

Mathematik für Physiker 3 / Analysis 2*

Wintersemester 2019/20

Stefan Teufel
Mathematisches Institut
Uni Tübingen

24. Januar 2020

*Diese vorläufige Version des Skriptums ist nur zum Gebrauch parallel zum Besuch der Vorlesung gedacht. Das Studium des Skripts kann den Besuch der Vorlesung **nicht** ersetzen! Falls Sie Fehler finden, teilen Sie mir diese (auch die offensichtlichen) bitte mit!

Inhaltsverzeichnis

1	Topologische, metrische und normierte Räume	1
2	Stetigkeit	11
3	Kompaktheit und Zusammenhang	19
4	Differenzierbarkeit	27
5	Taylorformel und lokale Extrema	47
6	Implizite Funktionen	55
7	Differentialrechnung in Banachräumen	65
8	Gewöhnliche Differentialgleichungen	67
9	Lineare Differentialgleichungen	77
10	Integration im \mathbb{R}^n : Eine kurze Einführung	89

1 Topologische, metrische und normierte Räume

Um Konzepte wie Stetigkeit, Differenzierbarkeit und Integrierbarkeit für Funktionen $f(x)$ mehrerer reeller Variable $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ zu definieren, muss man zunächst Begriffe wie Konvergenz von Folgen $x_m \rightarrow a \in \mathbb{R}$, offenes Intervall $(a, b) \in \mathbb{R}$, Vollständigkeit, etc. auf den \mathbb{R}^n verallgemeinern. Da wir schließlich aber auch über den \mathbb{R}^n hinausgehen werden, machen wir das alles gleich etwas allgemeiner auf sogenannten metrischen bzw. noch allgemeiner auf topologischen Räumen. Im Ergebnis werden wir nicht nur die genannten Konzepte aus der Analysis 1 verallgemeinern können, der Blick von einer höheren Warte aus wird auch zu einem tieferen Verständnis dieser Konzepte selbst im Spezialfall von Funktionen auf \mathbb{R} führen.

1.1 Definition. Metrik

Sei X eine Menge. Eine Abbildung

$$d : X \times X \rightarrow [0, \infty)$$

heißt **Metrik auf X** , wenn für alle $x, y, z \in X$ gilt:

$$(i) \quad d(x, y) = 0 \quad \Leftrightarrow \quad x = y, \quad (\text{Definitheit})$$

$$(ii) \quad d(x, y) = d(y, x), \quad (\text{Symmetrie})$$

$$(iii) \quad d(x, z) \leq d(x, y) + d(y, z). \quad (\text{Dreiecksungleichung})$$

Das Paar (X, d) heißt dann **metrischer Raum** und für zwei Punkte $x, y \in X$ heißt $d(x, y)$ der **Abstand** von x und y . \diamond

Ein metrischer Raum ist also eine Menge X , auf der man mit Hilfe einer Funktion $d : X \times X \rightarrow [0, \infty)$ jedem Paar von Punkten $\{x, y\}$ einen Abstand $d(x, y) \in [0, \infty)$ zuweist.

1.2 Beispiele. Euklidische und diskrete Metrik

(a) Sei $X = \mathbb{R}^n$. Dann ist für $x, y \in \mathbb{R}^n$ durch $d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$ eine Metrik definiert, die sog. **euklidische Metrik**. (\mathbb{R}^n, d) heißt **euklidischer Raum**.

Die Eigenschaften (i) und (ii) aus Definition 1.1 sind offensichtlich, die Dreiecksungleichung nicht (vgl. Aufgabe 3 auf Blatt 1). Falls nicht anders gesagt, sei die Menge \mathbb{R}^n im Folgenden immer mit der euklidischen Metrik versehen.

(b) Sei X eine beliebige Menge und

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

Dann ist d eine Metrik auf X und heißt **diskrete Metrik**. In diesem Fall sind die Eigenschaften (i)–(iii) aus Definition 1.1 in offensichtlicher Weise erfüllt.

(c) Die euklidische Einheitssphäre $S^2 := \{x \in \mathbb{R}^3 \mid d(x, 0) = 1\}$ im \mathbb{R}^3 ist mit der Metrik

$$d_S(x, y) := \arccos(\langle x, y \rangle)$$

ein metrischer Raum. Hier bezeichnet $\langle x, y \rangle := x_1y_1 + x_2y_2 + x_3y_3$ das euklidische Skalarprodukt. Gemäß dieser Metrik ist also der Abstand zweier Punkte auf der Sphäre durch die Länge des kürzesten Großkreisbogens auf der Sphäre gegeben, der die beiden Punkte verbindet.

1.3 Definition. Induzierte Metrik

Auf jeder Teilmenge $Y \subset X$ eines metrischen Raums (X, d) definiert die Einschränkung $d|_{Y \times Y} : Y \times Y \rightarrow [0, \infty)$ wieder eine Metrik, die sog. **induzierte Metrik**. Somit ist Y mit der induzierten Metrik selbst wieder ein metrischer Raum $(Y, d|_{Y \times Y})$. \diamond

1.4 Proposition. Das kartesische Produkt metrischer Räume

Seien (X_1, d_1) und (X_2, d_2) metrische Räume. Dann wird auf dem kartesischen Produkt $X_1 \times X_2 = \{(x_1, x_2) \mid x_1 \in X_1 \text{ und } x_2 \in X_2\}$ durch

$$d(x, y) := d_1(x_1, y_1) + d_2(x_2, y_2)$$

eine Metrik definiert. Man kann also das Produkt metrischer Räume immer in natürlicher Weise wieder als metrischen Raum betrachten.

Beweis. Übungsaufgabe. \square

1.5 Definition. Norm

Sei V ein Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}$ oder dem Körper $\mathbb{K} = \mathbb{C}$. Eine Abbildung

$$\|\cdot\| : V \rightarrow [0, \infty)$$

heißt eine **Norm** auf V , wenn für alle $x, y \in V$ und $\lambda \in \mathbb{K}$ gilt:

$$(i) \quad \|x\| = 0 \iff x = 0, \quad \text{(Definitheit)}$$

$$(ii) \quad \|\lambda x\| = |\lambda| \|x\|, \quad \text{(Homogenität)}$$

$$(iii) \quad \|x + y\| \leq \|x\| + \|y\|. \quad \text{(Dreiecksungleichung)}$$

Das Paar $(V, \|\cdot\|)$ heißt **normierter Raum**.

1.6 Proposition. Jede Norm induziert eine Metrik

Auf einem normierten Raum $(V, \|\cdot\|)$ wird durch

$$d : V \times V \rightarrow [0, \infty), \quad d(x, y) := \|x - y\|,$$

eine Metrik auf V definiert.

Beweis. (i) $d(x, y) = 0 \iff \|x - y\| = 0 \iff x - y = 0 \iff x = y,$

(ii) $d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |(-1)| \|y - x\| = \|y - x\| = d(y, x),$

(iii) $d(x, z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z).$ \square

1.7 Beispiele. (a) **Die euklidische Norm auf \mathbb{R}^n :**

Sei $V = \mathbb{R}^n$. Dann ist für $x \in \mathbb{R}^n$ durch

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

eine Norm definiert, die euklidische Norm. Die von ihr nach Proposition 1.6 induzierte Metrik ist die euklidische Metrik.

(b) Die **Maximumsnorm** auf \mathbb{R}^n ist

$$\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\}$$

und die **Summennorm** oder **1-Norm** auf \mathbb{R}^n ist

$$\|x\|_1 := \sum_{j=1}^n |x_j|.$$

Im allgemeinen ist für $x \in \mathbb{R}^n$ und $n > 1$ natürlich $\|x\|_1 \neq \|x\|_2 \neq \|x\|_\infty$. Beispielsweise findet man für $x = (1, 1) \in \mathbb{R}^2$, dass $\|x\|_1 = 2$, $\|x\|_2 = \sqrt{2}$ und $\|x\|_\infty = 1$.

Die Normeigenschaften von $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ zeigen Sie in den Übungen.

- (c) Sei X eine beliebige Menge und V der Vektorraum der beschränkten reell-wertigen Funktionen auf X ,

$$V = \{f : X \rightarrow \mathbb{R} \mid \sup_{x \in X} |f(x)| < \infty\}.$$

Dann ist

$$\|f\|_\infty := \sup_{x \in X} |f(x)|$$

eine Norm auf V .

Beweis. (i) $\|f\|_\infty = 0 \Leftrightarrow f(x) = 0 \forall x \in X \Leftrightarrow f = 0$.

(ii) $\|\lambda f\|_\infty = \sup_{x \in X} |\lambda| |f(x)| = |\lambda| \sup_{x \in X} |f(x)| = |\lambda| \|f\|_\infty$.

(iii) $\|f + g\|_\infty = \sup_{x \in X} |f(x) + g(x)| \leq \sup_{x \in X} (|f(x)| + |g(x)|) \leq \sup_{x \in X} |f(x)| + \sup_{x \in X} |g(x)| = \|f\|_\infty + \|g\|_\infty$. □

- (d) In (c) kann man \mathbb{R} durch einen beliebigen normierten Raum $(Y, \|\cdot\|)$ ersetzen und erhält, dass auf

$$V = \{f : X \rightarrow Y \mid \sup_{x \in X} \|f(x)\| < \infty\}.$$

durch

$$\|f\|_\infty := \sup_{x \in X} \|f(x)\|$$

eine Norm definiert wird.

- (e) Sei (Y, d) ein metrischer Raum, X eine Menge und $f : X \rightarrow Y$ eine injektive Abbildung. Dann wird auf X durch

$$d_f(x_1, x_2) := d(f(x_1), f(x_2))$$

eine Metrik definiert (vgl. Aufgabe 1 (a)).

1.8 Merke. Zusammenhang zwischen Metrik und Norm

Metrik $\hat{=}$ Abstand zwischen Punkten beliebiger Mengen

Norm $\hat{=}$ Länge eines Vektors

Jede Norm induziert auch eine Metrik, da "Abstand" gleich "Länge des Differenzvektors" gesetzt werden kann, vgl. Proposition 1.6. Wir können also jeden normierten Raum insbesondere auch als metrischen Raum auffassen und alle Definitionen und Sätze in metrischen Räumen gelten entsprechend für normierte Räume mit der induzierten Metrik.

1.9 Definition. Offene Mengen in metrischen Räumen

Sei (X, d) ein metrischer Raum.

- (a) Für $x_0 \in X$ und $r > 0$ heißt

$$B_r(x_0) := \{x \in X \mid d(x, x_0) < r\}$$

die **offene Kugel** (der offene Ball) um x_0 vom Radius r .

- (b) Eine Teilmenge $U \subset X$ heißt **Umgebung** des Punktes $x_0 \in X$, falls U auch eine offene Kugel um x_0 enthält, also falls ein $\varepsilon > 0$ existiert, so dass

$$B_\varepsilon(x_0) \subset U.$$

Falls U Umgebung von x_0 ist, so heißt x_0 **innerer Punkt** von U .

Insbesondere ist also für $r > 0$ die offene Kugel $B_r(x_0)$ selbst eine Umgebung von x_0 .

- (c) Eine Teilmenge $U \subset X$ heißt **offen**, wenn sie Umgebung jedes ihrer Punkte ist, d.h. wenn zu jedem $x \in U$ ein $\varepsilon > 0$ existiert, so dass $B_\varepsilon(x) \subset U$.

Eine Menge ist also offen, wenn sie nur innere Punkte enthält.

- 1.10 Beispiele.** (a) Die offene Kugel $B_r(x_0)$ ist offen: Sei $x \in B_r(x_0)$ beliebig, dann gilt nach Definition $d(x, x_0) < r$. Man setzt also $\varepsilon := r - d(x, x_0) > 0$ und hat dann wegen der Dreiecksungleichung $B_\varepsilon(x) \subset B_r(x_0)$:

$$y \in B_\varepsilon(x) \Rightarrow d(y, x_0) \leq d(y, x) + d(x, x_0) < \varepsilon + d(x, x_0) = r.$$

- (b) Ein offenes Intervall $(a, b) \subset \mathbb{R}$ ist eine offene Menge bzgl. der euklidischen Metrik $d(x, y) = |x - y|$, denn $(a, b) = B_{\frac{b-a}{2}}(\frac{a+b}{2})$.
- (c) Sei X beliebig und versehen mit der diskreten Metrik (Beispiel 1.2), dann ist jede Teilmenge $U \subset X$ offen. (Klar: $B_{\frac{1}{2}}(x) = \{x\} \forall x \in X$.)

1.11 Proposition. In einem metrischen Raum X gilt:

- (a) \emptyset und X sind offen.
 (b) Sind $U, V \subset X$ offen, so ist auch $U \cap V$ offen.
 (c) Sind $U_i \subset X$ offen ($i \in \mathcal{I}$, \mathcal{I} eine beliebige Indexmenge), so ist auch $\bigcup_{i \in \mathcal{I}} U_i$ offen.

Beweis. (a) Offenbar.

- (b) Sei $x \in U \cap V$. Dann gibt es $r, s > 0$ mit $B_r(x) \subset U$ und $B_s(x) \subset V$, da U und V ja offen sind. Setze $\varepsilon := \min(r, s) > 0$, dann gilt $B_\varepsilon(x) \subset U \cap V$, also ist auch $U \cap V$ offen.
- (c) Sei $x \in \bigcup U_i$. Dann gibt es ein $i_0 \in \mathcal{I}$ mit $x \in U_{i_0}$. Nun ist U_{i_0} offen, also existiert $\varepsilon > 0$ mit $B_\varepsilon(x) \subset U_{i_0} \subset \bigcup_i U_i$. □

1.12 Bemerkung. Mit Proposition 1.11 (b) sind natürlich auch Durchschnitte endlich vieler offener Mengen wieder offen. Durchschnitte unendlich vieler offener Mengen sind im Allgemeinen aber nicht mehr offen, z.B. ist $I_n = (-\frac{1}{n}, \frac{1}{n}) \subset \mathbb{R}$ für alle $n \in \mathbb{N}$ offen, aber $\bigcap_{n \in \mathbb{N}} I_n = \{0\}$ nicht. \diamond

1.13 Definition. Abgeschlossene Mengen

Sei X ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt **abgeschlossen**, wenn ihr Komplement $A^c := X \setminus A := \{x \in X \mid x \notin A\}$ offen ist.

- 1.14 Beispiele.** (a) $[a, b] \subset \mathbb{R}$ ist abgeschlossen.
 (b) $[a, \infty) \subset \mathbb{R}$ ist abgeschlossen.
 (c) $[a, b] \subset \mathbb{R}$ ist für $-\infty < a < b < \infty$ weder abgeschlossen noch offen.
 (d) Sei (X, d) metrischer Raum. Dann sind \emptyset und X sowohl offen als auch abgeschlossen.

1.15 Merke. Im Allgemeinen ist eine Teilmenge eines metrischen Raumes weder abgeschlossen noch offen, manchmal aber auch beides.

1.16 Proposition. In einem metrischen Raum X gilt:

- (a) \emptyset und X sind abgeschlossen.
 (b) Sind $U, V \subset X$ abgeschlossen, so ist auch $U \cup V$ abgeschlossen.
 (c) Beliebige Schnitte abgeschlossener Mengen sind abgeschlossen.

Beweis. Komplementbildung in Proposition 1.11. Man erinnere sich dazu an die De Morganschen Regeln: Für beliebige Teilmengen $V_i \subset X$, $i \in \mathcal{I}$, einer Obermenge X gelten $(\cup_{i \in \mathcal{I}} V_i)^c = \cap_{i \in \mathcal{I}} V_i^c$ und $(\cap_{i \in \mathcal{I}} V_i)^c = \cup_{i \in \mathcal{I}} V_i^c$. \square

Man kann nun viele der im Folgenden zu diskutierenden Konzepte wie Konvergenz, Stetigkeit, Kompaktheit etc. noch verallgemeinern, indem man vergisst, dass die offenen Mengen in Definition 1.9 mit Hilfe einer Metrik definiert wurden und stattdessen die in Proposition 1.11 hergeleiteten Eigenschaften offener Mengen in metrischen Räumen zur Definition erhebt.

1.17 Definition. Topologische Räume

Eine Menge X mit einem Teilmengensystem $\mathcal{T} \subset \mathcal{P}(X)$ (also einer Menge \mathcal{T} von Teilmengen von X , wobei $\mathcal{P}(X)$ die Potenzmenge, also die Menge aller Teilmengen bezeichnet) heißt **topologischer Raum** und \mathcal{T} eine **Topologie**, falls

- (a) $\emptyset, X \in \mathcal{T}$
 (b) $U, V \in \mathcal{T} \Rightarrow U \cap V \in \mathcal{T}$
 (c) $U_i \in \mathcal{T}$ für $i \in I \Rightarrow \bigcup_{i \in I} U_i \in \mathcal{T}$

Die Mengen $U \in \mathcal{T}$ heißen dann die **offenen Mengen**. Komplemente offener Mengen heißen wieder **abgeschlossen**.

Die Definition von innerer Punkt bzw. Umgebung lautet dann: x_0 ist **innerer Punkt** einer Menge A bzw. A ist **Umgebung** von x_0 , falls es eine offene Menge O gibt mit $x_0 \in O$ und $O \subset A$. \diamond

Zusammenfassend sollten Sie sich merken, dass man in einem topologischen Raum weiß, was die offenen Mengen, was Umgebungen und was innere Punkte sind. Ausgehend davon folgen, wie wir sehen werden, Begriffe wie Konvergenz, Stetigkeit, Kompaktheit, etc. In metrischen Räumen sind die offenen Mengen gemäß Definition 1.9 mithilfe offener Bälle durch die Metrik definiert. Metrische Räume sind also insbesondere auch topologische Räume.

1.18 Bemerkung. Topologische Räume in denen die Topologie nicht durch eine Metrik gegeben ist werden in dieser Vorlesung keine Rolle spielen. Trotzdem ist es nützlich, und in vielen Fällen auch einfacher, den Begriff der Metrik möglichst selten zu verwenden und stattdessen mit Umgebungen und offenen Mengen zu argumentieren.

Wir werden daher Definitionen und Aussagen wenn möglich für allgemeine topologische Räume formulieren. Im Umkehrschluss bedeutet das, dass alle Aussagen und Definitionen die wir nur für metrische bzw. normierte Räume formulieren in allgemeinen topologischen Räumen so nicht gelten. \diamond

1.19 Definition. Inneres, Abschluss, Rand

Sei X ein topologischer Raum und $Y \subset X$ eine Teilmenge.

- (a) Die Vereinigung aller in Y enthaltenen offenen Mengen

$$\overset{\circ}{Y} := \bigcup_{\substack{U \subset Y \\ U \text{ offen}}} U$$

heißt das **Innere** oder der **offene Kern** von Y .

- (b) Der Schnitt aller abgeschlossenen Mengen die Y enthalten

$$\overline{Y} := \bigcap_{\substack{A \supset Y \\ A \text{ abgeschlossen}}} A$$

heißt der **Abschluss** oder die **abgeschlossene Hülle** von Y .

- (c) Der Abschluss von Y ohne das Innere von Y

$$\partial Y = \overline{Y} \setminus \overset{\circ}{Y}$$

heißt der **Rand** von Y .

Es gilt also jeweils per Definition

- $\overset{\circ}{Y} \subset Y \subset \overline{Y}$
- $\overset{\circ}{Y}$ ist die größte in Y enthaltene offene Menge, insbesondere ist $\overset{\circ}{Y}$ offen.
- Y ist offen $\Leftrightarrow Y = \overset{\circ}{Y}$
- \overline{Y} ist die kleinste abgeschlossene Menge in der Y enthalten ist, insbesondere ist \overline{Y} abgeschlossen.
- Y ist abgeschlossen $\Leftrightarrow Y = \overline{Y}$
- $\overset{\circ}{Y}$ ist das Komplement von $\overline{X \setminus Y}$, also $(\overset{\circ}{Y})^c = \overline{(Y^c)}$ (folgt wieder mit De Morgan), bzw. $(Y^c)^\circ = (\overline{Y})^c$ (also die erste Behauptung für Y^c).

1.20 Proposition. Alternative Charakterisierungen

In einem topologischen Raum X gilt für jedes $Y \subset X$:

- (a) $\overset{\circ}{Y}$ ist die Menge der inneren Punkte von Y .
- (b) Ein Punkt $x \in X$ ist genau dann Randpunkt von Y , wenn jede Umgebung von x sowohl einen Punkt aus Y als auch einen Punkt aus $Y^c = X \setminus Y$ enthält.
- (c) $\overset{\circ}{Y} = Y \setminus \partial Y$
- (d) $\overline{Y} = Y \cup \partial Y$

Beweis. Übungsaufgabe □

1.21 Beispiele. (a) Sei $Y = (a, b) \subset \mathbb{R}$. Dann ist $\overset{\circ}{Y} = (a, b)$, $\overline{Y} = [a, b]$ und $\partial Y = \{a, b\}$.

- (b) Betrachte $\mathbb{Q} \subset \mathbb{R}$: Da für jedes $x \in \mathbb{R}$ der ε -Ball $B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$ um x sowohl rationale als auch irrationale Zahlen enthält, ist gemäß Proposition 1.20 (b) $\partial \mathbb{Q} = \mathbb{R}$. Mit Proposition 1.20 (c) und (d) folgt dann sofort, dass $\overset{\circ}{\mathbb{Q}} = \mathbb{Q} \setminus \partial \mathbb{Q} = \emptyset$ und $\overline{\mathbb{Q}} = \mathbb{Q} \cup \partial \mathbb{Q} = \mathbb{R}$.

Für Folgen (x_n) in \mathbb{R} oder \mathbb{C} wurde in Analysis 1 Konvergenz so definiert: Eine Folge (x_n) konvergiert gegen einen Punkt a , falls jedes ε -Intervall um a alle bis auf endlich viele Folgenglieder enthält. In metrischen Räumen ersetzt man in dieser Definition nun einfach “ ε -Intervall” durch

“ ε -Kugel” und in topologischen Räumen durch “Umgebung”. Wir beginnen also mit der allgemeinen Definition von Folgenkonvergenz in topologischen Räumen und zeigen dann, dass wir in metrischen Räumen und insbesondere in \mathbb{R} bzw. \mathbb{C} die Definition aus Analysis 1 als Spezialfall des allgemeinen Konzepts von Folgenkonvergenz wiederfinden.

1.22 Definition. Konvergenz von Folgen

Sei X ein topologischer Raum. Eine Folge (x_n) in X heißt **konvergent gegen** $a \in X$, geschrieben

$$\lim_{n \rightarrow \infty} x_n = a$$

oder kurz $x_n \rightarrow a$ für $n \rightarrow \infty$, wenn gilt:

Für jede Umgebung $U \subset X$ von a gibt es ein $N \in \mathbb{N}$, so dass $x_n \in U$ für alle $n \geq N$.

Eine Folge konvergiert also gegen einen Punkt, wenn jede (“noch so kleine”) Umgebung des Punktes alle bis auf endlich viele Folgenglieder enthält. \diamond

1.23 Bemerkung. Grenzwerte sind im Allgemeinen nicht mehr eindeutig

In allgemeinen topologischen Räumen kann eine Folge mehrere verschiedene Grenzwerte haben. Beispielsweise wird auf einer beliebigen Menge X durch $\mathcal{T} := \{\emptyset, X\}$ eine Topologie definiert, die sog. indiskrete Topologie. Und bezüglich dieser Topologie konvergiert jede Folge in X gegen jeden Punkt in X , da jeder Punkt $x \in X$ als einzige Umgebung die ganze Menge X besitzt. Daher ist die Schreibweise $\lim_{n \rightarrow \infty} x_n = a$ für Grenzwerte von Folgen in allgemeinen topologischen Räumen eigentlich nicht angebracht (aber durchaus üblich, siehe z.B. den deutschen Wikipedia-Eintrag), da das Gleichheitszeichen Eindeutigkeit suggeriert.

1.24 Proposition. Konvergenz in metrischen Räumen

Eine Folge (x_n) in einem metrischen Raum X konvergiert genau dann gegen $a \in X$, wenn $d(x_n, a)$ als Folge in \mathbb{R} gegen Null konvergiert, also

$$\lim_{n \rightarrow \infty} x_n = a \quad \Leftrightarrow \quad \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N : d(x_n, a) < \varepsilon,$$

oder anders gesagt, wenn jeder ε -Ball um a alle bis auf endlich viele Folgenglieder enthält. Der Grenzwert ist eindeutig.

Beweis. “ \Rightarrow ”: Da jeder ε -Ball um a insbesondere eine Umgebung von a ist, und da nach Voraussetzung jede Umgebung von a fast alle Folgenglieder der Folge enthält, enthält auch jeder ε -Ball um a fast alle Folgenglieder.

“ \Leftarrow ”: In einem metrischen Raum enthält jede Umgebung von a auch einen ε -Ball um a . Da jeder ε -Ball um a fast alle Folgenglieder der Folge enthält, enthält daher auch jede Umgebung von a fast alle Folgenglieder.

Eindeutigkeit des Grenzwertes: In jedem topologischen Raum gilt: Falls zwei Punkte a und b disjunkte Umgebungen U und V haben, so können nicht sowohl U als auch V fast alle Folgenglieder einer Folge enthalten und somit nicht sowohl a als auch b Grenzwert derselben Folge sein. Die Eindeutigkeit gilt also in jedem topologischen Raum in dem verschiedene Punkte immer auch disjunkte Umgebungen besitzen (sog. Hausdorffräume). Dies ist für metrische Räume immer der Fall, da für $a, b \in X$ mit $a \neq b$ gilt, dass $d(a, b) > 0$ und somit $B_{d(a,b)/2}(a)$ und $B_{d(a,b)/2}(b)$ disjunkte Umgebungen von a und b sind. \square

1.25 Definition. Häufungspunkt

Sei X ein topologischer Raum und (x_n) eine Folge in X . Ein Punkt $a \in X$ heißt **Häufungspunkt** von (x_n) , falls jede Umgebung U von a unendlich viele Folgenglieder enthält. \diamond

1.26 Satz. Folgenkriterium für den Abschluss in metrischen Räumen

Sei X ein metrischer Raum und $A \subset X$. Dann liegt ein Punkt $a \in X$ genau dann im Abschluss \overline{A} der Menge A , wenn es eine Folge (x_n) in A gibt, die gegen a konvergiert, also

$$a \in \overline{A} \quad \Leftrightarrow \quad \exists (x_n) \text{ in } A : \lim_{n \rightarrow \infty} x_n = a.$$

Eine Teilmenge $A \subset X$ ist also genau dann abgeschlossen, wenn für jede konvergente Folge (x_n) in A gilt, dass auch der Grenzwert $a = \lim_{n \rightarrow \infty} x_n$ in A liegt.

Beweis. „ \Rightarrow “: Sei $a \in \overline{A} = A \cup \partial A$. Falls $a \in A$, wähle $x_n = a$ für alle $n \in \mathbb{N}$. Falls $a \in \partial A$, so gilt nach Bemerkung 1.20 (b) für jedes $n \in \mathbb{N}$, dass $B_{\frac{1}{n}}(a) \cap A \neq \emptyset$. Wir können also für jedes $n \in \mathbb{N}$ ein $x_n \in B_{\frac{1}{n}}(a) \cap A$ auswählen. Die konstruierten Folgen liegen jeweils ganz in A und konvergieren gegen a .

„ \Leftarrow “: Diese Richtung gilt auch in topologischen Räumen. Sei (x_n) eine Folge in A mit $\lim_{n \rightarrow \infty} x_n = a$. Angenommen, $a \notin \overline{A}$, also $a \in X \setminus \overline{A}$. Dann wäre aber die offene Menge $X \setminus \overline{A}$ Umgebung von a und müßte alle bis auf endlich viele Folgenglieder x_n enthalten. Das ist aber ein Widerspruch zur Annahme, dass $x_n \in A$ für alle $n \in \mathbb{N}$. □

1.27 Definition. Cauchyfolge

Sei (X, d) ein metrischer Raum. Eine Folge (x_n) in X heißt **Cauchyfolge**, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n, m \geq N$ gilt

$$d(x_n, x_m) < \varepsilon.$$

1.28 Proposition. Jede konvergente Folge ist eine Cauchyfolge

Sei (x_n) eine Folge in einem metrischen Raum X . Falls (x_n) konvergiert, so ist (x_n) auch eine Cauchyfolge.

Beweis. Sei $\lim_{n \rightarrow \infty} x_n = a$. Dann gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ so, dass $d(x_n, a) < \frac{\varepsilon}{2}$ für alle $n \geq N$. Wegen der Dreiecksungleichung ist dann für $n, m \geq N$ aber $d(x_n, x_m) \leq d(x_n, a) + d(x_m, a) < \varepsilon$. □

Metrische Räume in denen auch die Umkehrung gilt heißen vollständig.

1.29 Definition. Vollständigkeit und Banachraum

- (a) Ein metrischer Raum (X, d) heißt **vollständig**, wenn in ihm jede Cauchyfolge konvergiert.
- (b) Ein vollständiger normierter Raum $(V, \|\cdot\|)$ heißt **Banachraum**.

1.30 Bemerkungen. (a) Der euklidische Raum $(\mathbb{R}^n, \|\cdot\|_2)$ bzw. auch der $(\mathbb{C}^n, \|\cdot\|_2)$ sind Beispiele für Banachräume. Wir zeigen das an dieser Stelle noch nicht da wir in Proposition 2.15 allgemeiner zeigen werden, dass jeder endlichdimensionale normierte Raum vollständig ist.

- (b) Konvergenz, Kompaktheit, Stetigkeit, Abschluss, Rand etc. sind topologische Begriffe. Die Konzepte Cauchyfolge und Vollständigkeit benötigen mehr Struktur, z.B. eine Metrik. ◇

Wir haben gesehen, dass jede Norm eine Metrik und jede Metrik eine Topologie (also eine Definition von offenen Mengen) induziert. Konvergenz von Folgen hängt allerdings nur von der Topologie ab. Deshalb ist es nützlich zu verstehen, wann verschiedene Normen den gleichen Konvergenzbegriff und auch den gleichen Vollständigkeitsbegriff liefern.

1.31 Definition. Äquivalenz von Normen

Zwei Normen $\|\cdot\|_a$ und $\|\cdot\|_b$ auf einem Vektorraum V heißen **äquivalent**, wenn es Konstanten $c, C > 0$ gibt, so dass für alle $x \in V$ gilt:

$$c \|x\|_a \leq \|x\|_b \leq C \|x\|_a. \quad (1.1)$$

1.32 Bemerkung. Es handelt sich bei der Äquivalenz von Normen tatsächlich um eine symmetrische Relation, denn mit (1.1) gilt dann offensichtlich auch

$$\frac{1}{C} \|x\|_b \leq \|x\|_a \leq \frac{1}{c} \|x\|_b.$$

1.33 Proposition. Seien $\|\cdot\|_a$ und $\|\cdot\|_b$ äquivalente Normen auf V . Für jedes $U \subset V$, jede Folge (x_n) in V und jedes $y \in V$ gilt dann:

- (a) U ist offen in $(V, \|\cdot\|_a)$ \Leftrightarrow U ist offen in $(V, \|\cdot\|_b)$
- (b) U ist abgeschlossen in $(V, \|\cdot\|_a)$ \Leftrightarrow U ist abgeschlossen in $(V, \|\cdot\|_b)$
- (c) $\lim_{n \rightarrow \infty} x_n = y$ in $(V, \|\cdot\|_a)$ \Leftrightarrow $\lim_{n \rightarrow \infty} x_n = y$ in $(V, \|\cdot\|_b)$.
- (d) (x_n) ist Cauchy in $(V, \|\cdot\|_a)$ \Leftrightarrow (x_n) ist Cauchy in $(V, \|\cdot\|_b)$.
- (e) $(V, \|\cdot\|_a)$ ist vollständig \Leftrightarrow $(V, \|\cdot\|_b)$ ist vollständig.

Beweis. (a) gilt offenbar, da jede ε -Kugel bzgl. $\|\cdot\|_b$ die entsprechende $c\varepsilon$ -Kugel bzgl. $\|\cdot\|_a$ enthält und jede ε -Kugel bzgl. $\|\cdot\|_a$ die entsprechende ε/C -Kugel bzgl. $\|\cdot\|_b$. (b) folgt wieder durch Komplementbildung. Für (c) sei $\lim_{n \rightarrow \infty} x_n = y$ in $(V, \|\cdot\|_a)$, also $\lim_{n \rightarrow \infty} \|x_n - y\|_a = 0$. Dann ist auch $\lim_{n \rightarrow \infty} \|x_n - y\|_b \leq \lim_{n \rightarrow \infty} C \|x_n - y\|_a = 0$. Für (d) argumentiert man analog und (e) folgt sofort aus (c) und (d). \square

1.34 Bemerkung. Wir werden im nächsten Kapitel zeigen, dass auf einem endlichdimensionalen Vektorraum alle Normen äquivalent sind (vgl. Satz 2.13). Auf unendlichdimensionalen Vektorräumen gilt das aber nicht mehr. So ist beispielsweise der Raum $C([0, 1])$ der stetigen Funktionen auf dem Intervall $[0, 1]$ versehen mit der Supremumsnorm $\|\cdot\|_\infty$ aus Beispiel 1.7 vollständig (vgl. Korollar 2.20) nicht aber mit der sogenannten L^2 -Norm $\|f\|_2 := \sqrt{\int_0^1 |f(x)|^2 dx}$.

2 Stetigkeit

Wir führen zunächst den Begriff der Folgenstetigkeit ein. Dieser ist in metrischen Räumen äquivalent zur Stetigkeit und folgt in topologischen Räumen zumindest aus der Stetigkeit. Folgenstetigkeit von Abbildungen erlaubt es uns, die Abbildung mit Grenzwertbildung zu vertauschen.

2.1 Definition. Folgenstetigkeit

Seien X und Y topologische Räume und $a \in X$. Eine Abbildung $f : X \rightarrow Y$ heißt **folgenstetig in a** , wenn für jede Folge (x_n) in X die gegen a konvergiert, ihre Bildfolge $(f(x_n))$ in Y gegen $f(a)$ konvergiert, also wenn für jede Folge (x_n) in X gilt

$$\lim_{n \rightarrow \infty} x_n = a \quad \Rightarrow \quad \lim_{n \rightarrow \infty} f(x_n) = f(a)$$

Es heißt $f : X \rightarrow Y$ **folgenstetig**, wenn f in jedem Punkt $x \in X$ folgenstetig ist. ◊

2.2 Notation. Grenzwert einer Funktion

Sei X ein metrischer Raum und $a \in X$. Wenn für eine Funktion $f : X \setminus \{a\} \rightarrow Y$ gilt, dass für jede Folge (x_n) in $X \setminus \{a\}$ die gegen a konvergiert, ihre Bildfolge $(f(x_n))$ gegen $b \in Y$ konvergiert, so schreibt man dafür kurz

$$\lim_{x \rightarrow a} f(x) = b$$

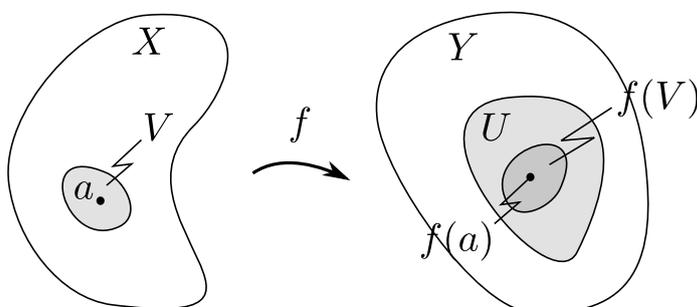
und spricht vom Grenzwert der Funktion an der Stelle a . Folgenstetigkeit von f in a bedeutet daher in dieser Schreibweise

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Die intuitive Idee hinter Stetigkeit ist allerdings eine etwas andere. Eine Funktion soll stetig bei $a \in X$ heißen, wenn sich bei "stetiger" Veränderung des Urbildpunktes a der Bildpunkt $f(a)$ ebenfalls "stetig" ändert, also nicht springt. Diese Idee kann man in topologischen Räumen folgendermaßen formalisieren.

2.3 Definition. Stetigkeit

Seien X und Y topologische Räume und $a \in X$. Eine Abbildung $f : X \rightarrow Y$ heißt **stetig in a** , wenn es zu jeder Umgebung U von $f(a) \in Y$ eine Umgebung V von $a \in X$ gibt, mit $f(V) \subset U$.



Eine Funktion, die an jedem Punkt in X stetig ist, heißt **stetig**.

2.4 Proposition. Stetigkeit impliziert Folgenstetigkeit

Ist eine Abbildung $f : X \rightarrow Y$ zwischen topologischen Räumen stetig am Punkt $a \in X$, so ist sie bei a auch folgenstetig.

2 Stetigkeit

Beweis. Sei (x_n) eine Folge in X mit $\lim_{n \rightarrow \infty} x_n = a$. Es ist zu zeigen, dass $\lim_{n \rightarrow \infty} f(x_n) = f(a)$ gilt, also, dass jede Umgebung von $f(a)$ fast alle Glieder der Folge $(f(x_n))$ enthält. Sei nun U eine Umgebung von $f(a)$. Aufgrund der Stetigkeit von f gibt es eine Umgebung V von a mit $f(V) \subset U$. Wegen $\lim_{n \rightarrow \infty} x_n = a$ enthält V aber fast alle Glieder der Folge (x_n) und somit U fast alle Glieder der Folge $(f(x_n))$. \square

Man mache sich die Einfachheit des Arguments an obigem Bild klar: wenn immer nur endlich viele Glieder der Folge außerhalb von V liegen, können auch jeweils nur endlich viele Glieder der Bildfolge außerhalb von U liegen. Da es zu jedem U ein passendes V gibt, ist man fertig.

2.5 Proposition. “ ε - δ -Stetigkeit” in metrischen Räumen

Seien X und Y metrische Räume und $a \in X$. Eine Abbildung $f : X \rightarrow Y$ ist genau dann stetig in a , wenn gilt: Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in X$ mit $d(x, a) < \delta$ gilt, dass $d(f(x), f(a)) < \varepsilon$.

Anders gesagt: Jede ε -Kugel um $f(a)$ enthält das Bild einer δ -Kugel um a .

Beweis. Es ist zu zeigen, dass die Aussage dazu äquivalent ist, dass jede Umgebung U von $f(a)$ das Bild $f(V)$ einer Umgebung V von a enthält. Diese Äquivalenz folgt aber sofort, da in metrischen Räumen einerseits jede Umgebung eines Punktes eine offene Kugel um diesen Punkt enthält und andererseits jede offene Kugel selbst Umgebung ihres Mittelpunkts ist: “ \Rightarrow ” Sei $B_\varepsilon(f(a)) =: U$ gegeben, dann enthält das entsprechende V eine Kugel $B_\delta(a)$, deren Bild in $B_\varepsilon(f(a))$ liegt. “ \Leftarrow ” Sei U gegeben, dann gibt es ein $B_\varepsilon(f(a)) \subset U$ und das entsprechende $B_\delta(a)$ dient als V . \square

2.6 Proposition. In metrischen Räumen impliziert Folgenstetigkeit auch Stetigkeit

Seien X und Y metrische Räume, $f : X \rightarrow Y$ und $a \in X$. Ist f bei a folgenstetig, so ist f bei a auch stetig.

Beweis. durch Kontraposition. Angenommen f ist nicht stetig. Dann gibt es ein $\varepsilon > 0$, sodass für alle $\delta > 0$ gilt $f(B_\delta(a)) \not\subset B_\varepsilon(f(a))$. Für $\delta = \frac{1}{n}$ wähle dann $x_n \in B_\delta(a) \setminus f^{-1}(B_\varepsilon(f(a)))$. Also $\lim_{n \rightarrow \infty} x_n = a$, aber $f(x_n) \notin B_\varepsilon(f(a))$ für alle $n \in \mathbb{N}$, und somit konvergiert $f(x_n)$ nicht gegen $f(a)$. Also ist f auch nicht folgenstetig. \square

2.7 Satz. Charakterisierung von Stetigkeit durch Urbilder offener Mengen

Seien X und Y topologische Räume. Eine Abbildung $f : X \rightarrow Y$ ist genau dann stetig, wenn das Urbild $f^{-1}(O)$ jeder offenen Menge $O \subset Y$ offen in X ist.

Beweis. “ \Rightarrow ”: Sei $O \subset Y$ offen und $a \in f^{-1}(O)$. Dann ist $f(a) \in O$ und aufgrund der Stetigkeit von f gibt es eine Umgebung V von a mit $f(V) \subset O$. Also ist $V \subset f^{-1}(O)$ und a somit innerer Punkt von $f^{-1}(O)$. Da a beliebig war, hat $f^{-1}(O)$ nur innere Punkte und ist somit offen.

“ \Leftarrow ” Sei $a \in X$ und U eine Umgebung von $f(a) \in Y$. Dann gibt es eine offene Menge $O \subset U$ mit $f(a) \in O$. Deren Urbild $f^{-1}(O) =: V$ enthält a und ist nach Voraussetzung offen, also eine Umgebung von a mit $f(V) \subset O \subset U$. \square

2.8 Beispiele. Stetige Abbildungen in metrischen Räumen

- (a) In einem metrischen Raum (X, d) ist für jedes $b \in X$ die **Abstandsfunktion**

$$d_b : X \rightarrow [0, \infty), d_b(x) := d(b, x)$$

stetig.

Beweis. Sei $a \in X$ beliebig und $\lim_{n \rightarrow \infty} x_n = a$. Aus der Dreiecksungleichung folgt

$$|d_b(x_n) - d_b(a)| = |d(b, x_n) - d(b, a)| \leq d(x_n, a) \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Also ist $\lim_{n \rightarrow \infty} d_b(x_n) = d_b(a)$ und somit ist d_b stetig. □

- (b) In einem normierten Raum $(V, \|\cdot\|)$ ist $\|\cdot\| : V \rightarrow \mathbb{R}$ stetig. *Beweis.* $\|x\| = d_0(x)$ in (a).
 (c) Sei $(V, \|\cdot\|)$ ein normierter Raum. Dann sind die Addition $\text{add}: V \times V \rightarrow V, (x, y) \mapsto x + y$ und die skalare Multiplikation $\text{mult}: \mathbb{K} \times V \rightarrow V, (\alpha, x) \mapsto \alpha x$ stetig.

Beweis. Sei $(x_n, y_n) \rightarrow (x, y)$ also $x_n \rightarrow x$ und $y_n \rightarrow y$ in V , dann gilt

$$\lim_{n \rightarrow \infty} \|\text{add}(x_n, y_n) - \text{add}(x, y)\| = \lim_{n \rightarrow \infty} \|x_n + y_n - x - y\| \leq \lim_{n \rightarrow \infty} (\|x_n - x\| + \|y_n - y\|) = 0.$$

Sei $(\alpha_n, x_n) \rightarrow (\alpha, x)$ also $\alpha_n \rightarrow \alpha$ in \mathbb{K} und $x_n \rightarrow x$ in V , dann gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} \|\text{mult}(\alpha_n, x_n) - \text{mult}(\alpha, x)\| &= \lim_{n \rightarrow \infty} \|\alpha_n x_n - \alpha x\| \\ &\leq \lim_{n \rightarrow \infty} (\|\alpha_n x_n - \alpha x_n\| + \|\alpha x_n - \alpha x\|) \\ &= \lim_{n \rightarrow \infty} (|\alpha_n - \alpha| \|x_n\| + |\alpha| \|x_n - x\|) = 0, \end{aligned}$$

da $\|x_n\|$ beschränkt ist. □

- (d) Die Verkettung (Komposition) stetiger Abbildungen ist stetig: Seien $f : Y \rightarrow Z$ und $g : X \rightarrow Y$ stetig, dann ist $f \circ g : X \rightarrow Z$ stetig.

Beweis. Sei $O \subset Z$ offen $\xrightarrow{f \text{ stetig}} f^{-1}(O) \subset Y$ offen $\xrightarrow{g \text{ stetig}} g^{-1}(f^{-1}(O)) = (f \circ g)^{-1}(O)$ ist offen in X . Also ist $f \circ g$ mit Satz 2.7 stetig. □

- (e) Eine Abbildung $f : X \rightarrow V \times \dots \times V, x \mapsto f(x) = (f_1(x), \dots, f_n(x))$, ist genau dann stetig, wenn alle $f_i : X \rightarrow V, i = 1, \dots, n$, stetig sind. *Beweis.* Übungsaufgabe.

- (f) Sind $f, g : X \rightarrow V$ sowie $h : X \rightarrow \mathbb{K}$ stetig, so sind auch die Funktionen $f + g$ und $h \cdot g : X \rightarrow V$ stetig.

Beweis. Es sind $f + g = \text{add} \circ (f, g)$ und $h \cdot g = \text{mult} \circ (h, g)$ mit (c), (d) und (e) stetig. □

- (g) Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen heißt **Lipschitz-stetig**, falls es eine Zahl $L \in [0, \infty)$ gibt, die sog. **Lipschitzkonstante**, so dass für alle $x_1, x_2 \in X$ gilt

$$d_Y(f(x_1), f(x_2)) \leq L d_X(x_1, x_2).$$

Jede Lipschitz-stetige Funktion ist insbesondere auch stetig (Übungsaufgabe). Die Umkehrung gilt nicht.

2.9 Definition. Homöomorphismen, Isometrien, isometrische Isomorphismen

- (a) Zwei topologische Räume X und Y heißen **homöomorph**, falls es eine Bijektion $f : X \rightarrow Y$ gibt so, dass f und f^{-1} beide stetig sind. Die Abbildung f heißt dann ein **Homöomorphismus**.
 (b) Zwei metrische Räume X und Y heißen **isometrisch**, falls es eine Bijektion $f : X \rightarrow Y$ gibt so, dass für alle $a, b \in X$ gilt

$$d_Y(f(a), f(b)) = d_X(a, b). \tag{2.1}$$

Die Abbildung f heißt dann ein **isometrischer Isomorphismus** metrischer Räume. Ein beliebiges $f : X \rightarrow Y$ welches (2.1) erfüllt heißt **Isometrie**.

2 Stetigkeit

- (c) Zwei normierte Räume V und W heißen **isometrisch isomorph**, falls es eine lineare Bijektion (Isomorphismus) $A : V \rightarrow W$ gibt so, dass für alle $v \in V$ gilt

$$\|Av\|_W = \|v\|_V.$$

Die Abbildung A heißt dann ein **isometrischer Isomorphismus** normierter Räume.

2.10 Bemerkungen. Folgende Aussagen ergeben sich unmittelbar aus den Definitionen:

- (a) Jeder isometrische Isomorphismus metrischer Räume ist insbesondere auch ein Homöomorphismus und jeder isometrische Isomorphismus normierter Räume ist insbesondere auch ein isometrischer Isomorphismus metrischer Räume.
- (b) Homöomorphismen bilden in beiden Richtungen konvergente Folgen auf konvergente Folgen ab.
- (c) Isometrische Isomorphismen bilden in beiden Richtungen Cauchyfolgen auf Cauchyfolgen ab.
- (d) Sind zwei metrische Räume X und Y isometrisch, so ist X genau dann vollständig wenn Y vollständig ist.

2.11 Bemerkung. zur Terminologie

Die griechische Vorsilbe *homöo* bedeutet "gleichartig", während *iso* "gleich auch in quantitativer Hinsicht" bedeutet. Das griechische Wort *morph* bedeutet Form oder Gestalt.

2.12 Beispiele. (a) Aus der linearen Algebra wissen wir, dass jeder endlichdimensionale Vektorraum V nach Wahl einer Basis isomorph zum \mathbb{K}^n ist. Sei $A : V \rightarrow \mathbb{K}^n$ so ein Vektorraumisomorphismus und $\|\cdot\|_V$ eine Norm auf V . Dann wird durch $\|x\|_A := \|A^{-1}x\|_V$ eine Norm auf \mathbb{K}^n definiert. Bezüglich dieser Normen ist A dann ein isometrischer Isomorphismus,

$$\|Av\|_A := \|A^{-1}Av\|_V = \|v\|_V.$$

- (b) Fassen wir \mathbb{R}^{2n} und \mathbb{C}^n jeweils als metrische Räume mit der euklidischen Metrik

$$d_{\mathbb{R}^{2n}}(a, b) := \left(\sum_{j=1}^{2n} (a_j - b_j)^2 \right)^{\frac{1}{2}} \quad \text{bzw.} \quad d_{\mathbb{C}^n}(z, w) := \left(\sum_{j=1}^n |z_j - w_j|^2 \right)^{\frac{1}{2}}$$

auf, so ist die Abbildung

$$J : \mathbb{C}^n \rightarrow \mathbb{R}^{2n}, \quad z \mapsto J(z) := (\operatorname{Re}z_1, \operatorname{Im}z_1, \dots, \operatorname{Re}z_n, \operatorname{Im}z_n)$$

ein isometrischer Isomorphismus, da für $z, w \in \mathbb{C}$ gilt $|z-w|^2 = (\operatorname{Re}z - \operatorname{Re}w)^2 + (\operatorname{Im}z - \operatorname{Im}w)^2$. Als metrische Räume (und somit auch als topologische Räume) können wir \mathbb{C}^n und \mathbb{R}^{2n} also identifizieren. Alle Eigenschaften des \mathbb{R}^{2n} die nur von dessen metrischer Struktur abhängen (z.B. Vollständigkeit) gelten automatisch auch für den \mathbb{C}^n .

Als Vektorräume sind \mathbb{C}^n und \mathbb{R}^{2n} allerdings nicht isomorph, da sie ja Vektorräume über verschiedenen Körpern sind.

- (c) Die Isometrien des euklidischen Raums \mathbb{R}^n auf sich selbst sind genau die Translationen, Rotationen und Spiegelungen sowie deren Kombinationen.

2.13 Satz. Auf einem endlichdimensionalen Vektorraum V sind alle Normen äquivalent.

Beweis. Sei $A : V \rightarrow \mathbb{K}^n$ ein Isomorphismus und für zwei Normen $\|\cdot\|_a$ und $\|\cdot\|_b$ auf V seien $\|\cdot\|_{a,A}$ und $\|\cdot\|_{b,A}$ die entsprechenden Normen auf \mathbb{K}^n aus Beispiel 2.12 (a), die A jeweils zu einem isometrischen Isomorphismus machen. Dann sind $\|\cdot\|_a$ und $\|\cdot\|_b$ genau dann äquivalent wenn $\|\cdot\|_{a,A}$ und $\|\cdot\|_{b,A}$ äquivalent sind. Daher genügt es zu zeigen, dass alle Normen auf dem \mathbb{K}^n äquivalent sind. Sei also o.B.d.A. $V = \mathbb{K}^n$. Da Äquivalenz von Normen eine Äquivalenzrelation ist, reicht es wiederum zu zeigen, dass jede Norm $\|\cdot\|$ zu $\|\cdot\|_1$ mit $\|x\|_1 := \sum_{j=1}^n |x_j|$ äquivalent ist. Sei $x = \sum_{j=1}^n x_j e_j \in \mathbb{K}^n$. Dann gilt

$$\|x\| = \left\| \sum_{j=1}^n x_j e_j \right\| \leq \sum_{j=1}^n |x_j| \|e_j\| \leq \max\{\|e_j\| \mid j = 1, \dots, n\} \sum_{j=1}^n |x_j| =: C \|x\|_1. \quad (2.2)$$

Angenommen es gibt kein c mit $c\|x\|_1 \leq \|x\|$ für alle $x \in \mathbb{K}^n$, dann existiert eine Folge (x_k) in \mathbb{K}^n mit

$$\frac{1}{k} \|x_k\|_1 > \|x_k\| \quad \text{und o.B.d.A.} \quad \|x_k\|_1 = 1. \quad (2.3)$$

Da die n Komponentenfolgen $(x_{k,j})$ (es ist $x_k = (x_{k,1}, \dots, x_{k,n}) \in \mathbb{K}^n$) beschränkte Folgen in \mathbb{K} sind, und somit nach dem Satz von Bolzano-Weierstraß konvergente Teilfolgen enthalten, kann man iterativ eine Teilfolge von (x_k) konstruieren, die bzgl. $\|\cdot\|_1$ konvergiert. Genauer wählt man zunächst eine Teilfolge von (x_k) für die $(x_{k,1})$ in \mathbb{K} konvergiert. Aus dieser Teilfolge wählt man dann wiederum eine Teilfolge für die auch $(x_{k,2})$ konvergiert und setzt dies fort, bis man schließlich eine Teilfolge erhält, für die alle Komponentenfolgen konvergieren. Diese Teilfolge, welche wir der Einfachheit halber wieder mit (x_k) bezeichnen, konvergiert dann aber auch bzgl. der $\|\cdot\|_1$ -Norm gegen $a = (a_1, \dots, a_n) := (\lim_{k \rightarrow \infty} x_{k,1}, \dots, \lim_{k \rightarrow \infty} x_{k,n})$, da

$$\lim_{k \rightarrow \infty} \|a - x_k\|_1 = \lim_{k \rightarrow \infty} \sum_{j=1}^n |a_j - x_{k,j}| = \sum_{j=1}^n \lim_{k \rightarrow \infty} |a_j - x_{k,j}| = 0.$$

Wegen (2.2) konvergiert die Folge (x_k) auch in der $\|\cdot\|$ -Norm gegen denselben Grenzwert a , da

$$\lim_{k \rightarrow \infty} \|a - x_k\| \leq C \lim_{k \rightarrow \infty} \|a - x_k\|_1 = 0.$$

Aufgrund der Stetigkeit der Normen ist dann einerseits $1 = \lim_{k \rightarrow \infty} \|x_k\|_1 = \|\lim_{k \rightarrow \infty} x_k\|_1 = \|a\|_1$, also $a \neq 0$. Andererseits ist $\|a\| = \lim_{k \rightarrow \infty} \|x_k\| \leq \lim_{k \rightarrow \infty} \frac{1}{k} \|x_k\|_1 = 0$, also $a = 0$. \square

2.14 Bemerkung. Wenn wir einen endlichdimensionalen Vektorraum V als normierten Raum auffassen, so hängen gemäß Satz 2.13 und Proposition 1.33 die Konzepte Konvergenz, Cauchyfolge und Vollständigkeit nicht von der konkreten Wahl der Norm ab. Wenn die konkrete Wahl der Norm auf einem endlichdimensionalen Vektorraum nicht aus anderen Gründen wichtig ist, verzichtet man daher oft darauf, sie zu spezifizieren.

2.15 Proposition. Vollständigkeit endlichdimensionaler normierter Räume

Jeder endlichdimensionale Vektorraum V über \mathbb{R} oder \mathbb{C} ist bezüglich jeder Norm vollständig.

Beweis. Gemäß Beispiel 2.12 ist V isometrisch isomorph zu \mathbb{K}^n mit einer geeigneten Norm und gemäß Bemerkung 2.10 (d) reicht es dessen Vollständigkeit nachzuweisen. Mit Bemerkung 2.14 genügt es, die Vollständigkeit von $(\mathbb{K}^n, \|\cdot\|_\infty)$ nachzuweisen. Sei dazu (x_k) eine Cauchyfolge in $(\mathbb{K}^n, \|\cdot\|_\infty)$, dann sind wegen $|x_{k,j} - x_{\ell,j}| \leq \|x_k - x_\ell\|_\infty$ auch die Komponentenfolgen $(x_{k,j})$ Cauchyfolgen in \mathbb{K} und somit wegen der Vollständigkeit von \mathbb{K} konvergent. Sei $\lim_{k \rightarrow \infty} x_{k,j} = a_j$ und $a = (a_1, \dots, a_n)$, dann ist $\|x_k - a\|_\infty = \max\{|x_{k,1} - a_1|, \dots, |x_{k,n} - a_n|\} < \varepsilon$ für k hinreichend groß und somit $\lim_{k \rightarrow \infty} x_k = a$. \square

2.16 Definition. Punktweise- und gleichmäßige Konvergenz von Funktionenfolgen

Seien X eine Menge, Y ein metrischer Raum und

$$f_n : X \rightarrow Y, n \in \mathbb{N}, \quad \text{sowie} \quad f : X \rightarrow Y \quad \text{Funktionen.}$$

(a) Die Folge (f_n) heißt **punktweise konvergent** gegen f , falls

$$\forall x \in X : \lim_{n \rightarrow \infty} d(f_n(x), f(x)) = 0,$$

oder, ausführlich,

$$\forall x \in X \quad \forall \varepsilon > 0 \quad \exists N = N(\varepsilon, x) \in \mathbb{N} \quad \forall n \geq N : d(f_n(x), f(x)) < \varepsilon.$$

(b) Die Folge (f_n) heißt **gleichmäßig konvergent** gegen f , falls

$$\lim_{n \rightarrow \infty} \sup_{x \in X} d(f_n(x), f(x)) = 0,$$

oder, ausführlich,

$$\forall \varepsilon > 0 \quad \exists N = N(\varepsilon) \in \mathbb{N} \quad \forall x \in X \quad \forall n \geq N : d(f_n(x), f(x)) < \varepsilon.$$

Eine punktweise konvergente Funktionenfolge konvergiert also gleichmäßig, wenn man N unabhängig von $x \in X$ wählen kann. \diamond

2.17 Bemerkung. Für einen normierten Raum $(Y, \|\cdot\|)$ ist gleichmäßige Konvergenz von $f_n : X \rightarrow Y$ gegen f gleichbedeutend mit Konvergenz in der Supremumsnorm. Also $f_n \rightarrow f$ gleichmäßig, genau dann wenn

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0.$$

2.18 Satz. Gleichmäßige Grenzwerte stetiger Funktionen sind stetig

Sei (f_n) eine Folge stetiger Funktionen $f_n : X \rightarrow Y$ zwischen metrischen Räumen X und Y die gleichmäßig gegen $f : X \rightarrow Y$ konvergiert. Dann ist auch f stetig.

Beweis. Sei (x_k) eine Folge in X mit $\lim x_k = a$. Zu zeigen ist, dass $\lim_{k \rightarrow \infty} f(x_k) = f(a)$ gilt. Sei dazu $\varepsilon > 0$ beliebig. Da $f_n \rightarrow f$ gleichmäßig, existiert ein $N \in \mathbb{N}$ so, dass für alle $x \in X$ gilt $d(f_N(x), f(x)) < \frac{\varepsilon}{3}$. Da f_N stetig ist, existiert ein $K \in \mathbb{N}$ so, dass für alle $k \geq K$ gilt $d(f_N(x_k), f_N(a)) < \frac{\varepsilon}{3}$. Dann gilt mit der Dreiecksungleichung, dass für alle $k \geq K$

$$d(f(x_k), f(a)) \leq d(f(x_k), f_N(x_k)) + d(f_N(x_k), f_N(a)) + d(f_N(a), f(a)) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

also $\lim f(x_k) = f(a)$, d.h. f ist stetig in a . \square

2.19 Bemerkung. Der vorausgegangene Beweis ist ein typisches “ $\frac{\varepsilon}{3}$ -Argument” nach folgendem Schema

$$(*) \quad \begin{array}{ccc} f(x_k) & \xrightarrow{\varepsilon} & f(a) \\ \frac{\varepsilon}{3} \downarrow & & \uparrow \frac{\varepsilon}{3} \\ f_N(x_k) & \xrightarrow{\varepsilon/3} & f_N(a) \end{array}$$

Im Schritt (*) braucht man die gleichmäßige Konvergenz, da $|f(x_k) - f_N(x_k)| < \frac{\varepsilon}{3}$ für alle x_k mit $k \geq K$ gelten muss. \diamond

2.20 Korollar. Sei (X, d) ein metrischer Raum, $(Y, \|\cdot\|)$ ein vollständiger normierter Raum und $C(X, Y)$ der Raum der stetigen beschränkten Funktionen von X nach Y mit der Supremumsnorm $\|f\|_\infty := \sup_{x \in X} \|f(x)\|$. Der normierte Raum $(C(X, Y), \|\cdot\|_\infty)$ ist vollständig.

Beweis. Sei (f_n) eine $\|\cdot\|_\infty$ -Cauchyfolge in $C(X, Y)$. Dann ist $(f_n(x))$ für jedes $x \in X$ eine Cauchyfolge in Y , denn $\|f_n(x) - f_m(x)\| \leq \|f_n - f_m\|_\infty$. Da Y vollständig ist, konvergiert $(f_n(x))$ in Y und wir nennen den Grenzwert $f(x)$. Es konvergiert also (f_n) punktweise gegen die so definierte Funktion f . Wenn wir zeigen, dass $f_n \rightarrow f$ gleichmäßig, dann folgt, mit Satz 2.18, dass f stetig ist und somit $f \in C(X, Y)$ und $\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0$.

Sei also $\varepsilon > 0$ beliebig. Da (f_n) eine $\|\cdot\|_\infty$ -Cauchyfolge ist, existiert ein $n_1 \in \mathbb{N}$ so, dass $\|f_n - f_m\|_\infty < \frac{\varepsilon}{2}$ für alle $n, m \geq n_1$. Zu jedem $x \in X$ sei nun $n_2(x) \geq n_1$ so gewählt, dass $\|f_{n_2}(x) - f(x)\| < \frac{\varepsilon}{2}$ ist (f_n konvergiert ja punktweise gegen f). Dann gilt für alle $n \geq n_1$ und alle $x \in X$

$$\|f_n(x) - f(x)\| \leq \|f_n(x) - f_{n_2}(x)\| + \|f_{n_2}(x) - f(x)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also $f_n \rightarrow f$ gleichmäßig. □

2.21 Merke. (a) Gleichmäßig konvergente Folgen stetiger Funktionen konvergieren gegen stetige Funktionen.

(b) Räume stetiger Funktionen mit Werten in Banachräumen und versehen mit der Supremumsnorm sind vollständig, also selbst Banachräume.

2.22 Satz. Banachscher Fixpunktsatz

Sei X ein vollständiger metrischer Raum, A eine abgeschlossene Teilmenge und $f : A \rightarrow A$ eine **Kontraktion**, d.h. es gibt eine Konstante $0 < \theta < 1$, so dass

$$d(f(x), f(y)) \leq \theta d(x, y) \quad \text{für alle } x, y \in A.$$

Dann hat f genau einen **Fixpunkt** $a \in A$, d.h. es gilt $f(a) = a$ für genau ein $a \in A$.

Beweis. Eindeutigkeit: Es kann höchstens einen Fixpunkt geben, denn ist $f(a) = a$ und $f(b) = b$, so ist $d(a, b) = d(f(a), f(b)) \leq \theta d(a, b)$, also $d(a, b) = 0$ und somit $a = b$.

Existenz: Sei $x_0 \in A$ beliebig. Betrachte die Iterationsfolge $x_{n+1} := f(x_n)$ für $n \geq 0$.

Wir zeigen, dass (x_n) eine Cauchyfolge ist. Zunächst ist (x_n) beschränkt, da für alle $n \in \mathbb{N}$ gilt

$$\begin{aligned} d(x_n, x_0) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \cdots + d(x_1, x_0) \\ &\leq \theta d(x_{n-1}, x_{n-2}) + \theta d(x_{n-2}, x_{n-3}) + \cdots + d(x_1, x_0) \\ &\leq (\theta^{n-1} + \theta^{n-2} + \cdots + 1) d(x_1, x_0) \\ &= \frac{1 - \theta^n}{1 - \theta} d(x_1, x_0) \leq \frac{1}{1 - \theta} d(x_1, x_0) =: M \end{aligned}$$

und damit $d(x_m, x_n) \leq 2M$ für alle $n, m \in \mathbb{N}$. Sei nun $\varepsilon > 0$ beliebig und $n, m \geq n_0$, so ist

$$\begin{aligned} d(x_n, x_m) &= d(f(x_{n-1}), f(x_{m-1})) \leq \theta d(x_{n-1}, x_{m-1}) \leq \dots \\ &\leq \theta^{n_0} d(x_{n-n_0}, x_{m-n_0}) \leq 2M \theta^{n_0} < \varepsilon \end{aligned}$$

wenn $n_0 = n_0(\varepsilon)$ groß genug ist.

Aufgrund der Vollständigkeit von X konvergiert die Cauchyfolge (x_n) gegen ein $a \in X$. Da alle x_n in A liegen und A abgeschlossen ist, muss auch $a \in A$ sein (vgl. Satz 1.26). Schließlich gilt wegen der Stetigkeit von f in a (jede Kontraktion ist insbesondere Lipschitz-stetig mit Lipschitzkonstante $L = \theta$), dass

$$f(a) = f\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = a.$$

Also ist a ein Fixpunkt. □

3 Kompaktheit und Zusammenhang

In Analysis 1 wurden kompakte Teilmengen von \mathbb{R} als die abgeschlossenen und beschränkten Teilmengen definiert. Aus dieser Definition ergaben sich beispielsweise der Satz von Bolzano-Weierstraß (jede Folge in einem Kompaktum K hat eine in K konvergente Teilfolge) oder der Satz vom Maximum (stetige reellwertige Funktionen auf Kompakta sind beschränkt und nehmen ihr Supremum an).

In allgemeinen metrischen bzw. topologischen Räumen müssen wir Kompaktheit etwas vorsichtiger definieren, um all die schönen Eigenschaften kompakter Mengen weiterhin sicherzustellen. Wir werden zwar sehen, dass nach der neuen Definition kompakte Mengen immer noch abgeschlossen und beschränkt sind, aber nur noch in speziellen Räumen (z.B. \mathbb{R}^n und \mathbb{C}^n) jede abgeschlossene und beschränkte Menge auch tatsächlich kompakt ist.

3.1 Definition. Offene Überdeckung und kompakte Menge

Sei X ein topologischer Raum und $Y \subset X$ eine Teilmenge.

- (a) Eine Familie von Teilmengen $U_i \subset X$, $i \in \mathcal{I}$, heißt eine **offene Überdeckung von Y** , wenn jedes U_i offen ist und

$$Y \subset \bigcup_{i \in \mathcal{I}} U_i.$$

- (b) Eine Teilmenge $K \subset X$ heißt **kompakt**, wenn folgendes gilt:

Zu *jeder* offenen Überdeckung $(U_i)_{i \in \mathcal{I}}$ von K gibt es eine **endliche Teilüberdeckung** von K , d.h. es gibt $i_1, \dots, i_n \in \mathcal{I}$ so, dass bereits

$$K \subset U_{i_1} \cup \dots \cup U_{i_n}.$$

Achtung: Es wird nicht gefordert, dass eine endliche offene Überdeckung existiert. Die existiert nämlich immer!

3.2 Beispiele. (a) Jede endliche Teilmenge $K = \{x_1, \dots, x_n\}$ eines topologischen Raumes ist kompakt.

- (b) Das offene Intervall $(0, 1) \subset \mathbb{R}$ ist nicht kompakt: Die offene Überdeckung $(0, 1) = \bigcup_{n=2}^{\infty} (\frac{1}{n}, 1)$ erlaubt offenbar keine endliche Teilüberdeckung.

- (c) $\mathbb{N} \subset \mathbb{R}$ ist ebenfalls nicht kompakt: Die offene Überdeckung $\mathbb{N} \subset \bigcup_{n=1}^{\infty} (n - \frac{1}{2}, n + \frac{1}{2})$ erlaubt offenbar keine endliche Teilüberdeckung.

- (d) Ist (x_n) eine konvergente Folge in X und $a = \lim_{n \rightarrow \infty} x_n$, so ist die Menge

$$K = \{x_n \mid n \in \mathbb{N}\} \cup \{a\} \subset X$$

kompakt.

Beweis. Sei $(U_i)_{i \in \mathcal{I}}$ eine offene Überdeckung von K , so wähle ein i_0 mit $a \in U_{i_0}$. Da $\lim_{n \rightarrow \infty} x_n = a$, enthält U_{i_0} alle Folgenglieder bis auf endlich viele. Die restlichen liegen aber in endlich vielen U_{i_1}, \dots, U_{i_n} und somit ist K kompakt. \square

3 Kompaktheit und Zusammenhang

- (e) Läßt man in (d) den Punkt a weg, betrachtet man also $Y = \{x_n \mid n \in \mathbb{N}\}$, so ist diese Menge nicht notwendigerweise kompakt. Beispielsweise ist $Y = \{\frac{1}{n} \mid n \in \mathbb{N}\}$ nicht kompakt, da die offene Überdeckung durch $U_n := \left(\frac{1}{n} - \frac{1}{2} \left(\frac{1}{n} - \frac{1}{n+1}\right), \frac{1}{n} + \frac{1}{2} \left(\frac{1}{n-1} - \frac{1}{n}\right)\right)$ keine endliche Teilüberdeckung zulässt. (Übungsaufgabe: Falls $\lim x_n = a$, dann ist Y kompakt $\Leftrightarrow a \in Y$).

3.3 Satz. Bolzano-Weierstraß

Ist X ein topologischer Raum und $K \subset X$ kompakt, so besitzt jede Folge in K einen Häufungspunkt in K .

Beweis. Sei (x_n) eine Folge in K , also $x_n \in K$ für alle $n \in \mathbb{N}$. Angenommen (x_n) hat keinen Häufungspunkt in K . Dann gibt es zu jedem $a \in K$ eine Umgebung U_a von a , die nur endlich viele Folgenglieder enthält. Es ist

$$K \subset \bigcup_{a \in K} U_a.$$

Da K kompakt ist, gibt es $a_1, \dots, a_n \in K$, so dass bereits

$$K \subset U_{a_1} \cup \dots \cup U_{a_n}.$$

Dann liegen aber nur endlich viele Folgenglieder in K , ein Widerspruch! □

In metrischen Räumen gilt auch die Umkehrung:

3.4 Proposition. Charakterisierung kompakter Mengen in metrischen Räumen

Sei K Teilmenge eines metrischen Raumes X . Falls gilt, dass jede Folge in K einen Häufungspunkt in K hat, dann ist K kompakt.

Beweis. Wir lassen den Beweis aus Zeitgründen weg, er ist aber leicht mit den vorhandenen Mitteln zu führen (siehe z.B. Satz 8.5.10 im Buch *Analysis* von Deitmar). □

3.5 Satz. Stetige Bilder kompakter Mengen sind kompakt

Seien X und Y topologische Räume und $f : X \rightarrow Y$ stetig. Mit $K \subset X$ kompakt ist auch das Bild $f(K) \subset Y$ kompakt.

Beweis. Sei $(V_i)_{i \in \mathcal{I}}$ eine offene Überdeckung von $f(K) \subset Y$. Da f stetig ist, sind die Urbilder $U_i := f^{-1}(V_i) \subset X$ offen und $K \subset \bigcup_{i \in \mathcal{I}} U_i$. Da K kompakt ist, gilt $K \subset U_{i_1} \cup \dots \cup U_{i_n}$ für geeignete $i_1, \dots, i_n \in \mathcal{I}$. Dann ist aber auch $f(K) \subset f(U_{i_1}) \cup \dots \cup f(U_{i_n}) \subset V_{i_1} \cup \dots \cup V_{i_n}$ und somit ist auch $f(K)$ kompakt. □

3.6 Definition. Beschränkte Mengen und ihr Durchmesser

Sei X ein metrischer Raum.

- (a) Eine Teilmenge $B \subset X$ heißt **beschränkt**, wenn es ein $C \in \mathbb{R}$ gibt, so dass für alle $x, y \in B$ gilt: $d(x, y) \leq C$.
- (b) Für eine beliebige Teilmenge $Y \subset X$ definiert man den **Durchmesser von Y** durch

$$\text{diam}(Y) := \sup\{d(x, y) \mid x, y \in Y\} \in [0, \infty) \cup \{\infty\}.$$

Also ist $B \subset X$ genau dann beschränkt, wenn $\text{diam}(B) < \infty$ ist. ◇

3.7 Proposition. Charakterisierung beschränkter Mengen

Sei X ein metrischer Raum und $A \subset X$. Dann sind folgende Aussagen äquivalent:

- (i) A ist beschränkt.
- (ii) Es gibt ein $x_0 \in X$ und ein $R > 0$ so, dass $A \subset B_R(x_0)$.
- (iii) Zu jedem $x_0 \in X$ existiert ein $R > 0$ so, dass $A \subset B_R(x_0)$.

Eine Teilmenge $A \subset V$ eines normierten Raums ist also genau dann beschränkt, wenn $\sup_{v \in A} \|v\| < \infty$.

Beweis. (iii) \Rightarrow (ii) ist klar. (ii) \Rightarrow (i) ebenfalls, da aufgrund der Dreiecksungleichung für $x, y \in A$ gilt, dass $d(x, y) \leq d(x, x_0) + d(x_0, y) < 2R$. Für (i) \Rightarrow (iii) sei A beschränkt und nichtleer und $x_0 \in X$ beliebig. Zu einem beliebigen $x \in A$ wähle $R > \text{diam}(A) + d(x, x_0)$. Dann ist für $y \in A$ wieder $d(y, x_0) \leq d(y, x) + d(x, x_0) \leq \text{diam}(A) + d(x, x_0) < R$, also $y \in B_R(x_0)$. \square

3.8 Satz. Kompakta sind abgeschlossen und beschränkt

Eine kompakte Teilmenge K eines metrischen Raumes X ist abgeschlossen und beschränkt.

Warnung: Die Umkehrung gilt im Allgemeinen nicht! (vgl. Satz 3.10)

Beweis. Beschränktheit: Sei $p \in X$ beliebig und fest. Die offenen Kugeln $U_n := B_n(p)$ um p mit Radius $n \in \mathbb{N}$ bilden eine offene Überdeckung von ganz X und somit auch von K . Also gibt es ein $n_0 \in \mathbb{N}$, so dass

$$K \subset U_1 \cup \dots \cup U_{n_0} = U_{n_0} = B_{n_0}(p).$$

Damit ist für alle $x, y \in K$

$$d(x, y) \leq d(x, p) + d(p, y) < n_0 + n_0 = 2n_0,$$

also ist K beschränkt.

Abgeschlossenheit: Sei (x_n) eine konvergente Folge in K , also $\lim_{n \rightarrow \infty} x_n = a \in X$ und $x_n \in K$ für alle $n \in \mathbb{N}$. Nach Bolzano-Weierstraß hat (x_n) einen Häufungspunkt \tilde{a} in K . Da eine konvergente Folge in einem metrischen Raum aber außer dem Grenzwert keine weiteren Häufungspunkte haben kann (Denn für jeden Punkt $b \neq a$ gilt, dass $B_{d(a,b)/2}(b)$ nur endlich viele Folgenglieder enthalten kann, da $B_{d(a,b)/2}(a)$ alle bis auf endlich viele enthält und $B_{d(a,b)/2}(b) \cap B_{d(a,b)/2}(a) = \emptyset$), gilt $a = \tilde{a}$, somit $a \in K$ und gemäß Satz 1.26 ist K dann abgeschlossen. \square

3.9 Proposition. Abgeschlossene Teilmengen kompakter Mengen sind kompakt

Sei X ein topologischer Raum und $K \subset X$ kompakt. Jede abgeschlossene Teilmenge $A \subset K$ von K ist ebenfalls kompakt.

Beweis. Sei $(U_i)_{i \in \mathcal{I}}$ eine offene Überdeckung von A . Setzt man $U := X \setminus A$, dann ist U offen und $(U, U_i)_{i \in \mathcal{I}}$ ist offene Überdeckung von K . Da K kompakt ist, existiert eine endliche Teilüberdeckung $(U, U_{i_1}, \dots, U_{i_n})$ von K . Da $A \subset K \setminus U$, ist

$$A \subset U_{i_1} \cup \dots \cup U_{i_n}$$

und somit ist auch A kompakt. \square

3.10 Satz. Heine-Borel

Eine Teilmenge K des euklidischen Raums \mathbb{R}^n ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis. " \Rightarrow ": Gilt nach Satz 3.8 in beliebigen metrischen Räumen.

" \Leftarrow ": Sei $K \subset \mathbb{R}^n$ beschränkt und abgeschlossen. Dann gibt es ein $R > 0$ so, dass K im Würfel W_R mit Kantenlänge $2R$ enthalten ist, $K \subset [-R, R]^n =: W_R$. Wir zeigen in Satz 3.13, dass W_R kompakt ist, also ist mit Proposition 3.9 auch K kompakt. \square

3 Kompaktheit und Zusammenhang

3.11 Korollar. Eine Teilmenge K eines endlichdimensionalen normierten Raums ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis. Für Teilmengen des \mathbb{C}^n folgt die Aussage sofort aus Beispiel 2.12 (b) und Satz 3.10. Für \mathbb{K}^n versehen mit einer beliebigen Norm gilt die Aussage ebenfalls, da ja alle Normen auf \mathbb{K}^n äquivalent sind. Denn gemäß Proposition 1.33 hängt Offenheit und damit auch Kompaktheit einer Menge in \mathbb{K}^n dann nicht von der Wahl der Norm ab und gleiches gilt für Abgeschlossenheit und Beschränktheit. Allgemeine endlichdimensionale normierte Räume sind schließlich nach Bemerkung 2.12 (a) isometrisch isomorph zu einem \mathbb{K}^n mit geeigneter Norm und die Eigenschaften Kompaktheit, Abgeschlossenheit und Beschränktheit bleiben unter Isometrien alle erhalten. \square

3.12 Proposition. Schachtelungsprinzip

Sei X ein vollständiger metrischer Raum. Zu jeder Folge $(A_n)_{n \in \mathbb{N}}$ nichtleerer, abgeschlossener Teilmengen mit $A_{n+1} \subset A_n$ für alle $n \in \mathbb{N}$ und $\lim_{n \rightarrow \infty} \text{diam}(A_n) = 0$ existiert genau ein $a \in X$ welches in allen A_n liegt, also

$$\bigcap_{n \in \mathbb{N}} A_n = \{a\}$$

erfüllt.

Beweis. Eindeutigkeit: Seien $a, b \in \bigcap_{n \in \mathbb{N}} A_n$. Dann ist $d(a, b) \leq \text{diam}(A_n) \xrightarrow{n \rightarrow \infty} 0$, also $d(a, b) = 0$ und somit $a = b$.

Existenz: Da alle A_n nichtleer sind, gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in A_n$. Wir zeigen, dass (x_n) eine Cauchy-Folge ist: Sei $\varepsilon > 0$ und $n_0 \in \mathbb{N}$ so groß, dass $\text{diam}(A_{n_0}) < \varepsilon$ ist. Dann gilt für alle $n, m \geq n_0$

$$d(x_n, x_m) < \varepsilon \quad \text{da} \quad x_n \in A_n \subset A_{n_0} \quad \text{und} \quad x_m \in A_m \subset A_{n_0}.$$

Da X vollständig ist, konvergiert (x_n) gegen ein $a \in X$. Da für jedes $n \in \mathbb{N}$ gilt, dass $(x_k)_{k=n}^{\infty}$ in A_n liegt, und da jedes A_n abgeschlossen ist, folgt $a \in A_n$ für alle $n \in \mathbb{N}$ und somit $a \in \bigcap_{n \in \mathbb{N}} A_n$. \square

3.13 Satz. Sei $R > 0$ und $W = \{x \in \mathbb{R}^n \mid \|x\|_{\infty} \leq R\}$ der abgeschlossene Würfel der Kantenlänge $2R$. Dann ist W kompakt.

Beweis. Sei $(U_i)_{i \in \mathcal{I}}$ eine offene Überdeckung von W und nehme an, dass es keine endliche Teilüberdeckung gibt. Zerlege $W_0 := W$ in 2^n abgeschlossene Würfel der Kantenlänge R . Dann erlaubt auch einer dieser kleineren Würfel keine endliche Teilüberdeckung. Diesen nennen wir W_1 und konstruieren so eine Folge W_k abgeschlossener Würfel der Kantenlänge $2^{-k} \cdot 2R$, die alle keine endliche Teilüberdeckung erlauben. Nach Konstruktion gilt dann $W_{k+1} \subset W_k$ und $\lim_{k \rightarrow \infty} \text{diam}(W_k) \leq \lim_{k \rightarrow \infty} 2^{-k} \cdot 2R = 0$. Mit dem Schachtelungsprinzip 3.12 und der Vollständigkeit von \mathbb{R}^n folgt $\bigcap_k W_k = \{a\}$. Da $a \in W \subset \bigcup_{i \in \mathcal{I}} U_i$, gibt es ein $i_0 \in \mathcal{I}$ mit $a \in U_{i_0}$. Da U_{i_0} offen ist, existiert ein $\varepsilon > 0$ so, dass $B_{\varepsilon}(a) \subset U_{i_0}$. Sei nun $k_0 \in \mathbb{N}$ so groß, dass $\text{diam}(W_{k_0}) < \varepsilon$, also $W_{k_0} \subset B_{\varepsilon}(a) \subset U_{i_0}$. Aber dann wird W_{k_0} schon von einem einzigen der U_i überdeckt, obwohl es nach der Konstruktion nicht mal von endlich vielen überdeckt wird. Widerspruch! \square

Nachdem wir nun die kompakten Teilmengen endlichdimensionaler normierter Räume charakterisiert haben, wollen wir an dieser Stelle auch noch ein wichtiges Resultat zu kompakten Mengen in Räumen stetiger Funktionen auf kompakten Definitionsbereichen besprechen. Auf unendlichdimensionalen Räumen gilt, wie gesagt, nicht mehr, dass jede abgeschlossene und beschränkte Menge auch kompakt ist. Die zusätzliche notwendige Bedingung ist für Räume stetiger Funktionen die gleichgradige Stetigkeit.

3.14 Definition. Gleichgradige Stetigkeit

Seien X und Y metrische Räume und $A \subset C(X, Y)$ eine Menge stetiger Funktionen von X nach Y . Es heißt A gleichgradig stetig im Punkt $x \in X$, falls

Für alle $\varepsilon > 0$ existiert ein $\delta > 0$ so, dass für alle $f \in A$ und $x' \in X$ gilt:

$$d_X(x, x') < \delta \Rightarrow d_Y(f(x), f(x')) < \varepsilon.$$

Es ist also A gleichgradig stetig in x , falls man zu gegebenem $\varepsilon > 0$ für alle $f \in A$ dasselbe δ wählen kann. Es heißt A gleichgradig stetig, falls A an allen Punkten $x \in X$ gleichgradig stetig ist.

3.15 Satz. Arzelà-Ascoli

Sei X ein kompakter metrischer Raum und $C(X, \mathbb{K})$ versehen mit der Supremumsnorm. Eine Teilmenge $K \subset C(X, \mathbb{K})$ ist genau dann kompakt, wenn sie abgeschlossen, punktweise beschränkt (d.h. für alle $x \in X$ gilt $\sup_{f \in K} |f(x)| < \infty$) und gleichgradig stetig ist.

Beweis. Für den Beweis verweisen wir aus Zeitgründen wieder auf Kapitel 8.6 des Buchs Analysis von Anton Deitmar. \square

3.16 Bemerkung. Der Satz von Arzelà-Ascoli spielt in der Analysis eine wichtige Rolle, auch wenn wir ihn in dieser Vorlesung nicht verwenden werden. Er gilt auch deutlich allgemeiner, nämlich für Funktionen auf kompakten Hausdorffräumen X mit Werten in metrischen Räumen Y , falls man gleichgradige Stetigkeit in offensichtlicher Weise verallgemeinert (statt $\delta > 0$ verlangt man die Existenz einer Umgebung U_x von x) und die punktweise Beschränktheit durch die Bedingung ersetzt, dass der Abschluss der Menge $\{f(x) \mid x \in X\} \subset Y$ für jedes $x \in X$ kompakt ist.

3.17 Satz. Weierstraß

Sei $K \subset X$ eine kompakte Teilmenge eines topologischen Raumes und $f : K \rightarrow \mathbb{R}$ eine stetige Abbildung. Dann nimmt f ihr Supremum und ihr Infimum an.

3.18 Bemerkung. Ist $Y \subset X$ beliebig und $f : Y \rightarrow \mathbb{R}$ eine Funktion, so ist i.A.

$$\sup_{x \in Y} f(x) \in \mathbb{R} \cup \{\infty\}, \quad \inf_{x \in Y} f(x) \in \mathbb{R} \cup \{-\infty\}.$$

Dass f ihr Supremum bzw. Infimum annimmt bedeutet, dass es ein $y_0 \in Y$ gibt, so dass

$$\sup_{y \in Y} f(y) = f(y_0) \quad \text{bzw.} \quad \inf_{y \in Y} f(y) = f(y_0).$$

Insbesondere ist dann f nach oben bzw. nach unten beschränkt, denn $f(y_0) \in \mathbb{R}$. \diamond

Beweis. von Satz 3.17: Sei $c := \sup_{x \in K} f(x) \in \mathbb{R} \cup \{\infty\}$. Es gibt dann, nach Definition des Supremums, eine Folge (x_n) in K mit $\lim_{n \rightarrow \infty} f(x_n) = c$. Da K kompakt ist, hat (x_n) einen Häufungspunkt $a \in K$. Also existiert eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$, sodass $\lim_{k \rightarrow \infty} x_{n_k} = a$. Wegen der Stetigkeit von f gilt aber

$$f(a) = \lim_{k \rightarrow \infty} f(x_{n_k}) = c.$$

Der Übergang von f zu $-f$ zeigt, dass auch das Infimum angenommen wird. \square

In der Analysis 1 waren die natürlichen Definitionsbereiche für Funktionen die Intervalle, insbesondere weil diese "zusammenhängend" sind. Funktionen die auf mehreren disjunkten Intervallen definiert sind (oder auf anderen "unzusammenhängenden" Teilmengen von \mathbb{R}) zerfallen in "unabhängige" Teile. Im Folgenden diskutieren wir das entsprechende Konzept von zusammenhängenden Mengen in allgemeinen topologischen Räumen.

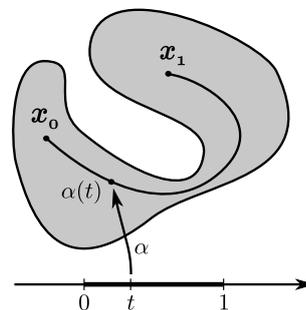
3.19 Definition. Zusammenhang

Sei X ein topologischer Raum.

Es heißt X **zusammenhängend**, falls es keine zwei nichtleeren offenen Mengen $O_1, O_2 \subset X$ gibt, so dass $O_1 \cap O_2 = \emptyset$ und $X = O_1 \cup O_2$.

Es heißt X **wegzusammenhängend**, wenn es zu je zwei Punkten $x_0, x_1 \in X$ einen **Weg** von x_0 nach x_1 in X gibt, d.h. eine stetige Abbildung $\alpha : [0, 1] \rightarrow X$ mit $\alpha(0) = x_0$ und $\alpha(1) = x_1$.

Eine Teilmenge $Y \subset X$ eines topologischen Raumes heißt **zusammenhängend** bzw. **wegzusammenhängend**, wenn sie als topologischer Raum mit der induzierten Topologie zusammenhängend bzw. wegzusammenhängend ist.



3.20 Bemerkung. (a) Auf jeder Teilmenge $Y \subset X$ eines topologischen Raumes X wird durch die Topologie des umgebenden Raumes eine Topologie induziert, die sogenannte **Relativtopologie**: Eine Teilmenge $U \subset Y$ ist offen bezüglich der Relativtopologie auf Y (und heißt dann relativ-offen), wenn es eine offene Menge $O \subset X$ gibt, sodass $U = O \cap Y$.

(b) Ein topologischer Raum X ist genau dann zusammenhängend, wenn keine nichtleere echte Teilmenge von X sowohl offen als auch abgeschlossen ist.

Beweis. Es existieren nichtleere offene Mengen $O_1, O_2 \subset X$ mit $O_1 \cap O_2 = \emptyset$ und $X = O_1 \cup O_2$ genau dann, wenn es ein offenes $O_1 \subsetneq X$ gibt, sodass $O_2 = X \setminus O_1$ ebenfalls offen ist. \square

3.21 Satz. Seien X und Y topologische Räume und sei $f : X \rightarrow Y$ stetig. Dann gilt: Ist $A \subset X$ (weg)zusammenhängend, so ist auch $f(A) \subset Y$ (weg)zusammenhängend.

Beweis. A **zusammenhängend**: Wir zeigen die Kontraposition. Angenommen, $f(A)$ ist nicht zusammenhängend. Dann existieren zwei disjunkte offene Mengen $U_1, U_2 \subset Y$ so, dass $f(A) \subset U_1 \cup U_2$, $f(A) \cap U_1 \neq \emptyset$ und $f(A) \cap U_2 \neq \emptyset$. Da f stetig ist, sind $O_1 := f^{-1}(U_1)$ und $O_2 := f^{-1}(U_2)$ offene nichtleere Mengen in X . Außerdem erfüllt die Urbildabbildung offenbar für beliebige Teilmengen $B, C \subset Y$ und $D \subset X$, dass $f^{-1}(B \cup C) = f^{-1}(B) \cup f^{-1}(C)$, $f^{-1}(B \cap C) = f^{-1}(B) \cap f^{-1}(C)$ und $D \subset f^{-1}(f(D))$. Also sind $O_1, O_2 \subset X$ nichtleere offene Mengen mit

$$A \subset f^{-1}(f(A)) \subset f^{-1}(U_1 \cup U_2) = f^{-1}(U_1) \cup f^{-1}(U_2) = O_1 \cup O_2$$

und

$$O_1 \cap O_2 = f^{-1}(U_1) \cap f^{-1}(U_2) = f^{-1}(U_1 \cap U_2) = f^{-1}(\emptyset) = \emptyset.$$

Sei $y \in f(A) \cap U_1 \neq \emptyset$. Dann existiert ein $x \in A$ mit $y = f(x) \in U_1$, d.h. $x \in f^{-1}(U_1) = O_1$. Also ist $A \cap O_1 \neq \emptyset$ und analog $A \cap O_2 \neq \emptyset$. Insgesamt folgt, dass auch A nicht zusammenhängend ist.

A **wegzusammenhängend**: Seien $y_0, y_1 \in f(A)$ beliebig und $x_0, x_1 \in A$ so, dass $f(x_0) = y_0$ und $f(x_1) = y_1$. Dann gibt es nach Annahme einen stetigen Weg $\alpha : [0, 1] \rightarrow A$ mit $\alpha(0) = x_0$ und $\alpha(1) = x_1$. Dann ist auch $\beta = f \circ \alpha$ stetig und damit ein Weg von $(f \circ \alpha)(0) = f(x_0) = y_0$ nach $(f \circ \alpha)(1) = f(x_1) = y_1$. Also ist $f(A)$ wegzusammenhängend. \square

3.22 Proposition. Wegzusammenhängend impliziert zusammenhängend

Jede wegzusammenhängende Teilmenge eines topologischen Raumes ist auch zusammenhängend.

Beweis. Der Beweis ist eine Übungsaufgabe und verwendet, dass Intervalle in \mathbb{R} zusammenhängend sind. \square

3.23 Bemerkung. Zusammenhängend impliziert nicht wegzusammenhängend

Im Allgemeinen sind zusammenhängende Mengen nicht notwendigerweise auch wegzusammenhängend. So ist beispielsweise die Menge

$$M := \{(x, y) \in \mathbb{R}^2 \mid x > 0, y = \sin(1/x)\} \cup \{(x, y) \in \mathbb{R}^2 \mid x = 0\}$$

zusammenhängend, aber nicht wegzusammenhängend. \diamond

3.24 Proposition. Für offene Teilmengen eines normierten Raumes V sind die Eigenschaften zusammenhängend und wegzusammenhängend äquivalent.

Beweis. Mit Proposition 3.22 ist eine Richtung bereits gezeigt. Sei nun $O \subset V$ offen und zusammenhängend und $p \in O$ beliebig. Setze

$$U_p := \{x \in O \mid \text{es existiert ein stetiger Weg } \alpha : [0, 1] \rightarrow O \text{ von } p \text{ nach } x\} \subset O.$$

Wir zeigen, dass U_p in O offen und abgeschlossen ist. Dann folgt mit Bemerkung 3.20 (a), dass $U_p = O$ und somit, dass O auch wegzusammenhängend ist.

Sei also $x \in U_p$. Da O offen ist, existiert ein $\delta > 0$ so, dass $B_\delta(x) \subset O$. Es gilt aber auch $B_\delta(x) \subset U_p$, da sich x mit jedem Punkt $y \in B_\delta(x)$ durch den stetigen Weg $\gamma_{x,y} : [0, 1] \rightarrow B_\delta(x)$, $\gamma_{x,y}(t) = x + t(y - x)$ verbinden lässt. Also ist U_p offen in V und somit auch offen in O . Sei nun (x_k) eine in O konvergente Folge in U_p mit $\lim_{k \rightarrow \infty} x_k = a \in O$. Es folgt auch die Abgeschlossenheit von U_p in O , sobald wir $a \in U_p$ gezeigt haben. Da O offen ist, existiert ein $\delta > 0$ so, dass $B_\delta(a) \subset O$. Für $k_0 \in \mathbb{N}$ groß genug liegt x_{k_0} in $B_\delta(a)$. Mit demselben Argument wie zuvor lässt sich x_{k_0} mit a durch einen stetigen Weg in $B_\delta(a)$ verbinden, woraus wegen $x_{k_0} \in U_p$ auch $a \in U_p$ folgt. \square

3.25 Bemerkungen. (a) Wir haben im vorangegangenen Beweis implizit verwendet, dass sich zwei stetige Wege $\alpha_1 : [0, 1] \rightarrow O$ und $\alpha_2 : [0, 1] \rightarrow O$ in einer Menge O mit $\alpha_1(1) = \alpha_2(0)$ zu einem stetigen Weg β von $\alpha_1(0)$ nach $\alpha_2(1)$ in O zusammensetzen lassen,

$$\beta : [0, 1] \rightarrow O, \quad \beta(t) := \begin{cases} \alpha_1(2t) & \text{falls } t \in [0, \frac{1}{2}] \\ \alpha_2(2t - 1) & \text{falls } t \in (\frac{1}{2}, 1] \end{cases}.$$

(b) Weiterhin genügt es im vorangegangenen Beweis zu fordern, dass jeder Punkt im Raum V eine wegzusammenhängende Umgebung hat. Daher gilt das vorangegangene Resultat in jedem topologischen Raum in dem jeder Punkt eine wegzusammenhängende Umgebung besitzt, also in sogenannten lokal wegzusammenhängende Räumen. Die Menge M aus Bemerkung 3.23 ist nicht lokal wegzusammenhängend, da keiner der Punkte in $\{(x, y) \mid x = 0 \text{ und } y \in [-1, 1]\} \subset M$ eine wegzusammenhängende Umgebung besitzt.

Normierte Räume sind aber lokal wegzusammenhängend, da Kugeln konvex sind: Seien $x, y \in B_r(a)$, dann gilt für $t \in [0, 1]$, dass

$$\|x + t(y - x) - a\| = \|(1 - t)(x - a) + t(y - a)\| \leq (1 - t)\|x - a\| + t\|y - a\| < (1 - t)r + tr = r, \\ \text{also } x + t(y - x) \in B_r(a) \text{ für } t \in [0, 1].$$

3.26 Definition. Gebiet

Sei X ein topologischer Raum. Eine nichtleere Teilmenge $U \subset X$ heißt **Gebiet in X** , falls U offen und zusammenhängend ist.

Mit Proposition 3.24 folgt, dass Gebiete im \mathbb{R}^n auch wegzusammenhängend sind. \diamond

Im Folgenden wird oft aus rein praktischen Gründen vorausgesetzt, dass Funktionen auf Gebieten definiert sind.

3.27 Bemerkung. Die Gebiete in \mathbb{R} sind genau die offenen Intervalle $(a, b) \subset \mathbb{R}$ mit $a < b$, $a \in \{-\infty\} \cup \mathbb{R}$ und $b \in \mathbb{R} \cup \{\infty\}$ (Übungsaufgabe).

4 Differenzierbarkeit

Für Funktionen einer reellen Variable mit Werten in einem normierten Raum, also $f : I \rightarrow W$, $I \subset \mathbb{R}$ ein offenes Intervall und $(W, \|\cdot\|)$ ein normierter \mathbb{K} -Vektorraum, definiert man die Ableitung an der Stelle $x \in I$

$$f'(x) := \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad \left(= \lim_{\tilde{x} \rightarrow x} \frac{f(\tilde{x}) - f(x)}{\tilde{x} - x} \right)$$

genau wie für reell- bzw. komplexwertige Funktionen (vgl. Analysis 1). Zur Erinnerung: Der Grenzwert $\lim_{h \rightarrow 0}$ existiert genau dann, wenn für alle Nullfolgen $(h_k)_{k \in \mathbb{N}}$ in $\mathbb{R} \setminus \{0\}$ mit $x + h_k \in I$ für alle $k \in \mathbb{N}$ die Folge

$$\left(\frac{f(x + h_k) - f(x)}{h_k} \right)_{k \in \mathbb{N}}$$

konvergiert, und zwar als Folge in dem normierten Raum $(W, \|\cdot\|)$. Der Grenzwert ist dann notwendigerweise für all diese Folgen gleich und wird mit $f'(x)$ bezeichnet. Beachte, dass $f'(x) \in W$ ein Vektor ist und, dass der Zähler $f(x+h) - f(x)$ im Differenzenquotienten überhaupt nur für vektorwertige Funktionen Sinn macht.

Man nennt Funktionen $f : I \rightarrow W$ die ein Intervall in einen Vektorraum (oder auch andere Mengen) abbilden meist **Kurven**. Stellen Sie sich dazu $\text{Bild } f := \{f(x) \in W \mid x \in I\}$ für ein stetiges f vor. Falls f differenzierbar ist, so nennt man den Vektor $f'(x)$ Tangentialvektor oder auch Geschwindigkeitsvektor an die Kurve f im Punkt $f(x)$.

Wir werden uns in diesem Kapitel mit der Frage beschäftigen, wie man Differenzierbarkeit für Funktionen definiert, deren Definitionsbereich eine Teilmenge eines reellen Vektorraums V ist. Um ein Gefühl für die Problematik zu bekommen (und weil nach Wahl einer Basis zumindest jeder endlichdimensionale reelle Vektorraum mit dem \mathbb{R}^n identifiziert werden kann), betrachten wir zunächst Funktionen auf dem \mathbb{R}^n .

4.1 Definition. Partielle Ableitung und Gradient

Sei $n \geq 1$, $G \subset \mathbb{R}^n$ offen und $(W, \|\cdot\|)$ ein normierter Raum. Für $x \in G$ und $j \in \{1, \dots, n\}$ heißt eine Funktion $f : G \rightarrow W$ im **Punkt x in die j -te Koordinatenrichtung partiell differenzierbar**, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x + h e_j) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_j + h, \dots, x_n) - f(x_1, \dots, x_n)}{h}$$

existiert, wobei e_j den j -ten kanonischen Basisvektor des \mathbb{R}^n bezeichnet. Wir schreiben dann

$$\frac{\partial f}{\partial x_j}(x) = D_j f(x) = \partial_j f(x) := \lim_{h \rightarrow 0} \frac{f(x + h e_j) - f(x)}{h}$$

und nennen diesen Vektor die **j -te partielle Ableitung von f in x** . Die drei angegebenen Schreibweisen sind alle gebräuchlich und wir werden sie auch alle verwenden.

Im Fall $W = \mathbb{K}$ nennt man den Vektor $\nabla f(x) = \text{grad } f(x) := (\partial_1 f(x), \dots, \partial_n f(x)) \in \mathbb{K}^n$ den **Gradient** von f and der Stelle x .

4 Differenzierbarkeit

4.2 Bemerkung. Man führt also in dieser Definition die Differenzierbarkeit einer Funktion mehrerer Veränderlicher auf die Situation einer reellen Veränderlichen zurück. Setzt man $g(x_j) := f(x_1, x_2, \dots, x_j, \dots, x_n)$ für festes x_i falls $i \neq j$, so ist $\frac{\partial f}{\partial x_j}(x) = g'(x_j)$. D.h., man berechnet die partielle Ableitung $\frac{\partial f}{\partial x_j}(x)$ indem man alle x_i mit $i \neq j$ als Konstanten behandelt und nur nach der einen reellen Variable x_j differenziert. \diamond

4.3 Beispiele. (a) Die Radiusfunktion

$$r : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto r(x) = \|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}$$

ist für alle $x \in \mathbb{R}^n \setminus \{0\}$ und $j \in \{1, \dots, n\}$ partiell in die j -te Richtung differenzierbar und es gilt

$$\partial_j r(x) = \frac{1}{2} \frac{2x_j}{\sqrt{x_1^2 + \dots + x_n^2}} = \frac{x_j}{\|x\|_2}.$$

An der Stelle $x = 0$ ist r in keine Richtung partiell differenzierbar, da $g(x_j) := r(0, \dots, x_j, \dots, 0) = \sqrt{x_j^2} = |x_j|$ bei $x_j = 0$ nicht differenzierbar ist.

(b) Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x) = \chi_{\mathbb{R} \times \mathbb{Q}}(x) := \begin{cases} 1 & \text{falls } x_2 \in \mathbb{Q} \\ 0 & \text{falls } x_2 \notin \mathbb{Q} \end{cases}$$

ist an jedem Punkt $x \in \mathbb{R}^2$ partiell differenzierbar in Richtung e_1 mit

$$\partial_1 f(x) = \lim_{h \rightarrow 0} \frac{f(x_1 + h, x_2) - f(x_1, x_2)}{h} = 0.$$

Sie ist an keinem Punkt partiell differenzierbar in Richtung e_2 , da die Funktion $x_2 \mapsto g(x_2) := \chi_{\mathbb{Q}}(x_2)$ nirgends differenzierbar ist. \diamond

4.4 Definition. Partielle und stetige partielle Differenzierbarkeit

Sei $G \subset \mathbb{R}^n$ offen, $(W, \|\cdot\|)$ ein normierter Raum und $f : G \rightarrow W$.

- Es heißt f **partiell differenzierbar im Punkt** $x \in G$, falls f bei x in allen Koordinatenrichtungen partiell differenzierbar ist.
- Es heißt f **partiell differenzierbar in** G , falls f in allen Punkten in G partiell differenzierbar ist.
- Es heißt f **stetig partiell differenzierbar in** G , falls f partiell differenzierbar in G ist und die partiellen Ableitungen $\partial_j f : G \rightarrow W$, $j = 1, \dots, n$, stetig sind.

4.5 Bemerkung. Man beachte, dass die partielle Differenzierbarkeit einer Funktion $f : G \rightarrow \mathbb{R}$ in einem Punkt $x \in G$ im Allgemeinen nicht, wie im Fall $n = 1$, die Stetigkeit von f in x nach sich zieht. Betrachte z.B. die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = \begin{cases} 0 & \text{wenn } xy = 0 \\ 1 & \text{wenn } xy \neq 0. \end{cases}$$

Dann ist f nämlich in $(x, y) = (0, 0)$ partiell differenzierbar mit $\partial_1 f(0) = \partial_2 f(0) = 0$, aber offenbar nicht stetig. Wir werden jedoch sehen, dass jede Funktion f die in einer offenen Menge stetig partiell differenzierbar ist, dort auch stetig ist. \diamond

4.6 Bemerkung. Rechenregeln für partielle Ableitungen

Da der Begriff der partiellen Differenzierbarkeit einer Funktion $f : G \rightarrow W$, $G \subset \mathbb{R}^n$ offen und W ein normierter \mathbb{K} -Vektorraum, auf den der Differenzierbarkeit einer Funktion in einer Veränderlichen zurückgeführt werden kann, gelten auch die bekannten Rechenregeln:

- (a) **Linearität:** Sind $f, g : G \rightarrow W$ (stetig) partiell differenzierbar und $\lambda \in \mathbb{K}$, dann sind auch $f + g : G \rightarrow W$ und $\lambda f : G \rightarrow W$ (stetig) partiell differenzierbar und es gilt für $j \in \{1, \dots, n\}$

$$\begin{aligned}\partial_j(f + g) &= \partial_j f + \partial_j g \\ \partial_j(\lambda f) &= \lambda \partial_j f.\end{aligned}$$

Insbesondere ist also

$$C^1(G) := \{f : G \rightarrow W \mid f \text{ stetig partiell differenzierbar}\}$$

ein \mathbb{K} -Vektorraum.

- (b) **Produktregel:** Sind $f, g : G \rightarrow \mathbb{K}$ (stetig) partiell differenzierbar, so auch ihr Produkt $fg : G \rightarrow \mathbb{K}$ und ihr Quotient $f/g : G \rightarrow \mathbb{K}$ (falls $g \neq 0$ auf G).

Für $j \in \{1, \dots, n\}$ gilt dann

$$\begin{aligned}\partial_j(fg) &= (\partial_j f)g + f(\partial_j g) \\ \partial_j\left(\frac{f}{g}\right) &= \frac{(\partial_j f)g - f(\partial_j g)}{g^2}.\end{aligned}$$

- (c) **Kettenregel:** Ist $f : G \rightarrow \mathbb{R}$ (stetig) partiell differenzierbar und $h : \mathbb{R} \rightarrow \mathbb{R}$ (stetig) differenzierbar, so ist auch $h \circ f : G \rightarrow \mathbb{R}$ (stetig) partiell differenzierbar und es gilt

$$\partial_j(h \circ f) = (h' \circ f) \partial_j f.$$

4.7 Beispiel. Ist $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ eine **rotationssymmetrische Funktion**, d.h. $f(x) = f(y)$ falls $\|x\| = \|y\|$, so gibt es eine Funktion $h : (0, \infty) \rightarrow \mathbb{R}$, so dass $f(x) = h(\|x\|)$ für alle $x \in \mathbb{R}^n \setminus \{0\}$. Wähle z.B. $h(r) = f(r, 0, 0, \dots, 0)$. Es ist also $f = h \circ r$ mit $r(x) = \|x\|$ wie zuvor. Ist nun f partiell differenzierbar, so ist auch h differenzierbar ($h'(r) = \partial_1 f(r, 0, \dots, 0)$) und es gilt für $j \in \{1, \dots, n\}$

$$\begin{aligned}\partial_j f(x) &= \partial_j(h \circ r)(x) = (h' \circ r)(x) \partial_j r(x) \\ &= h'(\|x\|) \frac{x_j}{\|x\|},\end{aligned}$$

also $\text{grad} f(x) = h'(\|x\|) \frac{x}{\|x\|}$. ◇

4.8 Definition. Vektorfeld

Sei $G \subset \mathbb{R}^n$. Eine Abbildung $f : G \rightarrow \mathbb{R}^n$ heißt ein **Vektorfeld** auf G . ◇

4.9 Beispiel. Gradient als Vektorfeld

Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{R}$ partiell differenzierbar. Dann ist $\text{grad} f : G \rightarrow \mathbb{R}^n$ ein Vektorfeld.

4.10 Definition. Divergenz und Laplace

Sei $G \subset \mathbb{R}^n$ offen.

- (a) Für ein partiell differenzierbares Vektorfeld $f : G \rightarrow \mathbb{R}^n$ heißt die Funktion

$$\begin{aligned}\text{div} f : G &\rightarrow \mathbb{R} \\ x &\mapsto \text{div} f(x) = \sum_{j=1}^n \frac{\partial f_j}{\partial x_j}(x)\end{aligned}$$

die **Divergenz** von f .

4 Differenzierbarkeit

- (b) Sei $f : G \rightarrow \mathbb{R}$ partiell differenzierbar und sei auch $\text{grad} f : G \rightarrow \mathbb{R}^n$ partiell differenzierbar, so heißt die Funktion

$$\begin{aligned} \Delta f : G &\rightarrow \mathbb{R} \\ x &\mapsto \Delta f(x) = \text{div}(\text{grad} f)(x) = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j^2} \end{aligned}$$

Laplace von f .

- 4.11 Beispiel.** (a) Die Identität $\text{id} : G \rightarrow \mathbb{R}^n$, $x \mapsto x$, ist ein partiell differenzierbares Vektorfeld auf G . Ihre Divergenz ist

$$\text{div}(\text{id})(x) = \sum_{j=1}^n \frac{\partial \text{id}_j}{\partial x_j}(x) = \sum_{j=1}^n \underbrace{\frac{\partial x_j}{\partial x_j}(x)}_{=1} = n.$$

- (b) Die Abbildung $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n$, $x \mapsto \frac{x}{\|x\|} = \text{grad} r$ ist ein partiell differenzierbares Vektorfeld auf $\mathbb{R}^n \setminus \{0\}$. Die Divergenz ist

$$\begin{aligned} \text{div} f(x) &= \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{x_j}{\|x\|} \right) = \sum_{j=1}^n \left(\frac{1}{\|x\|} - \frac{x_j^2}{\|x\|^3} \right) \\ &= n \frac{1}{\|x\|} - \frac{\|x\|^2}{\|x\|^3} = \frac{n-1}{\|x\|}. \end{aligned}$$

Also gilt auf $\mathbb{R}^n \setminus \{0\}$, dass $\Delta r = \text{div}(\text{grad} r) = \frac{n-1}{r}$

4.12 Definition. Die m -mal stetig partiell differenzierbaren Funktionen $C^m(G, W)$

Sei $G \subset \mathbb{R}^n$ offen, W ein normierter Raum und $m \in \mathbb{N}$. Eine Funktion $f : G \rightarrow W$ heißt **m -mal stetig partiell differenzierbar**, wenn für alle $j = (j_1, \dots, j_m)$ mit $j_1, \dots, j_m \in \{1, \dots, n\}$ gilt:

- f ist stetig partiell differenzierbar
- $\partial_{j_1} f$ ist stetig partiell differenzierbar
- $\partial_{j_2}(\partial_{j_1} f)$ ist stetig partiell differenzierbar.
- ⋮
- $\partial_{j_{m-1}} \cdots \partial_{j_1} f$ ist stetig partiell differenzierbar, also $\partial_{j_m} \cdots \partial_{j_1} f$ ist stetig.

Wiederholte Anwendung von 4.6 (a) liefert, dass

$$C^m(G, W) := \{f : G \rightarrow W \mid f \text{ ist } m\text{-mal stetig partiell differenzierbar}\}$$

ein \mathbb{K} -Vektorraum ist. Im Fall $W = \mathbb{K}$ schreibt man oft nur $C^m(G) := C^m(G, W)$. Welcher Körper \mathbb{K} dann gemeint ist, sollte sich aus dem Kontext ergeben. \diamond

- 4.13 Notation.** Ist f m -mal stetig partiell differenzierbar, so schreibt man auch

$$\frac{\partial^m f}{\partial x_{j_m} \cdots \partial x_{j_1}}(x) = \partial_{j_m} \cdots \partial_{j_1} f(x)$$

für jedes $j = (j_1, \dots, j_m) \in \{1, \dots, n\}^m$. \diamond

4.14 Beispiel. Sei $h \in C^2((0, \infty), \mathbb{R})$ und $r : \mathbb{R}^n \rightarrow [0, \infty)$, $r(x) = \|x\|$, die Radiusfunktion. Es ist $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$, $f := h \circ r$ also eine rotationssymmetrische Funktion.

Dann ist f zweimal stetig partiell differenzierbar und auch $\Delta f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ ist rotationssymmetrisch, d.h. $\Delta f = g \circ r$ für eine stetige Funktion $g : (0, \infty) \rightarrow \mathbb{R}$. Es gilt

$$g(r) = h''(r) + \frac{n-1}{r}h'(r).$$

Beweis. Übungsaufgabe □

4.15 Satz. von Schwarz

Sei $G \subset \mathbb{R}^n$ offen, W ein normierter Raum endlicher Dimension, $f : G \rightarrow W$ zweimal stetig partiell differenzierbar und $1 \leq i, j \leq n$. Dann vertauschen die partiellen Ableitungen, d.h. für alle $a \in G$ gilt

$$\partial_i \partial_j f(a) = \partial_j \partial_i f(a).$$

Beweis. Da jeder m -dimensionale normierte \mathbb{K} -Vektorraum nach Wahl einer Basis isometrisch isomorph zum \mathbb{K}^m mit einer geeigneten Norm ist (vgl. Beispiel 2.12), reicht es die Aussage für $W = \mathbb{K}^m$ zu zeigen. Für Funktionen mit Werten in \mathbb{K}^m können wir allerdings die einzelnen Komponenten getrennt betrachten, denn $\partial_j f = (\partial_j f_1, \dots, \partial_j f_m)$. Und für Funktionen mit Werten in \mathbb{C} können wir Real- und Imaginärteil getrennt betrachten, denn $\partial_j f = \partial_j \operatorname{Re}(f) + i \partial_j \operatorname{Im}(f)$. Insgesamt genügt es also, wenn wir die Aussage des Satzes für reellwertige Funktionen $f : G \rightarrow \mathbb{R}$ beweisen.

Nun seien o.B.d.A. $n = 2$, $i = 1$, $j = 2$ und $a = 0$. Wir wenden wiederholt den Mittelwertsatz aus Analysis 1 für differenzierbare Funktionen einer Variable an (die andere Variable halten wir jeweils fest) und erhalten für x in einer hinreichend kleinen Umgebung von 0, dass

$$\begin{aligned} \underbrace{(f(x_1, x_2) - f(x_1, 0))}_{=: F_1(x_1)} - \underbrace{(f(0, x_2) - f(0, 0))}_{=: F_1(0)} &= F'_1(\alpha) x_1 = \underbrace{(\partial_1 f(\alpha, x_2))}_{=: F_2(x_2)} - \underbrace{(\partial_1 f(\alpha, 0))}_{=: F_2(0)} x_1 \\ &= F'_2(\beta) x_2 x_1 = \partial_2 \partial_1 f(\alpha, \beta) x_1 x_2 \end{aligned}$$

mit $|\alpha(x_1, x_2)| \leq |x_1|$ und $|\beta(x_1, x_2)| \leq |x_2|$. (Die x -Abhängigkeit der Zwischenpunkte α und β haben wir der Übersichtlichkeit halber in der Gleichung nicht explizit gemacht, da wir im letzten Schritt den Mittelwertsatz bei festem α anwenden.) Anders geklammert erhalten wir für denselben Ausdruck

$$\begin{aligned} \underbrace{(f(x_1, x_2) - f(0, x_2))}_{=: G_1(x_2)} - \underbrace{(f(x_1, 0) - f(0, 0))}_{=: G_1(0)} &= G'_1(\gamma) x_2 = \underbrace{(\partial_2 f(x_1, \gamma))}_{=: G_2(x_1)} - \underbrace{(\partial_2 f(0, \gamma))}_{=: G_2(0)} x_2 \\ &= G'_2(\delta) x_1 x_2 = \partial_1 \partial_2 f(\delta, \gamma) x_1 x_2 \end{aligned}$$

mit $|\gamma(x)| \leq |x_2|$ und $|\delta(x)| \leq |x_1|$. Für $x_1 x_2 \neq 0$ gilt also $\partial_2 \partial_1 f(\alpha(x), \beta(x)) = \partial_2 \partial_1 f(\delta(x), \gamma(x))$. Da sowohl $\partial_2 \partial_1 f$ als auch $\partial_1 \partial_2 f$ nach Voraussetzung stetige Funktionen sind, erhalten wir schließlich

$$\partial_2 \partial_1 f(0, 0) = \lim_{x \rightarrow 0} \partial_2 \partial_1 f(\alpha(x), \beta(x)) = \lim_{x \rightarrow 0} \partial_1 \partial_2 f(\delta(x), \gamma(x)) = \partial_1 \partial_2 f(0, 0). \quad \square$$

4.16 Definition. Hesse-Matrix

Für $f \in C^2(G)$ und $x \in G$ nennt man die $n \times n$ -Matrix

$$\operatorname{Hess} f(x) = \begin{pmatrix} \partial_1 \partial_1 f(x) & \partial_1 \partial_2 f(x) & \cdots & \partial_1 \partial_n f(x) \\ \partial_2 \partial_1 f(x) & \partial_2 \partial_2 f(x) & \cdots & \partial_2 \partial_n f(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_n \partial_1 f(x) & \partial_n \partial_2 f(x) & \cdots & \partial_n \partial_n f(x) \end{pmatrix}$$

die **Hesse-Matrix von f in x** . Wegen Satz 4.15 ist $\operatorname{Hess} f(x)$ symmetrisch.

4.17 Beispiel. Für $f \in C^2(G)$ gilt

$$\Delta f(x) = \operatorname{div}(\operatorname{grad} f)(x) = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j^2}(x) = \operatorname{Spur}(\operatorname{Hess} f(x)).$$

4.18 Definition. Rotation eines Vektorfeldes

Sei $G \subset \mathbb{R}^3$ offen und $v : G \rightarrow \mathbb{R}^3$ ein partiell differenzierbares Vektorfeld, $v = (v_1, v_2, v_3)$. Man definiert die **Rotation von v** durch $\operatorname{rot} v : G \rightarrow \mathbb{R}^3$,

$$\operatorname{rot} v := \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right).$$

4.19 Notation. Nabla-Operator

Führt man ∇ als “vektorwertigen Operator” ein, den sogenannten **Nabla-Operator**,

$$\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right) = (\partial_1, \dots, \partial_n),$$

so schreiben sich Gradient, Divergenz, Rotation und Laplace folgendermaßen:

$$\begin{aligned} \operatorname{grad} f &= (\partial_1 f, \dots, \partial_n f) = \nabla f \\ \operatorname{div} v &= \sum_{j=1}^n \frac{\partial v_j}{\partial x_j} = \langle \nabla, v \rangle = \nabla \cdot v \\ \Delta f &= \operatorname{div}(\operatorname{grad} f) = \nabla \cdot \nabla f \\ \operatorname{rot} v &= \nabla \times v. \end{aligned}$$

4.20 Korollar. zu Satz 4.15. Sei $G \subset \mathbb{R}^3$ offen.

- (a) Für $f \in C^2(G)$ gilt $\operatorname{rot}(\operatorname{grad} f) = 0$.
- (b) Für $v \in C^2(G, \mathbb{R}^3)$ gilt $\operatorname{div}(\operatorname{rot} v) = 0$.

Beweis. Übungsaufgabe (man rechnet das einfach nach und verwendet den Satz von Schwarz). \square

Das Konzept der partiellen Ableitung für Funktionen auf dem \mathbb{R}^n lässt sich in natürlicher Weise als Spezialfall der **Richtungsableitung** für Funktionen auf beliebigen reellen Vektorräumen verstehen.

4.21 Definition. Richtungsableitung

Seien V ein reeller normierter Raum, W ein normierter Raum, $G \subset V$ offen und $f : G \rightarrow W$ eine Funktion. Unter der **Richtungsableitung** von f im Punkt $x \in G$ in Richtung $v \in V$ versteht man den Differentialquotienten

$$\partial_v f(x) = D_v f(x) := \frac{d}{dh} f(x + hv)|_{h=0} = \lim_{h \rightarrow 0} \frac{f(x + hv) - f(x)}{h}, \tag{4.1}$$

falls dieser Grenzwert existiert.

Für $V = \mathbb{R}^n$ und $v = e_j$ ist $\partial_v f$ also gleich der j -ten partiellen Ableitung $\partial_j f$. \diamond

4.22 Bemerkung. Der Grund dafür, dass wir im Folgenden nur reelle Vektorräume V als Definitionsbereiche betrachten, lässt sich leicht an (4.1) erläutern. Im Falle eines komplexen Vektorraums V würde es einen großen Unterschied machen, ob wir im Differentialquotienten (4.1) die Konvergenz nur für alle reellen Nullfolgen $h_n \in \mathbb{R} \setminus \{0\}$ oder auch für alle komplexen Nullfolgen $h_n \in \mathbb{C} \setminus \{0\}$ fordern. Im ersten Fall spricht man von reeller Differenzierbarkeit. Dann können wir V auch einfach als Vektorraum über \mathbb{R} betrachten und alles weitere gilt entsprechend. Im zweiten Fall ist die Bedingung der Konvergenz in (4.1) deutlich stärker und die sogenannten komplex differenzierbaren Funktionen haben viele zusätzliche Eigenschaften, welche Thema der Funktionentheorie (Englisch “complex analysis”) sind. \diamond

Im Folgenden werden wir zeigen, dass für sogenannte total differenzierbare Funktionen die Abbildung $Df|_x : V \rightarrow W$, $v \mapsto Df|_x v := D_v f(x)$, \mathbb{R} -linear¹ ist, der Wert der Richtungsableitung an einer festen Stelle $x \in G$ also linear von der Richtung v abhängt. Diese lineare Abbildung $Df|_x$ heißt dann die totale Ableitung oder das Differential von f am Punkt x .

Zur Motivation erinnern wir zunächst an den Zusammenhang zwischen Ableitung und **linearer Approximation**. Für $f : I \rightarrow W$, $I \subset \mathbb{R}$ ein offenes Intervall, bedeutet Differenzierbarkeit an der Stelle $x_0 \in I$, dass

$$\lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right) =: \lim_{x \rightarrow x_0} \frac{\varphi(x, x_0)}{x - x_0} = 0,$$

wobei

$$\varphi(x, x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0).$$

Umstellen der Terme liefert, dass sich der Wert von f an einem Punkt x in der Nähe von x_0 schreiben lässt als

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \varphi(x, x_0), \quad (4.2)$$

wobei der Term $\varphi(x, x_0)$ für $x \rightarrow x_0$ gegenüber den anderen vernachlässigt werden kann, da

$$\lim_{x \rightarrow x_0} \varphi(x, x_0) = \lim_{x \rightarrow x_0} \frac{\varphi(x, x_0)}{x - x_0} = 0.$$

Die Abbildung $\mathbb{R} \rightarrow W$, $x \mapsto f'(x_0)x$ ist eine \mathbb{R} -lineare Abbildung und die Abbildung $\mathbb{R} \rightarrow W$, $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ ist affin \mathbb{R} -linear (“affin”, da noch ein konstanter Term addiert wird). Ihr Graph ist die Tangente an den Graphen von f im Punkt x_0 . Daher spricht man in (4.2) von linearer Approximation der Funktion f in einer Umgebung von x_0 .

4.23 Definition. Totale Differenzierbarkeit

Seien V ein endlichdimensionaler reeller Vektorraum, W ein normierter Raum, $G \subset V$ offen, und $f : G \rightarrow W$ eine Abbildung. Es heißt f in einem Punkt $x_0 \in G$ **total differenzierbar** (oder einfach nur **differenzierbar**), wenn es eine \mathbb{R} -lineare Abbildung $A : V \rightarrow W$ gibt, sodass

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - A(x - x_0)}{\|x - x_0\|_V} = 0,$$

also für jede Folge (x_n) in $G \setminus \{x_0\}$ mit $\lim_{n \rightarrow \infty} x_n = x_0$

$$\lim_{n \rightarrow \infty} \frac{\|f(x_n) - f(x_0) - A(x_n - x_0)\|_W}{\|x_n - x_0\|_V} = 0 \quad (4.3)$$

gilt. Hierbei ist $\|\cdot\|_V$ eine beliebige Norm auf V . Da gemäß Satz 2.13 auf einem endlichdimensionalen Vektorraum alle Normen äquivalent sind, hängt die Gültigkeit von (4.3) nicht von der Wahl der Norm auf V ab.

¹Eine Abbildung $A : V \rightarrow W$ von einem \mathbb{R} -Vektorraum V in einen \mathbb{K} -Vektorraum W ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) heißt \mathbb{R} -linear, wenn für alle $v, u \in V$ und $\alpha \in \mathbb{R}$ gilt, dass $A(v + \alpha u) = Av + \alpha Au$.

4 Differenzierbarkeit

Falls f bei x_0 total differenzierbar ist, so ist die \mathbb{R} -lineare Abbildung $A : V \rightarrow W$ eindeutig bestimmt (Übungsaufgabe), wird mit $Df|_{x_0}$ bezeichnet und das **Differential** oder die **Ableitung** von f an der Stelle x_0 genannt.

Es heißt $f : G \rightarrow W$ **total differenzierbar**, wenn f in allen Punkten $x \in G$ total differenzierbar ist. Dann ist $Df : G \rightarrow \mathcal{L}(V, W)$, $x \mapsto Df(x) = Df|_x$ eine Funktion mit Werten im Raum der linearen Abbildungen von V nach W . \diamond

4.24 Bemerkung. Der schon in Bemerkung 4.22 angesprochene Unterschied zwischen reeller Differenzierbarkeit (dem Thema dieser Vorlesung) und komplexer Differenzierbarkeit (dem Thema der Funktionentheorie) macht sich daran fest, dass wir in Definition 4.23 durch \mathbb{R} -lineare Abbildungen approximieren und nicht durch \mathbb{C} -lineare.

Im Folgenden wird der Definitionsbereich G immer als Teilmenge eines reellen endlichdimensionalen Vektorraums V angenommen. Endlichdimensional, weil dann die Wahl einer Norm auf V für die Frage der Differenzierbarkeit keine Rolle spielt, und reell, weil wir wie gesagt Differenzierbarkeit im Sinne von \mathbb{R} -linearer Approximierbarkeit untersuchen wollen. Die Ergebnisse übertragen sich aber sofort auch auf n -dimensionale \mathbb{C} -Vektorräume V , sofern wir diese einfach als $2n$ -dimensionale \mathbb{R} -Vektorräume auffassen. Denn jeder n -dimensionale \mathbb{C} -Vektorraum ist isomorph zum \mathbb{C}^n und \mathbb{C}^n ist mittels $\mathbb{C}^n \rightarrow \mathbb{R}^{2n}$, $(z_1, \dots, z_n) \mapsto (x_1, y_1, \dots, x_n, y_n)$ \mathbb{R} -isomorph zum \mathbb{R}^{2n} . Hier ist $x_j := \operatorname{Re} z_j$ und $y_j := \operatorname{Im} z_j$. \diamond

4.25 Bemerkung. Ist $g : V \supset B_\delta(0) \rightarrow W$ eine Funktion und $k \in \mathbb{N}_0$, so sagt man, dass g von höherer als k -ter Ordnung in 0 verschwindet und schreibt $g = o(\|h\|_V^k)$, falls

$$\lim_{h \rightarrow 0} \frac{g(h)}{\|h\|_V^k} = 0.$$

Mit dieser Schreibweise ist also $f : G \rightarrow W$ differenzierbar in x_0 wenn sich f in x_0 bis auf einen Fehler der Ordnung $o(\|x - x_0\|_V)$ linear approximieren läßt, also wenn

$$f(x) = f(x_0) + Df|_{x_0}(x - x_0) + o(\|x - x_0\|_V).$$

4.26 Beispiele. Lineare und bilineare Abbildungen, kartesische Produkte

- (a) Sei $L : V \rightarrow W$ eine \mathbb{R} -lineare Abbildung, wobei wir im Folgenden die übliche Schreibweise $L(x) =: Lx$ ohne Klammern ums Argument für lineare Abbildungen verwenden. Dann ist

$$Lx = L(x_0 + x - x_0) = Lx_0 + L(x - x_0) =: Lx_0 + DL|_{x_0}(x - x_0),$$

also $DL|_x \equiv L$. Die lineare Approximation einer linearen Abbildung ist also, wenig überraschend, die Abbildung selbst.

- (b) Sei $f : V \supset G \rightarrow W_1 \times W_2$, $x \mapsto f(x) = (f_1(x), f_2(x))$, differenzierbar. Dann ist für $x \in G$

$$Df|_x v = (Df_1|_x v, Df_2|_x v).$$

Insbesondere ist für $g : V \supset G \rightarrow \mathbb{R}^m$

$$Dg|_x v = (Dg_1|_x v, \dots, Dg_m|_x v), \quad \text{also } (Dg|_x v)_j = Dg_j|_x v \text{ für } j = 1, \dots, m.$$

- (c) Sei $f : V_1 \times V_2 \supset G \rightarrow W$, $x = (x_1, x_2) \mapsto f(x) = f(x_1, x_2)$, differenzierbar. Dann ist für $x \in G$

$$Df|_x v = Df|_x(v_1, v_2) =: D_1 f|_x v_1 + D_2 f|_x v_2.$$

Denn für jede lineare Abbildung $L : V_1 \times V_2 \rightarrow W$ gilt

$$Lv = L(v_1, v_2) = L(v_1, 0) + L(0, v_2) =: L_1 v_1 + L_2 v_2.$$

(d) Sei $B : V = V_1 \times V_2 \rightarrow W$, $z := (x, y) \mapsto B(x, y)$ eine bilineare Abbildung. Dann ist

$$\begin{aligned} B(z) &= B(x, y) = B(x_0 + x - x_0, y_0 + y - y_0) \\ &= B(x_0, y_0) + B(x - x_0, y_0) + B(x_0, y - y_0) + B(x - x_0, y - y_0) \\ &= B(x_0, y_0) + D_1 B|_{y_0}(x - x_0) + D_2 B|_{x_0}(y - y_0) + B(z - z_0) \\ &= B(z_0) + DB|_{z_0}(z - z_0) + \mathcal{O}(\|z - z_0\|_V^2), \end{aligned}$$

wobei $D_1 B|_{y_0} : V_1 \rightarrow W$, $D_1 B|_{y_0} x := B(x, y_0)$ und entsprechend $D_2 B$ jeweils lineare Abbildungen auf V_j sind. Um zu sehen, dass tatsächlich

$$\|B(z - z_0)\|_W \leq C \|z - z_0\|_V^2 \quad (4.4)$$

gilt, wählen wir Basen $(e_{1,1}, \dots, e_{1,n})$ und $(e_{2,1}, \dots, e_{2,m})$ von V_1 und V_2 und stellen fest, dass

$$\begin{aligned} \|B(z)\|_W &= \|B(x, y)\|_W =: \left\| B\left(\sum_{j=1}^n v_{1,j} e_{1,j}, \sum_{i=1}^m v_{2,i} e_{2,i}\right) \right\|_W \\ &\leq \sum_{j=1}^n \sum_{i=1}^m |v_{1,j}| |v_{2,i}| \|B(e_{1,j}, e_{2,i})\|_W \\ &\leq \max_{i,j} \{ \|B(e_{1,j}, e_{2,i})\|_W \} \left(\sum_{j=1}^n |v_{1,j}| \right) \left(\sum_{i=1}^m |v_{2,i}| \right) \\ &=: c \|x\|_1 \|y\|_1 \leq c \|z\|_1^2, \end{aligned}$$

wobei $\|z\|_1 = \|(x, y)\|_1 := \|x\|_1 + \|y\|_1$. Da auf einem endlichdimensionalen Vektorraum alle Normen äquivalent sind, folgt (4.4). \diamond

Es mag zwar zunächst offensichtlich erscheinen, dass eine lineare Abbildung auch stetig ist. Tatsächlich ist das im Allgemeinen aber nur für lineare Abbildungen auf endlichdimensionalen Vektorräumen richtig. Da die Stetigkeit des Differentials im Folgenden immer wieder wichtig sein wird, diskutieren wir an dieser Stelle kurz die Stetigkeit bzw. äquivalent die Beschränktheit linearer Abbildungen.

4.27 Definition. Beschränkte lineare Abbildungen und ihre Norm

Sei $A : V \rightarrow W$ eine lineare Abbildung zwischen normierten Räumen. Es heißt A **beschränkt**, falls eine Konstante $C \in \mathbb{R}$ existiert, sodass für alle $v \in V$ gilt

$$\|Av\|_W \leq C \|v\|_V.$$

Die kleinste solche Konstante

$$\|A\| := \sup\{\|Av\|_W \mid v \in V \text{ mit } \|v\|_V \leq 1\}$$

heißt die **Operatornorm** von A . Es gilt also für alle $v \in V$, dass

$$\|Av\|_W = \left\| A \frac{v}{\|v\|_V} \right\|_W \|v\|_V \leq \|A\| \|v\|_V.$$

4.28 Satz. Für eine lineare Abbildung $A : V \rightarrow W$ zwischen normierten Räumen sind äquivalent:

- (i) A ist stetig bei $v = 0$.
- (ii) A ist stetig.
- (iii) A ist beschränkt.

Beweis. Übungsaufgabe. \square

4.29 Satz. Lineare Abbildungen auf endlichdimensionalen Räumen sind stetig

Seien V und W normierte Räume und V endlichdimensional. Dann ist jede lineare Abbildung $A : V \rightarrow W$ beschränkt und somit stetig.

Beweis. Sei (v_1, \dots, v_n) eine Basis von V und $\Phi : V \rightarrow \mathbb{K}^n$ der zugehörige Basisisomorphismus. Dann ist

$$\|\cdot\|_{V,\infty} : V \rightarrow [0, \infty), \quad v \mapsto \|v\|_{V,\infty} := \|\Phi(v)\|_\infty = \max\{|\Phi(v)_1|, \dots, |\Phi(v)_n|\}$$

eine Norm auf V und

$$\begin{aligned} \|Av\|_W &= \left\| A \sum_{j=1}^n c_j v_j \right\|_W = \left\| \sum_{j=1}^n c_j A(v_j) \right\|_W \leq \sum_{j=1}^n |c_j| \|A(v_j)\|_W \\ &\leq \underbrace{\sum_{j=1}^n \|A(v_j)\|_W}_{=:C} \cdot \max\{|c_1|, \dots, |c_n|\} = C \|v\|_{V,\infty}. \end{aligned}$$

Also ist A als Abbildung von $(V, \|\cdot\|_{V,\infty})$ nach $(W, \|\cdot\|_W)$ beschränkt. Da alle Normen auf V äquivalent sind, gibt es zu jeder anderen Norm $\|\cdot\|_V$ auf V ein $c \in \mathbb{R}$ mit $\|v\|_{V,\infty} \leq c\|v\|_V$ und somit gilt auch $\|Av\|_W \leq cC\|v\|_V$. \square

4.30 Korollar. Stetigkeit differenzierbarer Funktionen

Sei $f : V \supset G \rightarrow W$ an der Stelle $x_0 \in G$ total differenzierbar. Dann ist f bei x_0 auch stetig.

Beweis. Es ist

$$\lim_{x \rightarrow x_0} f(x) = f(x_0) + \lim_{x \rightarrow x_0} (Df|_{x_0}(x - x_0) + o(\|x - x_0\|)) = f(x_0),$$

wobei wir verwendet haben, dass $Df|_{x_0}$ gemäß Satz 4.29 stetig ist. \square

4.31 Bemerkung. Wir werden uns im Folgenden meist auf den Fall beschränken, dass auch W ein endlichdimensionaler Vektorraum ist. Aus der linearen Algebra wissen wir, dass V und W nach Wahl jeweils einer Basis mit dem \mathbb{R}^n bzw. \mathbb{K}^m identifiziert werden können und eine \mathbb{R} -lineare Abbildung $A : V \rightarrow W$ dann bezüglich dieser Basen durch eine $(m \times n)$ -Matrix (a_{ij}) mit Einträgen aus \mathbb{K} beschrieben wird.

Im Folgenden sind alle Aussagen die eine Basisdarstellung involvieren gleich für den \mathbb{R}^n bzw. \mathbb{K}^m mit der kanonischen Basis formuliert. D.h. wir machen die Wahl einer Basis in V bzw. W nicht explizit um die Notation nicht zu überladen. \diamond

4.32 Satz. und Definition. Die Jacobi-Matrix

Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{K}^m$ eine Abbildung, die im Punkt $x_0 \in G$ total differenzierbar ist, also

$$f(x) = f(x_0) + Df|_{x_0}(x - x_0) + o(\|x - x_0\|)$$

mit der Matrix $Df|_{x_0} = (a_{ij}) \in M(m \times n, \mathbb{K})$ erfüllt.

Dann sind alle Komponenten $f_i : G \rightarrow \mathbb{K}$, $1 \leq i \leq m$, von f in x_0 partiell differenzierbar mit

$$\frac{\partial f_i}{\partial x_j}(x_0) = a_{ij}.$$

Man nennt $Df|_{x_0}$ in diesem Fall auch die **Jacobi-Matrix** oder die **Funktionalmatrix** von f im Punkte x_0 und schreibt

$$Df|_{x_0} =: J_f(x_0) = \left(\frac{\partial f_i}{\partial x_j}(x_0) \right)_{ij}.$$

Beweis. Für $i = 1, \dots, m$ ist

$$f_i(x) = f_i(x_0) + \sum_{\ell=1}^n a_{i\ell}(x - x_0)_\ell + o(\|x - x_0\|),$$

also für $h \in \mathbb{R}$ hinreichend klein und $x = x_0 + he_j$

$$f_i(x_0 + he_j) = f_i(x_0) + ha_{ij} + o(\|he_j\|).$$

Damit folgt für die partielle Ableitung von f_i in Richtung e_j

$$\frac{\partial f_i}{\partial x_j}(x_0) := \lim_{h \rightarrow 0} \frac{f_i(x_0 + he_j) - f_i(x_0)}{h} = a_{ij} + \lim_{h \rightarrow 0} \frac{o(|h|)}{h} = a_{ij}. \quad \square$$

Wenn wir bereits wissen, dass eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{K}^m$ total differenzierbar ist, so können wir also durch Berechnung der partiellen Ableitungen auch das Differential Df in Form der Jacobi-Matrix bestimmen. Der folgende Satz gibt uns eine hinreichende Bedingung dafür, wie wir aus Kenntnis der partiellen Ableitungen auf die totale Differenzierbarkeit schließen können: Sind die partiellen Ableitungen stetig an einem Punkt x , so ist die Funktion dort total differenzierbar.

4.33 Satz. Stetig partiell differenzierbar \Rightarrow total differenzierbar

Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion. Falls alle partiellen Ableitungen $\partial_j f$ im Punkt $x \in G$ stetig sind, so ist f in x total differenzierbar.

Beweis. Für $h \in \mathbb{R}^n$ hinreichend klein sei

$$z^{(i)} := x + \sum_{j=1}^i h_j e_j \quad , \quad i = 0, \dots, n.$$

Es gilt $z^{(0)} = x$ und $z^{(n)} = x + h$. Die Punkte $z^{(i-1)}$ und $z^{(i)}$ unterscheiden sich nur in der i -ten Koordinate. Nach dem Mittelwertsatz für differenzierbare Funktionen einer Veränderlichen gibt es deshalb jeweils ein $\theta_i(h) \in [0, 1]$, so dass

$$f(z^{(i)}) - f(z^{(i-1)}) = \partial_i f(y^{(i)}) h_i$$

wobei

$$y^{(i)} = z^{(i-1)} + \theta_i(h) h_i e_i.$$

Daraus folgt

$$f(x + h) - f(x) = \sum_{i=1}^n \partial_i f(y^{(i)}) h_i$$

bzw.

$$f(x + h) = f(x) + \sum_{i=1}^n \partial_i f(x) h_i + \underbrace{\sum_{i=1}^n \left(\partial_i f(y^{(i)}) - \partial_i f(x) \right) h_i}_{=: \varphi(h)}.$$

Wegen der Beschränktheit von $\theta_i(h)$ gilt jeweils $\lim_{h \rightarrow 0} y^{(i)} = x$ und wegen der Stetigkeit von $\partial_i f$ in x gilt

$$\lim_{h \rightarrow 0} \left(\partial_i f(y^{(i)}) - \partial_i f(x) \right) = \partial_i f(\lim_{h \rightarrow 0} y^{(i)}) - \partial_i f(x) = \partial_i f(x) - \partial_i f(x) = 0.$$

Da $\frac{|h_i|}{\|h\|} \leq 1$ für alle $i = 1, \dots, n$, folgt

$$\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|} = 0. \quad \square$$

4 Differenzierbarkeit

4.34 Korollar. Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{K}^m$ stetig partiell differenzierbar. Dann ist f total differenzierbar und somit stetig.

Beweis. Sei $\mathbb{K}^m = \mathbb{R}^m$. Mit Satz 4.33 gilt für $j = 1, \dots, m$

$$f_j(x+h) = f_j(x) + \sum_{i=1}^n \partial_i f_j(x) h_i + \varphi_j(h),$$

wobei $\varphi_j(h) = o(\|h\|)$. Multiplikation mit $e_j \in \mathbb{R}^m$ und Summation über j liefert

$$f(x+h) = f(x) + \sum_{j=1}^m \sum_{i=1}^n \partial_i f_j(x) h_i e_j + \sum_{j=1}^m \varphi_j(h) = f(x) + Df|_x h + \varphi(h)$$

mit $\varphi(h) := \sum_{j=1}^m \varphi_j(h) = o(\|h\|)$. Im Fall $\mathbb{K}^m = \mathbb{C}^m$ fassen wir den \mathbb{C}^m wieder als reellen Vektorraum \mathbb{R}^{2m} auf und argumentieren analog. \square

4.35 Merke. Es gelten also die Implikationen:

$$\begin{array}{ccccc} \text{stetig partiell differenzierbar} & \Rightarrow & \text{total differenzierbar} & \Rightarrow & \text{partiell differenzierbar} \\ & & \Downarrow & & \\ & & \text{stetig} & & \end{array}$$

Die Umkehrungen gelten im Allgemeinen nicht! Dafür gilt aber

$$\text{stetig partiell differenzierbar} \Leftrightarrow \text{total differenzierbar mit stetiger Ableitung}$$

weshalb man "stetig partiell differenzierbar" und "stetig differenzierbar" synonym verwendet.

4.36 Satz. Kettenregel

Seien U, V endlichdimensionale reelle Vektorräume, W ein normierter Raum, $G \subset U$ und $H \subset V$ offen und $g : G \rightarrow V$ und $f : H \rightarrow W$ Abbildungen mit $g(G) \subset H$, also

$$U \supset G \xrightarrow{g} H \subset V \xrightarrow{f} W.$$

Die Abbildung g sei im Punkt $x \in G$ differenzierbar und die Abbildung f sei im Punkt $g(x) \in H$ differenzierbar. Dann ist die Komposition

$$f \circ g : G \rightarrow W$$

im Punkt x differenzierbar und für ihr Differential gilt:

$$D(f \circ g)|_x = \underbrace{Df|_{g(x)}}_{W \leftarrow V} \cdot \underbrace{Dg|_x}_{V \leftarrow U}.$$

4.37 Bemerkung. Man beachte die Natürlichkeit der Kettenregel: Die lineare Approximation $D(f \circ g)|_x$ der Verkettung $f \circ g$ ist also einfach die Verkettung der linearen Approximationen $Dg|_x$ und $Df|_{g(x)}$. \diamond

Beweis. (der Kettenregel). Sei $A := Dg|_x$ und $B = Df|_{g(x)}$. Es ist zu zeigen, dass $D(f \circ g)|_x = BA$. Nach Voraussetzung gelten

$$\begin{aligned} g(x+h) &= g(x) + Ah + \varphi(h) \\ f(g(x) + \eta) &= f(g(x)) + B\eta + \psi(\eta) \end{aligned}$$

mit $\varphi(h) = o(\|h\|)$ und $\psi(\eta) = o(\|\eta\|)$. Dann ist

$$\begin{aligned}(f \circ g)(x+h) &= f(g(x+h)) = f(g(x) + Ah + \varphi(h)) \\ &= f(g(x)) + BAh + B\varphi(h) + \psi(Ah + \varphi(h)) \\ &= (f \circ g)(x) + BAh + \chi(h)\end{aligned}$$

mit $\chi(h) = B\varphi(h) + \psi(Ah + \varphi(h))$. Es bleibt also zu zeigen, dass $\chi(h) = o(\|h\|)$. Mit $\varphi(h) = o(\|h\|)$ ist auch $B\varphi(h) = o(\|h\|)$, da B ja beschränkt ist. Außerdem gibt es eine Konstante $K > 0$, so dass $\|\varphi(h)\| \leq K\|h\|$ für alle hinreichend kleinen h . Wegen $\psi(\eta) = o(\|\eta\|)$ gilt sogar $\psi(\eta) =: \|\eta\|\psi_1(\eta)$ mit $\lim_{\eta \rightarrow 0} \psi_1(\eta) = 0$. Damit ergibt sich

$$\|\psi(Ah + \varphi(h))\| \leq (\|A\| + K)\|h\| \cdot \|\psi_1(Ah + \varphi(h))\|,$$

also

$$\lim_{h \rightarrow 0} \frac{\psi(Ah + \varphi(h))}{\|h\|} = 0. \quad \square$$

4.38 Satz. Richtungsableitung und Differential

Seien V ein endlichdimensionaler reeller Vektorraum, W ein normierter Raum, $G \subset V$ offen und $f : G \rightarrow W$ total differenzierbar. Dann gilt für jedes $x \in G$ und jeden Vektor $v \in V$

$$\partial_v f(x) = Df|_x v.$$

Beweis. Sei $\alpha : (-\delta, \delta) \rightarrow G$ definiert durch $\alpha(t) := x + tv$ und $h := f \circ \alpha : (-\delta, \delta) \rightarrow W$. Nach Definition der Richtungsableitung ist

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

Aus der Kettenregel folgt

$$\frac{d}{dt} h(t) = Df|_{\alpha(t)} \cdot D\alpha|_t = Df|_{\alpha(t)} v$$

also $\frac{d}{dt} h(0) = Df|_x v$. □

4.39 Beispiel. Für eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ist die Richtungsableitung in Richtung $v \in \mathbb{R}^n$ also durch das Skalarprodukt von v mit dem Gradienten ∇f gegeben,

$$\partial_v f(x) = Df|_x v = (\partial_1 f(x), \dots, \partial_n f(x)) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \langle \nabla f(x), v \rangle.$$

Ist $\nabla f(x) \neq 0$, so ist der Winkel θ zwischen den Vektoren v und $\nabla f(x)$ definiert und es gilt

$$\partial_v f(x) = \langle \nabla f(x), v \rangle = \|\nabla f(x)\| \|v\| \cos \theta.$$

Die Richtungsableitung ist also maximal, falls $\theta = 0$, d.h. falls v und $\nabla f(x)$ die gleiche Richtung haben. Der Vektor $\nabla f(x)$ gibt also die Richtung des stärksten Anstiegs von f an. ◇

4.40 Beispiel. Der Laplaceoperator in Polarkoordinaten

Die Abbildung

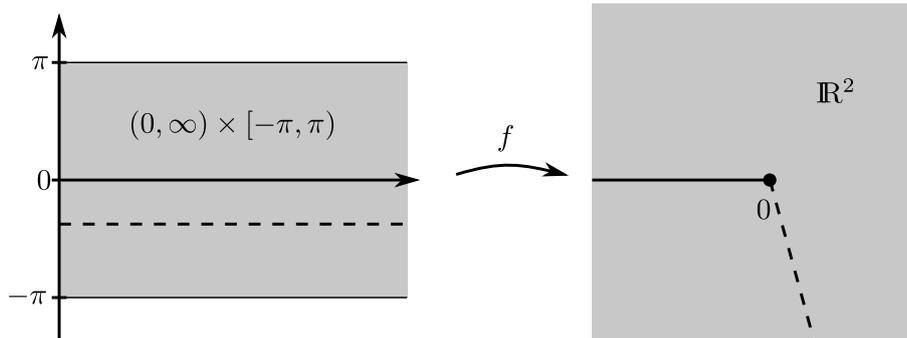
$$f : (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}^2 \setminus \{0\}, \quad (r, \varphi) \mapsto f(r, \varphi) = (r \cos \varphi, r \sin \varphi) =: (x(r, \varphi), y(r, \varphi))$$

4 Differenzierbarkeit

versieht den $\mathbb{R}^2 \setminus \{0\}$ mit Polarkoordinaten. Eingeschränkt z.B. auf $(0, \infty) \times [-\pi, \pi)$ ist f sogar bijektiv. Sei $G \subset \mathbb{R}^2$ offen und $u \in C^2(G, \mathbb{R})$, dann drückt man u in Polarkoordinaten aus, indem man $u \circ f$ betrachtet, also u als Funktion von (r, φ) schreibt. Genauso ist Δu in Polarkoordinaten durch $(\Delta u) \circ f$ gegeben. Ziel ist es nun, $(\Delta u) \circ f$ durch Differentiation an $u \circ f$ ausdrücken. Und tatsächlich gilt

$$(\Delta u) \circ f = \frac{\partial^2(u \circ f)}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2(u \circ f)}{\partial \varphi^2} + \frac{1}{r} \frac{\partial(u \circ f)}{\partial r}. \quad (4.5)$$

Beweis. Nachrechnen mit Kettenregel (Übungsaufgabe). □



EINSCHUB FÜR INTERESSIERTE: Wie kommt man auf den Ausdruck (4.5)? Dazu betrachten wir zunächst das Differential von u in Polarkoordinaten. Mit der Kettenregel gilt

$$D(u \circ f) = Du \circ f \cdot Df, \quad \text{also} \quad Du \circ f = D(u \circ f) \cdot (Df)^{-1}, \quad (4.6)$$

wobei $(Df)^{-1}$ in unserem Beispiel leicht berechnet werden kann:

$$\begin{aligned} (Df)^{-1} &= \begin{pmatrix} \frac{\partial x}{\partial r}(r, \varphi) & \frac{\partial x}{\partial \varphi}(r, \varphi) \\ \frac{\partial y}{\partial r}(r, \varphi) & \frac{\partial y}{\partial \varphi}(r, \varphi) \end{pmatrix}^{-1} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}^{-1} \\ &= \frac{1}{r} \begin{pmatrix} r \cos \varphi & r \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix}. \end{aligned}$$

Wir haben also in (4.6) eine allgemeine Formel für das Differential Du von u ausgedrückt in Koordinaten geben durch f . Dabei benötigen wir nur die Invertierbarkeit von Df . In unserem Beispiel ergibt sich

$$\begin{aligned} Du \circ f &= \left(\frac{\partial(u \circ f)}{\partial r}, \frac{\partial(u \circ f)}{\partial \varphi} \right) \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix} \\ &= \left(\cos \varphi \frac{\partial(u \circ f)}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial(u \circ f)}{\partial \varphi}, \sin \varphi \frac{\partial(u \circ f)}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial(u \circ f)}{\partial \varphi} \right) \\ &=: \underbrace{\left(\cos \varphi \frac{\partial u}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial u}{\partial \varphi} \right)}_{\frac{\partial u}{\partial x} \circ f} \underbrace{\left(\sin \varphi \frac{\partial u}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial u}{\partial \varphi} \right)}_{\frac{\partial u}{\partial y} \circ f} = \text{grad } u \circ f, \quad (4.7) \end{aligned}$$

wobei hier die Komponenten bezüglich der kanonischen Basis des \mathbb{R}^2 stehen und $\text{grad } u$ als Zeilenvektor aufgefasst wird. Im Folgenden unterdrücken wir manchmal f , d.h. an Stelle von $(u \circ f)(r, \varphi)$ schreiben wir einfach $u(r, \varphi)$ und entsprechend $\partial_r u(r, \varphi)$ statt $\partial_r(u \circ f)(r, \varphi)$. In dieser verkürzten

Notation können wir nun auch $\Delta u \circ f$ berechnen:

$$\begin{aligned} \frac{\partial^2}{\partial x^2} u &= \left(\cos \varphi \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi} \right)^2 u = \left(\cos \varphi \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi} \right) \left(\cos \varphi \frac{\partial u}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial u}{\partial \varphi} \right) \\ &= \cos^2 \varphi \frac{\partial^2 u}{\partial r^2} + \frac{\cos \varphi \sin \varphi}{r^2} \frac{\partial u}{\partial \varphi} - \frac{\cos \varphi \sin \varphi}{r} \frac{\partial^2 u}{\partial r \partial \varphi} \\ &\quad - \frac{\cos \varphi \sin \varphi}{r} \frac{\partial^2 u}{\partial r \partial \varphi} + \frac{\sin^2 \varphi}{r} \frac{\partial u}{\partial r} + \frac{\cos \varphi \sin \varphi}{r^2} \frac{\partial u}{\partial \varphi} + \frac{\sin^2 \varphi}{r^2} \frac{\partial^2 u}{\partial \varphi^2} \end{aligned}$$

und analog

$$\begin{aligned} \frac{\partial^2}{\partial y^2} u &= \left(\sin \varphi \frac{\partial}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial}{\partial \varphi} \right)^2 u = \left(\sin \varphi \frac{\partial}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial}{\partial \varphi} \right) \left(\sin \varphi \frac{\partial u}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial u}{\partial \varphi} \right) \\ &= \sin^2 \varphi \frac{\partial^2 u}{\partial r^2} - \frac{\cos \varphi \sin \varphi}{r^2} \frac{\partial u}{\partial \varphi} + \frac{\cos \varphi \sin \varphi}{r} \frac{\partial^2 u}{\partial r \partial \varphi} \\ &\quad + \frac{\cos \varphi \sin \varphi}{r} \frac{\partial^2 u}{\partial r \partial \varphi} + \frac{\cos^2 \varphi}{r} \frac{\partial u}{\partial r} - \frac{\cos \varphi \sin \varphi}{r^2} \frac{\partial u}{\partial \varphi} + \frac{\cos^2 \varphi}{r^2} \frac{\partial^2 u}{\partial \varphi^2}. \end{aligned}$$

Addiert man die beiden Ausdrücke, so ergibt sich (4.5).

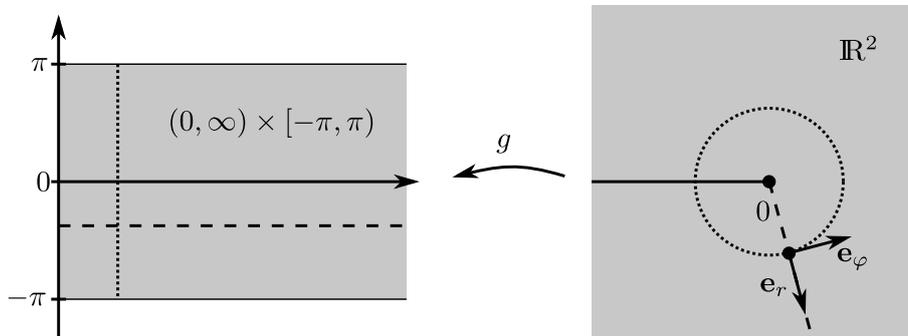
Nochmals zurück zu (4.7): Die Matrix $Du \circ f$ ist die Linearisierung der Abbildung $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ im Punkte $f(r, \varphi)$ bezüglich der kanonischen Basis des \mathbb{R}^2 . Man kann nun aber auch noch versuchen, die lineare Abbildung $Du \circ f$ bezüglich einer den Koordinaten angepassten Basis darzustellen, nämlich bezüglich der Basisvektoren \mathbf{e}_r und \mathbf{e}_φ . Um diese vernünftig zu definieren, schränken wir zunächst f auf $(0, \infty) \times (-\pi, \pi)$ ein und erhalten so einen **Diffeomorphismus**

$$f : G := (0, \infty) \times (-\pi, \pi) \rightarrow \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\} =: D$$

Seine Umkehrung ist gegeben durch $g : D \rightarrow G$, $g(x, y) = (r(x, y), \varphi(x, y))$ mit $r(x, y) = \sqrt{x^2 + y^2}$ und

$$\varphi(x, y) = \begin{cases} \arctan(x/y) & \text{für } x > 0, y > 0 \\ \pi/2 & \text{für } x = 0, y > 0 \\ \vdots & \end{cases}$$

Wir werden später zeigen, dass die Invertierbarkeit von Df ganz allgemein die lokale Invertierbarkeit von f impliziert.



Die Zeilenvektoren von Dg bilden nun eine Basis des \mathbb{R}^2 und stehen senkrecht auf den jeweiligen Koordinatenlinien, da sie ja gerade durch den Gradienten der Koordinatenfunktionen gegeben sind,

$$Dg = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} \end{pmatrix} = \begin{pmatrix} \text{grad } r \\ \text{grad } \varphi \end{pmatrix}.$$

4 Differenzierbarkeit

Um nun Dg auszurechnen, differenziert man nicht etwa g (was bei komplizierten f 's oft gar nicht explizit geht), sondern verwendet nochmals die Kettenregel. Weil $g \circ f = \text{id}$ ist, ist nach der Kettenregel

$$Dg|_{f(r,\varphi)} \cdot Df|_{(r,\varphi)} = D(\text{id})|_{(r,\varphi)} = E,$$

wobei E wie immer die Einheitsmatrix bezeichnet. Also ist

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

was wir für unser Beispiel oben schon berechnet haben. Der Basiswechsel von der neuen Basis in die kanonische ist durch $(Dg)^T = ((Df)^{-1})^T$ gegeben. Multiplikation von (4.7) mit $((Df)^{-1})^T$ von rechts (nur im Urbildraum wird die Basis gewechselt) liefert das Differential Du als Matrix bezüglich der durch f beschriebenen Basisvektoren

$$(Du \circ f)(Dg \circ f)^T = D(u \circ f) \cdot (Df)^{-1} \cdot ((Df)^{-1})^T.$$

Im Beispiel der Polarkoordinaten ergibt sich bzgl. der Basis $(\text{grad}r, \text{grad}\varphi)$

$$Du = \left(\cos \varphi \frac{\partial u}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial u}{\partial \varphi}, \sin \varphi \frac{\partial u}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial u}{\partial \varphi} \right) \begin{pmatrix} \cos \varphi & -\frac{\sin \varphi}{r} \\ \sin \varphi & \frac{\cos \varphi}{r} \end{pmatrix} = \left(\frac{\partial u}{\partial r}, \frac{1}{r^2} \frac{\partial u}{\partial \varphi} \right),$$

bzw. nach Normierung mit $\mathbf{e}_r = \text{grad}r$ und $\mathbf{e}_\varphi = r^{-1} \text{grad}\varphi$ und einem weiteren Basiswechsel in die Basis $(\mathbf{e}_r, \mathbf{e}_\varphi)$

$$Du = \left(\frac{\partial u}{\partial r}, \frac{1}{r^2} \frac{\partial u}{\partial \varphi} \right) \begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix} = \left(\frac{\partial u}{\partial r}, \frac{1}{r} \frac{\partial u}{\partial \varphi} \right).$$

Wir hätten natürlich auch gleich in die Basis $(\mathbf{e}_r, \mathbf{e}_\varphi)$ transformieren können,

$$Du = \left(\cos \varphi \frac{\partial u}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial u}{\partial \varphi}, \sin \varphi \frac{\partial u}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial u}{\partial \varphi} \right) \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} = \left(\frac{\partial u}{\partial r}, \frac{1}{r} \frac{\partial u}{\partial \varphi} \right),$$

Da $(\mathbf{e}_r, \mathbf{e}_\varphi)$ eine Orthonormalbasis ist, gilt schließlich

$$Du = \frac{\partial u}{\partial r} \mathbf{e}_r^T + \frac{1}{r} \frac{\partial u}{\partial \varphi} \mathbf{e}_\varphi^T.$$

ENDE DES EINSCHUBS

Wir kommen nun zur Verallgemeinerung der Produktregel. Während man \mathbb{K} -wertige Funktionen f, g einfach multiplizieren kann, indem man punktweise die Werte multipliziert, $(fg)(x) := f(x)g(x)$, macht das für vektorwertige Funktionen zunächst keinen Sinn mehr. Allerdings ist für die Gültigkeit der Produktregel lediglich notwendig, dass das "Produkt" linear in beiden Faktoren ist. Daher zeigen wir nun eine Produktregel für Abbildungen der Form

$$F(x) = B(f(x), g(x))$$

wobei f und g nun vektorwertig sind und B linear in beiden Argumenten ist.

4.41 Satz. Die verallgemeinerte Produktregel

Seien U, V_1, V_2 reelle endlichdimensionale Vektorräume und W ein normierter Raum. Seien weiterhin $B : V = V_1 \times V_2 \rightarrow W$, $z = (x, y) \mapsto B(x, y)$, eine bilineare Abbildung, $f_j : G \rightarrow V_j$, $j = 1, 2$, differenzierbare Abbildungen auf einem Gebiet $G \subset U$ und $f : G \rightarrow V_1 \times V_2$, $f(u) := (f_1(u), f_2(u))$. Dann ist $F : G \rightarrow W$, $u \mapsto F(u) = (B \circ f)(u) = B(f_1(u), f_2(u))$ differenzierbar und

$$DF|_u = \underbrace{D_1 B|_{f_2(u)}}_{W \leftarrow V_1} \underbrace{Df_1|_u}_{V_1 \leftarrow U} + \underbrace{D_2 B|_{f_1(u)}}_{W \leftarrow V_2} \underbrace{Df_2|_u}_{V_2 \leftarrow U}.$$

Hier ist $D_1B|_y : V_1 \rightarrow W$ die lineare Abbildung $x \mapsto D_1B|_y x := B(x, y)$ und $D_2B|_x : V_2 \rightarrow W$ entsprechend $y \mapsto D_2B|_x y := B(x, y)$.

Die Wirkung der linearen Abbildung $DF|_u$ auf einen Vektor $w \in U$ ist somit

$$DF|_u w = DB(f_1(\cdot), f_2(\cdot))|_u w = B(Df_1|_u w, f_2(u)) + B(f_1(u), Df_2|_u w).$$

Man kann also in einer bilinearen Abbildung “zuerst nur das eine und dann nur das andere Argument” ableiten und die beiden Terme addieren.

Die entsprechende Aussage gilt auch für multilineare Abbildungen $B : V_1 \times \cdots \times V_n \rightarrow W$ und differenzierbare $f_j : G \rightarrow V_j$,

$$DF|_u w = \sum_{j=1}^n D_j B|_{f(u)} Df_j|_u w = \sum_{j=1}^n B(f_1(u), \dots, Df_j|_u w, \dots, f_n(u)).$$

Beweis. Die Kettenregel zusammen mit Beispiel (4.26) (d) liefert sofort die Aussage für bilineare Abbildungen,

$$DF|_u = DB|_{f(u)} Df|_u = D_1 B|_{f_2(u)} Df_1|_u + D_2 B|_{f_1(u)} Df_2|_u.$$

Der allgemeine Fall geht ganz analog. □

4.42 Beispiele. (a) Die Multiplikation $\cdot : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$ in einem Körper \mathbb{K} ist eine bilineare Abbildung. Seien $f, g : \mathbb{R}^n \rightarrow \mathbb{K}$ differenzierbar. Dann lautet die Produktregel

$$D(f \cdot g)|_x = g(x) \cdot Df|_x + f(x) \cdot Dg|_x.$$

Für $n = 1$ ist das die bekannte Produktregel $(f \cdot g)'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$.

(b) Ein Skalarprodukt $\langle \cdot, \cdot \rangle : W \times W \rightarrow \mathbb{K}$ auf einem normierten Raum W ist eine bilineare Abbildung. Seien $f_1, f_2 : \mathbb{R}^n \rightarrow W$ differenzierbar, dann ist

$$D\langle f_1(\cdot), f_2(\cdot) \rangle|_x y = \langle Df_1|_x y, f_2(x) \rangle + \langle f_1(x), Df_2|_x y \rangle.$$

(c) Die Determinante $\det : \mathbb{R}^m \times \cdots \times \mathbb{R}^m \rightarrow \mathbb{R}$ ist eine multilineare Abbildung. Seien $f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}^m$ differenzierbar, dann ist

$$D \det(f_1(\cdot), \dots, f_m(\cdot))|_x y = \sum_{j=1}^m \det(f_1(x), \dots, Df_j|_x y, \dots, f_m(x)).$$

Erinnerung: Der wichtige Mittelwertsatz der Differentialrechnung in einer Veränderlichen lässt sich für eine **stetig** differenzierbare Funktion $f : [a, b] \rightarrow \mathbb{R}$ aus dem Hauptsatz der Integral- und Differentialrechnung wie folgt ableiten: Substitution $x = a + t(b - a)$ liefert

$$\begin{aligned} f(b) - f(a) &= \int_a^b f'(x) dx = \int_0^1 f'(a + t(b - a)) \cdot (b - a) dt \\ &= \int_0^1 f'(a + th) dt \cdot h = f'(a + \theta h) \cdot h, \end{aligned}$$

wobei die letzte Gleichung mit einer Zahl $\theta \in [0, 1]$ gemäß dem Mittelwertsatz der Integralrechnung gilt. Bis auf diese letzte Gleichheit gilt ein analoges Resultat auch für $f : \mathbb{R}^n \supset G \rightarrow \mathbb{K}^m$.

4.43 Definition. Für eine stetige Funktion $A : [a, b] \rightarrow M(m \times n, \mathbb{K})$, $t \mapsto A(t)$, sei

$$\left(\int_a^b A(t) dt \right)_{ij} := \int_a^b a_{ij}(t) dt.$$

4.44 Satz. Wegintegral der Ableitung (“Hauptsatz entlang von Kurven”)

Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{K}^m$ stetig partiell differenzierbar. Sei $\gamma : [a, b] \rightarrow G$ ein stetig differenzierbarer Weg. Dann gilt

$$f(\gamma(b)) - f(\gamma(a)) = \int_a^b Df|_{\gamma(t)} D\gamma|_t dt.$$

Beweis. Es ist $f_j \circ \gamma : [a, b] \rightarrow \mathbb{K}$ für jede Komponente $j = 1, \dots, m$ eine stetig differenzierbare Funktion. Weiterhin gilt mit der Kettenregel

$$(f_j \circ \gamma)'(t) = D(f_j \circ \gamma)|_t = Df_j|_{\gamma(t)} D\gamma|_t.$$

Also liefert der Hauptsatz aus Analysis 1

$$(f_j \circ \gamma)(b) - (f_j \circ \gamma)(a) = \int_a^b (f_j \circ \gamma)'(t) dt = \int_a^b Df_j|_{\gamma(t)} D\gamma|_t dt,$$

und somit die Behauptung. □

4.45 Korollar. Mittelwertsatz der Differentialrechnung

Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{K}^m$ stetig partiell differenzierbar. Sei $x \in G$ und $h \in \mathbb{R}^n$ derart, dass die Strecke $\{x + th \mid t \in [0, 1]\}$ ganz in G liegt. Dann gilt:

$$f(x + h) - f(x) = \left(\int_0^1 Df|_{x+th} dt \right) \cdot h$$

4.46 Bemerkung. Man beachte, dass man für $m \geq 2$ das Integral $\int_0^1 Df|_{x+th} h dt$ im Allgemeinen **nicht** durch den Wert des Integranden an einer Zwischenstelle $x + \theta h$ ersetzen kann. ◇

Beweis. des Mittelwertsatzes. Wählen wir in Satz 4.44 den Weg $\gamma : [0, 1] \rightarrow G, \gamma(t) = x + th$, so ist $f(\gamma(1)) = f(x + h)$, $f(\gamma(0)) = f(x)$ und $D\gamma|_t = h$. □

4.47 Bemerkung. Die wichtige Konsequenz aus dem Mittelwertsatz für Funktionen in einer Veränderlichen, dass nämlich die Differenz der Funktionswerte durch die Differenz der Argumente mal einer **Schranke** auf die erste Ableitung abgeschätzt werden kann, liefert auch noch diese integrierte Form des Mittelwertsatzes in n Veränderlichen. ◇

4.48 Korollar. Schrankensatz

Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{K}^m$ stetig partiell differenzierbar. Seien $x, y \in G$ und $\gamma : [0, 1] \rightarrow G$ ein stetig differenzierbarer Weg mit $\gamma(0) = x$ und $\gamma(1) = y$. Mit

$$M := \sup \{ \|Df|_{\gamma(t)}\| \mid t \in [0, 1] \} < \infty \tag{4.8}$$

gilt

$$\|f(y) - f(x)\| \leq M \underbrace{\int_0^1 \|\gamma'(t)\| dt}_{\text{Länge des Weges } \gamma}.$$

Der Satz gilt für beliebige Normen auf \mathbb{R}^n und \mathbb{K}^m und die zugehörige Operatornorm in (4.8), wir zeigen ihn aber nur für die euklidische Norm auf dem \mathbb{K}^m , da das folgende Lemma in diesem Fall sehr leicht zu zeigen ist.

4.49 Lemma. Ist $\beta : [0, 1] \rightarrow \mathbb{K}^m$ eine stetige Abbildung und $\|\cdot\|$ die euklidische Norm auf \mathbb{K}^m . Dann gilt:

$$\left\| \int_0^1 \beta(t) dt \right\| \leq \int_0^1 \|\beta(t)\| dt.$$

Beweis. Sei $I = \int_0^1 \beta(t) dt = \left(\int_0^1 \beta_1(t) dt, \dots, \int_0^1 \beta_n(t) dt \right) \in \mathbb{K}^m$. Dann gilt mit Cauchy-Schwarz und der Linearität des Skalarprodukts

$$\|I\|^2 = \langle I, I \rangle = \left\langle \int_0^1 \beta(t) dt, I \right\rangle = \int_0^1 \langle \beta(t), I \rangle dt \leq \int_0^1 \|\beta(t)\| \cdot \|I\| dt = \|I\| \cdot \int_0^1 \|\beta(t)\| dt. \quad \square$$

Beweis. des Schrankensatzes. Mit $D\gamma|_t = \gamma'(t)$ ist

$$\begin{aligned} \|f(y) - f(x)\| &= \left\| \int_0^1 Df|_{\gamma(t)} \cdot D\gamma|_t dt \right\| \leq \int_0^1 \|Df|_{\gamma(t)} \cdot D\gamma|_t\| dt \\ &\leq M \int_0^1 \|D\gamma|_t\| dt = M \int_0^1 \|\gamma'(t)\| dt. \end{aligned}$$

□

4.50 Korollar. Sei $G \subset \mathbb{R}^n$ ein Gebiet, $f : G \rightarrow \mathbb{K}^m$ stetig partiell differenzierbar und $Df|_x = 0$ für alle $x \in G$. Dann ist f auf G konstant.

Beweis. Seien $x, y \in G$ beliebig und $\gamma : [0, 1] \rightarrow G$ ein stetiger Weg mit $\gamma(0) = x$ und $\gamma(1) = y$. (So ein Weg existiert, da G nach Voraussetzung ein Gebiet ist und Gebiete im \mathbb{R}^n offen und wegzusammenhängend sind.) Wenn γ auch differenzierbar ist, so folgt mit dem Schrankensatz sofort

$$\|f(y) - f(x)\| \leq M \int_0^1 \|\gamma'(t)\| dt = 0,$$

da wegen $Df \equiv 0$ auch $M = 0$ ist. Also $f(x) = f(y)$. Nun garantiert uns die Tatsache, dass G als Gebiet auch wegzusammenhängend ist aber lediglich die Existenz eines stetigen Wegs von x nach y . Daher müssen wir entweder zeigen, dass es in einer offenen wegzusammenhängenden Menge im \mathbb{R}^n zu je zwei Punkten auch einen differenzierbaren Verbindungsweg gibt (was nicht allzu schwer ist). Oder wir modifizieren das obige Argument so, dass auch ein stetiger Weg ausreicht (was noch einfacher ist).

Setze $t_0 = \sup\{t \in [0, 1] \mid f(\gamma(\tau)) = f(x) \forall \tau \in [0, t]\}$. Wegen der Stetigkeit von f und γ ist dann auch $f(\gamma(t_0)) = f(x)$. Angenommen $t_0 < 1$, dann gibt es aufgrund der Offenheit von G ein $\delta > 0$ mit $B_\delta(\gamma(t_0)) \subset G$ und aufgrund der Stetigkeit von γ ein $0 < \tilde{\delta} < \delta$, so dass auch $\gamma(t_0 + \tilde{\delta}) \in B_\delta(\gamma(t_0))$ liegt. Da für $h := \gamma(t_0 + \tilde{\delta}) - \gamma(t_0)$ die ganze Strecke $[0, 1] \ni s \mapsto \gamma(t_0) + hs \subset B_\delta(\gamma(t_0)) \subset G$ liegt, gilt nach dem Schrankensatz

$$\|f(\gamma(t_0 + \tilde{\delta})) - f(\gamma(t_0))\| \leq M \|\gamma(t_0 + \tilde{\delta}) - \gamma(t_0)\|$$

mit

$$M = \sup_{s \in [0, 1]} \|Df|_{\alpha(t_0) + sh}\| = 0.$$

Also ist auch noch $f(\gamma(t_0 + \tilde{\delta})) = f(\gamma(t_0)) = f(x)$. Somit ist $t_0 = 1$, d.h. $f(x) = f(y)$ und f ist konstant. □

5 Taylorformel und lokale Extrema

In Analysis 1 wurde gezeigt, dass für eine zweimal stetig differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ die Taylorentwicklung an einem Punkt x_0

$$\begin{array}{cccc}
 f(x_0 + h) & = & f(x_0) & + & f'(x_0) \cdot h & + & \frac{1}{2} f''(x_0) \cdot h^2 & + & o(|h|^2) \\
 & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
 & & \text{0. Ordnung} & & \text{1. Ordnung} & & \text{2. Ordnung} & & \text{Fehler höherer} \\
 & & = \text{konst.} & & = \text{linear} & & = \text{quadratisch} & & \text{Ordnung}
 \end{array}$$

eine quadratische Approximation an die Funktion f in der Nähe von x_0 liefert.

Für Funktionen $f : V \rightarrow W$ zwischen normierten Räumen ist die Differenz $h = (x_0 + h) - x_0 \in V$ ein Vektor und die erste Ableitung $Df : V \rightarrow \mathcal{L}(V, W)$, $x \mapsto Df|_x$, nimmt Werte in den linearen Abbildungen von V nach W an. Die zweite Ableitung, also das Differential der ersten Ableitung, $D^2f|_x := D(Df)|_x$ ist dann eine lineare Abbildung von V nach $\mathcal{L}(V, W)$, was man wiederum als bilineare Abbildung von $V \times V$ nach W auffassen kann,

$$(v_1, v_2) \mapsto D(Df)|_x(v_1, v_2) = \underbrace{(D(Df)|_x v_1)}_{\in \mathcal{L}(V, W)} v_2.$$

Die naheliegende Verallgemeinerung der Taylorschen Formel für solche f ist daher

$$\begin{array}{cccccccc}
 f(x_0 + h) & = & f(x_0) & + & Df|_{x_0} h & + & \frac{1}{2} D^2f|_{x_0}(h, h) & + & \dots & + & \frac{1}{k!} D^k f|_{x_0}(h, \dots, h) \\
 & & \uparrow & & \uparrow & & \uparrow & & & & \uparrow \\
 & & \text{konstant} & & \text{linear} & & \text{bilinear} & & & & \text{k-multilinear} \\
 & & & & & & & & & & + o(\|h\|^k).
 \end{array}$$

Wir werden in dieser Vorlesung die Idee, dass die k -te Ableitung einer Funktion $f : V \rightarrow W$ eine Multilinearform mit k Argumenten ist, nicht weiter vertiefen. Stattdessen betrachten wir nun Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und drücken alle Terme der Taylorentwicklung mittels partieller Ableitungen aus. So sind erste Ableitung $Df|_x : \mathbb{R}^n \rightarrow \mathbb{R}$, $h \mapsto Df|_x h = \langle \nabla f(x), h \rangle =: h \cdot \nabla f(x)$ und die zweite Ableitung $\text{Hess}f(x)$ eine Matrix. Die naheliegende Verallgemeinerung der Taylorschen Formel für solche f ist

$$\begin{aligned}
 f(x_0 + h) &= f(x_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x_0) h_j + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) h_i h_j + o(\|h\|^2) \\
 &= f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle h, \text{Hess}f(x_0) h \rangle + o(\|h\|^2).
 \end{aligned}$$

Diese Formel und die allgemeine Form des Taylorpolynoms wollen wir im Folgenden verstehen.

5.1 Notation. Multiindices und iterierte Richtungsableitung

Um auch die höheren Terme der Taylorentwicklung günstig zu notieren, führt man folgende Schreibweisen ein:

5 Taylorformel und lokale Extrema

(a) **Multiindices:** Für $\alpha \in \mathbb{N}_0^n$, $\alpha = (\alpha_1, \dots, \alpha_n)$, seien

$$|\alpha| := \alpha_1 + \dots + \alpha_n = \sum_{j=1}^n \alpha_j, \quad \text{und} \quad \alpha! := \alpha_1! \cdots \alpha_n! = \prod_{j=1}^n \alpha_j!,$$

und für eine $|\alpha|$ -mal stetig partiell differenzierbare Funktion $f : G \rightarrow \mathbb{R}$ sei

$$\partial^\alpha f := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}},$$

und schließlich für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n} = \prod_{j=1}^n x_j^{\alpha_j}.$$

(b) **Iterierte Richtungsableitung:** Für $f \in C^k(G)$ und $v \in \mathbb{R}^n$ sei

$$(v \cdot \nabla)^k f(x) := \partial_v^k f(x) := \sum_{j_1=1}^n \cdots \sum_{j_k=1}^n v_{j_1} \cdots v_{j_k} \partial_{j_1} \cdots \partial_{j_k} f(x).$$

Wir betrachten in diesem Abschnitt nur Funktionen mit Werten in \mathbb{R} . Allerdings lassen sich die Aussagen wieder leicht auf Funktionen mit Werten in \mathbb{R}^m , \mathbb{C}^m oder beliebigen endlichdimensionalen normierten Räumen übertragen.

5.2 Lemma. Sei $G \subset \mathbb{R}^n$ offen, $f : G \rightarrow \mathbb{R}$ eine k -mal stetig partiell differenzierbare Funktion, $x_0 \in G$ und $h \in \mathbb{R}^n$ derart, dass die geradlinige Verbindung von x_0 nach $x_0 + h$ ganz in G verläuft, also $\{x_0 + ht \mid t \in [0, 1]\} \subset G$. Dann ist die Funktion $\varphi : [0, 1] \rightarrow \mathbb{R}$, $\varphi(t) = f(x_0 + th)$ auch k -mal stetig differenzierbar und es gilt

$$\frac{d^k \varphi}{dt^k}(t) = ((h \cdot \nabla)^k f)(x_0 + th) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x_0 + th) h^\alpha.$$

5.3 Bemerkung. (a) Die Notation $\sum_{|\alpha|=k}$ bedeutet, dass sich die Summe über alle n -Tupel $\alpha \in \mathbb{N}_0^n$ erstreckt, für die $|\alpha| = k$ ist. Davon gibt es $\binom{n-1+k}{k}$ Stück.

(b) Ist $\alpha = (\alpha_1, \dots, \alpha_n)$ mit $k = |\alpha| = \alpha_1 + \dots + \alpha_n$, so gibt es $\frac{k!}{\alpha!}$ Möglichkeiten, eine k -elementige Menge M in n disjunkte Teilmengen S_1, \dots, S_n zu zerlegen, $M = S_1 \dot{\cup} \cdots \dot{\cup} S_n$, so dass S_i gerade α_i Elemente hat ($i = 1, \dots, n$). Oder anders formuliert: es gibt $\frac{k!}{\alpha!}$ Möglichkeiten k verschiedene Kugeln auf n Urnen S_1, \dots, S_n zu verteilen, sodass in der j -ten Urne genau α_j Kugeln liegen. Für $n = 2$ ist beispielsweise $\frac{k!}{\alpha_1! \alpha_2!} = \frac{k!}{\alpha_1!(k-\alpha_1)!} = \binom{k}{\alpha_1}$.

Beweis. von Lemma 5.2. Nach der Kettenregel ist

$$\frac{d}{dt} \varphi(t) = Df|_{x_0+th} h = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x_0 + th) h_j.$$

Nochmals die Kettenregel liefert

$$\frac{d^2}{dt^2} \varphi(t) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0 + th) h_j h_i = \sum_{i_1, i_2=1}^n \partial_{i_1} \partial_{i_2} f(x_0 + th) h_{i_1} h_{i_2},$$

und k -malige Anwendung schließlich

$$\frac{d^k}{dt^k} \varphi(t) = \sum_{i_1, \dots, i_k=1}^n \partial_{i_1} \cdots \partial_{i_k} f(x_0 + th) h_{i_1} \cdots h_{i_k} = ((h \cdot \nabla)^k f)(x_0 + th).$$

Da die Reihenfolge der Differentiationen aber gemäß Satz 4.15 keine Rolle spielt, fassen wir Terme, in denen α_1 -mal nach der ersten Koordinate, α_2 -mal nach der zweiten Koordinate etc. abgeleitet wird, zusammen. Es gibt nun nach Bemerkung 5.3 gerade $\frac{k!}{\alpha_1! \cdots \alpha_n!}$ solche k -Tupel (i_1, \dots, i_k) in denen α_1 -mal der Wert 1, α_2 -mal der Wert 2, \dots , und α_n -mal der Wert n vorkommt. Durch Zusammenfassen der Summanden ergibt sich also

$$\frac{d^k}{dt^k} \varphi(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x_0 + th) h^\alpha. \quad \square$$

5.4 Satz. von Taylor

Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{R}$ eine $(k+1)$ -mal stetig partiell differenzierbare Funktion. Sei $x_0 \in G$ und $h \in \mathbb{R}^n$ derart, dass die Strecke $[x_0, x_0 + h] := \{x_0 + th \mid t \in [0, 1]\}$ ganz in G liegt. Dann gibt es ein $\theta \in [0, 1]$ so, dass

$$\begin{aligned} f(x_0 + h) &= \sum_{m=0}^k \frac{((h \cdot \nabla)^m f)(x_0)}{m!} + \frac{((h \cdot \nabla)^{k+1} f)(x_0 + \theta h)}{(k+1)!} \\ &= \sum_{m=0}^k \sum_{|\alpha|=m} \frac{\partial^\alpha f}{\alpha!}(x_0) h^\alpha + \sum_{|\alpha|=k+1} \frac{\partial^\alpha f}{\alpha!}(x_0 + \theta h) h^\alpha. \end{aligned}$$

5.5 Bemerkung. Für Funktionen $f : G \rightarrow \mathbb{R}^m$ erhält man eine Taylorentwicklung für jede Komponente f_j , $j = 1 \dots, m$.

5.6 Bemerkung. Man nennt

$$P_{f, x_0}^{(k)}(h) := \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x_0)}{\alpha!} h^\alpha$$

das **Taylorpolynom vom Grad k von f in x_0** . Schreiben wir

$$P_{f, x_0}^{(k)}(h) := P_0(h) + P_1(h) + \cdots + P_k(h),$$

so ist

$$\begin{aligned} P_0(h) &= f(x_0) \\ P_1(h) &= \partial_1 f(x_0) h_1 + \cdots + \partial_n f(x_0) h_n = \langle \nabla f(x_0), h \rangle \\ P_2(h) &= \sum_{|\alpha|=2} \frac{1}{\alpha!} \partial^\alpha f(x_0) h^\alpha \\ &= \frac{1}{2} \left(\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) h_i h_j \right) = \frac{1}{2} \langle h, \text{Hess} f(x_0) h \rangle. \end{aligned}$$

Achtung: Nach obiger Definition ist das Taylorpolynom ein Polynom in der Variable h , also im Abstandsvektor $h = x - x_0$ zum Entwicklungspunkt x_0 . Erst durch Einsetzen von $h = x - x_0$ erhält man eine Polynomfunktion in x , welche nach dem Satz von Taylor die Funktion $f(x)$ in der Nähe von x_0 approximiert, also

$$f(x) \approx P_{f, x_0}^{(k)}(x - x_0) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha$$

5 Taylorformel und lokale Extrema

erfüllt. Es macht aber meist keinen Sinn (und wird weder in den Übungen noch in der Klausur erwartet), diesen Ausdruck auszumultiplizieren und in der Form $\sum_{\alpha} c_{\alpha} x^{\alpha}$ zu schreiben, da man an diesem Ausdruck das Verhalten in der Nähe von x_0 nicht mehr ablesen kann. \diamond

Beweis. von Satz 5.4. Betrachte die Kurve $\gamma : (-\varepsilon, 1 + \varepsilon) \rightarrow G$, $\gamma(t) = x_0 + th$, für hinreichend kleines $\varepsilon > 0$ und setze $\varphi : (-\varepsilon, 1 + \varepsilon) \rightarrow \mathbb{R}$, $\varphi(t) = (f \circ \gamma)(t)$. Dann ist φ eine $(k + 1)$ -mal stetig differenzierbare Funktion, also existiert nach der Taylorformel für Funktionen in einer Veränderlichen (vgl. Analysis 1) ein $\theta \in [0, 1]$, so dass gilt

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

Nach Lemma 5.2 erhält man

$$\begin{aligned} f(x_0 + h) &= \varphi(1) = \sum_{m=0}^k \frac{((h \cdot \nabla)^m f)(x_0)}{m!} + \frac{((h \cdot \nabla)^{k+1} f)(x_0 + \theta h)}{(k+1)!} \\ &= \underbrace{\sum_{m=0}^k \sum_{|\alpha|=m} \frac{\partial^{\alpha} f(x_0)}{\alpha!} h^{\alpha}}_{\sum_{|\alpha| \leq k}} + \sum_{|\alpha|=k+1} \frac{\partial^{\alpha} f(x_0 + \theta h)}{\alpha!} h^{\alpha}. \quad \square \end{aligned}$$

5.7 Korollar. Sei $G \subset \mathbb{R}^n$ offen und $f \in C^k(G, \mathbb{R})$. Dann gilt für jedes $x_0 \in G$ und $\delta > 0$ mit $B_{\delta}(x_0) \subset G$, dass für $x \in B_{\delta}(x_0)$

$$f(x) = P_{f, x_0}^{(k)}(x - x_0) + o(\|x - x_0\|^k).$$

Beweis. Nach Taylors Satz gibt es für jedes $h \in B_{\delta}(0)$ ein $\theta = \theta(h) \in [0, 1]$, so dass

$$f(x_0 + h) = \sum_{|\alpha| \leq k-1} \frac{\partial^{\alpha} f(x_0)}{\alpha!} h^{\alpha} + \sum_{|\alpha|=k} \frac{\partial^{\alpha} f(x_0 + \theta h)}{\alpha!} h^{\alpha}.$$

Wir setzen $\varphi : B_{\delta}(0) \rightarrow \mathbb{R}$ als

$$\varphi(h) := \sum_{|\alpha|=k} \left(\frac{\partial^{\alpha} f(x_0 + \theta h)}{\alpha!} - \frac{\partial^{\alpha} f(x_0)}{\alpha!} \right) h^{\alpha}.$$

Wegen der Stetigkeit von $\partial^{\alpha} f$ ist

$$\lim_{h \rightarrow 0} \frac{1}{\alpha!} (\partial^{\alpha} f(x_0 + \theta h) - \partial^{\alpha} f(x_0)) = 0$$

und deshalb gilt

$$\frac{\|\frac{1}{\alpha!} (\partial^{\alpha} f(x_0 + \theta h) - \partial^{\alpha} f(x_0)) \cdot h^{\alpha}\|}{\|h\|^k} \leq \|\frac{1}{\alpha!} (\partial^{\alpha} f(x_0 + \theta h) - \partial^{\alpha} f(x_0))\| \frac{\|h^{\alpha}\|}{\|h\|^k} \rightarrow 0$$

für $h \rightarrow 0$, denn $\frac{\|h^{\alpha}\|}{\|h\|^k} = \frac{|h_1|^{\alpha_1} \dots |h_n|^{\alpha_n}}{\|h\|^{\alpha_1 + \dots + \alpha_n}} \leq 1$. Also ist

$$f(x_0 + h) = \sum_{|\alpha| \leq k} \frac{\partial^{\alpha} f(x_0)}{\alpha!} h^{\alpha} + \varphi(h) \quad \text{mit } \varphi(h) = o(\|h\|^k). \quad \square$$

5.8 Bemerkung. (a) Für $k = 0$ erhält man die Aussage, dass eine stetige Funktion stetig in x ist, denn

$$f(x+h) = \sum_{|\alpha| \leq 0} \frac{\partial^\alpha f}{\alpha!}(x) + o(\|h\|) = f(x) + o(1),$$

also $\lim_{h \rightarrow 0} f(x+h) = f(x)$.

(b) Für $k = 1$ erhält man die Aussage, dass eine stetig differenzierbare Funktion in x stetig differenzierbar ist, denn

$$f(x+h) = \sum_{|\alpha| \leq 1} \frac{\partial^\alpha f}{\alpha!}(x) + o(\|h\|) = f(x) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) h_i + o(\|h\|) = f(x) + Df(x)h + o(\|h\|).$$

(c) Aber für $k = 2$ erhält man nun die sehr nützliche Aussage, dass eine zweimal stetig differenzierbare Funktion die folgende Darstellung erlaubt:

$$f(x+h) = \sum_{|\alpha| \leq 2} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + o(\|h\|^2) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle h, \text{Hess} f(x) h \rangle + o(\|h\|^2).$$

5.9 Definition. Lokale Extrema

Sei $G \subset \mathbb{R}^n$ und $x \in G$. Man sagt, dass eine Funktion $f : G \rightarrow \mathbb{R}$ in x ein **lokales Maximum** hat, wenn es ein $\delta > 0$ gibt, sodass für alle $y \in B_\delta(x)$ gilt

$$f(y) \leq f(x).$$

Gilt sogar $f(y) < f(x)$ für alle $y \in B_\delta(x) \setminus \{x\}$, so spricht man von einem **strikten lokalen Maximum**.

Ist $f(y) \geq f(x)$ bzw. $f(y) > f(x)$ für alle $y \in B_\delta(x) \setminus \{x\}$, so spricht man von einem **(strikten) lokalen Minimum**. Hat f in x ein lokales Maximum oder Minimum, so spricht man von einem **lokalen Extremum**. \diamond

5.10 Proposition. Notwendiges Kriterium für lokales Extremum

Sei $G \subset \mathbb{R}^n$ offen, $f : G \rightarrow \mathbb{R}$, $x \in G$ und f habe in x ein lokales Extremum. Falls f in x partiell differenzierbar ist, so gilt

$$\nabla f(x) = 0.$$

Beweis. Definiert man für $\delta > 0$ klein genug die Funktionen $h_i : (-\delta, \delta) \rightarrow \mathbb{R}$ durch $h_i(t) = f(x + te_i)$, so ist

$$\frac{d}{dt} h_i(0) = \frac{\partial f}{\partial x_i}(x).$$

Mit f hat aber auch h_i in 0 ein lokales Extremum, also folgt (vgl. Analysis 1) $\frac{d}{dt} h_i(0) = 0$ für $i = 1, \dots, n$. Somit ist auch $\nabla f(x) = 0$. \square

5.11 Definition. Für eine stetig differenzierbare Funktion $f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}^m$ nennt man diejenigen Punkte $x \in G$, für die $Df|_x$ nicht surjektiv ist, die **kritischen Punkte** von f . Alle anderen Punkte heißen **reguläre Punkte**.

Im Fall $m = 1$ sind die kritischen Punkte also genau die Nullstellen des Gradienten ∇f . \diamond

Wie im eindimensionalen Fall ist $\nabla f(x) = 0$ nur ein notwendiges, aber kein hinreichendes Kriterium für das Vorliegen eines lokalen Extremums. Die kritischen Punkte von f , also die Lösungen der Gleichung $\nabla f(x) = 0$, liefern somit die Kandidaten für die lokalen Extrema. Um ein hinreichendes Kriterium zu finden, betrachtet man wie im Fall von $f : \mathbb{R} \rightarrow \mathbb{R}$ die zweite Ableitung.

5.12 Erinnerung. Eine reelle symmetrische $n \times n$ -Matrix $A = (a_{ij})$ heißt

- (i) **positiv definit**, wenn für alle $x \in \mathbb{R}^n \setminus \{0\}$ gilt, dass $\langle x, Ax \rangle_{\mathbb{R}^n} = \sum_{i,j=1}^n a_{ij} x_i x_j > 0$.
- (ii) **negativ definit**, wenn für alle $x \in \mathbb{R}^n \setminus \{0\}$ gilt, dass $\langle x, Ax \rangle_{\mathbb{R}^n} < 0$.
- (iii) **indefinit**, wenn es $x, y \in \mathbb{R}^n$ gibt mit $\langle x, Ax \rangle > 0$ und $\langle y, Ay \rangle < 0$.

In der linearen Algebra wird gezeigt: Jede symmetrische Matrix ist diagonalisierbar, besitzt also eine Orthonormalbasis aus Eigenvektoren. Deshalb ist $A \in \text{Sym}_n$ genau dann positiv bzw. negativ definit, wenn alle Eigenwerte positiv bzw. negativ sind. \diamond

5.13 Satz. Lokale Extrema und die zweite Ableitung

Sei $G \subset \mathbb{R}^n$ offen und $f \in C^2(G, \mathbb{R})$. Sei $x \in G$ eine Nullstelle von ∇f , also $\nabla f(x) = 0$.

- (a) Ist $\text{Hess}f(x)$ positiv definit, so hat f in x ein striktes lokales Minimum.
- (b) Ist $\text{Hess}f(x)$ negativ definit, so hat f in x ein striktes lokales Maximum.
- (c) Ist $\text{Hess}f(x)$ indefinit, so hat f in x kein lokales Extremum.

5.14 Lemma. Sei $A \in \text{Sym}_n$ positiv definit und $\lambda_0 > 0$ der kleinste Eigenwert von A . Dann gilt für alle $x \in \mathbb{R}^n$, dass

$$\langle x, Ax \rangle \geq \lambda_0 \|x\|^2.$$

Beweis. Stelle x in einer Orthonormalbasis aus Eigenvektoren dar. \square

Beweis. von Satz 5.13. (a) Sei $\delta > 0$ so klein, dass $B_\delta(x) \subset G$ ist. Nach Korollar 5.7 gilt dann für $h \in B_\delta(0)$

$$f(x+h) = f(x) + \frac{1}{2} \langle h, \text{Hess}f(x)h \rangle + \varphi(h)$$

mit $\varphi(h) = o(\|h\|^2)$. Wähle nun $0 < \delta' < \delta$ so klein, dass $|\varphi(h)| \leq \frac{\lambda_0}{4} \|h\|^2$ für alle $h \in B_{\delta'}(0)$ gilt, wobei $\lambda_0 > 0$ der kleinste Eigenwert von $\text{Hess}f(x)$ sei. Dann ist für alle $h \in B_{\delta'}(0) \setminus \{0\}$

$$\begin{aligned} f(x+h) &= f(x) + \frac{1}{2} \langle h, \text{Hess}f(x)h \rangle + \varphi(h) \geq f(x) + \frac{\lambda_0}{2} \|h\|^2 - \frac{\lambda_0}{4} \|h\|^2 \\ &= f(x) + \frac{\lambda_0}{4} \|h\|^2 > f(x). \end{aligned}$$

Also hat f an der Stelle x ein striktes lokales Minimum.

(b) Ist $\text{Hess}f(x)$ negativ definit, so ist $\text{Hess}(-f)(x) = -\text{Hess}f(x)$ positiv definit. Also hat $-f$ ein striktes lokales Minimum und somit f ein striktes lokales Maximum.

(c) Wähle $h \in \mathbb{R}^n \setminus \{0\}$ so, dass $\alpha := \langle h, \text{Hess}f(x)h \rangle > 0$. Dann gilt für alle $t > 0$ mit t klein genug, dass

$$f(x+th) = f(x) + \frac{1}{2} \langle th, \text{Hess}f(x)th \rangle + \varphi(th) = f(x) + \frac{1}{2} \alpha t^2 + \varphi(th).$$

Für $|t|$ klein genug ist aber $|\varphi(th)| \leq \frac{\alpha}{4} t^2$, also

$$f(x+th) \geq f(x) + \frac{\alpha}{4} t^2 > f(x)$$

für $0 < t < \delta$ und δ klein genug. Also gibt es in jeder Umgebung von x einen Punkt $y = x+th$, sodass $f(y) > f(x)$ ist. Das gleiche Argument für \tilde{h} mit $\tilde{\alpha} = \langle \tilde{h}, \text{Hess}f(x)\tilde{h} \rangle < 0$ zeigt, dass in jeder Umgebung von x auch ein Punkt \tilde{y} mit $f(\tilde{y}) < f(x)$ liegt. Also hat f in x kein lokales Extremum. \square

5.15 Beispiel. Minimum und Sattel

- (a) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x^2 + y^2$. Dann ist $\nabla f(x, y) = (2x, 2y)$, also $\nabla f(x, y) = 0 \Leftrightarrow (x, y) = (0, 0)$. Weiterhin ist

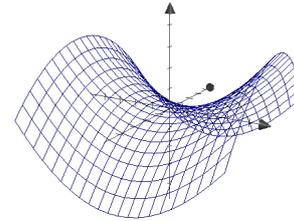
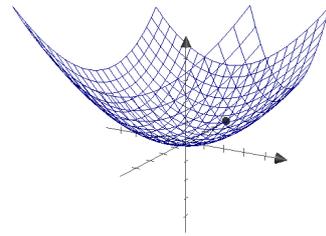
$$\text{Hess}f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \text{Hess}f(0, 0)$$

positiv definit. Somit hat f bei $(0, 0)$ ein striktes lokales Minimum.

- (b) Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x^2 - y^2$ erfüllt $\nabla f(0, 0) = (0, 0)$ und

$$\text{Hess}f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Also hat f in $(0, 0)$ kein lokales Extremum.



5.16 Bemerkung. Ist $\text{Hess}f(x)$ nur positiv semidefinit, gilt also nur $\langle h, \text{Hess}f(x)h \rangle \geq 0$ für alle $h \in \mathbb{R}^n \setminus \{0\}$ und $\nabla f(x) = 0$, so kann man noch nicht entscheiden, ob in x ein lokales Minimum vorliegt. Beispiele:

- (i) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x^2 + y^4$ hat lokales Minimum bei $(0, 0)$
- (ii) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x^2 + y^3$ hat kein lokales Minimum bei $(0, 0)$
- (iii) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x^2$ hat lokales entartetes Minimum bei $(0, 0)$.

Wegen Satz 5.13 (c) ist aber die Semidefinitheit ein notwendiges Kriterium für das Vorliegen eines Extremums.

6 Implizite Funktionen

Oft ist eine Funktion $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ nicht „explizit“ in der Form $y = g(x)$ gegeben, sondern nur „implizit“ durch eine Gleichung der Form

$$F(x, g(x)) = 0 \quad \text{mit} \quad F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad (x, y) \mapsto F(x, y)$$

bestimmt. Man möchte dann die implizite Gleichung $F(x, y) = 0$ explizit machen, d.h. „nach y auflösen“ und in der Form $y = g(x)$ schreiben.

6.1 Beispiele. (a) Betrachte $F : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^2 - 1$. Dann ist

$$C = \{(x, y) \in \mathbb{R}^2 \mid F(x, y) = 0\}$$

die Einheitskreislinie in \mathbb{R}^2 . C ist aber nicht der Graph einer Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$, denn

(i) Für $x \in \mathbb{R}$ mit $|x| > 1$ gibt es kein $y \in \mathbb{R}$ mit $(x, y) \in C$.

(ii) Für $x \in \mathbb{R}$ mit $|x| < 1$ gibt es gleich zwei $y \in \mathbb{R}$ mit $(x, y) \in C$, nämlich $y = \pm\sqrt{1-x^2}$.

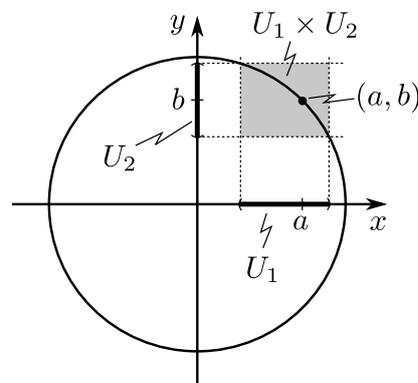
Gibt man allerdings einen Punkt $(a, b) \in C$ mit $b \neq 0$ vor, so kann man $F(x, y)$ lokal um (a, b) nach y auflösen, d.h. es gibt eine Umgebung $U_1 \subset \mathbb{R}$ von a und eine Umgebung U_2 von b und eine Funktion $g : U_1 \rightarrow U_2$ so, dass für alle $(x, y) \in U_1 \times U_2$ gilt

$$F(x, y) = 0 \quad \Leftrightarrow \quad y = g(x).$$

Ist etwa $b > 0$, so wähle $\varepsilon > 0$ klein genug, $U_1 = (a - \varepsilon, a + \varepsilon)$, $U_2 = (b - \varepsilon, b + \varepsilon)$ und

$$g(x) = +\sqrt{1-x^2}.$$

Aber: Für $b = 0$ kann man in keiner Umgebung von $(1, 0)$ oder $(-1, 0)$ nach y auflösen, wohl aber nach x , durch $x = \pm\sqrt{1-y^2}$. Die Punkte $(\pm 1, 0) \in C$ sind aber genau die Punkte, wo $\frac{\partial F}{\partial y} = 2y$ verschwindet, d.h. die Tangente an C vertikal ist.



(b) Eine lineare Abbildung $F : \mathbb{R}_x^n \times \mathbb{R}_y^m \rightarrow \mathbb{R}^m$ ist von der Form

$$F(x, y) = Ax + By$$

mit einer $m \times n$ -Matrix A und einer $m \times m$ -Matrix B . In diesem Fall kann man $F(x, y) = Ax + By \stackrel{!}{=} 0$ offenbar genau dann nach y auflösen, wenn B invertierbar ist und es gilt dann $y = -B^{-1}Ax = -(D_y F)^{-1}D_x Fx$.

Der folgende Satz über implizite Funktionen sagt nun, dass für eine stetig differenzierbare Funktion $F : \mathbb{R}_x^n \times \mathbb{R}_y^m \rightarrow \mathbb{R}^m$ die Gleichung $F(x, y) = 0$ zumindest in einer Umgebung eines Punktes (a, b) mit $F(a, b) = 0$ nach y auflösbar ist, wenn die lineare Approximation

$$F(x, y) \approx DF|_{(a,b)}(x, y) = D_x F|_{(a,b)}(x - a) + D_y F|_{(a,b)}(y - b) = 0$$

nach y auflösbar ist, also wenn $D_y F|_{(a,b)}$ invertierbar ist.

6.2 Satz. Satz über implizite Funktionen

Sei $G \subset \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ offen und $F : G \rightarrow \mathbb{R}^m$ eine stetig partiell differenzierbare Funktion. Sei $(a, b) \in G$ derart, dass $F(a, b) = 0$ ist und

$$D_y F|_{(a,b)} := \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_m} \end{pmatrix} (a, b)$$

invertierbar ist.

Dann existieren offene Umgebungen $U_1 \subset \mathbb{R}^n$ von a und $U_2 \subset \mathbb{R}^m$ von b mit $U_1 \times U_2 \subset G$ und eine stetig partiell differenzierbare Funktion

$$g : U_1 \rightarrow U_2,$$

sodass für alle $(x, y) \in U_1 \times U_2$ gilt:

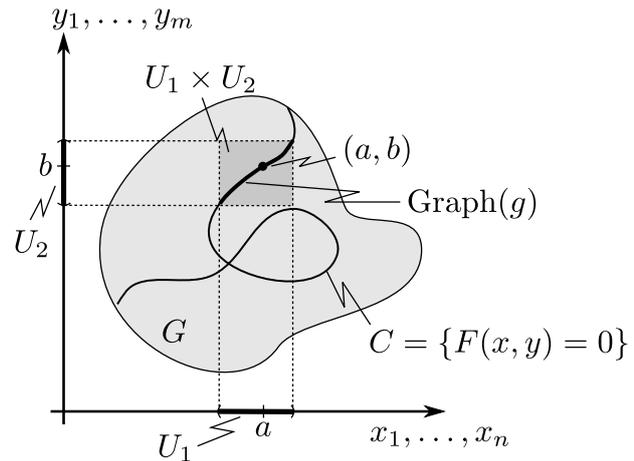
$$F(x, y) = 0 \iff y = g(x).$$

Das Differential von g ist

$$Dg|_x = - [D_y F|_{(x,g(x))}]^{-1} D_x F|_{(x,g(x))},$$

wobei

$$D_x F|_{(x,g(x))} := \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_n} \end{pmatrix} (x, g(x)).$$



6.3 Bemerkungen. (a) Die Formel für die Ableitung von g ergibt sich a posteriori (also wenn man die Existenz einer stetig partiell differenzierbaren Funktion g mit $F(x, g(x)) = 0$ annimmt) auch aus der Kettenregel.

(b) Während die explizite Berechnung der Funktion g in konkreten Anwendungen meist nicht möglich ist, so lässt sich zumindest die Ableitung von g am Punkt a via

$$Dg|_a = - [D_y F|_{(a,b)}]^{-1} D_x F|_{(a,b)}$$

explizit berechnen. Das liefert immerhin $g(x) = g(a) + Dg|_a(x - a) + o(\|x - a\|)$. Nimmt man an, dass F sogar k -mal stetig partiell differenzierbar ist, so kann man zeigen, dass auch g k -mal stetig partiell differenzierbar ist und eine explizite Berechnung des Taylorpolynoms von g bei a bis zur Ordnung k ist möglich.

Beweis. von Satz 6.2. Wir zerlegen den Beweis in die folgenden drei Schritte:

- (a) Wir werden zunächst für U_1, U_2 klein genug mit Hilfe des Banachschen Fixpunktsatzes zeigen, dass $g : U_1 \rightarrow U_2$ mit $F(x, g(x)) = 0$ existiert und eindeutig bestimmt ist.
- (b) Wiederum mit dem Banachschen Fixpunktsatz zeigen wir dann die Stetigkeit von g .
- (c) Die stetige Differenzierbarkeit von g folgt schließlich aus der stetigen Differenzierbarkeit von F .

Zu (a): Setze $B := D_y F|_{(a,b)}$ und betrachte die Funktion $\Phi : G \rightarrow \mathbb{R}^m$ gegeben durch

$$\Phi(x, y) = y - B^{-1} F(x, y).$$

Dann gilt

$$\begin{aligned} F(x, y) = 0 &\Leftrightarrow B^{-1}F(x, y) = 0 \\ &\Leftrightarrow y - B^{-1}F(x, y) = y \\ &\Leftrightarrow \Phi(x, y) = y. \end{aligned}$$

Also ist $F(x, y) = 0$, genau dann wenn y ein Fixpunkt der Abbildung $y \mapsto \Phi(x, y)$ ist.

Weil $D_y\Phi|_{(a,b)} = E_m - B^{-1}D_yF|_{(a,b)} = E_m - B^{-1}B = 0$ ist, ist $\|D_y\Phi|_{(x,y)}\| \leq \frac{1}{2}$ in einer hinreichend kleinen Umgebung von (a, b) , die wiederum das Produkt zweier abgeschlossener Kugeln $\overline{B}_{r_1}(a) \times \overline{B}_{r_2}(b)$ enthält. Der Schrankensatz impliziert dann, dass für (x, y) und (x, y') in dieser Umgebung

$$\|\Phi(x, y) - \Phi(x, y')\| \leq \frac{1}{2}\|y - y'\|$$

gilt. Daraus folgt, dass

$$\Phi(x, \cdot) : \overline{B}_{r_2}(b) \rightarrow \overline{B}_{r_2}(b)$$

für jedes feste x in einer geeigneten Kugel $\overline{B}_{r_3}(a) \subset \overline{B}_{r_1}(a)$ eine Kontraktion ist: für $y \in \overline{B}_{r_2}(b)$ ist

$$\|\Phi(x, y)\| \leq \|\Phi(x, y) - \Phi(x, b)\| + \|\Phi(x, b)\| \leq \frac{1}{2}\|y\| + \frac{1}{2}r_2 \leq r_2,$$

wenn man $\overline{B}_{r_3}(a)$ so klein wählt, dass $\|\Phi(x, b)\| \leq r_2/2$ ist für $x \in \overline{B}_{r_3}(a)$. Letzteres ist wegen $\Phi(a, b) = 0$ und der Stetigkeit von Φ immer möglich. Nach dem Banachschen Fixpunktsatz hat $\Phi(x, \cdot)$ genau einen Fixpunkt, den wir $g(x)$ nennen. Also gibt es zu jedem $x \in \overline{B}_{r_3}$ genau ein $y = g(x) \in \overline{B}_{r_2}$, sodass $F(x, y) = 0$ ist. Insbesondere ist $g(a) = b$.

Zu (b): Man betrachtet die gleiche Abbildung wie in (a), nur für alle $x \in \overline{B}_{r_3}$ gleichzeitig. Sei

$$\mathcal{A} := \{h \in C(\overline{B}_{r_3}, \mathbb{R}^m) \mid \text{Bild}(h) \subset \overline{B}_{r_2}, h(a) = b\}.$$

Es ist \mathcal{A} eine abgeschlossene Teilmenge des Banachraums $C(\mathbb{R}^n, \mathbb{R}^m)$ mit der $\|\cdot\|_\infty$ -Norm. Die Abbildung $\Psi : \mathcal{A} \rightarrow \mathcal{A}$

$$\Psi(h)(x) = \Phi(x, h(x))$$

ist dann wieder eine Kontraktion, denn

$$\|\Psi(h) - \Psi(h')\|_\infty = \sup_{x \in \overline{B}_{r_3}} \|\Phi(x, h(x)) - \Phi(x, h'(x))\| \leq \sup_{x \in \overline{B}_{r_3}} \frac{1}{2}\|h(x) - h'(x)\| = \frac{1}{2}\|h - h'\|_\infty.$$

Der Fixpunkt von Ψ ist

$$\begin{aligned} \Psi(h) = h &\Leftrightarrow \Phi(x, h(x)) = h(x) && \forall x \in \overline{B}_{r_3} \\ &\Leftrightarrow F(x, h(x)) = 0 && \forall x \in \overline{B}_{r_3} \\ &\Leftrightarrow h(x) = g(x) && \forall x \in \overline{B}_{r_3}. \end{aligned}$$

Also ist $g \in \mathcal{A}$ und somit stetig.

Zu (c): Da F stetig differenzierbar ist, gilt mit dem Mittelwertsatz für alle $x, x_0 \in \overline{B}_{r_3}$

$$\begin{aligned} 0 &= F(x, g(x)) - F(x_0, g(x_0)) = F(x, g(x)) - F(x_0, g(x)) + F(x_0, g(x)) - F(x_0, g(x_0)) \\ &= \left(\int_0^1 D_x F|_{(x_0+t(x-x_0), g(x))} dt \right) (x - x_0) + \left(\int_0^1 D_y F|_{(x_0, g(x_0)+t(g(x)-g(x_0)))} dt \right) (g(x) - g(x_0)) \\ &=: A(x, x_0)(x - x_0) + B(x, x_0)(g(x) - g(x_0)), \end{aligned}$$

wobei $A : \overline{B}_{r_3} \times \overline{B}_{r_3} \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ und $B : \overline{B}_{r_3} \times \overline{B}_{r_3} \rightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ stetige Abbildungen sind. (Beachte hierzu, dass die Integranden nach Voraussetzung stetige Funktionen auf dem Kompaktum $\overline{B}_{r_3} \times \overline{B}_{r_3} \times [0, 1]$ sind und somit gleichmäßig stetig. Satz 8.35 aus Analysis 1 zur

6 Implizite Funktionen

Vertauschbarkeit von Grenzwert und Integral impliziert daher die Stetigkeit von A und B .) Nun ist $B(a, b) = D_y F|_{(a,b)}$ nach Annahme invertierbar also $\det(B(a, b)) = d \neq 0$. Da auch $\det: \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m) \rightarrow \mathbb{R}$ stetig ist, gibt es eine Umgebung $B_{r_4} \times B_{r_4} \subset \mathbb{R}^n \times \mathbb{R}^n$ von (a, b) auf der $|\det(B)| > |d|/2$ ist, auf der also B invertierbar und B^{-1} beschränkt bleibt.

Für $x, x_0 \in B_{r_4}$ gilt somit

$$g(x) - g(x_0) = -B(x, x_0)^{-1} A(x, x_0) \cdot (x - x_0).$$

Aufgrund der Stetigkeit von A und B gilt

$$\lim_{x \rightarrow x_0} B^{-1}(x, x_0) A(x, x_0) = B^{-1}(x_0, x_0) A(x_0, x_0) = D_y F|_{(x_0, g(x_0))}^{-1} D_x F|_{(x_0, g(x_0))},$$

also

$$Dg|_{x_0} = -D_y F|_{(x_0, g(x_0))}^{-1} D_x F|_{(x_0, g(x_0))},$$

was schließlich die Existenz und Stetigkeit der partiellen Ableitungen von g zeigt. \square

6.4 Beispiel. Sei $G \subset \mathbb{R}^n$ offen und $f: G \rightarrow \mathbb{R}$ stetig differenzierbar. Wir betrachten die Niveaufläche

$$N_c = \{x \in G \mid f(x) = c\}$$

zu $c \in \mathbb{R}$ von f . Ist nun $a \in N_c$ ein regulärer Punkt von f , also $\nabla f(a) \neq 0$, so lässt sich N_c bei a lokal als Graph einer stetig differenzierbaren Funktion g in $n-1$ Veränderlichen darstellen. Denn sei z.B. $\frac{\partial f}{\partial x_n}(a) \neq 0$ und $\bar{x} = (x_1, \dots, x_{n-1})$ und $y = x_n$. Dann ist

$$x \in N_c \Leftrightarrow \underbrace{f(\bar{x}, y) - c}_{=: F(\bar{x}, y)} = 0$$

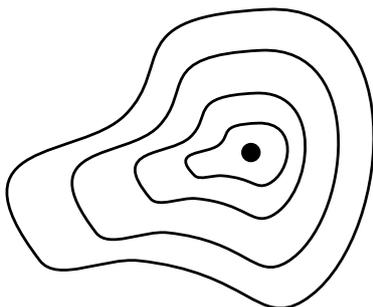
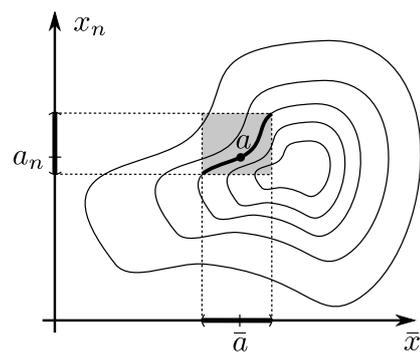
und

$$\frac{\partial F}{\partial y}(\bar{a}, a_n) = \frac{\partial f}{\partial x_n}(\bar{a}, a_n) \neq 0.$$

Nach dem Satz über implizite Funktionen existiert eine Umgebung $\bar{U} \times U_n$ von a und eine C^1 Funktion $g: \bar{U} \rightarrow U_n$ mit $F(\bar{x}, y) = 0 \Leftrightarrow g(\bar{x}) = y$. Somit ist

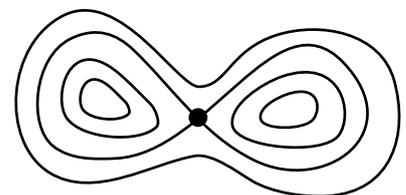
$$N_c \cap (\bar{U} \times U_n) = \text{Graph}(g) \cap (\bar{U} \times U_n).$$

An den kritischen Punkten von f , also dort wo $\nabla f = 0$ ist, gilt das nicht: z.B. ist bei strikten lokalen Extrema N_c ein einzelner Punkt und bei einem Sattelpunkt hat N_c eine Selbstdurchschneidung.



← striktes Extremum

Sattelpunkt →



Wir kommen nun zu der Frage nach der lokalen Umkehrbarkeit einer stetig differenzierbaren Funktion $f : \mathbb{R}^n \supset G \rightarrow D \subset \mathbb{R}^n$. Insbesondere ist es oft wichtig zu wissen, z.B. bei Koordinatentransformationen, dass auch $g = f^{-1} : D \rightarrow G$ wieder stetig differenzierbar ist.

6.5 Definition. Diffeomorphismus

Seien $G, D \subset \mathbb{R}^n$ offen. Eine stetig differenzierbare Abbildung $f : G \rightarrow D$ heißt **Diffeomorphismus**, wenn f bijektiv und $f^{-1} : D \rightarrow G$ stetig differenzierbar ist.

6.6 Bemerkung. Ist $f : G \rightarrow D$ ein Diffeomorphismus, so ist für jedes $x \in G$ die Ableitung $Df|_x$ invertierbar und mit $g = f^{-1}$ gilt $(Df|_x)^{-1} = Dg|_{f(x)}$.

Beweis. Wegen $g \circ f = \text{id}$ liefert die Kettenregel

$$E = D(\text{id})|_x = D(g \circ f)|_x = Dg|_{f(x)} Df|_x. \quad \square$$

6.7 Bemerkung. Nicht jede stetig differenzierbare Bijektion hat auch eine stetig differenzierbare Inverse. Beispielsweise ist $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3$, bijektiv und stetig differenzierbar, die Umkehrabbildung $g : \mathbb{R} \rightarrow \mathbb{R}, y \mapsto \sqrt[3]{y}$ ist aber im Nullpunkt nicht differenzierbar. Beachte, dass die nach Bemerkung 6.6 notwendige Bedingung $f' \neq 0$ für $x = 0$ nicht erfüllt ist, und somit f kein Diffeomorphismus sein kann. \diamond

6.8 Satz. Satz über die Umkehrabbildung

Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Funktion. Ist nun $a \in G$ so, dass $Df|_a \in M(n \times n, \mathbb{R})$ invertierbar ist (also a ein regulärer Punkte von f), so existiert eine offene Umgebung $U \subset G$ von a so, dass $f|_U : U \rightarrow V$ mit $V = f(U)$ ein Diffeomorphismus ist.

Beweis. Sei $b := f(a)$. Man möchte die Gleichung $y = f(x)$ in einer Umgebung von $(x, y) = (a, b)$ nach x auflösen, also $x = g(y)$ schreiben (beachte, dass die Rollen von x und y gegenüber der Notation im Satz über implizite Funktionen hier vertauscht sind!). Dazu betrachtet man die Abbildung

$$F : G \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$F(x, y) = y - f(x).$$

Es ist

$$D_x F|_{(a,b)} = -Df|_a$$

nach Voraussetzung invertierbar. Deshalb liefert der Satz über implizite Funktionen Umgebungen $U_2 \subset G$ von a und $U_1 \subset \mathbb{R}^n$ von b und ein stetig differenzierbares $g : U_1 \rightarrow U_2$ so, dass für alle $(x, y) \in U_2 \times U_1$ gilt

$$F(x, y) = 0 \Leftrightarrow x = g(y).$$

Setzt man noch $U := g(U_1)$, so gilt für alle $(x, y) \in U \times U_1$

$$f(x) = y \Leftrightarrow F(x, y) = 0 \Leftrightarrow x = g(y).$$

Also ist $f|_U : U \rightarrow U_1$ bijektiv, $g = f^{-1}$ und g stetig differenzierbar und somit $f|_U$ ein Diffeomorphismus. \square

6.9 Beispiel. Kugelkoordinaten

Kugelkoordinaten des \mathbb{R}^3 sind gegeben durch

$$f : (0, \infty) \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^3, \quad (r, \theta, \varphi) \mapsto (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta).$$

Eingeschränkt auf

$$G = (0, \infty) \times (0, \pi) \times (-\pi, \pi)$$

ist f injektiv und für die Funktionaldeterminante $\det(Df) : G \rightarrow \mathbb{R}$ gilt

$$\det(Df|_{(r,\theta,\varphi)}) = \begin{vmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{vmatrix} = r^2 \sin \theta \neq 0$$

für alle $(r, \theta, \varphi) \in G$. Deshalb ist f ein lokaler Diffeomorphismus. Da f injektiv ist, ist es aber auch global ein Diffeomorphismus auf sein Bild $D := f(G) \subset \mathbb{R}^3$. \diamond

Häufig sucht man Extrema von Funktionen $f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}$ unter einer Nebenbedingung, die man durch $h(x) = 0$ für eine geeignete Funktion $h : G \rightarrow \mathbb{R}$ beschreiben kann. Beispielsweise sucht man das Maximum einer Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ auf der Einheitskreislinie in \mathbb{R}^2 , also unter der Nebenbedingung $h(x) = \|x\| - 1 = 0$.

6.10 Definition. Extrema unter Nebenbedingungen

Sei $G \subset \mathbb{R}^n$ offen und seien $f, h : G \rightarrow \mathbb{R}$ stetig differenzierbar. Sei $M = \{x \in G \mid h(x) = 0\}$ und $a \in M$. Man sagt, f habe bei a ein **lokales Maximum (bzw. Minimum) unter der Nebenbedingung $h = 0$** , wenn es eine offene Umgebung $U \subset G$ von a gibt, sodass gilt

$$f(x) \leq f(a) \quad (\text{bzw. } f(x) \geq f(a)) \quad \text{für alle } x \in U \cap M.$$

Der Satz über implizite Funktionen liefert ein notwendiges Kriterium für das Vorliegen lokaler Extrema unter Nebenbedingungen.

6.11 Satz. Satz über Extrema unter Nebenbedingungen

Seien $G \subset \mathbb{R}^n$ offen, $f, h : G \rightarrow \mathbb{R}$ stetig differenzierbar und $a \in M = \{x \in G \mid h(x) = 0\}$. Es habe f ein lokales Extremum unter der Nebenbedingung $h = 0$ in a **und** es sei $\nabla h(a) \neq 0$ (also a ein regulärer Punkt von h). Dann gibt es ein $\lambda \in \mathbb{R}$ mit

$$\nabla f(a) = \lambda \nabla h(a).$$

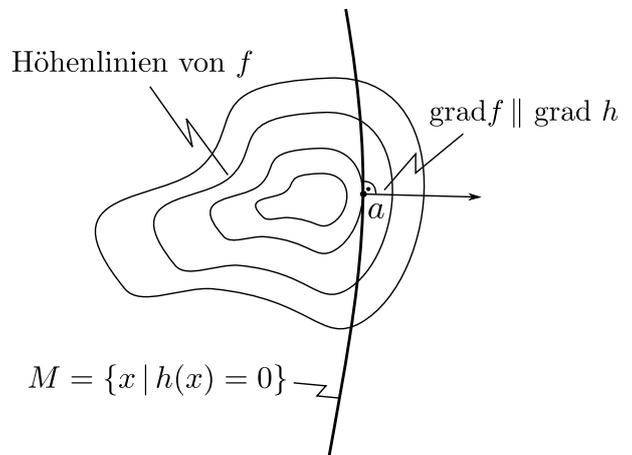
6.12 Bemerkung. Geometrische Bedeutung

Der Gradient von h in a steht senkrecht auf der Niveaufäche M von h . Damit f auf M ein lokales Extremum hat, müssen nur die Richtungsableitungen $\langle v, \nabla f(a) \rangle$ von f tangential an M verschwinden, also

$$\langle v, \nabla f(a) \rangle = 0 \quad \text{falls} \quad \langle v, \nabla h(a) \rangle = 0,$$

d.h. der Gradient von f muss auf M senkrecht stehen. Da M Kodimension 1 hat, folgt daraus

$$\nabla f(a) \parallel \nabla h(a).$$



Man beachte, dass Satz 6.11 das Analogon zu Proposition 5.10 für den Fall ohne Nebenbedingungen ist. Es ist $\nabla f(a) = \lambda \nabla h(a)$ eine notwendige, aber keine hinreichende Bedingung für das Vorliegen eines lokalen Extremums unter der Nebenbedingung $h = 0$.

Beweis. von Satz 6.11. Da $\nabla h(a) \neq 0$ ist, ist wenigstens eine partielle Ableitung von h in a von Null verschieden, sagen wir $\frac{\partial h}{\partial x_n}(a) \neq 0$. Sei $\bar{a} := (a_1, \dots, a_{n-1})$. Der Satz über implizite Funktionen liefert dann Umgebungen $V \subset \mathbb{R}^{n-1}$ von \bar{a} und $I \subset \mathbb{R}$ von a_n mit $V \times I \subset G$ und ein stetig differenzierbares $g : V \rightarrow I$, so dass für alle $(\bar{x}, x_n) \in V \times I$ gilt

$$h(\bar{x}, x_n) = 0 \quad \Leftrightarrow \quad x_n = g(\bar{x}).$$

Hat nun $f : G \rightarrow \mathbb{R}$ ein lokales Extremum unter der Nebenbedingung $h = 0$ in $a \in M$, so hat $f(\bar{x}, g(\bar{x})) =: f \circ \varphi(\bar{x})$ ein lokales Extremum in \bar{a} (ohne Nebenbedingung). Hierbei ist $\varphi : V \rightarrow M \subset \mathbb{R}^n$, $\bar{x} \mapsto \varphi(\bar{x}) = (\bar{x}, g(\bar{x}))$, und man nennt φ eine lokale Parametrisierung von M . Es gilt gemäß Proposition 5.10 also

$$\nabla(f \circ \varphi)(\bar{a}) = 0.$$

Für $i = 1, \dots, n-1$ ergibt die Kettenregel

$$0 = \frac{\partial(f \circ \varphi)}{\partial x_i}(\bar{a}) = \sum_{j=1}^n \partial_j f(\varphi(\bar{a})) \frac{\partial \varphi_j}{\partial x_i}(\bar{a}) = \partial_i f(a) + \partial_n f(a) \frac{\partial g}{\partial x_i}(\bar{a}). \quad (*)$$

Andererseits gilt wegen $(h \circ \varphi)(\bar{x}) = 0$ für alle $\bar{x} \in V$, dass

$$0 = \frac{\partial(h \circ \varphi)}{\partial x_i}(\bar{a}) = \partial_i h(a) + \partial_n h(a) \frac{\partial g}{\partial x_i}(\bar{a}),$$

also wegen $\frac{\partial h}{\partial x_n}(a) \neq 0$

$$\frac{\partial g}{\partial x_i}(\bar{a}) = - \left(\frac{\partial h}{\partial x_n}(a) \right)^{-1} \frac{\partial h}{\partial x_i}(a). \quad (**)$$

Setzen wir nun

$$\lambda := \frac{\frac{\partial f}{\partial x_n}(a)}{\frac{\partial h}{\partial x_n}(a)} \in \mathbb{R},$$

so folgt zunächst für $i = n$

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

aber dann auch für $i = 1, \dots, n-1$, wenn man $(**)$ in $(*)$ einsetzt, dass

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

also $\nabla f(a) = \lambda \nabla h(a)$. □

6.13 Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto y^2 - x^2$ und $K := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ die abgeschlossene Kreisscheibe vom Radius 1. Da f stetig ist und K kompakt, nimmt f auf K ihr Supremum $c := \sup_{(x,y) \in K} f(x, y)$ an. Wir wollen c und die Stellen $(x, y) \in K$ wo f den Wert c annimmt, berechnen. Dazu suchen wir zunächst lokale Extrema im Inneren und auf dem Rand.

(i) **Innere Punkte:** Da $\nabla f(x, y) = (-2x, 2y) \neq (0, 0)$ für $(x, y) \neq (0, 0)$ und

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

indefinit ist, hat f in $\overset{\circ}{K} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ kein lokales Extremum. Die Extrema von f müssen also auf dem Rand

$$M = \{(x, y) \mid x^2 + y^2 = 1\}$$

liegen.

6 Implizite Funktionen

(ii) **Randpunkte:** Betrachte die Nebenbedingung $h = 0$ für

$$h : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2 - 1.$$

Ein Maximum von f auf einem Randpunkt $a \in M$ ist insbesondere auch ein Maximum von f unter der Nebenbedingung $h = 0$. Es muss also gelten, dass

$$\nabla f(a) = \lambda \nabla h(a)$$

für ein $\lambda \in \mathbb{R}$. Beachte, dass $\nabla h(x, y) = (2x, 2y) \neq 0$ für $(x, y) \in M$. Wir erhalten somit die Gleichungen

$$\begin{aligned} \text{(I.)} \quad & -2x = \lambda(2x) \\ \text{(II.)} \quad & 2y = \lambda(2y) \\ \text{(III.)} \quad & x^2 + y^2 = 1 \end{aligned}$$

Ist nun $x \neq 0$, so folgt aus I., dass $\lambda = -1$ und dann aus II. und III., dass $y = 0$ und $x = \pm 1$. Ist dagegen $y \neq 0$, so folgt $x = 0$ und $y = \pm 1$. Man erhält also die vier Kandidaten

$$a \in \{(1, 0), (0, 1), (-1, 0), (0, -1)\}.$$

Da $f(\pm 1, 0) = -1$ und $f(0, \pm 1) = +1$ ist, gilt $c = +1$ und das Supremum wird genau in den Punkten $(0, 1)$ und $(0, -1)$ angenommen. \diamond

6.14 Bemerkung. Konkret findet man also alle Kandidaten für lokale Extrema einer Funktion f auf dem \mathbb{R}^n unter der Nebenbedingung $h = 0$, indem man die Lösungen der $n + 1$ Gleichungen

$$\partial_i f(x) + \lambda \partial_i h(x) = 0, \quad i = 1, \dots, n, \quad \text{und} \quad h(x) = 0$$

in den $n + 1$ Unbekannten $x = (x_1, \dots, x_n)$ und λ bestimmt. Die Zahl λ nennt man auch den **Lagrangeschen Multiplikator**. \diamond

6.15 Beispiel. Wir geben nun einen alternativen Beweis für die Tatsache, dass jede symmetrische Matrix mindestens einen reellen Eigenwert hat. Sei dazu $A \in M(n \times n, \mathbb{R})$ symmetrisch und

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \langle x, Ax \rangle, \quad \text{sowie} \quad h : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \langle x, x \rangle - 1.$$

Die Fläche $\{h = 0\}$ ist die Einheitssphäre S^{n-1} im \mathbb{R}^n und ist kompakt. Damit nimmt die stetige Funktion f auf S^{n-1} ihr Maximum und ihr Minimum an. Sei x_0 ein Punkt an dem f auf S^{n-1} maximal wird. Dann gilt nach Satz 6.11, dass

$$\nabla f(x_0) = 2Ax_0 = \lambda \nabla h(x_0) = \lambda 2x_0$$

für ein $\lambda \in \mathbb{R}$. Also ist λ Eigenwert zum Eigenvektor x_0 . \diamond

In Analogie zu Satz 5.13 kann man mithilfe der zweiten Ableitung auch ein hinreichendes Kriterium für Extrema unter Nebenbedingungen zeigen.

6.16 Satz. Hinreichendes Kriterium

Es seien $G \subset \mathbb{R}^n$ offen und $f, h \in C^2(G, \mathbb{R})$. Am Punkt $a \in M = \{x \in G \mid h(x) = 0\}$ sei die notwendige Bedingung aus Satz 6.11 erfüllt, also $\nabla h(a) \neq 0$ und es gibt ein $\lambda \in \mathbb{R}$ mit $\nabla F(a) = 0$ für $F := f + \lambda h$.

- (i) Ist $\langle v, \text{Hess } F(a) v \rangle > 0$ für alle $v \in \mathbb{R}^n \setminus \{0\}$ mit $\langle v, \nabla h(a) \rangle = 0$, so hat f bei a ein striktes lokales Minimum unter der Nebenbedingung $h = 0$.
- (ii) Ist $\langle v, \text{Hess } F(a) v \rangle < 0$ für alle $v \in \mathbb{R}^n \setminus \{0\}$ mit $\langle v, \nabla h(a) \rangle = 0$, so hat f bei a ein striktes lokales Maximum unter der Nebenbedingung $h = 0$.

- (iii) Gibt es $v, v' \in \mathbb{R}^n$ mit $\langle v, \nabla h(a) \rangle = \langle v', \nabla h(a) \rangle = 0$, $\langle v, \text{Hess } F(a) v \rangle < 0$ und $\langle v', \text{Hess } F(a) v' \rangle > 0$, so hat f bei a kein lokales Extremum unter der Nebenbedingung $h = 0$.

Es gilt also Satz 5.13 entsprechend, wenn man die Definitheit von $\text{Hess}F(a)$ eingeschränkt auf den Tangentialraum an M im Punkt a fordert, denn die Vektoren $v \in \mathbb{R}^n$ mit $\langle v, \nabla h(a) \rangle = 0$ sind ja gerade tangential an M im Punkt a .

Achtung: Die zunächst naheliegende Bedingung

- (i)' $\langle v, \text{Hess } f(a) v \rangle > 0$ für alle $v \in \mathbb{R}^n$ mit $\langle v, \nabla h(a) \rangle = 0$

ist **nicht** hinreichend für das Vorliegen eines Minimums.

Beweis. Da a ein regulärer Punkt von h ist, können wir wie im Beweis von Satz 6.11 auflösen und setzen $f(\bar{x}, g(\bar{x})) =: (f \circ \varphi)(\bar{x})$. Dann ist $h \circ \varphi \equiv 0$ und somit

$$0 = \frac{\partial(h \circ \varphi)}{\partial x_i}(\bar{x}) = \frac{\partial^2(h \circ \varphi)}{\partial x_j \partial x_i}(\bar{x}),$$

was wiederum sofort

$$\frac{\partial(f \circ \varphi)}{\partial x_i}(\bar{x}) = \frac{\partial(F \circ \varphi)}{\partial x_i}(\bar{x}) \quad \text{und} \quad \frac{\partial^2(f \circ \varphi)}{\partial x_j \partial x_i}(\bar{x}) = \frac{\partial^2(F \circ \varphi)}{\partial x_j \partial x_i}(\bar{x})$$

impliziert. Wir zeigen nun (i), indem wir mithilfe des hinreichenden Kriteriums ohne Nebenbedingung (Satz 5.13) nachweisen, dass $f \circ \varphi$ in \bar{a} ein lokales Minimum hat. Dann hat f offenbar ein lokales Minimum in a unter der Nebenbedingung $h = 0$. Für (ii) und (iii) verfährt man analog.

Zunächst gilt für $i = 1, \dots, n-1$

$$\frac{\partial(f \circ \varphi)}{\partial x_i}(\bar{a}) = \frac{\partial(F \circ \varphi)}{\partial x_i}(\bar{a}) = \langle \nabla F(a), \frac{\partial \varphi}{\partial x_i}(\bar{a}) \rangle = 0.$$

Weiterhin ist für $i, j = 1, \dots, n-1$

$$\begin{aligned} \frac{\partial^2(f \circ \varphi)}{\partial x_j \partial x_i}(\bar{a}) &= \frac{\partial^2(F \circ \varphi)}{\partial x_j \partial x_i}(\bar{a}) = \partial_j \left(\sum_{\ell=1}^n (\partial_\ell F \circ \varphi) \frac{\partial \varphi_\ell}{\partial x_i} \right) (\bar{a}) \\ &= \sum_{\ell=1}^n \left(\sum_{k=1}^n (\partial_k \partial_\ell F \circ \varphi) \frac{\partial \varphi_\ell}{\partial x_i} \frac{\partial \varphi_k}{\partial x_j} + (\partial_\ell F \circ \varphi) \frac{\partial^2 \varphi_\ell}{\partial x_j \partial x_i} \right) (\bar{a}) \\ &= \sum_{\ell=1}^n \sum_{k=1}^n (\partial_k \partial_\ell F \circ \varphi) \frac{\partial \varphi_\ell}{\partial x_i} \frac{\partial \varphi_k}{\partial x_j} \\ &= (D\varphi|_{\bar{a}}^T \text{Hess}F(a) D\varphi|_{\bar{a}})_{ji}, \end{aligned}$$

wobei wir in der dritten Gleichheit $(\nabla F \circ \varphi)(\bar{a}) = \nabla F(a) = 0$ verwendet haben (An dieser Stelle ist es wichtig, dass wir F statt f betrachten!). Nun gilt für jeden Vektor $\bar{v} \in \mathbb{R}^{n-1}$, dass

$$0 = D(h \circ \varphi)|_{\bar{a}} \bar{v} = Dh|_a D\varphi|_{\bar{a}} \bar{v} = \langle \nabla h(a), D\varphi|_{\bar{a}} \bar{v} \rangle$$

und somit folgt die positive Definitheit von $\text{Hess}(f \circ \varphi)(a)$ aus (i),

$$\langle \bar{v}, \text{Hess}(f \circ \varphi)(a) \bar{v} \rangle = \langle \bar{v}, D\varphi|_{\bar{a}}^T \text{Hess}F(a) D\varphi|_{\bar{a}} \bar{v} \rangle = \langle D\varphi|_{\bar{a}} \bar{v}, \text{Hess}F(a) D\varphi|_{\bar{a}} \bar{v} \rangle > 0. \quad \square$$

6.17 Bemerkung. Mehrere Nebenbedingungen

Liegen mehrere Nebenbedingungen h_1, \dots, h_k vor, sucht man also ein Extremum von f auf der $(n-k)$ -dimensionalen "Fläche"

$$\{h_1 = 0\} \cap \dots \cap \{h_k = 0\},$$

6 Implizite Funktionen

so ergibt sich die notwendige Bedingung, dass Richtungsableitungen $\langle v, \nabla f(a) \rangle$ von f die tangential an alle Hyperflächen $M_j := \{h_j = 0\}$ liegen, verschwinden. Also muss gelten

$$\langle v, \nabla h_j(a) \rangle = 0 \quad \forall j = 1, \dots, k \quad \Rightarrow \quad \langle v, \nabla f(a) \rangle = 0.$$

Falls a ein regulärer Punkt für die vektorwertige Funktion $h : \mathbb{R}^n \rightarrow \mathbb{R}^k$, $h = (h_1, \dots, h_k)$, ist, die Menge $\{\nabla h_1(a), \dots, \nabla h_k(a)\}$ also linear unabhängig ist (man spricht auch von unabhängigen Nebenbedingungen), so ergibt sich als notwendige Bedingung für das Vorliegen eines Extremums von $f|_{\cap_{j=1}^k M_j}$ somit, dass

$$\nabla f(a) \in \text{Bild}(Dh|_a) = \text{span}\{\nabla h_1(a), \dots, \nabla h_k(a)\}. \quad (*)$$

Es ist (*) wieder äquivalent zu

$$\nabla(f + \lambda \cdot h) = 0,$$

wobei der Lagrange-Multiplikator λ nun ein Vektor im \mathbb{R}^k ist. Auch das hinreichende Kriterium gilt entsprechend mit $F = f + \lambda \cdot h$. \diamond

7 Differentialrechnung in Banachräumen

Fast alle Resultate der vorhergehenden Kapitel gelten im Wesentlichen unverändert, wenn man den \mathbb{R}^n durch einen beliebigen Banachraum, also einen vollständigen normierten Raum $(X, \|\cdot\|)$, ersetzt. Wir werden das aus Zeitmangel nicht im Detail ausführen können, besprechen aber in diesem kurzen Kapitel die grundlegende Definition und eine Beispielanwendung.

7.1 Definition. Fréchet Ableitung

Seien X und Y Banachräume und $G \subset X$ offen. Eine Abbildung $f : G \rightarrow Y$ heißt differenzierbar im Punkt x in G , wenn es eine stetige lineare Abbildung $A : X \rightarrow Y$ gibt so, dass

$$f(x+h) = f(x) + Ah + o(\|h\|)$$

für h in einer hinreichend kleinen Umgebung der Null.

Wie im endlichdimensionalen Fall ist A eindeutig bestimmt und heißt **Fréchet Ableitung** von f bei x bezeichnet mit $Df|_x$. ◇

- 7.2 Bemerkung.** (a) Ein wichtiger Punkt in obiger Definition ist, dass die Stetigkeit von A gefordert wird. In unendlichdimensionalen normierten Räumen impliziert die Linearität einer Abbildung nämlich nicht ihre Stetigkeit.
- (b) Wieder impliziert die Differenzierbarkeit einer Funktion in einem Punkt die Stetigkeit der Funktion in diesem Punkt, vgl. Satz 4.32.
- (c) Der Mittelwertsatz, der Satz über implizite Funktionen, der Satz über die Umkehrabbildung und die Aussagen über lokale Extrema mit und ohne Nebenbedingungen gelten analog auch für differenzierbare Funktionen auf Banachräumen.

7.3 Beispiel. Euler-Lagrange-Gleichungen in der klassischen Mechanik

Ohne Beweis stellen wir zunächst fest, dass der Raum der zweimal stetig differenzierbaren Pfade $X := \{x : [0, T] \rightarrow \mathbb{R}^n \mid x \text{ ist zweimal stetig differenzierbar}\}$ mit der Norm

$$\|x\|_X := \|x\|_\infty + \|\dot{x}\|_\infty + \|\ddot{x}\|_\infty := \sup_{t \in [0, T]} \|x(t)\|_{\mathbb{R}^n} + \sup_{t \in [0, T]} \|\dot{x}(t)\|_{\mathbb{R}^n} + \sup_{t \in [0, T]} \|\ddot{x}(t)\|_{\mathbb{R}^n}$$

ein Banachraum ist, wobei die Punkte wie in der Physik üblich Zeitableitungen, also Ableitungen nach t , darstellen.

Man ordnet nun jedem Weg $x \in X$ eine sogenannte Wirkung zu, d.h. man definiert

$$S : X \rightarrow \mathbb{R}, \quad x \mapsto S(x) := \int_0^T L(x(t), \dot{x}(t)) dt,$$

wobei $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $(q, v) \mapsto L(q, v)$ eine zweimal stetig differenzierbare Funktion, die sogenannte Lagrangefunktion, ist. Für ein nichtrelativistisches Teilchen mit Masse $m > 0$ in einem differenzierbaren Potential $V : \mathbb{R}^n \rightarrow \mathbb{R}$ ist beispielsweise

$$L(q, v) := \frac{1}{2}m\|v\|_{\mathbb{R}^n}^2 - V(q). \tag{7.1}$$

7 Differentialrechnung in Banachräumen

Wir bestimmen zunächst die Ableitung DS von S . Dazu stellen wir fest, dass für $h \in X$

$$\begin{aligned} S(x+h) &= \int_0^t L(x(t)+h(t), \dot{x}(t)+\dot{h}(t)) dt \\ &= \int_0^T \left(L(x(t), \dot{x}(t)) + \left\{ D_q L|_{(x(t), \dot{x}(t))} h(t) + D_v L|_{(x(t), \dot{x}(t))} \dot{h}(t) \right\} \right) dt \\ &\quad + \mathcal{O}(\|h\|_X^2). \end{aligned}$$

Partielle Integration im zweiten Summanden liefert

$$\begin{aligned} S(x+h) - S(x) &= \int_0^T \left(D_q L|_{(x(t), \dot{x}(t))} h(t) + D_v L|_{(x(t), \dot{x}(t))} \dot{h}(t) \right) dt + \mathcal{O}(\|h\|_X^2) \\ &= D_v L|_{(x(T), \dot{x}(T))} h(T) - D_v L|_{(x(0), \dot{x}(0))} h(0) \\ &\quad + \int_0^T \left\{ D_q L|_{(x(t), \dot{x}(t))} - \left(\frac{d}{dt} D_v L|_{(x(t), \dot{x}(t))} \right) \right\} h(t) dt + \mathcal{O}(\|h\|_X^2), \end{aligned}$$

wobei $\mathcal{O}(\|h\|_X^2)$ bedeutet, dass der Restterm betragsmäßig durch eine von h unabhängige Konstante mal $\|h\|_X^2$ beschränkt ist. Die Ableitung $DS|_x : X \rightarrow \mathbb{R}$,

$$\begin{aligned} DS|_x h &= D_v L|_{(x(T), \dot{x}(T))} h(T) - D_v L|_{(x(0), \dot{x}(0))} h(0) \\ &\quad + \int_0^T \left\{ D_q L|_{(x(t), \dot{x}(t))} - \left(\frac{d}{dt} D_v L|_{(x(t), \dot{x}(t))} \right) \right\} h(t) dt \end{aligned}$$

existiert also an jedem Punkt $x \in X$ als stetige lineare Abbildung von X nach \mathbb{R} .

Die physikalisch realisierbaren Trajektorien x bekommt man nun als kritische Punkte der Wirkung bei festgehaltenen Endpunkten $x(0) = x_0$ und $x(T) = x_1$. Wir fordern also $DS|_x h = 0$ für alle $h \in X$ mit $h(0) = h(T) = 0$, was auf die Euler-Lagrange-Gleichung

$$D_q L|_{(x(t), \dot{x}(t))} - \frac{d}{dt} D_v L|_{(x(t), \dot{x}(t))} = 0$$

führt.

Im letzten Schritt hätten wir auch unter den jeweils n Nebenbedingungen

$$H_0(x) = x(0) - x_0 = 0 \quad \text{und} \quad H_1(x) = x(T) - x_1 = 0$$

minimieren können. Wegen

$$DH_0|_x h = h(0) \quad \text{und} \quad DH_1|_x h = h(T)$$

folgt dann aus

$$\begin{aligned} \left(DS|_x - \lambda_0 DH_0|_x - \lambda_1 DH_1|_x \right) h &= \left(D_v L|_{(x(T), \dot{x}(T))} - \lambda_1 \right) h(T) - \left(D_v L|_{(x(0), \dot{x}(0))} + \lambda_0 \right) h(0) \\ &\quad + \int_0^T \left\{ D_q L|_{(x(t), \dot{x}(t))} - \left(\frac{d}{dt} D_v L|_{(x(t), \dot{x}(t))} \right) \right\} h(t) dt \\ &\stackrel{!}{=} 0, \end{aligned}$$

dass $\lambda_0 = -D_v L|_{(x(0), \dot{x}(0))}$ und $\lambda_1 = D_v L|_{(x(T), \dot{x}(T))}$, sowie wiederum die Euler-Lagrange Gleichung

$$D_q L|_{(x(t), \dot{x}(t))} - \frac{d}{dt} D_v L|_{(x(t), \dot{x}(t))} = 0.$$

Für die spezielle Lagrangefunktion (7.1) lautet diese Differentialgleichung schließlich

$$-\nabla V(x(t)) - m \ddot{x}(t) = 0$$

oder umgestellt

$$m \ddot{x}(t) = -\nabla V(x(t)).$$

Dies ist die Newtonsche Differentialgleichung “Masse mal Beschleunigung = Kraft”, wobei die auf den Körper ausgeübte Kraft durch den negativen Gradienten des Potentials gegeben ist.

8 Gewöhnliche Differentialgleichungen

Differentialgleichungen sind Gleichungen für Funktionen, die sowohl die Funktion selbst als auch Ableitungen der Funktion beinhalten. Beispielsweise ist für $f : \mathbb{R} \rightarrow \mathbb{R}$ die Gleichung

$$f' = f \tag{8.1}$$

eine Differentialgleichung. Hier ist gemeint, dass f' und f als Funktionen gleich sind, also $f'(x) = f(x)$ für alle $x \in \mathbb{R}$ gilt. Jede Funktion $f \in C^1(\mathbb{R})$, die die Gleichung (8.1) erfüllt, heißt dann Lösung der Differentialgleichung. Eine Lösung von (8.1) ist die Exponentialfunktion, $f(x) = e^x$. Gibt es noch andere Lösungen?

Ein weiteres Beispiel ist die Differentialgleichung

$$\frac{\partial g}{\partial t} = \Delta_x g, \tag{8.2}$$

die sogenannte Wärmeleitungsgleichung. Dies ist eine Gleichung für Funktionen $g : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$, $(t, x) \mapsto g(t, x)$.

Differentialgleichungen für Funktionen die nur von einer reellen Variable abhängen nennt man **gewöhnliche Differentialgleichungen** (beispielsweise (8.1)), während Differentialgleichungen für Funktionen auf dem \mathbb{R}^n **partielle Differentialgleichungen** heißen (beispielsweise (8.2)).

Wir beschäftigen uns in dieser Vorlesung nur mit gewöhnlichen Differentialgleichungen. Für diese existiert eine sehr allgemeine und übersichtliche Lösungstheorie. Wie wir sehen werden kann man eine sehr große Klasse von gewöhnlichen Differentialgleichungen auf eine Standardform bringen, die eine natürliche geometrische bzw. dynamische Interpretation erlaubt.

8.1 Definition. Autonome Differentialgleichung erster Ordnung

Sei $G \subset \mathbb{R}^n$ offen, $v : G \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld und $I \subset \mathbb{R}$ ein offenes Intervall. Eine differenzierbare Kurve $\gamma : I \rightarrow G$ heißt **Lösung der autonomen Differentialgleichung erster Ordnung**

$$\gamma' = v(\gamma), \quad (\text{oder äquivalent } \gamma' = v \circ \gamma) \tag{8.3}$$

falls für alle $t \in I$ gilt, dass

$$\gamma'(t) = v(\gamma(t)).$$

Falls $0 \in I$ ist, so heißt γ eine **Lösung zum Anfangswert** $x_0 \in G$ falls $\gamma(0) = x_0$ ist.

In Komponenten lautet (8.3)

$$\begin{aligned} \gamma'_1(t) &= v_1(\gamma_1(t), \dots, \gamma_n(t)) \\ &\vdots \\ \gamma'_n(t) &= v_n(\gamma_1(t), \dots, \gamma_n(t)), \end{aligned}$$

weshalb man auch von einem **System autonomer gewöhnlicher Differentialgleichungen** für die n Funktionen $\gamma_1, \dots, \gamma_n : \mathbb{R} \rightarrow \mathbb{R}$ spricht. ◇

8.2 Bemerkung. Zur Terminologie: Wie schon zuvor erklärt bedeutet “gewöhnlich” hier, dass nur Ableitungen bezüglich einer reellen Variable auftauchen. “Erste Ordnung” bedeutet, dass nur die erste Ableitung von γ eingeht. “Autonom” bezeichnet schließlich die Eigenschaft, dass das

Vektorfeld v nicht explizit von der Zeit t abhängt. In der Physik wird oft die Notation $\dot{\gamma}$ statt γ' für die Ableitung übernommen, wo der Punkt üblicherweise eine Zeitableitung symbolisiert. Man hat also die Vorstellung, dass sich der Punkt $\gamma(t) \in G \subset \mathbb{R}^n$ mit der Geschwindigkeit $\dot{\gamma}(t)$ entlang der Kurve γ "bewegt", wobei er zum Zeitpunkt $t = 0$ im Anfangswert $x_0 = \gamma(0)$ startet. Die Differentialgleichung gibt dann also vor, dass eine Lösung am "Ort" $x \in G$ die "Geschwindigkeit" $v(x)$ haben muss. \diamond

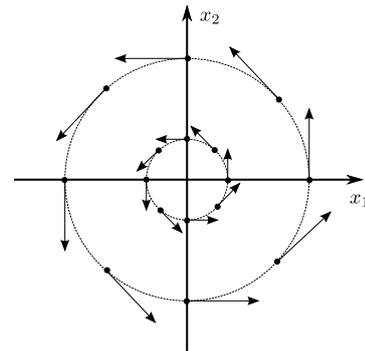
8.3 Bemerkung. Phasendiagramm

Man verdeutlicht sich das Vektorfeld $v : G \rightarrow \mathbb{R}^n$ oft durch das sogenannte **Phasendiagramm**, wobei man im **Phasenraum** G sowohl das Vektorfeld v als auch einige Lösungskurven γ skizziert.

Sei beispielsweise $v : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x_1, x_2) \mapsto (-x_2, x_1)$, die zugehörige DGL also

$$\gamma' = v(\gamma) = \begin{pmatrix} -\gamma_2 \\ \gamma_1 \end{pmatrix}.$$

Die Lösungskurven $\gamma(t)$ sind überall tangential an v und heißen oft auch **Integrialkurven** an das Vektorfeld v . Beim Lösen einer Differentialgleichung spricht man manchmal vom „Integrieren der Gleichung“.



Für dieses Beispiel findet man leicht explizit, dass $\gamma(t) := (r \cos(t + \varphi), r \sin(t + \varphi))$ eine Lösung zum Anfangswert $x_0 = (r \cos \varphi, r \sin \varphi)$ ist. \diamond

8.4 Definition. Autonome Differentialgleichung m -ter Ordnung

Sei $m \in \mathbb{N}$. Eine **autonome gewöhnliche Differentialgleichung m -ter Ordnung** auf einem Gebiet $D \subset \mathbb{R}^n$ ist gegeben durch eine stetige Funktion

$$f : D \times \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{(m-1)\text{-mal}} \rightarrow \mathbb{R}^n$$

und die Gleichung

$$\gamma^{(m)} = f\left(\gamma, \gamma', \dots, \gamma^{(m-1)}\right). \tag{8.4}$$

Unter einer Lösung von (8.4) versteht man eine m -mal stetig differenzierbare Kurve $\gamma : I \rightarrow D$, $I \subset \mathbb{R}$ ein offenes Intervall, so dass für alle $t \in I$ gilt

$$\gamma^{(m)}(t) = f\left(\gamma(t), \gamma'(t), \dots, \gamma^{(m-1)}(t)\right).$$

Falls $0 \in I$ ist, so heißt $\gamma : I \rightarrow D$ Lösung zum Anfangswert $(x_0, y_1, \dots, y_{m-1})$ falls $\gamma(0) = x_0$ und $\gamma^{(j)}(0) = y_j$ für $j = 1, \dots, m - 1$. \diamond

8.5 Bemerkung. Reduktion auf ein System erster Ordnung

Man kann eine autonome DGL m -ter Ordnung auf einem Gebiet $D \subset \mathbb{R}^n$,

$$\gamma^{(m)} = f(\gamma, \gamma', \dots, \gamma^{(m-1)}), \tag{8.5}$$

immer auf eine autonome DGL erster Ordnung auf dem Gebiet $G := D \times \mathbb{R}^{n(m-1)} \subset \mathbb{R}^{nm}$ zurückführen. Man definiert dazu das Vektorfeld $v : G \rightarrow \mathbb{R}^{nm}$ durch

$$v(x, y_1, \dots, y_{m-1}) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ f(x, y_1, \dots, y_{m-1}) \end{pmatrix}$$

und betrachtet für $\alpha : I \rightarrow G$ die Differentialgleichung erster Ordnung

$$\alpha' = v(\alpha). \quad (8.6)$$

Ist nun $\gamma : I \rightarrow D$, $t \mapsto \gamma(t)$ eine Lösung von (8.5), so ist

$$\alpha : I \rightarrow G, \quad t \mapsto \alpha(t) := (\gamma(t), \gamma'(t), \dots, \gamma^{(m-1)}(t)),$$

Lösung von (8.6).

Ist umgekehrt $\alpha : I \rightarrow G$, $\alpha(t) = (\alpha_1(t), \dots, \alpha_m(t))$ mit $\alpha_j(t) \in \mathbb{R}^n$, Lösung von (8.6), so ist

$$\gamma : I \rightarrow D, \quad t \mapsto \gamma(t) := \alpha_1(t),$$

Lösung von (8.5). ◇

8.6 Definition. Autonome und nicht-autonome Systeme

Sei $J \subset \mathbb{R}$ ein offenes Intervall. Ein stetiges **zeitabhängiges Vektorfeld** v auf einem Gebiet $D \subset \mathbb{R}^n$ ist eine stetige Abbildung

$$v : J \times D \rightarrow \mathbb{R}^n, \quad (t, x) \mapsto v(t, x).$$

Man nennt dann die gewöhnliche Differentialgleichung

$$\gamma' = v(t, \gamma) \quad (8.7)$$

nicht-autonom, wenn v explizit von $t \in J$ abhängt. Ist v (wie bisher) unabhängig von t , so heißt $\gamma' = v(\gamma)$ **autonom**. Sei $I \subset J$ ein Teilintervall. Eine stetig differenzierbare Kurve $\gamma : I \rightarrow D$ heißt **Lösung von** (8.7), wenn für alle $t \in I$ gilt

$$\gamma'(t) = v(t, \gamma(t)).$$

Für $t_0 \in I$ heißt γ Lösung zum Anfangswert $x_0 = \gamma(t_0)$ zur Anfangszeit t_0 . ◇

8.7 Bemerkung. Reduktion auf ein autonomes System

Man kann jedes nicht-autonome System $\gamma' = f(t, \gamma)$ auf einem Gebiet $D \subset \mathbb{R}^n$ auf ein autonomes System auf $G := J \times D \subset \mathbb{R}^{n+1}$ wie folgt zurückführen:

Sei

$$v : G \rightarrow \mathbb{R}^{1+n}, \quad v(s, y) = (1, f(s, y)),$$

und sei

$$\alpha : I \rightarrow G, \quad t \mapsto \alpha(t) := (\alpha_0(t), \alpha_1(t))$$

eine Lösung von $\alpha' = v(\alpha)$ zum Anfangswert $\alpha(0) = (t_0, x_0)$. Dann ist

$$\gamma : \{t \in \mathbb{R} \mid t - t_0 \in I\} \rightarrow \mathbb{R}^n, \quad t \mapsto \gamma(t) := \alpha_1(t - t_0)$$

eine Lösung von $\gamma' = f(t, \gamma)$ zum Anfangswert x_0 zur Anfangszeit t_0 : Es ist $\gamma(t_0) = \alpha_1(0) = x_0$ und es gilt

$$\gamma'(t) = \alpha_1'(t - t_0) = f(\underbrace{\alpha_0(t - t_0)}_{=t}, \alpha_1(t - t_0)) = f(t, \gamma(t)).$$

Hier haben wir verwendet, dass α_0 die Differentialgleichung

$$\alpha_0' = 1$$

zum Anfangswert $\alpha_0(0) = t_0$ löst, welche die eindeutige Lösung $\alpha_0(t) = t_0 + t$ hat. ◇

8.8 Bemerkung. (a) Im allgemeinen hat man keine Chance, ein System gewöhnlicher Differentialgleichungen

$$\gamma' = v(\gamma)$$

auf ein Gebiet $G \subset \mathbb{R}^n$ bei Vorgabe von $v : G \rightarrow \mathbb{R}^n$ zu integrieren, d.h. Lösungen explizit zu bestimmen. Man ist daher schon damit zufrieden, qualitative Aussagen über die Bahn $t \mapsto \gamma(t)$ zu bekommen, wie z.B. Antworten auf die Fragen:

- (i) Existiert zu gegebenem Vektorfeld v und Anfangswert $x_0 \in G$ eine Lösung und ist diese eindeutig?
- (ii) Konvergiert $t \mapsto \gamma(t)$ für $t \rightarrow \infty$ gegen eine Gleichgewichtslage $x_* \in G$, also einen Punkt mit $v(x_*) = 0$?
- (iii) Ist $t \mapsto \gamma(t)$ periodisch, also $\gamma(t + T) = \gamma(t)$ für ein $T > 0$ und alle $t \in \mathbb{R}$?

Man nennt das Studium dieser Fragen die „Qualitative Theorie gewöhnlicher Differentialgleichungen“.

- (b) Hat man spezielle Informationen über v oder G , z.B. Symmetrieaussagen oder $G \subset \mathbb{R}$ so kann man $\gamma' = v(\gamma)$ in manchen Fällen doch explizit integrieren.

8.9 Proposition. Integration eindimensionaler Differentialgleichungen

Seien $G \subset \mathbb{R}$ und $J \subset \mathbb{R}$ offene Intervalle und seien $a : J \rightarrow \mathbb{R}$ und $b : G \rightarrow \mathbb{R}$ stetige Funktionen. Dann ist

$$v : J \times G \rightarrow \mathbb{R}, \quad v(t, x) = a(t)b(x)$$

ein stetiges zeitabhängiges Vektorfeld. Sei $(t_0, x_0) \in J \times G$ mit $v(t_0, x_0) \neq 0$. Dann existieren offene Umgebungen $U_1 \subset J$ von t_0 und $U_2 \subset G$ von x_0 so, dass die implizite Gleichung

$$F(t, x) := \int_{t_0}^t a(s) ds - \int_{x_0}^x \frac{1}{b(u)} du \stackrel{!}{=} 0$$

eine eindeutige stetig differenzierbare Lösung $\gamma : U_1 \rightarrow U_2$ hat, also $F(t, \gamma(t)) = 0$ gilt. Es ist γ eine Lösungskurve von $\gamma' = v(t, \gamma)$ zum Anfangswert x_0 zur Anfangszeit t_0 .

Beweis. Es ist F auf hinreichend kleinen Umgebungen von t_0 und x_0 wohldefiniert und stetig differenzierbar. Da $F(t_0, x_0) = 0$ und $\partial_x F(t, x)|_{(t_0, x_0)} = \frac{1}{b(x_0)} \neq 0$, gibt es nach dem Satz über implizite Funktionen Umgebungen U_1 und U_2 und eine stetige differenzierbare Funktion $\gamma : U_1 \rightarrow U_2$ mit $\gamma(t_0) = x_0$, $F(t, \gamma(t)) = 0$ und

$$\gamma'(t) = -\partial_x F(t, \gamma(t))^{-1} \partial_t F(t, \gamma(t)) = b(\gamma(t)) a(t) = v(t, \gamma(t)). \quad \square$$

8.10 Bemerkung. Diese Aussage zeigt die Berechtigung der heuristischen Rechnung

$$\gamma' = v(\gamma) \Rightarrow \frac{d\gamma}{dt} = a(t)b(\gamma) \stackrel{(*)}{\Rightarrow} \frac{1}{b(\gamma)} d\gamma = a(t) dt \Rightarrow \int_{x_0}^{\gamma} \frac{1}{b(x)} dx = \int_{t_0}^t a(s) ds.$$

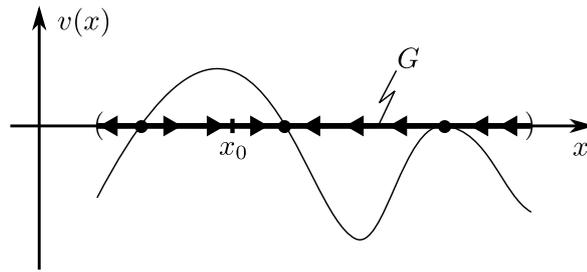
Den Schritt (*) nennt man oft **Trennung der Variablen**. ◇

8.11 Bemerkung. (a) Im Fall einer autonomen DGL erster Ordnung auf $G \subset \mathbb{R}$ und $t_0 = 0$ erhält man gemäß Proposition 8.9 eine Lösung zum Anfangswert $x_0 \in G$, indem man

$$t = \int_{x_0}^{\gamma} \frac{1}{v(u)} du$$

nach γ auflöst. Da die rechte Seite eine monotone Funktion von γ ist, hat das Phasendiagramm einer solchen DGL folgende Form:

In den Nullstellen von v ruht das System. Dazwischen läuft es monoton von einer Nullstelle zur nächsten. Ist v an den Nullstellen differenzierbar, so werden die Gleichgewichtslagen in endlicher Zeit nicht erreicht. An den Rändern des Intervalls G können die Lösungen in endlicher Zeit entweichen.



- (b) Ganz allgemein führen die Punkte $x_0 \in G$ an denen ein gegebenes Vektorfeld $v : G \rightarrow \mathbb{R}^n$ verschwindet, also $v(x_0) = 0$ gilt, zu besonders einfachen Lösungen der Differentialgleichung $\gamma' = v(\gamma)$: Es ist dann $\gamma : \mathbb{R} \rightarrow G$, $t \mapsto \gamma(t) \equiv x_0$, eine globale Lösung von $\gamma' = v(\gamma)$. Man nennt diese Punkte daher Gleichgewichtslagen oder stationäre Punkte.

8.12 Beispiele. (a) Sei $c \in \mathbb{R}$ und $v : \mathbb{R} \rightarrow \mathbb{R}$, $v(x) = cx$ ein lineares Vektorfeld. Die Lösung von $\gamma' = v(\gamma)$ auf $G = \mathbb{R}$ zum Anfangswert $x_0 > 0$ für $c \neq 0$ ist nach Proposition 8.8 für $x > 0$ gegeben durch Auflösen von

$$t = \int_0^t ds = \int_{x_0}^{\gamma} \frac{1}{cu} du = \frac{1}{c} (\ln \gamma - \ln x_0) = \frac{1}{c} \ln \frac{\gamma}{x_0}$$

nach γ , also

$$\ln \frac{\gamma}{x_0} = ct \Rightarrow \gamma(t) = x_0 e^{ct} \quad \forall t \in \mathbb{R}.$$

Diese Formel liefert auch für $x_0 \leq 0$ eine Lösung und somit ist zu jedem Anfangswert $x_0 \in \mathbb{R}$ die Funktion $\gamma : \mathbb{R} \rightarrow \mathbb{R}$, $t \mapsto \gamma(t) = e^{ct} x_0$ eine globale Lösung.

- (b) Sei $v : \mathbb{R} \rightarrow \mathbb{R}$, $v(x) = 1 + x^2$ quadratisch in x . Die Lösungskurve $t \rightarrow \gamma(t)$ zum Anfangswert $x_0 = 0$ erhält man durch Auflösen von

$$t = \int_0^t ds = \int_0^{\gamma} \frac{1}{1+u^2} du = \arctan(\gamma)$$

nach γ , also

$$\gamma(t) = \tan(t).$$

Somit ist das Existenzintervall der Lösung $I(0) = (-\frac{\pi}{2}, \frac{\pi}{2})$ und $t \mapsto \gamma(t)$ läuft „in endlicher Zeit nach Unendlich“.

- (c) Sei $v : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $v(x) = \sqrt{|x|}$. Dann erreichen die Lösungen zu Startwerten $x(0) < 0$ in endlicher Zeit die Ruhelage $x = 0$. Dort können Sie beliebig lange verweilen und dann nach rechts wieder aus der Ruhelage herauslaufen. Die Lösungen sind also nicht eindeutig. (vgl. Übungen).

Wir erinnern zunächst an die allgemeine Definition von Lipschitz-Stetigkeit.

8.13 Definition. Lipschitz-Stetigkeit

Seien X, Z metrische Räume. Eine Abbildung $f : X \rightarrow Z$ heißt **Lipschitz-stetig** mit **Lipschitz-Konstante** $L \in [0, \infty)$, falls für alle $x, y \in X$ gilt, dass

$$d_Z(f(x), f(y)) \leq L d_X(x, y).$$

Für zeitabhängige Vektorfelder werden wir im Folgenden nur eine lokale Lipschitz-Bedingung in der zweiten Variablen fordern.

8.14 Definition. Lipschitz-Bedingung an Vektorfelder

Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ und $v : U \rightarrow \mathbb{R}^n$ stetig.

- (a) Wir sagen, dass v einer **Lipschitz-Bedingung** genügt, wenn es ein $L \geq 0$ gibt, so dass für alle $(t, x), (t, y) \in U$ gilt

$$\|v(t, x) - v(t, y)\| \leq L\|x - y\|.$$

Es heißt dann L eine **Lipschitz-Konstante** für v .

- (b) Wir sagen, dass v einer **lokalen Lipschitz-Bedingung** genügt, wenn jedes $(t, x) \in U$ eine Umgebung $V \subset U$ besitzt, so dass $v|_V$ einer Lipschitz-Bedingung genügt.

8.15 Proposition. Differenzierbar \Rightarrow lokal Lipschitz

Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ offen und $v : U \rightarrow \mathbb{R}^n$, $(t, x) \mapsto v(t, x)$ stetig. Falls alle partiellen Ableitungen von $\partial_{x_j} v$ bezüglich der Variablen $x = (x_1, \dots, x_n)$ existieren und auf U stetige Funktionen definieren, so genügt v einer lokalen Lipschitz-Bedingung.

Beweis. Übungsaufgabe (Satz 3.17 und Schrankensatz). □

8.16 Proposition. Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ offen und $v : U \rightarrow \mathbb{R}^n$ genüge einer lokalen Lipschitz-Bedingung. Für jedes Kompaktum $K \subset U$ genügt $v|_K$ dann einer Lipschitz-Bedingung.

Beweis. Übungsaufgabe. □

Wir formulieren nun den für die Theorie gewöhnlicher Differentialgleichungen zentralen Existenz- und Eindeutigkeitsatz von Picard und Lindelöf.

8.17 Theorem. Existenz- und Eindeutigkeitsatz von Picard-Lindelöf

Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ ein Gebiet und $v : U \rightarrow \mathbb{R}^n$ sei stetig und genüge einer lokalen Lipschitz-Bedingung.

Lokale Existenz: Zu jedem Anfangswert $(t_0, x_0) \in U$ existieren ein $\delta > 0$ und eine stetig differenzierbare Kurve $\gamma : (t_0 - \delta, t_0 + \delta) \rightarrow \mathbb{R}^n$, die Lösung von $\gamma' = v(t, \gamma)$ zum Anfangswert (t_0, x_0) ist, also

$$\gamma'(t) = v(t, \gamma(t)) \quad \text{für alle } t \in (t_0 - \delta, t_0 + \delta) \quad \text{und} \quad \gamma(t_0) = x_0 \quad (8.8)$$

erfüllt.

Eindeutigkeit: Ist $J \subset \mathbb{R}$ ein Intervall mit $t_0 \in J$ und $\tilde{\gamma} : J \rightarrow \mathbb{R}^n$ ebenfalls Lösung von (8.8), so gilt

$$\tilde{\gamma}(t) = \gamma(t) \quad \text{für alle } t \in (t_0 - \delta, t_0 + \delta) \cap J.$$

Beweis. Die Grundidee des Beweises ist die Folgende: Erfüllt $\gamma : (t_0 - \delta, t_0 + \delta) \rightarrow \mathbb{R}^n$ die Gleichungen (8.8), so folgt durch Integration und den Hauptsatz

$$\gamma(t) - \underbrace{\gamma(t_0)}_{=x_0} = \int_{t_0}^t \gamma'(s) \, ds = \int_{t_0}^t v(s, \gamma(s)) \, ds,$$

also

$$\gamma(t) = x_0 + \int_{t_0}^t v(s, \gamma(s)) \, ds. \quad (8.9)$$

Jede Lösung von (8.8) ist also Fixpunkt der Abbildung Φ , die jeder Funktion φ die Funktion $\Phi[\varphi]$ zuordnet, gegeben durch

$$\Phi[\varphi](t) = x_0 + \int_{t_0}^t v(s, \varphi(s)) \, ds.$$

Ist umgekehrt γ Fixpunkt von Φ und stetig, so liefert Differentiation von (8.9), dass γ (8.8) erfüllt. Man muss nun lediglich den Funktionenraum auf dem Φ operiert so definieren, dass er ein Banachraum stetiger Funktionen ist und Φ eine Kontraktion darauf. Dann liefert der Banachsche Fixpunktsatz das Resultat.

Nun zu den Details. Da U offen ist, existieren $\delta_0, r > 0$ so, dass $K_1 \times K_2 \subset U$ für $K_1 := [t_0 - \delta_0, t_0 + \delta_0]$ und $K_2 := \overline{B}_r(x_0)$. Da $K_1 \times K_2$ kompakt und $v|_{K_1 \times K_2} : K_1 \times K_2 \rightarrow \mathbb{R}^n$ stetig ist, existiert ein $M \in \mathbb{R}$ so, dass für alle $(t, x) \in K_1 \times K_2$ gilt $\|v(t, x)\| \leq M$. Nach Proposition 8.16 genügt $v|_{K_1 \times K_2}$ einer Lipschitz-Bedingung mit Lipschitz-Konstante L .

Setze nun $\delta := \min\{\frac{1}{L}, \frac{r}{M}, \delta_0\}$ und sei $0 < \tilde{\delta} < \delta$ beliebig und $I := [t_0 - \tilde{\delta}, t_0 + \tilde{\delta}]$. Wir betrachten den Vektorraum

$$X = \{\varphi : I \rightarrow \mathbb{R}^n \mid \varphi \text{ ist stetig}\} = C(I, \mathbb{R}^n)$$

mit der Supremumsnorm

$$\|\varphi\|_\infty = \sup_{t \in I} \|\varphi(t)\|.$$

Es ist dann $(X, \|\cdot\|_\infty)$ ein Banachraum und die Teilmenge

$$\mathcal{A} = \{\varphi \in X \mid \varphi(t_0) = x_0 \text{ und } \varphi(I) \subset K_2\} \subset X$$

ist abgeschlossen. Für $\varphi \in \mathcal{A}$ sei nun $\Phi[\varphi] : I \rightarrow \mathbb{R}^n$ gegeben durch

$$\Phi[\varphi](t) := x_0 + \int_{t_0}^t v(s, \varphi(s)) \, ds.$$

Es ist dann $\Phi[\varphi]$ wieder in \mathcal{A} , denn $\Phi[\varphi]$ ist stetig (sogar stetig differenzierbar), $\Phi[\varphi](t_0) = x_0$ und $\Phi[\varphi](t) \in K_2$ für alle $t \in I$. Letzteres folgt aus

$$\|\Phi[\varphi](t) - x_0\| = \left\| \int_{t_0}^t v(s, \varphi(s)) \, ds \right\| \leq \int_{t_0}^t \underbrace{\|v(s, \varphi(s))\|}_{\substack{\in K_1 \times K_2 \\ \leq M}} \, ds \leq \tilde{\delta} M < \frac{r}{M} M = r.$$

Damit ist Φ eine Abbildung von \mathcal{A} nach \mathcal{A} , $\Phi : \mathcal{A} \rightarrow \mathcal{A}$. Schließlich ist Φ eine Kontraktion mit Kontraktionskonstante $0 < \theta < 1$, $\theta := \tilde{\delta} L$. Denn für alle $\varphi, \psi \in \mathcal{A}$ gilt

$$\begin{aligned} \|\Phi[\varphi] - \Phi[\psi]\|_\infty &= \sup_{t \in I} \left\| \int_{t_0}^t (v(s, \varphi(s)) - v(s, \psi(s))) \, ds \right\| \leq \sup_{t \in I} \int_{t_0}^t \|v(s, \varphi(s)) - v(s, \psi(s))\| \, ds \\ &\leq \sup_{t \in I} \int_{t_0}^t L \|\varphi(s) - \psi(s)\| \, ds \leq L \tilde{\delta} \|\varphi - \psi\|_\infty = \theta \|\varphi - \psi\|_\infty \end{aligned}$$

mit $\theta = \tilde{\delta} L < 1$, weil $\tilde{\delta} < \delta \leq \frac{1}{L}$ ist.

Nach dem Banachschen Fixpunktsatz existiert also genau ein $\varphi \in \mathcal{A}$ mit $\Phi[\varphi] = \varphi$.

Setze nun $\gamma : (t_0 - \delta, t_0 + \delta) \rightarrow \mathbb{R}^n$, $\gamma(t) = \varphi(t)$, wobei für $t \in (t_0 - \delta, t_0 + \delta)$ erst ein $\tilde{\delta}$ mit $|t - t_0| < \tilde{\delta} < \delta$ gewählt sei und φ dann der eindeutige Fixpunkt von $\Phi = \Phi_{\tilde{\delta}}$ sei. Da jeder Fixpunkt von $\Phi_{\tilde{\delta}}$ auch Fixpunkt von Φ_{δ_1} mit $\delta_1 \leq \tilde{\delta}$ ist, hängt diese Definition nicht von der Wahl von $\tilde{\delta}$ ab. Es folgt, dass γ stetig ist und für alle $t \in (t_0 - \delta, t_0 + \delta)$ gilt:

$$\gamma(t) = x_0 + \int_{t_0}^t v(s, \gamma(s)) \, ds.$$

Also ist γ sogar stetig differenzierbar mit

$$\gamma'(t) = v(t, \gamma(t)) \quad \text{und} \quad \gamma(t_0) = x_0,$$

8 Gewöhnliche Differentialgleichungen

d.h. ist γ Lösung von (8.8).

Die Eindeutigkeitsaussage folgt nun leicht: Es löst $\tilde{\gamma}$ die Gleichung (8.8) also auch (8.9). Daher lösen sowohl $\gamma|_{(t_0-\delta, t_0+\delta) \cap J}$ als auch $\tilde{\gamma}|_{(t_0-\delta, t_0+\delta) \cap J}$ Gleichung (8.9). Nun ist aber Φ auch für jedes kompakte Intervall $I \subset (t_0 - \delta, t_0 + \delta) \cap J$ mit $t_0 \in I$ auf

$$\tilde{A} := \{\varphi \in C(I, \mathbb{R}^n) \mid \varphi(t_0) = x_0 \text{ und } \varphi(I) \subset K_2\}$$

eine Kontraktion und somit ihr Fixpunkt eindeutig, also $\gamma|_{(t_0-\delta, t_0+\delta) \cap J} = \tilde{\gamma}|_{(t_0-\delta, t_0+\delta) \cap J}$. \square

8.18 Bemerkung. (a) Man beachte, dass der Existenzsatz nur die Existenz einer Lösung von $\gamma' = v(s, \gamma)$, $\gamma(t_0) = x_0$ „für kurze Zeiten“ liefert, $\gamma : (t_0 - \delta, t_0 + \delta) \rightarrow G$. Man spricht daher von lokaler- bzw. Kurzzeitexistenz.

(b) Der Beweis gibt auch eine untere Schranke für die „Lebensdauer“ der Lösung γ , nämlich

$$\delta(t_0, x_0) = \min \left\{ \frac{1}{L(t_0, x_0)}, \frac{r(t_0, x_0)}{M(t_0, x_0)}, \delta_0(t_0, x_0) \right\},$$

wobei $\delta_0(t_0, x_0)$ und $r(t_0, x_0)$ so gewählt sind, dass $K := [t_0 - \delta_0, t_0 + \delta_0] \times \overline{B_r}(x_0) \subset U$ ist und $M(t_0, x_0) > 0$ eine Schranke für die „Geschwindigkeit“ $\|v\|$ auf K und $L(t_0, x_0) > 0$ eine Lipschitzkonstante (z.B. eine Schranke an $\|D_x v\|$, falls v stetig differenzierbar bzgl. x ist) auf K ist.

(c) Für autonome Systeme kann man den Zeitnullpunkt einer Lösung beliebig verschieben: Ist $\gamma : I \rightarrow \mathbb{R}^n$ eine Lösung von $\gamma' = v(\gamma)$ zum Anfangswert $x_0 = \gamma(0)$ und $\tau \in I$, so ist $\tilde{\gamma} : I - \tau \rightarrow \mathbb{R}^n$, $\tilde{\gamma}(t) := \gamma(t + \tau)$ Lösung von $\gamma' = v(\gamma)$ zum Anfangswert $\tilde{x}_0 = \tilde{\gamma}(0) = \gamma(\tau)$, wobei $I - \tau := \{t \in \mathbb{R} \mid t + \tau \in I\}$. Das folgt sofort aus

$$\frac{d}{dt} \tilde{\gamma}(t) = \gamma'(t + \tau) = v(\gamma(t + \tau)) = v(\tilde{\gamma}(t)).$$

Der einfacheren Notation halber betrachten wir im Folgenden nur den Fall $U = J \times G$, wobei $J \subset \mathbb{R}$ ein offenes Intervall und $G \subset \mathbb{R}^n$ ein Gebiet ist.

8.19 Definition. Maximale Lösung

Sei $v : J \times G \rightarrow \mathbb{R}^n$ stetig und genüge einer lokalen Lipschitz-Bedingung. Eine Lösung $\gamma : I \rightarrow G$, $I \subset J$ ein offenes Intervall, von $\gamma' = v(t, \gamma)$ heißt **maximal**, wenn gilt: Ist $\tilde{I} \subset J$ ein Intervall mit $\tilde{I} \supset I$ und $\tilde{\gamma} : \tilde{I} \rightarrow G$ eine Lösung von $\gamma' = v(t, \gamma)$ mit $\tilde{\gamma}|_I = \gamma$, so gilt bereits $\tilde{I} = I$ und somit $\tilde{\gamma} = \gamma$.

Man sagt auch, eine maximale Lösung kann nicht fortgesetzt werden.

8.20 Proposition. Sei $J \subset \mathbb{R}$ ein offenes Intervall, $G \subset \mathbb{R}^n$ ein Gebiet und $v : J \times G \rightarrow \mathbb{R}^n$ sei stetig und genüge einer lokalen Lipschitz-Bedingung. Seien $I, \tilde{I} \subset J$ offene Intervalle mit $t_0 \in I \cap \tilde{I}$ und $\gamma : I \rightarrow G$ und $\tilde{\gamma} : \tilde{I} \rightarrow G$ Lösungen der Differentialgleichung $\gamma' = v(t, \gamma)$ mit Anfangswert $\gamma(t_0) = \tilde{\gamma}(t_0) = x_0 \in G$. Dann gilt für alle $t \in I \cap \tilde{I}$, dass $\gamma(t) = \tilde{\gamma}(t)$.

Beweis. Sei $A = \{t \in I \cap \tilde{I} \mid \gamma(t) = \tilde{\gamma}(t)\}$. Wir zeigen, dass $A = I \cap \tilde{I}$ ist. Sei dazu $I = (t_-, t_+)$ und $\tilde{I} = (s_-, s_+)$ und o.B.d.A. $t_+ \leq s_+$ und

$$t_* := \sup\{\tau \in (t_0, t_+) \mid \gamma(t) = \tilde{\gamma}(t) \forall t_0 \leq t < \tau\}.$$

Nach dem Eindeutigkeitsatz existiert ein $\delta_{x_0, t_0} > 0$, sodass $\gamma(t) = \tilde{\gamma}(t)$ für alle $t_0 \leq t < t_0 + \delta_{x_0, t_0}$. Also ist $t_0 + \delta_{x_0, t_0} \leq t_* \leq t_+$. Angenommen $t_* < t_+$. Die Stetigkeit von γ und $\tilde{\gamma}$ liefert dann $\gamma(t_*) = \tilde{\gamma}(t_*) =: x_*$. Wiederum nach Picard-Lindelöf existiert dann aber ein $\delta_{x_*, t_*} > 0$ so, dass $\gamma(t) = \tilde{\gamma}(t)$ für alle $t \in (t_* - \delta_{x_*, t_*}, t_* + \delta_{x_*, t_*})$, was im Widerspruch zur Definition vor t_* steht. Also ist $t_* = t_+$ und mit einem analogen Argument für die andere Intervallgrenze ergibt sich $A = I \cap \tilde{I}$. \square

8.21 Satz. Existenz und Eindeutigkeit der maximalen Lösung

Sei $J \subset \mathbb{R}$ ein offenes Intervall, $G \subset \mathbb{R}^n$ ein Gebiet und $v : J \times G \rightarrow \mathbb{R}^n$ sei stetig und genüge einer lokalen Lipschitz-Bedingung. Dann existiert zu jedem $(t_0, x_0) \in J \times G$ genau eine maximale Lösung $\gamma : I \rightarrow G$ von $\gamma' = v(t, \gamma)$ mit $\gamma(t_0) = x_0$.

Beweis. Ist $\tilde{\gamma} : \tilde{I} \rightarrow G$ eine Lösung von $\gamma' = v(t, \gamma)$ mit $\tilde{\gamma}(t_0) = x_0$ und $t_0 \in \tilde{I}$, so notieren wir diese mit $(\tilde{\gamma}, (t_0, x_0), \tilde{I})$. Sei nun

$$I := \bigcup_{(\tilde{\gamma}, (t_0, x_0), \tilde{I}) \text{ ist Lösung}} \tilde{I} \subset J$$

Dann ist $I \subset J$ ein offenes Intervall. Ist $t \in I$ beliebig, so wählen wir eine Lösung $(\tilde{\gamma}, (t_0, x_0), \tilde{I})$ mit $t \in \tilde{I}$ und setzen $\gamma(t) := \tilde{\gamma}(t)$. Wegen Proposition 8.20 ist dies unabhängig von der Wahl $(\tilde{\gamma}, (t_0, x_0), \tilde{I})$. Es ist dann die so definierte Funktion $\gamma : I \rightarrow G$ eine Lösung von $\gamma' = v(t, \gamma)$ mit $\gamma(t_0) = x_0$. Nach Konstruktion ist γ maximal und zwar die einzige maximale Lösung. \square

8.22 Bemerkung. (a) Wir schreiben für das Definitionsintervall I der maximalen Lösung von $\gamma' = v(t, \gamma)$, $\gamma(t_0) = x_0$,

$$I(t_0, x_0) = (t_-(t_0, x_0), t_+(t_0, x_0)),$$

wobei $t_-(t_0, x_0) \in \{-\infty\} \cup (-\infty, t_0)$ bzw. $t_+(t_0, x_0) \in (t_0, \infty) \cup \{\infty\}$ die linke bzw. rechte Intervallgrenze bezeichnet. Man sagt, die Lösung existiert **global**, wenn $I(t_0, x_0) = J$ ist, wobei J das Zeitintervall ist, auf dem das Vektorfeld definiert ist.

(b) Zu $v : J \times G \rightarrow \mathbb{R}^n$ und fest gewählten $t_0 \in J$ sei

$$\Omega := \{(t, x) \in J \times G \mid t \in I(t_0, x)\}$$

und $\varphi : \Omega \rightarrow G$, $(t, x) \mapsto \varphi(t, x) := \gamma_x(t)$, wobei $I(t_0, x) \rightarrow G$, $t \mapsto \gamma_x(t)$ die maximale Lösung von $\gamma' = v(t, \gamma)$ zum Anfangswert $(t_0, x) \in J \times G$ ist.

Man nennt die Abbildung $\varphi : \Omega \rightarrow G$ den **Fluss** oder die **Flussabbildung** zum Vektorfeld v . Für festes t bildet

$$\varphi_t : \Omega_t \rightarrow G, \quad x \mapsto \varphi_t(x) := \varphi(t, x)$$

also die Menge aller Anfangsdaten $\Omega_t := \{x \in G \mid (t, x) \in \Omega\}$ deren maximales Existenzintervall t enthält auf die jeweilige Lösung zum Zeitpunkt t ab. Die Flussabbildung erfüllt die Gleichung

$$\varphi_t \circ \varphi_s = \varphi_{t+s}$$

auf der Teilmenge von Ω_{t+s} auf der auch die linke Seite definiert ist. Für einen globalen Fluss mit $J = \mathbb{R}$, also wenn $\Omega = \mathbb{R} \times G$ gilt, ist $t \mapsto \varphi_t$ also eine Gruppenwirkung der Gruppe $(\mathbb{R}, +)$ auf der Menge G .

(c) Man kann nun zeigen, dass Ω offen ist und $\varphi : \Omega \rightarrow G$ stetig ist. Man spricht von „stetiger Abhängigkeit von den Anfangsdaten“. Falls $v : J \times G \rightarrow \mathbb{R}^n$ stetig differenzierbar ist, so ist auch $\varphi : \Omega \rightarrow G$ stetig differenzierbar und man spricht von „differenzierbarer Abhängigkeit von den Anfangsdaten“. Der Beweis der letzten Aussage ist allerdings etwas umfangreicher.

Wie schon gesagt, ist es im Allgemeinen nicht möglich, die Lösungen einer Differentialgleichung explizit zu bestimmen. Meistens weiß man nicht einmal, ob bei gegebenem $v : J \times G \rightarrow \mathbb{R}^n$ und $(t_0, x_0) \in J \times G$ das Ende $t_+(t_0, x_0) \in (t_0, \infty] \cap J$ des Existenzintervalls kleiner oder gleich der oberen Grenze j_+ von $J =: (j_-, j_+)$ ist. Der folgende Satz besagt aber immerhin, dass die Bahnkurve $t \mapsto \gamma(t)$ eines Anfangspunktes $(t_0, x_0) \in U$ mit $t_+(t_0, x_0) < j_+$ jedes Kompaktum K in G verlassen muss, wenn $t \rightarrow t_+(t_0, x_0)$ geht, d.h. $\gamma(t)$ strebt für $t \rightarrow t_+(t_0, x_0)$ zum Rand von G oder nach Unendlich. Anders gesagt, $t \mapsto \gamma(t)$ kann sich in endlicher Zeit „nicht einfach in Luft auflösen“.

8.23 Satz. Verhalten für $t \rightarrow t_+(t_0, x_0)$

Sei $J = (j_-, j_+) \subset \mathbb{R}$ ein offenes Intervall, $G \subset \mathbb{R}^n$ ein Gebiet und $v : J \times G \rightarrow \mathbb{R}^n$ erfülle eine lokale Lipschitz-Bedingung. Sei $\gamma : (t_-(t_0, x_0), t_+(t_0, x_0)) \rightarrow G$ die maximale Lösung von $\gamma' = v(t, \gamma)$ zum Anfangswert $(t_0, x_0) \in J \times G$.

Falls $t_+(t_0, x_0) < j_+$, so gibt es zu jedem Kompaktum $K \subset G$ ein $0 < \tau_K < t_+(t_0, x_0)$, sodass für alle $t \in (\tau_K, t_+(t_0, x_0))$ gilt, dass $\gamma(t) \notin K$.

Beweis. Da K kompakt ist, gibt es ein $\varrho > 0$, so dass $B_\varrho(x) \subset G$ für alle $x \in K$ gilt. Denn die Distanzfunktion zum Rand von G , $\text{dist} : \mathbb{R}^n \rightarrow [0, \infty)$,

$$\text{dist}(x, \partial G) := \inf \{ \|x - y\| \mid y \in \partial G \}$$

ist stetig und nimmt auf K ihr Minimum ϱ an. Es ist $\varrho > 0$, denn sonst wäre $K \cap \partial G \neq \emptyset$, also $K \not\subset G$. Seien weiter $\|v(t, x)\| \leq M$ für $(t, x) \in \tilde{K} := [t_0, t_+(t_0, x_0)] \times \overline{B}_{\varrho/2}(K)$ und L eine Lipschitzkonstante für v auf \tilde{K} . Dann gilt für $t \in (t_0, t_+(t_0, x_0))$ mit $\gamma(t) \in K$, dass

$$t_+(t_0, x_0) = t_+(t, \gamma(t)) + t \geq \delta + t$$

mit $\delta := \min \{ \frac{1}{L}, \frac{\varrho}{2M}, j_+ - t \} > 0$. Für alle $t \in (\tau, t_+(t_0, x_0))$ mit $\tau := t_+(t_0, x_0) - \delta$ muss also $\gamma(t) \notin K$ gelten. \square

- 8.24 Bemerkung.** (a) Eine entsprechende Aussage gilt natürlich, wenn $t_-(t_0, x_0) > -\infty$ ist.
 (b) Bleibt eine Lösungskurve $t \mapsto \gamma(t)$ in einem Kompaktum, z.B. wenn $\lim_{t \rightarrow t_+} \gamma(t) = p$ für ein $p \in G$ ist, so muss also $t_+(t_0, x_0) = j_+$ sein.

9 Lineare Differentialgleichungen

9.1 Definition. Lineare Systeme

Sei $J \subset \mathbb{R}$ ein offenes Intervall und $A : J \rightarrow M(n, \mathbb{R})$ eine stetige Abbildung mit Werten in den reellen $n \times n$ -Matrizen.

(a) Man nennt dann die Differentialgleichung

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

ein **nicht-autonomes, homogenes, lineares System**, da das Vektorfeld $v(t, x) := A(t)x$ jetzt linear von x abhängt.

(b) Ist $b : J \rightarrow \mathbb{R}^n$ stetig, so heißt

$$\gamma' = A(t)\gamma + b(t)$$

ein **nicht-autonomes, inhomogenes, lineares System**.

9.2 Beispiel. Im autonom homogenen Fall

$$\gamma' = A\gamma$$

ist das Vektorfeld $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto v(x) = Ax$, Lipschitz-stetig mit Lipschitzkonstante $\|A\|$. Die eindeutige globale Lösung zum Anfangswert $x_0 \in \mathbb{R}^n$ ist durch

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^n, \quad t \mapsto \gamma(t) = e^{At}x_0$$

gegeben, wobei die Matrix e^{At} für jedes $t \in \mathbb{R}$ durch die Exponentialreihe

$$e^{At} := \sum_{j=0}^{\infty} \frac{t^j A^j}{j!}$$

gegeben ist. (Dies wird in einer Übungsaufgabe gezeigt). \diamond

Es liegt nahe, dass auch im nicht-autonomen linearen Fall die Lösungen höchstens exponentiell wachsen können, falls $\|A(t)\|$ beschränkt bleibt. Da sie dann aber nicht in endlicher Zeit nach Unendlich laufen können, existieren sie gemäß Satz 8.23 für alle Zeiten in I .

9.3 Satz. Globale Existenz für lineare Systeme

Sei $J \subset \mathbb{R}$ offenes Intervall, $A : J \rightarrow M(n, \mathbb{R})$ und $b : J \rightarrow \mathbb{R}^n$ seien stetig. Dann existiert zu jedem $t_0 \in J$ und $x_0 \in \mathbb{R}^n$ eine eindeutige maximale Lösung $\gamma : J \rightarrow \mathbb{R}^n$ von

$$\gamma' = A(t)\gamma + b(t) \quad \text{mit} \quad \gamma(t_0) = x_0.$$

Beweis. Es erfüllt $v(t, x) := A(t)x + b(t)$ offenbar eine lokale Lipschitz-Bedingung. Daher liefert Satz 8.21 die Existenz einer eindeutigen maximalen Lösung auf dem Zeitintervall $I := (t_-, t_+) \subset J =: (j_-, j_+)$. Wir zeigen nun $t_+ = j_+$ (analog sieht man dann $t_- = j_-$, also $I = J$). Angenommen $t_+ < j_+$, dann muss nach Satz 8.23 die maximale Lösung $(t_-, t_+) \rightarrow \mathbb{R}^n$, $t \mapsto \gamma(t)$ jedes Kompaktum $K \subset \mathbb{R}^n$ verlassen, also $\|\gamma(t)\| \rightarrow \infty$ für $t \rightarrow t_+$ gelten. Wir zeigen nun, dass $\|\gamma(t)\|$ beschränkt bleibt für $t \rightarrow t_+$ falls $t_+ < j_+$. Insgesamt folgt dann $t_+ = j_+$.

9 Lineare Differentialgleichungen

Setze dazu $L := \max\{\|A(t)\| \mid t_0 \leq t \leq t_+\}$ und $M := \max\{\|b(t)\| \mid t_0 \leq t \leq t_+\}$, so gilt für die stetige Funktion $u : [t_0, t_+) \rightarrow [0, \infty)$, $u(t) = \|\gamma(t)\|$, dass

$$\begin{aligned} u(t) &= \|\gamma(t)\| = \|\gamma(t_0) + \int_{t_0}^t \gamma'(s) \, ds\| \leq \|\gamma(t_0)\| + \int_{t_0}^t \|\gamma'(s)\| \, ds \\ &= u(t_0) + \int_{t_0}^t \|A(s)\gamma(s) + b(s)\| \, ds \\ &\leq \left(u(t_0) + \int_{t_0}^{t_+} \|b(s)\| \, ds \right) + \int_{t_0}^t \underbrace{\|A(s)\|}_{\leq L} \cdot \underbrace{\|\gamma(s)\|}_{=u(s)} \, ds \\ &\leq C + L \int_{t_0}^t u(s) \, ds. \end{aligned}$$

9.4 Lemma. von Grönwall

Sei $a < b$ und $u : [a, b] \rightarrow [0, \infty)$ eine stetige Funktion. Es gebe Konstanten $L, C \geq 0$ so, dass für alle $t \in [a, b]$ gilt

$$u(t) \leq C + L \int_a^t u(s) \, ds. \quad (9.1)$$

Dann ist

$$u(t) \leq C e^{L(t-a)} \quad \text{für alle } t \in [a, b].$$

Beachte, dass (9.1) beispielsweise dann erfüllt ist, wenn u stetig differenzierbar ist mit $u' \leq Lu$ und $u(a) = C$.

Ende des Beweises von Satz 9.3. Also ist $u(t) \leq C e^{L(t-t_0)} \leq C e^{L(t_+-t_0)}$ für alle $t_0 \leq t \leq t_+$ und somit $t \mapsto \|\gamma(t)\|$ beschränkt für $t \rightarrow t_+$. \square

Beweis. des Lemma von Grönwall. Sei zunächst $C > 0$. Setze $U : [a, b] \rightarrow [0, \infty)$

$$U(t) := C + L \int_a^t u(s) \, ds.$$

Dann ist U stetig differenzierbar und $U'(t) = Lu(t) \geq 0$, also monoton steigend. Außerdem ist $U(a) = C > 0$, also $U(t) \geq C$ für alle $t \in [a, b]$. Es folgt

$$\frac{d}{dt} \ln(U(t)) = \frac{U'(t)}{U(t)} = \frac{Lu(t)}{U(t)} \leq L$$

da $u \leq U$. Also ist

$$\ln U(t) - \ln C = \ln U(t) - \ln U(a) = \int_a^t \frac{d}{ds} \ln U(s) \, ds \leq \int_a^t L \, ds = L(t-a).$$

und somit

$$u(t) \leq U(t) = \exp(\ln U(t)) \leq \exp(\ln C + L(t-a)) = C e^{L(t-a)}.$$

Falls $C = 0$, so liefert unser Argument, dass $u(t) \leq \tilde{C} e^{Lt}$ für jedes $\tilde{C} > 0$, also $u(t) \equiv 0$. \square

Die wichtigste Eigenschaft homogener linearer Systeme ist, dass Linearkombinationen von Lösungen wieder Lösungen sind.

9.5 Satz. Der Lösungsraum homogener linearer Systeme

Sei $J \subset \mathbb{R}$ ein offenes Intervall und $A : J \rightarrow M(n, \mathbb{R})$ stetig. Sei $L_h \subset C^1(J, \mathbb{R}^n)$ die Menge aller maximalen Lösungen des homogenen linearen Systems

$$\gamma' = A(t)\gamma \quad (9.2)$$

auf \mathbb{R}^n . Dann gilt:

- (a) L_h ist ein n -dimensionaler Unterraum von $C^1(J, \mathbb{R}^n)$.
- (b) Ist $t_0 \in J$, $r \in \mathbb{N}$ und $x_1, \dots, x_r \in \mathbb{R}^n$, so seien $\gamma_1, \dots, \gamma_r \in L_h$ die Lösungen von $\gamma' = A(t)\gamma$ mit $\gamma_j(t_0) = x_j$, $j = 1, \dots, r$. Dann sind äquivalent:
- (i) $\{\gamma_1, \dots, \gamma_r\}$ ist linear unabhängig in L_h .
 - (ii) $\{\gamma_1(t), \dots, \gamma_r(t)\}$ ist linear unabhängig in \mathbb{R}^n für alle $t \in J$.
 - (iii) $\{x_1, \dots, x_r\}$ ist linear unabhängig in \mathbb{R}^n .

Beweis. Zu (a): Seien $\gamma, \tilde{\gamma} \in L_h$ und $\lambda \in \mathbb{R}$. Dann ist

$$\frac{d}{dt}(\gamma + \tilde{\gamma}) = \gamma' + \tilde{\gamma}' = A(t)\gamma + A(t)\tilde{\gamma} = A(t)(\gamma + \tilde{\gamma})$$

und

$$\frac{d}{dt}(\lambda\gamma) = \lambda\gamma' = \lambda A(t)\gamma = A(t)(\lambda\gamma).$$

Also sind auch $\gamma + \tilde{\gamma}$ und $\lambda\gamma$ Lösungen von (9.2) und somit ist L_h ein Untervektorraum. Die Aussage zur Dimension folgt sofort aus Teil (b).

Zu (b): (i) \Rightarrow (ii): Seien $\{\gamma_1, \dots, \gamma_r\} \subset L_h$ linear unabhängig und sei $\tilde{t} \in J$ beliebig aber fest. Erfüllen dann $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ die Gleichung

$$\lambda_1\gamma_1(\tilde{t}) + \dots + \lambda_r\gamma_r(\tilde{t}) = 0 \tag{9.3}$$

als Gleichung in \mathbb{R}^n , so ist zu zeigen, dass aus (9.3) schon $\lambda_j = 0$ für $j = 1, \dots, r$ folgt. Setze dazu $\gamma : J \rightarrow \mathbb{R}^n$, $\gamma(t) := \lambda_1\gamma_1(t) + \dots + \lambda_r\gamma_r(t)$. Nach (a) ist $\gamma(t)$ Lösung von (9.2) und wegen (9.3) ist $\gamma(\tilde{t}) = 0$ und somit, wegen der Eindeutigkeit der Lösung, $\gamma(t) = 0$ für alle $t \in J$. Also ist

$$\lambda_1\gamma_1 + \dots + \lambda_r\gamma_r = 0$$

in $C^1(J, \mathbb{R}^n)$ und nach Voraussetzung $\lambda_1 = \dots = \lambda_r = 0$. Es ist also $(\gamma_1(\tilde{t}), \dots, \gamma_r(\tilde{t})) \subset \mathbb{R}^n$ linear unabhängig für alle $\tilde{t} \in J$.

(ii) \Rightarrow (iii): klar, weil $x_j = \gamma_j(t_0)$ ist,

(iii) \Rightarrow (i): auch klar, denn ist

$$\lambda_1\gamma_1 + \dots + \lambda_r\gamma_r = 0$$

so ist insbesondere

$$\lambda_1x_1 + \dots + \lambda_rx_r = \lambda_1\gamma_1(t_0) + \dots + \lambda_r\gamma_r(t_0) = 0.$$

Also muss nach Voraussetzung $\lambda_1 = \dots = \lambda_r = 0$ sein. □

9.6 Korollar. Der Propagator

Sei $A : J \rightarrow M(n, \mathbb{R})$ stetig, $t_0 \in J$ und $\varphi_t : \mathbb{R}^n \rightarrow \mathbb{R}^n$ der zu $\gamma' = A(t)\gamma$ gehörige Fluss (vgl. Bemerkung 8.22 (b)) zum Anfangszeitpunkt t_0 : Sei $\gamma : J \rightarrow \mathbb{R}^n$ die Lösung von $\gamma' = A(t)\gamma$ zum Anfangswert $\gamma(t_0) = x_0$, so ist also $\varphi_t(x_0) = \gamma(t)$.

Dann ist φ_t für jedes feste $t \in J$ ein linearer Isomorphismus des \mathbb{R}^n , welcher im Folgenden auch mit $\Phi(t)$ bezeichnet und **Propagator** genannt wird.

Die Abbildung $\Phi : J \rightarrow M(n, \mathbb{R})$ erfüllt die Matrix Differentialgleichung

$$\Phi'(t) = A(t)\Phi(t) \quad \text{mit Anfangsbedingung} \quad \Phi(t_0) = \text{Id}.$$

Beweis. Übungsaufgabe. □

9.7 Definition. Fundamentalsystem

Sei $A : J \rightarrow M(n, \mathbb{R})$ stetig. Ein n -Tupel $(\gamma_1, \dots, \gamma_n) \subset L_h$ von Lösungen der homogenen Gleichung $\gamma' = A(t)\gamma$ heißt **Lösungs-Fundamentalsystem**, wenn $(\gamma_1, \dots, \gamma_n)$ eine Basis von L_h ist.

9.8 Satz. Der Lösungsraum inhomogener linearer Systeme

Sei $J \subset \mathbb{R}$ ein offenes Intervall und $A : J \rightarrow M(n, \mathbb{R})$ und $b : J \rightarrow \mathbb{R}^n$ seien stetig. Mit $L_i \subset C^1(J, \mathbb{R}^n)$ werde die Teilmenge aller Lösungen der inhomogenen Gleichung

$$\gamma' = A(t)\gamma + b(t)$$

bezeichnet, also

$$L_i := \{\gamma \in C^1(J, \mathbb{R}^n) \mid \gamma' = A(t)\gamma + b(t)\}$$

und

$$L_h := \{\gamma \in C^1(J, \mathbb{R}^n) \mid \gamma' = A(t)\gamma\}.$$

(a) Ist $\gamma : J \rightarrow \mathbb{R}^n$ eine Lösung der inhomogenen Gleichung, so gilt

$$L_i = \gamma + L_h := \{\gamma + \xi \in C^1(J, \mathbb{R}^n) \mid \xi \in L_h\}.$$

(b) **Variation der Konstanten**

Sei $\Phi(t) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ der Propagator des homogenen Systems zur Anfangszeit $t_0 \in J$ und sei $x_0 \in \mathbb{R}^n$. Dann ist

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

die Lösung der inhomogenen Gleichung $\gamma' = A(t)\gamma + b(t)$ zum Anfangswert $\gamma(t_0) = x_0$.

Beweis. (a) Ist γ eine Lösung von $\gamma' = A(t)\gamma + b(t)$, so ist $\tilde{\gamma}$ genau dann eine weitere Lösung von $\gamma' = A(t)\gamma + b(t)$, wenn $\xi := \tilde{\gamma} - \gamma$ Lösung der homogenen Gleichung ist:

$$\tilde{\gamma}' = A(t)\tilde{\gamma} + b(t) \quad \Leftrightarrow \quad \xi' = \tilde{\gamma}' - \gamma' = A(t)\tilde{\gamma} - A(t)\gamma = A(t)\xi.$$

(b) Es gilt $\gamma(t_0) = \Phi(t_0)x_0 = x_0$ und

$$\begin{aligned} \gamma'(t) &= \Phi'(t) \left(x_0 + \int_{t_0}^t \Phi(s)^{-1} b(s) ds \right) + \Phi(t)\Phi(t)^{-1} b(t) \\ &= A(t)\Phi(t) \left(x_0 + \int_{t_0}^t \Phi(s)^{-1} b(s) ds \right) + b(t) \\ &= A(t)\gamma(t) + b(t). \end{aligned} \quad \square$$

9.9 Bemerkung. Die Bezeichnung „Variation der Konstanten“ kommt aus folgendem Ansatz für die Lösung: $\gamma(t) = \Phi(t)c(t)$ statt $\gamma(t) = \Phi(t)x_0$ wie für die homogene Gleichung. Dann muss nämlich $c(t)$ folgende Differentialgleichung erfüllen:

$$\gamma' = \Phi'c + \Phi c' = A\Phi c + \Phi c' = A\gamma + \Phi c' \stackrel{!}{=} A\gamma + b(t)$$

also $\Phi c' = b$ oder $c'(t) = \Phi^{-1}(t)b(t)$. Somit folgt

$$c(t) = x_0 + \int_{t_0}^t \Phi^{-1}(s)b(s) ds.$$

9.10 Bemerkung. Die Menge aller Lösungen des homogenen Systems $\gamma' = A(t)\gamma$ ist ein n -dimensionaler Unterraum von $C^1(J, \mathbb{R}^n)$. Die Menge aller Lösungen des inhomogenen Systems ist ein n -dimensionaler affiner Unterraum von $C^1(J, \mathbb{R}^n)$. Kennt man die vollständige Lösung des homogenen Systems (also zumindest n linear unabhängige Lösungen) und eine einzige Lösung des inhomogenen Systems, so kann man auch das inhomogene System vollständig lösen. \diamond

9.11 Beispiel. Betrachte die Differentialgleichung

$$\gamma' = 2t\gamma + t^3 \quad (*)$$

auf \mathbb{R} . Mit $a(t) = 2t$ und $b(t) = t^3$ wird das zu $\gamma' = a(t)\gamma + b(t)$. Die Lösung der homogenen Gleichung $\gamma' = a(t)\gamma$ mit $\gamma(t_0) = x_0$ ist

$$\gamma(t) = e^{\int_{t_0}^t a(s) ds} x_0,$$

also mit $t_0 = 0$

$$\gamma(t) = e^{\int_0^t 2s ds} x_0 = e^{t^2} x_0.$$

Da $\dim L_h = 1$ ist, ist $\gamma(t) = e^{t^2} x_0$ eine Basis für L_h und der Propagator ist $\Phi(t) = e^{t^2}$. Variation der Konstanten liefert nun die Lösung der inhomogenen Gleichung zum Anfangswert x_0 durch

$$\gamma_i(t) = e^{t^2} \left(x_0 + \int_0^t e^{-s^2} s^3 ds \right) = e^{t^2} \left(x_0 + \frac{1}{2} \right) - \frac{1}{2}(1 + t^2),$$

wobei

$$\int_0^t e^{-s^2} s^3 ds = \frac{1}{2} \int_0^{t^2} e^{-\tau} \tau d\tau = -\frac{1}{2} e^{-\tau} \tau \Big|_0^{t^2} + \frac{1}{2} \int_0^{t^2} e^{-\tau} d\tau = -\frac{1}{2} e^{-t^2} t^2 - \frac{1}{2} e^{-t^2} + \frac{1}{2}. \quad \diamond$$

Es stellt sich nun die Frage, wie man den Propagator der homogenen Gleichung auch im Fall $n \geq 2$ bestimmen kann.

9.12 Bemerkung. Der Propagator im autonomen Fall

Für den autonomen Fall $A(t) \equiv A$ erhält man die Lösung $\Phi : \mathbb{R} \rightarrow M(n, \mathbb{R})$ von

$$\Phi' = A\Phi$$

mit $\Phi(0) = E_n$ durch exponentieren

$$\Phi(t) = e^{At},$$

vgl. Beispiel 9.2. Insbesondere existieren die Lösungen für alle Zeiten. Um $e^{At} := \sum_{j=0}^{\infty} \frac{(At)^j}{j!}$ explizit auszurechnen, transformiert man A auf Jordansche Normalform (vgl. Lineare Algebra).

Wir betrachten hier nur den Fall, dass A diagonalisierbar ist, also ein $S \in GL_n(\mathbb{R})$ existiert mit

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = D.$$

Dann ist

$$e^{Dt} = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix},$$

und

$$\Phi(t) := Se^{Dt}S^{-1}$$

erfüllt die Differentialgleichung

$$\Phi'(t) = SDe^{Dt}S^{-1} = SDS^{-1}Se^{Dt}S^{-1} = A(t)\Phi(t)$$

und $\Phi(0) = E_n$. Also ist $\Phi(t) = e^{At}$.

9 Lineare Differentialgleichungen

Beachte, dass für $A(t) \in M(n, \mathbb{R})$ zwar S und $\lambda_1, \dots, \lambda_n$ komplex sein können, der Propagator

$$\Phi(t) = S \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} S^{-1}$$

ist aber immer noch reell. ◇

9.13 Bemerkung. Der Propagator im nicht-autonomen Fall: Die Dyson Reihe

Auch im nicht-autonomen Fall und $n \geq 2$ kann man noch eine explizite Formel für die Lösung der homogenen Gleichung angeben, die sogenannte Dyson Reihe. Sei $A : J \rightarrow M(n, \mathbb{R})$ stetig, dann ist die Lösung $\Phi : J \rightarrow M(n, \mathbb{R})$ von

$$\Phi' = A(t)\Phi$$

mit $\Phi(t_0) = E_n$ gegeben durch die absolut konvergente Reihe

$$\Phi(t) = E_n + \sum_{j=1}^{\infty} \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \cdots \int_{t_0}^{\tau_{j-1}} d\tau_j A(\tau_1) \cdots A(\tau_j).$$

Beweis. Übungsaufgabe. □

9.14 Bemerkung. Alles in diesem Kapitel Gesagte gilt analog für lineare Differentialgleichungen $\gamma' = A(t)\gamma$ auf \mathbb{C}^n statt auf \mathbb{R}^n . D.h. man sucht Lösungen $\gamma : J \rightarrow \mathbb{C}^n$ von

$$\gamma' = A(t)\gamma \quad \text{bzw.} \quad \gamma' = A(t)\gamma + b(t),$$

wobei $A : J \rightarrow M(n \times n, \mathbb{C})$ und $b : J \rightarrow \mathbb{C}^n$ stetig sind. Die Zeitvariable bleibt hier aber reell! Die Lösungen der homogenen Gleichung bilden dann einen Unterraum der komplexen Dimension n in $C^1(J, \mathbb{C}^n)$.

Die Resultate zu linearen Differentialgleichungssystemen erster Ordnung lassen sich direkt auf lineare (Systeme) m -ter Ordnung übertragen.

9.15 Definition. Lineare DGL m -ter Ordnung

Sei $J \subset \mathbb{R}$ ein offenes Intervall und $a_k : J \rightarrow \mathbb{K}$, $k = 1, \dots, m-1$, stetige Funktionen. Hier ist jetzt $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Dann heißt

$$\gamma^{(m)}(t) + a_{m-1}(t)\gamma^{(m-1)}(t) + \dots + a_1(t)\gamma'(t) + a_0(t)\gamma(t) = 0$$

homogene lineare Differentialgleichung m -ter Ordnung.

Ist $b : J \rightarrow \mathbb{K}$ stetig, so heißt

$$\gamma^{(m)}(t) + a_{m-1}(t)\gamma^{(m-1)}(t) + \dots + a_1(t)\gamma'(t) + a_0(t)\gamma(t) = b(t)$$

inhomogene lineare Differentialgleichung m -ter Ordnung. ◇

9.16 Satz. Lösungsraum linearer DGLen m -ter Ordnung

(a) Sei L_h die Menge aller Lösungen $\gamma : J \rightarrow \mathbb{K}$ der homogenen Gleichung

$$\gamma^{(m)} + a_{m-1}\gamma^{(m-1)} + \dots + a_1\gamma' + a_0\gamma = 0.$$

Dann ist L_h ein m -dimensionaler Unterraum von $C^m(J, \mathbb{K})$.

(b) Sei L_i die Menge aller Lösungen $\gamma : J \rightarrow \mathbb{K}$ der inhomogenen Differentialgleichung

$$\gamma^{(m)} + a_{m-1} \gamma^{(m-1)} + \dots + a_1 \gamma' + a_0 \gamma = b$$

und $\tilde{\gamma} \in L_i$ beliebig. Dann gilt

$$L_i = \tilde{\gamma} + L_h.$$

(c) Ein m -Tupel (ξ_1, \dots, ξ_m) in L_h ist genau dann linear unabhängig, wenn für ein und damit für alle $t \in J$ die **Wronski-Determinante**

$$W(t) := \det \begin{pmatrix} \xi_1(t) & \dots & \xi_m(t) \\ \xi_1'(t) & \dots & \xi_m'(t) \\ \vdots & & \vdots \\ \xi_1^{(m-1)}(t) & \dots & \xi_m^{(m-1)}(t) \end{pmatrix}$$

von Null verschieden ist.

Beweis. Die Differentialgleichung

$$\gamma^{(m)} + a_{m-1} \gamma^{(m-1)} + \dots + a_1 \gamma' + a_0 \gamma = b \tag{9.4}$$

ist äquivalent zu dem inhomogenen linearen System 1. Ordnung

$$\begin{aligned} \alpha_0' &= \alpha_1 \\ \alpha_1' &= \alpha_2 \\ &\vdots \\ \alpha_{m-2}' &= \alpha_{m-1} \\ \alpha_{m-1}' &= -a_0 \alpha_0 - a_1 \alpha_1 - \dots - a_{m-1} \alpha_{m-1} + b \end{aligned} \tag{9.5}$$

Jeder Lösung $\gamma : J \rightarrow \mathbb{K}$ von (9.4) entspricht eine Lösung

$$\begin{pmatrix} \gamma \\ \gamma' \\ \vdots \\ \gamma^{(m-1)} \end{pmatrix} : J \rightarrow \mathbb{K}^m$$

von (9.5) und umgekehrt. Entsprechendes gilt für die homogenen Gleichungen ($b = 0$). Damit folgen die Behauptungen aus Satz 9.8. \square

9.17 Beispiel. Die Differentialgleichung

$$\gamma'' - \frac{1}{2t} \gamma' + \frac{1}{2t^2} \gamma = 0$$

auf dem Intervall $J = (0, \infty)$ besitzt die Lösungen $\gamma_1(t) := t$ und $\gamma_2(t) := \sqrt{t}$, wovon man sich durch Einsetzen überzeugt.

Die Wronski-Determinante von (γ_1, γ_2) ist

$$W(t) = \det \begin{pmatrix} \gamma_1(t) & \gamma_2(t) \\ \gamma_1'(t) & \gamma_2'(t) \end{pmatrix} = \det \begin{pmatrix} t & \sqrt{t} \\ 1 & \frac{1}{2\sqrt{t}} \end{pmatrix} = -\frac{\sqrt{t}}{2}.$$

Da $W(t) \neq 0$ für $t \in J$, bilden γ_1 und γ_2 ein Lösungs-Fundamentalsystem. Die allgemeine Lösung der Differenzialgleichung ist also

$$\gamma(t) = c_1 \gamma_1(t) + c_2 \gamma_2(t) = c_1 t + c_2 \sqrt{t}$$

mit beliebige Konstanten $c_1, c_2 \in \mathbb{C}$. \diamond

9.18 Beispiele. (a) Die Legendresche Differentialgleichung

Die Legendresche Differentialgleichung auf $J = (-1, 1)$ zu $n \in \mathbb{N}$ ist

$$(1 - t^2)\gamma'' - 2t\gamma' + n(n + 1)\gamma = 0$$

bzw., da $(1 - t^2) \neq 0$,

$$\gamma'' - \frac{2t}{(1 - t^2)}\gamma' + \frac{n(n + 1)}{(1 - t^2)}\gamma = 0.$$

Das Legendre-Polynom der Ordnung n ist definiert durch

$$P_n(t) := \frac{1}{2^n n!} \left(\frac{d}{dt} \right)^n (t^2 - 1)^n$$

und löst die Legendresche Differentialgleichung der Ordnung n .

(b) Die Hermitesche Differentialgleichung

Die Hermitesche Differentialgleichung auf $J = \mathbb{R}$ zu $n \in \mathbb{N}$ ist

$$\gamma'' - 2t\gamma' + 2n\gamma = 0.$$

Das Hermite Polynom der Ordnung n ist definiert durch

$$H_n(t) = (-1)^n e^{t^2} \left(\frac{d}{dt} \right)^n e^{-t^2}$$

und löst die Hermitesche Differentialgleichung der Ordnung n .

(c) Die Laguerresche Differentialgleichung

Die Laguerresche Differentialgleichung auf $J = (0, \infty)$ zu $n \in \mathbb{N}$ ist

$$t\gamma'' + (1 - t)\gamma' + n\gamma = 0.$$

Das Laguerresche Polynom der Ordnung n ist definiert durch

$$L_n(t) := e^t \left(\frac{d}{dt} \right)^n (t^n e^{-t})$$

und löst die Laguerresche Differentialgleichung der Ordnung n .

In allen drei Fällen (a), (b) und (c) sind die Polynome jeweils nur eine spezielle Lösung und gemäß Satz 9.16 gibt es jeweils noch eine weitere linear unabhängige Lösung. \diamond

9.19 Satz. Sei $\gamma : J \rightarrow \mathbb{R}$ eine Lösung der homogenen linearen Differentialgleichung zweiter Ordnung

$$\gamma'' + a(t)\gamma' + b(t)\gamma = 0.$$

Im Intervall $I \subset J$ gelte $\gamma(t) \neq 0$. Dann erhält man über I eine zweite von γ linear unabhängige Lösung $\tilde{\gamma} : I \rightarrow \mathbb{R}$ durch den Ansatz

$$\tilde{\gamma}(t) = \gamma(t)u(t),$$

wobei u eine nicht-konstante Lösung der Differentialgleichung

$$u'' + \left(2 \frac{\gamma'(t)}{\gamma(t)} + a(t) \right) u' = 0 \tag{9.6}$$

ist.

9.20 Bemerkung. Die Gleichung (9.6) ist eine lineare Differentialgleichung erster Ordnung für u' , welche die Lösung

$$\begin{aligned} u'(t) &= u'(t_0) e^{-\int_{t_0}^t (2\frac{\gamma'(s)}{\gamma(s)} + a(s)) ds} \\ &= u'(t_0) e^{-2(\ln\gamma(t) - \ln\gamma(t_0))} e^{-\int_{t_0}^t a(s) ds} \\ &= u'(t_0) \frac{\gamma(t_0)^2}{\gamma(t)^2} e^{-\int_{t_0}^t a(s) ds} \end{aligned}$$

hat. Man erhält u dann durch eine weitere Integration. \diamond

Beweis. von Satz 9.19.

$$\tilde{\gamma} = \gamma u \quad \Rightarrow \quad \tilde{\gamma}' = \gamma' u + \gamma u' \quad \Rightarrow \quad \tilde{\gamma}'' = \gamma'' u + 2\gamma' u' + \gamma u''$$

Also

$$\tilde{\gamma}'' + a\tilde{\gamma}' + b\tilde{\gamma} = \gamma'' u + 2\gamma' u' + \gamma u'' + a\gamma' u + a\gamma u' + b\gamma u = 2\gamma' u' + \gamma u'' + a\gamma u',$$

da $\gamma'' + a\gamma' + b\gamma = 0$. Somit löst $\tilde{\gamma}$ die Differentialgleichung, wenn $\gamma u'' + 2\gamma' u' + a\gamma u' = 0$ bzw. wenn

$$u'' + \left(2\frac{\gamma'}{\gamma} + a\right) u' = 0.$$

Ist u nicht konstant, so sind $\tilde{\gamma} = u\gamma$ und γ auf I linear unabhängig. \square

9.21 Beispiel. Für $n = 1$ ist die Legendresche Differentialgleichung auf $J = (-1, 1)$

$$\gamma'' - \frac{2t}{1-t^2} \gamma' + \frac{2}{1-t^2} \gamma = 0.$$

Sie hat die Lösung $P_1(t) = t$. Also erhält man auf $I = (0, 1)$ eine zweite Lösung durch den Ansatz $\tilde{\gamma}(t) = t u(t)$, wobei

$$\begin{aligned} u'(t) &= u'(t_0) \frac{t_0^2}{t^2} e^{\int_{t_0}^t \frac{2s}{1-s^2} ds} = u'(t_0) \frac{t_0^2}{t^2} e^{-\ln(1-s^2)|_{t_0}^t} = u'(t_0) \frac{t_0^2}{t^2} \frac{1-t_0^2}{1-t^2} \\ &= u'(t_0) t_0^2 (1-t_0^2) \left(\frac{1}{t^2} + \frac{1}{2} \left(\frac{1}{1+t} + \frac{1}{1-t} \right) \right) \end{aligned}$$

somit ist

$$\begin{aligned} u(t) &= u(t_0) + \int_{t_0}^t u'(s) ds = u(t_0) + u'(t_0) t_0^2 (1-t_0^2) \left(-\frac{1}{s} \Big|_{t_0}^t + \frac{1}{2} (\ln(1+s) - \ln(1-s)) \Big|_{t_0}^t \right) \\ &= u(t_0) + u'(t_0) t_0^2 (1-t_0^2) \left(\frac{1}{t_0} - \frac{1}{t} + \frac{1}{2} \ln \frac{1+t}{1-t} \right). \end{aligned}$$

Wir wählen nun für beliebiges $t_0 \in (0, 1)$

$$u'(t_0) = \frac{1}{t_0^2(1-t_0^2)} \quad \text{und} \quad u(t_0) = -u'(t_0)t_0(1-t_0^2)$$

und erhalten so die Lösung

$$u(t) = \frac{1}{2} \ln \frac{1+t}{1-t} - \frac{1}{t}.$$

Es ist also

$$\tilde{\gamma}(t) = t u(t) = \frac{t}{2} \ln \frac{1+t}{1-t} - 1$$

eine von $P_1(t)$ linear unabhängige Lösung auf dem Intervall $I = (0, 1)$. Man rechnet nun aber direkt nach, dass $\tilde{\gamma}(t)$ die Differentialgleichung auf dem ganzen Intervall $J = (-1, 1)$ löst. \diamond

9.22 Bemerkung. Lineare Differentialgleichungen mit konstanten Koeffizienten

Eine homogene lineare Differentialgleichung m -ter Ordnung mit konstanten Koeffizienten ist von der Form

$$\sum_{j=0}^m a_j \gamma^{(j)}(t) = 0 \quad \text{mit } a_m = 1, \tag{9.7}$$

für $\gamma : \mathbb{R} \rightarrow \mathbb{R}$ oder $\gamma : \mathbb{R} \rightarrow \mathbb{C}$. Sie ist äquivalent zu dem System erster Ordnung

$$\alpha'(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ -a_0 & -a_1 & \cdots & -a_{m-2} & -a_{m-1} \end{pmatrix} \alpha(t) =: A \alpha(t)$$

für $\alpha : \mathbb{R} \rightarrow \mathbb{R}^m$ oder $\alpha : \mathbb{R} \rightarrow \mathbb{C}^m$. Jeder Eigenvektor v von A zum Eigenwert λ liefert eine Lösung $\alpha(t) = e^{\lambda t} v$ des Systems erster Ordnung. Das charakteristische Polynom von A ist

$$P_A(\lambda) = (-1)^m \sum_{j=0}^m a_j \lambda^j,$$

läßt sich also direkt aus der Differentialgleichung (9.7) ablesen. Damit haben wir ein Rezept zum Auffinden von Lösungen von (9.7) gefunden: zu jeder Nullstelle λ_0 von $P_A(\lambda)$ ist

$$\gamma_{\lambda_0}(t) := e^{\lambda_0 t}$$

eine Lösung von (9.7). Hat $P_A(\lambda)$ tatsächlich m verschiedene Nullstellen, so bilden die zugehörigen Lösungen ein Fundamentalsystem. Aber auch beim Vorliegen einer ℓ -fachen Nullstelle kann man direkt ℓ linear unabhängige Lösungen angeben: sei λ_0 ℓ -fache Nullstelle von $P_A(\lambda)$, dann sind

$$\gamma_{\lambda_0,0}(t) := e^{\lambda_0 t}, \quad \gamma_{\lambda_0,1}(t) := t e^{\lambda_0 t}, \quad \gamma_{\lambda_0,2}(t) := t^2 e^{\lambda_0 t}, \quad \dots, \quad \gamma_{\lambda_0,\ell-1}(t) := t^{\ell-1} e^{\lambda_0 t}$$

ℓ linear unabhängige Lösungen von (9.7). (Beweis in den Übungen).

9.23 Beispiel. Der gedämpfte harmonische Oszillator

Das charakteristische Polynom zur Differentialgleichung

$$\gamma'' + 2d\gamma' + \omega^2\gamma = 0, \quad d, \omega > 0,$$

hat die Form $P(\lambda) = \lambda^2 + 2d\lambda + \omega^2$ und somit die Nullstellen

$$\lambda_{\pm} = -d \pm \sqrt{d^2 - \omega^2}.$$

Wir unterscheiden drei Fälle:

Überdämpfte Bewegung: Sei $d > \omega$, also $\lambda_{\pm} \in \mathbb{R}$ und $\lambda_+ \neq \lambda_-$. Zwei linear unabhängige Lösungen sind somit

$$\gamma_1(t) = e^{\lambda_+ t} = e^{(-d + \sqrt{d^2 - \omega^2})t} \quad \text{und} \quad \gamma_2(t) = e^{\lambda_- t} = e^{(-d - \sqrt{d^2 - \omega^2})t},$$

und die allgemeine reelle Lösung lautet

$$\gamma(t) = \left(c_1 e^{\sqrt{d^2 - \omega^2} t} + c_2 e^{-\sqrt{d^2 - \omega^2} t} \right) e^{-dt} = \left((c_1 + c_2) \cosh(\sqrt{d^2 - \omega^2} t) + (c_1 - c_2) \sinh(\sqrt{d^2 - \omega^2} t) \right) e^{-dt}$$

mit $c_1, c_2 \in \mathbb{R}$. Aus

$$\gamma(0) = c_1 + c_2 \stackrel{!}{=} x_0 \quad \text{und} \quad \gamma'(0) = \sqrt{d^2 - \omega^2}(c_1 - c_2) - d(c_1 + c_2) \stackrel{!}{=} v_0$$

ergibt sich auch leicht

$$c_1 + c_2 = x_0 \quad \text{und} \quad c_1 - c_2 = \frac{v_0 + dx_0}{\sqrt{d^2 - \omega^2}},$$

also

$$\gamma(t) = \left(x_0 \cosh(\sqrt{d^2 - \omega^2} t) + \frac{v_0 + dx_0}{\sqrt{d^2 - \omega^2}} \sinh(\sqrt{d^2 - \omega^2} t) \right) e^{-dt}.$$

Kritische Dämpfung: Für $d = \omega$ ist $\lambda_* = -d$ zweifacher Eigenwert. Zwei linear unabhängige Lösungen sind diesmal

$$\gamma_1(t) = e^{\lambda_* t} = e^{-dt} \quad \text{und} \quad \gamma_2(t) = t e^{\lambda_* t} = t e^{-dt},$$

und die allgemeine reelle Lösung lautet

$$\gamma(t) = (c_1 + c_2 t) e^{-dt}$$

mit $c_1, c_2 \in \mathbb{R}$ bzw.

$$\gamma(t) = (x_0 + (v_0 + dx_0)t) e^{-dt}.$$

Gedämpfte Schwingung: Sei $d < \omega$, also $\lambda_{\pm} \in \mathbb{C}$ und $\lambda_+ \neq \lambda_-$. Zwei linear unabhängige Lösungen sind

$$\gamma_1(t) = e^{\lambda_+ t} = e^{-dt} e^{i\tilde{\omega} t} \quad \text{und} \quad \gamma_2(t) = e^{\lambda_- t} = e^{-dt} e^{-i\tilde{\omega} t} \quad \text{mit } \tilde{\omega} := \sqrt{\omega^2 - d^2},$$

und die allgemeine komplexe Lösung lautet

$$\gamma_{\mathbb{C}}(t) = (c_1 e^{i\tilde{\omega} t} + c_2 e^{-i\tilde{\omega} t}) e^{-dt}$$

mit $c_1, c_2 \in \mathbb{C}$. Wählt man $c_2 = \bar{c}_1$, so ist

$$\gamma_{\mathbb{R}}(t) := (c_1 e^{i\tilde{\omega} t} + \bar{c}_1 e^{-i\tilde{\omega} t}) e^{-dt} = 2\operatorname{Re}(c_1 e^{i\tilde{\omega} t}) e^{-dt} = (a \cos(\tilde{\omega} t) + b \sin(\tilde{\omega} t)) e^{-dt}$$

mit $a = 2\operatorname{Re}(c_1) \in \mathbb{R}$ und $b = -2\operatorname{Im}(c_1) \in \mathbb{R}$ wieder reell und zwar die allgemeine reelle Lösung. Drückt man wiederum a und b durch x_0 und v_0 aus, so ergibt sich

$$\gamma(t) = \left(x_0 \cos(\tilde{\omega} t) + \frac{v_0 + dx_0}{\tilde{\omega}} \sin(\tilde{\omega} t) \right) e^{-dt}.$$

◇

10 Integration im \mathbb{R}^n : Eine kurze Einführung

Sie werden in Analysis 3 eine sehr mächtige und vielseitig einsetzbare Integrationstheorie kennenlernen. Wir werden uns in diesem Kapitel daher sehr kurz fassen und nur einige einfache Überlegungen zur Integration stetiger Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit kompaktem Träger anstellen. Die meisten Aussagen, die wir zeigen, gelten für eine deutlich größere Klasse von Funktionen.

Sei $Q := [a, b] \times [c, d]$ ein Rechteck im \mathbb{R}^2 und $f : Q \rightarrow \mathbb{R}$ eine stetige Funktion. Es liegt nahe, das Integral von f über Q durch

$$\int_Q f := \int_a^b \left(\int_c^d f(x, y) dy \right) dx \quad (10.1)$$

zu definieren. Dann stellt sich aber sofort die Frage, ob der Wert des Integrals unabhängig von der Integrationsreihenfolge ist, also ob

$$\int_a^b \left(\int_c^d f(x, y) dy \right) dx \stackrel{?}{=} \int_c^d \left(\int_a^b f(x, y) dx \right) dy \quad (10.2)$$

gilt?

10.1 Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = \sin(x + 3y)$. Dann ist

$$\begin{aligned} \int_0^\pi \left(\int_0^{\pi/2} f(x, y) dx \right) dy &= - \int_0^\pi (\cos(\frac{\pi}{2} + 3y) - \cos(3y)) dy \\ &= -\frac{1}{3} (\sin(\frac{7\pi}{2}) - \sin(\frac{\pi}{2}) - \sin(3\pi) + \sin(0)) = \frac{2}{3} \end{aligned}$$

und

$$\begin{aligned} \int_0^{\pi/2} \left(\int_0^\pi f(x, y) dx \right) dy &= -\frac{1}{3} \int_0^{\pi/2} (\cos(x + 3\pi) - \cos(x)) dx \\ &= -\frac{1}{3} (\sin(\frac{7\pi}{2}) - \sin(3\pi) - \sin(\frac{\pi}{2}) + \sin(0)) = \frac{2}{3}, \end{aligned}$$

beide Integrationsreihenfolgen liefern also denselben Wert. Wir werden gleich sehen, dass das kein Zufall ist. \diamond

Normalerweise wird das Integral über Q (oder auch viel allgemeinere Gebiete) zunächst direkt definiert und die Frage nach der Vertauschbarkeit der Integrationsreihenfolge bzw. die Darstellbarkeit des Integrals über Q durch iterierte Integrale über $[a, b]$ und $[c, d]$ wird durch den *Satz von Fubini* beantwortet. Wir kürzen die Sache hier erheblich ab, indem wir lediglich zeigen, dass für stetige Funktionen in (10.2) tatsächlich Gleichheit gilt und somit (10.1) in diesem Fall eine sinnvolle Definition darstellt.

Wir benötigen dazu einige Vorüberlegungen, die wir in einzelnen Lemmas formulieren.

10.2 Lemma. Sei $U \subset \mathbb{R}^n$, $Q := [a, b] \times U$ und $f : Q \rightarrow \mathbb{R}$ stetig. Sei weiterhin (y_k) eine konvergente Folge in U mit $y_* := \lim_{k \rightarrow \infty} y_k \in U$. Dann konvergiert die Funktionenfolge

$$f_k : [a, b] \rightarrow \mathbb{R}, \quad x \mapsto f_k(x) := f(x, y_k)$$

gleichmäßig gegen die Funktion $f_* : [a, b] \rightarrow \mathbb{R}$, $x \mapsto f_*(x) := f(x, y_*)$.

10 Integration im \mathbb{R}^n : Eine kurze Einführung

Beweis. Sei $K \subset U$ ein Kompaktum mit $\{y_k \mid k \in \mathbb{N}\} \cup \{y_*\} \subset K$. In Aufgabe 13 (b) auf Blatt 4 wurde gezeigt, dass f als stetige Funktion auf dem Kompaktum $[a, b] \times K$ gleichmäßig stetig ist, d.h. zu jedem $\varepsilon > 0$ gibt es ein $\delta(\varepsilon) > 0$ so, dass für alle $(x, y), (\tilde{x}, \tilde{y}) \in [a, b] \times K$ gilt:

$$\|(x, y) - (\tilde{x}, \tilde{y})\|_{\mathbb{R}^{n+1}} < \delta(\varepsilon) \quad \Rightarrow \quad |f(x, y) - f(\tilde{x}, \tilde{y})| < \varepsilon.$$

Da $\lim_{n \rightarrow \infty} y_k = y_*$, gibt es zu jedem $\varepsilon > 0$ ein $k_0 \in \mathbb{N}$ so, dass für alle $k \geq k_0$

$$\|y_k - y_*\|_{\mathbb{R}^n} = \|(x, y_k) - (x, y_*)\|_{\mathbb{R}^{n+1}} < \delta(\varepsilon)$$

und somit

$$\|f_k - f_*\|_{\infty} = \sup_{x \in [a, b]} |f(x, y_k) - f(x, y_*)| < \varepsilon. \quad \square$$

10.3 Lemma. Parameterintegrale: Stetigkeit

Sei $U \subset \mathbb{R}^n$, $Q := [a, b] \times U$ und $f : Q \rightarrow \mathbb{R}$ stetig. Dann ist

$$g : U \rightarrow \mathbb{R}, \quad y \mapsto g(y) := \int_a^b f(x, y) \, dx$$

eine stetige Funktion.

Beweis. Sei $y_* \in U$ und (y_k) eine Folge in U mit $\lim_{k \rightarrow \infty} y_k = y_*$. Wir schränken uns wieder auf ein Kompaktum $K \subset U$ mit $\{y_k \mid k \in \mathbb{N}\} \cup \{y_*\} \subset K$ ein. Aufgrund der Stetigkeit von f folgt

$$\lim_{k \rightarrow \infty} g(y_k) = \lim_{k \rightarrow \infty} \int_a^b f(x, y_k) \, dx \stackrel{(*)}{=} \int_a^b \lim_{k \rightarrow \infty} f(x, y_k) \, dx = \int_a^b f(x, y_*) \, dx = g(y_*),$$

also die Stetigkeit von g , wobei wir die Vertauschung von Grenzwert und Integration in $(*)$ noch begründen müssen. Eine hinreichende Bedingung für die Gültigkeit von $(*)$ ist gemäß Satz 8.35 aus Analysis 1, dass die Folge der Integranden $f_k : [a, b] \rightarrow \mathbb{R}, x \mapsto f_k(x) := f(x, y_k)$, gleichmäßig gegen $f_* : [a, b] \rightarrow \mathbb{R}, x \mapsto f_*(x) := f(x, y_*)$ konvergiert. Aber das ist gemäß Lemma 10.2 der Fall. \square

10.4 Lemma. Parameterintegrale: Differenzierbarkeit

Sei $U \subset \mathbb{R}^n$ offen, $Q := [a, b] \times U$ und $f : Q \rightarrow \mathbb{R}$ stetig und bezüglich der letzten n Variablen auch stetig partiell differenzierbar. Dann ist

$$g : U \rightarrow \mathbb{R}, \quad y \mapsto g(y) := \int_a^b f(x, y) \, dx$$

stetig partiell differenzierbar und es gilt für alle $y \in U$

$$\partial_{y_j} g(y) = \int_a^b \partial_{y_j} f(x, y) \, dx.$$

Beweis. Es gilt für jede Nullfolge (h_k) in \mathbb{R} , dass

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{g(y + h_k e_j) - g(y)}{h_k} &= \lim_{k \rightarrow \infty} \int_a^b \frac{f(x, y + h_k e_j) - f(x, y)}{h_k} \, dx \\ &\stackrel{(*)}{=} \int_a^b \lim_{k \rightarrow \infty} \frac{f(x, y + h_k e_j) - f(x, y)}{h_k} \, dx = \int_a^b \partial_{y_j} f(x, y) \, dx, \end{aligned}$$

falls wir in (*) wieder Limes und Integration vertauschen können, also falls $\frac{f(x, y + h_k e_j) - f(x, y)}{h_k}$ gleichmäßig in x für $n \rightarrow \infty$ gegen $\partial_{y_j} f(x, y)$ konvergiert. Nach dem Mittelwertsatz ist aber

$$\frac{f(x, y + h_k e_j) - f(x, y)}{h_k} = \partial_{y_j} f(x, y + \theta(h_k) h_k) =: \partial_{y_j} f(x, y_k)$$

mit $\lim_{k \rightarrow \infty} y_k = y$. Da $\partial_{y_j} f$ als stetig vorausgesetzt ist, folgt die gleichmäßige Konvergenz wieder mit Lemma 10.2. \square

10.5 Satz. Das Integral stetiger Funktionen über Rechtecke

Sei $Q := [a, b] \times [c, d]$ ein Rechteck im \mathbb{R}^2 und $f : Q \rightarrow \mathbb{R}$ eine stetige Funktion. Dann sind

$$g : [a, b] \rightarrow \mathbb{R}, \quad x \mapsto g(x) := \int_c^d f(x, y) \, dy \quad \text{und} \quad h : [c, d] \rightarrow \mathbb{R}, \quad y \mapsto h(y) := \int_a^b f(x, y) \, dx$$

stetige Funktionen und es gilt

$$\int_a^b \left(\int_c^d f(x, y) \, dy \right) dx = \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy.$$

Beweis. Die Stetigkeit von g und h folgt jeweils sofort mit Lemma 10.3. Weiterhin ist die Funktion

$$\tilde{g} : [a, b] \times [c, d] \rightarrow \mathbb{R}, \quad (x, t) \mapsto \tilde{g}(x, t) := \int_c^t f(x, y) \, dy$$

stetig und nach dem Hauptsatz bezüglich t sogar stetig differenzierbar. Also gilt mit Lemma 10.4 für die Funktion $\phi : [c, d] \rightarrow \mathbb{R}$,

$$t \mapsto \phi(t) := \int_a^b \left(\int_c^t f(x, y) \, dy \right) dx = \int_a^b \tilde{g}(x, t) \, dx,$$

dass

$$\phi'(t) = \int_a^b \partial_t \tilde{g}(x, t) \, dx = \int_a^b f(x, t) \, dx,$$

und somit wegen $\phi(c) = 0$

$$\int_c^d \left(\int_a^b f(x, y) \, dx \right) dy = \int_c^d \phi'(y) \, dy = \phi(d) = \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx. \quad \square$$

Das Integral stetiger Funktionen über Rechtecke ist also unabhängig von der Integrationsreihenfolge. Daher können wir Integrale stetiger Funktionen über Rechtecke sinnvoll durch

$$\int_Q f(x, y) \, d(x, y) := \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy \quad \left(= \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx \right)$$

definieren. Normalerweise definiert man aber zuerst $\int_Q f(x, y) \, d(x, y)$ und zeigt dann, dass man dieses Integral durch Doppelintegrale in beliebiger Reihenfolge berechnen kann. Diese Aussage heißt dann Satz von Fubini.

Man möchte nun natürlich auch über andere Bereiche als Rechtecke integrieren, z.B. über Kreise. Eine naheliegende Definition ist die Folgende.

10.6 Definition. Normalbereiche

Eine Teilmenge $A \subset \mathbb{R}^2$ heißt **x -Normalbereich**, wenn es ein Intervall $[a, b] \subset \mathbb{R}$ und stetige Funktionen $g, h : [a, b] \rightarrow \mathbb{R}$ gibt, die auf (a, b) stetig differenzierbar sind und

$$A = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, g(x) \leq y \leq h(x)\}$$

erfüllen.

Entsprechend heißt $A \subset \mathbb{R}^2$ **y -Normalbereich**, wenn es ein Intervall $[c, d] \subset \mathbb{R}$ und stetige Funktionen $u, w : [c, d] \rightarrow \mathbb{R}$ gibt, die auf (c, d) stetig differenzierbar sind und

$$A = \{(x, y) \in \mathbb{R}^2 \mid c \leq y \leq d, u(y) \leq x \leq w(y)\}$$

erfüllen.

Ist A sowohl ein x -Normalbereich als auch ein y -Normalbereich, so nennen wir A einen **Normalbereich**. ◇

10.7 Beispiel. Rechtecke der Form $[a, b] \times [c, d]$ sind offenbar sowohl x - als auch y -Normalbereiche. Gleiches gilt für Kreisscheiben, denn

$$\begin{aligned} \overline{B_1(0)} &= \left\{ (x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq 1, -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2} \right\} \\ &= \left\{ (x, y) \in \mathbb{R}^2 \mid -1 \leq y \leq 1, -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2} \right\}. \end{aligned}$$

10.8 Definition. Das Integral stetiger Funktionen über Normalbereiche

Sei $A \subset \mathbb{R}^2$ und $f : A \rightarrow \mathbb{R}$ stetig. Falls A ein x -Normalbereich ist, so setzt man

$$\int_A f(x, y) \, d(x, y) := \int_a^b \left(\int_{g(x)}^{h(x)} f(x, y) \, dy \right) dx.$$

Falls A ein y -Normalbereich ist, so setzt man

$$\int_A f(x, y) \, d(x, y) := \int_c^d \left(\int_{u(y)}^{w(y)} f(x, y) \, dx \right) dy.$$

Für die Wohldefiniertheit von $\int_A f(x, y) \, d(x, y)$ muss man natürlich noch zeigen, dass für Mengen A die sowohl ein x - als auch ein y -Normalbereich sind, die beiden Werte übereinstimmen. Das sparen wir uns allerdings, da es ein nicht unerheblicher Aufwand wäre, der durch die allgemeine Integrationstheorie obsolet werden wird. ◇

10.9 Beispiel. Das Integral der Funktion $f(x, y) = 1$ über den Normalbereich $\overline{B_1(0)}$ ist

$$\begin{aligned} \int_{\overline{B_1(0)}} f(x, y) \, d(x, y) &= \int_{-1}^1 \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy \right) dx = \int_{-1}^1 2\sqrt{1-x^2} dx \\ &\stackrel{\text{Part. Int.}}{=} \left(x\sqrt{1-x^2} + \arcsin(x) \right) \Big|_{-1}^1 = \left(\frac{\pi}{2} + \frac{\pi}{2} \right) = \pi, \end{aligned}$$

also genau die Fläche von $\overline{B_1(0)}$.

10.10 Definition. Fläche eines Normalbereichs

Sei $A \subset \mathbb{R}^2$ ein Normalbereich. Dann heißt

$$F(A) := \int_A d(x, y)$$

die Fläche von A .

Nun kann man alle Definitionen und Sätze dieses Kapitels in natürlicher Weise vom \mathbb{R}^2 auf den \mathbb{R}^n verallgemeinern. Wir verzichten wieder auf die Details und bemerken lediglich, dass man dann im Allgemeinen vom Volumen eines Normalbereichs und nicht mehr von der Fläche spricht.

Um einige weitere Aspekte der Integration auf dem \mathbb{R}^n zu verdeutlichen und die wichtige Transformationsformel zu formulieren, definieren wir nun noch das Integral stetiger Funktionen mit kompaktem Träger über den ganzen \mathbb{R}^n .

10.11 Definition. Der Träger einer Funktion und der Raum $C_0(G)$

Sei $G \subset \mathbb{R}^n$ und $f : G \rightarrow \mathbb{R}$. Dann heißt die Menge

$$\text{supp}(f) := \overline{\{x \in G \mid f(x) \neq 0\}}^G \subset G$$

der **Träger** von f (englisch "support"). Beachte, dass hier der Abschluss in G und nicht in \mathbb{R}^n gemeint ist. Allgemeiner definiert man den Träger einer Funktion $f : X \rightarrow V$ auf einem topologischen Raum X mit Werten in einem Vektorraum V durch

$$\text{supp}(f) := \overline{\{x \in X \mid f(x) \neq 0\}}.$$

Es bezeichnet

$$C_0(X) := \{f \in C(X) \mid \text{supp}(f) \text{ ist kompakt}\}$$

den Raum der **stetigen Funktionen mit kompaktem Träger**. ◇

10.12 Proposition. Das Integral stetiger Funktionen mit kompaktem Träger

Sei $f \in C_0(\mathbb{R}^n)$. Dann hat

$$\int_Q f \, d(x_1, \dots, x_n) = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\dots \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) \, dx_n \right) \dots \right) dx_2 \right) dx_1$$

für jeden achsenparallelen Quader $Q = \prod_{j=1}^n [a_j, b_j] \subset \mathbb{R}^n$ mit $\text{supp}(f) \subset Q$ denselben Wert, welchen wir mit

$$\int_{\mathbb{R}^n} f \, dx^n$$

bezeichnen. Weiterhin schreiben wir

$$\int_{\mathbb{R}^n} f \, dx^n = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \left(\dots \left(\int_{-\infty}^{\infty} f(x_1, \dots, x_n) \, dx_n \right) \dots \right) dx_2 \right) dx_1,$$

wobei aufgrund der Kompaktheit des Trägers von f die Integralgrenzen immer hinreichend groß aber endlich gewählt werden können.

Beweis. Sei $\text{supp}(f) \neq \emptyset$ (für $f \equiv 0$ ist die Aussage offenbar wahr) und seien Q und \tilde{Q} zwei Quader, die $\text{supp}(f)$ enthalten. Dann ist

$$P := Q \cap \tilde{Q} =: \prod_{j=1}^n ([a_j, b_j] \cap [\tilde{a}_j, \tilde{b}_j]) = \prod_{j=1}^n [\max\{a_j, \tilde{a}_j\}, \min\{b_j, \tilde{b}_j\}] =: \prod_{j=1}^n [\bar{a}_j, \bar{b}_j]$$

ebenfalls ein Quader mit $\text{supp}(f) \subset P$. Daher gilt

$$\begin{aligned} \int_Q f \, d(x_1, \dots, x_n) &= \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\dots \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) \, dx_n \right) \dots \right) dx_2 \right) dx_1 \\ &= \int_{\bar{a}_1}^{\bar{b}_1} \left(\int_{\bar{a}_2}^{\bar{b}_2} \left(\dots \left(\int_{\bar{a}_n}^{\bar{b}_n} f(x_1, \dots, x_n) \, dx_n \right) \dots \right) dx_2 \right) dx_1 \\ &= \int_P f \, d(x_1, \dots, x_n) \end{aligned}$$

und analog auch $\int_{\tilde{Q}} f \, d(x_1, \dots, x_n) = \int_P f \, d(x_1, \dots, x_n)$. □

10.13 Definition. Bewegungen des \mathbb{R}^n

Zu jedem Vektor $w \in \mathbb{R}^n$ und jeder orthogonalen Matrix $A \in O(n)$ definieren wir die Abbildungen

$$T_w : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto T_w(x) := x + w \quad (\text{Translation um } w)$$

und

$$R_A : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto R_A(x) := Ax \quad (\text{Rotation + Spiegelung mit } A).$$

Kombinationen $B = R_A \circ T_w$ aus Translationen, Rotationen und Spiegelungen heißen **Bewegungen**. ◇

Da Bewegungen des \mathbb{R}^n Isometrien sind und zumindest anschaulich das Volumen von Mengen nicht ändern, erwarten wir, dass sich auch das Integral einer Funktion nicht ändert, wenn wir sie lediglich verschieben oder drehen. Dass dies so ist, besagt folgende Proposition.

10.14 Proposition. Invarianz des Integrals unter Bewegungen

Sei $f \in C_0(\mathbb{R}^n)$, $w \in \mathbb{R}^n$ und $A \in O(n)$. Dann gilt

$$\int_{\mathbb{R}^n} f \, dx^n = \int_{\mathbb{R}^n} f \circ T_w \, dx^n = \int_{\mathbb{R}^n} f \circ R_A \, dx^n.$$

Das Integral ist also invariant unter Translationen und Rotationen.

Beweis. Die Invarianz unter Translationen folgt direkt aus der Definition des Integrals. Die Invarianz unter orthogonalen Transformationen folgt aus der Transformationsformel für lineare Abbildungen, Satz 10.15, da für $A \in O(n)$ gilt, dass $|\det(A)| = 1$. □

Allgemeine lineare Abbildungen

$$L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto L_A(x) := Ax,$$

gegeben durch Matrizen $A \in M(n, \mathbb{R})$, ändern allerdings das Volumen von Mengen: Der Einheitswürfel $E := [0, 1]^n$ hat per Definition das Volumen $V(E) = 1$. Die Bildmenge $L_A(E) := \{Ax \mid x \in E\}$ ist das Parallelotop welches von den n Spaltenvektoren a_1, \dots, a_n der Matrix A aufgespannt wird, also

$$L_A(E) = \left\{ \sum_{j=1}^n \alpha_j a_j \mid \alpha_j \in [0, 1] \text{ für } j = 1, \dots, n \right\}.$$

Wie Sie aus der linearen Algebra wissen, ist das (signierte) Volumen dieses Parallelotops genau durch die Determinante von A gegeben, also

$$V(L_A(E)) = \det(A).$$

Es ist also nicht überraschend, dass sich auch das Integral (und somit auch Volumina von Normalbereichen) entsprechend transformiert.

10.15 Proposition. Transformationsformel für lineare Abbildungen

Sei $f \in C_0(\mathbb{R}^n)$ und $A \in M(n, \mathbb{R})$ invertierbar. Dann ist $f \circ L_A \in C_0(\mathbb{R}^n)$ und es gilt

$$\int_{\mathbb{R}^n} f \, dx^n = |\det(A)| \int_{\mathbb{R}^n} f \circ L_A \, dx^n.$$

Beweis. Um die Notation übersichtlich zu halten, beschränken wir uns auf den Fall \mathbb{R}^2 . Es wird aber offensichtlich sein, wie man das Argument auf höhere Dimensionen überträgt.

Sei zunächst $R \in M(2, \mathbb{R})$ eine obere Dreiecksmatrix, also $R = \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}$. Dann liefert die Substitutionsregel aus Analysis 1 sofort

$$\begin{aligned} \int_{\mathbb{R}^2} f \circ L_R \, d(x, y) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(r_{11}x + r_{12}y, r_{22}y) \, dx \right) dy \\ &= \frac{1}{r_{11}} \int_{-\infty}^{\infty} \left(\int_{-\operatorname{sgn}(r_{11})\infty}^{\operatorname{sgn}(r_{11})\infty} f(\tilde{x} + r_{12}y, r_{22}y) \, d\tilde{x} \right) dy \\ &\stackrel{(*)}{=} \frac{1}{r_{11}} \int_{-\infty}^{\infty} \left(\int_{-\operatorname{sgn}(r_{11})\infty}^{\operatorname{sgn}(r_{11})\infty} f(\tilde{x}, r_{22}y) \, d\tilde{x} \right) dy \\ &= \frac{1}{r_{11}r_{22}} \int_{-\operatorname{sgn}(r_{22})\infty}^{\operatorname{sgn}(r_{22})\infty} \left(\int_{-\operatorname{sgn}(r_{11})\infty}^{\operatorname{sgn}(r_{11})\infty} f(\tilde{x}, \tilde{y}) \, d\tilde{x} \right) d\tilde{y} \\ &= \frac{1}{|r_{11}||r_{22}|} \int_{\mathbb{R}^2} f \, d(x, y) = |\det(R)|^{-1} \int_{\mathbb{R}^2} f \, d(x, y), \end{aligned}$$

wobei wir in (*) die Translationsinvarianz des Integrals verwendet haben. Völlig analog zeigt man die Behauptung auch für untere Dreiecksmatrizen. Nun wissen wir aus der Linearen Algebra (oder der Numerik), dass sich jede invertierbare Matrix A in ein Produkt aus einer oberen Dreiecksmatrix R und einer unteren Dreiecksmatrix L zerlegen lässt, also $A = LR$, die sogenannte LR -Zerlegung. Dies folgt aus dem Gaußschen Algorithmus, wenn man die Umformungen in L "speichert". Damit gilt aber

$$\begin{aligned} \int_{\mathbb{R}^2} f \circ L_A \, d(x, y) &= \int_{\mathbb{R}^2} f \circ L_L \circ L_R \, d(x, y) = |\det(R)|^{-1} \int_{\mathbb{R}^2} f \circ L_L \, d(x, y) \\ &= |\det(L)|^{-1} |\det(R)|^{-1} \int_{\mathbb{R}^2} f \, d(x, y) = |\det(A)|^{-1} \int_{\mathbb{R}^2} f \, d(x, y), \end{aligned}$$

wobei wir im letzten Schritt den Produktsatz für Determinanten verwendet haben,

$$\det(L) \det(R) = \det(LR) = \det(A). \quad \square$$

Wir kommen schließlich zu der Frage, wie sich das Integral unter allgemeinen Diffeomorphismen Φ des \mathbb{R}^n bzw. von Teilmengen des \mathbb{R}^n transformiert, also insbesondere unter Koordinatentransformationen. Die folgende Transformationsformel ist die Verallgemeinerung der Substitutionsregel auf Integrale im \mathbb{R}^n . Sie besagt, dass nun die Volumenänderung vom Ort x abhängt und lokal durch die Determinante der Linearisierung $D\Phi|_x$ von Φ gegeben ist.

10.16 Satz. Transformationsformel

Sei $\Phi : U \rightarrow G$ ein Diffeomorphismus offener Teilmengen U und G des \mathbb{R}^n . Sei weiterhin $f \in C_0(G)$. Dann ist $f \circ \Phi \in C_0(U)$ und es gilt

$$\int_{\mathbb{R}^n} f \, dx^n = \int_{\mathbb{R}^n} (f \circ \Phi)(x) \cdot |\det(D\Phi|_x)| \, dx^n.$$

Bemerkung: Der Beweis dieser Aussage wäre zwar mit den vorhandenen Mitteln möglich, aber doch recht aufwendig. Außerdem wird dieser Satz in der Integrationstheorievorlesung für eine deutlich allgemeinere Klasse von sogenannten integrierbaren Funktionen bewiesen. Daher verzichten wir an dieser Stelle auf einen Beweis. Heuristisch ist es aber zumindest einleuchtend, dass die sogenannte Funktional- oder auch Jacobideterminante $|\det(D\Phi|_x)|$ der lokalen Linearisierung $D\Phi|_x$ von Φ die lokale Änderung des Volumens unter der Transformation Φ beschreibt.

10.17 Bemerkung. Im Fall $n = 1$ erhalten wir die Substitutionsregel aus Analysis 1:

$$\int_{\Phi(a)}^{\Phi(b)} f(x) dx = \int_a^b (f \circ \Phi)(x) \Phi'(x) dx .$$

Den Betrag von $\Phi'(x)$ bekommt man, sobald man auf der linken Seite ebenfalls “von kleiner nach größer” integriert. Denn es ist $\Phi(a) > \Phi(b)$ genau dann, wenn $\Phi'(x) < 0$ für alle $x \in [a, b]$. \diamond

10.18 Bemerkung. Die Klasse der im \mathbb{R}^n integrierbaren Funktionen ist weitaus größer als $C_0(\mathbb{R}^n)$ und die Transformationsformel gilt dementsprechend auch viel allgemeiner. Wir werden daher in den Anwendungsbeispielen nicht mehr darauf achten, dass $f \in C_0(\mathbb{R}^n)$ gilt.

10.19 Beispiel. Polarkoordinaten

Die Abbildung

$$\Phi : (0, \infty) \times (-\pi, \pi) \rightarrow \mathbb{R}^2 \setminus \{x_1 \leq 0\}, \quad (r, \varphi) \mapsto \Phi(r, \varphi) = (r \cos \varphi, r \sin \varphi) =: (x(r, \varphi), y(r, \varphi))$$

ist ein Diffeomorphismus und versieht den $\mathbb{R}^2 \setminus \{x_1 \leq 0\}$ mit Polarkoordinaten. Die Funktionaldeterminante von Φ ist

$$\det D\Phi|_{(r,\varphi)} = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r .$$

Damit lautet die Transformationsformel in diesem Fall

$$\int_{\mathbb{R}^2} f dx^2 = \int_{\mathbb{R}^2} (f \circ \Phi)(r, \varphi) \cdot |\det(D\Phi|_{(r,\varphi)})| d(r, \varphi) = \int_{-\pi}^{\pi} \int_0^{\infty} (f \circ \Phi)(r, \varphi) r dr d\varphi .$$

Man nennt $(f \circ \Phi)$ die “Funktion f in Polarkoordinaten” und $r dr d\varphi$ das “Flächenelement” in Polarkoordinaten.

Wir können jetzt beispielsweise die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad x \mapsto e^{-\|x\|^2},$$

integrieren. Denn $(f \circ \Phi)(r, \varphi) = e^{-r^2}$ und

$$\int_{\mathbb{R}^2} e^{-\|x\|^2} dx^2 = \int_{-\pi}^{\pi} \int_0^{\infty} e^{-r^2} r dr d\varphi = 2\pi \int_0^{\infty} r e^{-r^2} dr = 2\pi \left[-\frac{1}{2} e^{-r^2} \right]_0^{\infty} = \pi .$$

Da andererseits

$$\int_{\mathbb{R}^2} e^{-\|x\|^2} dx^2 = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{-x_1^2} dx_1 \right) e^{-x_2^2} dx_2 = \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 ,$$

haben wir auch

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

gezeigt. \diamond

Schließlich stellt sich noch die Frage, wie der Hauptsatz der Integral- und Differentialrechnung im Höherdimensionalen aussieht. Kann man das Integral über die Ableitung einer Funktion durch Auswerten der Funktion selbst auf dem Rand des Integrationsgebiets bestimmen? Dazu findet man zunächst einen ganzen Zoo von sogenannten Integralsätzen (Gauß, Stokes, Green), die sich wiederum als Spezialfälle des allgemeinen Satzes von Stokes für sogenannte Differentialformen auffassen lassen. Das ist dann Thema der Integrationstheorievorlesung bzw. der Differentialgeometrie. Wir stellen hier nur einen Spezialfall vor, nämlich den Satz von Gauß im \mathbb{R}^2 .

Der Rand eines “schönen” Gebiets im \mathbb{R}^2 ist eine Kurve. Daher diskutieren wir zunächst einige Aspekte des Integrierens entlang von Kurven etwas allgemeiner.

10.20 Definition. Kurvenintegrale

Sei $\gamma : [a, b] \rightarrow G \subset \mathbb{R}^n$ eine stetig differenzierbare Kurve und $v : G \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Dann heißt

$$\int_{\gamma} v \cdot ds := \int_a^b \langle v(\gamma(t)), \gamma'(t) \rangle dt$$

das Kurvenintegral von v entlang γ . Hier deutet der Punkt in der Notation $v \cdot ds$ das Skalarprodukt zwischen dem Vektorfeld v und dem "vektoriellen Längenelement" ds an. \diamond

10.21 Proposition. Invarianz unter Reparametrisierungen

Sei $\gamma : [a, b] \rightarrow G \subset \mathbb{R}^n$ eine stetig differenzierbare Kurve, $\phi : [c, d] \rightarrow [a, b]$ eine stetig differenzierbare Bijektion und $\tilde{\gamma} : [c, d] \rightarrow G$, $\tilde{\gamma} := \gamma \circ \phi$, also eine sogenannte Reparametrisierung von γ . Dann ist $\text{sgn}(\phi') \in \{-1, +1\}$ konstant und für jedes stetige Vektorfeld $v : G \rightarrow \mathbb{R}^n$ gilt, dass

$$\int_{\gamma} v \cdot ds = \text{sgn}(\phi') \int_{\tilde{\gamma}} v \cdot ds.$$

Beweis. Da ϕ bijektiv ist, ist $\phi'(t) \neq 0$ für alle $t \in [c, d]$. Also hat die stetige Funktion ϕ' ein festes Vorzeichen. Der Rest ist nun wieder einfach die Substitutionsregel aus Analysis 1,

$$\int_c^d \langle v(\tilde{\gamma}(t)), \tilde{\gamma}'(t) \rangle dt = \int_c^d \langle v(\gamma(\phi(t))), \gamma'(\phi(t)) \rangle \phi'(t) dt = \text{sgn}(\phi') \int_a^b \langle v(\gamma(s)), \gamma'(s) \rangle ds. \quad \square$$

Man kann die Invarianz des Integrals unter Reparametrisierung auch verstehen, indem man das Kurvenintegral als ein Integral über "geometrische" Objekte auffasst.

10.22 Definition. Tangential- und Normalenfeld an eine Kurve

Sei $\gamma : [a, b] \rightarrow G \subset \mathbb{R}^n$ eine stetig differenzierbare Kurve und zusätzlich injektiv und $\gamma'(t) \neq 0$ für alle $t \in [a, b]$, also ohne Selbstdurchschneidungen und mit nichtverschwindender Geschwindigkeit. Das Bild von γ , also $\gamma([a, b]) \subset \mathbb{R}^n$, nennen wir auch die **Spur** $\text{Sp}(\gamma)$ der Kurve. Dann heißt

$$\tau_{\gamma} : \text{Sp}(\gamma) \rightarrow \mathbb{R}^n, \quad x \mapsto \tau_{\gamma}(x) := \frac{\gamma'(\gamma^{-1}(x))}{\|\gamma'(\gamma^{-1}(x))\|}$$

das **normierte Tangentialvektorfeld** an $\text{Sp}(\gamma)$. Es ist offenbar bis auf ein globales Vorzeichen invariant unter Reparametrisierungen und hängt somit nur von der Richtung ab, in der $\text{Sp}(\gamma)$ durchlaufen wird.

Im Fall $n = 2$ erhält man ein **normiertes Normalenvektorfeld** an $\text{Sp}(\gamma)$ durch Drehung von τ ,

$$\nu_{\gamma} : \text{Sp}(\gamma) \rightarrow \mathbb{R}^2, \quad x \mapsto \nu_{\gamma}(x) := (\tau_2(x), -\tau_1(x)).$$

Man kann nun das Kurvenintegral von v entlang von γ auch als ein Integral der Funktion $\text{Sp}(\gamma) \rightarrow \mathbb{R}^n$, $x \mapsto \langle v(x), \tau(x) \rangle$ über $\text{Sp}(\gamma)$ verstehen, also

$$\int_{\text{Sp}(\gamma)} \langle v(x), \tau_{\gamma}(x) \rangle := \int_{\gamma} v \cdot ds = \int_a^b \langle v_{\gamma}(\gamma(t)), \tau_{\gamma}(\gamma(t)) \rangle \|\gamma'(t)\| dt$$

Dabei hängt der Wert wie zuvor gesehen nur von der Richtung ab, in der $\text{Sp}(\gamma)$ durchlaufen wird, nicht aber von der konkreten Parametrisierung.

10.23 Definition. Der Fluss durch eine Kurve im \mathbb{R}^2

Sei $\gamma : [a, b] \rightarrow G \subset \mathbb{R}^2$ eine stetig differenzierbare Kurve und $v : G \rightarrow \mathbb{R}^2$ ein stetiges Vektorfeld. Dann heißt

$$\int_{\text{Sp}(\gamma)} \langle v, \nu_{\gamma} \rangle := \int_a^b \langle v(\gamma(t)), \nu_{\gamma}(\gamma(t)) \rangle \|\gamma'(t)\| dt$$

der **Fluss von v durch die Kurve $\text{Sp}(\gamma)$** . \diamond

10 Integration im \mathbb{R}^n : Eine kurze Einführung

Wir kommen nun zur Formulierung des Gaußschen Satzes im \mathbb{R}^2 . Sei dazu $A \subset \mathbb{R}^2$ ein Normalbereich,

$$A = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, g(x) \leq y \leq h(x)\} = \{(x, y) \in \mathbb{R}^2 \mid c \leq y \leq d, u(y) \leq x \leq w(y)\}$$

für stetig differenzierbare Funktionen f, g, u, w . Der Rand von A lässt sich also als stetig differenzierbare Kurve

$$\gamma_{A,x} : [0, 2(b-a)] \rightarrow \mathbb{R}^2, \quad \gamma_{A,x}(t) := \begin{cases} (t+a, g(t+a)) & \text{falls } t \in [0, (b-a)] \\ (2b-a-t, h(2b-a-t)) & \text{falls } t \in [(b-a), 2(b-a)] \end{cases}$$

oder alternativ

$$\gamma_{A,y} : [0, 2(d-c)] \rightarrow \mathbb{R}^2, \quad \gamma_{A,y}(t) := \begin{cases} (w(t+c), t+c) & \text{falls } t \in [0, (d-c)] \\ (u(2d-c-t), 2d-c-t) & \text{falls } t \in [(d-c), 2(d-c)] \end{cases}$$

parametrisieren. Beide Kurven durchlaufen ∂A im positiven Sinne, also gegen den Uhrzeigersinn. Damit zeigt das zugehörige Einheitsnormalenfeld $\nu_{\gamma_{A,x}} = \nu_{\gamma_{A,y}}$ nach außen.

Zu einem stetigen Vektorfeld $v : A \subset G \rightarrow \mathbb{R}^2$ können wir den Fluss durch den Rand von A durch

$$\int_{\partial A} \langle v, \nu_{\gamma_{A,x}} \rangle = \int_0^{2(b-a)} \langle v(\gamma_{A,x}(t)), \nu_{\gamma_{A,x}}(\gamma_{A,x}(t)) \rangle \|\gamma'_{A,x}(t)\| dt$$

definieren. Nun hängt ist der Wert dieses Wegintegrals aber nach Proposition 10.21 invariant unter orientierungserhaltenden Reparametrisierungen, hat also z.B. für $\gamma_{A,y}$ denselben Wert. Daher schreiben wir im folgenden einfach

$$\int_{\partial A} \langle v, \nu \rangle$$

für den nach außen gerichteten Fluss durch ∂A .

Der Gaußsche Satz besagt nun, dass das Integral der "Ableitung" von v (genauer der Divergenz von v) über den Normalbereich A mit dem Integral von v über den Rand ∂A von A (genauer der Fluss von v durch ∂A) gegeben ist.

10.24 Satz. von Gauß

Sei $A \subset \mathbb{R}^2$ ein Normalbereich, $v : A \subset G \rightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_A \operatorname{div}(v) d(x, y) = \int_{\partial A} \langle v, \nu \rangle$$

Beweis. Es ist zu zeigen, dass

$$\int_A (\partial_x v_x(x, y) + \partial_y v_y(x, y)) d(x, y) = \int_\gamma (v_x(\gamma(t))\nu_x(\gamma(t)) + v_y(\gamma(t))\nu_y(\gamma(t))) \|\gamma'(t)\| dt,$$

wobei γ eine beliebige stetig differenzierbare Kurve ist, welche ∂A im positiven Sinne durchläuft. Für den ersten Summanden im Integral finden wir

$$\begin{aligned} \int_A \partial_x v_x(x, y) d(x, y) &= \int_c^d \left(\int_{u(y)}^{w(y)} \partial_x v_x(x, y) dx \right) dy \\ &= \int_c^d (v_x(w(y), y) - v_x(u(y), y)) dy \\ &= \int_0^{d-c} v_x(\gamma_{A,y}(t)) dt - \int_{d-c}^{2(d-c)} v_x(\gamma_{A,y}(t)) dt \\ &= \int_0^{2(d-c)} v_x(\gamma_{A,y}(t)) \frac{(\gamma'_{A,y}(t))_y}{\|\gamma'_{A,y}(t)\|} \|\gamma'_{A,y}(t)\| dt \\ &= \int_0^{2(d-c)} v_x(\gamma_{A,y}(t)) \nu_x(\gamma_{A,y}(t)) \|\gamma'_{A,y}(t)\| dt, \end{aligned}$$

und für den zweiten folgt mit $\gamma_{A,x}$ ganz analog

$$\int_A \partial_y v_y(x, y) \, d(x, y) = \int_0^{2(b-a)} v_y(\gamma_{A,x}(t)) \nu_y(\gamma_{A,x}(t)) \|\gamma'_{A,x}(t)\| \, dt. \quad \square$$

10.25 Bemerkung. Der Satz von Gauß in einer Dimension ist genau der Hauptsatz der Integral- und Differentialrechnung. In diesem Fall ist $A = [a, b]$ ein Intervall und $\partial A = \{a, b\}$ besteht nur aus den beiden Randpunkten a und b . Für eine stetig differenzierbare Funktion $v : \mathbb{R} \rightarrow \mathbb{R}$ ist $\operatorname{div}(v) = v'$ und der Fluss von v durch ∂A ist einfach die Auswertung von v an den Randpunkten, wobei der nach außen gerichtete Einheitsnormalenvektor am linken Randpunkt a gleich -1 ist. Also hat man insgesamt

$$\int_A \operatorname{div}(v) \, dx := \int_a^b v' \, dx \stackrel{\text{H.S.}}{=} v(b) - v(a) =: \int_{\partial A} \langle v, \nu \rangle.$$

10.26 Beispiele. Wir rechnen die beiden Seiten der Gleichung im Satz von Gauß für einige einfache Beispiele nochmals explizit aus. Sei $A = \overline{B_1(0)}$ die Einheitskreisscheibe und ∂A durch $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$, $t \mapsto \gamma(t) := (\cos(t), \sin(t))$ parametrisiert. Dann ist $\gamma'(t) = (-\sin(t), \cos(t))$, $\|\gamma'(t)\| = 1$ und, wenig überraschend, $\nu_\gamma(t) = (\cos(t), \sin(t)) = \gamma(t)$.

(a) Für $v(x, y) = (1, 0)$ ergibt sich einerseits $\operatorname{div}(v) \equiv 0$ und andererseits auch

$$\int_{\partial A} \langle v, \nu \rangle = \int_0^{2\pi} \langle v(\gamma(t)), \nu_\gamma(\gamma(t)) \rangle \|\gamma'(t)\| \, dt = \int_0^{2\pi} \cos(t) \, dt = 0.$$

Dass der Fluss von v durch ∂A verschwindet, ist auch geometrisch klar: Was links reinfließt, fließt rechts wieder raus.

(b) Für $v(x, y) = (x, y)$ ergibt sich einerseits $\operatorname{div}(v) \equiv 2$, also $\int_A \operatorname{div}(v) = 2\pi$, und andererseits

$$\int_{\partial A} \langle v, \nu \rangle = \int_0^{2\pi} \langle v(\gamma(t)), \nu_\gamma(\gamma(t)) \rangle \|\gamma'(t)\| \, dt = \int_0^{2\pi} (\cos(t)^2 + \sin(t)^2) \, dt = 2\pi.$$

(c) Für $v(x, y) = (-y, x)$ ergibt sich einerseits $\operatorname{div}(v) \equiv 0$ und andererseits

$$\int_{\partial A} \langle v, \nu \rangle = \int_0^{2\pi} \langle v(\gamma(t)), \nu_\gamma(\gamma(t)) \rangle \|\gamma'(t)\| \, dt = \int_0^{2\pi} 0 \, dt = 0.$$

Dass der Fluss von v durch ∂A verschwindet, ist auch geometrisch klar: Das Vektorfeld v ist überall tangential an ∂A .