A N A L Y S I S   II

2024

Patrick Dondl

Nach einem Skript von Prof. E. Kuwert

Abteilung für Angewandte Mathematik
Universität Freiburg

Inhaltsverzeichnis

Kapitel 1

Grundlegende Strukturen II

1.1 Topologische Räume

Definition 1.1.1 (Topologie)**.** Sei X eine Menge und sei $\mathcal{T}$ eine Menge von Teilmengen von X (Notation: $\mathcal{T} \subset 2^X$) mit den Eigenschaften

1. $\emptyset, X \in \mathcal{T}$
2. $\mathcal{T}' \subset \mathcal{T} \Rightarrow \bigcup_{U \in \mathcal{T}'} U \in \mathcal{T}$
3. $U_1, U_2 \in \mathcal{T} \to U_1 \cap U_2 \in \mathcal{T}$.

Dann wird $(X, \mathcal{T})$ ein *topologischer Raum* genannt, $\mathcal{T}$ wird *Topologie* genannt und die Mengen $U \in \mathcal{T}$ werden *offen* genannt.
Wenn zusätzlich gilt, dass für alle $x_1,\ x_2 \in X,\ x_1 \neq x_2$ offene Mengen $U_1, U_2 \in \mathcal{T}$ mit der Eigenschaft existieren, dass

$$x_1 \in U_1, \ \ x_2 \in U_2, \quad U_1 \cap U_2 = \emptyset, \tag{Trennungsaxiom}$$

dann wird $(X, \mathcal{T})$ *Hausdorff-Raum* genannt.

Definition 1.1.2. Eine Teilmenge $A \subset X$ eines topologischen Raumes heißt *abgeschlossen* wenn

$$U \in \mathcal{T} \quad \text{existiert, so dass} \quad A = X \setminus U = U^c.$$

Sei X fortan ein topologischer Raum.

Definition 1.1.3. Sei $A \subset X$.

1. $A^\circ := \{x \in X \mid \exists U \subset A,\ U \in \mathcal{T} \text{ so dass } x \in U\}$ ist das *(offene) Innere* von A.
2. $\overline{A} := \{x \in X \mid \forall U \in \mathcal{T}, x \in U \text{ haben wir } U \cap A \neq \emptyset\}$ ist der *Abschluss* von A.
3. $\partial A := \overline{A} \backslash A^\circ$ ist der *Rand* von A.

Definition 1.1.4. Eine Menge $A \subset X$ heißt *dicht* in X, wenn $\overline{A} = X$.

Proposition 1.1.5. *1.* $A^\circ \subset A \subset \overline{A}$.

2. $A = A^\circ \Leftrightarrow A$ *offen,* $\quad A = \overline{A} \Leftrightarrow A$ *abgeschlossen.*

3. *A° ist offen, $\overline{A}$ ist abgeschlossen.*

4. *$X \backslash \overline{A} = (X \backslash A)^\circ$.*

Beweis. Übung. □

Proposition 1.1.6. *Sei $A \subset X, (X, \mathcal{T})$ ein topologischer Raum. Dann ist $(A, \mathcal{T}_A)$ mit*

$$\mathcal{T}_A := \{U \cap A \mid U \in \mathcal{T}\}$$

ein topologischer Raum. $\mathcal{T}_A$ wird Unterraumtopologie *oder* Relativtopologie *genannt.*

Beweis. Das sollte ziemlich klar sein. □

Definition 1.1.7. Sei $(X, \mathcal{T})$ ein topologischer Raum.

1. $B \subset \mathcal{T}$ heißt *Basis* von $\mathcal{T}$, wenn jede Menge in $\mathcal{T}$ als eine Vereinigung von Mengen in B geschrieben werden kann.

2. $S \subset \mathcal{T}$ heißt eine *Subbasis* von $\mathcal{T}$, wenn die Menge aller endlichen Schnittmengen von Mengen in S eine Basis von $\mathcal{T}$ ist.

Proposition 1.1.8. *Sei X eine Menge, $S \subset 2^X$ eine Kollektion von Teilmengen von X. Sei nun $B \subset 2^X$ die Menge der Teilmengen von X, die durch beliebige endliche Schnitte von Mengen in S erzeugt werden. Dann ist die Menge, die durch beliebige Vereinigungen von Mengen in B erzeugt wird, zusammen mit der leeren Menge und X selbst, eine Topologie.*

Beweis. Es sind die Eigenschaften einer Topologie zu überprüfen. Seien also U_j, $j \in J$ mit beliebiger, nicht notwendigerweise abzählbarer Indexmenge J, Vereinigungen von Mengen aus B. Dann ist aber auch $\bigcup_{j \in J} U_j$ eine solche Vereinigung von Mengen aus B. Analog dazu seien U_1, U_2 Vereinigungen von Schnitten von Mengen aus S, also $U_1 = \bigcup_{j_1 \in J^1} \bigcap_{k=1}^{N^1_{j_1}} S^1_k$, $U_2 = \bigcup_{j_2 \in J^2} \bigcap_{k=1}^{N^2_{j_2}} S^2_k$ mit $S^{1,2}_k \in S$, $J_{1,2}$ Indexmengen, $N^{1,2}_{j_{1,2}} \in \mathbb{N}$. Dann gilt $U_1 \cap U_2 = \bigcup_{j_1 \in J^1} \bigcup_{j_2 \in J^2} \left(\left(\bigcap_{k=1}^{N^1_{j_1}} S^1_k \right) \cap \left(\bigcap_{k=1}^{N^2_{j_2}} S^2_k \right) \right)$. □

Definition 1.1.9. Seien $\mathcal{T}_1, \mathcal{T}_2$ Topologien auf X. $\mathcal{T}_2$ heißt *stärker* (oder *feiner*) als $\mathcal{T}_1$ und $\mathcal{T}_1$ *schwächer* (oder *gröber*) als $\mathcal{T}_2$, wenn

$$\mathcal{T}_1 \subset \mathcal{T}_2.$$

Beispiel. 1. Die indiskrete Topologie: $\mathcal{T} = \{\emptyset, X\}$,

2. Die diskrete Topologie: $\mathcal{T} = 2^X$,

3. Die kofinite Topologie: $X = \mathbb{N}, \mathcal{T} := \left\{U \in 2^{\mathbb{N}} \mid \#(X \backslash U) < \infty\right\} \cup \emptyset$,

4. Topologien, die durch eine Metrik induziert werden.

Definition 1.1.10 (Umgebung)**.** Sei A eine Teilmenge von X. Eine Menge N heißt eine Umgebung von A, wenn es eine offene Menge U gibt, so dass $A \subset U \subset N$. Man beachte, dass A eine einelementige Menge sein kann. Wir schreiben $N(A)$ oder N_a.

Definition 1.1.11 (Stetigkeit)**.** Seien $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ topologische Räume.

1. Eine Abbildung $f : X \to Y$ heißt *stetig*, wenn
$$\text{für alle} V \in \mathcal{T}_Y \text{ gilt } f^{-1}(V) \in \mathcal{T}_X.$$

2. Eine Abbildung $f : X \to Y$ heißt *stetig in* $x \in X$, wenn für jede Umgebung N_Y von $f(x)$ eine Umgebung N_X von x existiert, so dass $f(N_X) \subset N_Y$ ist.

Proposition 1.1.12. *$f : X \to Y$ ist stetig genau dann wenn $f : X \to Y$ stetig ist in allen $x \in X$.*

Beweis. **"⇒"** Ziemlich einfach, nehmen Sie einfach das Urbild der offenen Menge in N_Y als N_X.

"⇐" Betrachten wir die offene Teilmenge V_Y. Wenn deren Urbild leer ist, sind wir fertig. Andernfalls nehmen wir $x \in f^{-1}(V) =: A$. Nach der Definition der Stetigkeit in einem Punkt gibt es eine offene Menge U, die x enthält, so dass $U \subset A$ ist. Also ist $A = A^\circ$ und damit offen.

□

Definition 1.1.13 (Konvergenz)**.** Sei $(X, \mathcal{T})$ ein topologischer Raum, sei $x \in X$. Wir sagen, dass eine Folge $(x_k)_{k \in \mathbb{N}}$ *konvergiert gegen* x, wenn

für eine beliebige Umgebung N_x ein $K \in \mathbb{N}$ existiert, so dass für alle $k \geq K$ gilt $x_k \in N_x$.

Wir schreiben
$$x_k \xrightarrow{\mathcal{T}} x.$$

Proposition 1.1.14. *Wenn $(X, \mathcal{T})$ Hausdorffsch ist, dann ist der Grenzwert x eindeutig bestimmt. Wir schreiben in diesem Fall*
$$x = \lim_{k \to \infty} x_k \text{ bezüglich } \mathcal{T}.$$

Beweis. Angenommen $x_k \to x_1, x_k \to x_2 \neq x_1$. Dann
$$\exists U_1 \ni x_1, U_2 \ni x_2 \text{ mit } U_1 \cap U_2 = \emptyset.$$
Aber: $\exists K : \forall k \geq K$ gilt $x_k \in U_1$ und gleichzeitig $x_k \in U_2$. Dies ist ein Widerspruch. □

1.2 Metrik

Definition 1.2.1. Sei X eine Menge und $d : X\mathbf{x}X \to \mathbb{R}$, so, dass für alle $x, y, z \in X$ haben wir

1. $d(x, y) \geq 0$ und $d(x, y) = 0 \Leftrightarrow x = y$.

2. $d(x, y) = d(y, x)$.

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

Dann nennen wir (X, d) einen *metrischen Raum*, d eine *Metrik*, eine *Distanz* oder einen *Abstand*.

Bemerkung. Ohne die Bedingung, dass $d(x, y) = 0 \Leftrightarrow x = y$, nennen wir d eine Pseudo-Metrik. Indem man ein Modulo (also Äquivalenzklassen)

$$x \hat{=} y \quad \Leftrightarrow \quad d(x, y) = 0$$

bildet, kann man einen pseudometrischen Raum in einen metrischen Raum verwandeln.

Beispiel.

1. $X = \mathbb{R},\ d(x, y) = |x - y|$.
2. $X = \mathbb{R},\ d(x, y) = \frac{|x-y|}{1+|x-y|}$.
3. (X, d) metrischer Raum, $h : Y \to X$ injektiv,

   $$d_Y(y_1, y_2) := d(h(y_1), h(y_2))$$

   wird als die *Pullback-Metrik* bezeichnet.
4. Diskrete Metrik:

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

Definition 1.2.2. Sei (X, d) ein metrischer Raum,

$$B_r(x) := \{y \in X \mid d(x, y) < r\} \qquad \text{für } r > 0, x \in X.$$

Wir nennen $U \subset X$ *offen in Bezug auf die Metrik* d, wenn für alle $x \in U$ ein $r > 0$ existiert, so dass

$$B_r(x) \subset U.$$

Die leere Menge ist offen.

Proposition 1.2.3. *Sei (X, d) ein metrischer Raum, und sei*

$$\mathcal{T} := \{U \subset X \mid U \text{ offen bezüglich } d\}.$$

Dann ist $(X, \mathcal{T})$ ein Hausdorff-Raum. Wir nennen $\mathcal{T}$ die durch d induzierte Topologie.

Beweis.

1. $\emptyset, X \in \mathcal{T}$: klar.
2. Sei $\mathcal{T}' \subset \mathcal{T}$ und $W := \bigcup_{U \in \mathcal{T}'} U$. Wir müssen zeigen, dass es für alle $x \in W$ $r > 0$ mit $B_r(x) \subset W$ gibt. Dies ist jedoch offensichtlich, da es ein $U \in \mathcal{T}'$ mit $x \in U$ gibt und wir einfach das r für dieses U nutzen können.
3. Sei $U_1, U_2 \in \mathcal{T}, x \in U_1 \cap U_2$. Es gibt $r_1, r_2 > 0$ so dass

   $$B_{r_1}(x) \subset U_1 \quad \text{und} \quad B_{r_2}(x) \subset U_2.$$

   Aus der Dreiecksungleichung folgt, dass

   $$B_{\min(r_1, r_2)}(x) \subset U_1 \cap U_2.$$

4. Sei $x \neq y$. Es gilt $d(x,y) = c > 0$ und somit

$$B_r(x) \cap B_r(y) = \emptyset, \text{ für } r = \frac{c}{2} > 0. \qquad \square$$

Definition 1.2.4. Sei d_1, d_2 jeweils eine Metrik auf X.

- d_1 heißt *stärker* als d_2 bzw. d_2 *schwächer* als d_1, wenn dies auch für die induzierten Topologien gilt.

- d_1 wird *äquivalent* zu d_2 genannt, wenn die induzierten Topologien gleich sind.

Proposition 1.2.5 (Stetigkeit in metrischen Räumen)**.** *Seien* $(X, d_X), (Y, d_Y)$ *metrische Räume und sei* $f : X \to Y$*. Dann ist* f *stetig in* $x \in X$*, genau dann wenn*

$$\forall \epsilon > 0 : \exists \delta > 0 : \quad d_X(x,y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon.$$

Proposition 1.2.6 (Konvergenz in metrischen Räumen)**.** *Seien* $(X, d_X), (Y, d_Y)$ *metrische Räume.*

1. *Sei* $(x_j)_{j\in\mathbb{N}}$ *eine Folge in* X*, dann gilt*

$$x_k \xrightarrow{d} x \quad \Leftrightarrow \quad \forall \epsilon \, \exists K : d(x_k, x) \leq \epsilon \quad (\textit{ for } k \geq K\ ).$$

2. *Eine Abbildung* $f : X \to Y$ *ist stetig in* x*, wenn für alle Folgen* $(x_k)_{k\in\mathbb{N}}$ *mit* $x_k \to x$ *gilt*

$$f(x_k) \to f(x).$$

Bemerkung. Die zweite Bedingung wird als *Folgenstetigkeit* bezeichnet.
Warnung. In rein topologischen Räumen ist dies im Allgemeinen nicht äquivalent zur Stetigkeit.

Bemerkung. In Analogie zu Analysis I lassen sich nun auch Häufungspunkte von Mengen (Analysis I, Definition 4.2.3) sowie Grenzwerte von Funktionen (Analysis I, Definition 4.2.5) definieren.

Definition 1.2.7. Sei (X, d) ein metrischer Raum.

1. $(x_k)_{k\in\mathbb{N}}$ heißt eine *Cauchy-Folge*, wenn $\forall \varepsilon > 0 \; \exists K \in \mathbb{N} : \forall k, l > K$, gilt $d(x_k, x_l) < \varepsilon$.

2. Der Raum (X, d) heißt *vollständig*, wenn jede Cauchy-Folge einen Grenzwert in X besitzt.

Proposition 1.2.8 (Folgenkriterium für Abgeschlossenheit)**.** *Sei* $A \subset (X, d)$ *(metrischer Raum).* A *ist abgeschlossen, wenn für alle Folgen* $(x_j)_{j\in\mathbb{N}}$ *mit* $x_j \in A$ *und* $x_j \to x \in X$*, gilt* $x \in A$*.*

Beweis. $\Rightarrow$ Angenommen, es gäbe einen Grenzpunkt x in A^c, und A^c offen. Es existiert also eine Umgebung $N(x) \subset A^c$. Dann müssten Punkte der Folge in dieser Umgebung, also außerhalb von A liegen: ein Widerspruch.

$\Leftarrow$ Nehmen wir an, A sei nicht abgeschlossen und betrachten $x \in \overline{A} \setminus A$ (was denn nicht leer ist). Nach der Definition des Abschlusses gilt für jedes $r > 0$, dass $B_r(x) \cap A \neq \emptyset$. Man nehme nun eine Folge von Radien $r_j \to 0$ und wähle als x_k einen beliebigen Punkt in der Schnittmenge $B_{r_j}(x) \cap A$.

□

Proposition 1.2.9. *Sei (X, d) ein vollständiger metrischer Raum, und A sei eine abgeschlossene Teilmenge von X. Dann ist (A, d) ein vollständiger metrischer Raum.*

Beweis. Es ist klar, dass (A, d) ein metrischer Raum ist. Zur Vollständigkeit betrachten wir eine Cauchy-Folge in A. Durch die Vollständigkeit von (X, d) besitzt diese einen Grenzwert in X, dieser Grenzwert muss nach dem obigen Folgenkriterium auch in A liegen. □

1.3 Vektorräume und Normen

Wir betrachten nur Vektorräume über den Körpern $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und betrachten diese Körper als metrische (oder topologische) Räume mit dem üblichen Abstand.

Definition 1.3.1. Sei X ein Vektorraum über $\mathbb{K}$ und sei X zugleich ein topologischer Raum. Wenn Vektoraddition und Skalarmultiplikation stetig sind, dann heißt X *topologischer Vektorraum.*

Bemerkung. Man beachte, dass wir hier gegebenenfalls die Boxtopologie verwenden müssen. Die Basis der Boxtopologie auf dem Produktraum $X \times Y$ wird gebildet durch Mengen der Form

$$\{U_1 \times U_2 \mid U_1 \text{ offen in } X,\ U_2 \text{ offen in } Y\}.$$

Definition 1.3.2 (Norm)**.**

- Sei X ein Vektorraum über $\mathbb{K}$. Eine Abbildung $\|\cdot\| : X \to \mathbb{R}$ heißt *Norm*, wenn gilt

  1. $\|x\| \geq 0 \quad \forall x \in X.$
  2. $\|x\| = 0 \quad \Rightarrow \quad x = 0.$
  3. $\|\alpha x\| = |\alpha|\, \|x\|\,.$
  4. $\|x + y\|\, \|x\| + \|y\|\,.$

- Wenn $\|\cdot\|$ eine Norm auf X ist, dann wird $(X, \|\cdot\|)$ *normierter Raum* genannt.

- Eine Abbildung mit den Eigenschaften 1., 3., 4. wird *Seminorm* (oder *Halbnorm*) genannt.

Bemerkung. Wiederum können wir einen Raum mit einer Halbnorm in einen normierten Raum verwandeln, indem wir das Modulo nehmen.

Proposition 1.3.3. *Sei $(X, \|\cdot\|)$ ein normierter Raum. Indem wir*

$$d : X \times X \to \mathbb{R}, \quad d(x, y) := \|x - y\|\,,$$

setzen, erhalten wir einen metrischen Raum (X, d).

Beweis. Folgt direkt aus den Normeigenschaften. □

Definition 1.3.4. Ein vollständiger normierter Raum wird *Banachraum* genannt.

Proposition 1.3.5. *Ein normierter Vektorraum ist ein topologischer Vektorraum und die Topologie ist Hausdorffsch.*

Beweis. Übung. ☐

Proposition 1.3.6. *Sei $\|\cdot\|_1, \|\cdot\|_2$ sind zwei Normen auf X, d_1, d_2 die induzierten Metriken und $\mathcal{T}_1, \mathcal{T}_2$ die induzierten Topologien. Es gilt*

1. *d_2 ist stärker als d_1, genau dann wenn $\exists C > 0$:*
$$\|x\|_1 \leq C \|x\|_2 \quad \forall x \in X.$$

2. *Zwei Normen sind äquivalent (also $\Leftrightarrow \mathcal{T}_1 = \mathcal{T}_2$), genau dann wenn $\exists C, c > 0$:*
$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 .$$

Beweis. Punkt 2. folgt direkt aus 1., wir zeigen also nur 1.

$\Leftarrow$ Sei U offen bezüglich $\mathcal{T}_1$, und sei $x \in U$. Dann gibt es eine d_1-r-Kugel um x, die in U liegt. Die d_2-$\frac{r}{C}$-Kugel liegt dann ebenfalls in U. Somit ist U auch bzgl. d_2 offen.

$\Rightarrow$ Angenommen, die Ungleichung sei für kein $C > 0$ erfüllt. Sei dann $U = B_1^{d_2}(0)$ die bezüglich d_2 offene d_2-Einheitskugel um den Ursprung. Sei nun $\delta > 0$. Da die Ungleichung nicht erfüllt ist, existiert $y \in X$ mit $||y||_1 > \frac{1}{\delta}||y||_2$, und somit auch $x \in X$ mit $||x||_2 < 1$ aber $||x||_1 > \delta$. Also gibt es kein $\delta > 0$, so dass $B_\delta^{d_1}(0) \subset B_1^{d_2}(0)$ und $B_1^{d_2}(0)$ ist nicht offen bezüglich d_1.

☐

Satz 1.3.7. *Alle Normen auf einem endlichdimensionalen Vektorraum sind äquivalent.*

Beweis. Sei X ein n-dimensionaler Vektorraum mit Norm $|| \cdot ||$. Wir wählen eine beliebige Basis $\{e_1, \dots e_n\}$ von X. Wir zeigen, dass die Norm $|| \cdot ||$ äquivalent ist zur Norm[1]

$$|| \cdot ||_1 \colon X \to \mathbb{R}, \quad ||v||_1 = \sum_{k=1}^{n} |x_k| \text{ für } v = \sum_{k=1}^{n} x_k e_k.$$

Nachdem Äquivalenz von Normen eine transitive Eigenschaft ist (also aus $|| \cdot ||_a$ äquivalent zu $|| \cdot ||_b$ und $|| \cdot ||_b$ äquivalent zu $|| \cdot ||_c$ folgt $|| \cdot ||_a$ äquivalent zu $|| \cdot ||_c$) reicht es zu zeigen, dass $|| \cdot ||_1$ und $|| \cdot ||$ äquivalent sind.
Sei also zunächst $v \in X$. Wir rechnen (mit Dreiecksungleichung und Homogenität)

$$||v|| \leq \sum_{k=1}^{n} |x_k| \, ||e_k|| \leq C \sum_{k=1}^{n} |x_k| = C||v||_1 \quad \text{für } C = \max_{k=1\dots n} ||e_k||, \tag{1.1}$$

was eine der Abschätzungen zur Äquivalenz zeigt. Für die andere Abschätzung argumentieren wir durch Widerspruch.

[1] die Normeigenschaften dieser Abbildung lassen sich leicht überprüfen.

Wir nehmen also an, es existiert eine Folge $(w_j)_{j=1}^{\infty}$ in X, so dass $||w_j||_1 = 1$ aber $||w_j|| \to 0$. Da $||w_j||_1 = 1$ existiert eine Teilfolge $i(j) = i_j$ und ein $k_0 \in \{1, \ldots, n\}$, so dass

$$|x_{k_0}^{i_j}| \geq \frac{1}{n} \quad \text{für alle } j \in \mathbb{N}, \tag{1.2}$$

wobei wieder $w_j = \sum_{k=1}^{n} x_k^j e_k$.
Andererseits gilt aber auch $|x_k^j| \leq 1$ für alle $j \in \mathbb{N}$, $k \in \{1, \ldots, n\}$. Wir können also mit Hilfe des Satzes von Bolzano und Weierstraß Teilfolgen (von Teilfolgen ... insgesamt n-mal) von der Teilfolge $i(j)$ wählen (und die letztendliche Teilfolge wieder $i(j)$ nennen), so dass $x_k^{i_j} \to x_k$ in $\mathbb{K}$ für alle $k \in \{1, \ldots, n\}$. Wir setzen $v = \sum_{k=1}^{n} x_k e_k$ und es folgt $||w_{i_j} - v||_1 \to 0$. Mit (1.2) folgt $v \neq 0$.
Mit (1.1) folgt aber auch $||w_{i_j} - v|| \to 0$. Da aber $||w_j|| \to 0$ erhalten wir einen Widerspruch zu $v \neq 0$. ☐

Satz 1.3.8. *Endlichdimensionale Unterektorräume von normierten Räumen sind vollständig und abgeschlossen.*

Beweis. Vollständigkeit folgt aus Vollständigkeit von $\mathbb{K}^n$ und der Äquivalenz aller Normen aus Satz 1.3.7, Abgeschlossenheit aus dem Folgenkriterium in Proposition 1.2.8. ☐

Proposition 1.3.9. *Seien $(V, ||\,.\,||_V)$ und $(W, ||\,.\,||_W)$ normierte Vektorräume. Die Abbildungen $(x, y) \mapsto \left(||x||_V^p + ||y||_W^p\right)^{1/p}$ für $1 \leq p < \infty$ bzw. $(x, y) \mapsto \max\{||x||_V, ||y||_W\}$ stellen Normen auf $V \times W$ dar. Alle diese Normen erzeugen die Boxtopologie.*

Beweis. Übung. ☐

Bemerkung. Auf endlichen Produkten normierter Räume nehmen wir als Topologie stets die Boxtopologie an. Wir bemerken, dass eine Folge $(x_j, y_j)_{j \in \mathbb{N}}$ in $V \times W$ genau dann konvergiert, wenn x_j in V und y_j in W konvergiert.

1.4 Skalarprodukt

Definition 1.4.1. Sei X ein $\mathbb{K}$-Vektorraum. Eine Abbildung

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

heißt *Sesquilinearform*, wenn $\forall x, x_1, x_2, y, y_1, y_2 \in X, \alpha \in \mathbb{R}$ gilt

1. $(\alpha x, y) = \alpha\,(x, y) = (x, \bar{\alpha} y)$
2. $\begin{cases} (x_1 + x_2, y) = (x_1, y) + (x_2, y) \\ (x, y_1 + y_2) = (x, y_1) + (x, y_2) \end{cases}$

Eine Sesquilinearform ist *hermitsch* (oder, im reellen Fall, *symmetrisch*), wenn

$$(x, y) = \overline{(y, x)}.$$

Eine hermitsche Sesquilinearform heißt *positiv semidefinit*, wenn

$$(x, x) \geq 0 \quad \forall x \in X.$$

Eine positiv semidefinite Sesquilinearform heißt *positiv definit*, falls

$$(x, x) = 0 \quad \Rightarrow \quad x = 0.$$

Bemerkung. Wenn $\mathbb{K} = \mathbb{R}$, verwenden wir den Ausdruck *Bilinearform*.

Definition 1.4.2. Eine positiv definite (hermitsche) Sesquilinearform heißt *Skalarprodukt*. Das Paar $(X, (\cdot, \cdot))$ wird *Prä-Hilbertraum* genannt.

Lemma 1.4.3. *Sei $(\cdot, \cdot)$ ein Skalarprodukt und setze*

$$\|x\| := \sqrt{(x,x)} \quad \forall x \in X.$$

Dann ist $\|x\| := \sqrt{(x,x)}$ ist eine Norm auf X.
Weiters gilt

1. $|(x,y)| \leq \|x\| \cdot \|y\|$ *(Cauchy-Schwarz-Ungleichung)*

2. $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ *(Parallelogrammidentität)*

Beweis. Lineare Algebra bzw. Übung. □

Bemerkung. Zentrale neue Eigenschaft in Räumen mit Skalarprodukt: Orthogonalität.

Definition 1.4.4. 1. Zwei Vektoren x, y heißen *orthogonal*, wenn gilt $(x,y) = 0$.

2. Zwei Untervektorräume $U, V \subset X$ heißen *orthogonal*, wenn $(x,y) = 0 \quad \forall x \in U, y \in V$.

Definition 1.4.5. Ein vollständiger Prä-Hilbertraum wird *Hilbertraum* genannt.

Beispiel. $X = C^1([0,1])$ mit $(f,g) := \int_0^1 fg + \int_0^1 f'g'$ ist ein Prä-Hilbertraum.

1.5 Kompaktheit

Satz 1.5.1 (Kompaktheit in metrischen Räumen)**.** *Sei A eine Teilmenge eines metrischen Raumes (X, d). Dann sind die folgenden Aussagen äquivalent:*

1. *A ist (überdeckungs-)kompakt, d.h. jede offene Überdeckung von A beinhaltet eine endliche Teilüberdeckung.*

2. *A ist folgenkompakt, d.h. jede Folge in A besitzt eine konvergente Teilfolge.*

3. *$(A, d|_{A \times A})$ ist vollständig und präkompakt, d.h. für alle $\varepsilon > 0$ gibt es eine endliche Überdeckung von A mit ε-Kugeln.*

Bemerkung. Die Aussage in Punkt 1. ist die Definition von Kompaktheit in rein topologischen Räumen.

Beweis.

1. $\Rightarrow$ 2. Sei $(x_k)_{k \in \mathbb{N}}$ eine Folge in A ohne Häufungspunkt (als Annahme zum Widerspruch). Dann gilt $\forall y \in A : \exists r_y > 0 :$

$$N_y := \left\{ k \in \mathbb{N} \mid x_k \in B_{r_y}(y) \right\} \text{ ist endlich.}$$

Die Kugeln $B_{r_y}(y)$ bilden aber eine offene Überdeckung von A. Nach Annahme der Existenz einer endlichen Teilüberdeckung $\exists \{y_k\}_{k=1}^N$:

$$\bigcup_{k=1}^{N} B_{r_{y_k}}(y_k) \supset A.$$

Dies ist ein Widerspruch zu unserer Annahme, denn sonst könnten nur endlich viele Elemente in der Folge sein.

2. $\Rightarrow$ 3. Zunächst zeigen wir die Vollständigkeit: Wegen 2. gilt, dass jede Cauchy-Folge in A einen Häufungspunkt enthält. In jedem Fall besitzt eine eine Cauchy-Folge aber nur *höchstens* einen Häufungspunkt. Nach 2. dieser Häufungspunkt *in* A. Daher konvergiert die Cauchyfolge gegen ein Element von A.

   Für die Kompaktheit argumentieren wir wieder durch Widerspruch: Angenommen $\exists \varepsilon > 0$: existiert keine endliche Überdeckung mit ε-Kugeln. Dann existiert eine Folge $(x_k)_{k\in\mathbb{N}} \subset A$:

$$x_{k+1} \in A \setminus \bigcup_{j=1}^{k} B_\varepsilon(x_j).$$

   Diese Folge besitzt aber keinen Häufungspunkt.

3. $\Rightarrow$ 1. Sei $(U_i)_{i\in I}$ eine offene Überdeckung von A. Wir setzen

$$\mathcal{B} := \left\{ B \subset A \mid J \subset I, B \subset \bigcup_{i\in J} U_i \Rightarrow |J| = \infty \right\},$$

   also $\mathcal{B}$ sei die Menge aller Teilmengen B von A, so dass jede Teilüberdeckung von B durch Mengen aus $(U_i)_{i\in I}$ notwendigerweise unendlich sein muss.

   Zu zeigen ist nun $A \notin \mathcal{B}$. A ist präkompakt, also $\forall B \in \mathcal{B}, \varepsilon > 0$ gibt es eine Überdeckung

$$B \subset \bigcup_{i=1}^{n_\varrho} B_\varepsilon(x_i).$$

   Es folgt aber für mindestens ein i, dass gilt

$$B_\varepsilon(x_i) \cap B \in \mathcal{B},$$

   denn wenn keine dieser endlich vielen Schittmengen in $\mathcal{B}$ wäre, dann hätte B ja bereits eine endliche Teilüberdeckung.

   Nehmen wir nun zum Widerspruch and, dass $A \in \mathcal{B}$. Wir setzen induktiv $\varepsilon = \frac{1}{k}, k \in \mathbb{N}$ und erhalten die Existenz von $x_k \in X$ und Mengen:

$$B_1 := A, \quad B_k := B_{\frac{1}{k}}(x_k) \cap B_{k-1} \in \mathcal{B} \;\; \forall k \geq 2.$$

   Nehmen wir $y_k \in B_k, k \in \mathbb{N}$. Für $k \leq l$ gilt

$$y_k, y_l \in B_{\frac{1}{k}}(x_k) \Rightarrow d(y_k, y_l) \leq \frac{2}{k}.$$

Es gilt also, dass $(y_k)_{k\in\mathbb{N}}$ eine Cauchyfolge ist. Da A vollständig ist, $\exists y \in A$:

$$d(y_k, y) \xrightarrow[k\to\infty]{} 0.$$

Aber wir haben $y \in U_{i_0}$ für ein i_0 und damit folgt für k hinreichend groß

$$B_k \subset B_{\frac{1}{k}}(x_k) \subset B_{\frac{2}{k}}(y_k) \subset B_{\frac{2}{k}+d(y_k,y)}(y) \subset U_{i_0}.$$

Damit ist aber $B_k \notin \mathcal{B}$. Ein Widerspruch.

□

Satz 1.5.2 (Riesz)**.** *Sei X ein Banachraum. Dann ist $\overline{B_1(0)}$ kompakt genau dann wenn X endlichdimensional ist.*

Beweis. “$\Leftarrow$” Dank der Äquivalenz aller Normen genügt es die Aussage für $X = \mathbb{K}^n$ mit der $||\cdot||_1$-Norm zu zeigen. Dies folgt aber durch wiederholte Anwendung des Satzes von Bolzano-Weierstraß.

“$\Rightarrow$” Dank Präkompaktkeit von $\overline{B_1(0)}$, gibt es eine Überdeckung

$$\overline{B_1(0)} \subset \bigcup_{k=1}^{m} B_{\frac{1}{2}}(y_k).$$

Wir setzen

$$Y := \operatorname{span}\left(\{y_k\}_{k=1}^m\right).$$

Y ist endlichdimensional, also nach Satz 1.3.8 abgeschlossen in X. Wir nehmen an, dass gilt $Y \neq X$ und führen dies zu einem Widerspruch.

Behauptung: Wir haben $\forall \Theta \in (0,1) : \exists x_\theta \in X, \|x_\theta\| = 1$:

$$\operatorname{dist}(x_\theta, Y) \geq \theta.$$

Zum Beweis der Behauptung nehmen wir $x \in X \setminus Y$. Dann gilt

$$\operatorname{dist}(x, Y) = \inf_{y\in Y} ||x - y|| > 0$$

aufgrund der Offenheit von $X \setminus Y$. Weiters existiert $y_0 \in Y$:

$$0 < \|x - y_0\| \leq \frac{1}{\theta}\operatorname{dist}(x, Y).$$

Wir setzen

$$x_\theta := \frac{x - y_0}{\|x - y_0\|}.$$

Dann gilt für alle $y \in Y$:

$$\begin{aligned}
\|x_\theta - y\| &= \frac{1}{\|x - y_0\|}\,\|x - \overbrace{(y_0 + \|x - y_0\|\, y)}^{\in Y}\| \\
&\geq \frac{1}{\|x - y_0\|}\operatorname{dist}(x, Y) \\
&\geq \frac{\operatorname{dist}(x, Y)}{\frac{1}{\theta}\operatorname{dist}(x, Y)} \\
&= \theta.
\end{aligned}$$

Dies zeigt die Behauptung.

Nun muss aber für ein $j \in \{1, ..., n\}$ gelten, dass

$$x_\theta \in B_{\frac{1}{2}}(y_j).$$

Mit $\frac{1}{2} < \theta < 1$ ergibt sich aber ein Widerspruch zu

$$\operatorname{dist}(x_\theta, Y) \geq \theta. \qquad \square$$

Satz 1.5.3. *Seien $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ topologische Räume mit $(X, \mathcal{T}_X)$ kompakt im Sinne von Satz 1.5.1 Punkt 1. Sei $f : X \to Y$ stetig. Dann ist $f(X)$ kompakt.*

Beweis. Übung $\square$

Folgerung 1.5.4. *Abgeschlossene, beschränkte Teilmengen endlichdimensionaler normierter Vektorräume sind kompakt.*

Beweis. Sei A eine abgeschlossene, beschränkte Teilmenge des endlichdimensionalen, normierten Raumes X. Sei $R > 0$ hinreichend groß, so dass $A \subset B_R(0)$ Mir $\overline{B_1(0)}$ ist wegen Satz 1.5.3 auch $\overline{B_R(0)}$ kompakt und damit überdeckungskompakt. Somit ist auch A (überdeckungs)-kompakt. $\square$

Satz 1.5.5. *Stetige, reellwertige Funktionen auf Kompakta nehmen ihr Infimum und Supremum an.*

Beweis. Das Bild einer solchen Funktion ist eine kompakte – und damit wegen Vollständigkeit auch abgeschlossene – Teilmenge der reellen Zahlen. Diese enthält ihr Minimum und Maximum. $\square$

Kapitel 2

Differentialrechnung im $\mathbb{R}^n$

2.1 Die partiellen Ableitungen

Im Folgenden betrachten wir den $\mathbb{R}^n$ mit der Standardbasis $e_j, j = 1, \ldots, n$. Auf $\mathbb{R}^n$ nutzen wir das Standardskalarprodukt, also $(x, y) = \sum_{j=1}^n x_j y_j$ für $x = (x_1, \ldots, x_n)^T$, $y = (y_1, \ldots, y_n)^T$ und die zugehörige euklidische Norm $||x||_2 = \sum_{j=1}^n |x_j|^2$, welche wir auch als $|x|$ schreiben.

Bemerkung. Für Funktionen $f\colon I \to \mathbb{R}^m$, $f(x) = (f_1(x), \ldots, f_m(x))^T$ mit I offenes Intervall, $f_j\colon I \to \mathbb{R}$, $j = 1, \ldots, m$, setzen wir für $f'(x_0)$, $x_0 \in I$ die komponentenweise Ableitung $f'(x_0) = (f_1'(x_0), \ldots, f_m'(x_0))^T = \sum_{j=1}^m f_j'(x_0) e_j$, falls alle Komponenten in x_0 im Sinne von Analysis I differenzierbar sind. Sämtliche bekannten Eigenschaften der Ableitung übertragen sich.

Definition 2.1.1. Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^m$. Die partielle Ableitung von f nach x_j an der Stelle $x \in \Omega$ ist der Grenzwert (falls existent)

$$\partial_j f(x) = \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t} = \frac{d}{dt} f(x + te_j)\Big|_{t=0}.$$

Bemerkung.

1. Andere Bezeichnungen sind $\partial_{x_j} f(x)$, $\frac{\partial f}{\partial x_j}(x)$ oder $f_{x_j}(x)$.
2. Halten wir alle x_i mit $i \neq j$ fest, so ergibt sich eine Funktion in nur einer Variablen, also

$$\varphi(s) = f(x_1, \ldots, x_{j-1}, s, x_{j+1}, \ldots, x_n) \quad \text{für } s \in (x_j - \delta, x_j + \delta).$$

   Die Ableitung von φ im Punkt $s = x_j$ ist dann nichts anderes als die entsprechende partielle Ableitung, denn

$$\varphi'(x_j) = \lim_{t \to 0} \frac{\varphi(x_j + t) - \varphi(x_j)}{t} = \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t} = \partial_j f(x).$$

   Bei der Berechnung der partiellen Ableitung $\partial_j f$ können wir also die gewohnte eindimensionale Ableitung nach x_j bilden während wir die anderen Variablen konstant halten.

Die aus dem ersten Semester bekannten Differentiationsregeln lassen sich somit direkt übertragen.

Satz 2.1.2 (Ableitungsregeln)**.** *Sei* $\Omega \subset \mathbb{R}^n$ *offen und* $x \in \Omega$. *Die Existenz der partiellen Ableitungen* $\partial_j f(x)$ *und* $\partial_j g(x)$ *sei hier stets vorausgesetzt. Dann gilt*

1. *Linearität: für* $f, g : \Omega \to \mathbb{R}^m$ *und* $\alpha, \beta \in \mathbb{R}$ *gilt*

$$\partial_j(\alpha f + \beta g)(x) = \alpha \partial_j f(x) + \beta \partial_j g(x).$$

2. *Komponentenweise Differentiation: für* $f : \Omega \to \mathbb{R}^m$ *gilt, wenn eine der Seiten existiert,*

$$\partial_j f(x) = \sum_{i=1}^{m} \partial_j f_i(x) e_i.$$

3. *Produktregel: für* $f : \Omega \to \mathbb{R}^m$ *und* $g : \Omega \to \mathbb{R}$ *gilt*

$$\partial_j(fg)(x) = (\partial_j f)(x) g(x) + f(x) (\partial_j g)(x).$$

4. *Quotientenregel: für* $f : \Omega \to \mathbb{R}^m$ *und* $g : \Omega \to \mathbb{R}$ *mit* $g(x) \neq 0$ *gilt*

$$\partial_j \left(\frac{f}{g}\right)(x) = \frac{(\partial_j f)(x) g(x) - f(x) (\partial_j g)(x)}{g(x)^2}.$$

5. *Kettenregel: sei* f *reellwertig. Ist* I *offenes Intervall mit* $f(\Omega) \subset I$ *und* $\varphi : I \to \mathbb{R}$ *differenzierbar, so gilt*

$$\partial_j(\varphi \circ f)(x) = \varphi'(f(x)) \partial_j f(x).$$

Bemerkung. Höhere Ableitungen lassen sich nun ebenfalls bilden. Es sei die Funktion $f : \Omega \to \mathbb{R}^m$ auf ganz Ω partiell nach der Variablen x_j diffenenzierbar mit Ableitungsfunktion $\partial_j f$. Dann schreiben wir

$$\partial^2_{ij} f(x) := \partial_i (\partial_j f)(x) \quad \left(\text{alternative Notation } \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \text{ oder } f_{x_i x_j}(x)\right).$$

Entsprechend für Ableitungen k-ter Ordnung: seien $i_1, \dots, i_k \in \{1, \dots, n\}$. Ist die Funktion $\partial^{k-1}_{i_2 \dots i_k} f : \Omega \to \mathbb{R}^m$ bereits definiert so setzen wir, falls existent,

$$\partial^k_{i_1 \dots i_k} f(x) = \partial_{i_1} \left(\partial^{k-1}_{i_2 \dots i_k} f\right)(x).$$

Definition 2.1.3 (C^k-Räume)**.** Sei $\Omega \subset \mathbb{R}^n$ offen und $k \in \mathbb{N}_0 \cup \{\infty\}$. Wir bezeichnen mit $C^k(\Omega, \mathbb{R}^m)$ die Menge aller k-mal stetig differenzierbaren Funktionen auf Ω mit Werten im $\mathbb{R}^m$, das heißt alle partiellen Ableitungen $\partial^j_{i_1 \dots i_j} f$ der Ordnung $j \leq k$ (bzw. $j < \infty$ im Fall $k = \infty$) sind definiert und stetig auf Ω. Im reellwertigen Fall, also $m = 1$, schreiben wir $C^k(\Omega) = C^k(\Omega, \mathbb{R})$.

Satz 2.1.4 (Schwarz)**.** *Sei* $\Omega \subset \mathbb{R}^n$ *offen. Ist* $f \in C^2(\Omega)$, *so vertauschen für* $1 \leq i, j \leq n$ *die Ableitungen nach* x_i *und* x_j *:*

$$\partial_i \partial_j f = \partial_j \partial_i f \quad \text{auf } \Omega.$$

Folgerung 2.1.5. *Für eine Funktion* $f \in C^k(\Omega, \mathbb{R}^m)$ *vertauschen die partiellen Ableitungen bis zur Ordnung* k, *das heißt für jede Permutation* $\sigma \in S_k$ *gilt*

$$\partial_{i_{\sigma(1)}} \dots \partial_{i_{\sigma(k)}} f = \partial_{i_1} \dots \partial_{i_k} f$$

Definition 2.1.6 (Richtungsableitung)**.** Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^m$. Die Richtungsableitung von f an der Stelle $x \in \Omega$ in Richtung $v \in \mathbb{R}^n$ ist der Grenzwert (falls existent)

$$\partial_v f(x) = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t} = \frac{d}{dt} f(x+tv)\Big|_{t=0}.$$

Beispiel. Für die Funktion

$$f \colon \mathbb{R}^2 \to \mathbb{R} \quad f(x,y) = \begin{cases} \frac{2xy^2}{x^2+y^4} & \text{für } (x,y) \neq 0 \\ 0 & \text{sonst} \end{cases}$$

existieren im Ursprung alle Richtungsableitungen. Die Funktion ist allerdings nicht stetig.

2.2 Das totale Differential

Definition 2.2.1 (Ableitung)**.** Sei $\Omega \subset \mathbb{R}^n$ offen. $f : \Omega \to \mathbb{R}^m$ heißt differenzierbar in $x_0 \in \Omega$, falls $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ existiert, so dass

$$\lim_{x \to x_0} \frac{f(x) - (f(x_0) + A(x - x_0))}{|x - x_0|} = 0. \tag{2.1}$$

Mit der Substitution $h = x - x_0$ erhalten wir die äquivalente Formulierung

$$\lim_{h \to 0} \frac{f(x_0 + h) - (f(x_0) + Ah)}{|h|} = 0.$$

Die Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ heißt Ableitung oder totales Differential von f in x_0. Wir schreiben $Df(x_0) = A$.

Satz 2.2.2 (Berechnung und Eindeutigkeit der Ableitung)**.** *Die Funktion $f : \Omega \to \mathbb{R}^m$ sei in $x_0 \in \Omega$ differenzierbar. Dann hat f in x_0 die Richtungsableitungen*

$$\partial_v f(x_0) = Df(x_0)\, v \qquad \textit{für alle } v \in \mathbb{R}^n,$$

und $Df(x_0)$ hat bezüglich der Standardbasen die Matrixdarstellung (Jacobimatrix)

$$(\partial_j f_i(x_0)) = \begin{pmatrix} \partial_1 f_1(x_0) & \dots & \dots & \partial_n f_1(x_0) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \partial_1 f_m(x_0) & \dots & \dots & \partial_n f_m(x_0) \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

Insbesondere ist die Ableitung durch (2.1) *eindeutig bestimmt.*

Bemerkung. Typischerweise unterscheiden wir hier nicht zwischen einer linearen Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ und ihrer Matrixdarstellung in der Standardbasis.

Definition 2.2.3 (Gradient)**.** Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}$ differenzierbar in $x \in \Omega$. Der Gradient von f im Punkt x ist der Vektor

$$\operatorname{grad} f(x) = \sum_{j=1}^n \partial_j f(x) e_j = \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix} \in \mathbb{R}^n.$$

Satz 2.2.4 (Differenzierbare Funktionen sind stetig)**.** *Sei $f\colon \Omega \to \mathbb{R}^m$, $\Omega \subset \mathbb{R}^n$ offen, differenzierbar im Punkt $x_0 \in \Omega$. Dann ist f in x_0 stetig.*

Satz 2.2.5 (Kettenregel)**.** *Seien $f: U \to \mathbb{R}^m$ und $g: V \to \mathbb{R}^p$ mit $U \subset \mathbb{R}^n, V \subset \mathbb{R}^m$ offen und $f(U) \subset V$. Sind f in x_0 und g in $f(x_0)$ differenzierbar, so ist auch $g \circ f$ in x_0 differenzierbar und es gilt die Kettenregel*

$$D(g \circ f)(x_0) = Dg(f(x_0))\, Df(x_0).$$

Für die zugehörigen Jacobimatrizen bedeutet das mit $y_0 = f(x_0)$

$$\frac{\partial (g \circ f)_i}{\partial x_k}(x_0) = \sum_{j=1}^{m} \frac{\partial g_i}{\partial y_j}(y_0) \frac{\partial f_j}{\partial x_k}(x_0) \qquad \text{für } 1 \leq i \leq p, 1 \leq k \leq n.$$

Dies entspricht dem bekannten Matrixprodukt.

Satz 2.2.6 (komponentenweise Differentiation)**.** *$f: \Omega \to \mathbb{R}^m$ ist genau dann in $x_0 \in \Omega$ differenzierbar, wenn alle Komponenten $f_i: \Omega \to \mathbb{R}, i = 1, \ldots, m$, in x_0 differenzierbar sind. Ist $P_i: \mathbb{R}^m \to \mathbb{R}$ Projektion auf die i-te Koordinate, so gilt $Df_i(x_0) = P_i Df(x_0)$.*

Satz 2.2.7 (Ableitungsregeln)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen, und $f, g: \Omega \to \mathbb{R}$ seien differenzierbar im Punkt $x \in \Omega$. Dann gelten folgende Aussagen.*

1. *Linearität: für $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g$ in x differenzierbar mit Ableitung*

$$D(\alpha f + \beta g)(x) = \alpha Df(x) + \beta Dg(x).$$

2. *Produktregel: fg ist in x differenzierbar mit Ableitung*

$$D(fg)(x) = Df(x)g(x) + f(x)Dg(x).$$

3. *Quotientenregel: ist $g(x) \neq 0$, so ist f/g auf einer Umgebung von x definiert und*

$$D\left(\frac{f}{g}\right)(x) = \frac{Df(x)g(x) - f(x)Dg(x)}{g(x)^2}.$$

Satz 2.2.8 (Stetig partiell differenzierbare Funktionen sind differenzierbar)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen. Die Funktion $f: \Omega \to \mathbb{R}^m$ sei in Ω nach $x_1, \ldots, x_n$ partiell differenzierbar. Sind die Funktionen $\partial_j f: \Omega \to \mathbb{R}^m$ in $x \in \Omega$ stetig, so ist f in x differenzierbar.*

Folgerung 2.2.9. *Sei $k \in \mathbb{N}_0 \cup \{\infty\}$.*

1. *Mit $f, g \in C^k(\Omega, \mathbb{R}^m$, $\alpha, \beta \in \mathbb{R}$ ist auch $\alpha f + \beta g \in C^k(\Omega, \mathbb{R}^m)$.*

2. *Mit $f, g \in C^k(\Omega)$ ist auch $fg \in C^k(\Omega)$, sowie – falls $g \neq 0 auf \Omega$ – $f/g \in C^k(\Omega)$.*

3. *Sei $f \in C^k(\Omega, \mathbb{R}^m)$, $g \in C^k(\Omega, \mathbb{R}^l)$, $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen mit $f(u) \subset V$. Dann ist $g \circ f \in C^k(\Omega, \mathbb{R}^l)$.*

2.3 Minima, Maxima &c

Lemma 2.3.1. *Sei $\gamma\colon I \to \Omega$, $I = [a, b]$, $\Omega \subset \mathbb{R}^n$ stetig und stückweise C^1. Dann gilt für $f \in C^1(\Omega, \mathbb{R}^m)$, dass*

$$f(\gamma(b)) - f(\gamma(a)) = \int_a^b Df(\gamma(t))\gamma'(t)\,\mathrm{d}t.$$

Bemerkung. 1. Eine stückweise stetig differenzierbare Funktion auf einem Intervall $I = [a, b]$ ist eine Funktion f, so dass $x_0 = a < x_1 < \cdots < x_k = b$ existieren und $f|_{(x_{j-1}, x_j)}$ für $j = 1, \ldots, k$ stetig differenzierbar mit Existenz einseitiger Grenzwerte an den Grenzen des Teilintervalls für die Ableitung ist. Es sollte aufgrund der Gebietsadditivität des eindimensionalen Integrals offensichtlich sein, dass der Hauptsatz der Integral- und Differentialrechnung für solche Funktionen gilt.

2. Integrale – ebenso wie Ableitungen – von Funktionen $f\colon [a, b] \to \mathbb{R}^n$ betrachten wir komponentenweise. Die entsprechenden Integrations- und Differentiationsregeln übertragen sich.

Satz 2.3.2. *Für $f\colon \Omega \to \mathbb{R}^m$, Ω offen und zusammenhängend, gilt*

$$Df(x) = 0 \text{ für alle } x \in \Omega \quad \Rightarrow \quad f \text{ ist konstant.}$$

Satz 2.3.3 (Schrankensatz)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, $f \in C^1(\Omega, \mathbb{R}^m)$ mit $||Df(x)| \leq L$ für alle $x \in \Omega$ und ein $L \in \mathbb{R}$. Dann gilt*

$$|f(x) - f(y)| \leq L|x - y| \quad \text{für alle } x, y \in \Omega.$$

Folgerung 2.3.4. *Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega, \mathbb{R}^m)$. Sei weiters $K \subset \Omega$ kompakt. Dann existiert $L \in \mathbb{R}$ mit*

$$|f(x) - f(y)| \leq L|x - y| \quad \text{für alle } x, y \in K.$$

Definition 2.3.5. Eine Funktion $f\colon \Omega \to \mathbb{R}$ hat in $x \in \Omega$ ein lokales Minimum, falls $\delta > 0$ existiert mit

$$f(y) \geq f(x) \quad \text{für alle } y \in B_\delta(x) \cap \Omega.$$

Falls $f(y) > f(x)$ für alle $y \in B_\delta(x) \cap \Omega$, so heißt das Minimum isoliert.

Bemerkung. Lokale (isolierte) Maxima sind analog definiert. Ein Punkt der ein Maximum oder ein Minimum ist heißt Extremum.

Satz 2.3.6 (Notwendige Bedingung für Extrema)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen und $f\colon \Omega \to \mathbb{R}$ habe in $x \in \Omega$ ein lokales Extremum. Falls f in x differenzierbar ist, so folgt $Df(x) = 0$.*

Definition 2.3.7. Ein Punkt $x \in \Omega$ mit $Df(x) = 0$ heißt kritischer Punkt von f.

Definition 2.3.8. Sei $f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^n$ offen. Die zweite Ableitung f im Punkt $x \in \Omega$ ist die Bilinearform

$$D^2 f(x)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad D^2 f(x)(v, w) = \sum_{i,j=1}^n \partial_{ij}^2 f(x) v_i w_j.$$

Die Matrix $(\partial_{ij}^2 f(x))_{i,j=1}^n \in \mathbb{R}^{n \times n}$ heißt Hessematrix von f im Punkt x. Als Hesseform bezeichnet man die quadratische Form

$$v \mapsto D^2 f(x)(v, v) = \sum_{i,j=1}^n f(x) v_i v_j.$$

Bemerkung. Diese Definition ist konsistent mit der Standardinterpretation der zweiten Ableitung als Ableitung der Ableitungsfunktion. Diese wäre (in einem Punkt $x \in \Omega$ für Funktionen $f\colon \Omega \to \mathbb{R}$ ja ein Element von $L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}))$. Über das Standardskalarprodukt können wir $L(\mathbb{R}^n, \mathbb{R}))$ mit einem Vektor in $\mathbb{R}^n$ identifizieren (d.h., sei $\xi \in L(\mathbb{R}^n, \mathbb{R}))$, dann existiert ein eindeutiges $w \in \mathbb{R}^n$, so dass $\xi(v) = (w, v)$ für alle $v \in \mathbb{R}^n$).
Analog dazu identifizieren wir die dadurch entstandene lineare Abbildung in $L(\mathbb{R}^n, \mathbb{R}^n)$ mit einer Bilinearform $H(v, w) = (v, Lw)$. Auch diese Identifizierung ist eindeutig. Wir bemerken (aufgrund des Satzes von Schwarz (Satz 2.1.4), dass diese Bilinearform symmetrisch ist, also $(v, Lw) = H(v, w) = H(w, v) = (Lv, w)$.

Lemma 2.3.9. *Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^2(\Omega)$ und $\gamma \in C^2(I, \Omega)$. Dann gilt*

$$(f \circ \gamma)''(t) = D^2 f(\gamma(t))(\gamma'(t), \gamma'(t)) + Df(\gamma(t))\gamma''(t).$$

Lemma 2.3.10 (Taylorpolynom zweiter Ordnung)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^2(\Omega)$. Dann gilt*

$$\frac{f(x+h) - f(x) - \big(f(x) + Df(x)h + \frac{1}{2}D^2 f(x)(h,h)\big)}{|h|^2} \to 0 \quad \textit{für } |h| \to \mathbf{0}.$$

Wir benötigen im Weiteren einige Aussagen über Bilinearformen.

Satz 2.3.11. *Sei $b\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf dem n-dimensionalen, normierten Vektorraum $((V, ||\cdot||)$. Sei weiters*

$$\lambda = \inf\{b(x, x) : ||x|| = 1\}.$$

Dann gibt es ein $v \in V$ mit $||v|| = 1$, so dass $b(v, v) = \lambda$.

Satz 2.3.12. *Sei $f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^n$ offen und $x \in \Omega$.*

1. *Falls f in x ein lokales Minimum besitzt, so ist $D^2 f(x)$ positiv semidefinit.*
2. *Ist $Df(x) = 0$ und $D^2 f(x)$ positiv semidefinit, so hat f in x ein isoliertes Minimum.*

Definition 2.3.13 (konvexe Funktion)**.** Sei $K \subset \mathbb{R}^n$ konvex. Dann heißt $f\colon K \to \mathbb{R}$ konvex, falls für alle $x, y \in K$ gilt

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) \quad \text{für alle } t \in [0, 1].$$

Die Funktion f heißt strikt konvex, falls für $x \neq y$ und $t \in (0, 1)$ eine strikte Ungleichung gilt.

Bemerkung. Konvexität einer Funktion ist äquivalent zur Konvexität ihres Epigraphen

$$G^+(f) = \{(x, z) \in K \times \mathbb{R} : z \geq f(x)\} \subset \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}.$$

Satz 2.3.14. *Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, $f \in C^1(\Omega)$. Dann sind äquivalent*

1. *f ist konvex*
2. *$f(y) \geq f(x) + Df(x)(y - x)$ für alle $x, y \in \Omega$*
3. *$(Df(y) - Df(x))(y - x) \geq 0$ für alle $x, y \in \Omega$.*

Falls sogar $f \in C^2(\Omega)$ so sind die obigen Aussagen zusätzlich äquivalent zu $D^2 f(x)$ ist positiv semidefinit für alle $x \in \Omega$.

Bemerkung.

- Minima strikt konvexer Funktionen sind eindeutig.
- Wie im Eindimensionalen gelten einseitige Implikationen für strikte Konvexität, so folgt z.B. aus $f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^n$ offen und konvex, $D^2 f(x)$ positiv definit für alle $x \in \Omega$, dass f strikt konvex ist.
- Alle obigen Aussagen gelten analog für konkave Funktionen.

2.4 Taylorentwicklung im Mehrdimensionalen

Wir erinnern uns an den folgenden Satz aus der Analysis I.

Satz 2.4.1 (Taylorentwicklung)**.** *Sei $f \in C^{k+1}(I)$ für ein $k \in \mathbb{N}_0$ und $x_0 \in I$. Dann gilt*

$$f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j + R_k(x) \quad \text{für } x \in I,$$

wobei das Restglied $R_k(x)$ folgende Darstellungen besitzt:

$$\begin{aligned} R_k(x) &= \frac{1}{k!}\int_{x_0}^{x} (x-y)^k\, f^{(k+1)}(y)\, dy \qquad \textit{(Cauchy)} \\ &= \frac{f^{(k+1)}(\xi)}{(k+1)!}(x - x_0)^{k+1} \text{ für ein } \xi \in [x_0, x] \qquad \textit{(Lagrange).} \end{aligned} \tag{2.2}$$

Im Folgenden werden wir diese Aussage auf Funktionen in mehreren Veränderlichen verallgemeinern. Dazu zeigen wir zunächst eine Variante, welche die höheren Ableitungen in einer Darstellung als Multilinearformen nutzt.

Definition 2.4.2. Sei $\Omega \subset \mathbb{R}^n$ offen und konvex. Für $f \in C^k(\Omega)$ definieren wir die k-te Ableitung $D^k f(x)$ im Punkt $x \in \Omega$ als k-Linearform $D^k f(x) : \mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}$, wobei

$$D^k f(x)\,(v_1, \ldots, v_k) = \sum_{i_1,\ldots,i_k=1}^{n} \left(\partial^k_{i_1 \ldots i_k} f\right)(x)\,(v_1)_{i_1} \ldots (v_k)_{i_k}\,.$$

Bemerkung. Wie in Definition 2.3.8 ist dies konsistent mit der Definition der höheren Ableitungen als Ableitung von Ableitungen.

Proposition 2.4.3. *Sei $f \in C^k(\Omega)$, $\Omega \subset \mathbb{R}^n$. Für x_0, $x \in \Omega$ betrachten wir die C^k-Funktion*

$$\varphi : [0,1] \to \mathbb{R}, \quad \varphi(t) = f\,(x_0 + th) \qquad \text{mit } h = x - x_0.$$

Dann gilt

$$\varphi^{(k)}(t) = D^k f\,(x_0 + th)\,(h, \ldots, h).$$

Beweis. Für $k = 1$ gilt die Aussage nach Kettenregel und der Darstellung des totalen Differentials durch die Jacobimatrix, denn

$$\varphi'(t) = Df\,(x_0 + th)\,h = \sum_{i=1}^{n} \partial_i f\,(x_0 + th)\,h_i.$$

Für $k \geq 2$ ergibt sich induktiv mit dem Satz von Schwarz (Satz 2.1.4)

$$\begin{aligned}\varphi^{(k)}(t) &= \frac{d}{dt} \sum_{i_1,\dots,i_{k-1}=1}^{n} \left(\partial^{k-1}_{i_1\dots i_{k-1}} f\right)(x_0 + th)\,h_{i_1}\dots h_{i_{k-1}} \\ &= \sum_{i_1,\dots,i_{k-1}=1}^{n} \sum_{i=1}^{n} \left(\partial^{k}_{i_1\dots i_{k-1} i} f\right)(x_0 + th)\,h_{i_1}\dots h_{i_{k-1}} h_i \\ &= D^k f\,(x_0 + th)\,(h,\dots,h).\end{aligned}$$

□

Satz 2.4.1, angewandt auf die oben definierte Funktion φ, liefert sofort eine erste Fassung der mehrdimensionalen Taylorentwicklung.

Lemma 2.4.4. *Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, und sei $f \in C^{k+1}(\Omega)$ für ein $k \in \mathbb{N}_0$. Dann gibt es zu $x_0, x \in \Omega$ ein $\xi = (1-\tau)x_0 + \tau x$, $\tau \in [0,1]$, so dass mit $h = x - x_0$ gilt:*

$$f(x) = \sum_{j=0}^{k} \frac{D^j f\,(x_0)\,(h,\dots,h)}{j!} + \frac{D^{k+1} f(\xi)(h,\dots,h)}{(k+1)!}$$

Beweis. Wir wenden Satz 2.4.1 (Lagrangedarstellung) mit Entwicklungspunkt $t_0 = 0$ auf die C^{k+1}-Funktion $\varphi(t) = f(x_0 + th)$ an. Es ergibt sich Existenz von $\tau \in [0,1]$, so dass

$$\varphi(1) = \sum_{j=0}^{k} \frac{1}{j!}\varphi^{(j)}(0) + \frac{1}{(j+1)!}\varphi^{(j+1)}(\tau).$$

Mit Proposition 2.4.3 ergibt die Behauptung. □

Einer der größeren Hürden für das weitere Vorgehen stellt die Notation dar. Wir nennen $\alpha = (\alpha_0, \dots, \alpha_n) \in \mathbb{N}_0^n$ einen Multiindex – dieser wird angeben nach welcher Variable wie oft partiell abgeleitet wird. Sei nun α ein solcher Multiindex, $x \in \mathbb{R}^n$. Dann schreiben wir

$$\begin{aligned} |\alpha| &= \alpha_1 + \alpha_2 + \dots + \alpha_n && \text{Ordnung von } \alpha \\ \alpha! &= (\alpha_1)! \cdot (\alpha_2)! \dots (\alpha_n)! && \text{Fakultät von } \alpha \\ x^\alpha &= x_1^{\alpha_1} \cdot x_2^{\alpha_2} \dots x_n^{\alpha_n} && \text{Monom mit Exponent } \alpha \\ D^\alpha &= \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_n^{\alpha_n} && \text{Ableitung zum Multiindex } \alpha, \text{ wobei } D^0 = \mathrm{Id}. \end{aligned}$$

Insbesondere wird im Operator D^α genau α_i mal nach x_i differenziert.
Weiters benötigen wir grundlegende Eigenschaften von Polynomen, verallgemeinert auf den $\mathbb{R}^n$.

Definition 2.4.5. Eine Funktion $P : \mathbb{R}^n \to \mathbb{R}$ heißt Polynom vom Grad $k \geq 0$, wenn es $a_\alpha \in \mathbb{R}, |\alpha| \leq k$, gibt mit $a_\alpha \neq 0$ für mindestens ein $|\alpha| = k$, so dass gilt:

$$P(x) = \sum_{|\alpha| \leq k} a_\alpha x^\alpha \quad \text{für alle } x \in \mathbb{R}^n.$$

Separat betrachten wir die Nullfunktion als sogenanntes Nullpolynom vom Grad -1.

Proposition 2.4.6. *Für $x_0 \in \mathbb{R}^n$ beliebig bilden die Monome $(x - x_0)^\alpha$ mit $0 \leq |\alpha| \leq k$ eine Basis des Raums $\mathbb{P}_k$ der Polynome vom Grad $\leq k$.*

Beweis. Zunächst stellen wir fest, dass der Raum der Polynome vom Grad $\leq k$ auf $\mathbb{R}^n$ von den Monomen vom Grad $|\alpha| \leq k$ aufgespannt wird, die Polynome sind ja gerade Linearkombinationen der Monome. Davon gibt es aber insgesamt $\binom{k+n}{n}$ – Beweis durch Induktion über n.
Someit ist nur noch die lineare Unabhängigkeit der Monome zu zeigen. Diese folgt aber wie für $n = 1$ aus der Ableitungsregel

$$P(x) = \sum_{|\beta| \leq k} a_\beta (x - x_0)^\beta \Rightarrow D^\alpha P(x_0) = \alpha! a_\alpha \quad \text{für } |\alpha| \leq k.$$

Ist also $P = 0$ die Nullfunktion, so gilt $a_\alpha = 0$ für alle $|\alpha| \leq k$. $\square$

Satz 2.4.7 (Taylorentwicklung im $\mathbb{R}^n$)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, und sei $f \in C^{k+1}(\Omega)$ für ein $k \in \mathbb{N}_0$. Dann gibt es zu $x_0, x \in \Omega$ ein $\xi = (1-\tau)x_0 + \tau x$, $\tau \in [0,1]$, so dass gilt:*

$$f(x) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha + \sum_{|\alpha| = k+1} \frac{D^\alpha f(\xi)}{\alpha!} (x - x_0)^\alpha.$$

Beweis. Sei α Multiindex der Ordnung $|\alpha| = k$. Wieviele Tupel $(i_1, \ldots, i_k)$ gibt es, in denen jedes $i \in \{1, \ldots, n\}$ genau α_i mal vorkommt? Wähle α_1 Stellen für $i = 1$, aus den übrigen α_2 Stellen für $i = 2$, etc. Das ergibt die Zahl

$$\binom{k}{\alpha_1} \cdot \binom{k - \alpha_1}{\alpha_2} \cdot \ldots \cdot \binom{k - (\alpha_1 + \ldots + \alpha_{n-1})}{\alpha_n} = \frac{k!}{(\alpha_1)! \ldots (\alpha_n)!} = \frac{k!}{\alpha!}.$$

Die behauptete Entwicklung folgt nun aus Lemma 2.4.4 und der Vertauschbarkeit der partiellen Ableitungen (Satz 2.1.4), und der Ableitungsregeln für partielle Ableitungen. $\square$

Beispiel. Wir berechnen hier mit der Multiindexnotation die Taylorentwicklung erster Ordnung im Punkt $(1,1)$ für

$$f(x,y) = \frac{x-y}{x+y}$$

Es ist $f(1,1) = 0$, und die partiellen Ableitungen der Funktion lauten

$$D^{(1,0)} f(x,y) = \frac{2y}{(x+y)^2} \quad D^{(0,1)} f(x,y) = -\frac{2x}{(x+y)^2}$$
$$D^{(2,0)} f(x,y) = -\frac{4y}{(x+y)^3} \quad D^{(1,1)} f(x,y) = \frac{2(x-y)}{(x+y)^3} \quad D^{(0,2)} f(x,y) = \frac{4x}{(x+y)^3}$$

Das Taylorpolynom erster Ordnung ist somit

$$\begin{aligned}P_1(x,y) &= f(1,1) + D^{(1,0)}f(1,1)((x,y)-(1,1))^{(1,0)} + D^{(0,1)}f(1,1)((x,y)-(1,1))^{(0,1)}\\ &= \frac{1}{2}(x-1) - \frac{1}{2}(y-1) = \frac{1}{2}(x-y)\end{aligned}$$

Das Restglied lautet in Lagrangedarstellung mit Zwischenpunkt (ξ,η)

$$\begin{aligned}R_1(x,y) =& \frac{D^{(2,0)}f(\xi,\eta)}{2!0!}((x,y)-(1,1))^{(2,0)} + \frac{D^{(1,1)}f(\xi,\eta)}{1!1!}((x,y)-(1,1))^{(1,1)}\\ &+ \frac{D^{(0,2)}f(\xi,\eta)}{0!2!}((x,y)-(1,1))^{(0,2)}\\ =& \frac{2}{(\xi+\eta)^3}\left(-\eta(x-1)^2 + (\xi-\eta)(x-1)(y-1) + \xi(y-1)^2\right)\end{aligned}$$

2.5 Parameterabhängige Integrale

In diesem Abschnitt betrachten wir Integrale, in welchem die Integranden von einem Parameter abhängen, spricht Ausdrücke der Form

$$\phi : \Omega \to \mathbb{R}, \phi(x) = \int_I f(x,y)\mathrm{d}y,$$

wobei I ein Intervall und $\Omega \subset \mathbb{R}^n$ sowie $f\colon \Omega \times I \to \mathbb{R}$, $f = f(x,y)$ gegeben sind. Insbesondere sind wir an der Stetigkeit und Differenzierbarkeit dieser Integrale interessiert.

Bemerkung. Die Schreibweise $f = f(x,y)$ ist nützlich zur Notation von Funktionen auf Produkträumen um anzudeuten, dass das erste Argument aus dem ersten Raum des Produktes (hier Ω) und y aus dem zweiten Raum (hier I) gewählt werden.

Im folgenden wird sich folgende Größe als nützlich erweisen.

Definition 2.5.1. Als Oszillation einer Funktion $f\colon D \to \mathbb{R}$ bezeichnen wir den Ausdruck

$$\operatorname{osc}(f,\delta) = \sup\left\{\left|f(x) - f\left(x'\right)\right| : x, x' \in D, \left|x - x'\right| < \delta\right\}$$

Bemerkung. Man sieht, dass f genau dann gleichmäßig stetig ist, wenn gilt $\operatorname{osc}(f,\delta) \to 0$ für $\delta \to 0$.

Satz 2.5.2 (Stetigkeit von Parameterintegralen)**.** *Sei $f : \Omega \times I \to \mathbb{R}$, $f = f(x,y)$, wobei $\Omega \subset \mathbb{R}^n$ offen und $I = [a,b]$ kompakt. Ist $f \in C^0(\Omega \times I)$, so ist die Funktion*

$$\phi : \Omega \to \mathbb{R}, \phi(x) = \int_I f(x,y)dy$$

wohldefiniert und stetig.

Satz 2.5.3 (Differentiation unter dem Integral)**.** *Sei $f : \Omega \times I \to \mathbb{R}$, $f = f(x,y)$, wobei $\Omega \subset \mathbb{R}^n$ offen und $I = [a,b]$ kompakt. Es gelte:*

1. *$f(x,\cdot)$ ist Riemann-integrierbar für jedes $x \in \Omega$.*

2. Die partielle Ableitung $\frac{\partial f}{\partial x_j}$ existiert und ist stetig auf $\Omega \times I$.

Dann ist $\phi : \Omega \to \mathbb{R}, \phi(x) = \int_I f(x,y)dy$, nach x_j partiell differenzierbar, mit

$$\frac{\partial \phi}{\partial x_j}(x) = \int_I \frac{\partial f}{\partial x_j}(x,y)dy \quad \text{für alle } x \in \Omega$$

Sind f und $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ in $C^0(\Omega \times I)$, so ist $\phi \in C^1(\Omega)$.

Satz 2.5.4 (Kleiner Fubini)**.** *Seien $I = [a,b]$, $J = [\alpha, \beta]$ kompakte Intervalle. Dann gilt*

$$\int_\alpha^\beta \left(\int_a^b f(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int_a^b \left(\int_\alpha^\beta f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \quad \text{für } f \in C^0(I \times J).$$

2.6 Variationsrechnung

Wir betrachten in diesem Abschnitt Funktionen vom Typ

$$\mathcal{F} \colon C^1(I, \mathbb{R}^n) \to \mathbb{R}, \quad \mathcal{F}(u) = \int_a^b f\left(t, u(t), u'(t)\right)\,\mathrm{d}t$$

Es handelt sich hier um reellwertige Funktionen, welche auf einem unendlich dimensionalen Funktionenraum definiert sind. Solche Ausdrücke nennt man auch Funktionale. Das Funktional ist definiert durch die Das Funktional $\mathcal{F}$ ist dabei definiert durch die zugehörige Lagrangefunktion

$$f \colon I \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \;\; f = f(t,x,v).$$

Wir sind insbesondere interessiert an Minima bzw. stationären Punkten dieser Funktionale. Somit betrachten wir die Ableitung des Funktionals wenn sich die eingesetzte Funktion u ein wenig ändert, wir also eine parameterabhängige Schar von Funktionen

$$u : (-\epsilon_0, \epsilon_0) \times I \to \mathbb{R}^n, \quad u = u(\epsilon, t), \quad \text{wobei } u(0, \cdot) = u$$

in das Integral einsetzen. Dazu betrachten wir die Variation der Funktion $u(\epsilon, \cdot)$, also

$$\varphi : I \to \mathbb{R}^n, \;\; \varphi(t) = \frac{\partial u}{\partial \epsilon}(0,t),$$

so dass $u(\epsilon, t) \approx u(t) + \epsilon \varphi(t)$.

Bemerkung. Im Folgenden schreiben wir für Funktionen $f \in C^1(I \times \mathbb{R}^n \times \mathbb{R}^n)$, $f = f(t,x,v)$

$$D_x f = (\partial_{x_1} f, \ldots, \partial_{x_n} f) \in C(I \times \mathbb{R}^n \times \mathbb{R}^n; \mathbb{R}^n),$$
$$D_v f = (\partial_{v_1} f, \ldots, \partial_{v_n} f) \in C(I \times \mathbb{R}^n \times \mathbb{R}^n; \mathbb{R}^n).$$

Satz 2.6.1 (Erste Variation)**.** *Sei $f = f(t,x,v)$ eine Lagrangefunktion mit f und $D_v f$ stetig differenzierbar auf $I \times \mathbb{R}^n \times \mathbb{R}^n$. Für $u \in C^2\left((-\varepsilon_0, \varepsilon_0) \times I, \mathbb{R}^n\right)$ betrachten wir*

$$\psi(\epsilon) = \mathcal{F}(u(\epsilon, \cdot)) = \int_a^b f\left(t, u(\epsilon,t), \frac{\partial u}{\partial t}(\epsilon,t)\right)\mathrm{d}t$$

Dann gilt mit $\varphi = \frac{\partial u}{\partial \epsilon}(0, \cdot) \colon I \to \mathbb{R}^n$ die Formel

$$\frac{\mathrm{d}\psi}{\mathrm{d}\epsilon}(0) = \int_a^b (L_f(u), \varphi)\, dt + \left[\left(D_v f\left(t, u, u'\right), \varphi\right)\right]_{t=a}^{t=b} \tag{22.2}$$

Dabei ist $(\cdot, \cdot)$ das Standardskalarprodukt im $\mathbb{R}^n$ und $L_f(u) : I \to \mathbb{R}^n$ ist gegeben durch

$$L_f(u) = D_x f\left(t, u, u'\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left[D_v f\left(t, u, u'\right)\right].$$

Lemma 2.6.2 (Hauptsatz der Variationsrechnung)**.** *Sei $I = (a, b)$. Für $f \in C^0(I, \mathbb{R}^n)$ gelte*

$$\int_a^b (f, \varphi) = 0 \quad \text{für alle } \varphi \in C_c^\infty(I, \mathbb{R}^n) \tag{22.4}$$

Dann ist f die Nullfunktion.

Bemerkung. Sei $\Omega \subset \mathbb{R}^n$ offen. Mit $C_c^\infty(\Omega)$ bezeichnet man die Menge von Funktionen

$$C_c^\infty(\Omega) = \{f \in C^\infty(\Omega) : \exists K \subset \Omega, K \text{ kompakt mit } \operatorname{supp} f = \overline{\{f \neq 0\}} \subset K\}.$$

Satz 2.6.3 (Euler-Lagrange-Gleichungen)**.** *Sei $\mathcal{F}(u) = \int_a^b f(t, u, u')$ ein Funktional mit $f \in C^2(I \times \mathbb{R}^n \times \mathbb{R}^n)$, $f = f(t, x, v)$. Sei $u \in C^2(I, \mathbb{R}^n)$ stationärer Punkt, d.h.*

$$\left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F}(u + \varepsilon\varphi)\right|_{\varepsilon=0} = 0 \quad \text{für alle } \varphi \in C_c^\infty(I, \mathbb{R}^n)$$

Dann gelten die Euler-Lagrange-Gleichungen

$$L_f(u) = 0 \quad \Leftrightarrow \quad \frac{\partial f}{\partial x_i}\left(t, u, u'\right) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial f}{\partial v_i}\left(t, u, u'\right) = 0 \text{ für } i = 1, \ldots, n$$

Bemerkung.

1. Insbesondere sehen wir, dass das Erfüllen der Euler-Lagrange-Gleichungen für eine Funktion u eine notwendige Bedingung dafür ist, dass $\mathcal{F}(u)$ im Minimum ist innerhalb der Menge $A = \{u \in C^2(I = (a, b), \mathbb{R}^n) \cap C^0([a, b], \mathbb{R}^n) : u(a) = u_a, u(b) = u_b\}$.
2. Bei den Euler-Lagrange-Gleichungen handelt es sich um ein System von n Differentialgleichungen zweiter Ordnung, wie man durch Ausdifferenzieren des zweiten Terms sehen kann.

Beispiel. Die Bewegung eines Massenpunkts in einem konservativen Kraftfeld wird durch das Funktional

$$\mathcal{F}(u) = \int_a^b \left(\frac{m}{2}\left|u'\right|^2 - V(u(t))\right) \mathrm{d}t, \quad f(t, x, v) = \frac{m}{2}|v|^2 - V(x)$$

beschrieben, wobei $\frac{m}{2}\left|u'(t)\right|^2$ die kinetische Energie des Teilchens, und $V(u(t))$ die potentielle Energie des Teilchens zum Zeitpunkt t angeben.
Das zugehörige Kraftfeld ist gegeben durch $F(x) = -\operatorname{grad} V(x)$. Damit ergibt sich

$$\frac{\partial f}{\partial x_i}(t, x, v) = F_i(x), \quad \frac{\partial f}{\partial v_i}(t, x, v) = m v_i$$

Die Euler-Lagrange-Gleichungen lauten somit $F(u) - mu'' = 0$. Dies sind die Newtonschen Bewegungsgleichungen.

Kapitel 3

Diffeomorphismen

3.1 Inverse Funktionen

Definition 3.1.1. Eine Abbildung $f\colon U \to V$ zwischen offenen Mengen $U, V \subset \mathbb{R}^n$ heißt Diffeomorphismus der Klasse C^r, wobei $r \in \mathbb{N} \cup \{\infty\}$, falls f bijektiv ist und sowohl f als auch f^{-1} sind r-mal stetig differenzierbar.

Lemma 3.1.2 (Ableitung der Umkehrfunktion)**.** *Sei $f\colon U \to V$ bijektiv mit Umkehrabbildung $g\colon V \to U$, wobei $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offen. Ist f in x_0 und g in $y_0 = f(x_0)$ differenzierbar, so ist die lineare Abbildung $Df(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ invertierbar. insbesondere muss $m = n$ sein. Weiter gilt*

$$Dg(y_0) = Df(x_0)^{-1} \qquad \text{mit } y_0 = f(x_0)$$

Lemma 3.1.3 (Höhere Ableitungen der Umkehrfunktion)**.** *Seien $U, V \subset \mathbb{R}^n$ offen, und $f\colon U \to V$ bijektiv. Ist $f \in C^r(U, V)$ für ein $r \in \mathbb{N} \cup \{\infty\}$, und ist die Umkehrabbildung $g\colon V \to U$ differenzierbar, so folgt $g \in C^r(V, U)$.*

Satz 3.1.4 (Banach'scher Fixpunktsatz)**.** *Sei (X, d) ein vollständiger metrischer Raum, und $F\colon X \to X$ eine Kontraktion, das heißt es existiert ein $\theta \in [0, 1)$ mit*

$$d(F(x), F(y)) \leq \theta d(x, y) \qquad \text{für alle } x, y \in X$$

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

Bemerkung. Ein $x \in X$ mit $F(x) = x$ heißt Fixpunkt der Funktion F.

Satz 3.1.5 (über die inverse Funktion)**.** *Sei $f \in C^(\Omega, \mathbb{R}^n)$, $\Omega \subset \mathbb{R}^n$ offen. Sei $x_0 \in \Omega$ mit $Df(x_0) \in L(\mathbb{R}^n, \mathbb{R}^n)$ invertierbar. Dann existiert eine offene Umgebung U von x_0, so dass*

1. *$f(V)$ ist eine offene Umgebung von $y_0 = f(x_0)$,*
2. *$f|_U\colon U \to V$ ist ein Diffeomorphismus der Klasse C^1.*

Bemerkung. Falls $f \in C^r(\Omega)$, $r \in \mathbb{N} \cup \{\infty\}$, so ist auch $f^{-1} \in C^r(V, U)$.

Beweis. Schritt 1: Umformulierung als Fixpunktproblem $\phi_y(x) = x_0 + Df(x_0)^{-1}(y - f(x_0) - R_f(x)) = x$.

Schritt 2: Anwendung des Banach'schen Fixpunktsatzes auf ϕ_y in hinreichend kleinen Umgebungen, also $x \in B_\delta(x_0)$ und $y \in B_\epsilon(f(x_0))$.
Schritt 3: Invertierbarkeit im Punkt x_0, dann Invertierbarkeit auf ganz $U = B_\delta(x_0)$, nachdem Df auch nahe x_0 invertierbar sein muss. □

Folgerung 3.1.6. *Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in \mathbb{C}^1(\Omega, \mathbb{R}^n)$. Falls $Df(x)$ invertierbar ist für alle $x \in \Omega$, so ist $f(\Omega) \subset \mathbb{R}^n$ offen.*

3.2 Implizite Funktionen

Satz 3.2.1 (über die implizite Funktion)**.** *Sei $\Omega \subset \mathbb{R}^m \times \mathbb{R}^k$ offen, $f \in C^1(\Omega, \mathbb{R}^k)$ und sei $f(x_0, y_0) = z_0$. Falls $D_y f(x_0, y_0) \in L(\mathbb{R}^k, \mathbb{R}^k)$ invertierbar ist, so gibt es offene Umgebungen $U \subset \mathbb{R}^m$ von x_0 und $V \subset \mathbb{R}^k$ von y_0, sowie eine Funktion $g \in \mathbb{C}^1(U, V)$ mit*

$$\{(x, y) \in U \times V : f(x, y) = z_0\} = \{(x, g(x)) : x \in U\}.$$

Weiters gilt, dass für $r \in \mathbb{N} \cup \{\infty\}$ gilt

$$f \in C^r(\Omega, \mathbb{R}^k) \Rightarrow \in C^r(\Omega, \mathbb{R}^k).$$

3.3 Untermannigfaltigkeiten

Definition 3.3.1. Sei $1 \leq m \leq n$. Eine Menge $M \subset \mathbb{R}^n$ heißt m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ der Klasse C^r, wobei $r \in \mathbb{N} \cup \{\infty\}$, falls gilt: zu jedem $p \in M$ gibt es eine offene Umgebung $\Omega \subset \mathbb{R}^n$ und einen C^r-Diffeomorphismus $\phi \colon \Omega \to \phi(\Omega)$ mit

$$\phi(M \cap \Omega) = (\mathbb{R}^m \times \{0\}) \cap \phi(\Omega).$$

Bemerkung. Der Diffeomorphismus ϕ heißt auch (lokale) Plättung von M.

Satz 3.3.2. *Sei $M \subset \mathbb{R}^n$ und $m + k = n$. Dann sind äquivalent*

1. *M ist eine m-dimensionale Untermannigfaltigkeit der Klasse C^r*
2. *Niveaumengenkriterium: Zu jedem $p \in M$ gibt es eine offene Umgebung $\Omega \subset \mathbb{R}^n$ und eine Funktion $f \in C^r(\Omega, \mathbb{R}^k)$, so dass $M \cap \Omega = f^{-1}(\{0\})$ und $\operatorname{rank} Df(p) = k$*
3. *Graphenkriterium: Zu $p \in M$ gibt es eine offene Umgebung $U \times V \subset \mathbb{R}^m \times \mathbb{R}^k$ und $g \in C^r(U, V)$, so dass nach geeigneter Permutation der Koordinaten gilt*

$$M \cap (U \times V) = \{(x, g(x)) : x \in U\}.$$

Definition 3.3.3. Ein Vektor $v \in \mathbb{R}^n$ heißt Tangentialvektor von $M \subset \mathbb{R}^n$ im Punkt $p \in M$, falls es eine Abbildung $\gamma \colon (-\epsilon, \epsilon) \to M$ gibt mit $\gamma(0) = p$, $\gamma'(0) = v$. Die Menge der Tangentialvektoren von M im Punkt p ist der Tangentialraum T_pM.

Folgerung 3.3.4. *Sei $M \subset \mathbb{R}^n$ eine m-dimensionale C^1-Untermannigfaltigkeit, und $n = m + k$. Ist $p \in M \cap \Omega = f^{-1}(\{0\})$ für eine Funktion $f \in C^1(\Omega, \mathbb{R}^k)$ mit $\operatorname{rank} Df(p) = k$, so gilt*

$$T_pM = \ker Df(p).$$

Insbesondere ist T_pM eine m-dimensionaler Unterraum des $\mathbb{R}^n$.

3.4 Extrema mit Nebenbedingungen

Satz 3.4.1. *Sei* $f \in \mathbb{C}^1(\Omega, \mathbb{R}^k)$, $\Omega \subset \mathbb{R}^n$ *offen. Die Funktion* $\varphi \in C^1(\Omega)$ *habe in* p *ein Minimum unter der Nebenbedingung* $f(q) = z_0$, *also*

$$f(p) = z_0 \quad \textit{und} \quad \varphi(p) \leq \varphi(q) \quad \textit{für alle } q \in \Omega \textit{ mit } f(q) = z_0.$$

Falls $\operatorname{rank} Df(p) = k$, *so existieren* $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ *mit* $\operatorname{grad}\varphi(p) = \sum_{j=1}^{k} \lambda_j \operatorname{grad} f_j(p)$.

Kapitel 4

Gewöhnliche Differentialgleichungen

4.1 Das Anfangswertproblem

Definition 4.1.1. Sei G offen in $\mathbb{R} \times \mathbb{R}^n$ und $f \in C^0(G, \mathbb{R}^n)$, $f = f(t, x)$. Eine Funktion $x \in C^1(I, \mathbb{R}^n)$, $I \subset \mathbb{R}$ Intervall, heißt Lösung der Differentialgleichung $x' = f(\cdot, x)$, falls

$$x'(t) = f(t, x(t)) \quad \text{für alle } t \in i.$$

Insbesondere muss gelten $(t, x(t)) \in G$.
Gilt weiters

$$x(t_0) = x_0 \quad \text{für gegebene } (t_0, x_0) \in G,$$

so ist x eine Lösung des zugehörigen Anfangswertproblems (kurz AWP).

Satz 4.1.2. *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f \in C^0(G, \mathbb{R}^n)$ und $D_x f \in C^0(G, \mathbb{R}^{n \times n}$. Sei $I \ni t_0$ ein offenes Interval, so hat das Anfangswertproblem*

$$x' = f(\cdot, x) \quad \textit{auf } I \qquad x(t_0) = x_0$$

höchstens eine Lösung $x \in C^1(I, \mathbb{R}^n)$.

Lemma 4.1.3. *Sei $f \in C^0(G, \mathbb{R}^n)$ mit $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und sei $D_x f \in C^0(G, \mathbb{R}^{n \times n})$. Dann existiert zu $U_\epsilon(x_0, t_0) = (t_0 - \epsilon, t_0 + \epsilon) \times B_\epsilon(x_0) \subset K \subset G$, mit K kompakt eine Konstante $L > 0$ mit*

$$|f(t, x_0) - f(t, x_2)| \leq L|x_1 - x_2| \quad \textit{für alle } (t, x_1), (t, x_2) \in U_\epsilon(x_0, t_0).$$

Bemerkung. Funktionen $f \colon \Omega \to \mathbb{R}^n$, welche für ein $L > 0$ die Ungleichung $|f(x) - f(y)| \leq L|x - y|$ erfüllen nennt man Lipschitzstetig mit Lipschitzkonstante L.

Satz 4.1.4 (Picard-Lindelöf)**.** *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f \in C^0(G, \mathbb{R}^n)$ und $D_x f \in C^0(G, \mathbb{R}^{n \times n}$ sowie $(t_0, x_0) \in G$. Dann existiert $\delta > 0$, so dass das Anfangswertproblem*

$$x' = f(\cdot, x) \quad \textit{auf } I = (t_0 - \delta, t_0 + \delta) \qquad x(t_0) = x_0$$

eine Lösung $x \in C^1(I, \mathbb{R}^n)$ besitzt.

Satz 4.1.5 (C^0 ist ein Banachraum)**.** *Sei $I \subset \mathbb{R}$ ein Intervall. Dann ist $C^0(I, \mathbb{R}^n)$, zusammen mit der* sup-*Norm* $||u||_\infty = \sup_{x \in I} |u(x)|$, *ein Banachraum.*

Lemma 4.1.6. *Sei $f \in C^0(G, \mathbb{R}^n), (t_0, x_0) \in G$, und I offenes Intervall mit $t_0 \in I$. Für $x : I \to \mathbb{R}^n$ mit mit $\{(t, x(t)) : t \in I\} \subset G$ sind die folgenden Aussagen äquivalent.*

1. *$x \in C^1(I, \mathbb{R}^n)$ ist Lösung des Anfangswertproblems*

$$x'(t) = f(t, x(t)) \quad \text{für alle } t \in I,\ x(t_0) = x_0.$$

2. *$x \in C^0(I, \mathbb{R}^n)$ erfüllt die Gleichung*

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\,\mathrm{d}s \quad \text{für alle } t \in I.$$

Bemerkung. Setzen wir

$$\begin{aligned} t^+ &= \sup\{t \geq t_0 : \text{ es existiert eine Lösung des AWP auf } [t_0, t)\} \\ t^- &= \inf\{t \leq t_0 : \text{ es existiert eine Lösung des AWP auf } (t, t_0]\}, \end{aligned}$$

so folgt aus dem Satz von Picard-Lindelöf (Satz 4.1.4), dass $t^- < t_0 < t^+$. Weiters lässt sich auf dem Interval (t^-, t^+) mit Hilfe von Satz 4.1.2 eine eindeutige Lösung definieren. Diese Lösung wird auch maximale Lösung des AWP genannt.

Satz 4.1.7 (Lebensdauer)**.** *Sei $f : (a, b) \times D \to \mathbb{R}^n$ mit $f, D_x f$ stetig, wobei $D \subset \mathbb{R}^n$ offen und $-\infty \leq \alpha < t_0 < \beta \leq \infty$. Es sei $x : (t^-, t^+) \to D$ die maximale Lösung von*

$$x' = f(\cdot, x), x(t_0) = x_0 \in D.$$

Falls $t^+ < \beta$, so verlässt $x(t)$ für $t \nearrow t^+$ jedes Kompaktum, das heißt für jedes $K \subset D$ kompakt existiert $\delta > 0$, so dass $x(t) \notin K$ für $t > t^+ - \delta$. Entsprechendes gilt im Fall $t^- > a$.

Bemerkung. Im Fall $D = \mathbb{R}^n$ besagt der Satz

$$t^+ < b \Rightarrow \lim_{t \nearrow t^+} |x(t)| = \infty \quad \text{bzw.} \quad t^- > a \Rightarrow \lim_{t \searrow t^-} |x(t)| = \infty.$$

Beweis. Sei $t^+ < b$ und $K \subset D$ kompakt. Angenommen es gibt ein $\tau < t^+$ mit $x(t) \in K$ für alle $t \in (\tau, t^+)$. Da f stetig, ist $M = \sup\{|f(t, x)| : t \in [t_0, t^+], x \in K\} < \infty$ und es folgt

$$|x(t_2) - x(t_1)| = \int_{t_1}^{t_2} |f(t, x(t))| dt \leq M |t_2 - t_1| \qquad \text{für alle } t_{1,2} \in \left[0, t^+\right).$$

Somit existiert $x^+ = \lim_{t \nearrow t^+} x(t) \in K$. Nach dem Satz von Picard-Lindelöf (Satz 4.1.4) gibt es nun auf einem Intervall $I = (t^+ - \tau, t^+ + \tau)$ (mit gegebenenfalls kleinerem τ) eine Lösung y des Anfangswertproblems

$$y' = f(\cdot, y) \text{ auf } I, \quad y\left(t^+\right) = x^+.$$

Nach dem Eindeutigkeitssatz (4.1.2) stimmen x und y links von t^+überein, also ergibt Zusammensetzen eine Lösung $\tilde{x}$ der Differentialgleichung auf $[t_0, t^+ + \tau)$ mit $\tilde{x}(t_0) = x_0$, im Widerspruch zur Maximalität von t^+.

Der Satz ist damit nicht ganz gezeigt, denn die obige indirekte Annahme schließt nicht ein oszillierendes Verhalten aus, bei dem die Lösung für $t \nearrow t^+$ immer wieder nach K zurückkommt. Aber angenommen es gibt eine Folge $t_k \nearrow t^+$ mit $x(t_k) \in K$. Aus Kompaktheitsgründen gilt $\tilde{K} = \{x \in \mathbb{R}^n : \operatorname{dist}(x, K) \leq \varrho\} \subset D$ für $\varrho > 0$ klein, sowie $\tilde{M} = \sup\left\{|f(t,x)| : t \in [t_0, t^+], x \in \tilde{K}\right\} < \infty$. Wie schon gezeigt ist

$$\tilde{t}_k = \sup\left\{t \geq t_k : x([t_k, t]) \subset \tilde{K}\right\} < t^+,$$

und es folgt

$$0 < \varrho \leq \left|x\left(\tilde{t}_k\right) - x(t_k)\right| = \left|\int_{t_k}^{\tilde{t}_k} f(t, x(t))dt\right| \leq \tilde{M}\left(\tilde{t}_k - t_k\right).$$

Für t_k hinreichend nahe bei t^+ ist dies ein Widerspruch, und der Beweis ist komplett. □

Satz 4.1.8 (Langzeitexistenz)**.** *Sei $f : (\alpha, \beta) \times \mathbb{R}^n \to \mathbb{R}^n$ mit $f, D_x f$ stetig und $-\infty \leq \alpha < t_0 < \beta \leq \infty$. Es gebe $a, b \in C^0((\alpha, \beta))$ mit $a \geq 0$, so dass gilt*

$$|f(t,x)| \leq a(t)|x| + b(t) \quad \textit{für alle } t \in (\alpha, \beta),\ x \in \mathbb{R}^n.$$

Dann ist das Anfangswertproblem $x' = f(\cdot, x)$, $x(t_0) = x_0 \in \mathbb{R}^n$, auf ganz (α, β) lösbar.

Bemerkung. In den Beweisen zur Eindeutigkeit und Kurzzeitexistenz haben wir jeweils nur die Lipschitzstetigkeit in der zweiten Variable der Funktion $f \in C^0(G, \mathbb{R}^n)$, $f = f(t,x)$ benutzt. Auch die Wachstumsbedingung im Satz zur Langzeitexistenz folgt aus Lipschitzstetigkeit. Wir fassen also zusammen

1. Es sei $f \in C^0(G, \mathbb{R}^n)$, $G = (a,b) \times D$ offen, $a \geq -\infty$, $b \leq +\infty$, $(t_0, x_0) \in G$ und $f = f(t,x)$ *lokal* Lipschitzstetig in der zweiten Variablen, das heißt zu $x \in D$ existiert $L > 0$, $\delta > 0$, so dass

   $$|f(t,x) - f(t,y)| < L|x-y| \quad \text{für alle } y \in D,\ |x-y| < \delta,$$

   dann existiert $\tau > 0$, so dass das AWP $x' = f(\cdot, x)$, $x(t_0) = x_0$ auf $(t_0 - \tau, t_0 + \tau)$ eine eindeutige Lösung besitzt.

2. Es sei $f \in C^0(G, \mathbb{R}^n)$, $G = (a,b) \times D$ offen, $a \geq -\infty$, $b \leq +\infty$, $(t_0, x_0) \in G$ und $f = f(t,x)$ *global* Lipschitzstetig in der zweiten Variablen, das heißt es existiert $L \geq 0$ mit

   $$|f(t,x) - f(t,y)| < L|x-y| \quad \text{für alle } x, y \in D,$$

   eine eindeutig Lösung des AWP $x' = f(\cdot, x)$, $x(t_0) = x_0$ auf (a,b).

4.2 Lineare Differentialgleichungen

Wir betrachten das Problem

$$x'(t) = A(t)x(t) + b(t) \quad \text{für } t \in (\alpha, \beta),\ x(t_0) = x_0,$$

wobei $A : I \to \mathbb{R}^{n \times n}$, $b : I \to \mathbb{R}^n$, sowie $x : I \to \mathbb{R}^n$. Lineare Anfangswertprobleme besitzen nach Satz 4.1.8 immer eine eindeutige Lösung für alle Zeit, falls A und b stetig von der Zeit abhängen.

Für $b=0$ und $n=1$ bekommen wir die Lösung

$$x(t) = x_0 \exp \int_{t_0}^{t} A(s)\,\mathrm{d}s.$$

Satz 4.2.1 (Anfangswertisomorphismus)**.** *Sei $A \in C^0\left(I, \mathbb{K}^{n\times n}\right)$ für $\mathbb{K}=\mathbb{R}$ oder $\mathbb{K}=C$. Dann ist die Menge $L_A = \left\{x \in C^1\left(I, \mathbb{K}^n\right) : x' = Ax \ \text{ auf } I\right\}$ ein n-dimensionaler $\mathbb{K}$ Vektorraum, genauer ist für jedes $t_0 \in I$ die Abbildung*

$$\delta_{t_0} : L_A \to \mathbb{K}^n, \delta_{t_0}(x) = x\left(t_0\right)$$

ein Vektorraumisomorphismus.

Wir können nun ein sog. Fundamentalsystem der homogenen linearen gewöhnlichen Differentialgleichung $x'(t) = A(t)x(t)$ konstruieren, indem wir das Anfangswertproblem

$$x_j'(t) = A(t)x_j(t) \quad x_j(t_0) = e_j$$

lösen. Hier ist e_j der j-te Vektor der Standardbasis. Wir setzen die Lösungen $x_j(t)$ in einer Matrix zusammen, also

$$X(t) = (x_1(t), x_2(t), \ldots, x_n(t)).$$

Mithilfe dieses Fundamentalsystems lässt sich auch das inhomogene System lösen.

Satz 4.2.2 (Variation der Konstanten)**.** *Betrachte für $x \in C^1\left(I, \mathbb{K}^n\right)$ das Problem*

$$x'(t) = A(t)x(t) + b(t) \ \text{ auf } I, \quad x\left(t_0\right) = x_0 \in \mathbb{K}^n$$

wobei $A \in C^0\left(I, \mathbb{K}^{n\times n}\right)$ und $b \in C^0\left(I, \mathbb{K}^n\right)$. Sei $X \in C^1\left(I, \mathbb{K}^{n\times n}\right)$ ein Fundamentalsystem der homogenen Gleichung mit $X(t_0) = \mathrm{E}_n$. Dann lautet die eindeutige Lösung

$$x(t) = X(t)\left(x_0 + \int_{t_0}^{t} X(s)^{-1} b(s) ds\right)$$

Satz 4.2.3. *Die Matrix-Exponentialabbildung, definiert durch die absolut konvergente Reihe $\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$ für $A \in \mathbb{K}^{n\times n}$, hat folgende Eigenschaften:*

1. $\exp : \mathbb{K}^{n\times n} \to \mathrm{Gl}_n(\mathbb{K})$ *und* $\exp(0) = \mathrm{E}_n$.
2. $\det(\exp(A)) = \exp(\operatorname{tr}(A))$.
3. $\exp\left(SAS^{-1}\right) = S\exp(A)S^{-1}$ *für alle* $A \in \mathbb{K}^{n\times n}, S \in \mathrm{Gl}_n(\mathbb{K})$.
4. $\exp(A+B) = \exp(A)\exp(B)$, *falls* $[A,B] = AB - BA = 0$.

Satz 4.2.4 (Homogene Systeme mit konstanten Koeffizienten)**.** *Für $A \in \mathbb{K}^{n\times n}$ ist*

$$X : \mathbb{R} \to \mathbb{K}^{n\times n}, \quad X(t) = \exp(tA).$$

die Lösung des Anfangswertproblems $X' = AX, X(0) = \mathrm{E}_n$. Die Spaltenvektoren $x_j(t)$, $j = 1, \ldots, n$, sind ein Fundamentalsystem für $x' = Ax$ mit Anfangswerten $x_j(0) = e_j$.

Bemerkung. Falls A diagonalisierbar ist, folgt mit Punkt 3 des Satzes 4.2.3, dass sich eine Eigenbasis finden lässt, in welcher die Gleichung in n unabhängige skalare Gleichungen zerfällt. Deren Lösungen sind dann Exponentialabbildungen.
Im anderen Fall ist die Sache etwas komplizierter.