b : I \to \mathbb{R}$ stetig) mit der Methode der *Variation der Konstanten*.
Sei ϕ die Lösung der zugehörigen homogenen Differentialgleichung

$$y' = a(x)y.$$

Wir machen den Ansatz

$$\psi(x) = c(x)\phi(x).$$

Wenn der Anfangswert $y(x_0) = c \neq 0$, so sind alle Werte von ϕ ungleich null und jede Lösung lässt sich in dieser Form schreiben.
Wir erhalten die Bedingung:

$$\psi'(x) = c'(x)\phi(x) + c(x)\phi'(x) = a(x)c(x)\phi(x) + b(x) = a(x)\psi(x) + b(x).$$

Nach Voraussetzung ist

$$\phi'(x) = a(x)\phi(x),$$

daher erhalten wir

$$c'(x)\phi(x) = b(x) \Rightarrow c'(x) = \frac{b(x)}{\phi(x)} \Rightarrow c(x) = \int_{x_0}^{x} \frac{b(s)}{\phi(s)} \mathrm{d}s + c$$

Wir haben daher gezeigt:

Satz 14.8. *Seien I ein Intervall, $a, b : I \to \mathbb{R}$ stetig. Sei $x_0 \in I$, $c \in \mathbb{R}$. Dann gibt es genau eine Lösung der $\psi : I \to \mathbb{R}$ der Differentialgleichung*

$$y' = a(x)y + b(x), \qquad y(x_0) = c,$$

nämlich

$$\psi(x) = \phi(x) \left(c + \int_{x_0}^{x} \frac{b(s)}{\phi(s)} \mathrm{d}s \right)$$

wobei

$$\phi(x) = \exp\left(\int_{x_0}^{x} a(s) \mathrm{d}s \right).$$

Beispiel. Wir behandeln die inhomogene Gleichung

$$y' = 2xy + x^3.$$

Die homogene Gleichung hatten wir bereits gelöst: $\phi(x) = \exp(x^2)$ erfüllt

$$y' = 2xy.$$

Wir machen den Ansatz $\psi(x) = c(x)\exp(x^2)$ und erhalten

$$c'(x)\exp(x^2) + c(x)2x\exp(x^2) = 2xc(x)\exp(x^2) + x^3$$
$$\Rightarrow c'(x) = \exp(-x^2)x^3 \Rightarrow c(x) = c + \int_0^x t^3 \exp(-t^2)\mathrm{d}t.$$

Das Integral berechnet man mit der Substituion $s = t^2$ und partieller Integration und erhält

$$\int_0^x t^3 \exp(-t^2)\mathrm{d}x = \frac{1}{2} - \frac{1}{2}(x^2 + 1)\exp(-x^2)$$

Die allgemeine Lösung ist daher

$$\psi(x) = (c - \frac{1}{2}(x^2 + 1)\exp(-x^2))\exp(x^2) = c\exp(x^2) - \frac{1}{2}(x^2 + 1)$$

Exakte Differentialgleichungen

Wir betrachten Differentialgleichungen der Form

$$f(x,y) + g(x,y)y' = 0$$

und fassen sie auf als

$$f(x,y)\mathrm{d}x + g(x,y)\mathrm{d}y = 0.$$

Die Lösung wird gesucht über den Ansatz:

$$f(x,y) = \frac{\partial F(x,y)}{\partial x}, g(x,y) = \frac{\partial F(x,y)}{\partial y}.$$

Nach dem Satz über implizte Funktionen lässt sich dann y bestimmen aus der Gleichung

$$F(x,y) = c$$

mit einer Konstanten c, die von den Anfangsbedingungen abhängt. Notwendige Bedingung für die Existenz einer stetig differenzierbaren Funktion F ist die Gleichung

$$\frac{\partial}{\partial y}\frac{\partial F}{\partial x} = \frac{\partial}{\partial x}\frac{\partial F}{\partial y} \Leftrightarrow \frac{\partial f}{\partial y} = \frac{\partial g}{\partial x}.$$

Ist diese Gleichheit erfüllt, so bestimmen wir $F(x,y)$ durch Integration von f in Richtung x bis auf eine Funktion, die nur von y abhängt. Diese wird durch die zweite Bedingung bestimmt. Die Gleichheit der partiellen Ableitungen ist genau die Bedingung, die garantiert, dass das Verfahren funktioniert.

Beispiel. Wir behandeln

$$(2x + 4y + 2) + (4x + 12y + 8)y' = 0.$$

Wir überprüfen die Exaktheit:

$$\frac{\partial(2x + 4y + 2)}{\partial y} = 4 = \frac{\partial(4x + 12y + 8)}{\partial x}$$

Wir bilden

$$\int (2x + 4y + 2)\mathrm{d}x = x^2 + 4yx + 2x + h(y), \int (4x + 12y + 8)\mathrm{d}y = 4xy + 6y^2$$

Durch Verleichen sehen wir

$$F(x,y) = x^2 + 4xy + 2x + 6y^2 + c$$

Es handelt sich um eine Ellipsengleichung.

In physikalischer Sprache:
Das Vektorfeld $\begin{pmatrix} f \\ g \end{pmatrix}$ ist die Ableitung eines Potentials F, also $\nabla F = \begin{pmatrix} f \\ g \end{pmatrix}$, falls seine Rotation verschwindet.
Gelegentlich lässt sich die Methode anwenden, auch wenn die Differentialgleichung nicht exakt ist. Sei $\mu(x, y)$ eine Funktion, so dass

$$\mu(x, y)f(x, y) + \mu(x, y)g(x, y)y' = 0$$

eine exakte Differentialgleichung ist. Die Funktion μ heißt *integrierender Faktor*. Aus dem Potential für für diese Gleichung lässt sich dann wieder y bestimmen.

Beispiel.

$$(y^2 - 3xy - 2x^2) + (xy - x^2)y' = 0$$

also

$$p(x, y) = y^2 - 3xy - 2x^2, \qquad q(x, y) = xy - x^2$$

Wir haben

$$\frac{\partial p}{\partial y} = 2y - 3x, \qquad \frac{\partial q}{\partial x} = y - 2x$$

Die Differentialgleichung ist also nicht exakt. Wir wählen als integrierenden Faktor x und erhalten

$$\frac{\partial xp}{\partial y} = 2xy - 3x^2, \qquad \frac{\partial xq}{\partial x} = 2xy - 3x^2$$

Nun ist die Differentialgleichung exakt. Das Potential berechnet sich zu

$$F(x, y) = \int (y^2x - 3x^2y - 2x^3)\mathrm{d}x + a(y) = \frac{1}{2}y^2x^2 - x^3y - \frac{1}{2}x^4 + a(y)$$

bzw.

$$F(x, y) = \int (x^2y - x^3)\mathrm{d}y + b(x) = \frac{1}{2}x^2y^2 - x^3y + b(x)$$

und daher

$$F(x, y) = \frac{1}{2}y^2x^2 - x^3y - \frac{1}{2}x^4 + c$$

Kapitel 15

Lineare Differentialgleichungen

Definition 15.1. *Sei $I \subset \mathbb{R}$ ein Intervall,*

$$A = (a_{ij})_{i,j=1}^n : I \to M_{n\times n}(\mathbb{R})$$

eine stetige matrixwertige Funktion, d.h. alle a_{ij} sind stetig. Dann heißt

$$y' = A(x)y$$

homogenes lineares System von Differentialgleichungen*. Weiter sei*

$$b = (b_i)_{i=1}^n : I \to \mathbb{R}^n$$

stetige vektorwertige Funktion, d.h. alle b_i sind stetig. Dann heißt

$$y' = A(x)y + b$$

inhomogenes lineares Differentialgleichungssystem*.*

Bemerkung. Da in den Lösungsformeln Exponetialfunktionen vorkommen, ist es sinnvoll auch Funktionen A und b mit Werten in $M_{n\times n}(\mathbb{R})$ und $\mathbb{C}^n$ zu betrachten. Es ist aber weiterhin $x \in I \subset \mathbb{R}$ eine reelle Variable. Für Fragen der Differenzierbarkeit etc. fassen wir einfach $\mathbb{C}$ als $\mathbb{R}^2$ auf.

In der Sprache des letzten Kapitels handelt es sich um die Differentialgleichung

$$y' = F(x, y)$$

mit

$$F(x, y) = A(x)y + b(x).$$

Die Lipschitzbedingung ist lokal erfüllt, denn F ist stetig partiell nach y differenzierbar. Oder direkt: auf einer kompakten Umgebung J eines $x_0 \in I$ gilt

$$\|F(x,y_2)-F(x,y_1)\| = \|A(x)y_2+b(x)-A(x)y_2-b(x)\| = \|A(x)(y_2-y_1)\| \leq L\|y_2-y_1\|$$

mit

$$L = \sup_{x\in J} \|A(x)\| = \sup_{x\in J} \sum_{\|v\|=1} \|A(x)v\|.$$

Das Supremum existiert wegen der Stetigkeit von A als Funktion von x.

Aus der allgemeinen Theorie wissen wir also, dass das Anfangswertproblem eindeutig lösbar ist und Lösungen lokal existieren. Tatsächlich ist die Situation viel besser.

Satz 15.2 (Existenz- und Eindeutigkeitssatz)**.** *Seien I, A und b wie oben. Dann gibt es zu jedem $x_0 \in I$ und $c \in \mathbb{R}^n$ genau eine Lösung*

$$\phi : I \to \mathbb{R}^n$$

der linearen Differentialgleichung

$$y' = Ay + b$$

mit $y(x_0) = c$.

Beweis: Die Eindeutigkeit folgt wie oben aus dem allgemeinen Eindeutigkeitssatz. Für die Existenz können wir annehmen, dass I kompakt ist (im allgemeinen schreiben wir I als Vereinigung von immer größeren kompakten Intervallen; wegen der Eindeutigkeit passen die Lösungen zusammen). Wir behandeln die Existenz, indem wir das Argument aus dem Beweis des Existenzsatzes von Picard und Lindelöf wiederholen. Wir betrachten den vollständigen metrischen Raum der stetigen Funktionen $\phi : J \to \mathbb{R}^n$ bezüglich der Supremumsnorm. Die Lösung der Differentialgleichung finden wir als Fixpunkt des Operators $F : \phi \mapsto F(\phi)$

$$F(\phi)(x) = c + \int_{x_0}^{x} (A(t)\phi(t) + b(t))\mathrm{d}t.$$

Die Bildfunktion ist wieder stetig.

Wir überprüfen die Kontraktionseigenschaft. Es ist

$$F(\phi_2)(x) - F(\phi_1)(x) = \int_{x_0}^{x} A(t)(\phi_2(t) - \phi_1(t))\mathrm{d}t$$

und daher

$$\|F(\phi_2)(x) - F(\phi_1)(x)\| \leq L|x - x_0|\|\phi_2 - \phi_1(t)\| \ .$$

Dies ist eine Kontraktion für $L|x - x_0| < 1$, also auf $J \cap |x_0 - \delta, x_0 + \delta|$ mit eingem $\delta < L^{-1}$. Nach dem Banachschen Fixpunktsatz erhalten wir eine Lösung auf diesem Teilintervall. Man beachte, dass die Schranke weder von x_0 noch von c abhängt. Wir können also das Intervall J durch Teilintervalle der Breite δ überdecken und die Lösungen zusammenbauen zu einer globalen Lösung. □

Die Lösungen der homogenen Gleichung

$$y' = Ay$$

bilden einen $\mathbb{R}$-Vektorraum, denn sind y_1 und y_2 Lösungen, dann auch

$$(y_1 + y_2)' = y_1' + y_2' = Ay_1 + Ay_2 = A(y_1 + y_2).$$

Wir kennen auch seine Dimension, denn die Lösungen sind eindeutig durch den Wert in x_0 bestimmt.

Satz 15.3 (Fundamentallösungen)**.** *Sei V_H der Vektorraum der Lösungen der homogenen linearen Differentialgleichung*

$$y' = Ay$$

mit $A : I \to M_{n\times n}(\mathbb{R})$ stetig. Dann hat V_H die Dimension n. Für ein n-Tupel $(\phi_1, \dots, \phi_n)$ von Lösungen sind äquivalent:

(i) *$\phi_1, \dots, \phi_n$ sind linear unabhängig über $\mathbb{R}$.*

(ii) *Es existiert ein $x_0 \in I$, dass die Vektoren $\phi_1(x_0), \dots, \phi_n(x_0)$ linear unabhängig in $\mathbb{R}^n$ sind.*

(iii) *Für jedes $x_0 \in I$ sind die Vektoren $\phi_1(x_0), \dots, \phi_n(x_0)$ linear unabhängig in $\mathbb{R}^n$.*

Beweis: Sei $x_i \in I$. Die Auswertungsabbildung

$$V_H \to \mathbb{R}^n, \qquad \phi \mapsto \phi(x_0)$$

ist linear. Sie ist bijektiv nach dem Existenz- und Eindeutigkeitssatz für homogene lineare Differentialgleichungen. Die Urbilder einer Basis bilden also eine Basis und umgekehrt. □

Ein Tupel mit diesen Eigenschaften heißt *Lösungs-Fundamentalsystem* der Differentialgleichung. Es handelt sich um eine Basis von V_H. Schreibt man die ϕ_i als Vektoren

$$\begin{pmatrix} \phi_{i1} \\ \dots \\ \phi_{in} \end{pmatrix}$$

so erhalten wir die *Fundamentalmatrix*

$$\Phi = (\phi_{ij})_{i,j=1}^n : I \to M_{n\times n}(\mathbb{R}).$$

Die lineare Unabhängigkeit übersetzt sich in

$$\det(\Phi)(x_0) \neq 0.$$

Der Satz besagt, dass dies in allen $x \in I$ gilt, sobald es in einem $x_0 \in I$ erfüllt ist. Die Fundamentalmatrix erfüllt selbst die Gleichung

$$\Phi' = A\Phi$$

(diese Aussage fasst nur die Differentialgleichung für jedes ϕ_i zusammen.)

Korollar 15.4. *Sei $(\phi_1, \dots, \phi_n)$ ein Fundamentalsystem der homogenen linearen Differentialgleichung $y' = Ay$. Dann lässt sich jede Lösung eindeutig in der Form*

$$\phi = \sum_{i=1}^{n} \lambda_i \phi_i$$

schreiben.

Beweis: Definition einer Basis. □

Beispiel. Wir behandeln das System

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

(Es gehört zur linearen Gleichung 2. Grades $y'' = -y$). Es hat die Lösungen

$$\begin{pmatrix} \sin \\ \cos \end{pmatrix}, \qquad \begin{pmatrix} \cos \\ -\sin \end{pmatrix}.$$

Dies ist ein Fundamentalsystem, denn

$$\det \begin{pmatrix} \sin & \cos \\ \cos & -\sin \end{pmatrix} = -(\sin^2 + \cos^2) = -1 \neq 0$$

für jedes x.

Das sollte erinnern an entsprechenden Aussagen aus der linearen Algebra: Die Lösungsmenge eines homogenen linearen Gleichungssystems ist ein Vektorraum. Die Lösung eines inhomogenen linearen Gleichungssystems ist ein affiner Raum.

Satz 15.5. *Seien I, A, b wie oben. Die Lösungsmenge der inhomogenen linearen Differentialgleichung*

$$y' = Ay + b$$

ist ein affiner Raum über dem Vektorraum V_H der Lösungen der zugehörigen homogenen Differentialgleichung $y' = Ay$, d.h.: Sei ϕ_0 eine Lösung der inhomogenen Gleichung. Dann lässt sich jede Lösung eindeutig schreiben als

$$\psi = \psi_0 + \phi$$

wobei ϕ die homogene Gleichung löst.

Beweis: Seien ψ_0, ψ Lösungen von $y' = Ay + b$. Dann erfüllt $\phi = \psi - \psi_0$ die Gleicung

$$\phi' = \psi' - \psi_0' = A\psi + b - A\psi_0 - b = A(\psi - \psi_0) = A\phi.$$

Ist umgekehrt ϕ eine Lösung des homogenen Systems, so erfüllt $\psi = \psi_0 + \phi$ die inhomogene Gleichung. □

Im eindimensionalen Fall haben wir bereits aus der Lösung der homogenen Gleichung durch Variation der Konstanten die Lösung der homogenen bestimmt. Dasselbe funktioniert auch für $n > 1$.

Satz 15.6 (Variation der Konstanten)**.** *Sein I, A, b wie oben. Sei $\Phi = (\phi_1, \dots, \phi_n)$ ein Lösungs-Fundamentalsystem der homogenen linearen Gleichung*

$$y' = Ay \ .$$

Dann erhält man eine Lösung $\psi : I \to \mathbb{R}^n$ der Gleichung

$$y' = Ay + b$$

durch den Ansatz

$$\Psi(x) = \Phi(x)u(x) \ .$$

Dabei ist $u : I \to \mathbb{R}^n$ eine differenzierbare Funktion mit

$$\Phi(x)u'(x) = b(x) \Leftrightarrow u(x) = c + \int_{x_0}^{x} \Phi(t)^{-1} b(t) \mathrm{d}t \ .$$

Beweis: Sei ψ von der angegebenen Form, also mit $\Phi u' = b$. Dann folgt

$$\psi' = \Phi' u + \Phi u' = A\Phi u + b = A\psi + b.$$

Wir haben also eine Lösung gefunden. □

Lineare Differentialgleichungen höherer Ordnung

Wegen der besonderen Wichtigkeit wollen wir die Ergebnisse umschreiben auf Differentialgleichungen höherer Ordnung.

Definition 15.7. *Sei $I \subset \mathbb{R}$ ein Intervall, für $0 \leq k < n$ sein $a_k : I \to \mathbb{R}$ stetig. Dann heißt*

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_0 y = 0$$

homogene lineare Differentialgleichung der Ordnung n. *Ist $b : I \to \mathbb{R}$ eine weitere stetige Funktion, so heißt*

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_0 y = b$$

inhomogene lineare Differentialgleichung der Ordnung n.

Wir wenden unser Verfahren zum Übersetzen in ein Differentialgleichungssystem der Ordnung 1 an.

$$\begin{aligned} y_0 &= y, \\ y_1 &= y', \\ &\dots, \\ y_{n-1} = y^{(n-1)} \Rightarrow y'_{n-1} &= -a_{n-1} y_{n-1} - \dots - a_0 y_0. \end{aligned}$$

Satz 15.8 (Lineare Differentialgleichungen höherer Ordnung)**.** *Seien I, a_i, b, wie oben.*

(i) Sei $x_0 \in I$, $y_0, \dots, y_{n-1} \in \mathbb{R}$. Die inhomogene lineare Differentialgleichung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = b$$

hat eine eindeutige globale Lösung mit $y^{(i)}(x_0) = y_i$ für alle i.

(ii) Die Menge der Lösungen der homogenen Gleichung (mit $b = 0$) ist ein Vektorraum V_H der Dimension n.

(iii) Ein Tupel von Lösungen $(\phi_1, \dots, \phi_n)$ ist genau dann linear unabhängig, wenn die Wronski-Determinante

$$W(x) = \det \begin{pmatrix} \phi_1 & \dots & \phi_n \\ \phi_1' & \dots & \phi_n' \\ & \dots & \\ \phi_1^{(n-1)} & \dots & \phi_n^{(n-1)} \end{pmatrix}$$

in einem (und dann in jedem) $x \in I$ ungleich 0 ist.

(iv) Die Menge der Lösungen der inhomogenen Gleichung ist ein affiner Raum der Dimension n. Ist ψ_0 eine Lösung, so erhalten wir alle andere Lösungen eindeutig in der Form

$$\psi = \phi + \psi_0$$

wobei $\phi \in V_H$.

Beweis: Wir wenden die bisherigen Sätze über Differentialgleichungssysteme an. □

Eine Basis von V_H heißt *Lösungs-Fundamentalsystem.*

Beispiel. Wir betrachten

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0.$$

Diese homogene lineare Differentialgleichung der Ordnung 2 hat auf $I = (0, \infty)$ die Lösungen x und $\sqrt{x}$. Die Wronski-Determinante ist

$$W(x) = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{pmatrix} = \frac{\sqrt{x}}{2} - \sqrt{x} = \frac{\sqrt{x}}{2}$$

Dies ist ungleich 0 in $x = 1$, tatsächlich auf ganz $(0, \infty)$.

In der Physik treten besonders oft Differentialgleichungen für die Beschleunigung auf, also der Ordnung 2.

Satz 15.9 (Reduktion der Ordnung)**.** *Sei $I \subset \mathbb{R}$ ein Intervall, $a, b : I \to \mathbb{R}$ stetig. Sei $\phi : I \to \mathbb{R}$ eine Lösung der Differentialgleichung*

$$y'' + a(x)y' + b(x)y = 0.$$

Sei $\phi(x) \neq 0$ auf I. Dann erhält man eine zweite von ϕ linear unabhängige Lösung mit dem Ansatz

$$\psi(x) = u(x)\phi(x)$$

wobei u eine nicht-konstante Lösung der Differentialgleichung

$$u'' + \left(2\frac{\phi'(x)}{\phi(x)} + a(x)\right)u' = 0$$

ist.

Der Vorteil ist also, dass die Differentialgleichun für u' nur Ordnung 1. Solche Differentialgleichungen haben wir bereits gelöst.

Beweis: Es gibt eine von ϕ linear unabhängige Lösung. Da $\phi(x) \neq 0$, lässt sie sich wie im Ansatz schreiben. Es folgt

$$\psi'(x) = u'(x)\phi(x) + u(x)\phi'(x), \psi''(x) = u''(x)\phi(x) + 2u'(x)\phi'(x) + u(x)\phi''(x)$$

und mit der Differentialgleichung für ψ

$$u''(x)\phi(x)+2u'(x)\phi'(x)+u(x)\phi''(x) = a(x)\left(u'(x)\phi(x) + u(x)\phi'(x)\right)+b(x)u(x)\phi(x)$$

Da ϕ die Differentialgleichung erfüllt, folgt die Behauptung. $\square$

Beispiel. Wir behandeln die *Legendresche Differentialgleichung* auf $I = (-1, 1)$

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0$$

für $n \in \mathbb{N}$. Diese wird gelöst durch das *Legendre-Polynom*

$$P_n(x) = \frac{1}{2^n n!}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^n (x^2 - 1)^n.$$

Es handelt sich um ein Polynom vom Grad n. Der Vorfaktor erklärt sich aus Orthgonalitätsrelationen, die wir nicht diskutiere wollen. Wir verzichten auf den Beweis.
Sei speziell $n = 1$. Die Differentialgleichung lautet dann

$$(1 - x^2)y'' - 2xy' + 2y = 0 \Leftrightarrow y'' - \frac{2x}{1 - x^2}y' + \frac{2}{1 - x^2}y = 0$$

mit der Lösung $P_1(x) = x$. Die zweite, dazu lineare unabhängige Lösung hat die Form

$$\psi(x) = u(x)x$$

mit

$$u'' + \left(\frac{2}{x} - \frac{2x}{1-x^2}\right) u' = 0$$

Die Lösung erhält man als

$$\begin{aligned} u' &= \exp\left(-\int \frac{2}{x}\mathrm{d}x + \int \frac{2x}{1-x^2}\mathrm{d}x\right) \\ &= \exp(-2\log(x)) - \log(1-x^2)) \\ &= \frac{1}{x^2(1-x^2)} \end{aligned}$$

mit Stammfunktion

$$u(x) = -\frac{1}{x} + \frac{1}{2}\log\frac{1+x}{1-x}$$

Die zweite Lösung ist also

$$\psi = -1 + \frac{x}{2}\log\frac{1+x}{1-x}.$$

Beispiel. Wir behandlen die *Hermitesche Differentialgleichung*

$$xy'' - (1-x)y' + 2ny = 0$$

für $n \in \mathbb{N}$. Sie taucht bei der quantenmechanischen Behandlung des harmonischen Oszillators auf als zeitunabhängige Schrödingergleichung. Die Zahl $2n$ charakterisiert das (diskrete) Energieniveau. Sie wird gelöst durch das *Hermitsche Polynom*

$$H_n(x) = (-1)^n \exp(x^2) D^n \exp(-x^2).$$

Beweis: Die entscheidende Rechneregel ist $D^{n+1}(xf) = xD^{n+1}f + (n+1)D^n f$. Es gilt mit $y = \exp(x^2)D^n\exp(-x^2)$ und daher $\exp(-x^2)y = D^n\exp(-x^2)$:

$$\begin{aligned} D^2(\exp(-x^2)y) &= D^{n+2}\exp(-x^2) \\ &= D^{n+1}(-2x\exp(-x^2)) \\ &= -2xD^{n+1}\exp(-x^2) - 2(n+1)D^n\exp(-x^2) \\ &= -2xD(\exp(-x^2)y) - 2(n+1)\exp(-x^2)y \\ &= \exp(-x^2)(4x^2y - 2xy' - 2(n+1)y) \end{aligned}$$

wegen $D^{n+1}(xf) = xD^{n+1}f + (n+1)D^n f$ Andererseit ist

$$\begin{aligned} D^2\exp(-x^2y) &= \exp(-x^2)y'' + 2(D\exp(-x^2))y' + (D^2\exp(-x^2))y \\ &= \exp(-x^2)(y'' - 4xy' + (4x^2-2)y) \end{aligned}$$

Durch Vergleich erhalten wir

$$y'' - 2xy' + 2ny = 0.$$

□

Lineare Differentialgleichungen mit konstanten Koffizienten

Wir kehren zurück zur Differentialgleichung

$$y' = Ay$$

aber mit der Einschränkung, dass $A \in M_{n\times n}(\mathbb{R})$ und $b \in \mathbb{R}^n$ konstant sind. Für $n = 1$ ist die Lösung durch die Exponentialfunktion $\exp(Ax)$ gegeben.

Definition 15.10. *Sei $A \in M_{n\times n}(\mathbb{C})$. Wir definieren*

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

Sie erfüllt

$$\exp(A + B) = \exp(A)\exp(B)$$

falls $AB = BA$ und

$$\exp(SAS^{-1}) = S\exp(A)S^{-1}$$

Lemma 15.11. *Die matrixwertige Exponentialreihe konvergiert.*

Beweis: Wir betrachten die Operatornorm auf $M_{n\times n}(\mathbb{R})$. Es handelt sich um eine Norm, denn

$$\|A + B\| \leq \|A\| + \|B\|$$

da für alle $v \in \mathbb{R}^n$

$$\|(A + B)v\| = \|Av + Bv\| \leq \|Av\| + \|Bv\| \leq (\|A\| + \|B\|)\|v\|.$$

Es gilt

$$\|A^k\| \leq \|A\|^k$$

den für jedes $v \in \mathbb{R}^n$ mit $\|v\| = 1$ ist

$$\|A^2 v\| = \|AAv\| \leq \|A\|\|Av\| \leq \|A\|^2\|v\| = \|A\|^2.$$

Da die Exponentialfunktion in $\|A\|$ konvergiert, bilden die die Partialsummen der matrixwertigen Exponentialreihe eine Cauchyfolge bezüglich der Operatornorm. Da $M_{n\times n}(\mathbb{R}) \cong \mathbb{R}^{n^2}$ vollständig ist, konvergiert die Reihe.

Der Beweis der Formel für das Cauchy-Produkt folgt wie im eindimensionalen Fall, da die Reihe absolut konvergiert. Die Vertauschungsbedingung taucht beim Ausmultiplizieren der Partialsummen auf. □

Satz 15.12. *Sei $A \in M_{n\times n}(\mathbb{R})$, $c \in \mathbb{R}^n$. Dann löst*

$$y(t) = \exp(tA)c$$

das Anfangswertproblem

$$y' = Ay, \qquad y(0) = c.$$

Mit anderen Worten: $\exp(tA)$ ist eine Lösungsfundamentalsystem.

Beweis: Wir bestimmen die Ableitung von $t \mapsto y(t) = \exp(tA)c$.

$$y'(t) = \lim_{h\to 0} \frac{1}{h}(\exp((t+h)A) - \exp(tA)) = \lim_{h\to 0} \frac{1}{h}(\exp(hA) - \exp(0)) \exp(tA)$$

und berechnen den Grenzwert auf Reihenniveau:

$$\frac{1}{h}(\sum_{k=0}^{\infty} \frac{h^k}{k!} A^k - E_n) = \sum_{k=1}^{\infty} \frac{h^{k-1}}{k!} A^k = A$$

da auch diese Reihe absolut und lokal gleichmäßig konvergiert und daher der Grenwert mit der Reihe vertauscht. □

Die Lösung des Problem läuft also auf das Berechnen der Exponentialfunktion hinaus. In manchen Fällen ist das ganz einfach.

(i) Sei also A, wobei Λ eine Diagonalmatrix mit Einträgen $\lambda_1, \dots, \lambda_n$ ist. Diagonale 0. Dann ist $\exp \Lambda$ eine Diagonalmatrix mit Einträgen $\exp(\lambda_1), \dots, \exp(\lambda_n)$.

(ii) Sei $A = N$ eine obere Dreiecksmatrix mit Diagonale 0. Die Matrix N ist *nilpotent*, d.h. $N^n = 0$ und daher

$$\exp(N) = \sum_{k=0}^{n} \frac{1}{k!} N^k$$

(iii) Sei A eine diagonalisierbare Matrix, d.h. es gibt eine invertierbare Matrix S mit $A = S\Lambda S^{-1}$ mit einer Diagonalmatrix Λ. Dann gilt

$$\exp(A) = S \exp(\Lambda) S^{-1}.$$

(iv) Leider ist nicht jede Matrix diagonalisierbar. Dies ist der Punkt, an dem es von Vorteil ist, zu komplexen Koeffizienten überzugehen. Nach dem Satz über die Jordansche Normalform ist jedes A konjugiert zu einer Blockdiagonalmatrix aus Blöcken $J_{m,\lambda}$ wobei $J_{m,\lambda}$ eine obere Dreiecksmatrix ist mit allen Diagonaleinträgen λ. Die Matrix $\exp(A)$ erhalte wir durch Konjugation aus den Blockdiagonalmatrizen $\exp(J_{m,\lambda})$ Es ist also $J_{m,\lambda} = \lambda E_m + N$ mit einer nilpotenten Matrix N. In diesem Spezialfall vertauschen die Summanden, also

$$\exp(J_{m,\lambda}) = \exp(\lambda) E_n \exp(N)$$

wobei $\exp(N)$ nur ein Polynom in N ist.

Beispiel. Wir betrachten

$$y' = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} y$$

Die Matrix A ist eine Jordanmatrix. Die Diagonalmatrix ist E_2, die nilpotente ist $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Das Lösungs-Fundamentalsystem ist also

$$\exp(tA) = \exp(t)(E_2 + \frac{1}{1!}Nt) = \begin{pmatrix} \exp(t) & t\exp(t) \\ 0 & \exp(t) \end{pmatrix}.$$

Beispiel. Wir betrachten die Drehmatrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Sie hat die komplexen Eigenwerte $\pm i$ mit den Eigenvektoren $v_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$. Die zugehörigen Lösungen sind

$$\phi_1(t) = \exp(tA)v_1 = \begin{pmatrix} \exp(it) \\ i\exp(it) \end{pmatrix},$$
$$\phi_2(t) = \exp(tA)v_2 = \begin{pmatrix} \exp(-it) \\ -i\exp(-it) \end{pmatrix}.$$

Sie bilden ein Lösungsfundamentalsystem, da v_1, v_2 als Eigenvektoren zu unterschiedlichen Eigenwerte linear unabhängig sind. Die Lösung zum Anfangswert $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}(v_1 + v_2)$ ist dann

$$\psi_1(t) = \frac{1}{2}(\phi_1 + \phi_2) = \frac{1}{2}\begin{pmatrix} e^{it} + e^{-it} \\ i(e^{it} - e^{-it}) \end{pmatrix} = \begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix}$$

Genauso finden wir als Lösung zum Anfangswert $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$\begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}$$

und das Lösungsfundamentalsystem

$$\begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

Wir sind dieser Differentialgleichung inzwischen mehrfach begegnet. Sie gehört zur Gleichung 2. Ordnung

$$y'' + y = 0.$$

Bemerkung. Das Verfahren löst dann auch alle linearen Differentialgleichungen

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots a_1 y = 0$$

mit konstanten Koeffizienten. Wir haben bereits gesehen, dass man dann (mit Variation der Konstanten) auch die inhomogenen Gleichungen lösen kann.

Kapitel 16

Wegintegrale

Wir haben in Kapitel 11 in Zusammenhang mit Richtungsableitungen den Begriff eines Weges kennengelernt. Ein Weg in $\mathbb{R}^n$ ist eine stetige Abbildung

$$\gamma : I \to \mathbb{R}^n.$$

Ist $f : \mathbb{R}^n \to \mathbb{R}^m$ eine weitere stetige Abbildung, so können wir die Komposition $f \circ \gamma$ integrieren. Wir haben dies bereits oft ausgenutzt, z. B. beim Integrieren in Koordinatenrichtungen in Zusammenhang mit parameterabhängigen Integralen und dem Satz von Fubini.

Wir werden uns mit weiteren interessanten Beispielen beschäftigen.

Bogenlänge

Es geht darum, die Länge eines Weges zu definieren. Hierfür wählen wir eine Norm auf $\mathbb{R}^n$, meist $\|\cdot\|_2$. Für (stückweise) lineare Wege wissen wir dann, was die Länge ist. Einen beliebigen Weg approximieren wir durch solche und versuchen die Länge als Grenzwert zu definieren.

Sei also $\gamma : [a, b] \to \mathbb{R}^n$ ein Weg. Sei $a = t_0 < t_1 < \cdots < t_n = b$ eine Unterteilung. Die *Feinheit* dieser Unterteilung ist das Maximum der $t_{i+1} - t_i$. Wir approximieren γ durch den Polygonzug, der auf $[t_i, t_{i+1}]$ linear zwischen $\gamma(t_i)$ und $\gamma(t_{i+1})$ interpoliert:

$$P{:}[a, b] \to \mathbb{R}^n, \qquad t \mapsto \gamma(t_i) + \frac{\gamma(t_{i+1}) - \gamma(t_i)}{t_{i+1} - t_i}(t - t_i) \qquad t \in [t_i, t_{i+1}]$$

Er hat die Länge

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

Definition 16.1. *Ein Weg heißt* rektifizierbar *mit Länge L, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so das für jede Unterteilung $a = t_0 < t_1 < \cdots < t_n = b$*

der Feinheitheit kleiner δ gilt

$$|\sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\| - L| < \varepsilon.$$

Satz 16.2. *Sei $\gamma : [a,b] \to \mathbb{R}^n$ ein stetig differenzierbarer Weg. Dann ist γ rektifizierbar mit*

$$L = \int_a^b \|\gamma'(t)\| \mathrm{d}t.$$

Beweis: Als stetige Funktion ist $\|\gamma'(t)\|$ Riemann-integrierbar. Sei L das Integral. Wir approximieren das Integral durch Integrale über Treppenfunktionen der Form

$$t \mapsto \|\gamma'(t_i)\| \qquad t \in [t_i, t_{i+1})$$

mit Integralwert

$$\sum_{i=0}^{n-1} (t_{i+1} - t_i) \|\gamma'(t_i)\|.$$

Wie im Beweis der Riemann-Integrierbarkeit von stetigen Funktionen hängt die Genauigkeit der Approximation nur von Feinheit der Unterteilung ab, d.h. zu gegebenen ε gibt es ein $\delta > 0$, so dass

$$|L - \sum_{i=0}^{n-1} (t_{i+1} - t_i) \|\gamma'(t_i)\|| < \varepsilon$$

für alle Unterteilungen der Feinheit $< \delta$. Es genügt nun zu zeigen:

Behauptung. *Zu $\varepsilon > 0$ gibt es $\delta > 0$, so dass*

$$\|\frac{\gamma(t) - \gamma(\tau)}{t - \tau} - \gamma'(t)\| < \varepsilon$$

für alle t, τ mit $|t - \tau| < \delta$.

Da alle unsere Normen auf $\mathbb{R}^n$ äquivalent sind, genügt es, die Aussage in der Supremumsnorm zu beweisen. Dann wiederum folgt die Ausage sofort aus der komponentenweisen Abschätzung, also aus dem Fall $n = 1$. Die Ableitung ist gleichmäßig stetig, zu dem ε gibt es also $\delta > 0$, so dass

$$|\gamma'(t) - \gamma'(s)| < \varepsilon$$

für $|t - s| < \delta$. Sei nun t, τ mit $|t - \tau| < \delta$. Wir wenden den Mittelwertsatz der Differentialrechnung aus Analysis 1 an und erhalten

$$\frac{\gamma(t) - \gamma(\tau)}{t - \tau} = \gamma'(s)$$

für s zwischen t und τ. Dann ist $|t-s|<\delta$ und daher

$$\left|\frac{\gamma(t)-\gamma(\tau)}{t-\tau}-\gamma'(t)\right| = |\gamma'(s)-\gamma'(t)| < \varepsilon.$$

Dies beweist die Behauptung. Setzen wir die Behauptung in die vorherige Gleichung ein, so erhalten wir die gewünschte Approximation von L. $\square$

Beispiel. Wir betrachten den Kreisweg $t \mapsto (\cos(t), \sin(t))$ auf $[0, 2\pi]$. Wir erhalten

$$L = \int_0^{2\pi} \|(-\sin(t), \cos(t))\| \mathrm{d}t = \int_0^{2\pi} \sqrt{\sin^2 t + \cos^2 t} \mathrm{d}t = \int_0^{2\pi} \mathrm{d}t = 2\pi.$$

Die Bogenlänge hängt nicht von der konkreten Wahl der Parametrisierung des Weges ab.

Lemma 16.3. *Sei $\phi : [\alpha, \beta] \to [a, b]$ bijektive stetig differenzierbare Abbildung. Sei $\gamma : [a, b] \to \mathbb{R}^n$ ein stetig differenzierbarer Weg. Dann gilt*

$$L(\gamma) = L(\gamma \circ \phi).$$

Beweis: Es gilt nach der Kettenregel

$$(\gamma \circ \phi)' = (\gamma' \circ \phi)\phi'.$$

Da ϕ bijektiv und differenzierbar ist, bleibt das Vorzeichen von ϕ immer das gleiche. Sei zunächst $\phi'(t) \geq 0$, $\phi(\alpha) = a$, $\phi(\beta) = b$. Wir berechnen nach Definition und mit der Substitutionsregel

$$\begin{aligned} L(\gamma \circ \phi) &= \int_\alpha^\beta \|(\gamma \circ \phi)'\|(t) \mathrm{d}t \\ &= \int_\alpha^\beta \phi'(t) \|(\gamma' \circ \phi)(t)\| \mathrm{d}t \\ &= \int_a^b \|\gamma'(s)\| \mathrm{d}s \\ &= L(\gamma). \end{aligned}$$

Im Fall $\phi'(t) \leq 0$ erhalten wir ein Vorzeichen unter dem Integral, da $|\phi(t)| = -\phi(t)$. Wegen $\phi(\alpha) = b$, $\phi(\beta) = a$ erhalten wir ein zweites Mal ein Vorzeichen, insgesamt ebenfalls Gleichheit. $\square$

Integration von Vektorfeldern

Wir betrachten $U \subset \mathbb{R}^n$ und eine stetige Abbildung

$$f : U \to \mathbb{R}^n.$$

Jedem Punkt wird ein Vektor zugordnet. Wir nennen dies ein *Vektorfeld*. Sei nun $\gamma : [a, b] \to U$ ein stetig differenzierbarer Weg.

Definition 16.4. *Mit U, f, γ wie oben definieren wir das* Wegintegral

$$\int_\gamma F \cdot \mathrm{d}x = \int_a^b \langle F(\gamma(t)), \gamma'(t)\rangle \mathrm{d}t = \int_a^b \sum_{i=1}^n F_i(\gamma(t))\gamma_i'(t)\mathrm{d}t.$$

Durch Zusammensetzen setzt sich die Definition auch auf stetige, stückweise differenzierbare Wege fort.

Diese Definition ist unabhängig von der Wahl der Koordinaten auf $[a,b]$ (d.h. der Parametrisierung des Weges) und auf U.

Lemma 16.5. *Sei $\phi : [\alpha,\beta] \to [a,b]$ stetig differenzierbar mit $\phi(\alpha) = a$ und $\phi(\beta) = b$. Dann gilt*

$$\int_\gamma F \cdot \mathrm{d}x = \int_{\gamma\circ\phi} F \mathrm{d}x.$$

Beweis: Es gilt

$$F_i(\gamma(\phi(s))(\gamma\circ\phi)_i'(s) = F_i(\gamma(\phi(s))\gamma_i'(\phi(s))\phi'(s).$$

Dies setzen wir in die Definition ein. Die Formel folgt mit der Substitutionsregel für $t = \phi(s)$. □

Bemerkung. Bei der Berechnung der Länge eines Weges ist nicht erlaubt, zwischendruch rückwärts zu laufen - dies verlängert den Weg. Beim Wegintegral für Vektorfelder heben sich solche Beiträge weg.

Beispiel. Sei $F : U \to \mathbb{R}$ eine stetig differenzierbare Funktion, $f = \mathrm{grad} F$ das Gradientenvektorfeld. Dann gilt mit Kettenregel

$$\int_\gamma f \cdot \mathrm{d}x = \int_a^b \sum_{i=1}^n \frac{\partial F}{\partial x_i}(\gamma(t))\gamma_i'(t)\mathrm{d}t = \int_a^b (F(\gamma(t))'\mathrm{d}t = F(\gamma(b)) - F(\gamma(a)).$$

Dieses Beispiel motiviert die Definition des Wegintegrals. In dieser Situation heißt F auch *Stammfunktion* oder *Potential* von f.

Beispiel (Winkelvektorfeld)**.** Sei

$$W : \mathbb{R}^2 \smallsetminus \{0\} \to \mathbb{R}^2, \qquad (x,y) \mapsto \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\right).$$

Jedem Punkt wird ein Vektor zugeordnet, der senkrecht auf (x,y) steht, also ein Tangentialvektor an den Kreis um 0 durch (x,y). Auf dem Kreis mit Radius r ist die Länge $1/r$.

Sei $\gamma : [a,b] \to \mathbb{R}^2 \smallsetminus \}0\}$ ein Weg. Wir schreiben ihn in der Form $\gamma(t) = r(t)(\cos\theta(t), \sin\theta(t))$. Es folgt

$$\begin{aligned}\int_\gamma W \cdot \mathrm{d}x &= \int_a^b \langle \frac{1}{r}\begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix}, r'\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} + r\theta'\begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix}\rangle \mathrm{d}t \\ &= \int_a^b (0 + \theta'(t))\mathrm{d}t = \theta(b) - \theta(a)\end{aligned}$$

Das Integral misst den Winkel, den der Weg überstreicht.

Wir wollen nun die Frage klären, welche Vektorfelder Gradientenvektorfelder sind. Wir kennen bereits die notwendige Bedingung

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$$

(falls f stetig partiell diffbar). Eine andere notwendige Bedingung erhalten wir aus der Formel für Wegintegrale über Gradientevektorfelder: Der Wert hängt nur von Anfangs- und Endpunkt ab, nicht vom gewählten Weg!

Beispiel. Sei W das Winkelvektorfeld aus dem letzten Beispiel. Es ist

$$\frac{\partial \frac{-y}{x^2+y2}}{\partial y} = \frac{-(x^2+y^2)+2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$
$$\frac{\partial \frac{x}{x^2+y2}}{\partial x} = \frac{(x^2+y^2)-2x^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

Trotzdem handelt es sich nicht um ein Gradientenvektorfeld, denn der Kreisweg $\gamma : t \mapsto (\cos(t), \sin(t))$ auf $[0, 2\pi]$ liefert das Integral

$$\int_\gamma W \cdot \mathrm{d}x = 2\pi$$

während der konstante Weg von 1 nach 1 den Weg 0 liefert.
Das Problem entsteht durch den fehlenden Nullpunkt im Definitionsbereich von W.

Nun systematisch.

Definition 16.6. *Sei $U \subset \mathbb{R}^n$ offen. Wir sagen, U ist* wegzusammenhängend, *wenn es zu je zwei Punkten x_1, x_2 in U einen Weg γ von x_1 nach x_2 gibt.*

Beispiel. Bälle oder allgemeiner konvexe Mengen sind wegzusammenhängend. Auch $\mathbb{R}^n \smallsetminus \{0\}$ ist wegzusammenhängend, nicht jedoch $\mathbb{R}^n \smallsetminus \{x \in \mathbb{R}^n | \|x\| = 1\}$.

Lemma 16.7. *Sei $U \subset \mathbb{R}^n$ offen und wegzusammenhängend. Sei $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld. Dann ist die Stammfunktion von f eindeutig bestimmt (falls sie existiert) bis auf eine additive Konstante.*

Beweis: Sei $x_0 \in U$. Seien F_1 und F_2 Stammfunktionen mit $F_1(x_0) = F_2(x_0)$. Sei $x \in U$. Wir wählen eine Weg γ von x_0 nach x und erhalten

$$F_1(x) - F_1(x_0) = \int_\gamma f \cdot \mathrm{d}x = F_2(x) - F_2(x_0) \ .$$

Also ist $F_1 = F_2$. □

Satz 16.8. *Sei $U \subset \mathbb{R}^n$ offen und wegzusammenhängend, $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld. Dann sind äquivalent:*

(i) f ist ein Gradientenvektorfeld.

(ii) Für jeden geschlossenen Weg (d.h. Anfangs- und Endpunkt sind gleich) gilt

$$\int_\gamma f \cdot \mathrm{d}x = 0.$$

(iii) Für je zwei Wege γ_1, γ_2 von x_1 nach x_2 gilt

$$\int_{\gamma_1} f \cdot \mathrm{d}x = \int_{\gamma_2} f \cdot \mathrm{d}x.$$

Beweis: Wir haben bereits bemerkt, dass die dritte Aussage aus der ersten folgt. Die zweite ist ein Spezialfall ($\gamma_1 = \gamma$, γ_2 konstant). Umgekehrt folgt die dritte Aussage aus der zweiten, indem wir die beiden Wege hintereinanderhängen. In Formeln: $\gamma_1, \gamma_2 : [a,b] \to U$ mit gleichem Anfangs- und Endpunkt. Wir definieren $\gamma : [a, 2b-a] \to U$ als γ_1 auf $[a,b]$ und $\gamma_2(2b-t)$ auf $[b, 2b-a]$. Dies ist ein geschlossener Weg. Es ist

$$0 = \int_a f \cdot \mathrm{d}x = \int_{\gamma_1} f \cdot \mathrm{d}x - \int_{\gamma_2} f \cdot \mathrm{d}x.$$

Seien also die Integrale wegunabhängig. Sei also jetzt jedes Integral über einen geschlossenen Weg 0. Sei $x \in U$. Wir wählen einen Weg γ_x von x_0 nach x und definieren

$$F(x) = \int_{\gamma_x} f \cdot \mathrm{d}x.$$

Nach Voraussetzung ist dies wohldefiniert. Wir berechnen die partielle Ableitung in Richtung x_j in x, also

$$\lim_{h\to 0} \frac{1}{h} \left(\int_{\gamma_x} f \cdot \mathrm{d}x - \int_{\gamma_{x+h}} f \cdot \mathrm{d}x \right) = \lim_{h \to 0} \int_{\tilde{\gamma}_h} f \cdot \mathrm{d}x$$

wobei $\tilde{\gamma}_h$ ein Weg von x nach $x + he_j$ ist. Für kleine h können wir $\tilde{\gamma}_h$ als den linearen Weg $t \mapsto x + te_j$ auf $[0,h]$ wählen. Es ist dann $\gamma'(t) = e_j$.
Wir erhalten dann

$$\int_{\tilde{\gamma}_h} f \cdot \mathrm{d}x = \int_0^h f_j(x + te_j)\mathrm{d}t$$

Wie im Beweis Hauptsatz der Differential- und Integralrechnung erhalten wir den gewünschten Grenzwert

$$\lim_{h \to 0} \frac{1}{h} \int_0^h f_j(x + te_j)\mathrm{d}t = f_j(x).$$

□

Theorem 16.9. *Sei $U \subset \mathbb{R}^n$ offen und konvex. Sei $f : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld mit $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ für alle $i, j = 1, \dots, n$. Dann ist f ein Gradientenvektorfeld.*

Beispiel. Das Winkelvektorfeld W auf $\mathbb{R}^2 \setminus \{0\}$ ist auf jeder Kreisscheibe, die 0 nicht enthält, ein Gradientenvektorfeld. Auf jeder dieser Kreisscheiben lassen sich eindeutige Polarkoordinaten (r, θ) wählen. Es gilt dann $W = \mathrm{grad}\theta$. Die Einheitskreislinie lässt sich durch solche Kreisscheiben überdecken. Auf jeder der kleineren Kreisscheiben erhalten wir ein Potential. Auf den Schnitten stimmen je zwei Wahlen nur bis auf additive Konstante überein. Es ist *nicht* möglich, diese Konstanten global verträglich zu wählen. Die spiegel wider, dass arcsin und arccos nicht global stetig definierbar sind.

Beweis: Sei $\gamma : [0,1] \to U$ ein geschlossener Weg. Für $s \in [0,]$ definieren wir den Weg

$$\gamma_s : [0,1] \to U, \qquad \gamma_s(t) = s\gamma(t) + (1-s)\gamma(0).$$

Wir haben dann $\gamma_1 = \gamma$ und $\gamma_0 = \gamma(0)$. Alle diese Wege verlaufen in U, da U konvex ist. Außerdem sind sie alle geschlossen. Wir zeigen nun, dass die Wegeintegrale über alle γ_s gleich sind. Für $s = 0$ erhalten wir den Wert 0 wie gewünscht.
Die Abbildung $H : [0,1] \times [0,1] \to U$ mit $H(s,t) = \gamma_s(t)$ ist stetig partiell differenzierbar nach s und t. Tatsächlich gilt sogar

$$\frac{\partial^2 H}{\partial t \partial s} = \frac{\partial^2 H}{\partial s \partial t} = \frac{\partial \gamma}{\partial t}.$$

Wir differenzieren die Funktion

$$s \mapsto \int_{\gamma_s} f \cdot \mathrm{d}x$$

nach s und erhalten mit Differentiation unter dem Integral und partieller Integration im zweiten Summanden

$$\begin{aligned}
\frac{\partial}{\partial s} \int_{\gamma_s} f \cdot \mathrm{d}x = \frac{\partial}{\partial s} \int_0^1 \langle f(H(s,t)), \frac{\partial H}{\partial x}(s,t) \rangle \mathrm{d}t \\
&= \int_0^1 \langle (Df \circ H)(s,t) \frac{\partial H}{\partial s}(s,t), \frac{\partial H}{\partial t} \rangle + \int_0^1 \langle f \circ H(s,t), \frac{\partial^2 H}{\partial s \partial t} \rangle \mathrm{d}t \\
&= \int_0^1 \langle Df \circ H \frac{\partial H}{\partial s}, \frac{\partial H}{\partial t} \rangle \mathrm{d}t + \langle f \circ H, \frac{\partial H}{\partial s} \rangle \Big|_{t=0}^{t=1} - \int_0^1 \langle Df \circ H \frac{\partial H}{\partial t}, \frac{\partial H}{\partial s} \rangle \mathrm{d}t
\end{aligned}$$

Nach Voraussetzung ist DF symmetrisch, daher heben sich der erste und letzte Summand weg. Da alle γ_s geschlossen sind, verschwindet der mittlere Summand.
$\square$

Bemerkung. Eine Funktion H wie in dem Beweis heißt *Homotopie* von γ_0 nach γ_1. Zwei Wege von x_0 nach x_0, die durch eine Homotopie verbunden werden, heißen *homotop*. Der selbe Beweis zeigt, dass homotope Wege dasselbe Wegintegral haben, wenn das Vektorfeld die Voraussetzung an die partiellen Ableitungen erfüllt.

Sind in U je zwei geschlossene Wege homotop, so existiert eine Stammfunktion. Solche U heißen *einfach zusammenhängend*. Das Studium von topologischen Räumen bis auf Homotopie ist ein klassischer Gegendstand der *(algebraische) Topologie*. Die Homotopieklassen von Wegen von x_0 nach x_0 bilden eine Gruppe bezüglich des Aneinanderhängens. Dies ist die *Fundamentalgruppe*. Ein topologischer Raum ist einfach zusammenhängend, wenn seine Fundamentalgruppe verschwindet.

Wir sehen in dem Theorem oder seiner Verallgemeinerung, wie analytische und topologische Eigenschaften zusammenspielen.

Theorem 16.10 (Fundamentalsatz der Algebra)**.** *Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen, d.h. jedes nicht-konstante komplexe Polynom hat eine Nullstelle.*

Beweis: Wir identifizieren $\mathbb{R}^2$ und $\mathbb{C}$. Wir betrachten

$$p(z) = z^n + a_{n-1}z^{n-1} + \dots a_0$$

und schreiben es als $z^n + q(z)$ mit $q(z)$ vom Grad höchstens $n-1$. Es gilt dann $z^{-n}q(z) \to 0$ für $|z| \to \infty$. Für R genügend groß ist also

$$|q(Re^{i\theta}| \leq \frac{1}{2}R^n$$

für alle $\theta \in [0, 2\pi]$.
Wir betrachten die geschlossene Kurve

$$\gamma_0 : [0, 2\pi] \to \mathbb{R}^2, \qquad \gamma_0(\theta) = p(Re^{i\theta})$$

und die Homotopie

$$H : [0, 1] \times [0, 2\pi] \to \mathbb{R}^2 \smallsetminus \{0\}, \qquad H(s, \theta) = R^n e^{in\theta} + (1-s)q(Re^{i\theta}).$$

Für die Wohldefiniertheit nutzen wir

$$|H(s, \theta)| \geq |R^n e^{in\theta}| - (1-s)|q(Re^{i\theta}| \geq R^n - \frac{1}{2}R^n > 0.$$

Sie ist stetig differenzierbar. Für $s = 0$ erhalten wir γ_0. Für $s = 1$ erhalten wir $\gamma_1 : \theta \mapsto R^n e^{in\theta} = (R^n \cos(n\theta), R^n \sin(n\theta))$. Die beiden Wege sind also homotop. Wir betrachten das Winkelvektorfeld W aus dem vorherigen Beispiel. Es erfüllt die Voraussetzung an die partiellen Ableitungen, also gilt

$$\int_{\gamma_0} W \cdot \mathrm{d}x = \int_{\gamma_1} W \cdot \mathrm{d}x = 2\pi n$$

nach dem vorherigen Beispiel.

Nun geben wir eine zweite stetig differenzierbare Homotopie an:

$$K : [0, R] \times [0, 2\pi] \to \mathbb{R}^2 \setminus \{0\}, \qquad K(\varrho, \theta) = p(\varrho e^{i\theta})$$

Falls p keine Nullstellen hat, so ist sie wohldefiniert. Sie verbindet γ_0 und den konstanten Weg $\theta \mapsto p(0)$. Also folgt

$$\int_{\gamma_0} W \cdot \mathrm{d}x = \int_{K(0,\cdot)} W \cdot \mathrm{d}x = 0.$$

Damit ist $2\pi n = 0$. Dies ist ein Widerspruch. $\square$

Differentialformen

Der Nachteil unserer Zugangs ist, dass die Eigenschaft, Gradientenvektorfeld zu sein, koordinatenabhängig ist. Anders gesagt: die Werte der Ableitung einer Funktion sind koordinatenabhängig. Hier kommen Differentialformen ins Spiel.

Wir fixieren $U \subset \mathbb{R}^n$ offen mit den Koordinaten $x_1, \ldots, x_n$. Wir arbeiten mit der $\mathbb{R}$-Algebra (Ring und $\mathbb{R}$-Vektorraum) $C^\infty(U)$ der unendlich oft differenzierbaren Funktionen auf $\mathbb{R}$.

Definition 16.11. *Der Raum $A^1(U)$ der* glatten Differentialformen *besteht aus den Ausdrücken der Form*

$$\sum_{i=1}^n f_i dx_i$$

mit $f_i \in C^\infty(U)$. Wir definieren

$$d : C^\infty(U) \to A^1(U), \qquad f \mapsto \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i.$$

Bemerkung. In diesem Zugang ist dx_i einfach ein Symbol, der Name eines Basisvektors von $A^1(U)$. Unter der Abbildung d wird die Koordinatenfunktion x_i abgebildet auf $\sum_{j=1}^n \frac{\partial x_i}{\partial x_i} dx_j = dx_i$.

Es gibt eine alternative Definition von $A^1(U)$ mit einer universellen Eigenschaft bezüglich d. Dann ist es ein Satz, dass jede Differentialform die angegebene Form hat. Dies führt hier zu weit.

Bemerkung. Wir überprüfen die Wirkung von Koordinatenwechseln. Sei $y : U \to V \subset \mathbb{R}^n$ bijektiv und unendlich oft differenzierbar. Wir fassen $y_1, \ldots, y_n$ als Koordinaten auf U auf Sei $f : V \to \mathbb{R}$ glatt. Dann gilt

$$d(f \circ y) = \sum_{i=1}^n \frac{\partial f \circ y}{\partial x_i} dx_i = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial y_j} \frac{\partial y_j}{\partial x_i} dx_i = \sum_{j=1}^n \frac{\partial f}{\partial y_j} dy_j = df.$$

Lemma 16.12. *Die Abbildung d ist eine Derivation, d.h. $\mathbb{R}$-linear und*

$$d(fg) = fdg + gdf.$$

Beweis:

$$d(fg) = \sum_{i=1}^{n} \frac{\partial fg}{\partial x_i} dx_i = \sum_{i=1}^{n} f \frac{\partial g}{\partial x_i} dx_i + \sum_{i=1}^{n} g \frac{\partial f}{\partial x_i} dx_i = fdg + gdf.$$

Die $\mathbb{R}$-Linearität ist trivial. □

Lemma 16.13. *Es gilt $df = 0$ genau dann, wenn f konstant ist.*

Beweis: $df = 0$ bedeutet, dass das Gradientenvektorfeld zu f verschwindet. Dies ist für die konstanten Funktionen erfüllt. Wegen der Eindeutigkeit der Stammfunktion gilt nur für die konstanten Funktionen. □

Als nächstes wollen wir die Bedinung an die zweiten Ableitungen ebenfalls in dieser Sprache formulieren. Dafür benutzen wir Differentialformen von höherem Grad.

Definition 16.14. *Sei $p \geq 1$. Wir definieren*

$$A^p(U) = \bigwedge^{*} A^1(U)$$

als die Menge der Ausdrücke

$$\sum_{i_1 < i_2 < \cdots < i_p} f_{i_1 \ldots i_p} dx_{i_1} \wedge \cdots \wedge dx_{i_p}$$

mit $f_{i_1 \ldots i_p} \in C^\infty(U)$.

Bemerkung. In algebraischer Sprache: $A^1(U)$ ist ein freier $C^\infty(U)$-Modul und $A^p(U)$ seine p-te *äußere Potenz*. Laut Definition verschwindet $A^p(U)$ für $p > n$.

Wir führe eine Multiplikation ein

$$A^p(U) \times A^q(U) \to A^{p+q}(U)$$

die eindeutig festgelegt ist durch die Rechenregel

$$dx_i \wedge dx_j = -dx_j \wedge dx_i$$

für alle i, j. Insbesondere gilt $dx_i \wedge dx_i = -dx_i \wedge dx_i = 0$.
Auch die Differentiation setzt sich eindeutig fort zu einer $\mathbb{R}$-linearen Abbildung

$$d = d^p : A^p(U) \to A^{p+1}(U)$$

definiert durch

$$d(f dx_{i_1} \wedge \dots dx_{i_p} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \dots dx_{i_p}$$

$i_1 < \cdots < i_p$ und $f \in C^\infty(U)$.
Die beiden wesentlichen Eigenschaften sind:

Lemma 16.15. *(i) Es gilt* $d^{p+1} \circ d^p = 0$.

(ii) Für $\omega_p \in A^p(U)$ *und* $\omega_q \in A^q(U)$ *gilt*

$$d(\omega_p \wedge \omega_q) = d\omega_p \wedge \omega_q + (-1)^p \omega_p \wedge d\omega_q.$$

Beweis: Wir kürzen ab $\omega = dx_{i_1} \wedge \cdots \wedge dx_{i_p}$.

$$\begin{aligned} d(d(f\omega) &= d(\sum_{i=1} \frac{\partial f}{\partial x_i} dx_i \wedge \omega \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial x_j \partial x_i} dx_j \wedge dx_i \wedge \omega \end{aligned}$$

Wir spalten auf in die Summe mit $i < j$ und die Summe der $i > j$ (der Beitrag von $i = j$ ist 0 wegen $dx_i \wedge dx_i = 0$). In der zweiten Summe nutzen wir $dx_j \wedge dx_i = -dx_i \wedge dx_j$ aus. Der Summand hebt sich dann gegen den Beitrag mit vertauschtem i und j weg, da die gemischten partiellen Ableitunge der glatten Funktion f übereinstimmen.
Auch die Multiplikationsformel rechnet man einfach nach. Wir zeigen den wesentlichen Fall:

$$\begin{aligned} d(dx_1 \wedge f dx_2) = d(f x_1 \wedge dx_2) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i \wedge dx_1 \wedge dx_2 \\ &= \sum_{i=1}^{n} -\frac{\partial f}{\partial x_i} dx_1 \wedge dx_i \wedge dx_2 \\ &= ddx_1 \wedge f dx_2 - dx_1 \wedge (d(f dx_2)) \end{aligned}$$

□

Wir betrachten nun insbesondere das Differential d^1.

$$d^1(\sum_{i=1}^{n} f_i dx_i) = \sum_{i,j} \frac{\partial f_i}{\partial x_j} dx_j \wedge dx_i$$

Dieser Ausdruck verschwindet genau dann, wenn

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}.$$

Dies ist genau die Bedingung, die wir als notwendig für die Exitenz einer Stammfunktion erkannt hatten!

Definition 16.16. *Eine Differentialform $\omega \in A^q(U)$ heißt* geschlossen*, falls $d\omega = 0$. Sie heißt* exakt*, falls $\omega = d\omega'$.*

Wegen $dd = 0$ sind exakte Differentialformen geschlossen! Theorem 16.9 besagte für $q = 1$, dass jede geschlossene Differentialform exakt ist, falls U konvex (oder allgemeiner einfach zusammenhängend).

Theorem 16.17 (Poincaré Lemma)**.** *Sei U konvex, $\omega \in A^q(U)$ geschlossen. Dann ist ω exakt für $q \geq 1$ und konstant für $q = 0$.*

Beweis des Poincaré Lemmas: Der Beweis ist hauptsächlich geschicke Buchhaltung plus der Hauptsatz der Differential- und Integeralrechnung. Den Fall $p = 1$ haben wir bereits gezeigt. Sei also jetzt $p > 1$. Wir behandeln zunächst die Fälle $n = 1, 2$.

Im Fall $n = 1$ ist $U = I$ ein offenes Intervall. Jede 1-Form fdx ist geschlossen und hat eine Stammfunktion.

Sei nun $n = 2$. Wir arbeiten in den Koordinaten x, y. Für $p = 1$ haben wir

$$\omega = fdx + gdy, \qquad \frac{\partial f}{\partial y} = \frac{\partial g}{\partial x}.$$

Sei $F = \int f \mathrm{d}x$. Dann ist

$$dF = \partial F \partial x dx + \frac{\partial F}{\partial y} dy = fdx + hdy.$$

Wir ersetzen ω durch $\omega - dF$. Dies bleibt geschlossen, da ω und dF geschlossen sind. Ist $dG = \omega - dF$, so ist $G - F$ das gesuchte Urbild. Nach dieser Ersetzung ist

$$\omega = gdy, \qquad 0 = \frac{\partial g}{\partial x}.$$

D.h. die Funktion g hängt nicht von x ab. Aus dem Fall $n = 1$ wissen wir bereits, dass es ein Urbild gibt.

Für $p = 2$ haben wir

$$\omega = adx \wedge dy.$$

Sei $A = (\int a\mathrm{d}x)dy$. Dann ist

$$dA = \left[\left(\frac{\partial}{\partial x} \int a\mathrm{d}x\right) dx + \left(\frac{\partial}{\partial y} \int a\mathrm{d}x\right) dy\right] \wedge dy = adx \wedge dy.$$

Wir haben ein Urbild gefunden.

Sei nun n beliebig, $p > 1$. Wir haben

$$\omega = \sum_{i_1 < \ldots i_p} f_{i_1 \ldots i_p} dx_{i_1} \wedge \ldots dx_{i_p}$$

mit $d\omega = 0$. Wir berachten nur die Summanden mit $i_1 = 1$ und bilden

$$H(\omega) = \sum_{i_2 < \ldots i_p} \int f_{i_1 \ldots i_p} \mathrm{d}x_1 dx_{i_2} \wedge \ldots dx_{i_p}.$$

Dann hat $\omega' = \omega - dH$ die Eigenschaft, dass der Koffizient von $dx_1 \wedge dx_{i_2} \wedge \ldots dx_{i_p}$ verschwindet. Die Bedingung $d\omega' = 0$ impliziert, dass alle übrigen Koeffizienten von x_1 unabhängig sind. Mit vollständiger Induktion erhalten wir ein Urbild von ω'. □

Definition 16.18. *Sei $U \subset \mathbb{R}^n$ offen. Wir definieren die q-te* de Rham Kohomologie

$$H^q(U) = \frac{\mathrm{Kern} d^q}{\mathrm{Bild} d^{q-1}}$$

als Quotient der geschlossenen Formen modulo der exakten.

Korollar 16.19. *Sei U konvex. Dann gilt*

$$H^q(U) = \begin{cases} \mathbb{R} & q = 0 \\ 0 & q > 0 \end{cases}$$

Beispiel. Sei $U = \mathbb{R}^2 \setminus \{0\}$. Dann ist $H^1(U) \neq 0$, da W eine geschlossene Differentialform definiert, die nicht exakt ist. Tatsächlich ist $H^1(U) = \mathbb{R}$ erzeugt von der Klasse von W.

Bemerkung. Wir haben die Definition hier nur für offene Teilmengen des $\mathbb{R}^n$ angegeben. Da der Begriff der Differentialform unabhängig von Koordinatenwahlen ist, lässt sich alles auf Mannigfaltikeiten ausdehnen. Geometrisch sind Differentialformen Schnitte des *Kotantentialbündels*. Physiker sprechen von $(0, 1)$-Tensoren.

Kapitel 17

Ausblick

Nach der Analysis 2 verzeigt sich die Theorie in verschiedene Richtungen.

mehrdimensionale Integration

Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$. Wir wollen f integrieren. Es gibt einen recht naiven Zugang, bei der man eine Variable nach der nach der anderen integriert. Bei guten Funktionen führt dies auf vernünftige Ergebnisse. Garant ist der Satz von Fubini, den wir behandelt haben

In der Vorlesung *Mehrfachintegrale* wird diese Standpunkt systematisch studiert.

In *Analysis* wird ein abstrakterer Zugang verfolgt. Im Zentrum steht der Begriff des *Maßes* (=Volumen) einer Menge. Es ist nicht möglich jeder Teilmenge des $\mathbb{R}^n$ ein Maß zuzordnen, so dass vernünftige Rechenregeln gelten. Daher muss zuerst ein geeigneter Definitionsbereich gefunden werden.

Der abstrakte Zugang belohnt mit sehr starken Konvergenzsätzen. In der fortgeschrittenen *Wahrscheinlichkeitstheorie* wird immer mit Maßen gearbeitet. Auch in der Funktionalanalysis wird auf dieser Sprache aufgebaut.

Funktiontheorie

Dies ist die Differential- und Integralrechnung über $\mathbb{C}$. Die wichtigsten Sätze lassen sich schnell aus Analysis 2 und 3 herleiten. In der Regel wird jedoch ein anderer Standpunkt eingenommen und alles von Grund auf neu entwickelt. Die Sätze sind viel einfacher als im Reellen, da jede komplex differenzierbare Funktion aus Potenzreihe geschrieben werden kann und unendlich oft differenzierbar ist. Viele Integrale lassen sich über den Umweg ins Komplexe leicht berechnen. Unmengen von Formeln wie

$$\frac{2}{\pi} = \prod_{i=1}^{\infty} \cos(\pi/2^{i+1})$$

stammen aus der Funktionentheorie.

Funktionalanalysis

Hier geht es um unendlich dimensionale normierte Vektorräume, typischerweise um Räume von Funktionen. Differentiation definiert lineare Abbildungen. Die Sätze und die Sprache der Funktionalanalysis wird benötigt in der fortgeschrittenen Theorie von Differentialgleichungen.

Theorie und Numerik von gewöhnlichen und partiellen Differentialgleichungen

Selbsterklärend

Topologie

Dies greift den Begriff des topologischen Raums wieder auf. Zentraler Begriff ist Stetigkeit. Zur Klassifikation von topologischen Räumen verwendet man meist algebraische Invarianten wie die Fundamentalgruppe (Homotopieklassen von geschlossenen Wegen) oder Kohomologie. Besonders interessant ist das Zusammenspiel von analytischen und topologischen Eigenschaften von Mannigfaltigkeiten. An Vorkenntnissen genügt Linere Algebra und etwas Analysis.

Differentialgeometrie

Hier geht es um Mannigfaltigkeiten mit einem Längen- und Winkelbegriff, also einem Skalarprodukt auf dem Tangentialraum. Zentraler Begriff ist die Krümmung. Die Ebene ist flach, Kugeln sind positiv gekrümmt. Ein Beispiel für eine Fläche mit negativer Krümmung ist die hyperbolische Ebene. Als Menge nimmt man $\mathbb{R} \times \mathbb{R}_{>0}$ und als Geraden Halbkreise oder Geraden, die senkrecht auf der x-Achse stehen. Für elementare Differentialgeomtrie benötigt man neben LA und Analysis 1-2 auch etwas Integrationstheorie.

Inhaltsverzeichnis