Analysis II

Dozent: Prof. Dr. Markus Banagl
LaTeX-Mitschrift: Catharina Hock, Finn Schneider
Lektorat und Abbildungen: Friedrich Homann

26. Mai 2023

Inhaltsverzeichnis

Inhaltsverzeichnis

Organisatorisches

Siehe auch das Infoblatt auf Banagls Website. Die Übungsblätter können von Banagls Website heruntergeladen werden: Website von Professor Banagl Die Abgabe der Übungsblätter erfolgt physisch über die Zettelkästen. Die Klausur findet am 31.07 von 9:30-11:15 Uhr in INF 277 HS1, HS2, INF 308 HS1, HS2 statt. **Die Vorlesungen finden mittwochs schon um 9:15 Uhr statt.**

Themenüberblick

- Metrische und topologische Räume, Kompaktheit, Stetigkeit
- Differenzierbarkeit von Funktionen in mehreren Variablen. (Partielle Ableitungen, Extremwertaufgaben)
- Implizit definierte Funktionen
- Diffeomorphismen
- Untermannigfaltigkeiten in $\mathbb{R}^n$, Integration entlang von Untermannigfaltigkeiten (Wegintegrale, Potentiale, Flächenintegrale)
- Transformationssatz für Integrale im $\mathbb{R}^n$
- Integralsätze von Green, Stokes, Gauß
- Gewöhnliche Differentialgleichungen (Existenz und Eindeutigkeit von Lösungen)

Übungsbetrieb

Die Plenarübung findet zum ersten Mal am 27. April statt.

Zur Notation und Nummerierung

Definitioen, Propositionen, Beispiele, Beweise etc. werden im Schema a.b.c nummeriert. Wobei a das jeweilige Kapitel (section), b das jeweilige Unterkapitel (subsection) und c die Reihenfolge im jeweiligen Unterkapitel angibt.

Verteilung des Skriptes

Wir veröffentlichen unsere Versionen primär über unseren Discord-Server. Das hat den Vorteil, dass wir dort gute Möglichkeiten haben, Feedback zu bekommen und in klarer Struktur alle Versionen veröffentlichen zu können. Abgesehen davon werden wir versuchen, möglichst viele Versionen auch über die HeiBox von Dominic Wrazidlo zu teilen.

Feedback

Verbesserungs- und Korrekturvorschläge können gerne auf unserem Discord-Server an uns weitergegeben werden. Für fachliche Kommentare kann sich gerne auch direkt an catharina.hock@stud.uni-heidelberg.de gewendet werden.

Danksagungen

Vielen Dank an Ansel Beucker, Paul Obernolte, Dominic Wrazidlo und viele weitere, die bei der Erstellung dieses Skriptes geholfen haben.

Disclaimer

Auch wenn wir versuchen, möglichst viele Fehler durch aufmerksames Bearbeiten und Verwertung von Feedback zu finden und zu verbessern, stellen wir keinen Anspruch auf Korrektheit oder Fehlerfreiheit. Insbesondere stellt dieses Skript **keine Prüfungsgrundlage** dar.

Vorlesung 1

1 Metrische Räume

Beginnen wir mit einer kurzen Motivation. Sei $X := \mathbb{R}$ und $x, y \in X$. Wir wollen nun den Abstand von x und y messen. Intuitiv ist $d(x,y) := |y - x|$. Wir erhalten also eine Abb. $x : X \times X \to \mathbb{R} \quad (x,y) \mapsto d(x,y)$.
Diese Abb. hat folgende Eigenschaften:

1. $d(x,x) = 0$ oder genauer $d(x,y) = 0 \Leftrightarrow x = y$
2. $d(x,y) = d(y,x)$ ("Symmetrie")
3. $x, y, z \in X$, dann $d(x,y) + d(y,z) \geq d(x,z)$ ("Dreiecksungleichung")

Auf Basis dieses Beispiels definieren wir

Definition 1.0.1. Ein **metrischer Raum** ist ein Paar (X, d) wobei X eine Menge und $d : X \times X \to \mathbb{R}$ eine Abbildung ist, die die obigen drei Eigenschaften hat. d heißt **Metrik** auf X

In der Physik (z.B. für die Lorentz-Metrik) muss die Positivität (erstes Axiom) nicht zwingend erfüllt sein. In dieser Vorlesung werden wir aber immer die Positivität fordern.

Bemerkung 1.0.2. Aus den drei Axiomen folgt direkt $0 = d(x,x) \leq d(x,y) + d(y,x) = 2 \cdot d(x,y)$, also $d(x,y) \geq 0$ für alle $x, y \in X$.

Eine Klasse an Beispielen kennen wir schon aus der Linearen Algebra. Dazu wiederholen wir kurz ein paar Begriffe.
Sei V ein Vektorraum (VR) über $\mathbb{R}$ oder $\mathbb{C}$.

Definition 1.0.3. Eine **Norm** auf V ist eine Abb. $\|\cdot\| : V \to \mathbb{R}$ sodass

1. $\|x\| = 0 \Leftrightarrow x = 0 \quad \forall x, y \in V$
2. $\|\lambda \cdot x\| = |\lambda| \cdot \|x\| \quad \forall x \in V, \lambda \in \mathbb{C} \quad \text{bzw.} \quad \mathbb{R}$
3. $\|x + y\| \leq \|x\| + \|y\|$

Wie man leicht nachrechnet ist $d(x,y) := \|x - y\|$ eine Metrik auf V. Ein normierter VR ist also kanonisch ein metrischer Raum.

Beispiel 1.0.4.

1. Der euklidische Raum $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$, wobei $\langle \cdot, \cdot \rangle$ das kanonische innere Produkt auf $\mathbb{R}^n$ ist, d.h.

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$
$$x = (x_1, \ldots, x_n); y = (y_1, \ldots, y_n)$$
$$\|x\| := \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Dies ist eine Norm (die Euklidische Norm). Die ersten zwei Eigenschaften prüft man leicht nach. Die letzte Eigenschaft ist gerade die Minkowski-Ungleichung. $d(x,y) = \|x-y\| = \left(\sum_{i=1}^{n}(x_i-y_i)^2\right)^{1/2}$ bezeichnen wir als **Euklidische Metrik**. Immer wenn wir $\mathbb{R}^n$ mit Metrik verwenden, dann handelt es sich i.d.R um die euklidische Metrik.

2. Sei X eine Menge $B(X) := \{f : X \to \mathbb{R} | \quad \text{f beschränkt} \quad \}$. $B(X)$ ist ein $\mathbb{R}$-VR.

$$\|f\|_\infty := \sup\{|f(x)| : x \in X\} < \infty \quad \text{denn f ist beschränkt}$$

$\|\cdot\|_\infty$ ist eine Norm auf B(X). Wieder ist nur die Dreiecksungleichung interessant zu zeigen. Für $f, g \in B(X)$ ist

$$\begin{aligned}\|f+g\|_\infty &= \sup\{|f(x)+g(x)| : x \in X\} \leq \sup\{|f(x)| + |g(x)|; x \in X\} \\ &\leq \sup\{|f(x)|; x \in X\} + \sup\{|g(x)|; x \in X\} = \|f\|_\infty + \|g\|_\infty\end{aligned}$$

Insbesondere ist $B(X)$ ein metrischer Raum.

1.1 Umgebungen und offene Mengen

Sei (X,d) ein metrischer Raum, $x \in X, r \in \mathbb{R}, r > 0$.

$$B_r(x) := \{y \in X | d(x,y) < r\}$$

Dies nennen wir auch die **offene Kugel** (oder der offene Ball) mit Radius r um x.

Definition 1.1.1. $U \subset X$ heißt **Umgebung** von $x \in X$, wenn $\exists \epsilon > 0$ mit $B_\epsilon(x) \subset U$.

Definition 1.1.2. $U \subset X$ heißt **offen** (in X), wenn $\forall x \in U$ U eine Umgebung von x ist.

Definition 1.1.3. Sei (X,d) ein metrischer Raum und $A \subset X$ ein Teilmenge. Die **auf A induzierte Metrik** $d_A : A \times A \to \mathbb{R}$ ist gegeben durch

$$d_A(x,y) := d(x,y) \forall x, y, \in A$$

Der so entstehende metrische Raum (A, d_A) heißt **metrischer Unterraum** von (X,d).

Beispiel 1.1.4. $(a,b) \subset \mathbb{R}$ ist metrisch offen. (Mit der euklidischen Metrik). $(a,b] \subset \mathbb{R}$ ist nicht offen.

Beispiel 1.1.5. $B_r(x)$ ist selbst offen: Sei $y \in B_r(x)$. Dann ist $d(x,y) < r$. Sei $\epsilon := r - d(x,y) > 0$. Dann ist $B_\epsilon(y) \subset B_r(x)$. Denn für $z \in B_\epsilon(y)$, dann ist $d(y,z) < \epsilon$ und $d(x,y) \leq d(x,y) + d(y,z) < d(x,y) + \epsilon = r$.

Vorlesung 2

1.2 Eigenschaften offener Mengen

Proposition 1.2.1. *Sei (X,d) ein metrischer Raum.*

1. *$\emptyset$ und X sind offen.*

2. *$U, V \subset X$ und U, V offen. Dann ist auch $U \cap V$ offen.*
3. *$U_i \subset X$, und U_i sind offen $\forall i \in I$, dann ist $\bigcup_{i \in I} U_i$ offen.*

Beweis.

1. Für $\emptyset$ ist nichts zu zeigen. Für X wähle z.B. $B_1(x), x \in X$. Dann ist $B_1(x) \subset X$ und wir sind fertig.
2. Sei U, V offen. Sei $x \in U \cap V$. Da U offen ist, gibt es ein $\epsilon > 0$ mit $B_\epsilon(x) \subset U$. Analog existiert ein $\delta > 0$ mit $B_\delta(x) \subset V$. Sei nun $\gamma = \min\{\epsilon, \delta\} > 0$. Dann ist $B_\gamma(x) \subset B_\epsilon(x) \cap B_\delta(x) \subset U \cap V$.
3. Seien $U_i \subset X$ offen $\forall i \in I$. Sei $x \in \bigcup_{i \in I} U_i$. Dann $\exists i_0 \in I$ mit $x \in U_{i_0}$. Nun ist U_{i_0} offen, also existiert $\epsilon > 0$ mit $B_\epsilon(x) \subset U_{i_0} \subset \bigcup_i U_i$.

□

Diese Eigenschaften motivieren folgende Abstraktion:

Definition 1.2.2 (Hausdorff)**.** Sei X eine Menge. Eine **Topologie auf** X ist eine Menge τ von Teilmengen von X, so dass gilt:

1. $\emptyset, X \in \tau \rightarrow$ *(lies: "$\emptyset$ und X sind offen")*
2. $U, V \in \tau \Rightarrow U \cap V \in \tau$
3. $U_i \in \tau \; \forall i \in I \Rightarrow \bigcup_{i \in I} U_i \in \tau$

Die Elemente von τ heißen daher "**offene** Mengen". Das Paar (X, τ) heißt **topologischer Raum**.

Bemerkung 1.2.3. Unendliche Durchschnitte von offenen Mengen sind im Allgemeinen nicht offen. Beispiel: Sei $X = \mathbb{R}$ (mit der kanonischen euklidischen Metrik). Sei $i = 1, 2, 3\ldots$ Und $U_i := (-1/i, 1/i) \subset \mathbb{R}$ offen. Aber $\bigcap_{i=1}^{\infty} U_i = \{0\}$ ist nicht offen in $\mathbb{R}$.

Definition 1.2.4. Eine Teilmenge $A \subset X$ heißt abgeschlossen, wenn $X - A$ offen in X ist, d.h. $X - A \in \tau$.

Man kann auch für Komplemente eine entsprechende Axiomatik erhalten (mit den de-Morganschen Regeln). Analog muss eine unendliche Vereinigung von abgeschlossenen Mengen nicht abgeschlossen sein, aber der Durchschnitt von (auch unendlich vielen) abgeschlossenen Mengen ist wieder abgeschlossen.

Beispiel 1.2.5.

1. Sei (X, d) ein metrischer Raum, dann ist die Menge τ_d der metrisch offenen Mengen eine Topologie auf X, die sogenannte **metrische Topologie** (auch Standardtopologie). Alle Aussagen über topologische Räume lassen sich also auch auf metrische Räume anwenden. Insbesondere ist so $(\mathbb{R}^n, \tau_{\text{euklid.}})$ ein topologischer Raum. Ebenso ist $B(X)$ für eine Menge X selbst ein topologischer Raum. (X muss dabei selbst keine topologische Struktur besitzen).
2. Sei X eine Menge. $\tau := P(X)$ die Potenzmenge von X, dann sind die Axiome aus der Definition trivialerweise erfüllt. Also ist τ eine Topologie auf X, die sogenannte "**diskrete Topologie**". In der diskreten Topologie ist jede Teilmenge von X offen, z.B. auch einelementige Mengen.

3. $[a,b] \subset \mathbb{R} = X$ ist abgeschlossen, wohingegen $(a,b] \subset \mathbb{R}$ weder offen noch abgeschlossen ist. Es ist also nicht so, dass Mengen entweder abgeschlossen oder offen sind. Im Allgemeinen sind Mengen tatsächlich weder offen noch abgeschlossen.

Bemerkung 1.2.6. Wir werden später sehen, dass obige Begriffsbildung von Abgeschlossenheit in metrischen Räumen durch eine Grenzwerteigenschaft bzgl. Folgen charakterisiert werden kann. In Analysis 1 hatten wir Abgeschlossenheit so definiert, dass eine Folge, die in einer Menge $X \subset \mathbb{R}$ liegt und in $\mathbb{R}$ konvergiert, auch einen Grenzwert in X haben muss. Unsere neue Definition ist damit konsistent.

Definition 1.2.7 (Hausdorffeigenschaft)**.** Ein topologischer Raum (X,τ) heißt **Hausdorffraum**, wenn $\forall x,y \in X$ mit $x \neq y$ existieren offene $U,V \subset X$, sodass $x \in U, y \in V$ und $U \cap V = \emptyset$. Dies nennt man auch "Trennungseigenschaft".

Proposition 1.2.8. *Metrische Räume sind Hausdorffräume.*

Beweis. Sei $x,y \in X$, $x \neq y$ mit (X,d) metrischer Raum. Nun ist $\epsilon = \frac{1}{2}d(x,y) > 0$. Sei nun $U := B_\epsilon(x)$, und $V := B_\epsilon(y)$. Dann folgt $U \cap V = \emptyset$, denn wenn $z \in U \cap V$, dann ist $d(x,z) + d(y,z) < 2\epsilon = d(x,y)$ im Widerspruch zur Dreiecksungleichung $d(x,y) \leq d(x,z) + d(y,z)$. □

Insbesondere können Mengen, die die Hausdorffeigenschaft nicht erfüllen, keine Metrik besitzen.

Beispiel 1.2.9. Ein Raum, der die Hausdorffeigenschaft nicht erfüllt ist die Gerade $\mathbb{R}$, jedoch mit einer zweiten 0. Nun ist zwar $0_1 \neq 0_2$, aber es gibt keine Umgebungen von 0_1 und 0_2 die disjunkt sind. In diesem Raum hat nun auch eine Nullfolge keinen eindeutigen Grenzwert. Deswegen fordern wir i.d.R. die Hausdorffeigenschaft, da sonst konvergente Folgen mehr als einen Grenzwert haben können.

Wie sieht es jetzt mit Teilmengen von topologischen Räumen aus? Diese würden wir gerne auch als Räume erklären.

Definition 1.2.10. Sei (X,τ) ein topologischer Raum und $Y \subset X$. Sei

$$\tau_y := \{U \cap Y | U \subset X \quad offen\} \tag{1}$$

Dann ist τ_y eine Topologie auf Y. Wir nenn sie die von τ auf Y induzierte Topologie. (Y,τ_y) heißt Unterraum von (X,τ).

Bemerkung 1.2.11. Offenheit ist keine absolute, sondern eine relative Eigenschaft. Deswegen geben Sie bei Offenheit immer an, in welchem ambienten Raum die Menge offen ist. Wenn also $V \subset Y \subset X$, dann ist zu unterscheiden zwischen "V ist offen in Y" und "V ist offen in X" Wir verwenden daher ab jetzt die Schreibweise $V \underset{\text{offen}}{\subset} Y$.

Beispiel 1.2.12. Sei $(c,b] \subset (a,b] \subset \mathbb{R}$ und $a < b < c$. Dann ist $(c,b]$ offen in $(a,b]$, denn $(c,b] = (c,b+1) \cap (a,b]$, wobei $(c,b+1)$ offen in $\mathbb{R}$ ist. Andererseits ist $(c,b]$ nicht offen in $\mathbb{R}$.

Definition 1.2.13. Sei $Y \subset X$ eine Teilmenge.

$$\mathring{Y} := \bigcup \left\{ U \subset Y | U \underset{\text{offen}}{\subset} X \right\} \subset Y \tag{2}$$

Wir nennen diese Menge das "**Innere von Y**". $\mathring{Y}$ ist offen in X. $\mathring{Y}$ ist die größte offene Teilmenge (in X), die in X enthalten ist. Ist Y bereits offen, so ist $\mathring{Y} = Y$ Analog nennen wir

$$cl(Y) = \bar{Y} := \bigcap \left\{ A \supset Y | A \underset{\text{abgeschl.}}{\subset} X \right\} \supset Y \tag{3}$$

die **abgeschlossene Hülle von** Y. $\bar{Y}$ ist abgeschlossen in X, und sogar die kleinste abgeschlossene Menge, die Y enthält.

Beispiel 1.2.14. $\overline{(a,b)} = [a,b]$, und $[a, \overset{\circ}{} b] = (a,b)$.

Bemerkung 1.2.15. Der Rand einer Teilmenge U eines topologischen Raumes X ist die Differenzmenge zwischen Abschluss und Innerem von U. Der Rand einer Menge U wird üblicherweise mit ∂U bezeichnet, also:

$$\partial U = \overline{U} - \mathring{U} = \overline{U} \cap \overline{(X - U)}$$

Beweis. Wir zeigen $x \in \overline{U} - \mathring{U} \Leftrightarrow x \in \overline{U} \cap (\overline{X - U}$.
"$\Rightarrow$" Sei $x \in \overline{U}$, $x \notin \mathring{U}$. Es ist nun $X - U \subset \overline{X - U}$. Nun ist $U \supset X - \overline{X - U}$ offen, denn $\overline{X - U}$ ist abgeschlossen. Insbesondere ist dann $X - \overline{X - U} \subset \mathring{U}$. Da $x \notin \mathring{U}$ ist auch $x \notin X - \overline{X - U}$. Also muss $x \in \overline{X - U}$ und wir sind fertig.
"$\Leftarrow$" Sei $x \in \overline{U} \cap \overline{X - U}$. Es ist klar, dass $x \in \overline{U}$. Es bleibt noch zu zeigen, dass $x \notin \mathring{U}$ ist. Nun ist $\mathring{U}$ offen, d.h. $X - \mathring{U} = \overline{X - \mathring{U}}$ist bereits abgeschlossen. Da $X - M \subseteq X - \mathring{U} = \overline{X - \mathring{U}}$. Dann folgt aus der Definition des Abschlusses schon $\overline{X - U} \subseteq \overline{X - \mathring{U}} = X - \mathring{U}$. Also folgt aus $x \in \overline{X - U}$ bereits $x \notin \mathring{U}$. □

Definition 1.2.16. Sei $Y \subset X$ eine Teilmenge. Wir sagen Y ist **dicht** in X, wenn $\bar{Y} = X$. (Für metrische Räume. In Allgemeinen topologischen Räumen muss man aufpassen)

Beispiel 1.2.17. $\mathbb{Q} \subset \mathbb{R}$. Dann ist $\bar{\mathbb{Q}} = \mathbb{R}$, d.h. $\mathbb{Q}$ ist dicht in $\mathbb{R}$.

1.3 Konvergenz in metrischen Räumen

Sei (X,d) im Folgenden ein metrischer Raum.

Definition 1.3.1. Eine Folge (x_k) von Punkten x_k in X konvergiert gegen $a \in X$, wenn $\forall \epsilon > 0 \exists N : d(x_k, a) < \epsilon \forall k \geq N$. Insbesondere ist $x_k \to a$ für $k \to \infty \Leftrightarrow d(x_k, a) \to 0$ für $k \to \infty$.

Beispiel 1.3.2. Konvergenz in $\mathbb{R}^n = X$ mit euklidischer Metrik. Sei $(x_k) \subset \mathbb{R}^n$ eine Folge mit $x_k \to a \in \mathbb{R}^n$. Wir schreiben $x_k = (x_{k1}, \ldots, x_{kn})$. Nun ist

$$|x_{kj} - a_j|^2 \leq \sum_i |x_{ki} - a_i|^2 \Rightarrow |x_{kj} - a_j| \leq \|x_k - a\| = d(x_k, a) \to 0 \tag{4}$$

Also ist $x_k \to a \in \mathbb{R}^n$ äquivalent zu $x_{kj} \to a_j \in \mathbb{R}, \forall j = 1, \ldots n$

Proposition 1.3.3. *Sei (X,d) metrischer Raum, $A \subset X$. Dann ist*

$$A \text{ abgeschl. in } X \Leftrightarrow \forall (x_k) \subset A; x_k \to x \in X \Rightarrow x \in A \tag{5}$$

Achtung! Dies gilt i.d.R. in allgemeinen topologischen Räumen nicht.

Beweis. "$\Rightarrow$" Beweis durch Widerspruch. Sei $X - A$ offen, $(x_k) \subset A, x_k \to x \in X$. Wenn $x \notin A$, dann ist $x \in X - A$ offen. Also gibt es ein $\epsilon > 0$ mit $B_\epsilon(x) \subseteq X - A$. Dies ist aber ein Widerspruch zu $x_k \to x$.
"$\Leftarrow$" Z.Z. ist nun $X - A$ offen. Sei $x \in X - A$. Angenommen, $\nexists B_\epsilon(x)$ mit $B_\epsilon(x) \subset X - A$. D.h. $\forall k \in \mathbb{N}$ ist $B_{1/k}(x) \cap A \neq \emptyset$. Wähle $x_k \in B_{1/k}(x) \cap A$ Dann ist $(x_k) \subset A$ und $x_k \to x$. Nach Voraussetzung ist nun $x \in A$, im Widerspruch zur Annahme $x \in X - A$. Also muss es ein ϵ geben, sodass $B_{1/k}(x) \cap A = \emptyset$, und $X - A$ ist offen. □

Vorlesung 3

1.4 Cauchyfolgen und Vollständigkeit

Definition 1.4.1. Sei (X, d) ein metrischer Raum, sei $(x_k) \subset X$ eine Folge von Punkten. (x_k) heißt **Cauchyfolge**, wenn $\forall \epsilon > 0\, \exists N$ mit $d(x_k, x_l) < \epsilon\ \forall k, l \geq N$

Insbesondere brauchen wir keinen Grenzwert, um Cauchyfolgen zu definieren.

Proposition 1.4.2 (Cauchyfolgen und Konvergen)**.** *Jede konvergente Folge in einem metrischen Raum ist eine Cauchyfolge.*

Beweis. Sei $(x_k) \subset X$ konvergent, $x_k \to x \in X$. Sei $\epsilon > 0$. Aufgrund der Konvergenz existiert nun ein N mit $d(x_k, x) < \frac{\epsilon}{2}\ \forall k \geq N$. Dann gilt für $k, l \geq N$ mit der Dreiecksungleichung:

$$d(x_k, x_l) \leq d(x_k, x) + d(x, x_l) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \tag{6}$$

Also ist (x_k) eine Cauchyfolge. □

In Analysis 1 hatten wir gesehen, dass für $X = \mathbb{R}$ auch die Umkehrung gilt. Nun ist die Frage: GIlt die Umkehrung der Proposition für jeden metrischen Raum? Leider ist das nicht der Fall, wie folgendes Gegenbeispiel zeigt:

Beispiel 1.4.3. Sei $X = \mathbb{Q}$ mit der von ($\mathbb{R}$, euklidische Metrik) induzierten Metrik. Wähle eine Folge $(x_k) \subset \mathbb{Q}$, mit $x_k \to \sqrt{2}$ in $\mathbb{R}$. Dann ist (x_k) in $\mathbb{Q}$ eine Cauchyfolge, die nicht konvergiert.

Dies legt folgende Definition nahe:

Definition 1.4.4.

- Ein metrischer Raum heißt **vollständig**, wenn jede Cauchyfolge konvergiert.
- Ein vollständiger, normierter Vektorraum heißt **Banachraum**.

Beispiel 1.4.5. $\mathbb{R}^n$ ist ein Banachraum. Dies folgt aus der Vollständigkeit von $\mathbb{R}$, zusammen mit koordinatenweiser Konvergenz.

Satz 1.4.6 (Schachtelungsprinzip). *Sei (X,d) ein vollständiger metrischer Raum, $A_i \subset X$ abgeschlossen, $A_i \neq \emptyset$, und*

$$A_0 \supset A_1 \supset A_2 \supset \dots$$

Wenn $\operatorname{diam}(A_i) \to 0$ *für $i \to \infty$, dann ist $\bigcap_i A_i = \{a\}$, also genau ein Punkt. Ist $A \subset X$, dann setzen wir*

$$\operatorname{diam}(A) := \sup \{d(a_1,a_2) |\, a_1, a_2 \in A\} \in \mathbb{R} \cup \{\infty\} \tag{7}$$

Es ist A beschränkt $\Longleftrightarrow \operatorname{diam}(A) < \infty$

Beweis.

1. Wir zeigen zuerst $\bigcap_i A_i \neq \emptyset$. Wähle $a_i \in A_i$, $i = 1,2,3,\dots$. So erhält man eine Folge (a_i). Für $k < l$ ist $a_k, a_l \in A_k$, also $d(a_k,a_l) \leq \operatorname{diam}(A_k) \to 0$ für $k \to \infty$. Also ist (a_i) eine Cauchyfolge. Nun ist X vollständig, d.h. $\exists a \in X$ mit $a_i \to a$. Wir behaupten nun, dass a in $\bigcap_i A_i$ enthalten ist. Es ist per Konstruktion $(a_i)_{i\geq 0} \subset A_0$. Nach Voraussetzung ist A_0 abgeschlossen. Also ist $a \in A_0$. Analog ist $(a_i)_{i\geq 1} \subset A_1$. Auch A_i ist abgeschlossen, also ist $a \in A_1$. Dies kann man analog für alle anderen A_n zeigen. Also ist $a \in \bigcap_i A_i$.

2. Zu zeigen ist die Eindeutigkeit von a. Seien $a, a' \in \bigcap_i A_i$. Dann ist $\forall i\, d(a,a') \leq \operatorname{diam} A_i \to 0 \Longrightarrow d(a,a') = 0$, also $a = a'$.

$\square$

1.5 Stetigkeit

In Analysis 1 haben wir gesehen, dass eine Abb. $f : \mathbb{R} \to \mathbb{R}$ stetig ist genau dann, wenn $\forall V \underset{\text{offen}}{\subset} \mathbb{R} : f^{-1}(V) \underset{\text{offen}}{\subset} \mathbb{R}$. Insbesondere braucht man keine Abstandsbegriffe oder die Eigenschaften von Körpern, um Stetigkeit zu definieren.

Bemerkung 1.5.1. Wir geben ab jetzt bei topologischen Räumen nicht mehr immer die Topologien mit an. Die Elemente der τs bezeichnen wir als offene Mengen.

Definition 1.5.2. Seien X, Y topologische Räume und $f : X \to Y$ eine Abbildung. f heißt **stetig**, wenn $\forall V \underset{\text{offen}}{\subset} Y : f^{-1}(V)$ ist offen in X.

Definition 1.5.3. f heißt **stetig im Punkt** $x \in X$, genau dann wenn für jede Umgebung V von $f(x)$ eine Umgebung U von x existiert mit $f(U) \subset V$. Dann gilt f ist **stetig** $\Leftrightarrow$ f stetig in x $\forall x \in X$.

Bemerkung 1.5.4. Durch Komplementbildung erhält man: f stetig $\Leftrightarrow \forall A \underset{\text{abg.}}{\subset} Y : f^{-1}(A) \underset{\text{abg.}}{\subset} X$.

Satz 1.5.5. *Seien X, Y metrische Räume $f : X \to Y$ eine Abb. Dann gilt*

$$f \quad \textit{stetig} \quad \Longleftrightarrow \forall (x_n) \subset X : x_n \to x \Rightarrow f(x_n) \to f(x)$$

Beweis. "$\Rightarrow$": Sei f stetig. $(x_n) \subset X, x_n \to x$. Z.z.: $f(x_n) \to f(x)$. f ist stetig in x. $\epsilon > 0$, $V := B_\epsilon(f(x)) \underset{\text{offen}}{\subset} Y$. Wegen der Stetigkeit von f existiert eine Umgebung U von x mit $f(U) \subset V$. Aus

$x_n \to x$ folgt $\exists N : x_n \in U \forall n \geq N$. Also gilt auch $\forall n \geq N : f(x_n) \in V = B_\epsilon(f(x))$, oder äquivalent $d(f(x_n), f(x)) < \epsilon \; \forall n \geq N. \Rightarrow f(x_n) \to f(x)$.
"$\Leftarrow$": Voraussetzung ist nun $x_n \to x \Longrightarrow f(x_n) \to f(x)$. Sei $A \subset Y$ abgeschlossen. Z.z.: ist $f^{-1}(A) \underset{\text{abg.}}{\subset} X$. Sei $(x_n) \subset f^{-1}(A)$ mit $x_n \to x \in X$. Z.z. ist $x \in f^{-1}(A)$. Mit der Voraussetzung folgt daraus $f(x_n) \to f(x)$. Nun ist $f(x_n) \in A$ für alle n, und wegen der Abgeschlossenheit von A ist auch $f(x) \in A$. Also ist $x \in f^{-1}(A)$. □

Bemerkung 1.5.6. Die Richtung "$\Rightarrow$" gilt auch in allgemeinen topologischen Räumen.
Die Rückrichtung ist im Allgemeinen jedoch falsch für nicht metrisierbare topologische Räume.

1.6 Eigenschaften stetiger Abbildungen

Proposition 1.6.1. *Seien X, Y, Z topologische Räume und $f : X \to Y$, $g : Y \to Z$ stetige Abbildungen. Dann ist auch $g \circ f : X \to Z$ stetig.*

Beweis. Sei $V \subset Z$ offen. Aus der Stetigkeit von g folgt $V' := g^{-1}(V) \underset{\text{offen}}{\subset} Y$. Da auch f stetig ist, ist $f^{-1}(V') = f^{-1}(g^{-1}(V)) = (g \circ f)^{-1}(V) \underset{\text{offen}}{\subset} X$ □

Proposition 1.6.2 (Stetigkeit algebraischer Operationen)**.**

$$\mathbb{R}^2 \xrightarrow{+,\cdot} \mathbb{R}$$
$$(x, y) \mapsto x + y$$
$$(x, y) \mapsto x \cdot y$$

Und

$$\mathbb{R} \times (\mathbb{R} - \{0\}) \to \mathbb{R}$$
$$(x, y) \mapsto \frac{x}{y}$$

Sind stetige Abbildungen.

Beweis. Wir verwenden Satz 1.5.5. Sei $((x_k, y_k)) \subset \mathbb{R}^2$ eine konvergente Folge, $(x_k, y_k) \to (x, y) \in \mathbb{R}^2$. Koordinatenweise Konvergenz im $\mathbb{R}^n$ impliziert, dass $x_k \to x$ und $y_k \to y$. Aus der Analysis 1 wissen wir, dass dann auch $x_k + y_k \to x + k$. Ähnlich zeigt man Stetigkeit für $\cdot, /$. □

Korollar 1.6.3. *Sei X ein metrischer Raum. Seien $f, g : X \to \mathbb{R}$ stetig. Dann sind auch die Funktionen $f + g$, $f \cdot g : X \to \mathbb{R}$ stetig. Ist $g(x) \neq 0 \, \forall x \in X$, dann ist auch $f/g : X \to \mathbb{R}$ stetig.*

Beweis. Eine Koordinatenweise Betrachtung impliziert, dass

$$X \overset{(f,g)}{\to} \mathbb{R}^2$$
$$x \mapsto (f(x), g(x))$$

stetig ist. Die Abb. $\mathbb{R}^2 \overset{+}{\to} \mathbb{R}$ ist ebenfalls stetig nach Prop. 1.6.1 ist auch die Komposition $f + g$ stetig. □

Definition 1.6.4. Seien X, Y topologische Räume. Eine Bijektion $f : X \overset{=}{\to} Y$ heißt **Homöomorphismus**, wenn f stetig ist und die Umkehrabbildung $f^{-1} : Y \to X$ ebenfalls stetig ist. Gibt es einen Homöomorphismus zwischen X und Y schreiben wir auch $X \cong Y$ (X und Y sind "homöomorph").

Homöomorphismen induzieren also auch Bijektionen auf den Topologien von zwei topologischen Räumen.

Beispiel 1.6.5. Die Forderung nach der Stetigkeit von f^{-1} ist nicht redundant. Sei $f = \mathrm{id} : \mathbb{R}_d \to \mathbb{R}$ Wobei $\mathbb{R}_d$ $\mathbb{R}$ mit der diskreten Topologie bezeichnet. f ist bijektiv und stetig, f^{-1} ist jedoch unstetig.

Beispiel 1.6.6. Betrachte $B_1(0) \subset \mathbb{R}^n$. Nun können wir definieren $\mathbb{R}^n \overset{f}{\to} B_1(0)$; $f(x) := \frac{x}{1+\|x\|}$. f ist stetig. Die Umkehrabbildung $f^{-1}(x) = \frac{x}{1-\|x\|}$ ist ebenfalls stetig, wie man leicht nachrechnet. Also ist f ein Homöomorphismus, und $\mathbb{R}^n \cong B_1(0)$. Topologisch sind $B_1(0)$ und $\mathbb{R}^n$ nicht unterscheidbar. Als metrische Räume sind sie aber voneinander zu unterscheiden, da $B_1(0)$ beschränkt ist und $\mathbb{R}^n$ nicht.

1.7 Gleichmäßige Konvergenz

Definition 1.7.1. Sei X eine Menge und Y ein metrischer Raum. Seien $f_n : X \to Y$ Abbildungen. Die Folge (f_n) **konvergiert gleichmäßig** gegen $f : X \to Y$, wenn $\forall \epsilon > 0\ \exists N = N(\epsilon)$ (N ist unabhängig von dem Punkt, der ausgewählt wird), sodass

$$d(f_n(x), f(x)) < \epsilon \ \forall x \in X \, \forall n \geq N$$

Satz 1.7.2. *Sei X ein topologischer Raum, Y ein metrischer Raum, und $f_n : X \to Y$ stetig. Wenn $f_n \overset{gleichmäßig}{\longrightarrow} f$, dann ist f stetig.*

Beweis. Wie in Analysis 1, ersetze nur $|\cdot|$ durch $d(\cdot, \cdot)$. □

Vorlesung 4

1.8 Stetigkeit von linearen Abbildungen

Satz 1.8.1. *Seien V, W normierte Vektorräume und $A : V \to W$ eine lineare Abbildung. Dann gilt:*

$$A \quad stetig \quad \iff \exists C > 0 : \|A(x)\| \leq C \cdot \|x\| \ \forall x \in V$$

Beweis. "$\Rightarrow$" Sei A stetig in $0 \in V$. Für $\epsilon = 1\, \exists \delta > 0$ mit $\|y\| < \delta \Rightarrow \|A(y)\| < 1$. Sei $C := \frac{2}{\delta}$. Sei $x \in V - \{0\}$. $y := \frac{x}{C\|x\|}$. Dann

$$\|y\| = \left\| \frac{x}{C\|x\|} \right\| = \frac{1}{C} = \delta/2 < \delta$$

Also ist $\|A(y)\| < 1$, d.h. $\left\| A\left(\frac{x}{C \cdot \|x\|} \right) \right\| < 1$, also $\|A(x)\| < C \cdot \|x\|$.
"$\Leftarrow$" Sei $C > 0$ eine Konstante mit $\|A(x)\| \leq C \cdot \|x\| \forall x \in V$. Sei $\epsilon > 0$. $\delta := \frac{\epsilon}{C} > 0$. Wenn $\|x_1 - x_2\| < \delta$ dann ist

$$\begin{aligned} \|A(x_1) - A(x_2)\| &= \|A(x_1 - x_2)\| \leq C\|x_1 - x_2\| \\ &< C \cdot \delta = \epsilon \end{aligned}$$

□

Definition 1.8.2. Sei $A : V \to W$ eine stetige, lineare Abbildung. Dann ist die Menge $\{\|A(x)\| \,|\, \|x\| \leq 1\}$ beschränkt, denn $\|A(x)\| \leq C\|x\| \leq C$. Insbesondere existiert

$$\|A\| := \sup \{\|A(x)\| \,|\, \|x\| \leq 1\} \in \mathbb{R}$$

Dies nennen wir die **Operatornorm**. Es gilt $\forall x \in V$ $\left\|A(\frac{x}{\|x\|})\right\| \leq \|A\|$, also $\|A(x)\| \leq \|A\| \cdot \|x\|$ $\forall x \in V$.

Beispiel 1.8.3. Sei $A : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung mit Matrixdarstellung (a_{ij}) von A bzgl. der kanonischen Basen. Dann lässt sich zeigen, dass $C = (\sum_{i,j} a_{ij}^2)^{1/2}$ dem C aus Satz 1.8.1 entspricht. Dies lässt sich mit der Cauchy-Schwarz-Ungleichung beweisen. (Der Beweis ist eine Übung.) Lineare Abbildungen zwischen endlichdimensionalen Vektorräumen sind also stetig. Dies ist aber noch nicht "die" richtige Operatornorm. Um diese zu definieren braucht man etwas Lineare Algebra.

2 Kompakte Räume

Sei im Folgenden X ein topologischer Raum.

Definition 2.0.1. Eine **offene Überdeckung** von X ist eine Familie $\{U_i\}$ von offenen Mengen $U_i \subset X$, sodass $X = \bigcup_i U_i$.

Definition 2.0.2. X heißt **kompakt**, wenn **jede** offene Überdeckung $\{U_i\}$ von X eine endliche Teilüberdeckung besitzt, d.h. $\exists i_0, \dots i_n$ mit $X = U_{i_0} \cup \dots \cup U_{i_n}$. In dieser Definition fordern wir jedoch nicht die Hausdorffeigenschaft, wie es in einigen Lehrbüchern praktiziert wird.

Satz 2.0.3 (In dieser Vorlesung: Satz 1)**.** *Sei X ein metrischer Raum, $K \subset X$ und K eine kompakte Teilmenge. Dann ist K beschränkt.*

Beweis. Sei $K \subset X$ kompakt. Wenn $K = \emptyset$, dann ist K trivialerweise beschränkt. Sei also $K \neq \emptyset$. Wähle $a \in K$. Betrachte nun immer größer werdende Bälle um a, $\{B_n(a)\}_{n=1,2,\dots}$. Dies ist eine offene Überdeckung von X und somit auch von K: $K \subset \bigcup_{n=1}^{\infty} B_n(a)$. Denn $K = \bigcup_{n=1}^{\infty}(B_n(a) \cap K)$. Die $B_n(a) \cap K$ sind mit der von X geerbten Topologie automatisch offen (prüfe dazu einfach die Axiome für Topologien nach). Diese Technik werden wir in Zukunft häufig verwenden, ohne dies so ausführlich zu notieren. Da K kompakt ist, exisitert eine endliche Teilüberdeckung $K \subset B_{n_0}(a) \cup \dots \cup B_{n_k}(a)$. Nun gibt es ein $n := \max\{n_0, \dots, n_k\}$, d.h. $K \subset B_n(a)$ und damit ist K beschränkt. □

Satz 2.0.4 (In dieser Vorlesung: Satz 2)**.** *Sei X ein topologischer Raum, der die Hausdorffeigenschaft erfüllt. Sei $K \subset X$ eine kompakte Teilmenge. Dann ist K abgeschlossen.*

Beweis. Wir zeigen, dass $X - K$ offen ist. Sei $x_0 \in X - K$ ein beliebiger Punkt. Sei $x \in K$. Mit der Hausdorffeigenschaft existieren nun $U_x, V_x \underset{\text{offen}}{\subset} X, x \in U_x$, $x_0 \in V_x$ mit $U_x \cap V_x = \emptyset$. [Eine Zeichnung] Nun ist $K \subset \bigcup_{x \in K} U_x$ eine offene Überdeckung von K (Man beachte den Hinweis im vorherigen Satz zur Notation.) (siehe Abbildung 1).
Da K kompakt ist, besitzt diese Überdeckung eine endliche Teilüberdeckung $K \subset U_{x_1} \cup \dots \cup U_{x_n}$ mit

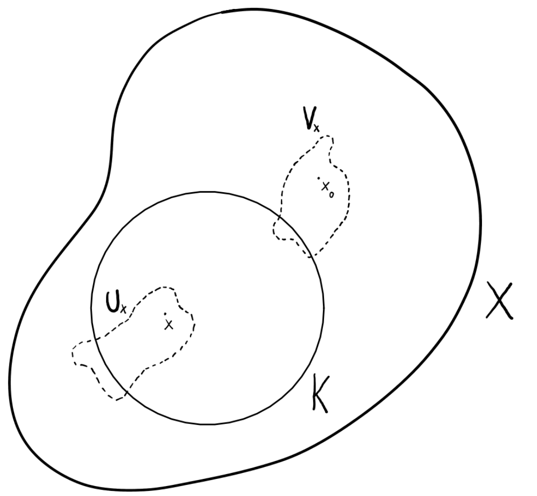

Abbildung 1: Umgebungen U_x und V_{x_0} um x und x_0.

$x_1, \dots, x_n \in K$. Sei

$$V := V_{x_1} \cap \dots \cap V_{x_n};\ x_0 \in V$$

Wieder ist V offen, da es sich um einen endlichen Schnitt handelt. V ist nun eine offene Umgebung von x_0 und $V \cap K = \emptyset$ wegen $U_x \cap V_x = \emptyset\ \forall x \in X$. Also ist $V \subset X - K$, und $X - K$ ist offen. □

Satz 2.0.5 (In dieser Vorlesung: Satz 3)**.** *Sei $A \subset X$, X kompakt und A abgeschlossen. Dann ist auch A kompakt.*

Beweis. Sei $A = \bigcup_i U_i$ eine offene Überdeckung von A, d.h. $U_i \underset{\text{offen}}{\subset} A$. Mit der von X induzierten Topologie existieren nun $U'_i \underset{\text{offen}}{\subset} X$ mit $U_i = A \cap U'_i$. Nun ist $A \subset \bigcup_i U'_i$, also $X \subseteq (X - A) \cup \bigcup_i U'_i$. Nun ist $X - A$ nach Voraussetzung offen (da A abgeschlossen ist). Da X kompakt ist besitzt die Überdeckung $(X - A) \cup \bigcup_i U'_i$ eine endliche Teilüberdeckung $X \subset (X - A) \cup U'_{i_0} \cup \dots \cup U'_{i_n}$. Daraus wollen wir eine endliche Überdeckung von A konstruieren. Das $X - A$ trägt offensichtlich nicht bei, also ist

$$A \subset U'_{i_0} \cup \dots \cup U'_{i_n} \text{ oder äquivalent } A = U_{i_0} \cup \dots \cup U_{i_n}.$$

□

Satz 2.0.6 (Kompakte Quader; in dieser Vorlesung: Satz 4)**.** *Seien $a_1, b_1, \dots, a_n, b_n \in \mathbb{R}$ mit $a_i \leq b_i\ \forall i$. Nun ist*

$$Q := [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$$

ein "Quader" in $\mathbb{R}^n$. Dann ist Q kompakt. (Wenn wir für $\mathbb{R}^n$ die euklidische Metrik verwenden.)

Beweis. Wir beweisen durch Widerspruch. Sei $\{U_i\}$ eine offene Überdeckung von Q. Angenommen,

$\{U_i\}$ besitzt keine endliche Teilüberdeckung. Dann konstruieren wir eine absteigende Folge von Quadern mit $Q = Q_0 \supset Q_1 \supset \dots$ mit folgenden Eigenschaften:

1. Für alle k gilt: Q_k wird nicht von endlich vielen U_i überdeckt.
2. $\operatorname{diam}(Q_{k+1}) = \frac{1}{2}\operatorname{diam}(Q_k)$, d.h. die Quader werden immer kleiner.

Induktiv sei $Q_k = [a_{k,1}, b_{k,1}] \times \dots \times [a_{k,n}, b_{k,n}]$ (siehe Abbildung 2). Um den nächsten Quader Q_{k+1} zu erhalten, teilen wir Q_k in 2^n kleinere, gleich große Quader, indem wir durch alle Seitenmitten schneiden. Nun ist $Q_k = Q^{(0)} \cup \dots \cup Q^{(2^n-1)}$. Nach 1. muss nun ein j existieren, sodass $Q^{(j)}$ nicht von endlich vielen U_i überdeckt wird, denn sonst müsste Q_k von endlich vielen U_i überdeckt werden.

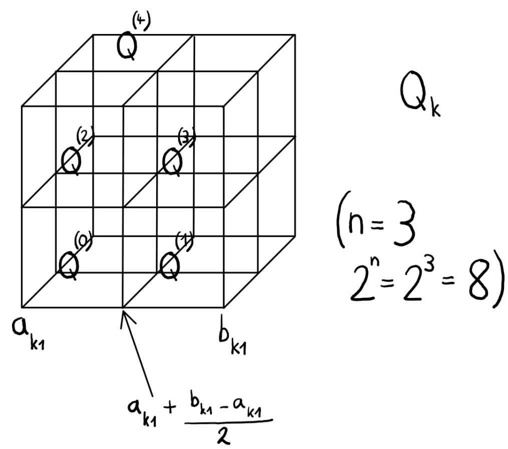

Abbildung 2: Quader, unterteilt in 8 kleinere Quader bei $n = 3$

Setze $Q_{k+1} := Q^{(j)}$. Dann gelten die geforderten Eigenschaften wieder. Nun ist

$$\operatorname{diam}(Q_k) = 2^{-k}\operatorname{diam}(Q) \to 0 \quad \text{für} \quad k \to \infty$$

Mit dem Schachtelungsprinzip ist $\bigcap_k Q_k = \{a\}$. Es existiert ein i_0 so dass $a \in U_{i_0}$. Da die U_i offen sind, exisitert nun ein $\epsilon > 0$ mit $B_\epsilon(a) \subset U_{i_0}$. Wähle nun k so, dass $\operatorname{diam}(Q_k) < \epsilon$. Dann ist $Q_k \subset B_\epsilon(a) \subset U_{i_0}$ im Widerspruch zu 1.! Also hat Q eine endliche Teilüberdeckung. □

Satz 2.0.7 (Heine-Borel)**.** *Ein Unterraum K von $\mathbb{R}^n$ ist kompakt, genau dann, wenn K abgeschlossen und beschränkt ist.*

Hier wird eine topologische und eine metrische Eigenschaft vermischt, um eine topologische Eigenschaft zu erhalten.

Beweis. "$\Rightarrow$"Sei K kompakt. Dann ist K beschränkt mit Satz 2.0.3. Außerdem ist mit Satz 2.0.4 K abgeschlossen, da $\mathbb{R}^n$ ein Hausdorffraum ist.
"$\Leftarrow$" Sei nun K beschränkt. Also existiert ein Quader $Q \subset \mathbb{R}^n$ mit $K \subset Q$. Mit Satz 2.0.6 ist Q kompakt. Nun folgt aus Satz 2.0.5, dass auch K kompakt sein muss, denn K ist abgeschlossen. □

Vorlesung 5

Bemerkung 2.0.8. Dieser Satz gilt nicht notwendigerweise in anderen metrischen Räumen. Dazu folgendes Beispiel: Betrachte ein Intervall $(a, b) =: Y \subset \mathbb{R}$. Dann ist Y mit der von $\mathbb{R}$ geerbten Metrik ein metrischer Raum. Analog ist $(a, \frac{a+b}{2}] =: X \subset Y$ ein metrischer Raum. Dann ist X abgeschlossen in Y und beschränkt, jedoch **nicht** kompakt.

2.1 Stetigkeit und Kompaktheit

Satz 2.1.1 (In dieser Vorlesung: Satz 1)**.** *Seien X,Y topologische Räume. Sei $f : X \to Y$ eine stetige Abbildung. Ist X kompakt, dann ist auch das Bild $f(X)$ kompakt.*

Beweis. Sei $f(X) \subset \bigcup_{i\in I} U_i$ eine offene Überdeckung von $f(X)$. Dann ist $\bigcup_{i\in I} f^{-1}(U_i)$ eine offene Überdeckung von X, denn f ist stetig, und jedes Urbild $f^{-1}(U_i)$ ist offen in X. Nun ist X kompakt, d.h. $X = f^{-1}(U_{i_1}) \cup \cdots \cup f^{-1}(U_{i_n})$ für endlich viele U_{i_k}. Also ist $f(X) \subset U_{i_1} \cup \cdots \cup U_{i_n}$. Also ist $f(X)$ kompakt. □

Satz 2.1.2 (In dieser Vorlesung: Satz 2)**.** *Sei X ein topologischer Raum, $f : X \to \mathbb{R}$ stetig. Ist X kompakt, dann nimmt f auf X ein Minimum und ein Maximum an.*

Beweis. Satz 2.1.1 impliziert, dass $f(X)$ kompakt ist. Nun ist $f(X) \subset \mathbb{R}$, d.h nach dem Satz von Heine-Borel (Satz 2.0.7) muss $f(X)$ abgeschlossen und beschränkt sein. Aufgrund der Beschränktheit ist $M := \sup f(X) \in \mathbb{R}$ und $m := \inf f(X) \in \mathbb{R}$. Da $f(X)$ auch abgeschlossen ist, müssen nun $M, m \in f(X)$ sein. (Angenommen, $M \notin f(X)$. Dann ist $Y - f(X)$ offen und $M \in Y - f(X)$, also ist $Y - f(X)$ eine Umgebung von M. Es gibt dann ein $\epsilon > 0$ sodass $B_\epsilon(M) \subset Y - f(X)$. Aber dann wäre z.B. auch $M > M - \epsilon/2 \in Y - f(X)$ ebenfalls eine obere Schranke für $f(X)$. Also kann es sich bei M nicht um die kleinste obere Schranke (das Supremum) von $f(X)$ handeln — im Widerspruch zur Annahme.) Also muss es $x_M, x_m \in X$ geben, sodass $M = f(x_M)$ und $m = f(x_m)$. □

Satz 2.1.3 (Bolzano-Weierstrass; in dieser Vorlesung: Satz 3)**.** *Sei X ein kompakter metrischer Raum. Dann besitzt jede Folge von Punkten $(x_n) \subset X$ eine konvergente Teilfolge.*

Beweis durch Widerspruch. Sei $(x_n) \subset X$ eine Folge, die **keine** konvergente Teilfolge besitzt. Sei $x \in X$ beliebig. Wir behaupten nun, dass eine offene Umgebung U_x von x exisitert, die nur endlich viele Folgenglieder enthält. Denn sonst könnten wir eine konvergente Teilfolge konstruieren: Angenommen, jede offene Umgebung von x enthält unendlich viele Folgenglieder. Dann enthält $B_1(x)$ ein Element der Folge x_{n_0}. Analog enthält $B_{1/2}(x)$ ein anderes Element x_{n_1} mit $n_1 > n_0$, denn $B_{1/2}(x)$ enthält nach Annahme unendlich viele Elemente von (x_n). Dieses Verfahren können wir beliebig oft wiederholen, um $n_0 < n_1 < n_2 < \ldots$ zu enthalten, sodass $x_{n_k} \in B_{1/n}(x)$ liegt. Dies ist aber eine Teilfolge $(x_{n_k})_k$ von (x_n), die nach Konstruktion gegen x konvergiert. Also muss die Behauptung wahr sein. Wähle nun jeweils U_x für alle $x \in X$ so, dass nur endlich viele Folgenglieder von (x_n) enthalten sind. Nun ist $X = \bigcup_{x\in X} U_x$ eine offene Überdeckung. Da X kompakt ist, gibt es endlich viele $x_1, \ldots, x_n \in X$, so dass $X = U_{x_1} \cup \cdots \cup U_{x_n}$. Dies impliziert jedoch, dass X nur endlich viele Folgenglieder enthält. Dies ist ein Widerspruch. □

Bemerkung 2.1.4. Die Umkehrung gilt in metrischen Räumen ebenfalls. Dies werden wir jedoch nicht beweisen.

Korollar 2.1.5. *Eine beschränkte Folge in $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.*

Beweis. Sei $(x_k) \subset \mathbb{R}^n$ beschränkt. Dann existiert ein Quader $Q \subset \mathbb{R}^n$, sodass $(x_k) \subset Q$. Nach Satz 2.0.6 ist Q kompakt. Also können wir den Satz von Bolzano-Weierstrass 2.1.3 anwenden, und (x_k) besitzt eine konvergente Teilfolge. □

Definition 2.1.6. Seien X, Y metrische Räume und $f : X \to Y$ eine Abbildung. Die Abbildung f heißt **gleichmäßig stetig**, wenn $\forall \epsilon > 0 \, \exists \delta = \delta(\epsilon) > 0$, sodass

$$d_X(x, x') < \delta \Longrightarrow d_Y(f(x), f(x')) < \epsilon \; \forall x, x' \in X$$

Satz 2.1.7 (In dieser Vorlesung: Satz 3)**.** *Ist f stetig und X kompakt, dann ist f sogar gleichmäßig stetig.*

Beweis. Sei $\epsilon > 0$, $x \in X$. Da f stetig in x ist, $\exists \, \delta(x) > 0 : \; x' \in B_{\delta(x)}(x) \Rightarrow f(x') \in B_{\epsilon/2}$. Es ist $X = \bigcup_{x \in X} B_{\delta(x)/2}(x)$ eine offene Überdeckung. Da X kompakt ist, existiert eine endliche Teilüberdeckung $X = B_{\delta(x_1)/2}(x_1) \cup \cdots \cup B_{\delta(x_n)/2}(x_n)$. Setze $\delta := \min \{\delta(x_1)/2, \ldots, \delta(x_n)/2\} > 0$. Es seien $x, x' \in X$ mit $d(x, x') < \delta$. Nun $\exists \, j : \; x \in B_{\delta(x_j)/2}(x_j)$. Dann gilt für $d(x', x_j) \leq d(x', x) + d(x, x_j) < \delta + \delta(x_j)/2 \leq \delta(x_j)$. Also ist $f(x') \in B_{\epsilon/2}(f(x_j))$. Aber auch $d(x, x_j) < \delta(x_j)/2$, also ist auch $f(x) \in B_{\epsilon/2}(f(x_j))$. Somit ist $d(f(x), f(x')) \leq d(f(x), f(x_j)) + d(f(x_j), f(x')) < \epsilon/2 + \epsilon/2 = \epsilon$ □

3 Partielle Ableitungen

3.1 Kurven im $\mathbb{R}^n$

Definition 3.1.1. Sei $I \subset \mathbb{R}$ ein Intervall. Eine **Kurve in** $\mathbb{R}^n$ ist eine **stetige** Abbildung $f : I \to \mathbb{R}^n$. Wir können dann schreiben $f(t) = (f_1(t), \ldots, f_n(t))$, wobei $f_i : I \overset{\text{stetig}}{\to} \mathbb{R}$ für alle $i = 1, \ldots, n$.

Beispiel 3.1.2. $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = (\cos(t), \; \sin(t))$. Dies entspricht einem Kreis. Wir können aber auch $g(t) := (\cos(2t), \; \sin(2t))$ setzen. f und g haben zwar dasselbe Bild, aber $f \neq g$ als Kurven, da für eine Kurve auch die Parametrisierung durch t relevant ist. g ist eine "Reparametrisierung" von f.

Definition 3.1.3. Eine Kurve f in $\mathbb{R}^n$ heißt **differenzierbar**, wenn ihre Komponenten f_i differenzierbar sind $\forall \, i = 1, \ldots, n$. Sei $f : I \to R^n$ differenzierbar. Dann heißt die Ableitung $f'(t) := (f_1'(t), \ldots, f_n'(t))$. Diesen Vektor nennen wir auch **Tangentialvektor** an die Kurve f zum Zeitpunkt t.

Bemerkung 3.1.4. Insbesondere sprechen wir nicht von einem Tangentialvektor in einem Punkt auf dem Bild der Kurve, sondern von einem Tangentialvektor zu einem bestimmten Zeitpunkt t. Denn wenn die Kurve sich selbst schneidet, wie z.B. in der Mitte des ∞-Zeichens, dann ist dort der Tangentialvektor nicht eindeutig definiert.

Beispiel 3.1.5. Sei $f(t) = (\cos(t), \sin(t)) \Rightarrow f'(t) = (-\sin(t), \cos(t))$

Definition 3.1.6. Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine Abbildung. Sei $x \in U$, $x = (x_1, \ldots, x_n)$ in Standardkoordinaten auf $\mathbb{R}^n$. Sei $\{e_1, \ldots, e_n\}$ die Standardbasis von $\mathbb{R}^n$, d.h. $(e_i)_j = \delta_{ij}$. f heißt im Punkt $x \in U$ **partiell differenzierbar** in der i-te Koordinatenrichtung, wenn der Limes

$$(\partial_i f)(x) := \lim_{t \to 0} \frac{f(x + t \cdot e_i) - f(x)}{t}$$

existiert. $(\partial_i f)(x)$ heißt dann partielle Ableitung von f im Punkt x in die i-te Koordinatenrichtung.

Bemerkung 3.1.7. Alternative Schreibweisen sind $\frac{\partial f}{\partial x_i}(x) = \left(\frac{\partial}{\partial x_i} f\right)(x) = (\partial_i f)(x)$.

Bemerkung 3.1.8 (Berechnung von partiellen Ableitungen). Setze $g(t) := f(x_1, x_2, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$. Dann ist

$$\begin{aligned} g'(x_i) := \frac{\mathrm{d}g}{\mathrm{d}t} &= \lim_{t \to 0} \frac{1}{t}(g(x_i + t) - g(x_i)) \\ &= \lim_{t \to 0} \frac{1}{t}(f(x + t \cdot e_i) - f(x)) = (\partial_i f)(x) \end{aligned}$$

Wir können also alle Ableitungsregeln aus Analysis 1 auch für partielle Ableitungen verwenden.

Bemerkung 3.1.9 (Geometrische Interpretation). Wir definieren den Graphen von f als

$$\mathrm{Gr}(f) := \left\{(x, f(x)) \in R^{n+1} \,|\, x \in U\right\}$$

Wir erhalten dann eine n-dimensionale Oberfläche im $\mathbb{R}^{n+1}$. Die partielle Ableitung in $x \in X$ in der i-ten Komponente ist dann die Steigung der Tangente an die Kurve in Richtung x_i, analog zum Fall $n = 1$ (siehe Abildung 3).

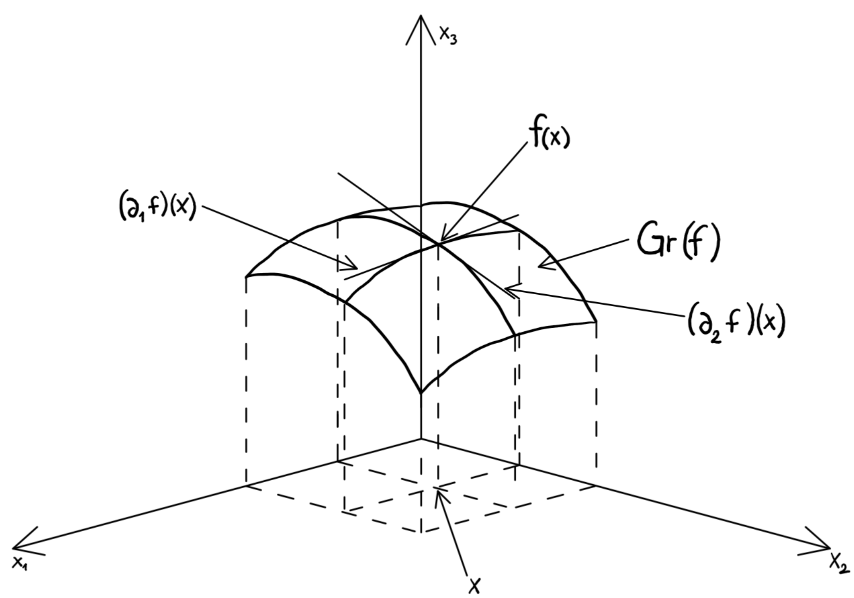

Abbildung 3: 2-dimensionale Oberfläche im $\mathbb{R}^3$

Definition 3.1.10. Ist f partiell differenzierbar in die i-te Richtung für alle $x \in U$, so sagen wir f ist **partiell differenzierbar** in die i**-te Richtung auf** U**.** Dann erhalten wir eine Abb. $\partial_i f : U \to \mathbb{R}$.

Beispiel 3.1.11. Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) := \frac{x_1 \cdot x_2}{x_1^2 + x_2^2}$ für $x \neq 0$, und $f(0,0) = 0$. Für $x \neq 0$ ist $(\partial_1 f)(x) = \frac{x_2(x_1^2+x_2^2)-(x_1x_2)2x_1}{(x_1^2+x_2^2)^2} = \frac{x_2^3 - x_1^2 x_2}{(x_1^2+x_2^2)^2}$ und analog für $\partial_2 f$. Für $x = 0$ ist

$$(\partial_1 f)(0) = \lim_{t\to 0} \frac{1}{t}(f(t,0) - f(0,0)) = 0$$

Analog erhält man $(\partial_2 f)(0) = 0$. Also ist f auf U partiell differenzierbar in beide Richtungen. Aber für die Folge $(a_n) \subset \mathbb{R}^2$, $a_k := (1/k,\ 1/k)$ ist $f(a_k) = \frac{1/k^2}{2/k^2} = \frac{1}{2}$, aber $\frac{1}{2} \neq 0 = f(0)$. D.h. eine partiell differenzierbare Funktion ist nicht wie im eindimensionalen Fall notwendigerweise stetig.

Vorlesung 6

Definition 3.1.12. Ein **Vektorfeld** ist eine Abbildung $V : U \to \mathbb{R}^n$, wobei $U \underset{\text{offen}}{\subset} \mathbb{R}^n$. Wir können V auch komponentenweise darstellen: $V = (V_1, \dots, V_n)$, mit $V_i : U \to \mathbb{R}$. Wir sagen V **ist partiell differenzierbar auf** U, wenn jede Komponentenfunktion V_i auf U partiell differenzierbar ist. Anschaulich kann man sich vorstellen, dass V jedem Punkt in U einen Vektorpfeil zuordnet.

Bemerkung 3.1.13. Sei $f : U \to \mathbb{R}$ eine Funktion, $U \underset{\text{offen}}{\subset} \mathbb{R}^n$. Sei f partiell differenzierbar auf U. Dann definiert

$$\begin{aligned} V : U &\to \mathbb{R}^n \\ u &\mapsto (\partial_1 f, \dots, \partial_n f)(u) \end{aligned}$$

ein Vektorfeld auf U. Dabei wird $\boldsymbol{\nabla}(f) := (\partial_1 f, \dots, \partial_n f)$ (sprich: "Nabla f") als **Gradient von** f bezeichnet. Man schreibt auch $\boldsymbol{\nabla} = (\partial_1, \dots, \partial_n)$.

Beispiel 3.1.14. Sei $f(x) = x_1^2 + \dots + x_n^2$. Dann ist $\mathrm{grad}(f) = (2x_1, \dots, 2x_n) = 2 \cdot x \in \mathbb{R}^n$.

Bemerkung 3.1.15. Sei nun $V : U \to \mathbb{R}^n$ ein partiell differenzierbares Vektorfeld. Dann definieren wir

$$\mathrm{div}(V) := \frac{\partial V_1}{\partial x_1} + \dots + \frac{\partial V_n}{\partial x_n}$$

Man nennt dies die **Divergenz von** V, und es handelt sich um eine Abbildung $\mathrm{div}(V) : U \to \mathbb{R}$. Formell lässt sich auch schreiben

$$\langle (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}), (V_1, \dots, V_n) \rangle = \langle \boldsymbol{\nabla}, V \rangle = \boldsymbol{\nabla} \cdot V$$

Wobei $\langle \cdot, \cdot \rangle$ das euklidische innere Produkt bezeichnet.

Beispiel 3.1.16. Sei $V(x) = (x_1, \dots, x_n)$. Dann ist $\mathrm{div}(V) = \sum_{i=1}^n \frac{\partial x_i}{\partial x_i} = n$.

Bemerkung 3.1.17. Wir können nun aus einer partiell differenzierbaren Funktion durch den Gradienten ein Vektorfeld konstruieren, und umgekehrt aus einem partiell differenzierbaren Vektorfeld eine Funktion, indem wir die Divergenz bilden.

3.2 Partielle Ableitungen höherer Ordnung

Definition 3.2.1. Sei $f : U \to \mathbb{R}$ partiell differenzierbar. Dann exisiteren $\partial_i f : U \to \mathbb{R}\ \forall i$. Sind alle $\partial_i f$ selbst wieder partiell differenzierbar, dann sagen wir f ist **zweimal partiell differenzierbar**. Es existieren dann $\partial_j(\partial_i f) : U \to \mathbb{R}\ \forall i, j$. Indunktiv lässt sich analog k-malige partielle Differenzierbarkeit definieren.

Satz 3.2.2 (Schwarz)**.** *Sei $f : U \to \mathbb{R}$, $U \underset{\text{offen}}{\subset} \mathbb{R}^n$ zweimal stetig partiell differenzierbar (d.h. f ist stetig, und die ersten und zweiten partiellen Ableitungen von f existieren und sind auch stetig). Dann gilt*

$$(\partial_j(\partial_i f))(a) = (\partial_i(\partial_j f))(a)\ \forall a \in U$$

d.h. die Operatoren der partiellen Ableitungen kommutieren.

Bemerkung 3.2.3. Man schreibt auch $\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2}{\partial x_j \partial x_i}$ unter Beachtung der Reihenfolge.

Beweis des Satzes von Schwarz. Für $i = j$ ist nichts zu zeigen. Sei $i \neq j$. Setze $g(x, y) := f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_{j-1}, y, a_{j+1}, \dots, a_n)$. Da U offen ist, $\exists \epsilon > 0$ so dass g auf $(a_i - \epsilon, a_i + \epsilon) \times (a_j - \epsilon, a_j + \epsilon)$ definiert ist. Zu zeigen ist für $a := (a_i, a_j)$ $\partial_2\partial_1 g(a) = \partial_1\partial_2 g(a)$. Sei $G_y(x) := g(x, y) - g(x, a_j)$. Mit dem Mittelwertsatz $\exists \xi$ zwischen x und a_i, sodass gilt

$$(x - a_i)G_y'(\xi) = G_y(x) - G_y(a_i)$$

Es gilt nun $G_y'(\xi) = \partial_1 g(\xi, y) - \partial_1 g(\xi, a_j)$. Betrachte die Funktion $y \mapsto \partial_1 g(\xi, y)$. Erneut gibt es wegen des Mittelwertsatzes ein η zwischen y un a_j, so dass

$$\begin{aligned}
(y - a_j)(\partial_2\partial_1 g)(\xi, \eta) &= \partial_1 g(\xi, y) - \partial_1 g(\xi, a_j) \\
(x - a_i) \cdot (y - a_j)(\partial_2\partial_1 g)(\xi, \eta) &= (x - a_i)(\partial_1 g(\xi, y) - \partial_1 g(\xi, a_j)) \\
&= (x - a_i) \cdot G_y'(\xi) = G_y(x) - G_y(a_i) \\
&= g(x, y) - g(x, a_j) - g(a_i, y) + g(a_i, a_j)
\end{aligned}$$

Besonders interessiert sind wir an

$$(x - a_i) \cdot (y - a_j)(\partial_2\partial_1 g)(\xi, \eta) = g(x, y) - g(x, a_j) - g(a_i, y) + g(a_i, a_j)$$

Durch Vertauschen der Rollen von x und y in dem obigen Argumnt erhält man $\hat{\xi}$ zwischen x und a_i sowie $\hat{\eta}$ zwischen y und a_j mit

$$(x - a_i) \cdot (y - a_j)(\partial_1\partial_2 g)(\hat{\xi}, \hat{\eta}) = g(x, y) - g(x, a_j) - g(a_i, y) + g(a_i, a_j)$$

Wenn $(x - a_i)(y - a_j) \neq 0$, dann ist $\partial_2\partial_1(\xi, \eta) = \partial_1\partial_2(\hat{\xi}, \hat{\eta})$. Wenn (x, y) gegen (a_i, a_j) konvergiert, dann konvergieren auch $(\xi, \eta), (\hat{\xi}, \hat{\eta})$ gegen (a_i, a_j). Nun sind $\partial_2\partial_1 g$, $\partial_1\partial_2 g$ nach Voraussetzung stetig. Also gilt auch

$$(\partial_2\partial_1 g)(a_i, a_j) = (\partial_1\partial_2 g)(a_i, a_j)$$

$\square$

Beispiel 3.2.4. Sei $V : U \to \mathbb{R}^3$, $U \underset{\text{offen}}{\subset} \mathbb{R}^3$ ein Vektorfeld. Wir nehmen an, dass V stetig partiell differenzierbar ist. Existiert nun ein "Potential" für V, d.h. eine Funktion $f : U \to \mathbb{R}$ mit $\boldsymbol{\nabla} f =$

V? Im eindimensionalen Fall handelt es sich bei dem Potential um die Stammfunktion. In höheren Dimensionen wird diese Frage jedoch schwieriger. Dafür wird der Satz von Schwarz noch nützlich sein.

3.3 Kreuzprodukt von Vektoren im $\mathbb{R}^3$

Seien $a, b \in \mathbb{R}^3$. Das Kreuzprodukt ist definiert als

$$\mathbb{R}^3 \times \mathbb{R}^3 \xrightarrow{\times} \mathbb{R}^3$$

$$a \times b = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2b_3 - a_3b_2,\, a_3b_1 - a_1b_3,\, a_1b_2 - a_2b_1)$$

Der Ausdruck mit der Matrixdeterminante dient dabei lediglich als Merkhilfe, weil er streng genommen keinen Sinn ergibt. (Allerdings sieht man in dieser Darstellung sehr schön, wieso $a \times b$ orthogonal zu a und b ist). Einige Eigenschaften des Kreuzprodukts sind

1. $(a \times b) \perp \mathbb{R}\langle a, b\rangle$
2. $e_1 \times e_2 = e_3$ und zyklische Permutationen davon.
3. $\|a \times b\|^2 = \|a\|^2 \cdot \|b\|^2 - \langle a, b\rangle^2$

Sei $V : U \to \mathbb{R}^3$, $U \underset{\text{offen}}{\subset} \mathbb{R}^3$ ein stetig partiell differenzierbares Vektorfeld. Setze $a = \nabla$, $b = V$. Dann ist

$$\boldsymbol{\nabla} \times V = (\partial_2 V_3 - \partial_3 V_2,\, \partial_3 V_1 - \partial_1 V_3,\, \partial_1 V_2 - \partial_2 V_1) =: \operatorname{rot}(V)$$

$\operatorname{rot}(V)$ ist wieder ein Vektorfeld auf U.
Kehren wir zurück zu unserer ursprünglichen Frage, nämlich, wann ein Vektorfeld V ein Potential besitzt. Nehmen wir zunächst an, V besitzt ein Potential f, d.h. $\boldsymbol{\nabla} f = V$. Dann ist

$$\text{rot(V)} = \operatorname{rot}(\boldsymbol{\nabla} f) = (\partial_2\partial_3 f - \partial_3\partial_2 f, \partial_3\partial_1 f - \partial_1\partial_3 f, \partial_1\partial_2 f - \partial_2\partial_1 f) = 0$$

Nach dem Satz von Schwarz (Satz 3.2.2). Wir haben also gezeigt: **Wenn V ein Potential besitzt, dann muss rot$(V) = 0$.**

Definition 3.3.1. Sei $f : U \to \mathbb{R}$, $U \underset{\text{offen}}{\subset} \mathbb{R}^n$, eine zweimal partiell differenzierbare Funktion. Wir definieren

$$\Delta f := \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2}$$

wobei $\Delta f : U \to \mathbb{R}$, $\Delta = \langle \boldsymbol{\nabla}, \boldsymbol{\nabla} \rangle = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$. Wir nennen Δ den **Laplaceoperator**.

Definition 3.3.2. f heißt **harmonisch**, wenn $\Delta(f) = 0$. Harmonische Funktionen spielen eine wichtige Rolle in der Geometrie, der Topologie und weiteren Bereichen der Mathematik.

Vorlesung 7

3.4 Totale Differenzierbarkeit

Bemerkung 3.4.1 (Motivation). Sei $f : \mathbb{R} \to \mathbb{R}$ in $x \in \mathbb{R}$ differenzierbar. Dann ist

$$f'(x) := \lim_{\xi \to 0} \frac{1}{\xi}(f(x+\xi) - f(x))$$

$f'(x)$ ist nun eine Konstante, die wir A nennen. Dann ist offenbar

$$\lim_{\xi \to 0} \frac{1}{\xi}(f(x+\xi) - f(x) - A\xi) = \lim_{\xi \to 0} r(\xi) = 0$$

Wobei wir den "Rest" von f definieren als $r(\xi) := \frac{1}{\xi}(f(x+\xi) - f(x) - A\xi)$. Nun gilt

$$f(x+\xi) = f(x) + A \cdot \xi + r(\xi) \quad \text{mit} \quad \lim_{\xi \to 0} r(\xi)/|\xi| = 0$$

Dies führt uns zu folgender Definition von (totaler) Differenzierbarkeit in höheren Dimensionen:

Definition 3.4.2. Sei $f : U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offene Teilmenge. Sei $x \in U$. Wir sagen f ist **total differenzierbar in** x (oder einfach "**differenzierbar**" im Punkt x), wenn eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ [1] existiert, sodass

$$f(x+\xi) = f(x) + A \cdot \xi + r(\xi) \quad \text{mit} \quad \lim_{\xi \to 0} \frac{r(\xi)}{\|\xi\|} = 0$$

Wir können ξ und f als $\xi = (\xi_1, \dots, \xi_n)^T$ und $f = (f_1, \dots, f_n)^T$ auffassen.

Bemerkung 3.4.3. Wir schreiben mir Landau-Notation daher auch $r(\xi)$"="$\mathrm{o}(\|\xi\|)$.

Satz 3.4.4. *$f : U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen. Sei f in $x \in U$ total differenzierbar. Dann gilt:*

1. *f ist stetig in x, und*
2. *Jede Komponentenfunktion $f_i : U \to \mathbb{R}$ ist partiell differenzierbar in x, und es gilt $A = (\partial_j f_i(x))$. Insbesondere ist A eindeutig.*

Beweis. 1. $f(x+\xi) = f(x) + A\xi + r(\xi)$, $r(\xi) = \mathrm{o}(\|\xi\|)$. Die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ ist stetig (siehe auch Übungsblatt 2). Also ist $\lim_{\xi \to 0} A\xi = 0$. Somit ist

$$\lim_{\xi \to 0} f(x+\xi) = f(x) + \lim_{\xi \to 0} \left(A\xi + \frac{r(\xi)}{\|\xi\|}\|\xi\| \right) = f(x)$$

denn $r(\xi)/\|\xi\|$ geht bereits gegen 0 für $\xi \to 0$, also insbesondere auch $\frac{r(\xi)}{\|\xi\|}\|\xi\|$. Also ist f stetig in x.
2. Es ist

$$f_i(x+\xi) = f_i(x) + \sum_j a_{ij}\xi_j + r_i\xi$$

[1] Wir schreiben statt $A(\xi)$ auch $A \cdot \xi$, denn wir können jederzeit eine Matrixdarstellung von A bezüglich der kanonischen Basen wählen, da es sich um eine lineare Abbildung von $\mathbb{R}^m$ nach $\mathbb{R}^n$ handelt.

Sei $\xi = \xi_j \cdot e_j$. Die partielle Ableitung von f_i (deren Existenz wir zeigen wollen) ist definiert als

$$(\partial_j f_i)(x) = \lim_{\xi \to 0} \frac{f_i(x + \xi_j e_j) - f_i(x)}{\xi_j} = \lim_{\xi \to 0} a_{ij} + \frac{r_i(\xi)}{\|\xi\|} = a_{ij}$$

Also ist ist f_i differenzierbar in x. □

Bemerkung 3.4.5. Die Matrix $A = (a_{ij}) = (\partial_j f_i(x))$ wird auch als **Differential von** f , **Jacobi-Matrix von** f oder **Funktionalmatrix von** f bezeichnet. Man schreibt auch $(Df)(x) = J_f(x) = \frac{\partial(f_1,\ldots,f_m)}{\partial(x_1,\ldots,x_n)}$.

Satz 3.4.6 (Hinreichende Bedingung für totale Differenzierbarkeit)**.** *Sei $f : U \to \mathbb{R}$ eine partiell differenzierbare Funktion, $U \underset{offen}{\subset} \mathbb{R}^n$. Sind die $\partial_j f$ stetig in $x \in U$ $\forall j$, dann ist f total differenzierbar in x.*

Beweis. Sei $x^{(0)} := (x_1, \ldots, x_n)^T$, $x^{(1)} := (x_1+\xi_1, x_2, \ldots, x_n)^T$, $x^{(2)} := (x_1+\xi_1, x_2+\xi_2, x_3, \ldots, x_n)^T$ und analog für $x^(i)$, $i > 2$. Nun existiert nach dem Mittelwertsatz ein $t \in [0, 1]$, sodass

$$f(x^{(1)}) - f(x^{(0)}) = (\partial_1 f)\Big(x^{(0)} + t_1\xi_1 e_1\Big) \cdot \xi_1$$

Analog existiert ein $t_2 \in [0, 1]$ mit

$$f(x^{(2)}) - f(x^{(1)}) = (\partial_2 f)(x^{(1)} + t_2\xi_2 e_2) \cdot \xi_2$$

Dies kann man für alle $0 < i \leq n$ formulieren, bis man schließlich ankommt bei

$$f(x^{(n)}) - f(x^{(n-1)}) = (\partial_n f)\Big(x^{(n-1)} + t_n\xi_n e_n\Big) \cdot \xi_n$$

Indem man alle Ausdrücke der Form $f(x^{(i)}) - f(x^{(i-1)})$ aufsummiert, erhält man

$$f(x+\xi) - f(x) = \Big(f(x^{(n)}) - f(x^{(n-1)})\Big) + \Big(f(x^{(n-1)}) - f(x^{(n-2)})\Big) + \cdots + \Big(f(x^{(1)}) - f(x^{(0)})\Big) \tag{8}$$

$$= \sum_{i=1}^{n} (\partial_i f)(x^{(i-1)} + t_i\xi_i e_i) \cdot x_i \tag{9}$$

Definiere $a_j := (\partial_j f)(x)$. Dann wird

$$f(x+\xi) - f(x) = \sum_{i=1}^{n} (\partial_i f)(x^{(i-1)} + t_i\xi_i e_i) \cdot x_i - \sum_j a_j\xi_j + \sum a_j\xi_j \tag{10}$$

$$= \sum_j a_j\xi_j + \sum_{i=1}^{n} \Big(\partial_j f(x^{(j-1)} + t_j\xi_j e_j) - a_j\Big)\xi_j \tag{11}$$

Wir haben jetzt fast schon die gewünschte Form $f(x+\xi) = f(x) + A\xi + r(\xi)$ erreicht. Es bleibt noch zu zeigen, dass

$$r(\xi) := \sum_{j=1}^{n} \Big(\partial_j f(x^{(j-1)} + t_j\xi_j e_j) - a_j\Big)\xi_j$$

die geforderten Eigenschaften eines Rests hat. D.h.

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

Nach Voraussetzung sind die $\partial_j f$ nun stetig, und $\lim_{\xi\to 0} x^{(j-1)} + t_j\xi_j e_j = \lim_{\xi\to 0} x^{(j-1)} = x^{(0)}$. Also ist

$$\lim_{\xi\to 0}(\partial_j)(x^{(j-1)} + t_j\xi_j e_j) - a_j = (\partial_j(x^{(0)})) - a_j = 0$$

und wie gefordert ist $\lim_{\xi\to 0} \frac{r(\xi)}{\|\xi\|} = 0$. □

Bemerkung 3.4.7 (Zusammenfassung)*.* Ist eine Funktion stetig partiell differenzierbar, so ist sie auch total differenzierbar. Ist eine Funktion total differenzierbar, so ist sie auch stetig und partiell differenzierbar. Die Umkehrungen gelten jedoch nicht.

3.5 Verallgemeinerte Kettenregel

Satz 3.5.1. *Sei $f : U \to V$ (total) differenzierbar im Punkt $x \in U$, wobei $U \underset{\text{offen}}{\subset} \mathbb{R}^n$, $V \underset{\text{offen}}{\subset} \mathbb{R}^m$, und $g : V \to \mathbb{R}^k$ differenzierbar im Punkt $y := f(x) \in V$. Dann ist $g \circ f : U \to \mathbb{R}^k$ differenzierbar im Punkt x und es gilt $D(gf)(x) = (Dg)(f(x)) \cdot (Df)(x)$.*

Beweis. f ist total differenzierbar in x, daher gilt $f(x+\xi) = f(x) + A\xi + r_f(\xi)$, $A := (Df)(x)$, $\lim_{\xi\to 0} r(\xi)/\|\xi\| = 0$. Analoges gilt für g: $g(y+\eta) = g(y)+B\eta+r_g(\eta)$, $B := (Dg)(y)$, $\lim_{\eta\to 0} r_g(\eta)/\|\eta\| = 0$. Betrachte $\eta = f(x+\xi) - f(x) = f(x+\xi) - y$. Dann ist

$$\begin{aligned} g(f(x+\xi)) &= g(y+\eta) = g(f(x)) + B(f(x+\xi) - y) + r_g(f(x+\xi) - y) \\ &= g(f(x)) + B(A\xi + r_f(\xi)) + r_g(A\xi + r_f(\xi)) \\ &= g(f(x)) + BA\xi + (Br_f(\xi) + r_g(A\xi + r_f(\xi))) =: g(f(x)) + BA\xi + r_{fg}(\xi) \end{aligned}$$

Es bleibt zu zeigen, dass $r_{fg}(\xi)/\|\xi\| = 0$ für $\xi \to 0$.

$$\lim_{\xi\to 0} \frac{\|Br_f(\xi)\|}{\|\xi\|} \leq \lim_{\xi\to 0} \frac{\|B\|\|r_f(\xi)\|}{\|\xi\|} = 0$$

wegen $\lim_{\xi\to 0} r_f(\xi)/\|\xi\| = 0$. Außerdem könnn wir schreiben $r_g(\eta) = \eta \cdot R(\eta)$ mit $R(\eta) \to 0$ für $\eta \to 0$ wegen der Bedinung für r_g. Also ist

$$\begin{aligned} r_g(A\xi + r_f(\xi)) &= (A\xi + r_f(\xi)) \cdot R(A\xi + r_f(\xi)) \\ &\leq (\|A\| + 1)\|\xi\| \cdot R(A\xi + r_f(\xi)) \quad \text{für } \xi \text{ hinreichend klein} \end{aligned}$$

Und da $A\xi + r_f(\xi)$ gegen 0 geht für $\xi \to 0$, muss auch $R(A\xi + r_f(\xi)) \to 0$. Somit ist

$$\frac{\|r_{gf}(\xi)\|}{\|\xi\|} \leq \frac{\|B\|\|r_f(\xi)\|}{\|\xi\|} + (\|A\| + 1)R(A\xi + r_f(\xi)) \to 0 \quad \text{für} \quad \xi \to 0$$

□

Beispiel 3.5.2. Sei $\gamma : \mathbb{R} \to \mathbb{R}^m$ eine Kurve, $f : \mathbb{R}^m \to \mathbb{R}$ eine Funktion. Mit Satz 3.5.1 ist

$$(f \circ \gamma)'(t) = (Df)(\gamma(t)) \cdot (D\gamma)(t) = (\boldsymbol{\nabla} f)(\gamma(t)) \cdot \gamma'(t) = \langle (\boldsymbol{\nabla} f)(\gamma(t)),\, \gamma'(t) \rangle$$

Vorlesung 8

Beispiel 3.5.3. Sei $f : \mathbb{R}^n \to \mathbb{R}$, $\gamma : \mathbb{R} \to \mathbb{R}^m$. Dann ist

$$\begin{aligned} D(\gamma \circ f)(x) &= \begin{pmatrix} \gamma_1'(f(x)) \\ \vdots \\ \gamma_m'(f(x)) \end{pmatrix} \cdot ((\partial_1 f)(x), \dots, (\partial_n f)(x)) \\ &= \begin{pmatrix} \gamma_1'(f(x))(\partial_1 f)(x) & \dots & \gamma_1'(f(x))(\partial_n f)(x) \\ \vdots & \ddots & \vdots \\ \gamma_m'(f(x))(\partial_1 f)(x) & \dots & \gamma_m'(f(x))(\partial_n f)(x) \end{pmatrix} \end{aligned}$$

3.6 Richtungsableitungen

Bis jetzt haben wir in Richtung der Einheitsvektoren der kanonischen Basis von $\mathbb{R}^m$ abgeleitet. Nun wollen wir beliebige Richtungsableitungen definieren.

Definition 3.6.1. Sei $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, $x \in U$. Sei $v \in R^n$ ein Vektor mit $\|v\| = 1$ (v ist also ein "Richtungsvektor"). Die Ableitung von f **in Richtung** v im Punkt x ist

$$(\partial_v f)(x) := \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}$$

(falls der Limes exisitert).

Bemerkung 3.6.2 (Berechnung von Richtungsableitungen). Betrache $\gamma : (-\epsilon, \epsilon) \to U$, $\gamma(t) := x + t \cdot v$, $f : U \to \mathbb{R}$ für $\epsilon > 0$. Dann ist

$$\begin{aligned} (\partial_v f)(x) &= \frac{\mathrm{d}}{\mathrm{d}t} f(x + tv)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} (f \circ \gamma)(t)\Big|_{t=0} \\ &= \langle (\boldsymbol{\nabla} f)(\gamma(0),\, \gamma'(0)) \rangle = \langle (\boldsymbol{\nabla} f)(x), v \rangle \end{aligned}$$

Wir halten fest: $(\partial_v f)(x) = \langle (\boldsymbol{\nabla} f)(x), v \rangle$. Diese Formel verwendet man in der Praxis, um Richtungsableitungen zu berechnen.

4 Taylor-Entwicklungen

Sei $f : U \to \mathbb{R}$, $U \underset{\text{offen}}{\subset} \mathbb{R}^n$, $(k+1)$-mal stetig partiell differenzierbar. Sei $x \in U$, $v \in \mathbb{R}^n$ ein Richtungsvektor ($\|v\| = 1$), sodass die Strecke $\{x + tv | t \in [0, 1]\}$ ganz in U liegt. Wir setzen $g(t) := f(x + t \cdot v)$, $g : [0, 1] \to \mathbb{R}$. Der Satz von Taylor aus Analysis 1 mit Langrangeschem Restglied, angewendet bei

$t = 0$, liefert nun

$$g(t) = \sum_{m=0}^{k} \frac{1}{m!} g^{(m)}(0) \cdot t^m + \frac{1}{(k+1)!} g^{(k+1)}(\xi) t^{k+1} \quad \text{für ein} \quad \xi \in [0,t]$$

Für $t = 1$ gilt also

$$f(x+v) = g(1) = \sum_{m=0}^{k} \frac{1}{m!} g^{(m)}(0) + \frac{1}{(k+1)!} g^{(k+1)}(\xi),\ \xi \in [0,1]$$

Jetzt gilt es noch herauszufinden, welche Formen die $g^{(m)}$ haben.

$$\begin{aligned} g'(t) &= \langle (\boldsymbol{\nabla} f), v \rangle = \sum_{i=1}^{n} (\partial_i f)(x+tv) \cdot v_i \\ g''(t) &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{i=1}^{n} (\partial_i f)(x+tv) \cdot v_i \right) = \sum_{i=1}^{n} \frac{\mathrm{d}}{\mathrm{d}t} ((\partial_i f)(x+tv)) \cdot v_i \\ &= \sum_{i,j=1}^{n} (\partial_j \partial_i f)(x+tv) \cdot v_j \cdot v_i \end{aligned}$$

Mittels Induktion nach m erhält man:

$$\begin{aligned} g^{(m)}(t) &= \sum_{i_1,\dots,i_m}^{n} (\partial_{i_m} \dots \partial_{i_1} f)(x+tv) \cdot v_{i_m} \cdot \dots \cdot v_{i_1} \\ g^{(m)}(0) &= \sum_{i_1,\dots,i_m}^{n} (\partial_{i_m} \dots \partial_{i_1} f)(x) \cdot v_{i_m} \cdot \dots \cdot v_{i_1} \end{aligned}$$

Einsetzen liefert

$$f(x+v) = \sum_{m=0}^{k} \frac{1}{m!} \sum_{i_1,\dots,i_m}^{n} (\partial_{i_m} \dots \partial_{i_1} f)(x) \cdot v_{i_m} \cdot \dots \cdot v_{i_1} + \frac{1}{(k+1)!} \sum_{i_1,\dots,i_{k+1}=1}^{n} (\partial_{i_{k+1}} \dots \partial_i f)(x+\xi v)$$

für $\xi \in [0,1]$.

4.1 Einschub: Multinomischer Lehrsatz

Erinnern wir uns an den **Bi**nomischen Lehrsatz:

$$\begin{aligned} (\lambda_1 + \lambda_2)^m &= \sum_{\alpha=0}^{m} \left(\binom{m}{\alpha} \right) \lambda_1^{\alpha} \lambda_2^{m-\alpha} = \sum_{\alpha_1+\alpha_2=m} \frac{m!}{\alpha_1! \alpha_2!} \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \\ &= \sum_{i_1,\dots,i_m=1}^{2} \lambda_{i_1} \cdot \dots \cdot \lambda_{i_m} \end{aligned}$$

Wir versuchen nun, den binomischen Lehrsatz auf n Summanden zu verallgemeinern:

$$\begin{aligned}(\lambda_1 + \cdots + \lambda_n)^m &= ((\lambda_1 + \cdots + \lambda_{n-1}) + \lambda_n)^m \\ &= \sum_{\beta_1 + \alpha_n = m} \frac{m!}{\beta_1! \alpha_n!} (\lambda_1 + \cdots + \lambda_{n-1})^{\beta_1} \cdot \lambda_n^{\alpha_n}\end{aligned}$$

wir machen nun die (Induktions-)Annahme

$$(\lambda_1 + \cdots + \lambda_{n-1})^{\beta_1} = \sum_{\alpha_1 + \cdots + \alpha_{n-1} = \beta_1} \frac{\beta_1!}{\alpha_1! \ldots \alpha_{n-1}!} \lambda_1^{\alpha_1} \ldots \lambda_{n-1}^{\alpha_{n-1}}$$

Dann wird der Ausdruck $(\lambda_1 + \cdots + \lambda_n)^m$ zu

Satz 4.1.1 (Binomischer Lehrsatz)**.**

$$(\lambda_1 + \cdots + \lambda_n)^m = \sum_{\alpha_1 + \cdots + \alpha_n = m} \frac{m!}{\alpha_1! \ldots \alpha_n!} \lambda_1^{\alpha_1} \ldots \lambda_n^{\alpha_n}$$

*Dies ist der **Multionomische Lehrsatz**, den wir soeben per Induktion bewiesen haben.*

Beweis durch Induktion. Siehe oben. □

Bemerkung 4.1.2. Eine kompaktere Schreibweise lautet

$$\begin{aligned}\alpha &:= (\alpha_1, \ldots, \alpha_n) \\ |\alpha| &:= \alpha_1 + \cdots + \alpha_2 \\ \alpha! &:= \alpha_1! \ldots \alpha_n! \\ \lambda &= (\lambda_1, \ldots, \lambda_n) \\ \lambda^\alpha &:= \lambda_1^{\alpha_1} \ldots \lambda_n^{\alpha_n} \\ \partial^\alpha &:= \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n}\end{aligned}$$

Der multinomische Lehrsatz wird dann sehr kompakt:

$$(\lambda_1 + \cdots + \lambda_n) = \sum_{|\alpha| = m} \frac{m!}{\alpha!} \lambda^\alpha$$

Der Satz von Taylor wird dann auch sehr kompakt zu

$$\begin{aligned}f(x+v) &= \sum_{m=0}^{k} \frac{1}{m!} \sum_{|\alpha|=m} \frac{m!}{\alpha!} (\partial^\alpha f)(x) v^\alpha + \frac{1}{(k+1)!} \sum_{|\alpha|=k+1} (\partial^\alpha f)(x + \xi v) v^\alpha \\ &= \sum_{|\alpha| \leq k} \frac{1}{\alpha!} (\partial^\alpha f)(x) v^\alpha + \sum_{|\alpha|=k+1} \frac{1}{\alpha!} (\partial^\alpha f)(x + \xi v) v^\alpha \quad \text{für} \quad \xi \in [0, 1]\end{aligned}$$

$$\begin{aligned}
f(x+v) &= \sum_{|\alpha|\leq k-1} \frac{1}{\alpha!}(\partial^\alpha f)(x)v^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!}(\partial^\alpha f)(x+\xi v)v^\alpha \\
&= \sum_{|\alpha|\leq k} \frac{1}{\alpha!}(\partial^\alpha f)(x)v^\alpha - \sum_{|\alpha|=k} \frac{1}{\alpha!}(\partial^\alpha f)(x)v^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!}(\partial^\alpha f)(x+\xi v)v^\alpha \\
&= \sum_{|\alpha|\leq k} \frac{1}{\alpha!}(\partial^\alpha f)(x)v^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!}((\partial^\alpha f)(x+\xi v) - (\partial^\alpha f)(x)) \cdot v^\alpha
\end{aligned}$$

Aufgrund der Stetigkeit der $\partial^\alpha f$ geht für $v \to 0$ auch $x+\xi v \to x$, also auch $(\partial^\alpha f)(x+\xi v)-(\partial^\alpha f)(x) \to 0$. Nun ist $\forall\, i\ |v_i| \leq \|v\|$

$$|v|^\alpha = |v_1^{\alpha_1} \dots v_n^{\alpha_n}| \leq \|v\|^{\alpha_1} \dots \|v\|^{\alpha_n} = \|v\|^{|\alpha|} = \|v\|^k$$

Insbesondere ist $|v^\alpha|/\|v\|^k \leq 1$. Insbesondere geht für $v \to 0$ auch $((\partial^\alpha f)(x+\xi v) - (\partial^\alpha f)(x)) \cdot v^\alpha \to 0$. Wir haben also gezeigt:

Satz 4.1.3 (Taylor)**.**

$$f(x+v) = \sum_{|\alpha|\leq k} \frac{1}{\alpha!}(\partial^\alpha f)(x)v^\alpha + o\Big(\|v\|^k\Big)$$

Dies ist die Taylor-Entwicklung von f im Mittelpunkt x mit Lagrangschem Restglied der Ordnung $k+1$.

Beispiel 4.1.4. Für $k = 1$ ist

$$\begin{aligned}
f(x+v) &= f(x) + \sum_{i=1}^{n} (\partial_i f)(x) v_i + o(\|v\|) \\
f(x) &= \langle (\boldsymbol{\nabla} f(x))(x),\, v\rangle + +o(\|v\|)
\end{aligned}$$

denn dann ist $|\alpha| = 1$, und alle α haben die Form $\alpha = (0, \dots, 0, 1, 0, \dots, 0)$. Für $k = 2$ ist analog $|\alpha| = 2$. Die α sind dann von der Form $\alpha_1 = (0, \dots, 1, \dots, 1, \dots, 0)$, dann ist $\alpha! = 1$ und $\alpha_2 = (0, \dots, 2, \dots, 0)$ und $\alpha_2! = 2$. Wir können dann schreiben

$$\begin{aligned}
f(x+v) &= f(x) + \sum_{i=1}^{n} (\partial_i f)(x) v_i + \sum_{i<j} (\partial_i \partial_j f)(x) v_i v_j + \frac{1}{2}\sum_{i=1}^{n} (\partial_i^2 f)(x) v_i^2 + o\Big(\|v\|^2\Big) \\
&= f(x) + \sum_{i=1}^{n} (\partial_i f)(x) v_i + \frac{1}{2}\sum_{i\neq j} (\partial_i \partial_j f)(x) v_i v_j + \frac{1}{2}\sum_{i=1}^{n} (\partial_i^2 f)(x) v_i^2 + o\Big(\|v\|^2\Big) \quad \text{Satz von Schwarz} \\
&= f(x) + \sum_{i=1}^{n} (\partial_i f)(x) v_i + \frac{1}{2}\sum_{i,j=1}^{n} (\partial_j \partial_i f)(x) v_i v_j + o\Big(\|v\|^2\Big) \\
&= f(x) + \langle (\boldsymbol{\nabla} f)(x), v\rangle + \frac{1}{2}\,\langle v, H(f)\cdot v\rangle + o\Big(\|v\|^2\Big)
\end{aligned}$$

Wobei wir definieren $H(f)(x) = ((\partial_i \partial_j f)(x))_{i,j}$. $H(f)(x)$ heißt **Hesse-Matrix** von f im Punkt $x \in U$. Aus dem Satz von Schwarz (Satz 3.2.2) folgt, dass $H(f)(x)$ sogar eine **symmetrische** Matrix ist.

Vorlesung 9

5 Lokale Extrema

Wie für den Fall $f : \mathbb{R} \to \mathbb{R}$ wollen wir nun für allgemeinere Funktionen lokale Extrema betrachten.

Definition 5.0.1. Sei $f : U \to \mathbb{R}$ partiell differenzierbar, $U \underset{\text{offen}}{\subset} \mathbb{R}^n$. f besitzt ein lokales Maximum (bzw. Minimum) im Punkt $x \in U$, wenn $\exists$ Umbgebung $V \subset U$ von x mit

$$f(x) \geq f(y)\ \forall\, y \in V \quad \text{bzw.} \quad \leq$$

Definition 5.0.2. f hat ein **lokales Extremum in** x, wenn f in x ein lokales Minimum oder Maximum hat.

Wie im eindimensionalen Fall sind für die Extremwertbetrachtung die kritischen Punkte (im Eindimensionalen $f'(x) = 0$) von Interesse:

Definition 5.0.3. x heißt **kritischer Punkt von** f, wenn $(\boldsymbol{\nabla} f)(x) = 0$.

Satz 5.0.4 (Notwendige Bedingung für lokale Extrema)**.** *Besitzt f ein lokales Extremum in x, dann ist x ein kritischer Punkt von f*

Beweis. Setze $g_i(t) := f(x + te_i)$ wobei e_i der i-te kanonische Basisvektor ist. Hat f ein lokales Extremum in x, dann hat g_i ein lokales Extremum in $t = 0$. Aus Analysis 1 folgt dann $g_i'(0) = 0$. Andererseits ist $\partial_i f(x) = g_i'(0) = 0$. Also ist

$$(\boldsymbol{\nabla} f)(x) = ((\partial_1 f)(x), \ldots, (\partial_n f)(x)) = 0$$

Also ist x ein kritischer Punkt con f. □

Jedoch ist nicht jeder kritische Punkt ist notwendigerweise ein lokales Extremum, denn für $f(x) = x^3$ ist $x = 0$ zwar ein kritscher Punkt, jedoch nur ein Sattelpunkt und kein Extremum. Um eine hinreichende Bedingung für die Existenz lokaler Extrema zu erhalten, betrachten wir die partiellen Ableitung 2. Ordnung von f, d.h. die Hesse-Matrix. Wir machen einen kurzen Ausflug in die Lineare Algebra:

Definition 5.0.5. Sei H eine symmetrische, relle $n \times n$-Matrix (z.B. die Hesse-Matrix). H heißt **positiv definit**, wenn $\forall\, v \in \mathbb{R}^n - 0 :\ \langle v, Hv \rangle > 0$. Analog heißt H **negativ definit**, wenn $-H$ positiv definit ist. H heißt indefinit, wenn $\exists\, v, w \in \mathbb{R}^n - 0$, so dass $\langle v, Hv \rangle > 0$, $\langle w, Hw \rangle < 0$.

Bemerkung 5.0.6. Ist H symmetrisch, so exisitert eine Orthonormalbasis $\{v_1, \ldots, v_n\}$ von $\mathbb{R}^n$ bestehend aus Eigenvektoren, d.h. $Hv_i = \lambda_i v_i,\ \lambda_i \in \mathbb{R}$. Dann lässt sich jedes $v \in \mathbb{R}^n$ schreiben als $v = \sum_i \alpha_i v_i$ mit $\alpha_i \in \mathbb{R}$. Dann ist

$$\langle v,\, Hv \rangle = \left\langle \sum_i \alpha_i v_i,\, \sum_j \alpha_j H v_j \right\rangle = \sum_{i,j} \alpha_i \alpha_j \left\langle v_i,\, \lambda_j v_j \right\rangle = \sum_i \alpha_i^2 \lambda_j$$

Denn die v_i sind orthogonal. Also ist H positiv definit genau dann, wenn $\forall\, j\ \lambda_j > 0$. Analog ist H negativ definit $\Leftrightarrow \lambda_j < 0\, \forall j$ und H indefinit $\Leftrightarrow\ \exists i, j :\ \lambda_i > 0 \quad \text{und} \quad \lambda_j < 0$.

In der Praxis kann es sehr umständlich sein, erst alle Eigenvektoren zu bestimmen, um auf positive Definitheit zu prüfen.

Proposition 5.0.7 (Hauptminorenkriterium)**.** *H ist genau dann positiv definit, wenn*

$$\det \begin{pmatrix} h_{11} & \dots & h_{1k} \\ \vdots & \ddots & \vdots \\ h_{k1} & \dots & h_{kk} \end{pmatrix} > 0 \,\forall\, k = 1, \dots, n$$

Zurück zu $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ offen. Sei $x \in U$ ein kritischer Punkt von f.

Satz 5.0.8 (Hinreichende Bedingung für lokale Extrema)**.**

1. *Ist die Hesse-Matrix $H(f)(x)$ positiv definit, dann besitzt f in x ein lokales Minimum.*
2. *Ist die Hesse-Matrix negativ definit, dann besitzt f in x ein lokales Maximum.*
3. *Ist die Hesse-Matrix indefinit, dann besitzt f kein lokales Extremum in x.*

Beweis. Betrachte die Taylor-Approximation zweiter Ordnung von f in x:

$$f(x+v) = f(x) + \langle (\boldsymbol{\nabla} f)(x), v \rangle + \frac{1}{2} \langle v, H(f) \cdot v \rangle + o\Big(\|v\|^2 \Big)$$

Da x ein kritischer Punkt ist, ist $(\boldsymbol{\nabla} f)(x) = 0$ und die Approximation vereinfacht sich zu

$$f(x+v) = f(x) + \frac{1}{2} \langle v, H(f) \cdot v \rangle + r(v)$$

mit $r(v) = o\Big(\|v\|^2 \Big)$.
Wir zeigen 1.: Sei H positiv definit. Sei $S := \{ v \in \mathbb{R}^n \mid \|v\| = 1 \}$ die Einheitssphäre. S ist abgeschlossen und beschränkt, also nach Heine-Borel (Satz 2.0.7) kompakt. Es ist $v \mapsto \langle v, \, Hv \rangle$ auf S stetig, denn letztlich ist $\langle v, \, Hv \rangle$ nur ein Polynom von den Einträgen von v. Also existiert nach Satz **??** ein

$$\mu := \min \{ \langle v, Hv \rangle \mid v \in S \} > 0$$

Denn $\langle v, Hv \rangle > 0 \ \forall v \in \mathbb{R}^n$ nach Voraussetzung. Sei $\bar{v} = \lambda v$ beliebig, $\|v\| = 1$, $\lambda \in \mathbb{R}$, $\|\bar{v}\| = |\lambda|$. Dann ist

$$\begin{aligned} \langle \bar{v}, \, H\bar{v} \rangle &= \langle \lambda v, \, \lambda H v \rangle = \lambda^2 \langle v, Hv \rangle \\ &\geq \lambda^2 \mu = \mu \|\bar{v}\|^2 \end{aligned}$$

Nun gilt es noch, den Restterm abzuschätzen. Da $r(v) = o\Big(\|v\|^2 \Big)$ existiert ein $\delta > 0 : \ \|v\| < \delta \to \frac{r(v)}{\|v\|^2} \leq \frac{\mu}{4}$ Also ist

$$\begin{aligned} f(x+v) &= f(x) + \frac{1}{2} \langle v, \, Hv \rangle + r(v) \\ &\geq f(x) + \frac{\mu}{2} \|v\|^2 - \frac{\mu}{4} \|v\|^2 \\ &= f(x) + \frac{\mu}{4} \|v\|^2 \geq f(x) \ \forall \, v \in B_\delta(0) \end{aligned}$$

Also ist $f(x)$ ein lokales Minimum.
2.: Betrachte $-f$ und verwende 1.
3.: Sei H indefinit. Also $\exists\, v, w \in \mathbb{R}^n - 0$ mit $\lambda := \langle v,\, Hv\rangle > 0$ und $\langle w,\, Hw\rangle < 0$. O.B.d.A. können wir annehmen $\|v\| = 1 = \|w\|$. (Falls dies nicht der Fall ist, können wir v und w entsprechend normieren.) Dann ist

$$\begin{aligned} f(x+tv) &\geq f(x) + \frac{\lambda}{2}t^2 - \frac{\lambda}{2}t^2 \ \forall 0 < t < \delta' \\ &= f(x) + \frac{\lambda}{4}t^2 > f(x) \end{aligned}$$

für ein wie analog zu 1. gewähltes δ'. Also hat f kein lokales Maximum in x. Analog ist $f(x+tw) < f(x)\ \forall 0 < t < \delta''$. Also hat f auch kein lokales Minimum in x. □

Beispiel 5.0.9. Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = c + x^2 + y^2$ für $c \in \mathbb{R}$. Dann ist

$$(\boldsymbol{\nabla} f)(x,y) = (2x,\, 2y) \stackrel{!}{=} (0,\, 0) \Rightarrow\ x = 0,\, y = 0$$

Also ist $(0,0) \in \mathbb{R}^2$ der einzige kritische Punkt von f. Es ist

$$H(f) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Offenbar ist $H(f)$ (in jedem Punkt) positiv definit, also auch im kritischen Punkt $(0,0)$. Nach Satz 5.0.8 hat f ein lokales Minimum in $(0,0)$.
Wir können obiges Beispiel auch abändern zu $f(x,y) = c - x^2 - y^2$. Wieder ist $(0,0)$ ein kritischer Punkt, diesmal ist $H(f) = -2\mathbb{I}_2$. Also hat f in $(0,0)$ ein lokales Maximum. Für $f(x,y) = c + x^2 - y^2$ ist wieder $(0,0)$ ein kritischer Punkt, aber wegen

$$H(f) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

ist $H(f)$ indefinit. f hat in $(0,0)$ also kein Extremum. Wir nennen $(0,0)$ dann einen **Sattelpunkt**.

Bemerkung 5.0.10. Ist $H(f)(x)$ weder positiv definit, noch negativ definit, noch indefinit, dann muss man Ableitungen höherer Ordnung betrachten. (Dies kann zum Beispiel eintreten, wenn H nur positive Eigenwerte plus zusätzlich den Eigenwert 0 hat.) Konkret hat $f_1(x,y) = x^2 + y^4$ einen kritischen Punkt bei $(0,0)$, aber $H(f_1)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Analog ist für $f_2(x,y) = x^2 + y^3$ auch $(0,0)$ ein kritischer Punkt und $H(f_2)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Allerdings hat f_1 ein lokales Minimum bei $(0,0)$, f_2 jedoch kein Extremum bei $(0,0)$.

Vorlesung 10

6 Implizite Funktionen

6.1 Einführung

Definition 6.1.1. Sei $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ eine stetig partiell differenzierbare Abbildung. Wir werden häufig $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ schreiben, wobei x die ersten n Variablen enthält, und y die letzten m Variablen. Angenommen, für jedes $x \in \mathbb{R}^n$ existiert genau ein $y \in \mathbb{R}^m$, sodass $F(x, y) = 0$. Dann definiert $g(x) := y$ eine Abbildung $g : \mathbb{R}^n \to \mathbb{R}^m$ für die gilt $F(x, g(x)) = 0 \;\forall x$. Wir sagen dann: Die Funktion g ist **implizit** durch die Gleichung $F(x, y) = 0$ bestimmt.

Beispiel 6.1.2. Sei $A : \mathbb{R}^n \to \mathbb{R}^m$ und $B : \mathbb{R}^m \to \mathbb{R}^n$ lineare Abbildungen [2]. Sei $F(x, y) := Ax - By$. Also ist $F(x, y) = 0$ äquivalent zu $Ax = By$. Dies legt die Invertierbarkeit von B als Voraussetzung nahe. In diesem Fall gilt nämlich $y = B^{-1}Ax = g(x)$. Nun ist

$$\frac{\partial F}{\partial y} := \frac{\partial(F_1, \ldots, F_m)}{(y_1, \ldots, y_m)} = \left(\frac{\partial F_i}{\partial y_j}\right)_{i,j}$$

eine $m \times m$-Matrix, nämlich $-B = \frac{\partial F}{\partial y}$. Dies suggeriert die Invertierbarkeit der Matrix $\frac{\partial F}{\partial y}$.

Die Lösbarkeit des Problems hängt auch von der Wahl eines Punktes $(a, b) \in \mathbb{R}^n \times \mathbb{R}^m$ mit $F(a, b) = 0$, wie etwa in diesem Beispiel:

Beispiel 6.1.3. Sei $F(x, y) := x^2 + y^2 - 1$, $x \in \mathbb{R}$, $y \in \mathbb{R}$. Geometrisch sit die Lösungsmenge $\left\{(x, y) \in \mathbb{R}^2 \,|\, F(x, y) = 0\right\}$ der Einheitskreis $\left\{x^2 + y^2 = 1\right\}$. Es gelte $F(a, b) = 0$.

1. $-1 < a < 1$. Dann exisitert eine Umgebung V_1 von a in $\mathbb{R}$ und eine eindeutige, stetige Funktion $g : V_1 \to \mathbb{R}$ mit $g(a) = b$ und $F(x, g(x)) = 0 \;\forall x \in V_1$. Denn $y^2 = 1 - x^2$, $y = \pm\sqrt{1 - x^2}$, es gibt also zwei Lösungen; $g(a) = b$ legt eindeutig den Zweig der Wurzelfunktion fest, der für g zu wählen ist. Um noch einmal auf die Jacobi-Matrix zurückzukommen: $\frac{\partial F}{\partial y} = 2y$. Wenn $a \in (-1, 1)$, dann ist $\frac{\partial F}{\partial y} \neq 0$, d.h $\frac{\partial F}{\partial y}$ ist invertierbar.
2. $(a, b) = (\pm 1, 0)$. Es gilt $F(a, b) = 0$. Aber $\frac{\partial F}{\partial y} = 0$ ist nicht invertierbar. Es existiert nun keine Umgebung von a, so dass wir nicht implizit einen Zweig der Wurzel gewählt haben, wodurch sich g eindeutig definieren ließe.

Wir bereiten den Beweis des impliziten Funktionensatzes durch Resultate über Banachräume vor, insbesondere durch den Fixpunktsatz. Zur Erinnerung: Ein Banachraum ist ein vollständiger, normierter Vektorraum $(V, \|\cdot\|)$.

Satz 6.1.4 (Banachscher Fixpunktsatz)**.** *Sei $(V, \|\cdot\|)$ ein Banachraum und $\emptyset \neq A \subset V$ eine abgeschlossene Teilmenge. Sei $\phi : A \to A$ eine sogenannte* ***Kontraktion****, d.h. es existiert eine Konstante $0 < \lambda < 1$ mit*

$$\|\phi(f) - \phi(g)\| \leq \lambda \|f - g\| \;\forall f, g \in A$$

Dann hat ϕ einen Fixpunkt in A, also einen Punkt f_0 mit $\phi(f_0) = f_0$.

[2] Wir schreiben wieder $Ax = A(x)$, $By = B(y)$.

Beweis. Wähle einen Punkt[3] $f^0 \in A \neq \emptyset$, $f^1 := \phi(f^0) \in A$. Induktiv definieren wir $f^k := \phi(f^{k-1}) \in A$. Wir erhalten so eine Folge $(f^k)_k \subset A$. Wir behaupten nun, dass es sich dabei um eine Cauchyfolge handelt.

$$\left\|f^{k+1} - f^k\right\| = \left\|\phi(f^k) - \phi(f^{k+1})\right\| \leq \lambda\left\|f^k - f^{k-1}\right\| \leq \cdots \leq \lambda^k \cdot \left\|f^1 - f^0\right\|$$

Sei $C := \left\|f^1 - f^0\right\|$. Wir können dann schreiben

$$\begin{aligned}\left\|f^l - f^k\right\| &= \left\|\sum_{j=k}^{l-1}(f^{j+1} - f^j)\right\| \leq \sum_{j=k}^{l-1} \left\|f^{j+1} - f^j\right\| \\ &\leq C\sum_{j=k}^{\infty} \lambda^j = C\lambda^k \sum_{j=0}^{\infty} \lambda^j = \frac{C\lambda^k}{1-\lambda}\end{aligned}$$

Und $\lim_{k\to\infty} \frac{C\lambda^k}{1-\lambda} = 0$. Also ist $(f^k)_k$ eine Cauchyfolge in V. Da V vollständig ist, konvergiert f^k in V gegen einen Grenzwert $f_0 \in V$. Da A abgeschlossen ist, liegt f_0 sogar in A. Es bleibt zu zeigen, dass f_0 ein Fixpunkt ist. Es ist

$$f_0 = \lim_{k\to\infty} f^k$$

Kontraktionen sind per Definition stetig (tatsächlich sogar Lipschitz-stetig). Also ist

$$\phi(f_0) = \lim_{k\to\infty} \phi(f^k) = \lim_{k\to\infty} f^{k+1} = f_0$$

f_0 ist eindeutig, denn angenommen, f_0, f_1 seien beide Fixpunkte von ϕ. Dann ist

$$\|f_0 - f_1\| = \|\phi(f_0) - \phi(f_1)\| \leq \lambda\|f_0 - f_1\|$$

Dies kann nur der Fall sein, wenn $\|f_0 - f_1\| = 0$, also $f_0 = f_1$. □

Sei $U \subset \mathbb{R}^n$ beliebig. Wir setzen[4]

$$\mathcal{C}_b(U, \mathbb{R}^m) := \{f : U \to \mathbb{R}^m \mid \quad \text{f stetig und } \textbf{b}\text{eschränkt} \quad\}$$

Wir statten $\mathcal{C}_b(U, \mathbb{R}^m)$ mit der Supremumsnorm $\|\cdot\|_\infty$ aus:

$$\|f\|_\infty := \sup\{\|f(x)\| \mid x \in U\} \in \mathbb{R}$$

Satz 6.1.5. *$(\mathcal{C}_b(U, \mathbb{R}^m), \|\cdot\|_\infty)$ ist ein Banachraum.*

Beweis. Dies ist ein normierter Vektorraum. Wir zeigen, dass dieser Raum vollständig ist. Sei (f_n) eine Cauchyfolge in $\mathcal{C}_b(U, \mathbb{R}^m)$. Sei $\epsilon > 0$. Dann existiert ein $N = N(\epsilon)$, so dass $\|f_k - f_l\|_\infty < \epsilon \; \forall \, k, l \geq N$.

[3] Die Superskripte werden hier nur als Indizes verwendet, nicht als Ableitungen oder Potenzen (deren Existenz nicht zwingend gegeben ist).

[4] Das $\mathcal{C}$ steht für "continuous", also "stetig".

Es ist dann $\forall x \in U$

$$\|f_k(x) - f_l(x)\|_{\mathbb{R}^m} \leq \|f_k - f_k\|_\infty < \epsilon$$

Also ist $(f_k(x))_k$ eine Cauchyfolge in $\mathbb{R}^m$. Da $\mathbb{R}^m$ vollständig ist, konvergiert $(f_k(x))_k$ für ein festes x und $k \to \infty$ gegen einen Grenzwert $f(x) \in R^m$. Sei $\epsilon' > 0$. Nun existiert ein M so dass $\|f_l(x) - f(x)\| < \epsilon'$ für alle $l \geq M$. Wähle $l \geq \max\{M, N\}$. Dann ist

$$\|f_k(x) - f(x)\| \leq \|f_k(x) - f_l(x)\| + \|f_l(x) - f(x)\| < \epsilon + \epsilon'$$

Da aber $\epsilon' > 0$ beliebig war, ist

$$\|f_k(x) - f(x)\| \leq \epsilon \; \forall \, k \geq N, \; \forall \, x \in U$$

Also ist f beschränkt und $\|f_k - f\|_\infty \to 0$. Da die f_k alle stetig sind und gleichmäßig gegen f konvergieren, ist auch f stetig. Somit liegt f in $\mathcal{C}_b(U, \mathbb{R}^m)$ und f_k konvergiert gegen f in $\|\cdot\|_\infty$. □

Zurück zum impliziten Funktionensatz.

Satz 6.1.6 (über implizite Funktionen)**.** *Seien $U_1 \subset \mathbb{R}^n$ und $U_2 \subset \mathbb{R}^m$ offen und $F : U_1 \times U_2 \to \mathbb{R}^m$ eine stetig partiell differenzierbare Abbildung. Sei $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ ein Punkt mit $F(a, b) = 0$. Wenn $\frac{\partial F}{\partial y} \in \mathbb{R}^{m \times m}$ invertierbar ist, dann existiert eine Funktion $g : V_1 \to V_2$, wobei V_1 eine Umgebung von a in U_1 und V_2 eine Umgebung von b in U_2 ist, sodass g eindeutig durch die folgenden Eigenschaften bestimmt ist:*

1. *$g(a) = b$*
2. *$F(x, g(x)) = 0 \; \forall x \in V_1$*
3. *g ist stetig partiell differenzierbar.*

Beweis. Hier zunächst nur eine Beweisskizze:

1. Wir transformieren das Problem als Fixpunktproblem auf einem Banachraum.
2. Wir lösen das Fixpunktproblem mithilfe des Banachschen Fixpunktsatzes.
3. Wir zeigen: g ist stetig
4. Wir zeigen: g ist differenzierbar

□

Vorlesung 11

1 Implizite Funktionen

Bemerkung 1. Vorab eine kurze Notiz zur Notation. Sei $A(t) = (a_{i,j})(t)_{i,j}$ eine Matrix mit funktionalen Einträgen: $a_{i,j} : [t_o, t_1] \to \mathbb{R}$. Wir schreiben dann

$$\int_{t_o}^{t_1} A(t)\mathrm{d}t := (s_{i,j})_{i,j}$$

Wobei

$$s_{i,j} = \int_{t_o}^{t_1} a_{i,j}(t)\mathrm{d}t$$

Satz 1.1: Verallgemeinerter Mittelwertsatz

Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ stetig partiell differenzierbar, $x \in U$. Sei $v \in \mathbb{R}^n$ ein Vektor mit $\{x + tv | \, t \in [0,1]\} \subset U$. Dann gilt

$$f(x+v) - f(x) = \int_0^1 (Df)(x+tv)\mathrm{d}t \cdot v$$

Beweis. Zu zeigen ist:

$$f_i(x+v) - f_i(x) = \sum_j \int_0^1 (\partial_i f_i)(x+tv)\mathrm{d}t v_j$$

Wir setzen $g_i(t) := f_i(x+tv)$. Dann ist

$$f_i(x+v) - f_i(x) = g_i(t) - g_i(0) = \int_0^1 g_i'(t)dt = \int_0^1 \sum_j (\partial_j f_i)(x+tv)\mathrm{d}t \cdot v_j$$

Wobei wir im vorletzten Schritt den Hauptsatz der Differential- und Integralrechnung verwendet haben. □

Korollar 1.2

Sei

$$M := \sup\left\{\left\|(Df)(x+tv)\right\|, t \in [0,1]\right\}$$

Dann gilt:

$$\left\|f(x+v) - f(x)\right\| = \left\|\int_0^1 (Df)(x+tv)\mathrm{d}t \cdot v\right\| \leq M\|v\|$$

Mit der Cauchy-Schwarz-Ungleichung.

Zurück zum impliziten Funktionensatz:

Satz 1.3: Implizite Funktionen

Sei $F: U_1 \times U_2 \to \mathbb{R}^m$ eine stetige partiell differenzierbare Abbildung, $U_1 \subset \mathbb{R}^m$ offen, $U_2 \subset \mathbb{R}^m$ offen. Sei $(a,b) \in U_1 \times U_2$ ein Punkt mit $F(a,b) = 0$. Ist $\frac{\partial F}{\partial y}(a,b)$ invertierbar, dann $\exists$ Umgebung $V_1 \subset U_1$, von a und existiert eine Umgebung $V_2 \subset U_2$ von b und ein eindeutiges $g: V_1 \to V_2$, sodass gilt:

1. $g(a) = b$
2. $F(x, g(x)) = 0 \; \forall x \in V_1$
3. g ist stetig partiell differenzierbar

Beweis. O.B.d.A dürfen wir annehmen, dass $(a,b) = (0,0)$. Ansonsten können wir definieren $\tilde{F}(x,y) := F(x+a, y+b)$ und dann gilt per Konstruktion $\tilde{F}(0,0) = 0$. Angenommen, es gibt nun ein $\tilde{g}$, das die im Satz geforderten Eigenschaften hat, d.h. $\tilde{g}(0) = 0$ und $\tilde{F}(x, \tilde{g}(x)) = 0$. Dann ist für $g(x) := \tilde{g}(x-a) + b$, und es gilt $g(a) = \tilde{g}(a) + b = b$. Also folgt

$$F(x, g(x)) = F((x-a)+a, \tilde{g}(x-a)+b) = \tilde{F}(x-a, \tilde{g}(x-a)) = 0$$

Wenn wir also für $\tilde{F}$ den Satz bewiesen haben, folgt dessen Gültigkeit automatisch auch für F. Zunächst transformieren wir das Problem in ein Fixpunktproblem auf Banachräumen: Sei

$$G(x,y) := y - B^{-1}F(x,y) \quad \text{mit} \quad B = \frac{\partial F}{\partial y}(0,0)$$

Dann ist

$$\frac{\partial G}{\partial y}(x,y) = \mathbb{I}_{m\times m} - B^{-1}\frac{\partial F}{\partial y}(x,y) \tag{1}$$

$$\frac{\partial G}{\partial y}(0,0) = \mathbb{I}_{m\times m} - B^{-1} \cdot B = 0 \tag{2}$$

Nun ist $\frac{\partial G}{\partial y}$ stetig, denn B^{-1} ist als Matrix aus $\mathbb{R}^{m\times m}$ stetig, und $\frac{\partial F}{\partial y}$ nach Voraussetzung stetig. Also existiert eine Umgebung $W_1 \subset U_1$ von 0 und eine Umgebung $W_2 \subset U_2$ von 0, so dass

$$\left\| \frac{\partial G}{\partial y}(x,y) \right\| \leq \frac{1}{2} \quad \forall x \in W_1, y \in W_2$$

Sei $r > 0$ so gewählt, dass $V_2 := \{y \in \mathbb{R}^m, \|y\| \leq r\} \subset W_2$. Nun ist V_2 abgeschlossen in $\mathbb{R}^m$. Außerdem ist $G(0,0) = 0 - B^{-1}F(0,0) = 0$. Da G stetig ist, existiert nun eine Umgebung $V_1 \subset W_1$ von 0 mit $\|G(x,0)\| \leq \frac{r}{2} \; \forall x \in V_1$. Es gilt:

$$F(x,y) = 0 \iff G(x,y) = y$$

Was ein Fixpunktproblem für ein festes x ist.

Aus dem verallgemeinerten Mittelwertsatz folgt nun

$$\left\|G(x,y)-G(x,y_0)\right\| \leq \frac{1}{2}\left\|y-y_0\right\| \ \forall x \in V_1, \forall y, y_0 \in V_2 \quad (*)$$

Für $y_0 = 0$ folgt insbesondere

$$\left\|G(x,y)\right\| \leq \left\|G(x,y)-G(x,0)\right\| + \|G(x,0)\| \leq \frac{1}{2}\left\|y\right\| + \frac{r}{2} \leq \frac{r}{2} + \frac{r}{2} = r \quad \forall x \in V_1, \forall y \in V_2 \quad (**)$$

Mit anderen Worten: Ist x fest in V_1 so ist die Abbildung $\varphi : A \to A,\ y \mapsto G(x,y))$ für $A := V_2$ ist eine Kontraktion wegen (*) ($\lambda = \frac{1}{2} < 1$)
Nun können wir den Banchschen Fixpunktsatz anwenden. Dazu erinnern wir uns, dass $\mathbb{R}^m$ ein Banachraum ist. Also existiert ein eindeutiges $y_0 \in A : G(x,y) = y_0$. Setzte nun: $g(x) = y_0$. So erhalten wir eine Funktion $g : V_1 \to V_2$ mit $g(0) = 0$ und $G(x,y_0) = y_0$ Also auch $F(x,y_0) = 0 = F(x,g(x))\ \forall x \in V_1$ Es bleibt noch zu ziegen, dass g stetig partiell differenzierbar ist. Wir beginnen mit der Stetigkeit.

Betrachte den Funktionenraum $V = \mathscr{C}_b(V_1,\mathbb{R}^m)$. Dieser ist zusammen mit $\|\cdot\|_\infty$ ein Banachraum. Wir definieren eine Abbildung

$$\begin{aligned} \phi : V &\to V \\ f &\mapsto \phi(f)(x) = G(x,f(x)) \end{aligned}$$

Diese ist steig und beschränkt, denn f ist per Definition stetig und beschränkt und G ist ebenfalls stetig. Sei

$$A := \left\{f \in V : \|\cdot\|_\infty \leq r\right\}$$

ist abgeschlossen in V. Sei $f \in A$. Dann ist $\left\|f(x)\right\|_{\mathbb{R}^m} \leq r \forall x \in V_1$ Also ist $f(x) \in V_2$. Insbesondere ist

$$\left\|G(x,f(x))\right\|_{\mathbb{R}^m} \leq r \quad \text{wegen } (**)$$

Also ist $\left\|\phi(f)\right\|_\infty \leq r$, d.h. $\phi(f) \in A$. Also ist ϕ eine Abbildung $\phi : A \to A$ Nun können wir schreiben

$$\begin{aligned} \left\|\phi(f_1)(x)-\phi(f_2)(x)\right\|_{\mathbb{R}^m} &= \left\|G(x,f_1(x))-G(x,f_2(x))\right\|_{\mathbb{R}^m} \\ &\leq \frac{1}{2}\left\|f_1(x)-f_2(x)\right\|_{\mathbb{R}^m} \leq \frac{1}{2}\left\|f_1-f_2\right\|_\infty \end{aligned}$$

Also ist

$$\left\|\phi(f_1)-\phi(f_2)\right\|_\infty \leq \frac{1}{2}\left\|f_1-f_2\right\|_\infty$$

Somit ist $\phi : A \to A$ eine Kontraktion. Nach dem Banachschen Fixpunktsatz existiert dann ein Fixpunkt $f_0 \in A$, so dass $\phi(f_0) = f_0$. Aufgrund der Eindeutigkeit des Fixpunktes ist dann bereits $f_0 = g$,

und g ist stetig, da f_0 stetig ist.
Nun zeigen wir die (totale) Differenzierbarkeit von g in 0. Sei dazu $A := \frac{\partial F}{\partial x}(0,0)$, dann ist $(DF)(0,0) = (A|B)$
DieTaylor Approximation von F erster Ordnung wir dann zu

$$\begin{aligned}
F(x,y) = F(0,0) + (Df)\begin{pmatrix} x \\ y \end{pmatrix} + r(x,y) \quad \text{mit} \quad r = o(\|(x,y)\|) \\
F(x,y) = (A|B)\begin{pmatrix} x \\ y \end{pmatrix} + r = Ax + By + r(x,y) \\
0 = F(x,g(x)) = Ax + Bg(x) + r(x,g(x)) \\
Bg(x) = -Ax - r(x,g(x)) \\
g(x) = -B^{-1}Ax - \underbrace{B^{-1}r(x,g(x))}_{=:R(x)}
\end{aligned}$$

Zu zeigen ist noch $R(x) = o(\|x\|)$. Angenommen es exisitert eine Umgebung $V' \subset V$ von 0 und eine Konstante K mit $\|g(x)\| \leq K\|x\| \; \forall x \in V'$. Dann sind wir fertig, denn:

$$\begin{aligned}
\frac{\|R(x)\|}{\|x\|} &= \frac{\|B^{-1}r(x,g(x))\|}{\|x\|} \\
&\leq \frac{\|B\|^{-1}\|(x,g(x))\|\|R(x,g(x))\|}{\|x\|\|(x,g(x))\|} \\
&\leq \frac{\|B^{-1}\|(\|x\| + \|g(x)\|)}{\|x\|} \cdot o(1) \\
&\leq \frac{\|B^{-1}\|\|x\|(1+K)}{\|x\|} \cdot o(1) \\
&\leq \|B^{-1}\| \cdot (K+1) \cdot o(1) \to 0 \\
\Rightarrow R(x) = o(\|x\|)
\end{aligned}$$

Es ist noch zu zeigen, dass eine solche Konstante auch existiert. Es ist

$$\|g(x)\| \leq \|B^{-1}\|\|A\|\|x\| + \|B^{-1}\|\|r(x,g(x)\|$$

Da $r(x,y) = o(\|x,y\|)$ können wir ein ϵ wählen, so dass

$$\begin{aligned}
\|g(x)\| &\leq \|B^{-1}\|\|A\|\|x\| + \|B^{-1}\|\epsilon\|(x,g(x))\| \\
&\leq \|B^{-1}\|\|A\|\|x\| + \|B^{-1}\|\epsilon\big(\|x\| + \|g(x)\|\big)
\end{aligned}$$

Folglich

$$\|g(x)\| - \|B^{-1}\| \cdot \epsilon \cdot \|g(x)\| \leq \|B^{-1}\|\|A\|\|x\| + \|B^{-1}\|\epsilon\|x\|$$

Wenn wir x als hinreichend klein fordern, können wir wählen $\epsilon := \frac{1}{2\|B^{-1}\|}$. Dann ist

$$\Rightarrow \frac{1}{2}\|g(x)\| \leq \|B^{-1}\| \cdot \|A\| \cdot \|x\| + \frac{1}{2}\|x\|$$
$$\|g(x)\| \leq \underbrace{2\left(\|B^{-1}\| \cdot \|A\| + 1\right)}_{:=K} \|x\|$$

Zuletzt zeigen wir, dass g auch stetig partiell differenzierbar ist.

$$\frac{\partial}{\partial x} F(x, g(x)) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{\mathrm{d}g}{\mathrm{d}x} = 0$$
$$\frac{\mathrm{d}g}{\mathrm{d}x} = -\left(\frac{\partial F}{\partial y}^{-1} \frac{\partial F}{\partial x}\right)$$

Da die beiden Terme rechts stetig sind, folgt auch, dass die partielle Ableitung $\frac{\partial g}{\partial x}$ stetig ist. Für dieses Argument benötigen wir die totale Differenzierbarkeit, denn sonst wäre nichteinmal sichergestellt, dass die partiellen Ableitungen existieren. □

Vorlesung 12

Vorlesung 12

Definition 1.4

Sei $U \subset \mathbb{R}^n$ offen, $r = 0, 1, 2, \dots$. Eine Abbildung $f : U \to \mathbb{R}^m$ heißt C^r**-Abbildung**, wenn alle partiellen Ableitungen von f der Ordnung $\leq r$ exisitieren und stetig sind.
Ist f eine C^r-Abbildung für alle r, dann heißt f C^∞-Abbildung. Wir nennen C^∞ -Abbildungen auch **glatt**.

Bemerkung 2. D.h. $f \in C^0$ bedeutet, dass f stetig ist. Für $f \in C^1$ ist f stetig partiell differenzierbar.

Im weiteren wollen wir den äquidimensionalen Fall $(m = n)$ betrachten; $U \subset \mathbb{R}^n$ offen

Definition 1.5: Diffeomorphismus

Eine Bijektion $f : U \to V$ heißt C^r**-Diffeomorphismus**, wenn f und f^{-1} C^r-Abbildungen sind. Da $r \geq 0$, verlangen wir also stets die Stetigkeit von f und f^{-1}, d.h. f ist bereits ein Homöomorphismus.

Definition 1.6: Reguläre Punkte

Sei $U \subset \mathbb{R}^n$, $V \underset{\text{offen}}{\subset} \mathbb{R}^n$, $f : U \xrightarrow{C^1} V$. Sei $p \in U$ ein Punkt im Definitionsbereich von f. Wir bezeichen p als **regulären** Punkt der Abbildung f, wenn $(Df)(p)$ invertierbar ist (d.h. vollen Rang hat). (Also: p ist ein regulärer Punkt $\Rightarrow \det\big(Df(p)\big) \neq 0$)

Bemerkung 3. Sei $f: U \xrightarrow{C^1} V$ ein Diffeomorphismus und $p \in U$. Nun ist $f^{-1} \circ f = id$, also ist

$$D(f^{-1} \circ f) = D(id) = \mathbf{I}_{n \times n} = (Df^{-1})(f(p)) \circ (Df)(p) \quad \text{mit der Kettenregel}$$

Also ist $(Df)(p)$ invertierbar, d.h. p ist ein regulärer Punkt von f. Somit ist jeder Punkt im Definitionsbereich eines Diffeomorphismus ein regulärer Punkt. Die Umkehrung gilt jedoch nicht.

$$p \text{ regulär} \quad \forall p \in U \not\Rightarrow f \textit{ Diffeomorphismus}$$
$$(\det(Df)(p) \neq 0) (f \textit{ ist im Allgemeinen nicht global injektiv})$$

Lokal gilt die Umkehrung jedoch. Dies führt uns zum Satz über umkehrbare Funktionen

Satz 1.7: Umkehrbare Funktionen

Sei $f: U \to V$ eine C^1-Abbildung mit $U, V \underset{\text{offen}}{\subset} \mathbb{R}^n$. Sei $p \in U$ ein **regulärer Punkt** in f. Dann ist f ein lokaler Diffeomorphismus um den Punkt p.

Definition 1.8: lokale Diffeomorphismen

$f: U \to V$ heißt lokaler Diffeomorphismus um den Punkt $p \in U$, wenn eine offene Umgebung $U' \subset U$ von p existiert, sodass $f|_{U'}: U' \to f(U')$ ein (C^1-) Diffeomorphismus ist.

Beweis. (Zum Satz über umkehrbare Funktionen) Wir folgern dies aus dem Satz über implizite Funktionen. Wir setzen

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

Es sei $a := f(b)$, wobei $b = p$ der reguläre Punkt ist. Dann ist

$$F(a, b) = a - f(b) = f(b) - f(b) = 0$$

$$\frac{\partial F}{\partial y} = -(Df)(b)$$

ist invertierbar, da b ein regulärer Punkt ist. Somit sind die Vorraussetzugen zur Anwendung des Satztes über impliziete Funktionen erfüllt. Also

$$\exists \quad \text{offene Umgebung} \quad V_1 \subset \mathbb{R}^n \textit{ von } a, V_2 \subset U \textit{ von } b \tag{3}$$
$$\tag{4}$$

und es existiert eine C^1-Funktion $g: V_1 \to V_2$, sodass $F(x, g(x)) = 0$ für alle $x \in V_1$. Nun ist

$$F(x, g(x)) = x - f(g(x)) \quad \text{bzw.} \quad f(g(x)) = x \; \forall x \in V_1 \quad (*)$$

Da f stetig in b ist, exisitert eine offene Umgebung $V \subset V_2$ von b, sodass $f(V) \subset V_1$. Wir behaupten nun $f(V) = g^{-1}(V)$. Beginnen wir mit $g^{-1}(V) \subset f(V)$. Für $x \in g^{-1}(V)$ ist $g(x) \in V$. Aus (*) folgt dann $x = f(g(x))$. Die andere Richtung $f(V) \subset g^{-1}(V)$ folgt aus der Eindeutigkeit der Lösung von $F(x, y) = 0$. Denn für $x = f(y)$ folgt $x - f(y) = 0 = F(x, y)$. Und aus $F(x, g(x)) = 0$ folgt $y = g(x) = g(f(y))$. Also ist für $U' := f(V)$ $f|_{U'} : V \to U'$ ein Diffeomorphismus mit C^1-Umkehrfunktion. □

Bemerkung 4. Ist f eine C^r-Abbildung, dann ist auch f^{-1} eine C^r Abbildung. Dies gilt auch für $r = \infty$.

Beispiel 1. Sei

$$\gamma : \mathbb{R} \to \mathbb{R}^2$$
$$\gamma(t) := (t^3 - 4t,\ t^2 - 4)$$
$$\gamma'(t) = (3t^2 - 4,\ 2t) \neq 0\ \forall t \in \mathbb{R}$$

Jedoch ist $\gamma(2) = (0,0) = \gamma(-2)$. γ ist global nicht injektiv, aber lokal ein Diffeomorphismus.

1.1 Immersionen und Untermannigfaltigkeiten

Definition 1.9: Immersionen

Sei $T \subset \mathbb{R}^k$ offen. Eine C^1-Abbildung $\varphi : T \to \mathbb{R}^n$ heißt C^1-**Immersion**, wenn $\mathrm{Rg}(D_\varphi)(t) = k$ ist für alle $t \in T$. Insbesondere ist $k \leq n$. Die Spaltenvektoren $\frac{\partial \varphi}{\partial t_1}, \ldots, \frac{\partial \varphi}{\partial t_k}$ sind dann linear unabhängig. Also existieren Indizes $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ sodass $\det \frac{\partial(\varphi_{i_1},\ldots,\varphi_{i_k})}{\partial t_1,\ldots,t_k} \neq 0$. Die Determinante ist ein Polynom in den Einträgen der eingegebenen Matrix, ist insbesondere also stetig. Deswegen gilt für $t \in U$, U eine (hinreichend kleine) Umgebung von t_0

$$\det \frac{\partial(\varphi_{i_1}, \ldots, \varphi_{i_k})}{\partial t_1, \ldots, t_k} \neq 0$$

O.B.d.A. können wir $i_1 = 1, \ldots, i_k = k$ setzen. Wir schreiben ab jetzt

$$\varphi_{\leq k} := (\varphi_1, \ldots, \varphi_k) : R \to \mathbb{R}^k \subset \mathbb{R}^n$$

und es ist $(D_{\varphi \leq k} = \det(\ldots)$ invertierbar in $t_0 \in T$. Jetzt können wir den Satz über umkehrbare Funktionen anwenden. Es gibt nun eine Umbegbung $V \subset T$ von t_0 und eine Umgebung $U \subset \mathbb{R}^k$, sodass $\varphi_{\leq k}|_V : V \to U$ ein Diffeomorphismus ist. Also ist auch $\varphi : V \to \varphi(V)$ injektiv, da φ bereits in den ersten k Komponenten injektiv ist. Außerdem ist φ stetig und per Konstruktion surjektiv. Und φ hat die stetige Umkehrfunktion

$$\varphi(V) \to V$$
$$(x_1, \ldots, x_n) \mapsto (\varphi_{\leq k})^{-1}(x_1, \ldots, x_k)$$

Wobei wir die restlichen Terme $x_{k+1}, \ldots, x_n$ umweltfreundlich entsorgen. Folgich ist $\varphi|_V : V \to \varphi(V)$ ein Homöomorphismus.

Bemerkung 5. Wir bezeichnen T auch als **Paramterbereich** und φ als **Parametrisierung** durch t.

Definition 1.10: Untermannigfaltigkeiten

$M \subset \mathbb{R}^n$ heißt k**-dimensionale Untermannigfaltigkeit** (des $\mathbb{R}^n$), wenn jeder Punkt $p \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ besitzt zusammen mit einer C^1 -Immersion $T \underset{\text{offen}}{\subset} \mathbb{R}^k$, $\varphi : T \to \mathbb{R}^n$, sodass

$$\varphi : T \to M \cap U$$

ein Homöomorphismus ist. (φ heißt **Karte** um p auf M)

Beispiel 2. Im vorigen Beispiel hatten wir uns mit $\gamma(t) = (t^3 - 4t, t^2 - 4)$ beschäftigt. Dies definiert keine Untermannigfaltigkeit, da wir für den Punkt $(0,0)$ (dort, wo die Kurve sich selbst schneidet) keine lokale Parametrisierung finden können.

Beispiel 3. Der Kreis $S_1 = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\right\}$ ist eine Untermannigfaltigkeit vom $\mathbb{R}^2$. Etwa mit der Parametrisierung

$$\varphi(t) = (\cos t, \sin t)$$
$$\varphi'(t) = (-\sin t, \cos t)$$

Also ist $\left\|\varphi'(t)\right\| = 1$ und $\varphi'(t) \neq 0$.