# ANALYSIS II - SOMMERSEMESTER 2011

STEFAN FRIEDL

## Inhaltsverzeichnis

## Literatur

Zum Erlernen des Stoffes und zur Bearbeitung der Übungsaufgaben reicht das Skriptum. Nachdem sich das Skriptum an Forster: Analysis II orientiert, kann es allerdings hilfreich sein, dieses Buch begleitend zu studieren.

# 1. Metrische Räume

In diesem Semester wollen wir uns hauptsächlich mit Funktionen im $\mathbb{R}^n$ beschäftigen. Für die Diskussion von stetigen Funktionen bietet es sich allerdings an unser Blickfeld etwas zu erweitern und eine Verallgemeinerung des $\mathbb{R}^n$ zu betrachten, nämlich die metrischen Räume.

## 1.1. **Definition von metrischen Räumen.**

*Definition.* Ein *metrischer Raum* ist ein Paar $(X, d)$ bestehend aus einer Menge $X$ und einer Metrik $d$ auf $X$, d.h. einer Abbildung

$$\begin{array}{rcl} X \times X & \to & \mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \,|\, x \geq 0\}, \\ (x,y) & \mapsto & d(x,y), \end{array}$$

mit folgenden Eigenschaften:

(1) $d(x,y) = 0$ genau dann, wenn $x = y$,

(2) für alle $x, y \in X$ gilt

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

(3) für alle $x, y, z \in X$ gilt

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

Wir nennen $d(x,y)$ manchmal den *Abstand von $x$ zu $y$.*

*Beispiel.* (1) $X = \mathbb{R}^n$ und

$$d((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2},$$

dies ist die euklidische Metrik auf $\mathbb{R}^n$,

(2) $X = \mathbb{R}^n$ und

$$d((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \max \left\{ 1, \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} \right\},$$

dies ist die euklidische Metrik auf $\mathbb{R}^n$,

(3) $X$ eine beliebige Menge, z.B. $X = \{1, 2, 3, \ldots, k\}$, $X = \mathbb{R}^n$, und [1]

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

(4) $X = \mathbb{R}$ und $d(x,y) := |x - y|$. (Das ist nur Beispiel (1) mit $n = 1$.)

---

[1] Überlegen Sie sich, warum das eigentlich eine Metrik ist.

*Beispiel.* Es sei $(X, d)$ ein metrischer Raum und $A \subset X$ eine Teilmenge. Wir definieren

$$\begin{array}{rcl} d_A : A \times A & \to & \mathbb{R} \\ (a,b) & \mapsto & d(a,b), \end{array}$$

dann ist $(A, d_A)$ ebenfalls ein metrischer Raum. Wir nennen $d_A$ die *induzierte Metrik*.

## 1.2. **Normierte Vektorräume.**

*Definition.* Ein *normierter Vektorraum* ist ein Paar $(V, \|-\|)$ bestehend aus einem Vektorraum $V$ über $\mathbb{R}$ und einer *Norm auf* $V$, d.h. einer Abbildung

$$\begin{array}{rcl} \| - \| : V & \to & \mathbb{R}_{\geq 0} \\ x & \mapsto & \|x\|, \end{array}$$

mit folgenden Eigenschaften:

(1) $\|x\| = 0$ genau dann, wenn $x = 0$,
(2) $\|\lambda x\| = |\lambda| \cdot \|x\|$ für alle $\lambda \in \mathbb{R}$ und $x \in V$,
(3) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in V$.

*Beispiel.* (1) $V = \mathbb{R}^n$ und

$$\|(x_1, \dots, x_n)\| := \sqrt{\sum_{k=1}^{n} x_k^2} \qquad \text{(Euklidische Norm)},$$

bilden einen normierten Vektorraum.

(2) $V = \mathbb{R}^n$ und

$$\|(x_1, \dots, x_n)\| := \sum_{k=1}^{n} |x_k| \qquad \text{(Maximumsnorm)}$$

bilden einen normierten Vektorraum. [2]

(3)

$$V = M(n \times m, \mathbb{R}) = \text{Vektorraum der } m \times n\text{–Matrizeun}$$

und

$$\|(a_{ij}\| = \max_{ij}\{|a_{ij}|\}$$

bilden einen normierten Vektorraum.

[2]Diese Norm hat übrigens den Spitznamen ‘Manhattan-Norm’. Warum? Wenn sie zu Fuss von der Ecke ‘25th Street 6th Avenue’ zu ‘59th Street, 3rd Avenue’ gehen wollen, wie messen Sie die Distanz bzgl. der Koordinaten $(25, 6)$ und $(59, 3)$?

(4) Es sei $D \subset \mathbb{R}$, dann bilden
$$V = \text{alle beschränkten Funktionen } D \to \mathbb{R},$$
und
$$\|f\| := \sup\{|f(x)| \mid x \in D\}$$
einen normierten Vektorraum. (Siehe Forster Seite 5 für einen Beweis der Dreiecksungleichung.)

(5) Es seien $a, b \in \mathbb{R}, a < b$, dann bilden
$$C([a,b]) = \text{Menge aller stetigen Funktionen } [a,b] \to \mathbb{R},$$
und
$$\|f\| := \sqrt{\int_a^b f(x)^2\,dx}$$
einen normierten Vektorraum. Der Nachweis, dass dies ein normierter Vektorraum ist, ist allerdings nicht ganz trivial.

Folgender Satz folgt sofort aus den Definitionen:

**Satz 1.1.** *Sei* $(V, \|-\|)$ *ein normierter Vektorraum. Wir definieren*
$$d(v,w) := \|v - w\|,$$
*dann ist* $(V, d)$ *ein metrischer Raum.*

Im Folgenden betrachten wir $\mathbb{R}^n$ durchgehend (außer wir sagen explizit etwas anderes) als normierten Vektorraum bzgl. der Euklidischen Norm, insbesondere betrachten wir $\mathbb{R}^n$ als metrischen Raum bzgl. der Euklidischen Metrik.

### 1.3. **Offene Mengen in metrischen Räumen.**

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

(1) Es sei $x \in X$ und $r > 0$, dann heißt
$$B_r(x) := \{y \in X \mid d(x,y) < r\}$$
die *offene Kugel* um $x$ mit Radius $r$.

(2) Eine Teilmenge $A \subset X$ heißt *offen*, wenn es zu jedem $a \in A$ ein $r > 0$ gibt, so dass
$$B_r(a) \subset A.$$

(3) Es sei $x \in X$. Eine *Umgebung von* $x$ ist eine Teilmenge von $X$, so dass es ein $r > 0$ gibt, mit der Eigenschaft, dass $B_r(x) \subset U$.

*Beispiel.* (1) Wenn $X = \mathbb{R}$ mit $d(x,y) = |x - y|$, dann ist $B_r(x) = (x - r, x + r)$. Es folgt leicht, dass ein Intervall $(a,b) \subset \mathbb{R}$ eine offene Teilmenge von $\mathbb{R}$ ist.

(2) Ein Intervall von der Form $(a, \infty)$ oder $(-\infty, a)$ ist offen.
(3) Ein Intervall von der Form $[a, b)$ ist nicht offen, weil es für $x = a$ kein geeignetes $r > 0$ gibt.
(4) Es sei $(X, d)$ ein metrischer Raum. Dann ist $X$ eine offene Menge und die leere Menge ist ebenfalls eine offene Menge.
(5) Folgende Mengen sind beispielsweise Umgebungen von $1 \in \mathbb{R}$: $(1-\varepsilon, 1+\varepsilon), [0, 2), (-\infty, 3]$ und $\mathbb{R}$.

*Bemerkung.* Es sei $x \in X$. Eine Teilmenge $U \subset X$ ist eine Umgebung von $x$, genau dann, wenn es eine offene Teilmenge $V$ von $X$ gibt, mit der Eigenschaft, dass $x \in V$ und $V \subset U$. [3]

**Lemma 1.2.** *Es sei $(X, d)$ ein metrischer Raum, $x \in X$ und $r > 0$. Dann ist $B_r(x) \subset X$ eine offene Teilmenge.*

*Beweis.* Es sei $y \in B_r(x)$. Wir setzen $s := d(x, y)$. Per Definition von $B_r(x)$ gilt $s < r$. Wir setzen nun $t := r - s$. Nachdem $t > 0$ genügt es folgende Behauptung zu beweisen:

*Behauptung.*
$$B_t(y) \subset B_r(x).$$

Sei also $z \in B_t(y)$. Mithilfe der Dreiecksungleichung erhalten wir
$$d(x, z) \leq d(x, y) + d(y, z) < s + t = r.$$
□

**Satz 1.3.** (Hausdorffsches Trennungsaxiom) *Es sei $X$ ein metrischer Raum* [4] *und $x, y \in X$ zwei verschiedene Punkt. Dann existieren Umgebungen $U$ von $x$ und $V$ von $y$, welche disjunkt sind, d.h. $U \cap V = \emptyset$.*

*Beweis.* Wir setzen $\varepsilon := \frac{1}{2} d(x, y)$. Dann sind $U := B_\varepsilon(x)$ und $V := B_\varepsilon(y)$ Umgebungen von $x$ bzw. $y$. Wir müssen nun zeigen, dass diese disjunkt sind.

Nehmen wir an, es gibt einen Punkt $z \in U \cap V$. Dann folgt aus der Dreiecksungleichung, dass
$$2\varepsilon = d(x, y) \leq d(x, z) + d(z, y) < \varepsilon + \varepsilon = 2\varepsilon,$$
also $2\varepsilon < 2\varepsilon$. Wir erhalten also einen Widerspruch. □

**Satz 1.4.** *Es sei $X$ ein metrischer Raum.*

[3] Diese Aussage folgt leicht aus den Definitionen, wenn man diese mal verdaut hat.

[4] Wenn nichts anderes angegeben ist, dann wird die Metrik immer mit $d$ bezeichnet, d.h. 'Sei $X$ ein metrischer Raum' ist kurz für 'Sei $(X, d)$ ein metrischer Raum'.

(1) *Wenn $U_1, \dots, U_k$ offene Teilmengen von $X$ sind, dann ist $U_1 \cap U_2 \cap \dots \cap U_k$ ebenfalls offen.*
(2) *Es sei $U_i, i \in I$ eine Familie* [5] *von offenen Teilmengen, dann ist $\cup_{i\in I} U_i$ ebenfalls offen.*

*Bemerkung.* (1) Anders ausgedrückt, der Satz besagt, dass die Schnittmenge von *endlich* vielen offenen Mengen wieder offen ist, und dass die Vereinigung von beliebig vielen offenen Mengen wieder offen ist.
(2) Man beachte, dass die Schnittmenge von unendlich vielen offenen Mengen i.a. nicht offen ist. Beispielsweise sind die Mengen $U_n := (-\frac{1}{n}, \frac{1}{n}) \subset \mathbb{R}, n \in \mathbb{N}$ offen, aber die Schnittmenge $\cap_{n=1}^{\infty} U_n$ besteht nur aus $\{0\}$, aber $\{0\} \subset \mathbb{R}$ ist nicht offen.

*Beweis.* (1) Es sei $x \in U_1 \cap U_2 \cap \dots \cap U_k$. Nach Voraussetzung existieren $\varepsilon_i > 0, i = 1, \dots, k$, so dass $B_{\varepsilon_i}(x) \subset U_i$. Dann folgt, dass

$$B_{\min(\varepsilon_1,\dots,\varepsilon_k)}(x) \in U_1 \cap U_2 \cap \dots \cap U_k.$$

(2) Es sei $x \in \cup_{i\in I} U_i$. Dann existiert insbesondere ein $i \in I$, so dass $x \in U_i$. Nach Voraussetzung existiert ein $\varepsilon > 0$, so dass $B_\varepsilon(x) \subset U_i$. Dann gilt aber auch, dass

$$B_\varepsilon(x) \subset U_i \subset \cup_{i\in I} U_i.$$

□

*Definition.* Es sei $X$ ein metrischer Raum. Wir sagen $U \subset X$ ist *abgeschlossen*, wenn $X \setminus U$ offen ist.

*Beispiel.* (1) Es seien $a, b \in \mathbb{R}$, $a \le b$, dann ist $[a, b] \subset \mathbb{R}$ abgeschlossen, denn das Komplement

$$\mathbb{R} \setminus [a, b] = \underbrace{(-\infty, a)}_{\text{offen}} \cup \underbrace{(b, \infty)}_{\text{offen}}$$

ist die Vereinigung zweier offener Mengen, also ebenfalls offen.
(2) Es sei $X$ ein metrischer Raum, dann sind $X$ und die leere Menge wie oben gesehen offen. Ihre Komplemente sind per Definition abgeschlossen, also sind die leere Menge und $X$ auch abgeschlossen.
(3) Es sei $X$ ein metrischer Raum und $x \in X$. In den Übungen werden wir sehen, dass $\{x\} \subset X$ abgeschlossen ist.

[5] auf gut Deutsch heißt das Folgendes: $I$ ist eine Menge (z.B. $\{1, \dots, k\}$ oder $\mathbb{N}$, aber $I$ kann auch z.B. überabzählbar sein), und jedem Element $i \in I$ ist eine Menge $U_i$ zugeordnet. Siehe auch: `http://de.wikipedia.org/wiki/Familie_(Mathematik)`

**Lemma 1.5.** (1) *Die Vereinigung von endlich vielen abgeschlossen Teilmengen ist abgeschlossen.*
(2) *Der Durchschnitt von beliebig vielen abgeschlossenen Teilmengen ist wiederum abgeschlossen.*

*Beweis.* Es seien also $A_1, \dots, A_k \subset X$ abgeschlossene Mengen. Dann gilt

$$X \setminus (A_1 \cup \dots \cup A_k) = \cap_{i=1}^{k} X \setminus A_i,$$

aber nach Voraussetzung sind die Mengen $X \setminus A_i$ offen, und nach Satz 1.4 (2) ist die Schnittmenge von endlich vielen offenen Mengen wieder offen. Der Beweis der zweiten Aussage des Lemmas verläuft nach dem gleichen Schema. □

*Definition.* Es sei $Y$ eine Teilmenge eines metrischen Raumes $X$. Wir sagen $x \in X$ ist *Randpunkt von* $Y$, falls jede Umgebung von $x$ sowohl einen Punkt von $Y$ als auch einen Punkt von $X \setminus Y$ enthält. Die Menge aller Randpunkte von $Y$ heißt der *Rand* von $Y$ und wird mit $\partial Y$ bezeichnet.

*Bemerkung.* Wie in der Bemerkung vor Lemma 1.2 kann man Folgendes zeigen: Ein Punkt $x \in X$ ist ein Randpunkt von $Y$, genau dann, wenn jede Kugel $B_r(x)$ um $x$ sowohl einen Punkt von $Y$ als auch einen Punkt von $X \setminus Y$ enthält.

*Beispiel.* (1) Es seien $a, b \in R$, $a < b$. Die Randpunkte der Teilmengen

$$(a, b), (a, b] \text{ und } [a, b) \subset \mathbb{R}$$

sind $a$ und $b$. [6]
(2) Der Rand der Kugel

$$\{v \in \mathbb{R}^n \mid \|v\| \leq 1\},$$

ist die Sphäre

$$S^{n-1} := \{v \in \mathbb{R}^n \mid \|v\| = 1\}.$$

**Satz 1.6.** *Es sei $X$ ein metrischer Raum und $Y \subset X$. Dann gilt:*
(1) $Y \setminus \partial Y$ *ist offen,*
(2) $Y \cup \partial Y$ *ist abgeschlossen,*
(3) $\partial Y$ *ist abgeschlossen.*

[6] In der Tat, jedes Intervall $(a-r, a+r)$ enthält einen Punkt von $(a, b)$ und einen Punkt außerhalb von $(a, b)$, also ist $a$ ein Randpunkt von $(a, b)$. Genauso zeigt man, dass $b$ ein Randpunkt von $(a, b)$. Wenn $x \in (a, b)$, dann gibt es ein $r > 0$, so dass $(x - r, x + r) \subset (a, b)$, also ist $x \in (a, b)$ kein Randpunkt von $(a, b)$, und genauso zeigt man, dass ein Punkt $x \in (-\infty, a) \cup (b, \infty)$ kein Randpunkt von $(a, b)$ ist.

*Beweis.* (1) Es sei also $y \in Y \setminus \partial Y$. Wir müssen zeigen, dass es ein $\varepsilon > 0$ gibt, so dass $B_\varepsilon(y) \subset Y \setminus \partial Y$.

Nachdem jede Umgebung von $y$ insbesondere $y$ enthält, und nachdem $y \notin \partial Y$ existiert eine Umgebung $U$ von $y$, welche keinen Punkt von $X \setminus Y$ enthält. Per Definition von Umgebung gibt es nun ein $\varepsilon > 0$, so dass $B_\varepsilon(y) \subset U \subset Y$.

Wir müssen jetzt noch zeigen, dass $B_\varepsilon(y) \cap \partial Y = \emptyset$. Sei $z \in B_\varepsilon(y)$. Nachdem $B_\varepsilon(y)$ offen ist (Lemma 1.2), ist $B_\varepsilon(y)$ eine offene Umgebung von $z$. Aber die Umgebung $B_\varepsilon(y)$ von $z$ enthält keinen Punkt von $X \setminus Y$, also kann $z$ kein Randpunkt sein.

(2) Wir setzen $Y' := X \setminus Y$. Aus der Definition folgt sofort, dass $\partial Y = \partial Y'$. Es folgt, dass

$$Y \cup \partial Y = (X \setminus Y') \cup \partial Y' = X \setminus \underbrace{(Y' \setminus \partial Y')}_{\text{offen nach (a)}}$$

abgeschlossen ist.

(3) Wir setzen $Y' := X \setminus Y$. Aus (a) folgt, dass

$$X \setminus \partial Y = (Y \cup Y') \setminus \partial Y = (Y \setminus \partial Y) \cup (Y' \setminus \partial Y')$$

offen ist, insbesondere ist $\partial Y \subset X$ abgeschlossen.

□

1.4. **Topologische Räume.** Wir betrachten jetzt die topologischen Räume, welche man als eine Verallgemeinerung des Begriffes des metrischen Raumes auffassen kann. Topologische Räume werden in dieser Vorlesung eine eher untergeordnete Rolle spielen, aber für die meisten höheren Vorlesungen in der reinen Mathematik (z.B. Funktionalanalyis, Topologie, algebraische Geometrie, Zahlentheorie) ist dies ein sehr wichtiger Begriff.

*Definition.* Ein *topologischer Raum* ist ein Paar $(X, \mathcal{T})$ wobei $X$ eine Menge ist und $\mathcal{T}$ eine *Topologie auf* $X$, d.h. $\mathcal{T}$ ist eine Menge von Teilmengen von $X$ mit folgenden Eigenschaften:

(1)
$$\emptyset, X \in \mathcal{T},$$

(2)
$$U_1, \ldots, U_k \in \mathcal{T} \Rightarrow U_1 \cap \cdots \cap U_k \in \mathcal{T},$$

(3)
$$U_i \in \mathcal{T}, i \in I \Rightarrow \cup_{i \in I} U_i \in \mathcal{T}.$$

*Beispiel.*
(1) Wenn $(X, d)$ ein metrischer Raum ist, dann ist
$$\mathcal{T} := \{A \subset X \mid A \text{ offen in } (X, d)\}$$
eine Topologie auf $X$.
(2) Sei $X$ eine Menge, dann ist $\mathcal{T} = \{\emptyset, X\}$ eine Topologie auf $X$.
(3) Sei $X$ eine Menge und $\mathcal{T}$ die Menge aller Untermengen von $X$. Dann ist $\mathcal{T}$ eine Topologie auf $X$.
(4) Sei $X = \mathbb{R}$ und es sei $\mathcal{T}$ wie folgt definiert:
$$U \in \mathcal{T} \quad :\Leftrightarrow \quad \begin{array}{l} \text{es gibt } x_1, \ldots, x_n \in \mathbb{R}, \text{ so dass } U = \mathbb{R} \setminus \{x_1, \ldots, x_n\} \\ \text{oder } U = \emptyset. \end{array}$$
Dann ist $\mathcal{T}$ eine Topologie auf $X = \mathbb{R}$.

*Definition.* Sei $(X, \mathcal{T})$ ein topologischer Raum.
(1) Die Mengen in $\mathcal{T}$ werden *offen* genannt.
(2) Sei $(X, \mathcal{T})$ ein topologischer Raum und $x \in X$. Eine Teilmenge $U \subset X$ heißt *Umgebung von* $x$, falls es eine offene Menge $V$ gibt, so dass $x \in V \subset U$.

*Definition.* Ein topologischer Raum heißt *Hausdorff*, wenn es zu zwei verschiedenen Punkten $x \neq y$ disjunkte Umgebungen $U$ von $x$ und $V$ von $y$ gibt.

Es folgt leicht aus den Definitionen, dass ein topologischer Raum *Hausdorff* ist, genau dann, wenn es zu zwei verschiedenen Punkten $x \neq y$ offene Teilmengen $U$ und $V$ gibt, so dass $x \in U, y \in V$ und $U \cap V = \emptyset$.

*Beispiel.* Wir überprüfen ob die obigen Beispiele Hausdorff sind:
(1) Es folgt aus Satz 1.3, dass der einem metrischen Raum zugeordnete topologische Raum Hausdorff ist.
(2) Wenn $X$ mindestens zwei Elemente hat, dann ist $(X, \mathcal{T})$ nicht Hausdorff.
(3) Diese Topologie ist Hausdorff, in der Tat, für gegebene $x \neq y$ können wir $U = \{x\}$ und $V = \{y\}$ wählen.
(4) Diese Topologie ist nicht Hausdorff, denn die Schnittmenge von zwei nichtleeren offenen Mengen in $(X, \mathcal{T})$ ist nichtleer, es gibt also keine disjunkten nichtleeren offene Teilmengen in $(X, \mathcal{T})$.

*Bemerkung.* Es gibt viele verschiedene Normen auf $\mathbb{R}^n$, z.B.
$$\begin{array}{rcl} \|(x_1, \ldots, x_n)\|_1 & = & \sqrt{\sum_{k=1}^n x_i^2}, \\ \|(x_1, \ldots, x_n)\|_2 & = & \sum_{k=1}^n |x_i|, \\ \|(x_1, \ldots, x_n)\|_3 & = & \max\{|x_1|, \ldots, |x_n|\}. \end{array}$$

Man kann zeigen, dass alle Normen auf $\mathbb{R}^n$ die gleiche Topologie definieren, d.h. wenn eine Menge offen ist bzgl. einer Norm auf $\mathbb{R}^n$, dann ist sie auch offen bzgl. jeder anderen Norm auf $\mathbb{R}^n$. [7]

## 2. Konvergente Folgen und stetige Abbildungen

### 2.1. Konvergenz von Folgen.

*Definition.* Es sei $X$ ein metrischer Raum und $(x_k)_{k\in\mathbb{N}}$ eine Folge von Punkten aus $X$. Wir sagen $(x_k)_{k\in\mathbb{N}}$ *konvergiert gegen* $x \in X$ wenn

$$\underset{\varepsilon>0}{\forall}\ \underset{K\in\mathbb{N}}{\exists}\ \underset{k\geq K}{\forall}\ d(x_k, x) < \varepsilon.$$

Wie üblich schreiben wir in diesem Fall $\lim_{k\to\infty} x_k = x$.

Der Beweis des folgenden Lemmas ist eine Übungsaufgabe.

**Lemma 2.1.** *Es sei $X$ ein metrischer Raum und $(x_k)_{k\in\mathbb{N}}$ eine Folge von Punkten aus $X$. Dann gilt $\lim_{k\to\infty} x_k = x$ genau dann, wenn*

$$\underset{\substack{\text{Umgebung}\, U\\ \text{von}\, x}}{\forall}\ \underset{K\in\mathbb{N}}{\exists}\ \underset{k\geq K}{\forall}\ x_k \in U.$$

*Bemerkung.*
1. Das Lemma besagt unter anderem, dass es bei der Konvergenz von Folgen nicht auf die Metrik ankommt, sondern auf die offenen Mengen. Z.B. haben wir gesehen, dass es auf dem $\mathbb{R}^n$ verschiedene Normen gibt, welche die gleiche Topologie definieren, es folgt, dass eine Folge von Punkten in $\mathbb{R}^n$ genau dann bzgl. einer Norm konvergiert wenn sie bzgl. jeder anderen Norm konvergiert.
2. Die Formulierung für Konvergenz, welche im Lemma gegeben ist, hat den Vorteil, dass sie auch Sinn für topologische Räume macht.

*Konvention.* Wenn $x$ ein Punkt in $\mathbb{R}^n$ ist, dann bezeichnen wir die Koordinaten von $x$ mit $(x_1, \dots, x_n)$.

**Satz 2.2.** *Es sei $(x_k)_{k\in\mathbb{N}}$ eine Folge von Punkten in $\mathbb{R}^n$.* [8] *Dann gilt*

$$\lim_{k\to\infty} x_k = y \ \Leftrightarrow\ \lim_{k\to\infty} x_{ki} = y_i \text{ für alle } i = 1, \dots, n.$$

[7] Diese Aussage gilt ganz allgemeiner für Normen auf endlich dimensionalen Vektorräumen (nachdem diese alle isomorph zu einem $\mathbb{R}^n$ sind), aber die Aussage ist falsch für Normen auf unendlich dimensionalen Vektorräumen.

[8] Wie oben erwähnt, wenn nichts anderes angegeben ist, dann betrachten wir $\mathbb{R}^n$ mit der euklidischen Norm.

Dieser Satz erlaubt es uns die Konvergenz von Folgen im $\mathbb{R}^n$ mithilfe der Methoden aus der Analysis I zu studieren.

*Beweis.* Nehmen wir zuerst an, dass $\lim_{k\to\infty} x_k = y$. Es gilt also:

$$\underset{\varepsilon>0}{\forall}\ \underset{K\in\mathbb{N}}{\exists}\ \underset{k\geq K}{\forall}\ \|x_k - y\| < \varepsilon.$$

Es sei nun $i \in \{1,\dots,n\}$. Wir müssen zeigen:

$$\underset{\eta>0}{\forall}\ \underset{L\in\mathbb{N}}{\exists}\ \underset{l\geq L}{\forall}\ |x_{li} - y_i| < \varepsilon.$$

Sei also $\eta > 0$. Wenn wir $\varepsilon = \eta$ setzen, erhalten wir ein $K \in \mathbb{N}$, so dass für alle $k \geq K$ gilt, $\|x_k - y\| < \varepsilon$. Dann gilt aber für alle $l \geq L := K$, dass

$$|x_{li} - y_i| \leq \sqrt{\sum_{j=1}^{n}(x_{lj} - y_j)^2} = \|x_k - y\| < \varepsilon = \eta.$$

(Die erste Ungleichung ist einfach Algebra, und hat mit der speziellen Situation nichts zu tun.)

Nehmen wir nun an, dass für alle $i = 1,\dots,n$ gilt: $\lim_{k\to\infty} x_{ki} = y_i$. Wir wollen zeigen, dass $\lim_{k\to\infty} x_k = y$, d.h. wir wollen zeigen, dass

$$\underset{\varepsilon>0}{\forall}\ \underset{K\in\mathbb{N}}{\exists}\ \underset{k\geq K}{\forall}\ \|x_k - y\| < \varepsilon.$$

Sei also $\varepsilon > 0$. Zur Erinnerung, es gilt:

$$\|x_k - y\| = \sqrt{\sum_{i=1}^{n}(x_{ki} - y_i)^2}.$$

Sei nun $i \in \{1,\dots,n\}$. Nachdem $\lim_{k\to\infty} x_{ki} = y_i$ existiert insbesondere ein $L_i \in \mathbb{N}$, so dass für alle $k \geq L_i$ gilt: $|x_{ki} - y_i| < \eta = \frac{\varepsilon}{\sqrt{n}}$. Wenn wir nun $K := \max\{L_1,\dots,L_n\}$ setzen, dann gilt für alle $k \geq K$, dass

$$\|x_k - y\| = \sqrt{\sum_{i=1}^{n}(x_{ki} - y_i)^2} \leq \sqrt{\sum_{j=1}^{n}\left(\frac{\varepsilon}{\sqrt{n}}\right)^2} = \varepsilon.$$

(Die Ungleichung folgt, weil für $k \geq K$ gilt $k \geq K \geq L_i$, also $(x_{ki} - y_)^2 \leq \eta^2 = \frac{\varepsilon}{\sqrt{n}}$.) ∎

**Satz 2.3.** *Es sei $X$ ein metrischer Raum und $A \subset X$. Dann gilt:*

$$A \text{ ist abgeschlossen} \Leftrightarrow \begin{array}{l} \text{für jede Folge } (a_k)_{k\in\mathbb{N}} \text{ von Punkten in } A, \\ \text{welche in } X \text{ konvergiert} \\ \text{gilt: } \lim_{k\to\infty} a_k \in A. \end{array}.$$

*Beispiel.* Betrachten wir $U = (-1,1) \subset \mathbb{R}$, dann ist $a_k = 1 - \frac{1}{k}$ eine Folge von Punkten in $U$, welche konvergiert, aber der Grenzwert $\lim_{k\to\infty} a_k = 1$ liegt nicht mehr in $U$. Es folgt aus dem obigen Satz, dass $U \subset \mathbb{R}$ nicht abgeschlossen ist.

*Beweis.* "$\Rightarrow$" Nehmen wir zuerst an, dass $A$ abgeschlossen ist. Es sei $(a_k)_{k\in\mathbb{N}}$ eine Folge von Punkten in $A$, welche gegen $x \in X$ konvergiert. Wir müssen zeigen, dass $x \in A$.

Nehmen wir an, dass dies nicht der Fall ist, d.h. nehmen wir an, dass $x \in X \setminus A$. Aus der Definition von 'abgeschlossen' folgt, dass $X \setminus A$ offen ist, d.h. es existiert ein $\varepsilon > 0$, so dass $B_\varepsilon(x) \subset X \setminus A$. Anders ausgedrückt, es existiert ein $\varepsilon > 0$, so dass $B_\varepsilon(x) \cap A = \emptyset$.

Andererseits folgt aus $\lim_{k\to\infty} a_k = x$, dass es ein $K \in \mathbb{N}$ gibt, so dass $d(a_k, x) < \varepsilon$ für alle $k \geq K$. Anders ausgedrückt, es existiert ein $K \in \mathbb{N}$, so dass $a_k \in B_\varepsilon(x)$ für alle $k \geq K$. Dies ist nun aber ein Widerspruch nachdem wir oben gesagt hatten, dass $B_\varepsilon(x) \cap A = \emptyset$, d.h. $B_\varepsilon(x)$ kann keine Folgenglieder $x_k$ enthalten.

"$\Leftarrow$" Für den Beweis der Umkehrung siehe Forster, Seite 15. $\square$

*Definition.* Es sei $X$ ein metrischer Raum. Eine Folge $(x_k)_{k\in\mathbb{N}}$ heißt *Cauchy-Folge*, wenn gilt:
Zu jedem $\varepsilon > 0$ existiert ein $K \in \mathbb{N}$, so dass

$$d(x_k, x_l) < \varepsilon \text{ für alle } k, l \geq K.$$

Genau wie in Analysis I, Satz 4.1, zeigt man, dass jede konvergente Folge auch eine Cauchyfolge ist.

*Definition.* Ein metrischer Raum heißt *vollständig*, wenn jede Cauchyfolge konvergiert.

*Beispiel.*
1. Aus Analysis I wissen wir, dass $\mathbb{R}$ vollständig ist.
2. Mithilfe von Satz 2.2 kann man leicht zeigen, dass $\mathbb{R}^n$ ebenfalls vollständig ist. [9]
3. Der metrische Raum
$$\begin{aligned} X &:= \mathbb{R} \setminus \{0\}, \\ d(x,y) &:= |x-y|, \end{aligned}$$
ist nicht vollständig, denn $x_k = \frac{1}{k}$ ist eine Cauchyfolge, welche in $X$ nicht konvergiert.

[9] Genauer gesagt: Es sei $(x_k)_{k\in\mathbb{N}}$ eine Cauchy–Folge in $\mathbb{R}^n$, dann kann man ähnlich wie in Satz 2.2 zeigen, dass für jedes $i = 1, \ldots, n$ die Folge von reellen Zahlen $(x_{ki})_{k\in\mathbb{N}}$ eine Cauchyfolge ist. Nachdem $\mathbb{R}$ vollständig ist, existiert ein $y_i \in \mathbb{R}$, so dass $\lim_{k\to\infty} x_{ki} = y_i$. Dann folgt aber aus Satz 2.2, dass $\lim_{k\to\infty} x_k = (y_1, \ldots, y_n)$.

(4) Der metrische Raum

$$\begin{array}{rcl} X & := & \mathbb{Q}, \\ d(x,y) & := & |x-y|, \end{array}$$

ist nicht vollständig, nachdem es z.B. Folgen von rationalen Zahlen gibt, welche gegen die irrationale Zahl $\sqrt{2}$ konvergieren.

(5) Wir betrachten den metrischen Raum

$$\begin{array}{rcl} X & = & \text{alle stetigen Funktionen } [-1,1] \to \mathbb{R}, \\ d(f,g) & = & \int_{-1}^{1} |f(x)-g(x)|\,dx. \end{array}$$

Betrachten wir die Folge von Funktionen

$$\begin{array}{rcl} f_n : [-1,1] & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} 0, & \text{wenn } x \in [-1,0), \\ nx, & \text{wenn } x \in [0,\frac{1}{n}), \\ 1, & \text{wenn } x \in [\frac{1}{n},1]. \end{cases} \end{array}$$

Dann gilt für $n \geq m$, dass

$$d(f_n,f_m) = \int_{-1}^{1} |f_n(x)-f_m(x)|\,dx \leq \frac{1}{m}.$$

(Dies folgt, weil $f_n$ und $f_m$ die gleichen Werte außerhalb vom Intervall $[0,\frac{1}{m}]$ haben, und auf dem Intervall $[0,\frac{1}{m}]$ gilt $|f_n(x) - f_m(x)| \leq 1$.) Es folgt, dass $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, aber diese konvergiert nicht in $X$, denn die Grenzfunktion

$$f(x) = \begin{cases} 0, & \text{wenn } x \in [-1,0], \\ 1, & \text{wenn } x \in (0,1]. \end{cases}$$

ist nicht stetig, liegt also nicht in $X$.

### 2.2. **Stetige Abbildungen.**

*Definition.* Es sei $f\colon X \to Y$ eine Abbildung zwischen zwei metrischen Räumen. Wir definieren:

$$f \text{ ist stetig im Punkt } x_0 \;:\Leftrightarrow \underset{\varepsilon>0}{\forall}\; \underset{\delta>0}{\exists}\; \underset{x\in B_\delta(x_0)}{\forall} d(f(x),f(x_0)) < \varepsilon.$$

Wir sagen $f$ ist stetig, wenn $f$ in jedem Punkt stetig ist.

*Beispiel.* (1) Für Abbildungen $\mathbb{R} \to \mathbb{R}$ erhalten wir den üblichen Stetigkeitsbegriff.

(2) Es sei $X$ eine Menge und $d$ die Metrik

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

Es sei $(Y, e)$ ein weiterer metrischer Raum, dann ist jede Abbildung $X \to Y$ stetig. In der Tat, für jedes $\varepsilon > 0$ hat $\delta = \frac{1}{2}$ die gewünschte Eigenschaft.

Wie in Analysis I, Lemma 6.1 zeigt man, dass die Identitätsabbildung $f(x) = x$ stetig ist, und dass eine konstante Abbildung ebenfalls stetig ist. Zudem gilt (siehe den Beweis von Analysis I, Satz 6.2):

**Lemma 2.4.** *Es sei $X$ ein metrischer Raum und es seien $f, g\colon X \to \mathbb{R}$ zwei stetige Funktionen* [10]. *Dann gilt:*

(1)
$$\begin{aligned} f + g\colon X &\to \mathbb{R} \\ x &\mapsto f(x) + g(x) \end{aligned}$$
*ist stetig,*

(2)
$$\begin{aligned} f \cdot g\colon X &\to \mathbb{R} \\ x &\mapsto f(x)g(x) \end{aligned}$$
*ist stetig,*

(3) *falls $g(x) \neq 0$ für alle $x \in X$, dann ist*
$$\begin{aligned} \tfrac{f}{g}: X &\to \mathbb{R} \\ x &\mapsto \tfrac{f(x)}{g(x)} \end{aligned}$$
*ebenfalls stetig.*

Für die Stetigkeit gibt es verschiedene äquivalente Definitionen:

**Satz 2.5.** *Es sei $f\colon X \to Y$ eine Abbildung zwischen zwei metrischen Räumen. Dann gilt:*

*$f$ ist stetig*
$\Leftrightarrow$ *für jede konvergente Folge $(x_k)_{k\in\mathbb{N}}$ in $X$ gilt:* $\lim_{k\to\infty} f(x_k) = f(\lim_{k\to\infty} x_k)$
$\Leftrightarrow$ *für jede offene Menge $U$ in $Y$ gilt: $f^{-1}(U)$ ist offen*
$\Leftrightarrow$ *für jede abgeschlossene Menge $A$ in $Y$ gilt: $f^{-1}(A)$ ist abgeschlossen.*

*Bemerkung.*
(1) Der Beweis der Äquivalenz der ersten beiden Aussagen erfolgt fast genauso wie in Satz 6.5 und Satz 6.6 in Analysis I. (Siehe auch Forster Analysis II, Seite 19). Der Beweis der Äquivalenz der ersten und der dritten Aussage ist eine Übungsaufgabe. Die Äquivalenz der dritten und der vierten Aussage ist ebenfalls eine Übungsaufgabe.
(2) Die Definition der Stetigkeit über die offenen Mengen macht auch Sinn für Abbildungen zwischen topologischen Räumen.

[10] Wir verwenden normalerweise das Wort 'Funktion' für Abbildungen nach $\mathbb{R}$ oder $\mathbb{R}^n$

**Korollar 2.6.** *Es seien $f\colon X \to Y$ und $g\colon Y \to Z$ stetige Abbildungen zwischen metrischen Räumen. Dann ist die Verknüpfung $g \circ f\colon X \to Z$ ebenfalls stetig.*

*Beweis.* Aus Satz 2.5 folgt, dass es genügt zu zeigen, dass $(g \circ f)^{-1}(U) \subset X$ offen ist für jede offene Menge $U \subset Z$.

Sei also $U \subset Z$ offen, dann gilt

$$(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)),$$

aber $g^{-1}(U) \subset Y$ ist offen, weil $g$ stetig und $U$ offen, und dann ist $f^{-1}(g^{-1}(U))$ offen, weil $f$ stetig und $g^{-1}(U)$ offen ist. $\square$

Man kann das Korollar auch genau wie in Analysis I, Satz 6.4 zeigen, man muss nur in dem Beweis von Satz 6.4 die Absolutbeträge durch die Metriken auf $X, Y$ und $Z$ ersetzen.

*Beispiel.* Wir betrachten

$$\begin{array}{rcl} M(n,\mathbb{R}) & = & \text{alle reellen } n \times n\text{–Matrizen,} \\ GL(n,\mathbb{R}) & = & \text{alle invertierbaren } n \times n\text{–Matrizen.} \end{array}$$

Die Abbildung

$$\begin{array}{rcl} \det : M(n,\mathbb{R}) & \to & \mathbb{R} \\ A & \mapsto & \det(A) \end{array}$$

ist stetig (dies folgt leicht aus der obigen Diskussion, nachdem die Determinante ein Polynom in den Einträgen der Matrix ist). Man beachte, dass $GL(n,\mathbb{R}) = \{A \subset M(n,\mathbb{R}) \mid \det(A) \neq 0\}$. Nachdem det stetig ist, und nachdem $\mathbb{R} \setminus \{0\} \subset \mathbb{R}$ offen ist, folgt aus Satz 2.5, dass

$$GL(n,\mathbb{R}) = \{A \subset M(n,\mathbb{R}) \mid \det(A) \neq 0\} = \det^{-1}(\mathbb{R} \setminus \{0\}) \subset M(n,\mathbb{R})$$

ebenfalls offen ist.

Aus Satz 2.2 und Satz 2.5 folgt:

**Lemma 2.7.** *Es sei $X$ ein metrischer Raum und*

$$f = (f_1, \dots, f_n) : X \to \mathbb{R}^n$$

*eine Funktion. Dann gilt:*

$$f\colon X \to \mathbb{R}^n \text{ ist stetig} \Leftrightarrow f_i : X \to \mathbb{R} \text{ ist stetig für alle } i = 1, \dots, n.$$

**Lemma 2.8.** *Es sei $A$ eine reelle $m \times n$–Matrix. Dann ist die Abbildung*

$$\begin{array}{rcl} \mathbb{R}^n & \to & \mathbb{R}^m \\ v & \mapsto & Av \end{array}$$

*stetig.*

*Beweis.* Sei also $A = (a_{ij})$ eine Matrix. Aus Lemma 2.7 folgt, dass es genügt zu zeigen, dass für jedes $i$ die Abbildung

$$\begin{array}{rcl} \mathbb{R}^n & \to & \mathbb{R} \\ (v_1, \ldots, v_n) & \mapsto & a_{i1}v_1 + \cdots + a_{in}v_n, \end{array}$$

stetig ist. In Übungsblatt 3, Aufgabe 1 zeigen wir, dass die Abbildungen $(v_1, \ldots, v_n) \mapsto v_j$ für jedes $j$ stetig sind. Es folgt nun aus Lemma 2.4, dass die Abbildungen

$$(v_1, \ldots, v_n) \mapsto a_{i1}v_1 + \cdots + a_{in}v_n$$

wie gewünscht stetig sind. □

*Definition.* Es sei $f: X \to Y$ eine Abbildung zwischen metrischen Räumen. Wir sagen $f$ ist ein *Homöomorphismus*, wenn gilt:

(1) $f$ ist stetig,
(2) $f$ ist bijektiv,
(3) die Umkehrabbildung $f^{-1}$ ist stetig.

Wenn es einen solchen Homöormorphismus zwischen $X$ und $Y$ gibt, dann sagen wir, dass $X$ und $Y$ *homöomorph* sind.

*Beispiel.* (1) $\mathbb{R}$ und $(-1, 1)$ sind homöomorph, in der Tat, die Abbildung

$$\begin{array}{rcl} f: \mathbb{R} & \to & (-1, 1) \\ x & \mapsto & \frac{2}{\pi} \arctan(x) \end{array}$$

ist stetig und bijektiv, und die Umkehrabbildung $x \mapsto \tan(\frac{\pi}{2})$ ist ebenfalls stetig. D.h. $f$ ist in der Tat ein Homöomorphismus.

(2) Die Abbildung

$$\begin{array}{rcl} f: X := [0, 2\pi) & \to & Y := \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\} \\ t & \mapsto & (\cos(t), \sin(t)) \end{array}$$

ist stetig und bijektiv, aber die Umkehrabbildung $f^{-1}$ ist im Punkt $(1, 0)$ nicht stetig (siehe Forster Seite 21). [11]

*Bemerkung.* Man kann relativ leicht zeigen, dass

$$X := [0, 2\pi) \text{ und } Y := \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\}$$

nicht homöomorph sind. Im Allgemeinen ist es jedoch eine schwierige Aufgabe zu zeigen, dass gewisse Räume nicht homöomorph sind. Beispielsweise gibt es für jedes $n$ eine stetige surjektive Abbildung $\mathbb{R} \to \mathbb{R}^m$ ('space filling curve'), aber man benötigt die Methoden der 'algebraischen Topologie' um zu zeigen, dass für $n \neq m$ die metrischen Räume $\mathbb{R}^n$ und $\mathbb{R}^m$ nicht homöomorph sind.

[11] Wenn nichts anderes angegeben ist, dann betrachten wir $\mathbb{R}^n$ immer mit dem euklidischen Abstand, und jede Teilmenge von $\mathbb{R}^n$ mit der induzierten Metrik.

2.3. **Stetigkeit von linearen Abbildungen.** In Lemma 2.8 hatten wir gesehen, dass alle lineare Abbildungen $f\colon \mathbb{R}^n \to \mathbb{R}^m$ stetig sind. Es folgt, dass jede linear Abbildung zwischen endlich dimensionalen normierten Vektorräumen ebenfalls stetig ist. Wir werden nun sehen, dass lineare Abbildungen zwischen unendlich dimensionalen Vektorräumen nicht notwendigerweise stetig sind.

**Satz 2.9.** *Es sei $\phi : V \to W$ eine lineare Abbildung zwischen normierten Vektorräumen. Dann sind die folgenden Aussagen äquivalent:*

(1) *$\phi$ ist stetig,*
(2) *$\phi$ ist stetig im Nullpunkt,*
(3) *es existiert ein $C \geq 0$, so dass $\|\phi(x)\| \leq C\|x\|$ für alle $x \in V$.*

*Beweis.* (1) $\Rightarrow$ (2) ist trivial.
(2) $\Rightarrow$ (3): Nehmen wir an $\phi$ ist stetig im Nullpunkt. Dann existiert insbesondere zu $\varepsilon = 1$ ein $\delta > 0$, so dass

$$\|\phi(z)\| < 1 \text{ für alle } z \in V \text{ mit } \|z\| < \delta. \tag{2.1}$$

Wir behaupten, dass $C := \frac{2}{\delta}$ die gewünschte Eigenschaft besitzt. Sei also $x \in V \setminus \{0\}$ beliebig. Wir setzen $\lambda := C \cdot \|x\|$. Es gilt

$$\|\frac{1}{\lambda} \cdot x\| = \frac{1}{C \cdot \|x\|}\|x\| = \frac{1}{C} = \frac{\delta}{2} < \delta.$$

Es folgt also aus (2.1), angewandt auf $\frac{1}{\lambda}x$, dass

$$\|\phi(\frac{1}{\lambda} \cdot x)\| < 1.$$

Nachdem $\phi$ eine lineare Abbildung ist, folgt auch, dass

$$\|\phi(x)\| = \|\phi(\lambda \cdot \frac{1}{\lambda} \cdot x)\| = C \cdot \|x\| \cdot \|\phi(\frac{1}{\lambda} \cdot x)\| < C \cdot \|x\| \cdot 1 = C \cdot \|x\|.$$

(3) $\Rightarrow$ (1): Es sei $x_0 \in V$. Dann gilt für alle $x \in V$, dass

$$\|\phi(x) - \phi(x_0)\| = \|\phi(x - x_0)\| \leq C\|x - x_0\|.$$

(Die zweite Gleichheit gilt, weil $\phi$ eine lineare Abbildung ist.) Nun kann man leicht zeigen, dass $\phi$ stetig ist. (Für gegebenes $\varepsilon$ wählt man $\delta = \frac{\varepsilon}{C}$.) □

*Beispiel.* Wir betrachten

$$V := C^\infty([0,1]) := \text{Vektorraum aller } C^\infty\text{–Funktionen } [0,1] \to \mathbb{R}$$

mit der Supremumsnorm

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

Die Abbildung

$$\begin{array}{rcl} \phi : V & \to & V \\ f & \mapsto & f' = \frac{df}{dx} \end{array}$$

ist eine lineare Abbildung. Für die Funktion $f_n(x) := x^n$ gilt $\|f_n\| = 1$, aber es gilt $\|f_n'\| = \|nx^{n-1}\| = n$. Insbesondere gibt es kein $C \geq 0$, so dass $\|f\| \leq C\|f\|$ für alle $f \in V$. Es folgt aus dem vorherigen Satz, dass $\phi : V \to V$ nicht stetig ist.

## 3. Kompakte Mengen

In Analysis I nannten wir Intervalle vom Typ $[a, b]$, also Intervalle bei denen die Endpunkte enthalten sind, kompakte Intervalle. Zudem hatten wir gesehen, dass wenn $I$ ein kompaktes Intervall ist, und $f : I \to \mathbb{R}$ stetig, dann gilt:

(1) $f$ ist beschränkt, d.h. es existiert ein $C \geq 0$, so dass $|f(x)| \leq C$ für alle $x \in I$ (siehe Analysis I, Lemma 8.1),
(2) $f$ nimmt sein Maximum und sein Minimum ein, d.h. es existieren $x_0, x_1 \in I$, so dass

$$f(x_0) \leq f(x) \leq f(x_1) \text{ für alle } x \in I,$$

(siehe Analysis I, Satz 8.2),
(3) $f$ ist gleichmäßig stetig (siehe Analysis I, Satz 8.4).

Wenn $X$ nun ein beliebiger metrischer Raum ist, dann werden wir im Folgenden sogenannte 'kompakte Teilmengen' einführen. Wenn $A \subset X$ kompakt ist, und wenn $f : A \to \mathbb{R}$ stetig ist, dann werden wir sehen, dass $f$ ebenfalls Eigenschaften (1), (2) und (3) besitzt.

### 3.1. **Definition von kompakten Mengen.**

*Definition.* Es sei $X$ ein metrischer Raum und $A \subset X$ eine Teilmenge. Eine *offene Überdeckung von $A$* ist eine Familie $\{U_i\}_{i \in I}$ von offenen Teilmengen von $X$, so dass

$$A \subset \bigcup_{i \in I} U_i.$$

*Beispiel.* Wir betrachten $A = (0, 1] \subset \mathbb{R}$. Für $i \in \mathbb{N}$ betrachten wir zudem

$$U_n := (\frac{1}{n}, 2).$$

Dann ist $\{U_n\}_{n \in \mathbb{N}}$ eine offene Überdeckung von $A$.

*Definition.* Es sei $X$ ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt *kompakt*, wenn es zu jeder offenen Überdeckung $\{U_i\}_{i\in I}$ von $A$ endlich viele Indizes $i_1, \dots, i_k \in I$ gibt, so dass

$$A \subset U_{i_1} \cup U_{i_2} \cup \dots \cup U_{i_k}.$$

Anders ausgedrückt, $A$ heißt kompakt wenn man für *jede* offene Überdeckung $\{U_i\}_{i\in I}$ von $A$ es endlich viele Mengen gibt, so dass diese schon $A$ überdecken.

*Beispiel.* Betrachten wir obiges Beispiel, dann sehen wir, dass $(0, 1]$ *nicht* kompakt ist. In der Tat, wenn wir

$$U_n := (\frac{1}{n}, 2),$$

setzen, dann ist $\{U_n\}_{n\in\mathbb{N}}$ eine offene Überdeckung von $(0, 1]$, aber $(0, 1]$ kann nicht durch endlich viele dieser Mengen $\{U_n\}_{n\in\mathbb{N}}$ überdeckt werden. In der Tat, wenn wir nur endliche viele $i_1, \dots, i_k \in \mathbb{N}$ betrachten, dann ist

$$\frac{1}{\max\{i_1, \dots, i_k\}} \notin U_{i_1} \cup U_{i_2} \cup \dots \cup U_{i_k},$$

aber

$$\frac{1}{\max\{i_1, \dots, i_k\}} \in (0, 1].$$

*Beispiel.* Wir betrachten

$$A := \{0\} \cup \{\frac{1}{n} \,|\, n \in \mathbb{N}\} \subset \mathbb{R}.$$

Wir behaupten, dass $A$ kompakt ist. Sei also $\{U_i\}_{i\in I}$ eine offene Überdeckung von $A$.

(1) Nachdem $\{U_i\}_{i\in I}$ eine offene Überdeckung von $A$ ist, existiert insbesondere ein $i \in I$, so dass $0 \in U_i$.
(2) Nachdem $U_i$ offen ist, existiert ein $\varepsilon > 0$, so dass $(-\varepsilon, \varepsilon) \subset U_i$.
(3) Nachdem $\varepsilon > 0$, existiert ein $N \in \mathbb{N}$, so dass $\frac{1}{N} \in (0, \varepsilon)$.
(4) Dann gilt aber auch, dass $\frac{1}{n} \in (0, \varepsilon)$ für alle $n \geq N$. Insbesondere enthält $U_i$ alle Punkte $\frac{1}{N}, \frac{1}{N+1}, \dots$.
(5) Nachdem $\{U_i\}_{i\in I}$ eine offene Überdeckung ist, existiert zu jedem $n \in \{1, \dots, N-1\}$ ein $i_n \in I$, so dass $\frac{1}{n} \in U_{i_n}$.

Es folgt nun, dass $A \subset U_i \cup U_{i_1} \cup \dots \cup U_{i_{N-1}}$.

## 3.2. **Der Satz von Heine–Borel.**

**Lemma 3.1.** *Es seien $a, b \in \mathbb{R}, a < b$. Dann ist $[a, b] \subset \mathbb{R}$ kompakt.*

*Beweis.* O.B.d.A. sei $a = 0, b = 1$. Wir schreiben $I := [0, 1]$. Nehmen wir an, $I$ ist nicht kompakt. Dann existiert eine offene Überdeckung $\{U_j\}_{j\in J}$, so dass $I$ nicht durch endlich viele $U_j$ überdeckt werden kann.

*Behauptung.* Es gibt eine Folge von abgeschlossenen Intervallen

$$I \supset I_1 \supset I_2 \supset \ldots$$

so dass für jedes $m$ gilt:

(1) $I_m$ kann nicht durch endlich viele $U_j$ überdeckt werden,
(2) die Länge von $I_m$ ist $2^{-m}$.

Wir zeigen die Behauptung durch vollständige Induktion. Wir setzen $I_0 := I$. Nehmen wir an, dass wir $I_m$ schon gefunden haben. Wir teilen $I_m$ in zwei abgeschlossene Intervalle $I'$ und $I''$ der halben Länge. Da $I_m$ nicht durch endlich viele $U_j$ überdeckt werden kann, kann zumindest eines der Intervalle $I'$ oder $I''$ nicht durch endlich viele $U_j$ überdeckt werden. Dieses Intervall wählen wir als $I_{m+1}$.

*Behauptung.* Es gibt ein $x \in I$ mit der Eigenschaft, $x \in \cap_{m\in\mathbb{N}} I_m$.

.

Für jedes $m$ wählen wir ein $a_m \in I_m$. Wenn $m \geq n$ gegeben sind, dann folgt, dass $a_m, a_n \in A_n$, insbesondere gilt $|a_m - a_n| \leq \frac{1}{2^n}$. Es folgt dass $(a_m)_{m\in\mathbb{N}}$ eine Cauchyfolge ist. [12] Wir setzen $x := \lim_{m\to\infty} a_m$. [13]

Wir müssen noch zeigen, dass $x \in \cap_{m\in\mathbb{N}} I_m$. Anders ausgedrückt, wir müssen zeigen, dass für jedes $m \in \mathbb{N}$ gilt: $x \in I_m$. Sei also $m \in \mathbb{N}$. Dann gilt für alle $n \geq m$, dass $a_n \in I_n \subset I_m$. Insbesondere ist $(a_n)_{n\geq m}$ eine konvergente Folge von Punkten in $I_m$. Nachdem $I_m$ abgeschlossen ist, folgt aus Satz 2.3, dass $\lim_{n\to\infty} a_n \in I_m$. Wir haben also damit die Behauptung bewiesen.

Nachdem $\{U_j\}_{j\in J}$ eine Überdeckung von $I$ ist existiert ein $j \in J$, so dass $x \in U_j$. Wegen der Offenheit von $U_j$ existiert ein $\varepsilon > 0$, so dass $(x - \varepsilon, x + \varepsilon) \subset U_j$. Wir wählen nun ein $m \in \mathbb{N}$, so dass $2^{-m} < \varepsilon$. Nachdem $x \in I_m$ und nachdem die Länge von $I_m$ genau $2^{-m}$ ist, folgt, dass

$$I_m \subset [x - 2^{-m}, x + 2^{-m}].$$

---

[12] Dies sieht man wie folgt: Sei also $\varepsilon > 0$. Dann existiert ein $N \in \mathbb{N}$, so dass $\frac{1}{2^N} < \varepsilon$, dann folgt aber, dass für alle $n, m \geq N$ gilt $|a_m - a_n| \leq \frac{1}{2^N} < \varepsilon$.

[13] Bis zu diesem Schritt ist der Beweis sehr ähnlich dem Beweis vom Satz von Bolzano–Weierstraß.

Dann folgt aber, dass

$$I_m \subset [x - 2^{-m}, x + 2^{-m}] \subset (x - \varepsilon, x + \varepsilon) \subset U_j,$$

d.h. $I_m$ ist in einem einzigen $U_j$ enthalten, im Widerspruch zur Eigenschaft (1) der Intervalle $I_m$. □

Ganz ähnlich beweist man auch folgenden Satz (siehe Forster Seite 26).

**Satz 3.2.** *Für jedes $a > 0$ ist der Würfel*

$$[-a, a]^n := \{(x_1, \dots, x_n) \,|\, x_i \in [-a, a]\} \subset \mathbb{R}^n$$

*kompakt.*

Mithilfe des folgenden Satzes erhalten wir weitere Beispiele von kompakten Mengen:

**Satz 3.3.** *Es sei $X$ ein metrischer Raum, $K \subset X$ eine kompakte Teilmenge und $A \subset K$ eine abgeschlossene Teilmenge. Dann ist auch $A$ kompakt.*

*Beweis.* Es sei $\{U_i\}_{i \in I}$ eine offene Überdeckung von $A$. Nachdem $A$ abgeschlossen ist, folgt, dass $X \setminus A$ offen ist. Insbesondere ist

$$(X \setminus A) \cup \{U_i\}_{i \in I}$$

eine offene Überdeckung von $K$. Nachdem $K$ kompakt ist, gibt es $i_1, \dots, i_k$, so dass

$$K \subset (X \setminus A) \cup U_{i_1} \cup \dots \cup U_{i_k}.$$

Nachdem $A \subset K$, folgt, dass

$$A \subset (X \setminus A) \cup U_{i_1} \cup \dots \cup U_{i_k}.$$

Aber nachdem $A \cap (X \setminus A) = \emptyset$ erhalten wir, dass

$$A \subset U_{i_1} \cup \dots \cup U_{i_k}.$$

□

Im Folgenden sagen wir, dass eine Teilmenge $A$ eines metrischen Raumes *beschränkt* ist, wenn es ein $C \geq 0$ gibt, so dass

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

**Satz 3.4.** *Es sei $X$ ein metrischer Raum. Wenn $A \subset X$ kompakt ist, dann ist $A$ beschränkt und abgeschlossen.*

*Beweis.* Wir zeigen zuerst, dass $A$ beschränkt ist. Sei $a \in A$. Für $n \in \mathbb{N}$ betrachten wir

$$U_n := B_n(a).$$

Nach Lemma 1.2 ist $U_n$ offen. Nachdem jeder Punkt einen endlichen Abstand von $a$ hat, folgt, dass $\cup_{n\in N} U_n = X$. Insbesondere ist $\{U_n\}_{n\in\mathbb{N}}$ eine offene Überdeckung von $A$. Nachdem $A$ kompakt ist, existieren $n_1 < n_2 < \cdots < n_k$, so dass

$$A \subset U_{n_1} \cup \cdots \cup U_{n_k},$$

insbesondere gilt, dass $d(a, x) < n_k$ für alle $x \in A$. Wir setzen nun $C := 2n_k$. Für alle $x, y \in A$ gilt dann

$$d(x, y) \leq d(x, a) + d(a, y) < n_k + n_k = C.$$

Wir haben damit bewiesen, dass $A$ beschränkt ist.

Wir wollen nun zeigen, dass $A$ abgeschlossen ist, d.h. wir wollen zeigen, dass $X \setminus A$ offen ist. Sei also $x \in X \setminus A$. Wir müssen zeigen, dass es ein $\varepsilon > 0$ gibt, so dass $B_\varepsilon(x) \subset X \setminus A$. Für $n \in \mathbb{N}$ betrachten wir die offene Teilmenge

$$U_n := \{y \in X \mid d(y, x) > \frac{1}{n}\}.$$

Wie in Lemma 1.2 zeigt man, das $U_n$ offen ist. Offensichtlich gilt $\cup_{n\in\mathbb{N}} U_n = X \setminus \{x\}$. Nachdem $x \in X \setminus A$ folgt, dass $\{U_n\}_{n\in\mathbb{N}}$ eine offene Überdeckung von $X \setminus A$ ist. Nachdem $A$ kompakt ist, existieren $n_1 < n_2 < \cdots < n_k$, so dass

$$A \subset U_{n_1} \cup \cdots \cup U_{n_k},$$

insbesondere gilt, dass $d(a, x) \geq \frac{1}{n_k}$ für alle $x \in A$. Es folgt, dass

$$B_{\frac{1}{n_k}}(a) \subset X \setminus A.$$

Insbesondere ist $X \setminus A$ offen.

□

**Satz 3.5. Satz von Heine–Borel** *Es sei $A$ eine Teilmenge von $\mathbb{R}^n$.* [14] *Dann gilt:*

$$A \textit{ ist kompakt } \Leftrightarrow \; A \textit{ ist beschränkt und abgeschlossen.}$$

*Beweis.* Wenn $A$ kompakt ist, dann besagt Satz 3.4, dass $A$ auch beschränkt und abgeschlossen ist.

Nun nehmen wir an, dass $A$ beschränkt und abgeschlossen ist. Nachdem $A$ beschränkt ist existiert ein $a > 0$, so dass $A$ eine Teilmenge des

[14] Zur Erinnerung, hier betrachten wir $\mathbb{R}^n$ mit der euklidischen Metrik. Die Aussage ist i.a. nicht richtig, wenn wir $\mathbb{R}^n$ mit einer anderen Metrik betrachten.

Würfels $[-a, a]^n \subset \mathbb{R}^n$ ist. [15] Der Würfel ist nach Satz 3.2 kompakt. Nachdem $A$ abgeschlossen ist, folgt dann aus Satz 3.3, dass auch $A$ kompakt ist. □

*Beispiel.* Wir betrachten die $n$–dimensionale Sphäre

$$S^n := \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \dots + x_{n+1}^2 = 1\} = \{x \in \mathbb{R}^n \mid d(x, O) = 1.$$

Wir wollen zeigen, dass $S^n$ kompakt ist. Nach dem Satz von Heine–Borel genügt es zu zeigen, dass $S^n$ beschränkt und abgeschlossen ist.

Für alle $x, y \in S^n$ gilt

$$d(x, y) \le d(x, O) + d(O, y) = 1 + 1 = 2,$$

d.h. $S^n$ ist beschränkt mit $C = 2$. Wir zeigen nun, dass $S^n$ abgeschlossen ist. Wir betrachten

$$\begin{array}{rcl} f\colon \mathbb{R}^{n+1} & \to & \mathbb{R} \\ (x_1, \dots, x_{n+1}) & \mapsto & x_1^2 + \dots + x_{n+1}^2. \end{array}$$

Die Funktion ist stetig und es gilt $S^n = f^{-1}(1)$. Nachdem $\{1\} \subset \mathbb{R}$ abgeschlossen ist, folgt aus Satz 2.5, dass auch $S^n$ abgeschlossen ist. [16]

## 3.3. **Eigenschaften von kompakten Mengen.**

**Satz 3.6.** *Es sei $f\colon X \to Y$ eine stetige Abbildung zwischen metrischen Räumen. Ist $K \subset X$ kompakt, dann ist auch $f(K)$ kompakt.*

*Beweis.* Sei also $\{U_i\}_{i \in I}$ eine offene Überdeckung von $f(K)$. Wir müssen zeigen, dass $f(K)$ in der Vereinigung von endlich vielen $U_i$'s enthalten ist.

Man beachte, dass

$$K \subset f^{-1}(f(K)) \subset f^{-1}(\{U_i\}_{i \in I}) = \cup_{i \in I} f^{-1}(U_i).$$

[17] Nachdem $f$ stetig folgt aus Satz 2.5, dass die Urbilder von den $U_i$ ebenfalls offen, insbesondere ist $\{f^{-1}(U_i)\}_{i \in I}$ eine offene Überdeckung

[15] Dies sieht man wie folgt: Nachdem $A$ beschränkt ist, existiert ein $C \ge 0$, so dass $d(x, y) \le C$ für alle $x, y \in A$. Sei nun $x \in A$ beliebig. Wir bezeichnen den Nullpunkt von $\mathbb{R}^n$ mit $O$. Dann gilt für alle $y \in A$, dass

$$d(O, y) \le d(O, x) + d(x, y) \le d(O, x) + C.$$

Wenn wir nun $a := d(O, x) + C$ setzen, dann gilt $A \subset [-a, a]^n$.

[16] Man könnte natürlich auch versuchen 'per Hand' mithilfe der $\varepsilon$–Definition zu zeigen, dass $\mathbb{R}^{n+1} \setminus S^n$ offen ist. Aber dies ist gelinde gesagt umständlich. In vielen Fällen kann man die Offenheit bzw. Abgeschlossenheit einer Menge durch geschicktes anwenden von Satz 2.5 zeigen.

[17] Hier verwenden wir folgende Tatsache: Es sei $f\colon X \to Y$ eine beliebige Abbildung zwischen Mengen und es sei $U \subset Y$ eine Teilmenge. Dann gilt

$$f(f^{-1}(U)) = U.$$

von $K$. Nachdem $K$ kompakt ist, existieren Indizes $i_1, \ldots, i_k$, so dass

$$K \subset (f^{-1}(U_{i_1}) \cup \cdots \cup f^{-1}(U_{i_k})).$$

Dann gilt aber auch, dass

$$\begin{aligned} f(K) &\subset f\left(f^{-1}(U_{i_1}) \cup \cdots \cup f^{-1}(U_{i_k})\right) \\ &= f\left(f^{-1}(U_{i_1})\right) \cup \cdots \cup f\left(f^{-1}(U_{i_k})\right) = U_{i_1} \cup \cdots \cup U_{i_k}. \end{aligned}$$

□

Das folgende Korollar ist eine Verallgemeinerung von Satz 8.2 aus Analysis I.

**Korollar 3.7.** *Es sei $K$ eine kompakte Teilmenge eines metrischen Raumes $X$.* [18] *Es sei $f\colon K \to \mathbb{R}$ eine stetige Abbildung. Dann ist $f$ beschränkt, und $f$ nimmt ein Minimum und ein Maximum an, d.h. es existieren $x_0, x_1 \in K$, so dass*

$$f(x_0) \leq f(x) \leq f(x_1)$$

*für alle $x \in K$.*

*Beweis.* Nachdem $K$ kompakt und $f$ stetig ist, folgt aus Satz 3.6, dass das Bild $f(K) \subset \mathbb{R}$ kompakt ist. Aus Satz 3.4 folgt, dass $f(K) \subset \mathbb{R}$ beschränkt und abgeschlossen ist. Da $f(K)$ beschränkt ist, existiert insbesondere $y_0 := \inf(f(K))$. Nachdem $f(K)$ abgeschlossen ist, gilt $y_0 \in f(K)$. [19] Es existiert also ein $x_0 \in K$, so dass $f(x_0) = y_0 = \inf(f(K))$. Dann gilt aber für alle $x \in K$, dass

$$f(x_0) \leq f(x).$$

Die Existenz von $x_1$ zeigt man ganz analog. □

---

Andererseits gilt für eine beliebige Teilmenge $U \subset X$, dass

$$U \subset f^{-1}(f(U)),$$

aber in diesem Falle gilt i.a. nicht, dass $U = f^{-1}(f(U))$. Betrachten wir z.B. die Abbildung $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$, dann ist $f^{-1}(f(2)) = \pm 2$.

[18] Hier, und in den anderen Sätzen ist natürlich der Fall $K = X$ zugelassen, wenn $X$ schon selber kompakt ist.

[19] Zur Erinnerung, das Supremum $y_0 := \sup(f(K)) \in \mathbb{R}$ ist durch folgende zwei Eigenschaften eindeutig charakterisiert:

(1) für alle $a \in f(K)$ gilt $y_0) \geq a$,
(2) für alle $\varepsilon > 0$ gibt es ein $a \in A$, so dass $a > y_0 - \varepsilon$.

Wir wollen nun zeigen, dass $y_0 \in f(K)$. Nehmen wir an, dies wäre nicht der Fall, dann ist $y_0 \in \mathbb{R} \setminus f(K)$. Nachdem $f(K)$ abgeschlossen ist, ist $\mathbb{R} \setminus f(K)$ offen, d.h. es existiert ein $\varepsilon > 0$, so dass $(y_0 - \varepsilon, y_0 + \varepsilon)$ ebenfalls in $\mathbb{R} \setminus f(K)$ liegt. Aber dies ist ein Widerspruch zu (2).

*Beispiel.* Betrachten wir $X := \{v \in \mathbb{R}^2 \,|\, 0 < \|v\| \leq 1\}$. Dann ist $X$ beschränkt aber nicht abgeschlossen. In diesem Fall existiert eine unbeschränkte stetige Funktion auf $X$, z.B.

$$\begin{array}{rcl} f\colon X & \to & \mathbb{R} \\ x & \mapsto & \frac{1}{\|x\|}. \end{array}$$

Betrachten wir hingegen die stetige Funktion

$$\begin{array}{rcl} g\colon X & \to & \mathbb{R} \\ x & \mapsto & \|x\|, \end{array}$$

dann ist diese zwar beschränkt, aber es gibt kein $x_0 \in X$, so dass $f(x_0) \leq f(x)$ für alle $x \in X$.

**Satz 3.8.** *(Bolzano–Weierstraß) Es sei $K$ eine kompakte Teilmenge eines metrischen Raumes $X$. Es sei $(x_n)_{n\in\mathbb{N}}$ eine Folge von Punkten in $K$. Dann gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, welche in $K$ konvergiert.*

*Beweis.* Nehmen wir an, es existiert keine Teilfolge, welche in $K$ konvergiert. Wir wollen diese Annahme zu einem Widerspruch führen. Wir beweisen zuerst folgende Behauptung.

*Behauptung.* Es sei $a \in K$ beliebig. Dann existiert ein $l \in \mathbb{N}$, so dass $B_{\frac{1}{l}}(a)$ höchstens endlich viele Folgenglieder enthält.

Nehmen wir an, dies sei nicht der Fall. Dann würde zu jedem $l \in \mathbb{N}$ unendlich viele Folgenglieder existiert, welche sich in $B_{\frac{1}{l}}(a)$ befinden. Insbesondere könnte man dann

$$n_1 < n_2 < n_3 < \dots$$

finden, so dass $x_{n_l} \in B_{\frac{1}{l}}(a)$. Dann hätten wir aber eine Teilfolge konstruiert, welche gegen $a \in K$ konvergiert. Wir hatten aber angenommen, dass es solche Teilfolgen nicht gibt.

Wir setzen nun $U_a := B_{\frac{1}{l}}(a)$. Wir haben nun also für jedes $a \in K$ eine offene Menge $U_a$ mit folgenden Eigenschaften gefunden:

(1) $U_a$ enthält den Punkt $a$ enthält,
(2) $U_a$ enthält höchstens endlich viele Folgenglieder.

Aus (1) folgt, dass

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

Nachdem jedes $U_a$ offen ist, ist also $\{U_a\}_{a\in A}$ eine offene Überdeckung von $K$. Nachdem $K$ kompakt ist, gibt es nun aber $a_1, \dots, a_n$, so dass

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

Nachdem jedes $U_{a_i}$ nach (2) aber höchstens endlich viele Folgenglieder enthält, enthält $K$ höchstens endlich viele Folgenglieder. Das kann aber nicht sein, weil $K$ alle Folgenglieder enthält. $\square$

Wir erhalten nun folgende Verallgemeinerung von Satz 4.13 aus Analysis I.

**Korollar 3.9.** *Jede beschränkte Folge in $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.*

*Beweis.* Es sei $(x_k)_{k\in\mathbb{N}}$ eine beschränkte Folge, d.h. es gibt ein $C \in \mathbb{R}$, so dass $\|x_k\| \leq C$ für alle $k \in \mathbb{N}$. Dann ist $(x_k)_{k\in\mathbb{N}}$ eine Folge in

$$K := \{v \in \mathbb{R}^n \mid \|v\| \leq C\},$$

aber dies ist eine kompakte Teilmenge von $\mathbb{R}^n$. (Siehe z.B. Übungsblatt 3). Aus dem Satz von Bolzano–Weierstrass folgt nun, dass $(x_k)_{k\in\mathbb{N}}$ eine konvergente Teilfolge besitzt. $\square$

Es sei nun $f\colon X \to Y$ eine Abbildung zwischen metrischen Räumen $X$ und $Y$. Wir definieren:

$$f \text{ gleichmäßig stetig } :\Leftrightarrow \underset{\varepsilon>0}{\forall}\ \underset{\delta>0}{\exists}\ \underset{x\in X}{\forall}\ \underset{y\in B_\delta(x)}{\forall} d(f(x), f(y)) < \varepsilon.$$

Zur Erinnerung: wenn $x_0 \in D$, dann hatten wir definiert:

$$f \text{ stetig im Punkt } x_0 \ :\Leftrightarrow \underset{\varepsilon>0}{\forall}\ \underset{\delta>0}{\exists}\ \underset{x\in B_\delta(x_0)\cap D}{\forall} d(f(x), f(x_0)) < \varepsilon.$$

Diese Definition von gleichmäßiger Stetigkeit verallgemeinert den Begriff der gleichmäßig stetigen Funktionen $f\colon D \to \mathbb{R}$.

Satz 8.4 aus Analysis I besagt, dass jede stetige Funktion $f\colon [a, b] \to \mathbb{R}$ auch gleichmäßig stetig ist. Der folgende Satz ist nun einer Verallgemeinerung:

**Satz 3.10.** *Es sei $f\colon X \to Y$ eine stetig Abbildung. Wenn $X$ kompakt ist, dann ist $f$ auch gleichmäßig stetig.*

Wir verweisen auf Forster, Analysis II, Seite 34 für den Beweis.

## 4. Kurven in $\mathbb{R}^n$

4.1. **Definitionen und Beispiele.** Wir wenden uns in diesem Kapitel wieder deutlich 'konkreteren' Objekten zu.

*Definition.* (1) Eine *Kurve* in $\mathbb{R}^n$ ist eine stetige Abbildung

$$f = (f_1, \ldots, f_n) : I \to \mathbb{R}^n,$$

wobei $I \subset \mathbb{R}$ ein Intervall ist.

(2) Wir sagen die Kurve ist *(stetig) differenzierbar*, wenn die $n$ Koordinatenfunktionen $f_i : I \to \mathbb{R}$, $i = 1, \dots, n$ (stetig) differenzierbar sind.

*Beispiel.* (1) Die Bildmenge der Kurve

$$\begin{array}{rcl} f\colon [0, 2\pi] & \to & \mathbb{R}^2, \\ t & \mapsto & (\sin(t), \cos(t)) \end{array}$$

ist der Einheitskreis um den Ursprung. Man beachte, dass die Kurve

$$\begin{array}{rcl} f\colon \mathbb{R} & \to & \mathbb{R}^2, \\ t & \mapsto & (\sin(t), \cos(t)) \end{array}$$

die gleiche Bildmenge besitzt.[20]

(2) Es seien $r > 0$ und $c \neq 0$ gegeben. Die Kurve

$$\begin{array}{rcl} f\colon \mathbb{R} & \to & \mathbb{R}^3 \\ t & \mapsto & (r\cos(t), r\sin(t), ct) \end{array}$$

beschreibt eine unendliche Helix.

*Definition.* Es sei

$$f = (f_1, \dots, f_n) : I \to \mathbb{R}^n$$

eine differenzierbare Kurve und $t \in I$. Wir nennen

$$f'(t) := (f_1'(t), \dots, f_n'(t)) \in \mathbb{R}^n$$

den *Tangentialvektor* der Kurve $f$ zum Parameterwert $t$.

*Beispiel.* Die Kurve

$$\begin{array}{rcl} f\colon \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & (t^2 - 1, t^3 - t) \end{array}$$

ist nicht injektiv, in der Tat, für die Parameterwerte $t = \pm 1$ gilt

$$f(1) = f(-1) = (0, 0).$$

und die Tangentialvektoren sind

$$f'(-1) = (-2, 2) \text{ und } f'(1) = (2, 2).$$

*Definition.* Es sei $f\colon I \to \mathbb{R}^n$ eine stetig differenzierbare Kurve.

(1) $t \in I$ heißt *singulär*, wenn $f'(t) = 0$,
(2) wenn $f$ keine singulären Punkte besitzt, dann heißt $f$ *regulär*.

[20]Eine Kurve in unserem Sinne ist also wirklich eine Abbildung, und nicht nur eine 'ein-dimensionale' Teilmenge von $\mathbb{R}^n$.

*Beispiel.* Die Kurve

$$\begin{array}{rcl} f\colon \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & (t^2, t^3) \end{array}$$

heißt Neilsche Parabel. Es gilt $f'(t) = (2t, 3t^2)$, d.h. $t = 0$ ist der einzige singuläre Punkt der Neilschen Parabel.

### 4.2. **Rektifizierbare Kurven.**

*Definition.* Eine Kurve $f\colon [a,b] \to \mathbb{R}^n$ heißt *Polygonzug*, wenn es Punkte $P_0, \dots, P_k \in \mathbb{R}^n$ und eine Unterteilung

$$a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b$$

des Intervalls $[a,b]$ gibt, mit folgender Eigenschaft: Für jedes $i \in \{0, \dots, k-1\}$ und $t \in [t_i, t_{i+1}]$ gilt:

$$f(t) = P_i \cdot \frac{t_{i+1} - t}{t_{i+1} - t_i} + P_{i+1} \cdot \frac{t - t_i}{t_{i+1} - t_i}.$$

Bildlich ausgedrückt, ein Polygonzug bewegt sich für $t \in [t_i, t_{i+1}]$ mit konstanter Geschwindigkeit vom Punkt $P_i$ zum Punkt $P_{i+1}$.

Wenn $f$ ein Polygonzug wie oben ist, dann definieren wir die *Länge des Polygonzuges* $f$ als

$$L(f) := \sum_{i=0}^{k-1} \|P_{i+1} - P_i\|.$$

(Wie üblich verwenden wir die euklidische Norm auf $\mathbb{R}^n$.)

Wenn $f\colon [a,b] \to \mathbb{R}^n$ eine beliebige Kurve ist, dann wollen wir auch einen Längenbegriff einführen. Die Idee ist zu sagen, dass $f$ eine Länge $L$ hat, wenn man $f$ durch Polygonzüge annähern kann, deren Länge gegen $L$ konvergiert. Wenn man versucht diese Idee mathematisch sauber zu formulieren, muss man etwas arbeiten:

*Definition.* Eine *Zerlegung vom Intervall* $[a,b]$ *der Feinheit* $\delta$ besteht aus Zahlen

$$a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b,$$

so dass $t_{i+1} - t_i < \delta$ für alle $i$. Wenn $f\colon [a,b] \to \mathbb{R}^n$ eine Kurve ist, dann definieren wir

$$L(f, t_0, \dots, t_k) := \sum_{i=0}^{k-1} \|f(t_{i+1}) - f(t_i)\|.$$

Dies ist also die Länge des Polygonzuges, welcher durch die Punkte $f(t_0), f(t_1), \dots, f(t_k)$ definiert ist.

*Definition.* Wir sagen eine Kurve $f\colon [a,b] \to \mathbb{R}^n$ ist *rektifizierbar* wenn es ein $L$ gibt mit folgender Eigenschaft:
Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für jede Zerlegung

$$a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b,$$

der Feinheit $\delta$ gilt:

$$|L(f, t_0, t_1, \ldots, t_k) \ - \ L| < \varepsilon.$$

Wenn $f$ rektifizierbar ist, dann nennen wir $L$ die *Länge* von $f$.

Man kann (relativ) leicht zeigen, dass Polygonzüge rektifizierbar sind. Wir haben nun zwei Definitionen von Längen von Polygonzügen (die Definition nur für Polygonzüge, und dann allgemein für alle rektifizierbaren Kurven), und man kann leicht zeigen, dass beide das gleiche Ergebnis liefern. Eine Übungsaufgabe dazu befindet sich auf dem 4. Übungsblatt.

Deutlich interessanter ist aber folgender Satz:

**Satz 4.1.** *Es sei $f\colon [a,b] \to \mathbb{R}^n$ eine stetig differenzierbare Kurve, dann ist $f$ rektifizierbar, und es gilt*

$$\textit{Länge}(f) = \int_a^b \|f'(t)\|\, dt.$$

Betrachten wir zuerst ein Geradenstück ist, d.h. wenn $f$ eine Kurve vom Typ

$$\begin{array}{rcl} f\colon [a,b] & \to & \mathbb{R}^n \\ t & \mapsto & P + tv, \end{array}$$

für einen Punkt $P \in \mathbb{R}^n$ und einen Vektor $v \in \mathbb{R}^n$. Dann ist dies insbesondere ein Polygonzug mit den Eckpunkten $f(a)$ und $f(b)$ , d.h. die Länge ist $\|f(b) - f(a)\| = \|(b-a)v\| = (b-a)\|v\|$. Andererseits gilt $f'(t) = v$, d.h. $\int_a^b \|f'(t)\|\, dt = (b-a)\|v\|$. Die Aussage des Satzes gilt also in diesem Fall. Der allgemeine Fall kann letztendlich durch geschicktes argumentieren auf diesen Fall zurückgeführt werden. Die (nicht ganz einfachen) Details finden Sie in Forster, Analysis II, §4.

*Beispiel.* Wir betrachten die Kurve

$$\begin{array}{rcl} f\colon [0,2\pi] & \to & \mathbb{R}^2 \\ t & \mapsto & (\cos(t), \sin(t)), \end{array}$$

dann gilt

$$\begin{array}{rcl} \text{Länge}(f) & = & \int_0^{2\pi} \|f'(t)\|\, dt = \int_0^{2\pi} \|(-\sin(t), \cos(t))\|\, dt \\ & = & \int_0^{2\pi} \sqrt{(-\sin(t))^2 + \cos(t)^2}\, dt \\ & = & \int_0^{2\pi} 1\, dt = 2\pi. \end{array}$$

*Beispiel.* Man beachte jedoch, dass nicht jede Kurve rektifizierbar ist. Beispielsweise kann man zeigen, dass die Kurve

$$\begin{array}{rcl} f\colon [0,1] & \to & \mathbb{R}^2 \\ t & \mapsto & \begin{cases} (t, t\cdot\cos(1/t)), & \text{falls } t>0, \\ (0,0), & \text{falls } t=0. \end{cases} \end{array}$$

stetig, aber nicht rektifizierbar ist.

*Beispiel.* Ein weiteres Beispiel ist durch die Peano–Kurve gegeben:
`http://de.wikipedia.org/wiki/Peano-Kurve`
diese Kurve ist eine stetige Abbildung $f\colon [0,1]\to\mathbb{R}^2$, so dass die Bildmenge aus dem Quadrat $[0,1]^2$ besteht. Die Peano–Kurve ist nicht rektifizierbar.

4.3. **Parametertransformationen.** Es sei $f\colon [a,b]\to\mathbb{R}^n$ eine Kurve und $\varphi\colon [c,d]\to[a,b]$ eine stetige bijektive Abbildung. Dann ist

$$\begin{array}{rcl} f\circ\varphi : [c,d] & \to & \mathbb{R}^n \\ t & \mapsto & f(\varphi(t)) \end{array}$$

ebenfalls eine Kurve, welche genau die gleichen Punkte annimmt wie $f$. Wir sagen, diese Kurve geht aus der Kurve $f$ durch *Parametertransformation* $\varphi$ hervor.

**Lemma 4.2.** *Nehmen wir an, dass $f\colon [a,b]\to R^n$ eine stetig differenzierbare Kurve ist, und dass sowohl $\varphi$ als auch $\varphi^{-1}$ stetig differenzierbar sind. Dann gilt*

$$\textit{Länge}(f) = \textit{Länge}(f\circ\varphi).$$

*Beweis.* Nachdem $\varphi$ stetig und bijektiv ist, folgt, dass entweder $\varphi$ streng monoton steigend oder streng monoton fallend ist. Wir betrachten zuerst den Fall, dass $\varphi$ streng monoton steigend ist, d.h. $\varphi'(s)\geq 0$ für alle $s\in[c,d]$. Dann gilt

$$\begin{array}{rcl} \text{Länge}(f\circ\varphi) & = & \int_c^d \|(f\circ\varphi)'(s)\|\,ds \\ & = & \int_c^d \|f'(\varphi(s))\cdot\varphi'(s)\|\,ds \\ & = & \int_c^d \|f'(\varphi(s))\|\cdot\varphi'(s)\,ds \\ & = & \int_{\varphi(c)}^{\varphi(d)} \|f'(u)\|\,du \\ & = & \int_a^b \|f'(u)\|\,du \\ & = & \text{Länge}(f). \end{array}$$

Die zweite Gleichheit folgt aus der Kettenregel für Ableitungen (angewandt auf die Komponenten $f_i\circ\varphi$) und die vierte folgt aus der Substitution $u=\varphi(s)$.

Jetzt betrachten wir den Fall, dass $\varphi$ streng monoton fallend ist. Es folgt, dass $\varphi'(s) \leq 0$ für alle $s$. Wir wenden das obige Argument noch einmal an, allerdings mit zwei kleinen Modifikationen:

$$\begin{aligned}
\text{Länge}(f \circ \varphi) &= \int_c^d \|(f \circ \varphi)'(s)\| \, ds \\
&= \int_c^d \|f'(\varphi(s)) \cdot \varphi'(s)\| \, ds \\
&= \int_c^d \|f'(\varphi(s))\| \cdot |\varphi'(s)| \, ds \\
&= \int_c^d \|f'(\varphi(s))\| \cdot (-\varphi'(s)) \, ds \\
&= -\int_{\varphi(c)}^{\varphi(d)} \|f'(u)\| \, du \\
&= -\int_b^a \|f'(u)\| \, du \\
&= \int_a^b \|f'(u)\| \, du \\
&= \text{Länge}(f).
\end{aligned}$$

$\square$

Das Lemma besagt also, dass eine Parametertransformation die Länge nicht verändert (zumindest unter der Annahme, dass sowohl $\varphi$ als auch $\varphi^{-1}$ stetig differenzierbar sind. [21]

## 5. Partielle Ableitungen

Im folgenden bezeichnen wir mit

$$e_i := (0, \ldots, 0, 1, 0, \ldots, 0) \subset \mathbb{R}^n$$

(mit der '1' in der $i$–ten Koordinate), den $i$–ten Einheitsvektor. Die Vektoren $e_1, \ldots, e_n$ bilden die Standardbasis des $\mathbb{R}^n$. Die Basis $\{e_1, \ldots, e_n\}$ wird manchmal auch die kanonische Basis des $\mathbb{R}^n$ genannt.

---

[21] Man kann das Lemma auch direkt über die Definition von der Länge einer Kurve zeigen. In der Tat, sei $L$ die Länge von $f$ und es sei $\varepsilon > 0$ gegeben. Es sei $\delta > 0$ so gewählt wie in der Definition von der Länge von $f$. Wir setzen

$$C := \sup\{|\varphi'(t)| \mid t \in [a, b]\}.$$

Dann kann man leicht sehen, dass für das gleiche $L$ und $\varepsilon$ jetzt $\delta' := \frac{1}{C}\delta$ die gewünschte Eigenschaft für $f \circ \varphi$ besitzt.

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ eine Funktion. Es sei $x \in U$. Wenn der Grenzwert [22]

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

existiert, dann heißt $f$ *partiell in der* $i$*–ten Koordinate differenzierbar* und wir nennen $\frac{\partial}{\partial x_i} f$ die $i$*–te partielle Ableitung von* $f$ *am Punkt* $x$. [23]

*Definition.* Wir sagen $f$ ist in $x$ partiell differenzierbar, wenn $f$ in allen Koordinaten partiell differenzierbar ist. Die Funktion $f\colon U \to \mathbb{R}$ heißt partiell differenzierbar ist, wenn $f$ in allen Punkten $x \in U$ partiell differenzierbar ist. Wir sagen, $f$ ist *stetig partiell differenzierbar*, wenn alle partiellen Ableitungen stetig sind.[24]

Sei nun $(x_1, \dots, x_n) \in U$ ein fest gewählter Punkt. Dann ist

$$\frac{\partial}{\partial x_i} f(x) := \lim_{h \to 0} \frac{f(x_1, \dots, x_i + he_i, \dots, x_n) - f(x_1, \dots, x_i, \dots x_n)}{h}$$

nichts anderes als die Ableitung der Funktion

$$x_i \mapsto f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$$

bei der wir $x_1, \dots, x_{i-1}$ und $x_{i+1}, \dots, x_n$ als Konstanten auffassen. Wir können deswegen partielle Ableitungen wie gewöhnliche Ableitungen von Funktionen einer Variable bestimmen.

*Beispiel.* (1) Wenn $f(x_1, x_2, x_3) = 3x_1^2 x_2 + x_2^5 + \sin(x_1 x_3)$, dann ist

$$\begin{aligned} \tfrac{\partial}{\partial x_1} f &= 6x_1 x_2 + x_3 \cos(x_1 x_3), \\ \tfrac{\partial}{\partial x_2} f &= 3x_1^2 + 5x_2^4 \\ \tfrac{\partial}{\partial x_3} f &= x_1 \cos(x_1 x_3). \end{aligned}$$

---

[22] Wir bestimmen hier den Grenzwert der Funktion

$$\begin{array}{ccc} (-\varepsilon, \varepsilon) & \to & \mathbb{R} \\ h & \mapsto & \frac{f(x+he_i)-f(x)}{h}, \end{array}$$

d.h. wir bestimmen den Grenzwert einer reellwertigen Funktion in einer Variablen, wie in Analysis I schon definiert.

[23] Der Grenzwert

$$\frac{f(x + he_i) - f(x)}{h}$$

macht nur Sinn, wenn es ein $\varepsilon > 0$ gibt, so dass $x + he_i$ für $h \in (-\varepsilon, \varepsilon)$ sich noch im Definitionsbereich von $f$ befindet. Wir betrachten daher von nun an zumeist Funktionen und Abbildungen, welche auf *offenen* Teilmengen des $\mathbb{R}^n$ definiert sind, dann können wir um jeden Punkt des Definitionsbereiches und für jedes $i$ immer solch ein $\varepsilon > 0$ finden.

[24] Die partiellen Ableitungen sind wiederum Funktion $\frac{\partial}{\partial x_i} f\colon U \to \mathbb{R}$, welche wir eben auf Stetigkeit überprüfen können.

(2) Für Funktionen $\mathbb{R}^2 \to \mathbb{R}$ verwendet man oft auch $x$ und $y$ anstatt $x_1$ und $x_2$. Beispielsweise gilt
$$\frac{\partial}{\partial y} e^{x^2+y^2} = 2ye^{x^2+y^2}.$$

Wir haben oben festgestellt, dass partielle Ableitungen nichts anderes sind, als gewöhnliche Ableitungen von Funktionen einer Variable. Wir erhalten damit folgende Rechenregeln: Es sei $U \subset \mathbb{R}^n$ offen und es seien $f, g\colon U \to \mathbb{R}$ zwei partiell differenzierbare Funktionen, dann gilt
$$\begin{aligned} \tfrac{\partial}{\partial x_i}(f+g)(x) &= \tfrac{\partial}{\partial x_i} f(x) + \tfrac{\partial}{\partial x_i} g(x), \\ \tfrac{\partial}{\partial x_i}(f \cdot g)(x) &= f(x) \cdot \tfrac{\partial}{\partial x_i} g(x) + \tfrac{\partial}{\partial x_i} f(x) \ \cdot g(x). \end{aligned}$$
Zudem, wenn $g(x) \neq 0$ für alle $x \in U$, dann gilt
$$\frac{\partial}{\partial x_i}\left(\frac{f}{g}\right)(x) = \frac{g(x) \cdot \frac{\partial}{\partial x_i} f(x) - \frac{\partial}{\partial x_i} g(x) \ \cdot f(x)}{g(x)^2}.$$

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ eine partiell differenzierbare Funktion. Es sei $x \in U$. Dann heißt der Vektor
$$\nabla f(x) := \operatorname{grad} f(x) := \left(\frac{\partial}{\partial x_1} f(x), \dots, \frac{\partial}{\partial x_n} f(x)\right)$$
der *Gradient* von $f$ am Punkt $x$. [25]

*Beispiel.* Wir betrachten $f\colon \mathbb{R}^2 \to R$, $f(x,y) = \frac{1}{6}(x^2+y^2)$, dann ist der Gradient am Punkt $(x,y)$ gegeben durch
$$\operatorname{grad} f(x,y) = \left(\frac{\partial}{\partial x} f(x,y), \frac{\partial}{\partial y} f(x,y)\right) = (\frac{1}{3}x, \frac{1}{3}y).$$
In Skizzen zeichnet man normalerweise $\operatorname{grad} f(x,y)$ als Vektor, welcher am Punkt $(x,y)$ beginnt.

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ eine partiell differenzierbare Funktion.

(1) Wenn die partiellen Ableitungen $\frac{\partial}{\partial x_i} f\colon U \to \mathbb{R}, i = 1, \dots, n$ wieder partiell differenzierbar sind, dann heißt $f$ *zweimal partiell differenzierbar*. Analog definieren wir *$k$-mal partiell differenzierbar*.
(2) Wir sagen $f$ ist *$k$–mal stetig partiell differenzierbar*, wenn $f$ $k$–mal partiell differenzierbar ist, und alle partiellen Ableitungen
$$\frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_l}} f\colon U \to \mathbb{R}$$
mit $l \leq k$ stetig sind.

[25] $\nabla f$ wird 'Nabla $f$' gesprochen.

Wenn $f(x_1,x_2,x_3)=3x_1^2x_2+x_2^5+\sin(x_1x_3)$, dann ist beispielsweise $f$ beliebig oft stetig partiell differenzierbar, und es gilt

$$\begin{aligned}\tfrac{\partial}{\partial x_3}\tfrac{\partial}{\partial x_1}f &= \tfrac{\partial}{\partial x_3}(6x_1x_2+x_3\cos(x_1x_3))\\ &= 1\cdot\cos(x_1x_3)-x_3\cdot x_1\sin(x_1x_3), \text{ und}\\ \tfrac{\partial}{\partial x_1}\tfrac{\partial}{\partial x_3}f &= \tfrac{\partial}{\partial x_1}(x_1\cos(x_1x_3))\\ &= 1\cdot\cos(x_1x_3)-x_1\cdot x_3\sin(x_1x_3).\end{aligned}$$

Wir sehen also, dass

$$\frac{\partial}{\partial x_3}\frac{\partial}{\partial x_1}f=\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_3}f,$$

dies ist kein Zufall, sondern gilt allgemein für zweimal stetig partiell differenzierbare Funktionen:

**Satz 5.1.** *Es sei $U\subset\mathbb{R}^n$ offen und $f\colon U\to\mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Dann gilt für alle $a\in U$ und $i,j\in\{1,\dots,n\}$, dass*

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

Im Übungsblatt 5 werden wir sehen, dass die Aussage i.a. nicht für Funktionen gilt, welche zwar zweimal partiell differenzierbar sind, aber nicht zweimal stetig partiell differenzierbar sind.

*Beweis.* O.B.d.A. nehmen wir an, dass $U=\mathbb{R}^2, i=1, j=2$ und $a=(0,0)$. Sei also $f\colon\mathbb{R}^2\to\mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Wir wollen zeigen, dass

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

*Behauptung.* Es seien $x>0$ und $y>0$, dann existieren $\xi\in(0,x)$ und $\eta\in(0,y)$, so dass

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

Für gegebenes $y$ betrachten wir nun

$$\begin{aligned}F_y:\mathbb{R} &\to \mathbb{R}\\ s &\mapsto F_y(s):=f(s,y)-f(s,0).\end{aligned}$$

Nach dem Mittelwertsatz (Analysis I, Kapitel 13) [26] existiert nun ein $\xi\in(0,x)$, so dass

$$F_y'(\xi)=\frac{F_y(x)-F_y(0)}{x-0}.$$

[26] Der Mittelwertsatz kann in unserem Falle wie folgt formuliert werden:

Aus der Definition von der partiellen Ableitung folgt aber sofort, dass

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

Wir wenden jetzt den Mittelwertsatz auf die (nach Voraussetzung) differenzierbare Funktion

$$\begin{array}{rcl} [0, y] & \to & \mathbb{R} \\ t & \mapsto & \frac{\partial}{\partial x} f(\xi, t) \end{array}$$

an, und erhalten ein $\eta \in (0, y)$, so dass

$$\underbrace{\frac{\partial}{\partial y}\frac{\partial}{\partial x} f(\xi, \eta)}_{\substack{\text{Ableitung von} \\ t \mapsto \frac{\partial}{\partial x} f(\xi, t) \\ \text{am Punkt}\, y = \eta}} \quad = \frac{\frac{\partial}{\partial x} f(\xi, y) - \frac{\partial}{\partial x} f(\xi, 0)}{y}.$$

Kombinieren wir jetzt die obigen Gleichungen, dann erhalten wir, dass

$$\begin{array}{rcl} \frac{\partial}{\partial y}\ \frac{\partial}{\partial x} f(\xi, \eta) xy & = & \frac{1}{y} \left( \frac{\partial}{\partial x} f(\xi, y) - \frac{\partial}{\partial x} f(\xi, 0) \right) \\ & = & \frac{1}{y} F_y'(\xi) \\ & = & \frac{1}{y}\frac{1}{x} (F_y(x) - F_y(0)) \\ & = & \frac{1}{xy} \left( f(x, y) - f(x, 0) - f(0, y) + f(0, 0) \right). \end{array}$$

Wir haben damit die Behauptung bewiesen.

Wenn man die Rollen von $x$ und $y$ im obigen Beweis vertauscht (d.h. wir betrachten zuerst $G_x(t) := f(x, t) - f(0, t)$ etc.) dann erhalten wir den Beweis für folgende Aussage:

*Behauptung.* Es seien $x > 0$ und $y > 0$, dann existieren $\tilde{\xi} \in (0, x)$ und $\tilde{\eta} \in (0, y)$, so dass

$$f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = \frac{\partial}{\partial x}\ \frac{\partial}{\partial y} f(\tilde{\xi}, \tilde{\eta}) xy.$$

Vergleichen wir nun die Behauptungen erhalten wir folgende Aussage:

---

**Satz 5.2.** *(Mittelwertsatz der Differentialrechnung) Es sei $h\colon [a, b] \to \mathbb{R}$ eine differenzierbare Funktion. Dann gibt es ein $\xi \in (a, b)$, so dass*

$$h'(\xi) = \frac{h(b) - h(a)}{b - a}.$$

*Aussage.* Es seien $x > 0$ und $y > 0$, dann existieren $\xi \in (0, x)$ und $\eta \in (0, y)$ sowie $\tilde{\xi} \in (0, x)$ und $e\tilde{t}a \in (0, y)$, so dass

$$\frac{\partial}{\partial y}\frac{\partial}{\partial x} f(\xi, \eta) = \frac{\partial}{\partial x}\frac{\partial}{\partial y} f(\tilde{\xi}, \tilde{\eta}).$$

Nach Voraussetzung ist $f$ zweimal stetig partiell differenzierbar, d.h. die Funktionen

$$\frac{\partial}{\partial y}\frac{\partial}{\partial x} f \text{ und } \frac{\partial}{\partial x}\frac{\partial}{\partial y} f$$

sind stetig. Der Satz folgt also nun aus folgender Behauptung, welche in den Übungen bewiesen wird:

*Behauptung.* Es seien $A, B : \mathbb{R}^2 \to \mathbb{R}$ zwei stetige Funktionen. Machen wir folgende Annahme: für alle $x > 0$ und $y > 0$, existieren $\xi, \tilde{\xi} \in (0, x)$ und $\eta, \tilde{\eta} \in (0, y)$, so dass

$$A(\xi, \eta) = B(\tilde{\xi}, \tilde{\eta}).$$

Dann gilt $A(0, 0) = B(0, 0)$.

□

Es sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine $k$–mal stetig partiell differenzierbare Funktion. Es seien $i, j \in \{1, \dots, n\}$, dann schreiben wir

$$\begin{aligned} \tfrac{\partial^2}{\partial x_i \partial x_j} f &= \tfrac{\partial}{\partial x_i} \tfrac{\partial}{\partial x_j} f \\ \tfrac{\partial^2}{\partial x_i^2} f &= \tfrac{\partial}{\partial x_i} \tfrac{\partial}{\partial x_i} f. \end{aligned}$$

Allgemeiner, wenn $i_1, \dots, i_k \in \{1, \dots, n\}$, dann schreiben wir:

$$\frac{\partial^k}{\partial x_{i_1} \dots \partial x_{i_k}} f = \frac{\partial}{\partial x_{i_1}} \frac{\partial}{\partial x_{i_2}} \dots \frac{\partial}{\partial x_{i_k}} f.$$

Es folgt aus Satz 5.1, dass es egal ist, in welcher Reihenfolge wir die partiellen Ableitungen bilden, also gilt beispielsweise [27]

$$\begin{aligned} \tfrac{\partial^3}{\partial x\, \partial y\, \partial z} f(x, y, z) &= \tfrac{\partial^3}{\partial z\, \partial y\, \partial x} f(x, y, z), \text{ oder auch} \\ \tfrac{\partial^4}{\partial x^2\, \partial z\, \partial y} f(x, y, z) &= \tfrac{\partial^3}{\partial x\, \partial z\, \partial x\, \partial y} f(x, y, z). \end{aligned}$$

[27] Aus Satz 5.1 folgt, dass man partielle Ableitungen immer paarweise vertauschen kann, aber dadurch kann man auch jede Permutation der partiellen Ableitungen erhalten, z.B. gilt

$$\frac{\partial^3}{\partial x\, \partial y\, \partial z} f = \frac{\partial^3}{\partial y\, \partial x\, \partial z} f = \frac{\partial^3}{\partial y\, \partial z\, \partial x} f = \frac{\partial^3}{\partial z\, \partial y\, \partial x} f.$$

## 6. Totale Differenzierbarkeit

6.1. **Differenzierbarkeit und partielle Ableitungen.** Wir erinnern zuerst an die Definition des Grenzwertes einer Abbildung: Es sei $V \subset \mathbb{R}^n$ offen und es sei

$$\begin{array}{rcl} g\colon V & \to & \mathbb{R}^m \\ v & \mapsto & g(v) \end{array}$$

eine Abbildung, dann definieren wir, wie in Analysis I:

$$\lim_{v\to v_0} g(v) = w :\Leftrightarrow \underset{\varepsilon>0}{\forall}\ \underset{\delta>0}{\exists}\ \underset{v\in B_\delta(v_0)\setminus\{v_0\}}{\forall} \|f(v)-w\| < \varepsilon.$$

Es folgt sofort aus den Definition, dass

$$g \text{ stetig im Punkt } v_0 \ \Leftrightarrow \lim_{v\to v_0} g(v) = g(v_0).$$

Wir können nun den Begriff der differenzierbaren Abbildung einführen.

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^m$ eine Abbildung. Es sei $x \in U$. Wir sagen $f$ *ist im Punkt $x$ total differenzierbar* (oder kurz *differenzierbar*), wenn es eine $m \times n$–Matrix $A$ gibt, so dass

$$\lim_{v\to 0} \frac{f(x+v)-f(x)-Av}{\|v\|} = 0.$$

Wir nennen dann $A$ *Differential von $f$*. Wir sagen $f$ *ist differenzierbar*, wenn $f$ in allen Punkten differenzierbar ist.

*Bemerkung.*
1. Wenn $m = n = 1$, dann erhalten wir die gleiche Definition von Differenzierbarkeit wie in Analysis I (in diesem Fall ist dann $A$ die $1 \times 1$–Matrix, welche durch $f'(x)$ gegeben ist). [28]
2. Wenn $B$ eine $m \times n$–Matrix ist und $C \in \mathbb{R}^m$, dann nennen wir

$$\begin{array}{rcl} f\colon \mathbb{R}^n & \to & \mathbb{R}^m \\ x & \mapsto & Bx + C \end{array}$$

[28] Genauer gesagt, wenn $f\colon (a,b) \to \mathbb{R}$ eine Funktion ist, dann hatten wir in Analysis I definiert, dass $f$ im Punkt $x$ differenzierbar heißt, wenn der Grenzwert

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

existiert. Wenn nun also $f$ diese Definition aus Analysis I erfüllt, dann gilt auch, dass

$$\begin{array}{rcl} \lim_{h\to 0} \frac{f(x+h)-f(x)-f'(x)\cdot h}{h} & = & \lim_{h\to 0} \left(\frac{f(x+h)-f(x)}{h} - f'(x)\right) \\ & = & \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} - f'(x) = f'(x) - f'(x) = 0, \end{array}$$

d.h. $f$ erfüllt die obige Definition der Differenzierbarkeit wenn wir $A = (f'(x))$ setzen.

eine affine Abbildung. (Die Abbildung ist eine lineare Abbildung, wenn $C = 0$.) Eine affine Abbildung ist differenzierbar. In der Tat, wenn wir $A = B$ setzen, dann gilt

$$\begin{aligned} \lim_{v\to 0} \tfrac{f(x+v)-f(x)-Av}{\|v\|} &= \lim_{v\to 0} \tfrac{(B(x+v)+C)-(Bx+C)-Bv}{\|v\|} \\ &= \lim_{v\to 0} \tfrac{(Bx+Bv+C)-(Bx+C)-Bv}{\|v\|} \\ &= \lim_{v\to 0} \tfrac{0}{\|v\|} = 0. \end{aligned}$$

(3) Die Intuition ist wie folgt: $f$ heißt differenzierbar im Punkt $x_0$ mit Differential $A$ wenn man $f$ im Punkt $x_0$ durch die affine Abbildung $x \mapsto f(x_0) + A(x - x_0)$ 'approximieren' kann. Das Ziel der Analysis ist den vagen Begriff 'approximieren' mathematisch sauber zu definieren.

**Satz 6.1.** *Es sei $U \subset \mathbb{R}^n$ offen und*

$$f = (f_1, \dots, f_m) : U \to \mathbb{R}^m$$

*eine Abbildung, welche im Punkt $x \in U$ differenzierbar mit Differential $A = (a_{ij})$ ist. Dann gilt:*

(1) *$f$ ist stetig im Punkt $x$,*

(2) *für alle $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$ existiert die partielle Ableitung $\frac{\partial f_i}{\partial x_j}(x)$ und es gilt*

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

*Bemerkung.* (1) Man beachte, dass dieser Satz insbesondere besagt, dass wenn $f$ im Punkt $x$ differenzierbar ist, dann ist das Differential eindeutig bestimmt. Das Differential wird auch *Funktional–Matrix* oder *Jacobi–Matrix* genannt. Die folgenden Schreibweisen werden üblicherweise verwendet:

$$Df(x) := J_f(x) := \left( \frac{\partial f_i}{\partial x_j}(x) \right).$$

(2) Wenn $f : \mathbb{R}^n \to \mathbb{R}$ eine Funktion ist, d.h. wenn $m = 1$, dann ist das Differential an einem Punkt $x$ gerade die $1 \times n$–Matrix

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

Wenn man diese Matrix als Vektor auffasst, dann ist das Differential der Gradient am Punkt $x$.

*Beweis von Satz 6.1.* (1) Wenn

$$\lim_{v\to 0} \frac{f(x+v) - f(x) - Av}{\|v\|} = 0,$$

dann folgt insbesondere, dass

$$\lim_{v\to 0}\left(f(x+v)-f(x)-Av\right)=0,$$

daraus folgt dann, dass [29]

$$\lim_{v\to 0} f(x+v)=\lim_{v\to 0}(f(x)+Av)=f(x)+\lim_{v\to 0} Av=f(x).$$

(2) Wenn

$$\lim_{v\to 0}\frac{f(x+v)-f(x)-Av}{\|v\|}=0,$$

dann gilt insbesondere für alle $j\in\{1,\dots,n\}$, dass

$$\lim_{h\to 0}\frac{f(x+he_j)-f(x)-Ahe_j}{\|he_j\|}=0.$$

Es folgt wie in Satz 2.2, dass dann auch die Grenzwerte für die $m$ Koordinaten gegen 0 konvergieren. Die $i$–te Koordinate von $A\cdot he_j$ ist $ha_{ij}$, folgt also, dass

$$\lim_{h\to 0}\frac{f_i(x+he_j)-f_i(x)-ha_{ij}}{\|he_j\|}=0.$$

Nachdem $\|he_j\|=h$, folgt nun, dass

$$\lim_{h\to 0}\frac{f_i(x+he_j)-f_i(x)-ha_{ij}}{\|h\|}=0.$$

Nachdem der Grenzwert Null ist, ist der Absolutbetrag im Nenner irrelevant, und es folgt, dass sogar

$$\lim_{h\to 0}\frac{f_i(x+he_j)-f_i(x)-ha_{ij}}{h}=0,$$

also

$$\lim_{h\to 0}\frac{f_i(x+he_j)-f_i(x)}{h}=a_{ij},$$

aber die linke Seite ist ja gerade die Definition von

$$\frac{\partial f_i}{\partial x_j}(x).$$

□

Bevor wir den nächsten Satz beweisen, formulieren wir das folgende Lemma, welche in Übungsblatt 4 bewiesen wurde.

[29]Hier verwenden wir, dass nach Lemma 2.8 gilt:

$$\lim_{v\to 0} Av=A(\lim_{v\to 0} v)=A\cdot 0=0.$$

**Lemma 6.2.** *Wir betrachten folgende Normen auf $\mathbb{R}^n$:*

$$\begin{array}{rcl} \|(v_1,\dots,v_n)\|_1 & := & \max\{|v_1|,\dots,|v_n|\} \text{ (Maximumsnorm)} \\ \|(v_1,\dots,v_n)\|_2 & := & \sum_{i=1}^n |v_i| \\ \|(v_1,\dots,v_n)\|_3 & := & \sqrt{\sum_{i=1}^n v_i^2} \text{ (Euklidische Norm).} \end{array}$$

*Dann gilt für alle $v \in \mathbb{R}^n$, dass*

$$\begin{array}{rcccl} \|v\|_1 & \leq & \|v\|_2 & \leq & n \cdot \|v\|_1 \\ \|v\|_1 & \leq & \|v\|_3 & \leq & \sqrt{n} \cdot \|v\|_1. \end{array}$$

*Insbesondere gilt für alle $i, j \in \{1, 2, 3\}$, dass es $C, D > 0$ gibt, so dass*[30]

$$C \cdot \|v\|_i \leq \|v\|_j \leq D \cdot \|v\|_i.$$

Wir können jetzt folgenden wichtigen Satz beweisen:

**Satz 6.3.** *Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ eine stetig partiell differenzierbare Funktion. Dann ist $f$ total differenzierbar.*

Im Folgenden werden wir anstatt '$f$ ist stetig partiell differenzierbar' einfach nur sagen '$f$ ist stetig differenzierbar'.

*Beweis.* Wir müssen also zeigen, dass $f$ in jedem Punkt $x$ total differenzierbar ist. O.B.d.A. sei $U = \mathbb{R}^n$ und $x = 0$.

Wir setzen

$$D = \left( \frac{\partial f}{\partial x_1}(0), \dots, \frac{\partial f}{\partial x_n}(0) \right).$$

Wir wollen nun zeigen, dass

$$\lim_{v \to 0} \frac{f(v) - f(0) - Dv}{\|v\|} = 0.$$

Bevor wir uns dem Grenzwert explizit zuwenden, wollen wir zuerst den Ausdruck $f(v) - f(0) - Dv$ genauer studieren. Sei im Folgenden $\delta > 0$ beliebig und es sei $v = (v_1, \dots, v_n)$ ein beliebiger Vektor mit $\|v\| < \delta$. Wir schreiben

$$\begin{array}{rcl} w_0 & := & (0,\dots,0), \\ w_1 & := & (v_1, 0, \dots, 0), \\ w_2 & := & (v_1, v_2, 0, \dots, 0), \\ \vdots & & \\ w_n & := & (v_1, \dots, v_n). \end{array}$$

[30] Es gilt sogar folgende Verallgemeinerung: Es seien $\|-\|_1$ und $\|-\|_2$ zwei beliebige Normen auf einem endlich dimensionalen reellen Vektorraum, dann existieren $C, D > 0$, so dass

$$C \cdot \|v\|_1 \leq \|v\|_2 \leq D \cdot \|v\|_1.$$

Diese Aussage gilt i.a. aber nicht mehr, wenn man Normen auf unendlich dimensionalen Vektorräumen betrachtet.

Dann gilt offensichtlich[31], dass

$$f(v)-f(0)-Dv=\sum_{i=1}^{n}\left(f(w_i)-f(w_{i-1})\right)\;-\;\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(0)v_i. \tag{6.1}$$

*Behauptung.* Es sei $i\in\{1,\dots,n\}$. Dann gibt es ein $u_i\in B_\delta(0)$, so dass

$$f(w_i)-f(w_{i-1})=\frac{\partial f}{\partial x_i}(u_i)\cdot v_i.$$

Wir betrachten die Funktion

$$\begin{array}{rcl} F:[0,v_i] & \to & \mathbb{R} \\ t & \mapsto & f(v_1,\dots,v_{i-1},t,0,\dots,0).\end{array}$$

Dann gilt $F(0)=f(w_{i-1})$ und $F(v_i)=f(w_i)$. Aus dem Mittelwertsatz der Analysis I folgt, dass es ein $\xi\in(0,v_i)$ gibt, so dass

$$F'(\xi)=\frac{F(v_i)-F(0)}{v_i-0},$$

d.h. es gilt

$$\begin{array}{rcl} f(w_i)-f(w_{i-1}) & = & F(v_i)-F(0) \\ & = & v_i\cdot F'(\xi)=v_i\cdot\frac{\partial f}{\partial x_i}(v_1,\dots,v_{i-1},\xi,0,\dots,0).\end{array}$$

(Hier die letzte Gleichheit folgt sofort aus der Definition der partiellen Ableitung.) Aber nachdem $\|(v_1,\dots,v_n)\|<\delta$ und $\xi\in(0,v_i)$, folgt auch, dass

$$\|(v_1,\dots,v_{i-1},\xi,0,\dots,0)\|<\delta.$$

Es folgt, dass $u_i:=(v_1,\dots,v_{i-1},\xi,0,\dots,0)$ die gewünschte Eigenschaft besitzt. Wir haben damit die Behauptung bewiesen.

Kombinieren wir diese Behauptung mit der Gleichung (6.1), dann haben wir folgende Aussage bewiesen:

*Aussage.* Es sei $\delta>0$ und $v\in B_\delta(0)$. Dann existieren $u_1,\dots,u_n\in B_\delta(0)$, so dass

$$f(v)-f(0)-Dv=\sum_{i=1}^{n}v_i\left(\frac{\partial f}{\partial x_i}(u_i)-\frac{\partial f}{\partial x_i}(0)\right).$$

Wir wenden uns jetzt dem Grenzwert $\lim_{v\to 0}\frac{f(v)-f(0)-Dv}{\|v\|}$ zu, d.h. wir wollen jetzt zeigen, dass

$$\lim_{v\to 0}\frac{f(v)-f(0)-Dv}{\|v\|}=0.$$

[31]Man beachte, dass alle Terme $f(w_i), i=1,\dots,n-1$ zweimal mit umgekehrtem Vorzeichen vorkommen, und sich deswegen in der Summe wieder wegheben.

Sei also $\varepsilon > 0$. Nach Voraussetzung sind die partiellen Ableitungen stetig, d.h. es existiert ein $\delta > 0$, so dass für alle $i \in \{1, \ldots, n\}$ gilt:

$$u \in B_\delta(0) \;\Rightarrow\; \left|\frac{\partial f}{\partial x_i}(u) - \frac{\partial f}{\partial x_i}(0)\right| < \frac{\varepsilon}{n}$$

Dann gilt aber für $v \in B_\delta(0)$, dass

$$\begin{aligned} |f(v) - f(0) - Dv| &= \left|\sum_{i=1}^n v_i \left(\tfrac{\partial f}{\partial x_i}(u_i) - \tfrac{\partial f}{\partial x_i}(0)\right)\right| \\ &\leq \sum_{i=1}^n |v_i| \cdot \tfrac{\varepsilon}{n} \\ &\leq \tfrac{\varepsilon}{n} \sum_{i=1}^n |v_i| \leq \tfrac{\varepsilon}{n} \cdot n\|v\| = \varepsilon\|v\|. \end{aligned}$$

(In der letzten Ungleichung haben wir Lemma 6.2 verwendet). Es folgt, also, dass für alle $v \in B_\delta(0)$ gilt:

$$\left|\frac{f(v) - f(0) - Dv}{\|v\|}\right| < \varepsilon.$$

□

Aus Satz 6.1 und Satz 6.3 (angewandt auf die Koordinatenfunktionen $f_1, \ldots, f_m$) folgt nun folgender Satz:

**Satz 6.4.** *Es sei $U \subset \mathbb{R}^n$ offen und $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ eine Abbildung, so dass alle partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}$ definiert und stetig sind. Dann ist $f$ differenzierbar, und an jedem Punkt $P$ ist das Differential die $m \times n$–Matrix*

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(P) & \ldots & \frac{\partial f_1}{\partial x_n}(P) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(P) & & \frac{\partial f_m}{\partial x_n}(P) \end{pmatrix}.$$

*Bemerkung.* Ein Ziel der Analysis ist es das Studium von Funktionen auf Lineare Algebra zurückzuführen, denn lineare (oder affine) Abbildungen sind deutlich leichter zu verstehen, als ganz allgemeine Abbildungen zwischen $\mathbb{R}^n$ und $\mathbb{R}^m$. Wenn eine Abbildung $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ in einem Punkt $x$ differenzierbar ist, dann können wir die Abbildung in der Nähe von $x$ durch die affine Abbildung $v \mapsto f(x) + Df(x) \cdot (v - x)$ 'approximieren'. Unser Interesse besteht daher darin, das Differential $Df$ zu bestimmen. Satz 6.3 besagt nun, dass man das Differential mithilfe der partiellen Ableitungen von den Koordinatenfunktionen $f_i$ bestimmen kann.

*Definition.* Die Determinante der Jacobi–Matrix wird die *Funktionaldeterminante* genannt. Zur Erinnerung, für $2 \times 2$–Matrizen gilt

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc,$$

und für $3 \times 3$–Matrizen gilt

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}.$$

6.2. **Die Kettenregel.** Zur Erinnerung, wenn $f, g\colon \mathbb{R} \to \mathbb{R}$ differenzierbare Funktionen sind, dann besagt die Kettenregel, dass $g \circ f$ ebenfalls differenzierbar ist, und es gilt

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x).$$

D.h.

$$\begin{array}{rcl} \text{Ableitung von } g \circ f \text{ am Punkt } x & = & \text{Ableitung von } g \text{ am Punkt } f(x) \\ & & \text{mal Ableitung von } f \text{ am Punkt } x. \end{array}$$

Für differenzierbare Abbildungen zwischen Abbildungen in mehreren Variablen gilt eine ganz analoge Kettenregel:

**Satz 6.5.** ***(Kettenregel)*** *Es seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen und es seien*

$$f\colon U \to \mathbb{R}^m \text{ und } g\colon V \to \mathbb{R}^k$$

*differenzierbare Abbildungen, so dass $f(U) \subset V$. Dann ist die Abbildung*

$$g \circ f\colon U \to \mathbb{R}^k$$

*ebenfalls differenzierbar, und für jeden Punkt $x \in U$ gilt:*

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

Die Kettenregel besagt also, dass

$$\begin{array}{rcl} \text{Differential von } g \circ f \text{ am Punkt } x & = & \text{Differential von } g \text{ am Punkt } f(x) \\ & & \text{mal Differential von } f \text{ am Punkt } x. \end{array}$$

Beachten Sie, dass $(Dg)(f(x))$ eine $m \times k$–Matrix ist, und dass $Df(x)$ eine $m \times n$–Matrix ist, und dass das Matrixprodukt $(Dg)(f(x)) \cdot Df(x)$ eine $k \times n$–Matrix ist.

Für den Beweis verweise ich auf Forster, Analysis II, Seite 66. Folgende kurze ‘Rechnung’ soll dazu dienen, sich die Kettenregel herleiten zu können: Zuerst beachten wir, dass wenn $h$ im Punkt $y$ differenzierbar ist, dann gilt für ‘kurze’ Vektoren $u$, dass

$$h(y + u) \approx h(y) + Dh(y) \cdot u.$$

Nun sei $x \in U$ und $v \in \mathbb{R}^n$ ein 'kurzer' Vektor. Dann gilt

$$g(f(x+v)) \approx g(f(x) + \underbrace{Df(x) \cdot v}_{:=u}) \approx g(f(x)) + Dg(x) \cdot \underbrace{Df(x) \cdot v}_{=u}.$$

Nachdem aber auch gilt

$$g(f(x+v)) \approx g(f(x)) + D(g \circ f)(x) \cdot v$$

'folgt', dass

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

Der Beweis der Kettenregel besteht nun darin, dieses kurze Argument in formale Mathematik zu übersetzen.

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen und $f \colon U \to \mathbb{R}$ eine Funktion. Sei $x \in U$ und $v \in \mathbb{R}^n$ ein Vektor mit $\|v\| = 1$. Dann heißt der Grenzwert (falls er existiert)

$$D_v f(x) := \frac{d}{dt} f(x+tv)\big|_{t=0} = \lim_{h \to 0} \frac{f(x+hv) - f(x)}{h}$$

die *Richtungsableitung von* $f$ *im Punkt* $x$ *in Richtung* $v$.

Man beachte, dass für $v = e_i$, d.h. $v$ der $i$–te Einheitsvektor, direkt aus den Definitionen folgt, dass

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

Wir erinnern nun an das Skalarprodukt von zwei Vektoren $v = (v_1, \ldots, v_n)$ und $w = (w_1, \ldots, w_n)$ in $R^n$, welches wie folgt definiert ist:

$$\langle v, w \rangle := v \cdot w := \sum_{i=1}^{n} v_i \cdot w_i.$$

Es sei $\varphi$ der Winkel zwischen $v$ und $w$, dann gilt

$$\langle v, w \rangle = \cos(\varphi) \cdot \|v\| \cdot \|w\|.$$

Wir erinnern auch an den Gradienten einer Funktion $f \colon U \to \mathbb{R}$ am Punkt $x$, welcher wie folgt definiert ist:

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

**Satz 6.6.** *Es sei $U \subset \mathbb{R}^n$ offen und $f \colon U \to \mathbb{R}$ eine stetig differenzierbare Funktion. Sei $x \in U$ und $v \in \mathbb{R}^n$ ein Vektor mit $\|v\| = 1$. Dann gilt*

$$D_v f(x) = \langle v, \operatorname{grad} f(x) \rangle.$$

*Beweis.* O.B.d.A. sei $U = \mathbb{R}^n$. Wir betrachten

$$\begin{aligned} \varphi : \mathbb{R} &\to \mathbb{R}^n \\ t &\mapsto x + tv = (x_1 + tv_1, \ldots, x_n + tv_n). \end{aligned}$$

Per Definition gilt

$$D_v f(x) := \frac{d}{dt} f(x + tv)|_{t=0} = \frac{d}{dt} f(\varphi(t))|_{t=0}.$$

Dann folgt:

$$\begin{aligned} \tfrac{d}{dt} f(\varphi(t))|_{t=0} &= D(f \circ \varphi)(0) \\ &= (Df)(\varphi(0)) \cdot D\varphi(0) \\ &= \begin{pmatrix} \frac{\partial f}{\partial x_1}(\varphi(0)) & \ldots & \frac{\partial f}{\partial x_n}(\varphi(0)) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\ &= \textstyle\sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) \cdot v_i \\ &= \langle \operatorname{grad}(f)(x), v \rangle. \end{aligned}$$

Die Gleichheiten können wie folgt begründet werden:

(1) für Funktion $\mathbb{R} \to \mathbb{R}$ ist das Differential gerade die Ableitung,
(2) dies ist die Kettenregel,
(3) dies folgt aus der Berechnung von $Df$ mithilfe von Satz 6.3, und der Berechnung von $D\varphi$,
(4) dies ist die Definition vom Produkt der zwei Matrizen,
(5) dies folgt aus der Definition des Gradienten und der Definition des Skalarproduktes.

□

Es sei $\varphi \in [0, 2\pi)$ der Winkel zwischen $v$ und $\operatorname{grad} f(x)$. Dann gilt also, dass

$$D_v f(x) = \langle v, \operatorname{grad} f(x) \rangle = \cos(\varphi) \cdot \|v\| \cdot \| \operatorname{grad} f(x)\| = \cos(\varphi) \cdot \| \operatorname{grad} f(x)\|.$$

Dies ist maximal für $\varphi = 0$. Wir erhalten also folgendes Korollar:

**Korollar 6.7.** *Es sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine stetig differenzierbare Funktion und $x \in U$ ein Punkt, so dass* $\operatorname{grad} f(x) \neq 0$. *Dann ist*

$$v := \frac{\operatorname{grad} f(x)}{\| \operatorname{grad} f(x)\|}$$

*die Richtung mit maximaler Richtungsableitung.*

Anders ausgedrückt, 'der Gradient zeigt in die Richtung in welche $f$ am schnellsten ansteigt'.

6.3. **Der Mittelwertsatz und die Norm von Matrizen.** Wenn $f = (f_1, \ldots, f_m) : [a, b] \to \mathbb{R}^m$ eine stetige Abbildung ist, dann definieren wir

$$\int_a^b f(t)\,dt := \left( \int_a^b f_1(t)\,dt, \ldots, \int_a^b f_m(t)\,dt \right),$$

d.h. wir integrieren Komponentenweise. Ganz analog, in dem wir jeden Matrixeintrag separat integrieren, definieren wir das Integral einer stetigen Abbildung $f\colon [a, b] \to M(n \times m, \mathbb{R})$ in den Raum der $n \times m$–Matrizen.

Wir beweisen zuerst folgendes Lemma (welches man als eine Verallgemeinerung von Lemma 14.11 aus Analysis I auffassen kann.)

**Lemma 6.8.** *Es sei $v : [a, b] \to \mathbb{R}^m$ eine stetige Abbildung. Dann gilt*

$$\left\| \int_a^b v(t)\,dt \right\| \le \int_a^b \|v(t)\|\,dt \in \mathbb{R}^m.$$

*Beweis.* Wir setzen

$$u := \int_a^b v(t)\,dt \in \mathbb{R}^n.$$

Dann gilt

$$\begin{aligned}
\|u\|^2 = \langle u, u\rangle &= \left\langle \textstyle\int_a^b v(t)\,dt, u \right\rangle \\
&= \left\langle \left( \textstyle\int_a^b v_1(t)\,dt, \ldots, \int_a^b v_m(t)\,dt \right), (u_1, \ldots, u_m) \right\rangle \\
&= \textstyle\sum_{i=1}^m \int_a^b v_i(t)\,dt \cdot u_i \\
&= \textstyle\sum_{i=1}^m \int_a^b v_i(t) \cdot u_i\,dt \\
&= \textstyle\int_a^b \langle v(t), u\rangle\,dt \\
&\le \textstyle\int_a^b \|v(t)\| \cdot \|u\|\,dt \\
&= \|u\| \cdot \textstyle\int_a^b \|v(t)\|\,dt.
\end{aligned}$$

(Die Ungleichung folgt aus der Tatsache, dass für alle Vektoren $u$ und $v$ gilt: $\langle u, v\rangle = \cos(\varphi)\|u\| \cdot \|v\| \le \|u\| \cdot \|v\|$.) Das Lemma folgt nun indem wir durch $\|u\|$ kürzen. □

**Satz 6.9.** ***(Mittelwertsatz)*** *Es sei $U \subset \mathbb{R}^n$ offen und $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ eine stetig differenzierbare Abbildung. Es sei $x \in U$ und $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ ein Vektor, so dass $x + tv \in U$ für alle $t \in [0, 1]$.*

*Dann gilt* [32]

$$f(x+v)-f(x)=\left(\int_0^1 Df(x+tv)\,dt\right)\cdot v.$$

*Beweis.* Wir beweisen den Mittelwertsatz koordinatenweise. Für $i=1,\dots,m$ betrachten wir

$$\begin{array}{rcl} g_i:[0,1] & \to & \mathbb{R} \\ t & \mapsto & g_i(t):=f_i(x+tv). \end{array}$$

Man beachte, dass aus der Kettenregel[33] folgt, dass

$$\begin{array}{rcl} g_i(t)' & = & Dg_i(t)=Df_i(x+tv)\begin{pmatrix} v_1 \\ \vdots \\ v_n\end{pmatrix} \\ & = & \begin{pmatrix} \frac{\partial f_i}{\partial x_1}(x+tv) & \dots & \frac{\partial f_i}{\partial x_n}(x+tv)\end{pmatrix}\begin{pmatrix} v_1 \\ \vdots \\ v_n\end{pmatrix} \\ & = & \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x+tv)v_j. \end{array}$$

Dann folgt aus dem Hauptsatz der Differential- und Integralrechnung aus Analysis I, dass

$$\begin{array}{rcl} f_i(x+v)-f_i(x) & = & g_i(1)-g_i(0) \\ & = & \int_0^1 g_i'(t)\,dt \\ & = & \int_0^1 \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x+tv)v_j\ \ dt \\ & = & \sum_{j=1}^n \int_0^1 \frac{\partial f_i}{\partial x_j}(x+tv)v_j\ \ dt. \end{array}$$

---

[32] Man beachte, dass die rechte Seite das Produkt einer $m\times n$–Matrix mit einem Vektor in $\mathbb{R}^n$ ist.

[33] Das Differential der Funktion

$$\begin{array}{rcl} g_i:[0,1] & \to & \mathbb{R} \\ t & \mapsto & g_i(t):=f_i(x+tv) \end{array}$$

ist das Produkt des Differentials der Abbildung $f_i$ am Punkt $x+tv$ mit dem Differential der Abbildung $t\mapsto x+tv$. Das Differential der letzteren Abbildung ist gerade der Vektor $v$. Siehe auch den Beweis von Satz 6.6 für ein ganz ähnliches Argument.

Aber das ist gerade die $i$–te Zeile von dem Vektor [34]

$$\left(\int_0^1 Df(x+tv)\,dt\right)\cdot v \in \mathbb{R}^m.$$

$\square$

Es sei nun $A \in M(m \times n, \mathbb{R})$ eine $m \times n$–Matrix. Wir definieren [35]

$$\|A\| := \sup\{\|Av\| \mid v \in \mathbb{R}^n \text{ mit } \|v\| = 1\}.$$

Dies ist in der Tat eine Norm auf dem reellen Vektorraum $M(m \times n, \mathbb{R})$. (Siehe Forster, Analysis II, Seite 23.) Man beachte, dass für alle Vektoren $w \in \mathbb{R}^n$ gilt:[36]

$$\|Aw\| \leq \|A\| \cdot \|w\|.$$

*Bemerkung.* Wir hatten in Kapitel 1.2 auch schon folgende Norm für $n \times m$–Matrizen eingeführt:

$$\|(a_{ij})\|' := \sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}|.$$

Es gilt, dass

$$\|A\|' \leq \|A\| \leq \sqrt{nm}\|A\|'.$$

(Siehe Forster, Analysis II, Seite 23 und siehe auch Fußnote 30.) Die Norm $\|A\|'$ ist relativ leicht zu berechnen, während die Norm $\|A\|$ zwar schwieriger zu berechnen ist, aber dafür für Anwendungen deutlich wichtiger ist (siehe den nächsten Satz). Es ist deshalb hilfreich, $\|A\|$ durch $\|A\|'$ abschätzen zu können.

Wir erhalten jetzt folgendes Korollar zu den obigen Ergebnissen:

**Korollar 6.10.** *Es sei $U \subset \mathbb{R}^n$ offen und $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ eine stetig differenzierbare Abbildung. Es sei $x \in U$ und $v \in \mathbb{R}^n$ ein Vektor, so dass $x + tv \in U$ für alle $t \in [0,1]$. Wir setzen*

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

---

[34] In der Tat, denn $\int_0^1 Df(x+tv)\,dt$ ist per Definition die Matrix

$$\begin{pmatrix} \int_0^1 \frac{\partial f_1}{\partial x_1}(x+tv)\,dt & \ldots & \int_0^1 \frac{\partial f_1}{\partial x_n}(x+tv)\,dt \\ \vdots & & \vdots \\ \int_0^1 \frac{\partial f_m}{\partial x_1}(x+tv)\,dt & \ldots & \int_0^1 \frac{\partial f_m}{\partial x_n}(x+tv)\,dt \end{pmatrix}.$$

[35] Hier ist $\|v\|$ die euklidische Norm von $v \in \mathbb{R}^n$ und $\|Av\|$ ist die euklidische Norm von $Av \in \mathbb{R}^m$.

[36] Dies zeigt man, indem man $w = \|w\| \cdot \frac{1}{\|w\|} w$ schreibt und die Definition von $\|A\|$ auf $v = \frac{1}{\|w\|} w$ anwendet.

*Dann gilt*

$$\|f(x+v)-f(x)\| \le M\|v\|.$$

*Beweis.* Aus dem Mittelwertsatz und aus Lemma 6.8 folgt, dass

$$\begin{aligned}
\|f(x+v)-f(x)\| &= \left\|\left(\int_0^1 Df(x+tv)\,dt\right)\cdot v\right\| \\
&= \left\|\int_0^1 Df(x+tv)\cdot v\,dt\right\| \\
&\le \int_0^1 \|Df(x+tv)\cdot v\|\,dt \\
&\le \int_0^1 \|Df(x+tv)\|\cdot\|v\|\,dt \\
&\le \int_0^1 M\cdot\|v\|\,dt \\
&= M\cdot\|v\|.
\end{aligned}$$

□

## 7. Der Satz über implizite Funktionen

### 7.1. **Der Banachsche Fixpunktsatz.**

*Definition.*

(1) Ein metrischer Raum heißt *vollständig*, wenn jede Cauchy–Folge konvergiert.
(2) Ein normierter Vektorraum heißt vollständig, wenn der dazugehörige metrische Raum[37] vollständig ist.
(3) Ein vollständiger normierter Vektorraum heißt *Banachraum.*

*Beispiel.*

(1) Der Vektorraum $\mathbb{R}^n$ mit der euklidischen Norm ist vollständig.
(2) In Kapitel 2.1 hatten wir gesehen, dass der normierte Vektorraum

$$\begin{aligned}
C([-1,1],\mathbb{R}) &:= \text{alle stetigen Funktionen } [-1,1]\to\mathbb{R}, \\
\|f\| &:= \int_{-1}^1 |f(x)|\,dx, \text{ (Integralnorm)}
\end{aligned}$$

nicht vollständig ist. In der Tat, wenn wir die Folge von Funktionen

$$\begin{aligned}
f_n : [-1,1] &\to \mathbb{R} \\
x &\mapsto \begin{cases} 0, & \text{wenn } x\in[-1,0), \\ nx, & \text{wenn } x\in[0,\frac{1}{n}), \\ 1, & \text{wenn } x\in[\frac{1}{n},1]. \end{cases}
\end{aligned}$$

betrachten, dann gilt für $n \ge m$, dass

$$\|f_n - f_m\| = \int_{-1}^1 |f_n(x)-f_m(x)|\,dx \le \frac{1}{m}.$$

---

[37] Zur Erinnerung: Wenn $(V,\|-\|)$ ein normierter Vektorraum ist, dann definiert $d(x,y) := \|x-y\|$ eine Metrik auf $V$.

Es folgt, dass $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, aber diese konvergiert nicht in $C([-1,1],\mathbb{R})$, denn die Grenzfunktion

$$f(x) = \begin{cases} 0, & \text{wenn } x \in [-1,0], \\ 1, & \text{wenn } x \in (0,1]. \end{cases}$$

ist nicht stetig, liegt also nicht in $C([-1,1],\mathbb{R})$.

Der folgende Satz gibt ein wichtiges Beispiel für einen unendlich dimensionalen Banachraum:

**Satz 7.1.** *Es sei $U \subset \mathbb{R}^n$ eine Teilmenge. Wir betrachten den Vektorraum*

$$\begin{aligned} C_b(U,\mathbb{R}^m) \;:=\; & \textit{Vektorraum aller beschränkten} \\ & \textit{stetigen Abbildungen } f\colon U \to \mathbb{R}^m, \end{aligned}$$

*mit der Supremumsnorm, d.h.*

$$\|f\| := \sup\{\|f(x)\| \,|\, x \in U\}.$$

*Dann ist $(C_b(U,\mathbb{R}^m), \|-\|)$ ein Banachraum.*

Der Satz wird für den Fall $m = 1$ in den Übungen bewiesen. Den allgemeinen Fall für beliebiges $m$ zeigt man fast genauso.

*Bemerkung.* Betrachten wir den Fall $U = [-1,1] \subset \mathbb{R}$ und $m = 1$, dann ist jede stetige Funktion auf $[-1,1]$ beschränkt, also gilt $C([-1,1],\mathbb{R}) = C_b([-1,1],\mathbb{R})$, d.h. es folgt aus Satz 7.1, dass

$$\begin{aligned} C([-1,1],\mathbb{R}) &:= \text{alle stetigen Funktionen } [-1,1] \to \mathbb{R}, \\ \|f\| &:= \sup\{\|f(x)\| \,|\, x \in U\} \text{ (Supremumsnorm)}, \end{aligned}$$

vollständig ist. Der Unterschied zum Beispiel als wir $C([-1,1],\mathbb{R})$ mit der Integralnorm betrachtet haben ist folgender: eine Folge von $g_n \in C([-1,1],\mathbb{R})$ welche bzgl. der Integralnorm eine Cauchyfolge bildet, ist nicht notwendigerweise eine Cauchyfolge bzgl. der Supremumsnorm. Wenn man beispielsweise $f_n$ wie oben definiert ist, dann gilt für alle $n$, dass

$$\|f_n - f_{2n}\| \sup\{\|f_n(x) - f_m(x)\| \,|\, x \in U\} = \frac{1}{2},$$

insbesondere bilden die $f_n$ keine Cauchyfolge bzgl. der Supremumsnorm.

Bevor wir den nächsten Satz formulieren führen wir noch zwei Begriffe ein:

(1) Es sei $(X,d)$ ein metrischer Raum, eine Abbildung $\Phi : X \to X$ heißt *Kontraktion* wenn es eine Konstante $\Theta \in [0,1)$ (genannt Kontraktionsfaktor) gibt, so dass

$$\|\Phi(x) - \Phi(y)| \leq \Theta\|x - y\| \text{ für alle } x, y \in X.$$

(2) Es sei $\Phi : X \to X$ eine Abbildung. Ein Punkt $x \in X$ heißt *Fixpunkt von* $\Phi$, wenn $\Phi(x) = x$.

*Beispiel.* (1) Betrachten wir $f \colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$, dann sind $x = 0$ und $x = 1$ die einzigen Fixpunkte.

(2) Es sei $P \in \mathbb{R}^n$ ein Punkt. Dann ist die Abbildung

$$\begin{array}{rcl} \Phi : \mathbb{R}^n & \to & \mathbb{R}^n \\ x & \mapsto & P + \frac{1}{2}(x - P) \end{array}$$

eine Kontraktion, und der Punkt $P$ ist (der einzige) Fixpunkt von $\Phi$. Geometrisch bildet $\Phi$ jeden Punkt $x$ auf den Mittelpunkt von $x$ und $P$ ab. Anders ausgedrückt, die Distanz von $f(x)$ zu $P$ ist halb so großwie die Distanz von $x$ zu $P$.

(3) Die Abbildung

$$\begin{array}{rcl} \Phi : \mathbb{R}^n \setminus \{0\} & \to & \mathbb{R}^n \setminus \{0\} \\ x & \mapsto & \frac{1}{3}x \end{array}$$

ist eine Kontraktion, aber die Abbildung besitzt keinen Fixpunkt.

Wir können jetzt den folgenden wichtigen Satz formulieren:

**Satz 7.2. (Banachscher Fixpunktsatz)** *Es sei $A$ eine abgeschlossene Teilmenge eines Banachraumes $(V, \| - \|)$ und es sei $\Phi : A \to A$ eine Kontraktion. Dann besitzt $\Phi$ genau einen Fixpunkt.*

Dieser unscheinbare Satz spielt eine wichtige Rolle in der Mathematik, wir werden ihn auch später wieder treffen, wenn wir die Lösungen von Differentialgleichungen studieren.

*Beweis.* Es sei $\Theta \in [0, 1)$ Kontraktionsfaktor von $\Phi$. Wir zeigen zuerst, dass $\Phi$ höchstens einen Fixpunkt besitzt: Es seien also $v, w \in V$ mit $\Phi(v) = \Phi(w)$, dann gilt

$$\|v - w\| = \|\Phi(v) - \Phi(w)\| < \Theta \cdot \|v - w\|.$$

Nachdem $\Theta \in [0, 1)$ ist dies nur möglich, wenn $\|v - w\| = 0$, d.h. wenn $v = w$.

Wir zeigen nun, dass $\Phi$ einen Fixpunkt besitzt. Genauer gesagt, wir werden folgende Behauptung beweisen:

*Behauptung.* Es sei $v_0 \in A$. Dann konvergiert die durch $v_k := \Phi(v_{k-1})$ rekursiv definiert Folge gegen einen Fixpunkt von $\Phi$. [38]

[38] Dies sieht man auch sehr gut an Beispiel (2) oben, in diesem Fall konvergiert für jedes $x \in \mathbb{R}^2$ die Folge von Punkten $x, \Phi(x), \Phi(\Phi(x)), \ldots$ offensichtlich gegen den Fixpunkt $P$.

Wir setzen $c := \|v_1 - v_0\|$. Für $i \geq 1$ gilt dann:

$$\begin{array}{rcl} \|v_{i+1} - v_i\| & = & \|\Phi(v_i) - \Phi(v_{i-1})\| \\ & \leq & \Theta \cdot \|v_i - v_{i-1}\| \\ & \leq & \Theta^2 \cdot \|v_{i-1} - v_{i-2}\| \dots \\ & \leq & \Theta^i \cdot \|v_1 - v_0\| = \Theta^i \cdot c. \end{array}$$

Für beliebige $m > k \in \mathbb{N}$ folgt also

$$\begin{array}{rcl} \|v_m - v_k\| & = & \left\| \sum_{i=k}^{m-1} (v_{i+1} - v_i) \right\| \\ & \leq & \sum_{i=k}^{m-1} \|v_{i+1} - v_i\| \\ & \leq & \sum_{i=k}^{m-1} \Theta^i c \\ & = & c\Theta^k \sum_{i=0}^{m-1-k} \Theta^i \\ & < & c\Theta^k \sum_{i=0}^{\infty} \Theta^i = c\Theta^k \frac{1}{1-\Theta}. \end{array}$$

Nachdem $c\Theta^k \frac{1}{1-\Theta} = \frac{c}{1-\Theta} \cdot \Theta^k$ sieht man leicht, dass $(v_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $V$ ist. [39] Nachdem $V$ vollständig ist, konvergiert die Folge gegen ein $v \in V$. Nachdem $A$ abgeschlossen ist, und nachdem alle Folgenglieder $v_n$ in $A$ liegen, folgt, dass $v \in A$.

Wir wollen nun zeigen, dass $v$ ein Fixpunkt von $\Phi$ ist. Jede Kontraktion ist stetig (siehe Übungsblatt 6), daher gilt

$$\Phi(v) = \Phi(\lim_{n\to\infty} v_n) = \lim_{n\to\infty} \Phi(v_n) = \lim_{n\to\infty} v_{n+1} = v.$$

□

7.2. **Satz über implizite Funktionen.** Bevor wir den Satz über implizite Funktionen formulieren, erinnern wir noch an zwei Definitionen:

(1) Es sei $x \in \mathbb{R}^n$, dann heißt $U \subset \mathbb{R}^n$ *offene Umgebung von* $x$, wenn $U$ offen ist, und wenn $x \in U$.
(2) Es sei $x \in \mathbb{R}^n$, dann heißt $U \subset \mathbb{R}^n$ *Umgebung von* $x$, wenn $U$ eine offene Menge $V$ enthält, so dass $x \in V$.

Beispielsweise ist

$$U := \{y \in \mathbb{R}^n \mid \|y\| \leq r\}$$

für $r > 0$ eine Umgebung von $x = 0$ (wir könnten beispielsweise $V = B_r(0) = \{y \in \mathbb{R}^n \mid \|y\| < r\}$ wählen), aber keine offene Umgebung von $x = 0$.

Der folgende Satz ist wahrscheinlich der auf den ersten Blick unverdaulichste Satz der Analysis II:

[39] In der Tat, für gegebenes $\varepsilon > 0$ existiert ein $N$, so dass $\frac{c}{1-\Theta} \cdot \Theta^N < \varepsilon$. Für alle $k, m \geq N$ folgt dann: $\|v_m - v_k\| < \frac{c}{1-\Theta} \cdot \Theta^N < \varepsilon$.

**Satz 7.3. (Satz über implizite Funktionen)** *Es seien* $U_x \subset \mathbb{R}^k$ *und* $U_y \subset \mathbb{R}^m$ *offene Teilmengen und es sei*

$$\begin{array}{rcl} F : U_x \times U_y & \to & \mathbb{R}^m \\ (x,y) & \mapsto & F(x,y) \end{array}$$

*eine stetig differenzierbare Abbildung. Es sei* $(a,b) \in U_x \times U_y$ *ein Punkt mit den folgenden Eigenschaften:*

(1) $F(a,b) = 0$, *und*
(2) *die* $m \times m$*–Matrix*

$$\frac{\partial F}{\partial y} := \frac{\partial(F_1,\dots,F_m)}{\partial(y_1,\dots,y_m)} := \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}$$

*ist im Punkt* $(a,b)$ *invertierbar.*

*Dann existieren*

(1) *eine offene Umgebungen* $V_x \subset U_x$ *von* $a$,
(2) *eine Umgebung* $V_y \subset U_y$ *von* $b$, *sowie*
(3) *eine stetig differenzierbare Abbildung* $g \colon V_x \to V_y$ *mit* $g(a) = b$,

*so dass Folgendes gilt:*

(1) *für alle* $x \in V_x$ *ist*

$$F(x, g(x)) = 0,$$

(2) *und für alle* $(x,y) \in V_x \times V_y$ *mit* $F(x,y) = 0$ *gilt, dass* $y = g(x)$.

*Bemerkung.* Etwas salopp ausgedrückt sagt der Satz Folgendes: Es sei $(a,b)$ ein Punkt ist, welcher die Gleichung $F(x,y) = 0$ erfüllt, und so dass die Matrix $\frac{\partial F}{\partial y}$ am Punkt $(a,b)$ invertierbar ist. Dann kann man eine kleine Umgebung um den Punkt $(a,b)$ finden, so dass die $y$–Koordinaten von den Lösungen der Gleichung $F(x,y) = 0$ eine Funktion der $x$–Koordinaten sind.
Anders ausgedrückt, zu jedem $x \in V_x$ gibt es genau ein $y = g(x) \in V_y$, so dass $F(x,y) = 0$.

Betrachten wir den Fall $k = m = 1$, $U_x = \mathbb{R}$ und $U_y = \mathbb{R}$ und die Funktion

$$\begin{array}{rcl} F : \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (x,y) & \mapsto & x^2 + y^2 - 4. \end{array}$$

Dann gilt

$$\frac{\partial F}{\partial y}(a,b) = 2b.$$

Betrachten wir den Punkt $(a,b) = (0,2)$. Dann gilt $F(0,2) = 0$ und am Punkt $(0,2)$ ist $\frac{\partial F}{\partial y} = (4)$, insbesondere eine invertierbare $1 \times 1$–Matrix.

Die Bedingungen des Satzes sind also erfüllt. Dann hat $V_x = (-1,1)$ und $V_y = (1,3)$ und

$$\begin{array}{rcl} g\colon V_x & \to & V_y \\ x & \mapsto & \sqrt{4-x^2} \end{array}$$

die gewünschte Eigenschaft:

(1) für alle $x \in (-1,1)$ gilt
$$F(x,g(x)) = F(x,\sqrt{4-x^2}) = x^2 + (\sqrt{4-x^2})^2 - 4 = 0,$$
(2) wenn $(x,y) \in V_x \times V_y$ die Gleichung $F(x,y) = 0$ erfüllt, dann gilt auch $y = \sqrt{4-x^2}$.

Beachte, dass wenn wir $V_y = \mathbb{R}$ gewählt hätten, dann würde die zweite Bedingung nicht mehr gelten: dann wäre die $y$–Koordinate von einer Lösung zu $F(x,y) = 0$ in $V_x \times \mathbb{R}$ nicht mehr eindeutig durch die $x$–Koordinate bestimmt. Beispielsweise erfüllen $(0,-2)$ und $(0,2)$ die Gleichung $F(x,y) = 0$.

*Beweis des Satzes über implizite Funktionen.* O.B.d.A. sei $(a,b) = (0,0)$. Wir setzen
$$B := \frac{\partial F}{\partial y}(0,0) \in \mathrm{GL}(m,\mathbb{R})$$
und betrachten
$$\begin{array}{rcl} G : U_x \times U_y & \to & \mathbb{R}^m \\ (x,y) & \mapsto & y - B^{-1}F(x,y). \end{array}$$
Man beachte, dass
$$F(x,y) = 0 \Leftrightarrow y = G(x,y). \tag{7.1}$$
Für fest gewähltes $x$ haben wir also das Studium der Lösungen von $F(x,y) = 0$ in ein Problem über Fixpunkte der Abbildung $y \mapsto G(x,y)$ übergeführt. Folgende Behauptung erlaubt es uns den Banachschen Fixpunktsatz anwenden zu können.

*Behauptung.* Es gibt eine offene Umgebung $V_x \subset U_x$ von $0 \in \mathbb{R}^k$ und eine Umgebung $V_y \subset U_y$ um $0 \in \mathbb{R}^m$ mit folgenden Eigenschaften:

(1) $V_y$ ist eine abgeschlossene Teilmenge von $\mathbb{R}^m$,
(2) für alle $x \in V_x, y \in V_y$ gilt: $G(x,y) \in V_y$,
(3) für $x \in V_x$ ist die Abbildung
$$\begin{array}{rcl} V_y & \to & V_y \\ y & \mapsto & G(x,y) \end{array}$$
kontrahierend.

Für den Beweis der Behauptung bestimmen wir zuerst, dass

$$\frac{\partial G}{\partial y}(x,y) = \mathrm{id}_m - B^{-1}\frac{\partial F}{\partial y},$$

insbesondere folgt, dass

$$\frac{\partial G}{\partial y}(0,0) = \mathrm{id}_m - B^{-1}\cdot\frac{\partial F}{\partial y}(0,0) = 0.$$

Nachdem $F$ stetig differenzierbar ist, sind auch alle Komponenten von $\frac{\partial G}{\partial y}$ stetig, es gibt also Umgebungen $W_x \subset U_x$ von $0 \in R^k$ und $W_y \subset U_y$ von $0 \in \mathbb{R}^m$, so dass [40]

$$\left\|\frac{\partial G}{\partial y}(x,y)\right\| \le \frac{1}{2} \text{ für alle } (x,y) \in W_x \times W_y. \tag{7.2}$$

Wir wählen nun ein $r > 0$, so dass

$$V_y := \{y \in \mathbb{R}^m \mid \|y\| \le r\} \subset W_y.$$

Nachdem $G(0,0) = 0$, gibt es auch eine offene Umgebung $V_x \subset W_x$ von $0 \in \mathbb{R}^k$, so dass

$$\sup_{x\in V_x} \|G(x,0)\| \le \frac{r}{2}. \tag{7.3}$$

Wir wollen jetzt zeigen, dass $V_x$ und $V_y$ die gewünschten Eigenschaften haben. Es ist klar, dass $V_y$ abgeschlossen ist. Sei nun $x \in V_x$ fest gewählt. Aus der Abschätzung (7.2) und aus Korollar 6.10 folgt, dass für alle $x \in V_x$ und $y, y' \in V_y$ gilt:

$$\|G(x,y) - G(x,y')\| \le \frac{1}{2}\|y - y'\|. \tag{7.4}$$

[40] Hier betrachten wir wieder folgende Norm von $m \times m$–Matrizen:

$$\|A\| := \sup\{\|Av\| \; |v \in \mathbb{R}^m \text{ mit } \|v\| = 1\}.$$

Für eine Matrix $A$ hatten wir auch folgende Norm eingeführt:

$$\|(a_{ij})\|' := \sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}|.$$

und darauf hingewiesen, dass

$$\|A\|' \le \|A\| \le \sqrt{nm}\|A\|'.$$

Nachdem die partiellen Ableitungen $\frac{\partial F_i}{\partial x_j}$ Null sind am Punkt $(0,0)$, können wir Umgebungen $W_x, W_y$ finden, so dass

$$\left\|\frac{\partial G}{\partial y}(x,y)\right\|' \le \frac{1}{2}\frac{1}{m} \text{ für alle } (x,y) \in W_x \times W_y..$$

Angewandt auf $y' = 0$ und unter Verwendung von (7.3) erhalten wir, dass

$$\begin{array}{lrcl} (7.5) & \|G(x,y)\| & = & \|G(x,y) - G(x,0) + G(x,0)\| \\ & & \leq & \|G(x,y) - G(x,0)\| + \|G(x,0)\|\frac{1}{2}\|y\| + \frac{r}{2} \leq r \end{array}$$

für alle $\|y\| \leq r$. Es folgt, dass $\|G(x,y) \in V_y$ für alle $V_y$. Aus (7.4) folgt zudem, dass die Abbildung $y \mapsto G(x,y)$ kontrahierend ist. Wir haben damit die Behauptung bewiesen.

Es sei $x \in V_x$. Nachdem $V_y$ eine abgeschlossene Kugel im Banachraum $\mathbb{R}^m$ ist, können wir den Banachschen Fixpunktsatz auf die kontrahierende Abbildung

$$\begin{array}{rcl} V_y & \to & V_y \\ y & \mapsto & G(x,y) \end{array}$$

anwenden. Es gibt als für gegebenes $x \in V_x$ genau ein $y \in V_y$, so dass $G(x,y) = 0$. Wir bezeichnen dieses eindeutig bestimmte Element mit $g(x)$. Aus (7.1) folgt nun, dass

$$F(x, g(x)) = 0 \text{ für alle } x \in V_x,$$

und ferner gilt:

$$F(x,y) = 0 \text{ und } x \in V_x, y \in V_y \;\Rightarrow\; y = g(x).$$

(Dies ist eine Umformulierung der obigen Aussage, dass es für gegebenes $x \in V_x$ *genau* ein $y \in V_y$ gibt, so dass $G(x,y) = 0$.) Zusammenfassend haben wir gezeigt, dass es genau eine Abbildung $V_x \to V_y$ gibt, mit der Eigenschaft, dass $F(x, g(x)) = 0$ für alle $x \in V_x$.

Wir müssen nun noch zeigen, dass $g\colon V_x \to V_y$ stetig differenzierbar ist. Wir werden zeigen, dass $g$ stetig ist, für den Beweis der Differenzierbarkeit verweise ich auf Forster, Analysis II, Seite 93.

Wir werden nun den Banachschen Fixpunktsatz erneut geschickt anwenden, um zu beweisen, dass die Abbildung $g$ stetig ist. Aus Satz 7.1 wissen wir, dass

$$\begin{array}{rcl} C_b(V_x, \mathbb{R}^m) & := & \text{Vektorraum aller beschränkten} \\ & & \text{stetigen Abbildungen } f\colon V_x \to \mathbb{R}^m, \end{array}$$

mit der Supremumsnorm, d.h. mit

$$\|f\| := \sup\{\|f(x)\| \,|\, x \in V_x\}$$

ein Banachraum. Wir betrachten die abgeschlossene [41] Teilmenge

$$A := \{f \in C_b(V_x, \mathbb{R}^m) \,|\, \|f\| \le r\} = \{f \in C_b(V_x, \mathbb{R}^m) \,|\, f(V_x) \subset V_y\}$$

und die Abbildung

$$\begin{array}{rcl} \Phi : A & \to & A \\ f & \mapsto & (x \mapsto G(x, f(x))). \end{array}$$

Aus (7.5) folgt, dass $\|G(x, f(x))\| \le r$ für alle $x \in V_x$, d.h. die Abbildung $x \mapsto G(x, f(x))$ liegt in der Tat in $A$. Aus (7.4) folgt, dass für alle $f_1, f_2 \in A$ gilt:

$$\begin{array}{rcl} \|\Phi(f_1) - \Phi(f_2)\| & = & \sup_{x \in V_x} \|G(x, f_1(x)) - G(x, f_2(x))\| \\ & \le & \frac{1}{2} \sup_{x \in V_x} \|f_1(x) - f_2(x)\| \\ & = & \frac{1}{2} \|f_1 - f_2\|. \end{array}$$

Wir haben nun also bewiesen, dass die Abbildung $\Phi : A \to A$ kontrahierend ist. Aus dem Banachschen Fixpunktsatz folgt, dass $\Phi$ *genau* einen Fixpunkt besitzt, d.h. es gibt genau ein $h \in A \subset C_b(V_x, \mathbb{R}^m)$, so dass $\Phi(h) = h$, d.h. so dass $G(x, h(x)) = h(x)$. Anders ausgedrückt, es gibt genau ein $h \in A \subset C_b(V_x, \mathbb{R}^m)$, so dass $F(x, h(x)) = 0$ für alle $x \in V_x$.

Wir hatten weiter oben schon gezeigt, dass es genau eine Abbildung $V_x \to V_y$ gibt, mit der Eigenschaft, dass $F(x, g(x)) = 0$ für alle $x \in V_x$. Nachdem $h$ ebenfalls diese Eigenschaft hat, folgt, dass $g = h$, insbesondere ist $g$ stetig. [42] □

---

[41] Hier ist der Beweis, dass

$$A = \{f \in C_b(V_x, \mathbb{R}^m) \,|\, f(V_x) \subset V_y\} = \{f \in C_b(V_x, \mathbb{R}^m) \,|\, \|f(x)\| \le r \text{ für alle } x \in V_x\}$$

im normierten Vektorraum $(C_b(V_x, \mathbb{R}^m), \| - \|)$ abgeschlossen ist: Wir zeigen, dass $C_b(V_x, \mathbb{R}^m) \setminus A$ offen ist. Sei $g \notin A$. Dann gilt $\sup\{\|f(x)\| \,|\, x \in V_x\} > r$. Insbesondere existiert ein $x \in V_x$, so dass $\|f(x)\| > r$. Wir setzen $s := \|f(x)\|$ und $\varepsilon = s - r$. Es genügt nun zu zeigen, dass $B_\varepsilon(f) \in C_b(V_x, \mathbb{R}^m) \setminus A$. Sei also $h \in B_\varepsilon(f) \subset C_b(V_x, \mathbb{R}^m)$. Dann gilt

$$\|f(x) - h(x)\| \le \|f - h\| < \varepsilon,$$

also folgt, dass $h(x) > f(x) - \varepsilon = r$. Insbesondere ist $h \notin A$.

[42] Der Banachsche Fixpunktsatz angewandt auf $\Phi : A \to A$ hat gezeigt, dass es genau eine *stetige* Abbildung $h$ gibt, so dass $F(x, h(x)) = 0$ für alle $x$. Diese Aussage ist aber schwächer als die Aussage, dass es genau eine Abbildung $g$ gibt, so dass $F(x, g(x)) = 0$, und das diese Abbildung stetig ist. Die erste Aussage schließt nicht aus, dass es noch weitere, nichtstetige Abbildungen $h'$ gibt, die zweite Aussage dagegen schließt dies aus.

7.3. **Umkehrabbildungen.** Es sei $f\colon U \to V$ eine Abbildung zwischen zwei offenen Mengen $U, V \subset \mathbb{R}^n$. Wir sagen, $g\colon V \to U$ ist eine *Umkehrabbildung von* $f$, wenn $g(f(x)) = x$ für alle $x \in U$ und wenn $f$ bijektiv ist.

Wenn $g$ eine differenzierbare Umkehrabbildung von einer differenzierbaren Abbildung $f$ ist, dann gilt für alle $a \in U$, dass $g(f(a)) = a$, d.h. $g \circ f$ ist die Identitätsabbildung auf $U$. Für beliebiges $a$ folgt dann aus der Kettenregel, dass

$$\mathrm{id}_n = D(\mathrm{id})(a) = D(g \circ f)(a) = Dg(f(a)) \cdot Df(a),$$

d.h. die $n \times n$–Matrix $Df(a)$ ist invertierbar, mit inverser Matrix $Dg(f(a))$.

Es ist sehr leicht zu überprüfen, ob eine lineare Abbildung invertierbar ist, beispielsweise genügt es die Determinante zu berechnen. Andererseits ist es i.a. sehr schwer festzustellen, ob eine beliebige Abbildung $f\colon U \to V$ umkehrbar ist. Der folgende Satz besagt nun, dass wenn das Differential $Df(a)$ einer stetig differenzierbaren an einem Punkt $a$ invertierbar ist, dann ist $f$ 'lokal umkehrbar', d.h. es gibt eine kleine Umgebung $U$ von $a$, so dass die Einschränkung von $f$ auf $U$ umkehrbar ist.

**Satz 7.4. (Satz über die Umkehrabbildung)** *Es sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung und $a \in U$. Wenn $Df(a)$ invertierbar ist, dann existieren offene Umgebungen $U' \subset U$ von $a$ und $V'$ von $b := f(a)$, so dass $f\colon U' \to V'$ bijektiv ist, und so dass die Umkehrabbildung $g = f^{-1} : V' \to U'$ ebenfalls stetig differenzierbar ist.*

*Beispiel.* Wir betrachten die stetig differenzierbare Abbildung

$$\begin{array}{rcl} f\colon \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2, \end{array}$$

und den Punkt $a = 3$ in $\mathbb{R}$. Dann ist $Df(a) = f'(a) = 2a = 6$ ungleich null. Die offenen Umgebungen $U' = (2, 4)$ von $a = 3$ und $V' = (4, 16)$ von $b = f(a) = 9$ und die Abbildung $g\colon V' \to U', x \mapsto \sqrt{x}$ haben dann die geforderten Eigenschaften.

*Beweis.* O.B.d.A. sei $U = \mathbb{R}^n$. Die Idee des Beweises ist die Aussage geschickt auf den Satz über implizite Funktionen zurück zu führen. [43]

[43] Die Anwendung wird (etwas) leichter, wenn man den Satz über implizite Funktionen mit einer anderen Notation formuliert: Es sei

$$\begin{array}{rcl} F : \mathbb{R}^k \times \mathbb{R}^m & \to & \mathbb{R}^m \\ (x, y) & \mapsto & F(x, y) \end{array}$$

eine stetig differenzierbare Abbildung. Es sei $(c, d) \in \mathbb{R}^k \times \mathbb{R}^m$ ein Punkt mit den folgenden Eigenschaften:

Wir betrachten die Abbildung

$$\begin{array}{rcl} F:\mathbb{R}^n\times\mathbb{R}^n & \to & \mathbb{R}^n \\ (x,y) & \mapsto & F(x,y):=x-f(y). \end{array}$$

Man beachte, dass

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

Es gilt $F(b,a)=0$ und

$$\frac{\partial F}{\partial y}(b,a)=-Df(a) \text{ ist nach Voraussetzung invertierbar.}$$

Aus dem Satz über implizite Funktionen, angewandt auf den Punkt $(b,a)$ folgt: es existieren

(1) eine offene Umgebung $V$ von $b$,
(2) eine Umgebung $U$ von $a$, sowie
(3) eine stetig differenzierbare Abbildung $g\colon V\to U$ mit $g(b)=a$,

so dass Folgendes gilt:

(i) für alle $x\in V$ ist

$$F(x,g(x))=0, \text{ d.h. } f(g(x))=x.$$

(ii) und für alle $(x,y)\in V\times U$ mit $F(x,y)=0$ gilt, dass $y=g(x)$.

Wir sind nun fast am Ziel, die Abbildung $g$ hat die Eigenschaft, dass $f(g(x))=x$ für alle $x$, aber die Abbildung $f\colon U\to\mathbb{R}^n$ ist nicht notwendigerweise bijektiv und es gilt auch nicht notwendigerweise, dass $f(U)\subset V$, d.h. i.a. ist $g(f(x))$ noch nicht einmal definiert. Wir müssen nun $U$ auf eine offene Umgebung $U'$ von $a$ einschränken, so dass $f(U)\subset V$. Das dies in der Tat möglich ist, wird im nächsten Absatz ausgeführt".

Nachdem $U$ eine Umgebung von $a$ ist, existiert eine offene Teilmenge $U''\subset U$ welche $a$ enthält. Wir setzen nun $U':=U''\cap f^{-1}(V)$, diese Menge ist der Durchschnitt zweier offener Mengen, also wieder offen. Wir betrachten nun $V':=f(U')\subset V$. Aus (ii) folgt, dass $V'=g^{-1}(U')$,

---

(1) $F(c,d)=0$, und
(2) die $m\times m$–Matrix $\frac{\partial F}{\partial y}$ ist im Punkt $(c,d)$ invertierbar.

Dann existieren

(1) eine offene Umgebungen $C$ von $c$,
(2) eine Umgebung $D$ von $d$, sowie
(3) eine stetig differenzierbare Abbildung $g\colon C\to D$ mit $g(c)=d$,

so dass Folgendes gilt:

(1) für alle $x\in C$ ist $F(x,g(x))=0$,
(2) und für alle $(x,y)\in C\times D$ mit $F(x,y)=0$ gilt, dass $y=g(x)$.

d.h. $V'$ ist das Urbild einer offenen Menge unter einer stetigen Abbildung, also ist $V'$ ebenfalls offen. Es folgt nun, dass $f\colon U' \to V'$ bijektiv ist mit Umkehrabbildung $g$. $\square$

*Definition.* (1) Es sei $U \subset \mathbb{R}^n$ offen und

$$f = (f_1, \dots, f_m) : U \to \mathbb{R}^m$$

eine Abbildung. Wir nennen $f$ eine $C^k$–Abbildung, wenn die Koordinatenfunktionen $f_1, \dots, f_m$ alle $k$–mal stetig partiell differenzierbar sind.

(2) Wir sagen, $f$ ist eine $C^\infty$ Abbildung, wenn $f$ für alle $k$ eine $C^k$–Abbildung ist, d.h. $f$ ist eine $C^\infty$–Abbildung, wenn die Koordinatenfunktionen $f_1, \dots, f_m$ beliebig oft stetig partiell differenzierbar sind.

Die Umkehrabbildung einer bijektive $C^k$–Abbildung ist nicht notwendigerweise eine $C^k$–Abbildung, wie man schon an dem Beispiel

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^3 \end{aligned}$$

sieht, den die Umkehrabbildung $x \mapsto \sqrt[3]{x}$ ist nicht differenzierbar, d.h. $f^{-1}$ ist noch nicht einmal eine $C^1$–Funktion.

**Lemma 7.5.** *Es sei $f\colon U \to V$ eine bijektive $C^k$–Abbildung zwischen offenen Teilmengen von $\mathbb{R}^n$. Wenn die Umkehrabbildung eine $C^1$–Abbildung ist, dann ist die Umkehrabbildung sogar eine $C^k$–Abbildung.*

Das Lemma besagt also, dass wenn wir zeigen wollen, dass die Umkehrabbildung einer $C^k$–Abbildung wiederum eine $C^k$–Abbildung ist, dann genügt es zu zeigen, dass die Umkehrabbildung eine $C^1$–Abbildung ist.

*Beweis.* Wir beweisen das Lemma für den Fall $k = 2$, der allgemeine Fall folgt dann durch Induktion.

Wir bezeichnen mit $g\colon V \to U$ die Umkehrabbildung von $f$. Nachdem $g$ stetig differenzierbar ist, folgt, dass die Einträge von $Dg(y)$ gerade die partiellen Ableitungen sind. Wir müssen zeigen, dass diese stetig differenzierbar sind. Nachdem $f \circ g = \mathrm{id}$ und nachdem $g$ nach Voraussetzung differenzierbar ist, folgt aus der Kettenregel, dass

$$Df(g(y)) \cdot Dg(y) = D\,\mathrm{id}(y) = \mathrm{id} \text{ für alle } y \in V,$$

d.h. es gilt

$$Dg(y) = (Df(g(y)))^{-1} \text{ für alle } y \in V.$$

Nachdem $g$ eine $C^1$–Abbildung ist, und nachdem die partiellen Ableitungen von $f$ nach Voraussetzung stetig differenzierbar sind, folgt, dass die Einträge von $y \mapsto Df(g(y))$ stetig differenzierbar sind. Wir müssen aber wirklich zeigen, dass die Einträge von $(Df(g(y)))^{-1}$ stetig differenzierbar sind. Dies folgt nun aber aus Lemma 7.6, welches im Anschluss bewiesen wird. $\square$

**Lemma 7.6.** *Es sei $U \subset \mathbb{R}^n$ offen und*

$$\begin{array}{rcl} U & \to & GL(n,\mathbb{R}) \\ x & \mapsto & A(x) = (a_{ij}(x)) \end{array}$$

*eine Abbildung, so dass die Matrixeinträge stetig differenzierbar sind, d.h. die Funktionen*

$$\begin{array}{rcl} U & \to & \mathbb{R} \\ x & \mapsto & a_{ij}(x) \end{array}$$

*sind stetig differenzierbar. Dann sind auch die Einträge von*

$$\begin{array}{rcl} U & \to & GL(n,\mathbb{R}) \\ x & \mapsto & A(x)^{-1} \end{array}$$

*stetig differenzierbar.*

*Beweis.* Aus der linearen Algebra ('Cramersche Regel') wissen wir, dass für eine beliebige Matrix $A$ gilt:

$$A^{-1} = \frac{1}{\det(A)} A^{adj},$$

wobei $A^{adj}$ die Adjunkte der Matrix $A$ ist, d.h. der $ij$–Eintrag von $A^{adj}$ ist gegeben durch

$$(-1)^{i+j} \cdot \begin{array}{l} \text{Determinante der Matrix welche man aus } A \\ \text{durch Streichen der } i\text{–ten Spalte und } j\text{–ten Zeile erhält.} \end{array}$$

Es folgt also, dass die Einträge von $A^{-1}$ rationale Funktionen in den Einträgen von $A$ sind. Insbesondere, wenn die Einträge von $x \mapsto A(x)$ stetig differenzierbar sind, dann sind auch die Einträge von $x \mapsto A(x)^{-1}$ stetig differenzierbar. $\square$

*Definition.* Es sei $f : U \to V$ eine bijektive $C^\infty$ Abbildung. Wenn die Umkehrfunktion ebenfalls eine $C^\infty$–Abbildung ist, dann nennen wir $f$ einen *Diffeomorphismus.*

Wir erhalten nun folgendes Korollar aus dem Satz über die Umkehrabbildung) und aus Lemma 7.5

**Korollar 7.7.** *Es sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ eine $C^\infty$–Abbildung und $a \in U$. Wenn $Df(a)$ invertierbar ist, dann existieren offene Umgebungen $U' \subset U$ von $a$ und $V'$ von $b := f(a)$, so dass $f : U' \to V'$ ein Diffeomorphismus ist.*

7.4. **Polarkoordinaten.** Wir beschließen das Kapitel über die impliziten Funktionen und den Umkehrsatz mit einem ausführlichem Beispiel. Wir betrachten die Abbildung

$$\begin{aligned} f\colon \mathbb{R}_{>0}\times\mathbb{R} &\to \mathbb{R}^2\setminus\{0,0\}\\ (r,\varphi) &\mapsto (r\cos\varphi, r\sin\varphi). \end{aligned}$$

Diese Abbildung ist eine $C^\infty$–Abbildung aber keine Bijektion, nachdem $f(r,\varphi) = f(r,\varphi+2\pi)$. Es gilt sogar folgende Aussage für $(r,\varphi)$ und $(r',\varphi') \in R_{>0}\times\mathbb{R}$:

$$f(r,\varphi) = f(r',\varphi') \Leftrightarrow r = r' \text{ und } \varphi = \varphi' + k\cdot 2\pi \text{ für ein } k\in\mathbb{Z}.$$

Die Jacobi-Matrix der Abbildung $f$ ist gegeben durch

$$Df(r,\varphi) = \begin{pmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial \varphi} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial \varphi}\end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi)\end{pmatrix}.$$

Die Determinante der Jacobi–Matrix ist $r$, nachdem $r > 0$, ist die Determinante überall nicht null. [44] Aus dem Satz über die Umkehrabbildung folgt nun, dass $f$ an allen Punkten $(r,\varphi)\in R_{>0}\times\mathbb{R}$ 'lokal invertierbar' ist.

Betrachten wir beispielsweise den Punkt $(1,0)$. Dann ist die Einschränkung von $f$ auf

$$U' := \mathbb{R}_{>0}\times(-\frac{\pi}{2},\frac{\pi}{2}) \to V' := \mathbb{R}_{>0}\times\mathbb{R}$$

umkehrbar mit Umkehrabbildung

$$\begin{aligned} g\colon V' = \mathbb{R}_{>0}\times\mathbb{R} &\to U' = \mathbb{R}_{>0}\times(-\tfrac{\pi}{2},\tfrac{\pi}{2})\\ (x,y) &\mapsto (\sqrt{x^2+y^2}, \arctan(\tfrac{y}{x})). \end{aligned}$$

## 8. Untermannigfaltigkeiten

### 8.1. Definition und Beispiele.

*Definition.* Es sei $T\subset\mathbb{R}^k$ eine offene Teilmenge. Eine stetig differenzierbare Abbildung

$$\begin{aligned} \varphi : T &\to \mathbb{R}^n\\ (t_1,\dots,t_k) &\to \varphi(t_1,\dots,t_k) \end{aligned}$$

[44] Die Abbildung $f\colon \mathbb{R}_{>0}\times\mathbb{R}\to\mathbb{R}^2\setminus 0$ hat also die Eigenschaft, dass die Jacobi-Matrix überall invertierbar ist, aber $f$ ist trotzdem nicht bijektiv.
Dies ist im Kontrast zur Situation von Funktionen $g\colon (a,b)\to(c,d)$ in einer Variablen: wenn die Ableitung überall nicht null ist, dann ist $g$ bijektiv.

heißt *Immersion*, wenn in jedem Punkt der Rang der Funktional–Matrix

$$D\varphi := \left(\frac{\partial \varphi_i}{\partial t_j}\right)_{\substack{i=1,\dots,n \\ j=1,\dots,k}} = \begin{pmatrix} \frac{\partial \varphi_1}{\partial t_1} & \dots & \frac{\partial \varphi_1}{\partial t_k} \\ \vdots & & \vdots \\ \frac{\partial \varphi_n}{\partial t_1} & \dots & \frac{\partial \varphi_n}{\partial t_k} \end{pmatrix}$$

gleich $k$ ist, d.h. wenn die Spalten der Funktional–Matrix linear unabghängig sind.

*Beispiel.* Wir betrachten die Abbildung

$$\begin{aligned} F : (-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \times \mathbb{R} &\to \mathbb{R}^3 \\ (\theta, \varphi) &\mapsto \begin{pmatrix} \cos\theta \cdot \cos\varphi \\ \cos\theta \cdot \sin\varphi \\ \sin\theta \end{pmatrix}. \end{aligned}$$

Die Bildmenge von $F$ ist die Sphäre mit Radius 1 im $\mathbb{R}^3$, ohne den 'Nordpol' und den 'Südpol', d.h. ohne die Punkte $(0, 0, 1)$ und $(0, 0, -1)$. Dann gilt

$$DF = \begin{pmatrix} -\sin\theta\cos\varphi & -\sin\theta\sin\varphi \\ \cos\theta\sin\varphi & \sin\theta\cos\varphi \\ \cos\theta & 0 \end{pmatrix}.$$

Man kann (z.B. mithilfe des Kreuzproduktes der Spaltenvektoren) zeigen, dass die Spaltenvektoren linear unabhängig sind für alle $(\theta, \varphi)$ mit $\cos(\theta) \neq 0$, insbesondere für alle $(\theta, \varphi)$ in unserem Definitionsbereich. Die Abbildung $F$ ist also eine Immersion.

*Bemerkung.*
(1) Der Rang von der $n \times k$–Matrix $D\varphi$ ist genau dann $k$, wenn die $k$ Spaltenvektoren linear unabhängig sind.
(2) Wenn $k = 1$, d.h. wenn $T$ eine Teilmenge von $\mathbb{R}$ ist, dann ist eine Immersion nichts anderes als eine reguläre Kurve.
(3) Aus der linearen Algebra wissen wir, dass eine $n \times k$–Matrix $A$ mit genau dann Rang $k$ hat, wenn es $k$ Zeilen gibt, so dass diese Zeilen eine $k \times k$–Matrix bilden, mit nicht verschwindender Determinante. Beispielsweise ist der Rang der Matrix

$$\begin{pmatrix} 2 & -1 \\ -4 & 2 \\ 0 & 1 \end{pmatrix}$$

gleich zwei, weil die $2 \times 2$–Matrix, welche wir aus der ersten und der letzten Zeile bilden, eine nicht verschwindende Determinante besitzt.

**Satz 8.1.** *Es sei $T \subset \mathbb{R}^k$ offen, $\varphi : T \to \mathbb{R}^n$ eine Immersion und $t_0 \in T$ ein Punkt. Dann existiert eine offene Umgebung $V \subset T$ von $t_0$, so dass die Einschränkung von $\varphi$ auf*

$$\varphi|_V : V \to \varphi(V) \subset \mathbb{R}^n$$

*ein Homöomorphismus* [45] *von $V$ auf $\varphi(V)$ darstellt.*

*Beispiel.* Wir fahren mit dem obigen Beispiel fort. Die Abbildung $F$ ist kein Homöomorphismus, nachdem $F$ noch nicht einmal injektiv ist (z.B. ist $F(\theta, \varphi) = F(\theta, \varphi + 2\pi)$). Aber für einen gegebenen Punkt $(\theta, \varphi)$ ist die Einschränkung von $F$ auf $V := (-\frac{\pi}{2}, \frac{\pi}{2}) \times (\varphi - \pi, \varphi + \pi)$ ein Homöomorphismus.

*Beweis.* Der Spezialfall $k = n$ ist gerade der Satz über die Umkehrabbildung, wir können uns deswegen nun auf den Fall $k < n$ beschränken. Wir werden den Fall $k < n$ ebenfalls auf den Satz über die Umkehrabbildung zurückführen.

Nehmen wir also an, dass der Rang der $n \times k$–Matrix

$$D\varphi = \begin{pmatrix} \frac{\partial \varphi_1}{\partial t_1} & \cdots & \frac{\partial \varphi_1}{\partial t_k} \\ \vdots & & \vdots \\ \frac{\partial \varphi_n}{\partial t_1} & \cdots & \frac{\partial \varphi_n}{\partial t_k} \end{pmatrix}$$

am Punkt $t_0$ gerade $k$ ist. Aus der Bemerkung (3) vor dem Satz wissen wir, dass es $k$ Zeilen von $D\varphi$ gibt, so dass die dazugehörige $k \times k$–Matrix eine nicht verschwindende Determinante besitzt. O.B.d.A. können wir annehmen, dass dies gerade die ersten $k$ Zeilen von $D\varphi$ sind.

Wir schreiben nun

$$\varphi = (\underbrace{\varphi_1, \ldots, \varphi_k}_{\varphi'}, \underbrace{\varphi_{k+1}, \ldots, \varphi_n}_{\varphi''}).$$

Aus der obigen Diskussion folgt, dass das Differential der Abbildung

$$\begin{array}{rcl} \varphi' : T & \to & \mathbb{R}^k \\ t & \mapsto & (\varphi_1(t), \ldots, \varphi_k(t)) \end{array}$$

am Punkt $t_0$ invertierbar ist, wir können also den Satz über die Umkehrabbildung auf $\varphi'$ anwenden. Wir erhalten eine offene Umgebung $V$

[45] Zur Erinnerung: Eine Abbildung $f \colon X \to Y$ zwischen metrischen Räumen heißt *Homöomorphismus*, wenn gilt:

(1) $f$ ist stetig,
(2) $f$ ist bijektiv,
(3) die Umkehrabbildung $f^{-1}$ ist stetig.

von $t_0$ und eine offene Umgebung $U$ von $\varphi'(t_0)$, so dass

$$\varphi' : V \to U$$

bijektiv ist und eine stetig differenzierbare Umkehrabbildung $\psi' : U \to V$ besitzt.

Die Abbildung $\varphi = (\varphi', \varphi'') : V \to U \times \mathbb{R}^{n-k}$ ist nun ebenfalls injektiv und daher bijektiv auf ihr Bild $\varphi(V)$. Eine Umkehrabbildung ist gegeben durch die Einschränkung der stetigen Abbildung

$$\begin{array}{rcl} U \times \mathbb{R}^{n-k} & \to & V \\ (x_1, \dots, x_k, x_{k+1}, \dots, x_n) & \mapsto & \psi'(x_1, \dots, x_k) \end{array}$$

auf $\varphi(V) \subset U \times \mathbb{R}^{n-k}$. □

*Definition.* Eine Teilmenge $M \subset \mathbb{R}^n$ heißt *$k$–dimensionale Untermannigfaltigkeit von* $\mathbb{R}^n$, wenn es zu jedem Punkt $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$, eine offene Teilmenge $T \subset \mathbb{R}^k$, sowie eine Immersion

$$\varphi : T \to \mathbb{R}^n$$

gibt, so dass $\varphi(T) = M \cap U$, und so dass $\varphi : T \to M \cap U$ ein Homöomorphismus ist.

Eine solche Abbildung $\varphi : T \to M \cap U$ wird *Parameterdarstellung*, oder *Parametrisierung* von $M$ in der Umgebung $U$ von $a$ genannt.

*Bemerkung.*
(1) In der ersten Vorlesung hatten wir gesehen, dass wir jede Teilmenge eines metrischen Raumes wieder als metrischen Raum auffassen können. Insbesondere sind, für $M, U$ und $T$ wie in der Definition, die Mengen $T$ und $M \cap U$ wieder metrische Räume, und es macht daher Sinn zu fragen, ob eine gegebene Abbildung $\varphi : T \to M \cap U$ ein Homöomorphismus ist.
(2) Nehmen wir an, wir haben schon eine stetige bijektive Abbildung $\varphi : T \to M \cap U$, und wir wollen noch zeigen, dass $\varphi$ sogar ein Homöomorphismus ist. Wir behaupten nun, dass es genügt eine stetige Abbildung $U \to T$ zu finden, so dass $\psi(\varphi(t)) = t$ für alle $t$. In der Tat, die Einschränkung von $\psi$ auf $M \cap U$ ist gerade die Umkehrabbildung von $\varphi$, aber die Einschränkung der stetigen Abbildung $\psi : U \to T$ auf $M \cap U$ ist weiterhin stetig. Es folgt, dass die Umkehrabbildung von $\varphi$ stetig ist, d.h. $\varphi$ ist ein Homömorphismus.

*Beispiel.* Wir betrachten zuerst den 'unendlichen Zylinder'

$$M := \{(x, y, z) \,|\, x^2 + y^2 = 1\}.$$

Wir wollen zeigen, dass $M$ eine 2–dimensionale Untermannigfaltigkeit von $\mathbb{R}^3$ ist. Sei $a = (a_x, a_y, a_z) \in M$ beliebig. Wir betrachten zuerst den Fall, dass $a_x > 0$. Wir setzen

$$\begin{array}{rcl} T & := & (-\frac{\pi}{2}, \frac{\pi}{2}) \times \mathbb{R} \\ U & := & \{(x, y, z) \,|\, x > 0\} \end{array}$$

und wir betrachten die Abbildung

$$\begin{array}{rcl} \varphi : T & \to & \mathbb{R}^3 \\ (\theta, z) & \mapsto & (\cos(\theta), \sin(\theta), z). \end{array}$$

Man kann leicht nachrechnen, dass das Differential immer Rang zwei hat, d.h. $\varphi$ ist eine Immersion. Zudem ist $\varphi$ stetig und $\varphi : T \to M \cap U$ ist offensichtlich bijektiv. Wir müssen nun noch zeigen, dass $\varphi$ ein Homömorphismus ist. Wir betrachten die Abbildung

$$\begin{array}{rcl} \psi : U & \to & T \\ (x, y, z) & \mapsto & (\arctan(\frac{y}{x}), z). \end{array}$$

Die Abbildung $\psi$ ist stetig und hat die Eigenschaft, dass $\psi(\varphi(\theta, z)) = (\theta, z)$ für alle $(\theta, z) \in T$. Es folgt jetzt aus der obigen Bemerkung, dass $\varphi$ in der Tat ein Homömorphismus ist.

Für andere Punkte $(a_x, a_y, a_z)$ wandelt man die obige Abbildung $\varphi$ geeignet ab.

*Beispiel.* Wir wollen zeigen, dass

$$S^2 = \{(x, y, z) \,|\, x^2 + y^2 + z^2 = 1\}$$

eine 2–dimensionale Untermannigfaltigkeit von $\mathbb{R}^3$ ist. Sei $a = (a_x, a_y, a_z) \in S^2$ ein beliebiger Punkt. Nehmen wir zuerst an, dass $a_x > 0$. Dann hat

$$U = \mathbb{R}_{>0} \times \mathbb{R} \times (-1, 1) = \{(x, y, z) \,|\, x > 0, y \in \mathbb{R}, z \in (-1, 1)\}$$

und

$$\begin{array}{rcl} F : T := (-\frac{\pi}{2}, \frac{\pi}{2}) \times (-\frac{\pi}{2}, \frac{\pi}{2}) & \to & \mathbb{R}^3 \\ (\theta, \varphi) & \mapsto & \begin{pmatrix} \cos\theta \cdot \cos\varphi \\ \cos\theta \cdot \sin\varphi \\ \sin\theta \end{pmatrix}, \end{array}$$

die geforderten Eigenschaften. Für andere Punkte $b = (b_x, b_y, b_z)$ mit $b_z \in (-1, 1)$ kann man ganz ähnliche Parametrisierungen verwenden, man muss nur den Wertebereich von $\varphi$ und $\theta$ verändern. Für $a = (0, 0, 1)$ verfahren wir ganz analog, wir müssen nun aber die Rollen der $y$ und der $z$–Koordinaten vertauschen. Genauer gesagt, die Umgebung

$$U = \mathbb{R}_{>0} \times (-1, 1) \times \mathbb{R} = \{(x, y, z) \,|\, x > 0, y \in (-1, 1), z \in \mathbb{R}\}$$

und die Immersion

$$\begin{aligned} F : T := (-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \times \mathbb{R} &\to \mathbb{R}^3 \\ (\theta, \varphi) &\mapsto \begin{pmatrix} \cos\theta \cdot \cos\varphi \\ \sin\theta \\ \cos\theta \cdot \sin\varphi \end{pmatrix}, \end{aligned}$$

haben die geforderten Eigenschaften.

*Beispiel.* Es seien $R > r > 0$ reelle Zahlen. Wir betrachten die Abbildung

$$\begin{aligned} F : \mathbb{R}^2 &\to \mathbb{R}^3 \\ (\theta, \varphi) &\mapsto \begin{pmatrix} (R + r\cos\theta)\cos\varphi \\ (R + r\cos\theta)\sin\varphi \\ r\sin\theta \end{pmatrix}. \end{aligned}$$

Die Bildmenge dieser Abbildung ist ein Torus $T$. Beispielsweise kann man die Bildmenge wie folgt veranschaulichen: Man betrachtet den Kreis mit Radius $r$ in der $xz$–Ebene um den Punkt $(R, 0, 0)$ und rotiert diesen Kreis in der $xz$–Ebene um die $z$–Achse. [46] Wir wollen nun zeigen, dass der Torus $T$ eine zwei–dimensionale Untermannigfaltigkeit des $\mathbb{R}^3$ ist. Sei also $(x, y, z)$ ein Punkt auf $T$. Wir wählen $(\theta, \varphi) \in \mathbb{R}^2$ mit $F(\theta', \varphi') = (x, y, z)$, dann hat

$$\begin{aligned} \varphi : (\theta' - \tfrac{\pi}{2}, \theta' + \tfrac{\pi}{2}) \times (\varphi' - \tfrac{\pi}{2}, \varphi' + \tfrac{\pi}{2}) &\to \mathbb{R}^3 \\ (\varphi, \theta) &\mapsto F(\varphi, \theta) \end{aligned}$$

die gewünschte Eigenschaft.

*Beispiel.* Betrachten wir folgende Teilmenge des $\mathbb{R}^2$:

$$M := \{(x, y) \in \mathbb{R}^2 \,|\, x = 0 \text{ oder } y = 0\}.$$

Dann ist $M$ *keine* 1–dimensionale Untermannigfaltigkeit, weil für jede Umgebung $U \subset \mathbb{R}^2$ von $(0, 0)$ die Menge $U \cap M$ ein ‘Kreuz’ enthält, und solch ein ‘Kreuz’ nicht homöomorph zu einem Teilgebiet von $\mathbb{R}^1$ ist.[47]

[46] Eine alternative Veranschaulichung ist wie folgt gegeben: Wir betrachten zuerst den Volltorus, d.h. die Menge aller Punkte in $\mathbb{R}^3$ von der Form

$$\begin{pmatrix} (R + s\cos\theta)\cos\varphi \\ (R + s\cos\theta)\sin\varphi \\ s\sin\theta \end{pmatrix}, \theta, \varphi \in \mathbb{R}, s \in [0, r].$$

Für $s = 0$ erhalten wir den Kreis von Radius $R$ in der $xy$–Ebene, der Volltorus ist dann die Menge aller Punkte welche sich um höchstens den Wert $r$ von diesem Kreis entfernen. Der Torus $T$ ist dann die Menge der Randpunkte des Volltorus.

[47] Diese Aussage ist zwar intuitiv klar, aber es ist nicht evident, wie man das mathematisch sauber beweist. Hier ist ein möglicher Ansatz zu zeigen, dass es

*Beispiel.* Betrachten wir die Abbildung

$$\begin{array}{rcl} \varphi : (-2, 2\pi) & \to & \mathbb{R}^2 \\ t & \mapsto & \begin{cases} (1, -t), & \text{wenn } t \in (-2, 0], \\ (\cos(t), -\sin(t)), & \text{wenn } t \in (0, 2\pi). \end{cases} \end{array}$$

Diese Abbildung ist stetig differenzierbar (selbst für $t = 0!$), und $\varphi$ ist eine Immersion. Anderseits ist die Bildmenge von $\varphi$ keine 1-dimensionale Untermannigfaltigkeit. In der Tat, die Bildmenge besteht aus einem Geradenstück und einem Kreis welche sich am Punkt $(1, 0)$ treffen. Aber an dem Punkt $(1, 0)$ kann man, ähnlich wie im vorherigen Beispiel, keine geeignete Umgebung $U$ finden.

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^k$ eine stetig differenzierbare Abbildung.

(1) Ein *regulärer Punkt von $f$* ist ein Punkt $x \in U$, an dem das Differential $Df(x)$ den Rank $k$ besitzt.
(2) Ein *regulärer Wert von $f$* ist ein $z \in \mathbb{R}$, so dass alle Punkte in $f^{-1}(z)$ regulär sind.

Ein Punkt (bzw. Wert), welcher nicht regulär ist, wird singulär genannt.

*Bemerkung.* (1) Das Differential $Df(x)$ ist eine $k \times n$–Matrix, d.h. das Differential hat den Rang $k$ genau dann, wenn die Matrix einen Epimorphismus $\mathbb{R}^n \to \mathbb{R}^k$ definiert, d.h. wenn die Spaltenvektoren von $Df(x)$ den $\mathbb{R}^k$ aufspannen.
(2) Der Fall $k = 1$ ist ein wichtiger Spezialfall: In diesem Fall ist das Differential nichts anderes als der Gradient, und das Differential hat Rang $k = 1$, genau dann, wenn es nicht die Nullmatrix ist. Wir sehen also, dass ein Punkt $x \in U$ ein regulärer Punkt einer Abbildung $f : U \to \mathbb{R}$ ist, genau dann, wenn $\operatorname{grad} f(x) \neq 0$.

*Beispiel.* Es seien $a_1, \ldots, a_n$ positive reelle Zahlen. Wir betrachten

$$\begin{array}{rcl} f : \mathbb{R}^n & \to & \mathbb{R} \\ (x_1, \ldots, x_n) & \mapsto & a_1 x_1^2 + \cdots + a_n x_n^2, \end{array}$$

---

keinen Hömömorphismus $\varphi$ von der Teilmenge

$$M' = \{(x, y) \in \mathbb{R}^2 \mid x, y \in (-\varepsilon, \varepsilon) \text{ und } x = 0 \text{ oder } y = 0\}$$

auf eine Teilmenge $N' \subset \mathbb{R}$ geben kann: Nachdem $M'$ zusammenhängend ist, müsste auch $N' = \varphi(M') \subset \mathbb{R}$ zusammenhängend sein. Andererseits besteht $M' \setminus \{(0, 0)\}$ aus vier Komponenten, die Teilmenge $\varphi(M' \setminus \{(0, 0)\}) = N' \setminus \varphi(0, 0)$ müsste also auch aus vier Komponenten bestehen. Aber es gibt keine zusammenhängende Teilmenge von $\mathbb{R}$, so dass durch herausnehmen eines Punktes die Teilmenge in vier Komponenten zerbricht.

dann gilt für jeden Punkt $x = (x_1, \dots, x_n)$, dass

$$\operatorname{grad} f(x_1, \dots, x_n) = (2a_1x_1, \dots, 2a_nx_n),$$

insbesondere ist der Gradient ungleich null für jeden Punkt $x \neq 0$. D.h. $\mathbb{R}^n \setminus \{0\}$ ist gerade die Menge der regulären Punkte von $f$. Nachdem $f(x) = 0$ genau dann, wenn $x = 0$ folgt, dass alle Werte in $\mathbb{R} \setminus \{0\}$ regulär sind.

Wir können jetzt folgenden nützlichen Satz formulieren:

**Satz 8.2.** (Satz vom regulären Wert) *Es sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^k$ eine stetig differenzierbare Abbildung. Wenn $z \in \mathbb{R}^k$ ein regulärer Wert ist, dann ist $f^{-1}(z) \subset \mathbb{R}^n$ eine $(n-k)$–dimensionale Untermannigfaltigkeit.*

Der Fall $k = 1$ ist eine Umformulierung der Übungsaufgabe 3 von Übungsblatt 7 (welche wiederum eine Anwendung des Satzes über die impliziten Funktionen ist). Wir geben hier nur eine Beweisskizze für den Satz.

*Beweisskizze.* O.B.d.A. sei $U = \mathbb{R}^n$ und sei $z = 0$ der reguläre Wert. Wir wollen zeigen, dass $M := f^{-1}(0)$ eine $(n-k)$–dimensionale Untermannigfaltigkeit ist.

Es sei also $c \in M = f^{-1}(z)$. Nach Voraussetzung ist $c$ ein regulärer Punkt. Das Differential $Df(c)$ hat also den Rang $k$, o.B.d.A. können wir annehmen, dass die letzten $k$ Spalten schon linear unabhängig sind. [48] Wir betrachten jetzt $f$ als eine Abbildung

$$f : \mathbb{R}^{n-k} \times \mathbb{R}^k \to \mathbb{R}^k.$$

Nachdem die Matrix der partiellen Ableitungen von $f = (f_1, \dots, f_k)$ bzgl. der letzten $k$ Variablen invertierbar ist, können wir den Satz über implizite Funktionen anwenden und erhalten eine offene Menge $T \subset \mathbb{R}^{n-k}$ und eine stetig differenzierbare Abbildung $g : T \to \mathbb{R}^k$, so dass $f(x, g(x)) = 0$ für alle $x \in T$, d.h. $(x, g(x)) \in f^{-1}(0)$ für alle $x \in T$. Man kann nun leicht nachweisen, dass $\varphi : T \to \mathbb{R}^n, x \mapsto (x, g(x))$ die gewünschten Eigenschaften besitzt. □

[48] Zur Erinnerung: Es sei $A$ eine $k \times n$–Matrix, dann sind die folgenden Aussagen äquivalent:

(1) $A$ hat Rang $k$.
(2) Die Spaltenvektoren von $A$ erzeugen den $\mathbb{R}^k$.
(3) Es gibt $k$ Spaltenvektoren, welche eine Basis für den $\mathbb{R}^k$ sind.

*Beispiel.* Betrachten wir das obige Beispiel mit $a_1 = \cdots = a_n = 1$ und dem regulären Wert $z = 1$, dann erhalten wir einen weiteren Beweis dafür, dass

$$S^{n-1} := \{(x_1, \ldots, x_n) \,|\, x_1^2 + \cdots + x_n^2 = 1\}$$

eine $(n-1)$–dimensionale Untermannigfaltigkeit ist. Im Allgemeinfall sehen wir, dass die Ellipsoiden

$$\{(x_1, \ldots, x_n) \,|\, a_1 x_1^2 + \cdots + a_n x_n^2 = 1\}$$

$(n-1)$–dimensionale Untermannigfaltigkeit sind.

*Beispiel.* Wir betrachten nun $M(n, \mathbb{R})$ den Raum der reellen $n \times n$–Matrizen. Nachdem eine $n \times n$–Matrix durch die $n^2$ Einträge bestimmt ist, können wir $M(n, \mathbb{R})$ mit dem Raum $\mathbb{R}^{n^2}$ identifizieren. Die Abbildung

$$\begin{array}{rcl} \det : M(n, \mathbb{R}) = \mathbb{R}^{n^2} & \to & \mathbb{R} \\ A & \mapsto & \det(A) \end{array}$$

ist ein Polynom in den Einträgen der Matrix $A$, insbesondere ist diese Abbildung stetig und stetig differenzierbar. In der $*$–Aufgabe von Übungsblatt 6 haben wir gesehen, dass jede Matrix $A \in M(n, \mathbb{R}) = \mathbb{R}^{n^2}$ mit $\det(A) \neq 0$ ein regulärer Punkt der Determinantenabbildung ist. Es folgt, dass jedes $z \in \mathbb{R} \setminus \{0\}$ ein regulärer Wert der Determinantenabbildung ist. Insbesondere folgt aus dem Satz vom regulären Wert, dass

$$SL(n, \mathbb{R}) := \{A \in M(n, \mathbb{R}) \,|\, \det(A) = 1\} \subset M(n, \mathbb{R}) = \mathbb{R}^{n^2}$$

eine $(n^2 - 1)$–dimensionale Untermannigfaltigkeit ist.

Es gibt verschiedene äquivalente Definitionen von Untermannigfaltigkeiten. Insbesondere folgende Äquivalenz ist wichtig für viele Anwendungen.

**Satz 8.3.** *Es sei $M$ eine Teilmenge von $\mathbb{R}^n$. Dann ist $M$ eine $k$–dimensionale Untermannigfaltigkeit genau dann, wenn es zu jedem Punkt $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$, eine offene Menge $V \subset \mathbb{R}^n$, sowie eine $C^1$–invertierbare Abbildung $\Phi : U \to V$ gibt, so dass*

$$\Phi(M \cap U) = V \cap E_k,$$

*wobei*

$$E_k := \{(x_1, \ldots, x_k, 0, \ldots, 0) \in \mathbb{R}^n \,|\, x_1, \ldots, x_k \in \mathbb{R}\}.$$

Die Abbildungen $\Phi$ werden 'Karten von $M$' genannt. Parametrisierungen sind also Immersionen von offenen Teilmengen von $\mathbb{R}^k$ welche $M$

beschreiben, Karten hingegen sind Abbildungen von offenen Teilmengen des $\mathbb{R}^n$ nach $\mathbb{R}^n$ welche '$M$ gerade biegen'. Beide Gesichtspunkte sind hilfreich und werden oft verwendet.

*Beweis.* Sei $a \in M$ ein Punkt. Nehmen wir zuerst an, dass es eine $C^1$–invertierbare Abbildung $\Phi : U \to V$ mit den genannten Eigenschaften gibt. Wir werden sehen, dass die Einschränkung von $\Phi^{-1}$ auf $V \cap E_k$ (aufgefasst als Teilmenge des $\mathbb{R}^k$), die gewünschte Parametrisierung ist.

Wir setzen also

$$T := \{(x_1, \dots, x_k) \in \mathbb{R}^k \,|\, (x_1, \dots, x_k, 0, \dots, 0) \in V \cap E_k\},$$

und wir betrachten die Abbildung

$$\begin{array}{rcl} \psi : T & \to & \mathbb{R}^n \\ (x_1, \dots, x_k) & \mapsto & (x_1, \dots, x_k, 0, \dots, 0) \end{array}$$

und definieren

$$\begin{array}{rcl} \varphi : T & \to & \mathbb{R}^n \\ (x_1, \dots, x_k) & \mapsto & \Phi^{-1}(\psi(x_1, \dots, x_k)). \end{array}$$

Wir behaupten, dass $\varphi$ die gewünschten Eigenschaften besitzt. [49]

Aus der Kettenregel folgt, dass

$$D\varphi(x) = D(\Phi^{-1} \circ \psi)(x) = D\Phi^{-1}(\psi(x)) \cdot D\psi(x).$$

Das Differential $D\psi(x)$ hat offensichtlich Rang $k$ und das Differential $D\Phi^{-1}(\varphi(x))$ ist ein Isomorphismus[50], die Verknüpfung der beiden Abbildungen hat dann ebenfalls Rang $k$. Wir haben damit gezeigt, dass $\varphi$ eine Immersion ist. Wir müssen noch zeigen, dass $\varphi : T \to M \cap U$ ein Homöomorphismus ist. Nach Voraussetzung ist $\Phi : M \cap U \to V \cap E_k$ ein Homöomorphismus. Nachdem $\varphi : T \to M \cap U$ die Verknüpfung der beiden Homöomorphismen $\psi : T \to V \cap E_k$ und $\Phi^{-1} : V \cap E_k \to M \cap U$ ist, folgt, dass auch $\varphi : T \to M \cap U$ ein Homöomorphismus ist.

Nehmen wir nun an, dass $M$ eine $k$–dimensionale Untermannigfaltigkeit ist. Es sei $a \in M$ ein Punkt. Nach Voraussetzung existiert eine offene Umgebung $U \subset \mathbb{R}^n$, eine offene Teilmenge $T \subset \mathbb{R}^k$, sowie eine Immersion

$$\varphi : T \to \mathbb{R}^n,$$

[49] Kurz gesagt, ist $\varphi$ die Einschränkung von $\Phi^{-1}$ auf den Durchschnitt von $V$ mit $\mathbb{R}^k$, wenn wir den $\mathbb{R}^k$ als Teilmenge des $\mathbb{R}^n$ betrachten.

[50] Die Umkehrabbildung ist gegeben durch das Differential von $\Phi$ am Punkt $\Phi^{-1}(\varphi(x))$.

so dass $\varphi(T) = M \cap U$, und so dass $\varphi : T \to M \cap U$ ein Homöomorphismus ist. Nach Voraussetzung sind die folgenden $k$ Vektoren in $\mathbb{R}^n$ linear unabhängig:

$$v_1 := \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1}(a) \\ \vdots \\ \frac{\partial \varphi_n}{\partial x_1}(a) \end{pmatrix}, \dots, v_k := \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_k}(a) \\ \vdots \\ \frac{\partial \varphi_n}{\partial x_k}(a) \end{pmatrix}.$$

Aus dem Basisergänzungssatz erhalten wir $n-k$ Vektoren $w_1, \dots, w_{n-k}$, so dass die Vektoren $v_1, \dots, v_k, w_1, \dots, w_{n-k}$ eine Basis des $\mathbb{R}^n$ bilden. Wir betrachten jetzt die Abbildung

$$\begin{array}{rcl} \Psi : T \times \mathbb{R}^{n-k} & \to & \mathbb{R}^n \\ (x_1, \dots, x_k, y_1, \dots, y_{n-k}) & \mapsto & \varphi(x_1, \dots, x_k) + y_1 w_1 + \dots + y_{n-k} w_{n-k}. \end{array}$$

Das Differential am Punkt $(x_1, \dots, x_k, 0, \dots, 0)$ ist gegeben durch die Matrix

$$\begin{pmatrix} v_1 & \dots & v_k & w_1 & \dots & w_{n-k} \end{pmatrix},$$

welche invertierbar ist. Nachdem Satz über die Umkehrabbildung besitzt $\Psi$ eine lokale Umkehrung $\Phi$. Dieses $\Phi$ hat dann die gewünschten Eigenschaften.

□

### 8.2. **Tangential– und Normalenvektoren.**

*Definition.* Es sei $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $a \in M$.

(1) Ein Vektor $v \in \mathbb{R}^n$ heißt *Tangentialvektor* an $M$ im Punkt $a$, wenn es eine stetig differenzierbare Kurve

$$\begin{array}{rcl} \gamma \colon (-\varepsilon, \varepsilon) & \to & M \\ t & \mapsto & \gamma(t) \end{array}$$

gibt, so dass $\gamma(0) = a$ und $\gamma'(0) = v$.

(2) Die Menge aller Tangentialvektoren am Punkt $a$ wird mit $T_a M$ bezeichnet. Wir nennen $T_a M$ den *Tangentialraum* am Punkt $a$.

(3) Ein Vektor $v \in \mathbb{R}^n$ heißt *Normalenvektor* von $M$ in $a$, wenn $v$ orthogonal ist zu allen Vektoren in $T_a M$.

*Beispiel.* (1) Wir betrachten

$$S^2 := \{(x, y, z) \,|\, x^2 + y^2 + z^2 = 1\}$$

und den Punkt $a = (0, 0, 1)$. Dann ist $v = (1, 0, 0)$ ein Tangentialvektor, nachdem $v$ die Ableitung der Kurve

$$\begin{array}{rcl} \gamma \colon (-\pi, \pi) & \to & S^2 \\ t & \mapsto & (\sin(t), 0, \cos(t)) \end{array}$$

durch den Punkt $a$ ist. Allgemeiner, für $r \geq 0$ und $\varphi \in [0, 2\pi]$ können wir die Kurve

$$\begin{array}{rcl} \gamma\colon (-\pi, \pi) & \to & S^2 \\ t & \mapsto & (\cos(\varphi)\cdot\sin(rt), \sin(\varphi)\cdot\sin(rt), \cos(rt)) \end{array}$$

betrachten.[51] Wir berechnen, dass $\gamma'(0) = (r\cos(\varphi), r\sin(\varphi), 0)$. Wir sehen also, dass alle Vektoren vom Typ $(x, y, 0)$ Tangentialvektoren sind. Es ist nicht ganz offensichtlich (siehe den nächsten Satz), aber dies sind in der Tat alle Tangentialvektoren, und es folgt, dass

$$T_{(0,0,1)}S^2 = \{(x, y, 0) \,|\, x, y \in \mathbb{R}\}.$$

Zudem sind alle Normalenvektor im Punkt $a$ von der Form $(0, 0, t)$, $t \in \mathbb{R}$.

(2) Betrachten wir

$$E_k := \{(x_1, \dots, x_k, 0, \dots, 0) \,|\, x_1, \dots, x_k \in \mathbb{R}\} \subset \mathbb{R}^n.$$

Dann gilt für jede Kurve $\gamma\colon (-\varepsilon, \varepsilon) \to E_k$, dass $\gamma'(0) \in E_k$, und man sieht nun leicht, dass für jedes $a \in E_k$ gilt:

$$T_a(E_k) = E_k.$$

**Satz 8.4.** *Es sei $M \subset \mathbb{R}^n$ eine $k$–dimensionale Untermannigfaltigkeit und $a \in M$. Dann gilt:*

(1) *$T_aM$ ist ein $k$–dimensionaler Vektorraum,*
(2) *wenn $\varphi : V \to M$ eine Parametrisierung von $M$ am Punkt $a$ ist mit $\varphi(b) = a$, dann bilden die Vektoren* [52]

$$\frac{\partial\varphi}{\partial t_1}(b), \dots, \frac{\partial\varphi}{\partial t_k}(b)$$

*eine Basis von $T_aM$.*

*Beweis.* Wir schreiben

$$v_i := \frac{\partial\varphi}{\partial t_i}(b), i = 1, \dots, k,$$

und bezeichnen mit $V$ den von $v_1, \dots, v_k$ aufgespannten Raum in $\mathbb{R}^n$. Wir beweisen zuerst, dass $V \subset T_aM$.

Sei also $v = \sum_{i=1}^k c_iv_i \in V$. Wir betrachten die Kurve

$$\begin{array}{rcl} \eta : (-\varepsilon, \varepsilon) & \to & \mathbb{R}^k \\ s & \mapsto & b + s\cdot(c_1, \dots, c_k). \end{array}$$

[51]Diese Kurve beschreibt den Großkreis durch den Nordpol von 'Längengrad' $\varphi$.
[52]Dies sind natürlich nur die Spaltenvektoren des Differentials von $\varphi$

Dann ist $s \mapsto \varphi(\eta(s))$ eine Kurve durch $a$, und aus der Kettenregel folgt, dass

$$\begin{aligned}(\varphi \circ \eta)'(0) &= D\varphi(\eta(0)) \cdot \eta'(0) \\ &= \begin{pmatrix} v_1 & \dots & v_k \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} \\ &= \textstyle\sum_{i=1}^k c_i v_i = v.\end{aligned}$$

Wir haben nun gezeigt, dass der $k$–dimensionale Vektorraum $V$ in $T_aM$ *enthalten* ist.

Wir wählen nun $\Phi : U \to V$ wie in Satz 8.3.

*Behauptung.* Das Differential $D\Phi(a)$ induziert eine Abbildung zwischen $T_aM$ und $T_bE_k$. [53]

Es sei $v \in T_aM$ ein Vektor. Wir müssen zeigen, dass $(D\Phi)(a) \cdot v \in T_bE_k$. Es sei $\gamma$ eine Kurve durch $a$ mit $\gamma'(0) = v$. Aus der Kettenregel (rückwärts angewandt) folgt:

$$D\Phi(a)(v) = D\Phi(a)(\gamma'(0)) = (\Phi \circ \gamma)'(0),$$

aber $\Phi \circ \gamma$ ist eine Kurve in $E_k$ durch $b$, d.h. $(\Phi \circ \gamma)'(0)$ liegt in $T_bE_k$. Wir haben damit die Behauptung bewiesen.

Wir haben also bewiesen, dass $D\Phi(a)(T_aM) \subset T_b(E_k)$. Nachdem $T_b(E_k) = E_k$ (siehe das obige Beispiel), folgt, dass

$$T_aM \subset ((D\Phi)(a))^{-1}(E_k).$$

Nachdem $D\Phi(a)$ ein Isomorphismus ist, ist auch $((D\Phi)(a))^{-1}(E_k)$ ein $k$–dimensionaler Vektorraum. Wir erhalten also, dass

$$V \subset T_aM \subset ((D\Phi)(a))^{-1}(E_k),$$

nachdem $V$ und $((D\Phi)(a))^{-1}(E_k)$ jeweils $k$–dimensionale Vektorräume sind, muss auch $T_aM$ ein $k$–dimensionaler Vektorraum sein. □

Wenn wir Untermannigfaltigkeiten betrachten, welche als Urbild eines regulären Wertes beschrieben einer Funktion $f : U \to \mathbb{R}$ sind, dann ist es besonders einfach den Tangentialraum zu bestimmen:

**Satz 8.5.** *Es sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine stetig differenzierbare Abbildung und $z \in \mathbb{R}$ ein regulärer Wert. Dann ist $M := f^{-1}(z) \subset \mathbb{R}^n$*

[53] Das Differential ist eine $n \times n$–Matrix, also eine lineare Abbildung $\mathbb{R}^n \to \mathbb{R}^n$. Der Tangentialraum $T_aM$ ist ein Unterraum von $\mathbb{R}^n$, und die Aussage der Behauptung ist nun, dass das Bild von $T_aM \subset \mathbb{R}^n$ unter der linearen Abbildung $D\Phi(a) : \mathbb{R}^n \to \mathbb{R}^n$ in $T_bE_k$ liegt.

*eine* $(n-1)$*–dimensionale Untermannigfaltigkeit, und für jedes* $a \in M$ *gilt:*

$$T_aM := \operatorname{grad} f(a)^\perp = \{v \in \mathbb{R}^n \,|\, v \textit{ orthogonal zu } \operatorname{grad} f(a)\}.$$

Die Aussage des Satzes ist eine Umformulierung von Aufgabe 2 aus dem Übungsblatt 6.

*Beweis.* Nach Satz 8.2 ist $M = f^{-1}(z)$ eine $(n-1)$–dimensionale Untermannigfaltigkeit. Sei nun $a \in M$. Nach Satz 8.4 ist $T_aM$ ein $(n-1)$–dimensionaler Untervektorraum. Nachdem $z$ ein regulärer Wert ist, ist $a$ ein regulärer Punkt von $f$, d.h. $\operatorname{grad} f(a)$ ist nicht der Nullvektor. Es folgt, dass

$$\operatorname{grad} f(a)^\perp = \{v \in \mathbb{R}^n \,|\, v \text{ orthogonal zu } \operatorname{grad} f(a)\}$$

ebenfalls ein $(n-1)$–dimensionaler Untervektorraum ist.

Es genügt nun zu zeigen, dass $T_aM \subset \operatorname{grad} f(a)^\perp$. Es sei also $v \in T_aM$. Wir wählen eine stetig differenzierbare Kurve

$$\begin{array}{rcl} \gamma\colon (-\varepsilon,\varepsilon) & \to & M \\ t & \mapsto & \gamma(t) \end{array}$$

mit $\gamma(0) = a$ und $\gamma'(0) = v$. Dann gilt $f(\gamma(t)) = z$ für alle $t \in (-\varepsilon,\varepsilon)$, insbesondere ist $f \circ \gamma$ eine konstante Funktion, und es folgt aus der Kettenregel, dass

$$0 = (f \circ \gamma)'(0) = \operatorname{grad} f(\gamma(0)) \cdot g'(0) = \operatorname{grad} f(a) \cdot v,$$

d.h. $v$ ist orthogonal zu $\operatorname{grad} f(a)$. □

*Beispiel.* Betrachten wir das Ellipsoid

$$M := \{(x,y,z) \,|\, f(x,y,z) = 9\}$$

wobei $f(x,y,z) = 2x^2 + 3y^2 + 4z^2$. Dann gilt

$$\operatorname{grad} f(x,y,z) = (4x, 6y, 8z),$$

und der Tangentialraum von $M$ am Punkt $(1,1,1)$ ist gegeben durch

$$T_{(1,1,1)}M = (4,6,8)^\perp = \{\lambda(3,-2,0) + \mu(-2,0,1) \,|\, \lambda,\mu \in \mathbb{R}\}.$$

## 9. Die Taylorformel und lokale Extrema

9.1. **Die Landau–Symbole.** Wir führen jetzt die Landau–Symbole $O$ und $o$ ein, welche in Forster Analysis I auf Seite 121 eingeführt werden.

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen, es seien $f, g\colon U \to \mathbb{R}$ Funktionen und es sei $c \in U$. Wir schreiben[54]

$$f(x) = o(g(x)) \text{ für } x \to c,$$

[54] Gesprochen wird das wie folgt: '$f(x)$ ist klein-oh von $g(x)$'.

wenn es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass

$$|f(x)| \leq \varepsilon |g(x)| \text{ für alle } x \in B_\delta(c).$$

Wenn $f, g, h : U \to \mathbb{R}$ Funktionen sind, dann schreiben wir auch

$$f(x) = g(x) + o(h(x)) \text{ wenn } f(x) - g(x) = o(h(x)).$$

*Beispiel.* Wenn wir uns auf $U = \mathbb{R}$ beschränken, dann gilt beispielsweise, dass

$$x^3 = o(x^2) \text{ für } x \to 0,$$

denn für jedes $\varepsilon > 0$ setzen wir $\delta = \varepsilon$, und dann gilt für $x \in (-\delta, \delta)$, dass

$$|x^3| = |x^2| \cdot |x| < |x^2| \cdot \delta = |x^2| \cdot \varepsilon.$$

In Übungsblatt 9 werden wir folgendes Lemma beweisen:

**Lemma 9.1.** *Es sei $U \subset \mathbb{R}^n$ offen, es seien $f, g \colon U \to \mathbb{R}$ Funktionen, so dass $f(c) = 0$ und $g(x) \neq 0$ für alle $x \in U$. Es sei $c \in U$. Dann gilt:*

$$\lim_{x \to c} \frac{f(x)}{g(x)} = 0 \;\Leftrightarrow\; f(x) = o(g(x)) \text{ für } x \to c.$$

Der vollständigkeitshalber führen wir noch das Symbol $O$ ein: Es seien $f, g \colon U \to \mathbb{R}$ Funktionen und $c \in U$. Wir schreiben[55]

$$f(x) = O(g(x)) \text{ für } x \to c$$

wenn es ein $K > 0$ und ein $\delta > 0$ gibt, so dass

$$|f(x)| \leq K \cdot |g(x)| \text{ für alle } x \in B_\delta(c).$$

*Beispiel.* Es gilt

$$e^x - 1 = O(x) \text{ für } x \to 0,$$

dies sieht man wenn man $K = 2$ und $\delta = \frac{1}{2}$ wählt.

Wenn $f, g \colon (a, \infty) \to \mathbb{R}$ Funktionen sind, dann schreiben wir auch

$$f(x) = o(g(x)) \text{ für } x \to \infty$$

wenn es für alle $\varepsilon > 0$ ein $R > 0$ gibt, so dass

$$|f(x)| \leq \varepsilon |g(x)| \text{ für alle } x \geq R.$$

Wir schreiben

$$f(x) = O(g(x)) \text{ für } x \to \infty$$

wenn es ein $K > 0$ und ein $R > 0$ gibt, so dass

$$|f(x)| \leq K \cdot |g(x)| \text{ für alle } x \geq R.$$

[55]Gesprochen wird das wie folgt: '$f(x)$ ist groß-oh von $g(x)$'.

9.2. **Die Taylorformel aus der Analysis I.** Bevor wir die Taylorformel für Funktionen $\mathbb{R}^k \to \mathbb{R}$ einführen, erinnern wir an das Taylorpolynom aus der Analysis I.

Es sei $g\colon (a,b) \to \mathbb{R}$ eine $C^k$-Funktion und $c \in (a,b)$. Dann heißt

$$p_k(t) := p_{k,c}(g)(t) := \sum_{l=0}^{k} \frac{g^{(l)}(c)}{l!}(t-c)^l$$

das *$k$-te Taylorpolynom von $g$ bei $c$*. Zur Erinnerung, das Polynom $p_k$ zeichnet sich dadurch aus, dass gür alle $l = 0, \dots, k$ gilt:

$$p_k^{(l)}(c) = g^{(l)}(c),$$

d.h. die ersten $k$ Ableitungen von $p_k$ und $g$ stimmen am Punkt $c$ überein.

Der folgende Satz ist nun Satz 16.2 aus der Analysis I:

**Satz 9.2.** *Es sei $g\colon (a,b) \to \mathbb{R}$ eine $C^k$-Funktion und $c \in (a,b)$, dann gilt*

$$\lim_{t \to c} \frac{g(t) - p_k(t)}{(t-c)^k} = 0.$$

In der Landau-Notation können wir auch schreiben, dass

$$g(t) = p_k(t) + o((t-c)^k) \text{ für } t \to c.$$

Eine etwas genauere Formulierung ist durch den folgenden Satz gegeben (Analysis I, Satz 16.4):

**Satz 9.3.** *(Restgliedformel von Lagrange) Sei $g\colon (a,b) \to \mathbb{R}$ eine $C^k$-Funktion, $c \in (a,b)$ und $t \in (a,b)$. Dann existiert ein $\theta \in [c,t]$, so dass*

$$g(t) = p_{k-1}(t) \ + \ \frac{g^{(k)}(\theta)}{(k)!} \cdot (t-c)^k.$$

Man kann Satz 9.2 leicht aus Satz 9.3 herleiten.

9.3. **Die Taylorformel im $\mathbb{R}^n$.** Die Idee ist nun eine Funktion $\mathbb{R}^n \to \mathbb{R}$ durch ein Polynom in $n$ Variablen zu approximieren. Damit wir die Taylorformel in mehreren Variablen formulieren können, müssen wir zuerst einige Notationen einführen.

*Notation.* Für ein $n$–Tupel $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ definieren wir

$$\begin{aligned} |\alpha| &:= \alpha_1 + \dots + \alpha_n, \\ \alpha! &:= \alpha_1!\,\alpha_2! \dots \alpha_n! \end{aligned}$$

Wenn $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, dann schreiben wir

$$x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}.$$

Wenn $f$ eine $|\alpha|$–mal stetig differenzierbare Funktion ist, so setzen wir

$$D^\alpha f\colon = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ eine $k$-mal stetig differenzierbare Funktion. Es sei $x \in U$. Für $m = 0, \dots, k$ definieren wir [56]

$$q_m(\xi) := \sum_{|\alpha|=m} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha$$

und wir bezeichnen

$$p_k(\xi) := \sum_{m=0}^{k} q_m(\xi)$$

als das *$k$–te Taylorpolynom von $f$ im Punkt $x$.*

*Bemerkung.*
(1) Man beachte, dass $q_m(\xi)$ ein homogenes Polynom $m$-ten Grades in den Variablen $\xi_1, \dots, \xi_n$ ist, und $p_k(\xi)$ ist ein nicht homogenes Polynom $k$–ten Grades in den Variablen $\xi_1, \dots, \xi_n$.
(2) Wenn $n = 1$, dann ist gibt es für jedes $m$ genau ein $\alpha$ mit $|\alpha| = m$, nämlich $\alpha = (m)$. Wir erhalten also, dass

$$p_k(\xi) := \sum_{m=0}^{k} \frac{f^{(m)}(x)}{m!} \xi^m.$$

Wenn $x = 0$, dann ist dies gerade das übliche Taylorpolynom aus der Analysis I, wenn $x \neq 0$, dann erhalten wir das Taylorpolynom aus der Analysis I um '$x$ auf der $x$-Achse verschoben'.
(3) Die partiellen Ableitungen von $f$ bis zur Ordnung $k$ am Punkt $x$ entsprechen den partiellen Ableitungen von $p_k$ am Punkt 0, d.h. für alle $\alpha$ mit $|\alpha| \leq k$ gilt:

$$(D^\alpha f)(x) = (D^\alpha p_k)(0).$$

Der Spezialfall $k = 2$ wird im Übungsblatt 9 behandelt, der allgemeine Fall ist ebenfalls nicht schwierig, aber die Notation wird etwas unübersichtlich.

Es gibt nur ein $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 0$, nämlich $\alpha = (0, \dots, 0)$, es folgt also, dass

$$q_0(\xi) = f(x).$$

Die einzigen Vektoren $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 1$, sind die Vektoren

$$\alpha = e_i = (0, \dots, 0, 1, 0, \dots, 0) \text{ mit 1 in der } i\text{-ten Koordinate.}$$

---

[56] Wir nehmen also die Summe über alle $\alpha \in \mathbb{N}^n$, so dass $|\alpha| = m$.

Es folgt, dass

$$\begin{array}{rcl} q_1(\xi) & = & \sum_{|\alpha|=1} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha = \sum_{i=1}^n \frac{D^{e_i} f(x)}{e_i!}\xi^{e_i} = \sum_{i=1}^n = \frac{\partial f}{\partial x_i}(x)\cdot \xi_i \\ & = & \langle \operatorname{grad} f(x), \xi\rangle. \end{array}$$

Die Bestimmung von $q_2(\xi)$ ist etwas aufwändiger, und wir schreiben dies deshalb als Lemma auf:

**Lemma 9.4.** *Es gilt*

$$q_2(\xi) = \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} f(x)\cdot \xi_i\xi_j.$$

Dieses Lemma wird in Forster Analysis II, Seite 77 bewiesen. Es ist auch eine Übungsaufgabe im 9. Übungsblatt.

Wir können jetzt den wichtigsten Satz dieses Kapitels formulieren:

**Satz 9.5. (Satz von Taylor)** *Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ eine $k$-mal stetig differenzierbare Funktion. Es sei $x \in U$. Dann gilt*

$$f(x+\xi) = p_k(\xi) \ + \ o(\|\xi\|^k) \text{ für } \xi \to 0.$$

Der Beweis des Satzes benötigt verschiedene Hilfssätze. Die Grundidee dabei ist Satz 9.5 auf Satz 9.3 zurück zu führen.

**Satz 9.6.** *Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ eine $k$-mal stetig differenzierbare Funktion. Es sei $x \in U$ und $\xi \in \mathbb{R}^n$ ein Vektor, so dass $x + t\xi \in U$ für alle $t \in [0,1]$. Dann ist die Funktion*

$$\begin{array}{rcl} g\colon [0,1] & \to & \mathbb{R} \\ t & \mapsto & f(x+t\xi) \end{array}$$

*$k$–mal stetig differenzierbar, und es gilt*

$$g^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!}(D^\alpha f)(x+t\xi)\cdot \xi^\alpha.$$

*Beweis.* Wir zeigen zuerst folgende Behauptung:

*Behauptung.* Es gilt [57]

$$g^{(k)}(t) = \sum_{i_1,\dots,i_k=1}^n \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x+t\xi)\xi_{i_1}\cdots\xi_{i_k}.$$

---

[57] Die Notation $\sum_{i_1,\dots,i_k=1}^n$ ist eine Abkürzung von $\sum_{i_1=1}^n \sum_{i_2=1}^k \cdots \sum_{i_k=1}^n$.

Wir beweisen die Behauptung durch Induktion nach $k$. Für $k = 1$ folgt aus der Kettenregel[58], dass

$$g'(t) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x + t\xi)\xi_i.$$

Der Induktionsschritt $k - 1 \to k$ folgt ebenfalls aus der Kettenregel: [59]

$$\begin{aligned}
g^{(k)}(t) &= \tfrac{d}{dt}\Big(\textstyle\sum_{i_1,\ldots,i_{k-1}=1}^{n} \frac{\partial}{\partial x_{i_{k-1}}} \cdots \frac{\partial}{\partial x_{i_1}} f(x+t\xi)\xi_{i_1}\cdots\xi_{i_{k-1}}\Big)\\
&= \textstyle\sum_{i_1,\ldots,i_{k-1}=1}^{n} \frac{d}{dt}\Big(\frac{\partial}{\partial x_{i_{k-1}}} \cdots \frac{\partial}{\partial x_{i_1}} f(x+t\xi)\Big)\cdot \xi_{i_1}\cdots\xi_{i_{k-1}}\\
&= \textstyle\sum_{i_1,\ldots,i_{k-1}=1}^{n} \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\Big(\frac{\partial}{\partial x_{i_{k-1}}} \cdots \frac{\partial}{\partial x_{i_1}} f(x+t\xi)\Big)\xi_j \cdot \xi_{i_1}\cdots\xi_{i_{k-1}}\\
&= \textstyle\sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x+t\xi)\xi_{i_1}\cdots\xi_{i_k}.
\end{aligned}$$

Die erste Gleichheit folgt aus der Induktionsannahme, die zweite Folge aus der Kettenregel wie in Fußnote 58, und die dritte Folgt in dem wir die Laufvariable '$j$' durch '$i_k$' ersetzen. Wir haben damit die Behauptung bewiesen.

*Behauptung.* Es gilt [60]

$$\sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x+t\xi)\xi_{i_1}\cdots\xi_{i_k} = \sum_{|\alpha|=k} \frac{k!}{\alpha!}(D^\alpha f)(x+t\xi)\cdot\xi^\alpha.$$

Es seien $i_1, \ldots, i_k \in \{1, \ldots, n\}$. Wenn unter den Indizes $(i_1, \ldots, i_k)$ der Index 1 genau $\alpha_1$-mal, der Index 2 genau $\alpha_2$-mal etc. vorkommt, so folgt aus der Vertauschbarkeit der partiellen Ableitungen, dass

$$\frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(y)\xi_{i_1}\cdots\xi_{i_k} = \frac{\partial^{|\alpha|} f(y)}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}\xi_1^{\alpha_1}\cdots\xi_n^{\alpha_n} = D^\alpha f(y)\cdot\xi^\alpha.$$

---

[58] Wir verwenden folgende Version der Kettenregel: es sei $h : \mathbb{R}^n \to \mathbb{R}$ eine Funktion und $v \in \mathbb{R}^n$ ein Vektor. Wir definieren $g(t) = x + tv$. Dann gilt

$$\begin{aligned}
\tfrac{d}{dt}h(x+tv) &= \tfrac{d}{dt}h(g(t)) = Df(g(t))g'(t) = \left(\tfrac{\partial h}{\partial x_1}(g(t))\ldots\tfrac{\partial h}{\partial x_n}(g(t))\right)\begin{pmatrix} v_1\\ \vdots \\ v_n\end{pmatrix}\\
&= \textstyle\sum_{i=1}^{n} \frac{\partial h}{\partial x_i}(g(t))\cdot v_i\\
&= \textstyle\sum_{i=1}^{n} \frac{\partial h}{\partial x_i}(x+tv)\cdot v_i.
\end{aligned}$$

[59] Dieses Mal angewandt auf

$$\begin{aligned}
h : \mathbb{R}^n &\to \mathbb{R}\\
x &\mapsto \tfrac{\partial}{\partial x_{i_{k-1}}}\cdots\tfrac{\partial}{\partial x_{i_1}} f(x),
\end{aligned}$$

und $g(t) = x + tv$.

[60] Es ist eine hilfreiche Übung sich davon zu überzeugen, dass im Falle $k = 2$ diese Behauptung gerade der Aussage von Lemma 9.4 darstellt.

Für gegebenes $\alpha$ gibt es genau

$$\frac{k!}{\alpha_1!\cdots\alpha_n!}=\frac{k!}{\alpha!}$$

$k$–tupel $(i_1,\dots,i_k)$, bei denen die $l$ genau $\alpha_l$–mal vorkommt. Es folgt, dass für gegebenes $\alpha\in\mathbb{N}^n$ mit $|\alpha|=k$, in dem Ausdruck

$$\sum_{i_1,\dots,i_k=1}^{n}\frac{\partial}{\partial x_{i_k}}\cdots\frac{\partial}{\partial x_{i_1}}f(x+t\xi)\xi_{i_1}\cdots\xi_{i_k}$$

die partielle Ableitung $D^\alpha f(x+t\xi)$ genau $\frac{k!}{\alpha!}$ erscheint. Es folgt die Behauptung. □

**Satz 9.7.** *Es sei $U\subset\mathbb{R}^n$ offen und $f\colon U\to\mathbb{R}$ eine $k$–mal stetig differenzierbare Funktion. Es sei $x\in U$ und $\xi\in\mathbb{R}^n$ ein Vektor, so dass $x+t\xi\in U$ für alle $t\in[0,1]$. Dann existiert ein $\theta\in[0,1]$, so dass*

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

Dieser Satz folgt sofort aus Satz 9.6 und Satz 9.3 angewandt auf die Funktion [61]

$$\begin{array}{rcl} g\colon[0,1] & \to & \mathbb{R}\\ t & \mapsto & g(t):=f(x+t\xi),\end{array}$$

und $c=0$. Man beachte, dass sich dabei die Faktoren $k!$ wegheben. (Siehe Forster Analysis II, Seite 75.)

Wir können nun Satz 9.5 beweisen.

*Beweis von Satz 9.5.* Es sei $U\subset\mathbb{R}^n$ offen und $f\colon U\to\mathbb{R}$ eine $k$-mal stetig differenzierbare Funktion. Es sei $x\in U$. Wir wollen zeigen, dass

$$f(x+\xi)=p_k(\xi)\ +\ o(\|\xi\|^k)\text{ für }\xi\to 0.$$

Sei also $\varepsilon>0$. Wir müssen zeigen, dass es ein $\delta>0$ gibt, so dass

$$|f(x+\xi)-p_k(\xi)|\le\varepsilon\|\xi\|^k\text{ für alle }\xi\in B_\delta(0).$$

Nachdem $U$ offen ist, existiert ein $\eta>0$, so dass $B_\eta(x)\subset U$. Nach Satz 9.7 existiert zu jedem $\xi\in\mathbb{R}^n$ mit $\|\xi\|<\eta$ ein $\theta\in[0,1]$, so dass

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

[61] Um ganz genau zu sein, man wählt ein $\varepsilon>0$, so $x+t\xi\in U$ für alle $t\in(-\varepsilon,1+\varepsilon)$, dies ist möglich, weil $U$ offen ist. Man betrachtet dann die Funktion $g\colon(-\varepsilon,1+\varepsilon)\to\mathbb{R},t\mapsto f(x+t\xi)$ und wendet Satz 9.3 auf diese, auf einem offenen Intervall definierte Abbildung, an

Es folgt aus der Definition von $p_k(\xi)$, dass

$$\begin{aligned}
|f(x+\xi)-p_k(\xi)| &= \left|\sum_{|\alpha|=k}\tfrac{1}{\alpha!}(D^\alpha f)(x)\cdot\xi^\alpha \;-\; \sum_{|\alpha|=k}\tfrac{1}{\alpha!}(D^\alpha f)(x+\theta\xi)\cdot\xi^\alpha\right| \\
&= \left|\sum_{|\alpha|=k}\tfrac{1}{\alpha!}((D^\alpha f)(x)-(D^\alpha f)(x+\theta\xi))\cdot\xi^\alpha\right| \\
&\le \sum_{|\alpha|=k}\left|\tfrac{1}{\alpha!}((D^\alpha f)(x)-(D^\alpha f)(x+\theta\xi))\right|\cdot|\xi^\alpha|\,.
\end{aligned}$$

Für $\alpha$ mit $|\alpha|=k$ gilt aber, dass

$$|\xi^\alpha| = |\xi_1^{\alpha_1}\cdots\xi_n^{\alpha_n}| = |\xi_1|^{\alpha_1}\cdots|\xi_n|^{\alpha_n} \le \|\xi\|^{\alpha_1}\cdots\|\xi\|^{\alpha_k} = \|\alpha\|^{|\alpha|} = \|\xi\|^k.$$

Aus der Stetigkeit der partiellen Ableitungen folgt zudem, dass es ein $\delta\in(0,\eta)$ gibt, so dass für alle $\alpha$ mit $|\alpha|=k$ und $y\in B_\delta(0)$ gilt:

$$\left|\frac{1}{\alpha!}\left((D^\alpha f)(x)-(D^\alpha f)(x+y)\right)\right\| < \frac{\varepsilon}{N},$$

wobei

$$N := \sum|\alpha| = k\frac{1}{\alpha!}.$$

Für alle $\xi\in B_\delta(0)$ folgt also, dass

$$\begin{aligned}
|f(x+\xi)-p_k(\xi)| &\le \sum_{|\alpha|=k}\tfrac{1}{\alpha!}\,|((D^\alpha f)(x)-(D^\alpha f)(x+\theta\xi))|\cdot|\xi^\alpha| \\
&\le \sum_{|\alpha|=k}\tfrac{1}{\alpha!}\tfrac{\varepsilon}{N}\cdot\|\xi\|^k \\
&= \varepsilon\|\xi\|^k\cdot\tfrac{1}{N}\cdot\sum_{|\alpha|=k}\tfrac{1}{\alpha!} \\
&= \varepsilon\|\xi\|^k.
\end{aligned}$$

□

9.4. **Spezialfälle der Taylorformel und die Hessesche Matrix.** Wir wollen jetzt mehrere Spezialfälle des Satzes von Taylor diskutieren. Zur Erinnerung, der Satz von Taylor lautet wie folgt:

**Satz. (Satz von Taylor)** *Es sei $U\subset\mathbb{R}^n$ offen und $f\colon U\to\mathbb{R}$ eine $k$-mal stetig differenzierbare Funktion. Es sei $x\in U$. Dann gilt*

$$f(x+\xi) = p_k(\xi) \;+\; o(\|\xi\|^k) \text{ für } \xi\to 0,$$

*wobei*

$$p_k(\xi) = \sum_{m=0}^{k}\sum_{|\alpha|=m}\frac{D^\alpha f(x)}{\alpha!}\xi^\alpha.$$

Wenn $k=0$, dann haben wir im vorherigen Kapitel schon gesehen, dass $p_0(\xi)=q_0(\xi)=f(x)$, und wir erhalten, dass

$$f(x+\xi) = f(x) \;+\; o(1) \text{ für } \xi\to 0.$$

Wenn $k=1$, dann ist

$$p_1(\xi) = q_0(\xi)+q_1(\xi) = f(x)+\langle\operatorname{grad} f(x),\xi\rangle,$$

und wir sehen, dass

$$f(x+\xi) = f(x) + \langle \operatorname{grad} f(x), \xi \rangle \ + o(\|\xi\|) \text{ für } \xi \to 0.^{62}$$

Der Fall $k = 2$ ist wohl der interessanteste Spezialfall. Wie wir oben gesehen habe, gilt

$$q_2(\xi) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \cdot \xi_i \xi_j.$$

Es folgt, also, dass
(9.1)

$$f(x+\xi) = f(x) + \langle \operatorname{grad} f(x), \xi \rangle \ + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \cdot \xi_i \xi_j + o(\|\xi\|^2) \text{ für } \xi \to 0.$$

Um die Notation etwas zu vereinfachen führen wir die Hessesche Matrix ein:

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen, $x \in U$ ein Punkt und $f\colon U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Die $n \times n$–Matrix

$$(\operatorname{Hess} f)(x) := \left( \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right)_{i,j=1,\dots,n}$$

wird als die *Hessesche Matrix von $f$ im Punkt $x$* bezeichnet.

Man beachte, dass aus der Vertauschbarkeit der partiellen Ableitungen (siehe Satz 5.1) folgt, dass $(\operatorname{Hess} f)(x)$ eine symmetrische Matrix ist.

Aus den Definitionen erhalten wir nun folgendes Korollar zum Satz von Taylor indem wir die Gleichung (9.1) mithilfe der Hesseschen Matrix umformulieren:

**Korollar 9.8.** *Es sei $U \subset \mathbb{R}^n$ offen, $x \in U$ ein Punkt und $f\colon U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Es sei $\xi \in \mathbb{R}^n$, so dass $x + t\xi \in U$ für alle $t \in [0,1]$. Dann gilt*

$$f(x+\xi) = f(x) + \langle \operatorname{grad} f(x), \xi \rangle \ + \frac{1}{2} \langle \xi, (\operatorname{Hess} f)(x)\xi \rangle + o(\|\xi\|^2) \textit{ für } \xi \to 0.$$

[62] Man beachte, dass diese Aussage äquivalent ist zu der Aussage, dass

$$\lim_{\xi \to 0} \frac{f(x+\xi) - f(x) - \langle \operatorname{grad} f(x), \xi \rangle}{|\xi\|} = 0,$$

welches gerade die Definition der Differenzierbarkeit am Punkt $x$ mit Differential $\operatorname{grad} f(x)$ ist.

Der Beweis folgt sofort durch explizites hinschreiben des Ausdruckes $\langle \xi, (\mathrm{Hess}\, f)(x)\xi \rangle$.

Aus Satz 9.6 (oder noch direkter, aus der ersten Behauptung im Beweis von Satz 9.6) folgt sofort folgende hilfreiche Interpretation von den Werten $\langle v, (\mathrm{Hess}\, f)(x)v \rangle$:

**Lemma 9.9.** *Es sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und es sei $f: U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Sei $v \in \mathbb{R}^n$ ein Vektor und $\varepsilon > 0$, so dass $x + tv \in U$ für alle $t \in (-\varepsilon, \varepsilon)$. Wir betrachten die Funktion*

$$\begin{array}{rcl} F : (-\varepsilon, \varepsilon) & \to & \mathbb{R} \\ t & \mapsto & f(x + tv). \end{array}$$

*Dann gilt:*

$$F''(0) = \langle v, Hess(f)(x)v \rangle.$$

9.5. **Lokale Extrema.**

*Definition.* Es sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine Funktion.

(1) Wir sagen, $x \in U$ ist ein *lokales Maximum*, wenn es eine offene Umgebung $V$ von $x$ gibt, so dass $f(y) \le f(x)$ für alle $y \in V$.
(2) Wir sagen, $x \in U$ ist ein *striktes lokales Maximum*, wenn es eine offene Umgebung $V$ von $x$ gibt, so dass $f(y) < f(x)$ für alle $y \neq x \in V$.

Ganz analog definieren wir *(striktes) lokales Minimum*. Ein (striktes) lokales Extremum ist ein (striktes) lokales Maximum oder Minimum.

Zur Erinnerung, wenn $f: (a, b) \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion ist, dann gilt für $x \in (a, b)$, dass

$$\begin{array}{rcl} x \text{ lokales Minimum} & \Rightarrow & f'(x) = 0 \text{ und } f'(x) \ge 0, \\ f'(x) = 0 \text{ und } f'''(x) > 0 & \Rightarrow & x \text{ striktes lokales Minimum} . \end{array}$$

Das Ziel ist nun, diese Aussagen aus der Analysis I auf zweimal stetig differenzierbare Funktionen $f: U \to \mathbb{R}$, $U$ offen in $\mathbb{R}^n$, zu verallgemeinern.

Es ist offensichtlich, dass das Differential von $f: U \to \mathbb{R}$ (oder äquivalent, der Gradient), die Rolle der 1. Ableitung spielt, und wir erhalten, nicht überraschend, folgenden Satz:

**Satz 9.10.** *Es sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine partiell differenzierbare Funktion. Wenn $x \in U$ ein lokales Extremum von $f$ ist, dann gilt*

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

*Beweis.* Es sei $i \in \{1, \ldots, n\}$. Wir betrachten die Funktion

$$\begin{array}{rcl} g_i : (-\varepsilon, \varepsilon) & \to & \mathbb{R} \\ t & \mapsto & f(x + te_i), \end{array}$$

dann ist $g_i$ differenzierbar und $g_i$ hat ein Extremum in $t = 0$, es folgt also, dass $g_i'(t) = 0$. Andererseits folgt aus der Definition der partiellen Ableitung, dass

$$\frac{\partial f}{\partial x_i} f(x) = g_i'(0).$$

Es folgt also, dass alle partiellen Ableitungen von $f$ am Punkt $x$ Null sind, d.h. $\operatorname{grad} f(x) = 0$. □

Die Rolle der 2. Ableitung einer Funktion $f\colon (a, b) \to \mathbb{R}$ übernimmt im mehrdimensionalen die Hessesche Matrix. Es ist auf den ersten Blick vielleicht nicht offensichtlich, was die Verallgemeinerung von $g''(x) > 0$ sein soll. Wir benötigen dazu folgende Definition:

*Definition.* Es sei $A \in M(n \times n, \mathbb{R})$ eine symmetrische Matrix.

(1) Die Matrix $A$ heißt *positiv definit*, falls

$$\langle \xi, A\xi \rangle > 0 \text{ für alle } \xi \in \mathbb{R}^n \setminus 0.$$

(2) Die Matrix $A$ heißt *positiv semidefinit*, falls

$$\langle \xi, A\xi \rangle \geq 0 \text{ für alle } \xi \in \mathbb{R}^n \setminus 0.$$

Ganz analog definieren wir negativ definit und negativ semidefinit.

(c) Die Matrix $A$ heißt *indefinit*, falls $A$ weder positiv semidefinit noch negativ semidefinit ist.

*Bemerkung.* (1) Man beachte, dass eine Matrix $A$ indefinit ist, genau dann, wenn es Vektoren $\xi, \eta \in \mathbb{R}^n$ gibt, so dass

$$\langle \xi, A\xi \rangle < 0 \text{ und } \langle \eta, A\eta \rangle > 0.$$

(2) Eine $1 \times 1$–Matrix $A = (a)$ ist positiv definit genau dann, wenn $a > 0$.

Eine symmetrische Matrix ist bekanntermaßen diagonalisierbar, und man kann an den Eigenwerten ablesen, ob $A$ definit oder indefinit ist:

**Satz 9.11.** *Es sei $A$ eine symmetrische Matrix. Dann gilt:*

$$\begin{array}{rcl} A \text{ ist positiv definit} & \Leftrightarrow & \text{alle Eigenwerte sind positiv} \\ A \text{ ist positiv semidefinit} & \Leftrightarrow & \text{alle Eigenwerte sind } \geq 0 \\ A \text{ ist negativ definit} & \Leftrightarrow & \text{alle Eigenwerte sind negativ} \\ A \text{ ist negativ semidefinit} & \Leftrightarrow & \text{alle Eigenwerte sind } \leq 0, \end{array}$$

*zudem gilt*

$$A \text{ ist indefinit} \Leftrightarrow A \text{ hat positive und negative Eigenwerte.}$$

*Beweis.* Es sei $A$ eine symmetrische Matrix. Aus der linearen Algebra wissen wir, dass es eine orthogonale Matrix $P$ gibt, so dass

$$P^tAP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Es sei nun $v \in \mathbb{R}^n$. Wir schreiben $w = P^{-1}v$, d.h. $v = Pw$. Es folgt, dass

$$\begin{aligned} \langle v, Av \rangle &= v^tAv = (Pw)^tAPw = w^t(P^tAP)w \\ &= \begin{pmatrix} w_1 & \dots & w_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \sum_{i=1}^n \lambda_i w_i^2. \end{aligned}$$

Dies ist genau dann positiv für alle $v \neq 0$, d.h. für alle $w \neq 0$, wenn alle Eigenwerte positiv sind. Die anderen Aussagen beweist man ganz analog. □

Folgender Satz folgt sofort aus Lemma 9.9 und den Resultaten aus der Analysis I:

**Satz 9.12.** *Es sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Es sei $x \in U$ ein Punkt mit* $\operatorname{grad} f(x) = 0$.

(1) *Wenn $x$ ein lokales Minimum ist, dann ist* $(\operatorname{Hess} f)(x)$ *positiv semidefinit.*
(2) *Wenn $x$ ein lokales Maximum ist, dann ist* $(\operatorname{Hess} f)(x)$ *negativ semidefinit.*
(3) *Wenn* $(\operatorname{Hess} f)(x)$ *indefinit ist, dann ist $x \in U$ weder ein lokales Minimum noch ein lokales Maximum.* [63]

*Beweis.* Sei $x$ ein lokales Minimum und sei $\xi \in \mathbb{R}^n$. Wir müssen zeigen, dass $\langle \xi, (\operatorname{Hess} f)(x)\xi \rangle \geq 0$. Wir betrachten die Funktion

$$\begin{aligned} (-\varepsilon, \varepsilon) &\to \mathbb{R} \\ t &\mapsto f(x + tv). \end{aligned}$$

Die Funktion nimmt ein Minimum in $t = 0$ an, d.h. $F''(0) \geq 0$. Aus Lemma 9.9 folgt nun, dass $\langle \xi, (\operatorname{Hess} f)(x)\xi \rangle = F''(0) \geq 0$. □

**Satz 9.13.** *Es sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Es sei $x \in U$ ein Punkt mit* $\operatorname{grad} f(x) = 0$.

[63] In diesem Fall sagen wir, dass $x$ ein Sattelpunkt ist.

*Dann gilt:*

$$\begin{array}{rcl} (\operatorname{Hess} f)(x) \text{ positiv definit} & \Rightarrow & x \text{ ist striktes lokales Minimum,} \\ (\operatorname{Hess} f)(x) \text{ negativ definit} & \Rightarrow & x \text{ ist striktes lokales Maximum.} \end{array}$$

*Beweis.* Wir schreiben $H := (\operatorname{Hess} f)(x)$. Wir betrachten den Fall, dass $H$ positiv definit ist, d.h. $\langle \xi, A\xi\rangle > 0$ für alle $\xi \in \mathbb{R}^n \setminus 0$. Nachdem $\operatorname{grad} f(x) = 0$ folgt aus dem Satz von Taylor, dass

$$f(x+\xi) = f(x) + \frac{1}{2}\langle \xi, H\xi\rangle + \varphi(\xi),$$

wobei $\varphi(\xi) = o(\|\xi\|^2)$ für $\xi \to 0$.

Wir müssen nun zuerst den Ausdruck $\langle \xi, H\xi\rangle$ genauer studieren.

*Behauptung.* Es gibt ein $\alpha > 0$, so dass

$$\langle \xi, H\xi\rangle \geq \alpha\|\xi\|^2 \text{ für alle } \xi \in \mathbb{R}^n.$$

Für den Beweis der Behauptung betrachten wir die Funktion

$$\begin{array}{rcl} S := \{\xi \in \mathbb{R}^n \mid \|\xi\| = 1\} & \to & \mathbb{R} \\ \xi & \mapsto & \langle \xi, H\xi\rangle. \end{array}$$

Diese Funktion ist stetig (weil sie ein Polynom in $\xi_1, \dots, \xi_n$ ist). Nachdem $S$ kompakt ist (siehe Kapitel 3.2), folgt aus Korollar 3.7, dass die Funktion ein Minimum annimmt. Es gibt also ein $\xi_0 \in S$, so dass

$$\langle \xi, H\xi\rangle \geq \langle \xi_0, H\xi_0\rangle \text{ für alle } \xi \in S.$$

Wir setzen $\alpha := \langle \xi_0, H\xi_0\rangle$. Man beachte, dass $\alpha > 0$, weil $H$ nach Voraussetzung positiv definit ist.

Sei nun $\xi \in \mathbb{R}^n \setminus 0$ beliebig. Wir setzen $\lambda := \|\xi\|$ und $\xi' = \frac{1}{\lambda}\|\xi\|$. Dann gilt $\xi = \lambda \cdot \xi'$ wobei $\xi' \in S$. Dann folgt

$$\langle \xi, H\xi\rangle = \langle \lambda \cdot \xi', H(\lambda \cdot \xi')\rangle = \lambda^2\langle \xi', H \cdot \xi'\rangle \geq \|\xi\|^2 \cdot \alpha.$$

Wir haben damit die Behauptung bewiesen.

Nachdem $\varphi(\xi) = o(\|\xi\|^2)$ für $\xi \to 0$ gibt es insbesondere ein $\delta > 0$, so dass

$$|\varphi(\xi)| \leq \frac{\alpha}{4}\|\xi\|^2 \text{ für alle } \xi \text{ mit } \|\xi\| < \delta.$$

Dann folgt für alle $\xi \in B_\delta(x)$, dass

$$\begin{array}{rcl} f(x+\xi) & = & f(x) + \frac{1}{2}\langle \xi, H\xi\rangle + \varphi(\xi) \\ & \geq & f(x) + \frac{1}{2}\alpha\|\xi\|^2 + \varphi(\xi) \\ & \geq & f(x) + \frac{1}{2}\alpha\|\xi\|^2 - \frac{\alpha}{4}\|\xi\|^2 \\ & = & f(x) + \frac{\alpha}{4}\|\xi\|^2. \end{array}$$

Wir sehen also, dass $f$ ein striktes lokales Minimum in $x$ annimmt.

Die zweite Aussage des Satzes wird ganz analog bewiesen (oder man kann die Aussage durch betrachten der Funktion $-f$ auf die erste Aussage zurückführen.) $\square$

*Beispiel.* Die Hessesche Matrix der Funktion

$$\begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x^2+y^2 \end{array}$$

am Punkt $(0,0)$ ist

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

also positiv definit. Zudem ist $\operatorname{grad} f(0,0) =$ der Nullvektor. Die Funktion $f$ erfüllt also alle Bedingungen von Satz 9.13, der Punkt $(0,0)$ ist also ein lokales Minimum. In diesem Fall kann man dies natürlich auch direkt zeigen: es ist offensichtlich, dass $f(x,y) > 0 = f(0,0)$ für alle $(x,y) \neq (0,0)$.

### 9.6. **Lokale Extrema mit Nebenbedingungen.** [64]

*Definition.* Es sei $M \subset \mathbb{R}^n$ eine $k$–dimensionale Untermannigfaltigkeit und $f\colon M \to \mathbb{R}$ eine Funktion. Wir sagen $x \in M$ ist ein lokales Maximum, wenn es eine offene Menge $U \subset \mathbb{R}^n$ gibt, so dass $f(y) \leq f(x)$ für alle $y \in M \cap U$.

Ganz analog führen wir auch striktes lokales Maximum und (striktes) lokales Minimum einer Funktion $f\colon M \to \mathbb{R}$ ein.

Wir wollen jetzt die lokalen Extreme von Funktionen auf Untermannigfaltigkeiten bestimmen.

Es sei $U \subset \mathbb{R}^n$ offen, $g\colon U \to \mathbb{R}$ eine $C^1$–Funktion und $M := g^{-1}(z)$ das Urbild eines regulären Wertes $z \in \mathbb{R}$. Es sei zudem $f\colon U \to \mathbb{R}$ eine weitere Funktion. Ein lokales Extremum von $f\colon M = g^{-1}(z) \to \mathbb{R}$ nennt man auch *lokales Extremum von $f\colon U \to \mathbb{R}$ unter der Nebenbedingung $g = z$.*

**Satz 9.14.** *Es sei $U \subset \mathbb{R}^n$ offen, $g\colon U \to \mathbb{R}$ eine $C^1$–Funktion und $M := g^{-1}$ das Urbild eines regulären Wertes $z \in \mathbb{R}$. Es sei $f\colon U \to \mathbb{R}$ eine $C^1$–Funktion. Wenn $x \in M$ ein lokales Extremum von $f$ auf $M$ ist, dann existiert ein $\lambda \in \mathbb{R}$, so dass*

$$\operatorname{grad} f(x) = \lambda \cdot \operatorname{grad} g(x).$$

Die reelle Zahl $\lambda$ wird *Lagrangemultiplikator* genannt.

[64] Der Stoff dieses Kapitels basiert auf Forster, Analysis II, Kapitel 9.

*Beweis.* Sei also $x \in M$ ein lokales Extremum der Einschränkung von $f$ auf $M$. In Satz 8.5 hatten wir gesehen, dass

$$T_xM = (\operatorname{grad} g(x))^\perp.$$

Anders ausgedrückt, es gilt

$$(T_xM)^\perp = \mathbb{R} \cdot \operatorname{grad} g(x) = \{\lambda \cdot \operatorname{grad} g(x) \mid \lambda \in \mathbb{R}\}. \tag{9.2}$$

Wir wollen zeigen, dass $\operatorname{grad} f(x) = \lambda \cdot \operatorname{grad} g(x)$ für ein $\lambda \in \mathbb{R}$. Es folgt aus (9.2), dass es genügt zu zeigen, dass $\operatorname{grad} f(x)$ orthogonal zu $T_xM$ ist.

Sei also $v \in T_xM$. Dann existiert eine stetig differenzierbare Kurve $\gamma\colon (-\varepsilon, \varepsilon) \to M$, so dass $\gamma(0) = x$ und $\gamma'(0) = v$. Die Funktion

$$\begin{array}{rcl} F : (-\varepsilon, \varepsilon) & \to & \mathbb{R} \\ t & \mapsto & f(\gamma(t)) \end{array}$$

ist stetig differenzierbar und nimmt ein lokales Extremum in $t = 0$ ein. Es folgt, dass $F'(0) = 0$. Aus der Kettenregel folgt, dass

$$0 = F'(0) = \frac{d}{dt} f(\gamma(t))|_{t=0} = \operatorname{grad} f(\gamma(0)) \cdot \gamma'(0) = \operatorname{grad} f(x) \cdot v,$$

d.h. $\operatorname{grad} f(x)$ ist orthogonal zu $v$. Wir haben also bewiesen, dass $\operatorname{grad} f(x)$ orthogonal zu allen Vektoren in $T_xM$ ist, d.h. $\operatorname{grad} f(x)$ ist orthogonal zu $T_xM$. □

*Bemerkung.* Der Satz kann auch verallgemeinert werden zu Funktionen auf Mannigfaltigkeiten, welche durch mehrere Gleichungen

$$g_1 = 0, g_2 = 0, \dots, g_s = 0$$

beschrieben sind. Siehe Forster, Analysis II, Seite 110 für Details.

*Beispiel.* Wir betrachten die Funktionen

$$\begin{array}{rcl} g\colon \mathbb{R}^3 & \to & \mathbb{R} \\ g(x, y, z) & = & x^2 + y^2 + z^2 \end{array}$$

und

$$\begin{array}{rcl} f\colon \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y, z) & \mapsto & xy. \end{array}$$

Wir wollen das absolute Maximum von $f$ auf $S^2 := g^{-1}(1)$ bestimmen. (Oder anders ausgedrückt, das absolute Maximum von $f$ unter der Nebenbedingung $g = 1$.)

Es ist

$$\begin{array}{rcl} \operatorname{grad} g & = & (2x, 2y, 2z), \\ \operatorname{grad} f & = & (y, x, 0). \end{array}$$

Die beiden Vektoren sind parallel, genau dann, wenn $z = 0$ und $x = y$ oder $x = -y$. Es gibt genau 4 Punkte auf $S^2$, d.h. genau 4 Punkte mit $x^2 + y^2 + z^2 = 1$, so dass $z = 0$ und $x = \pm y$, nämlich

$$\begin{aligned} P_+ &= (\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}, 0) \\ P_- &= (-\tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2}, 0) \\ Q_+ &= (\tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2}, 0) \\ Q_- &= (-\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}, 0). \end{aligned}$$

Satz 9.3 besagt, dass das absolute Maximum (bzw. Minimum) von $f \colon S^2 \to \mathbb{R}$ an einem der Punkte $P_\pm, Q_\pm$ angenommen wird.

Es gilt $f(P_\pm) = \frac{1}{2}$ und $f(Q_\pm) = -\frac{1}{2}$. Also ist $\frac{1}{2}$ das absolute Maximum von $f$ und $-\frac{1}{2}$ das absolute Minimum von $f \colon S^2 \to \mathbb{R}$.

## 10. Differentialgleichungen: Definition und Beispiele

### 10.1. **Definition von Differentialgleichungen.**

*Definition.* Es sei $A \subset \mathbb{R}^n$ eine Teilmenge. Ein Vektorfeld auf $A$ ist eine stetige Abbildung $f : A \to \mathbb{R}^n$.

Ein Vektorfeld auf $A$ ordnet also jedem Punkt in $A \subset \mathbb{R}^n$ einen Vektor (oder Richtung) in $\mathbb{R}^n$ zu.

*Definition.* Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge [65] und

$$\begin{aligned} f : G &\to \mathbb{R}^n \\ (t, y) &\mapsto f(t, y) \end{aligned}$$

eine stetige Funktion. Wir nennen eine differenzierbare Funktion $\varphi : I \to \mathbb{R}^n$, $I \subset \mathbb{R}$, eine *Lösung der Differentialgleichung* $y' = f(t, y)$, wenn gilt

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

(Damit die rechte Seite definiert ist, muss gelten, dass $(t, \varphi(t)) \in G$ für alle $t \in I$.)

*Bemerkung.* Nehmen wir an, dass $G = \mathbb{R} \times A$ mit $A \subset \mathbb{R}^n$. Eine Abbildung

$$\begin{aligned} f : \mathbb{R} \times A &\to \mathbb{R}^n \\ (t, y) &\mapsto f(t, y) \end{aligned}$$

definiert also zu für fest gewähltes $t \in \mathbb{R}$ ('Zeitpunkt $t$') ein Vektorfeld $y \mapsto f(t, y)$. Anders ausgedrückt, $f$ ist ein Vektorfeld auf $A$ welches sich mit der Zeit $t$ verändern kann. Eine Lösung der Differentialgleichung $y' = f(t, y)$ ist dann eine Kurve $\varphi$, so dass zu jedem Zeitpunkt $t$ die

[65] In den meisten Anwendungen ist $G = I \times A$, wobei $I \subset \mathbb{R}$ ein Intervall ist und $A \subset \mathbb{R}^n$ eine Teilmenge.

Ableitung gerade durch den Vektor des Vektorfeldes am Punkt $\varphi(t)$ zur Zeit $t$ gegeben ist.

*Beispiel.* Es sei $A \subset \mathbb{R}^3$ eine 'Teilmenge eines Sees' und $f(t,a)$ die Geschwindigkeit zum Zeitpunkt $t$ am Punkt $a$. Dann entspricht die Lösung der Differentialgleichung $y' = f(t,y)$ gerade der Bahn eines Wassermoleküls.

*Beispiel.* Es sei $B = B_1(0) \in \mathbb{R}^3$ die offene Kugel von Radius $r$ um den Nullpunkt. Wir betrachten[66]

$$\begin{array}{rcl} f : \mathbb{R}_{>0} \times B & \to & \mathbb{R}^3 \\ (t,y) & \mapsto & \frac{1}{t}y \end{array}$$

und die dazugehörige Differentialgleichung $y' = f(t,y) = \frac{1}{t}y$. Für jedes $w \in \mathbb{R}^n \neq \{0\}$ ist dann

$$\begin{array}{rcl} \varphi : (0, \frac{1}{\|w\|}) & \to & \mathbb{R}^3 \\ t & \mapsto & tw \end{array}$$

eine Lösung der Differentialgleichung $y' = f(t,y)$. In der Tat, denn es gilt

$$\varphi'(t) = w = \frac{1}{t}tw = f(t,tw) = f(t,\varphi(t)),$$

und $\varphi(t) \in B$ für alle $t \in (0, \frac{1}{\|w\|})$. Man beachte, dass alle diese Lösungen der Differentialgleichung nur auf einem endlichen Intervall definiert sind, außer der konstanten Lösungen $\varphi(t) = 0$ gibt es keine Lösung, welche auf ganz $\mathbb{R}_{>0}$ definiert ist.

Es gibt kein allgemeines Verfahren um Differentialgleichungen zu lösen. Wir werden in diesem Kapitel verschiedene Verfahren kennen lernen um bestimmte Typen von Differentialgleichung zu lösen. In den folgenden Kapiteln werden wir dann die Theorie der Differentialgleichungen behandeln.

10.2. **Differentialgleichungen mit getrennten Variablen.** Wir betrachten jetzt nur Differentialgleichungen im Fall $n = 1$. Es seien $I, J \subset \mathbb{R}$ offene Intervalle und $f : I \to \mathbb{R}$ sowie $g : J \to \mathbb{R}$ zwei stetige Funktionen, so dass $g(y) \neq 0$ für alle $y \in J$. Eine Differentialgleichung vom Typ

$$y' = f(t)g(y)$$

heißt *Differentialgleichung mit getrennten Variablen.*

---

[66]Das Vektorfeld zeigt an einem Punkt $x \in B$ 'nach außen' und hat Länge $\frac{1}{t}\|x\|$, d.h. mit zunehmender Zeit werden die Vektoren immer kürzer.

**Satz 10.1.** *Es seien $I, J \subset \mathbb{R}$ offene Intervalle und $f : I \to \mathbb{R}$ sowie $g : J \to \mathbb{R}$ zwei stetige Funktionen, so dass $g(y) \neq 0$ für alle $y \in J$. Sei $(t_0, y_0) \in I \times J$ ein Punkt. Wir definieren*

$$\begin{array}{rcl} F : I & \to & \mathbb{R} \\ t & \mapsto & \int_{t_0}^t f(s)ds \end{array} \quad \text{und} \quad \begin{array}{rcl} G : J & \to & \mathbb{R} \\ y & \mapsto & \int_{y_0}^y \frac{1}{g(s)}ds. \end{array}$$

*Es sei $I' \subset I$ ein Intervall, so dass $t_0 \in I'$ und $F(I') \subset G(J)$. Dann ist* [67]

$$\begin{array}{rcl} \varphi : I' & \to & \mathbb{R} \\ t & \mapsto & G^{-1}(F(t)) \end{array}$$

*die einzige Lösung der Differentialgleichung $y' = f(t)g(y)$ auf dem Intervall $I'$, welche die Anfangsbedingung $\varphi(t_0) = y_0$ erfüllt.*

*Beweis.* Wir zeigen zuerst, dass $\varphi : I' \to \mathbb{R}$ in der Tat eine Lösung ist. Es ist [68]

$$\begin{aligned} \varphi'(t) &= (G^{-1}(F(t)))' \\ &= (G^{-1})'(F(t)) \cdot F'(t) \\ &= \frac{1}{G'(G^{-1}(F(t)))} \cdot f(t) \\ &= g(G^{-1}(F(t))) \cdot f(t) = g(\varphi(t)) \cdot f(t). \end{aligned}$$

Wir zeigen nun die Eindeutigkeit der Lösung. Sei also $\psi : I' \to \mathbb{R}$ eine beliebige Lösung der Differentialgleichung $y' = f(t)g(y)$ mit $\psi(t_0) = y_0$. Wir wollen zeigen, dass $G(\psi(t)) = F(t)$ für alle $t \in I'$. Es gilt

$$\begin{aligned} G(\psi(t)) = \int_{\psi(t_0)}^{\psi(t)} \frac{1}{g(u)}\, du &= \int_{t_0}^t \frac{\psi'(s)}{g(\psi(s))}\, ds \\ &= \int_{t_0}^t f(s)\, ds = F(t). \end{aligned}$$

Die erste Gleichheit folgt aus der Definition von $G$, die zweite aus der Substitutionsregel (hier 'rückwarts' angewandt auf $u = \psi(s)$), und die dritte folgt aus der Tatsache, dass $\psi'(s) = f(s)g(\psi(s))$. $\square$

*Bemerkung.* Man kann sich die Lösung der Differentialgleichung $y' = f(t)g(y)$ mit folgender suggestiver, allerdings sinnloser 'Rechnung' herleiten:

$$\begin{aligned} \frac{dy}{dt} = f(t)g(y) \quad &\Rightarrow \quad \frac{1}{g(y)}dy = f(t)dt \\ &\Rightarrow \quad \int \frac{dy}{g(y)} = \int f(t)dt + \text{const} \\ &\Rightarrow \quad G(y) = F(t) \Rightarrow y = G^{-1}(F(t)). \end{aligned}$$

[67] Man beachte, dass die Voraussetzung, dass $g(y) \neq 0$ impliziert, dass $G'(y) = \frac{1}{g(y)}$ durchgehend das gleiche Vorzeichen besitzt, d.h. $G$ ist eine streng monotone Funktion, insbesondere besitzt $G$ eine Umkehrfunktion.

[68] Hier verwenden wir, dass i.a. gilt: $(h^{-1})'(t) = \frac{1}{h'(h^{-1}(t))}$.

*Beispiel.* Wir betrachten die Differentialgleichung

$$y' = 2ty^2,$$

welche auf $G = \mathbb{R}\times\mathbb{R}$ definiert ist. Es sei $c \neq 0$. Wir suchen die Lösung der Differentialgleichung $y' = 2ty^2 = 2t \cdot y^2$ mit der Anfangsbedingung $\varphi(0) = c$. Dann ist

$$\begin{aligned} F(t) &:= \int_0^t 2s\,ds = t^2, \\ G(y) &:= \int_c^y \tfrac{1}{s^2}\,ds = -\tfrac{1}{y} + \tfrac{1}{c}. \end{aligned}$$

Dann ist

$$\varphi(t) = G^{-1}(F(t)) = \frac{1}{\frac{1}{c} - t^2}$$

die Lösung der Differentialgleichung mit $\varphi(0) = c$. Man beachte, dass der maximale Definitionsbereich von $\varphi$ gegeben ist durch $I = \mathbb{R}$, falls $c < 0$ und

$$I = \{t \in \mathbb{R} \,|\, |t| < \sqrt{1/c} = c^{-1/2}\}.$$

Wenn $c < 0$ divergiert also die Lösung für $t \to \pm c^{-1/2}$, insbesondere gibt es keine Lösung welche auf ganz $\mathbb{R}$ definiert ist.

10.3. **Lineare homogene Differentialgleichungen.** Es sei $I \subset \mathbb{R}$ ein Intervall und $a : I \to \mathbb{R}$ eine stetige Funktion. Wir nennen

$$y' = a(t)y$$

*homogene lineare Differentialgleichung*. Eine homogene Differentialgleichung ist ein wichtiger Spezialfall einer Differentialgleichung mit getrennten Variablen.

**Satz 10.2.** *Es sei $I \subset \mathbb{R}$ ein Intervall und $a : I \to \mathbb{R}$ eine stetige Funktion, es sei $t_0 \in I$ und $c \in \mathbb{R}$. Dann besitzt die homogene Differentialgleichung*

$$y' = a(t)y$$

*genau eine Lösung $\varphi : I \to \mathbb{R}$, welche die Anfangsbedingung $\varphi(t_0) = c$ erfüllt, nämlich*

$$\varphi(t) = c \cdot \exp\left(\int_{t_0}^t a(s)\,ds\right).$$

Es sei $a \in \mathbb{R}$. Dann besagt der Satz insbesondere, dass es genau eine differenzierbare Funktion $\varphi : \mathbb{R} \to \mathbb{R}$ mit $\varphi'(t) = a\varphi(t)$ und $\varphi(t_0) = c$ gibt, nämlich $\varphi(t) = c \cdot e^{a(t-t_0)}$.

*Beweis.* Wir setzen

$$\varphi(t) = c \cdot \exp\left(\int_{t_0}^t a(s)\,ds\right),$$

dann folgt sofort aus $\exp(g(t))' = g'(t) \cdot \exp(t)$ und dem Hauptsatz der Integral– und Differentialrechnung, dass $\varphi$ in der Tat eine Lösung der Differentialgleichung $y' = a(t)y$ ist. Zudem ist offensichtlich, dass $\varphi(t_0) = c$.

Nehmen wir nun an, $\psi : I \to \mathbb{R}$ sei eine beliebige Lösung der Differentialgleichung mit $\psi(t_0) = c$. Wir wollen zeigen, dass $\psi(t) = \varphi(t)$ für alle $t$. Nachdem $\psi(t_0) = \varphi(t_0)$ genügt es zu zeigen, dass die Quotientenfunktion $t \mapsto \frac{\psi(t)}{\varphi(t)}$ konstant ist. Dies ist aber der Fall, nachdem

$$\left(\psi(t) \cdot \frac{1}{\varphi(t)}\right)' = \frac{\psi'(t)\varphi(t) - \psi(t)\varphi'(t)}{\varphi(t)^2} = \frac{a(t)\psi(t)\varphi(t) - \psi(t)a(t)\varphi(t)}{\varphi(t)^2} = 0.$$

□

10.4. **Lineare inhomogene Differentialgleichungen.** Es sei $I \subset \mathbb{R}$ ein Intervall und $a, b : I \to \mathbb{R}$ stetige Funktionen. Wir nennen

$$y' = a(t)y + b(t)$$

*inhomogene lineare Differentialgleichung.*

Sei nun $t_0 \in I$ und $c \in \mathbb{R}$. Das Ziel ist eine Lösung $\varphi$ der inhomogenen linearen Differentialgleichung zu finden, welche die Anfangsbedingung $\varphi(t_0) = c$ erfüllt. Sei dazu $\psi : I \to \mathbb{R}$ eine Lösung der *homogenen* linearen Differentialgleichung $y' = a(t)y$ zu der Anfangsbedingung $\psi(t_0) = 1$. Sei $u : I \to \mathbb{R}$ nun eine beliebige differenzierbare Funktion. Wir wollen untersuchen, unter welchen Bedingungen $\varphi(t) = \psi(t)u(t)$ eine Lösung der inhomogenen linearen Differentialgleichung ist, welche die Anfangsbedingung $\varphi(t_0) = c$ erfüllt.

Es ist

$$\varphi' = (\psi \cdot u)' = \psi' \cdot u + \psi \cdot u' = a \cdot \psi \cdot u + \psi \cdot u' = a\varphi + \psi \cdot u'.$$

Wir sehen also, dass $\varphi' = a\varphi + b$ genau dann, wenn $b = \psi \cdot u'$, d.h. wenn

$$u(t) = \int_{t_0}^{t} \psi(s)^{-1} b(s)\, ds + \text{const}.$$

Die Anfangsbedingung $\varphi(t_0) = c$ ist zudem genau dann erfüllt, wenn wir die Integrationskonstante gleich $c$ setzen. Wir haben also folgenden Satz bewiesen:

**Satz 10.3.** *Es sei $I \subset \mathbb{R}$ ein Intervall und $a, b : I \to \mathbb{R}$ eine stetige Funktion. Es sei $t_0 \in I$ und $c \in \mathbb{R}$. Wir setzen*

$$\psi(t) = \exp\left(\int_{t_0}^{t} a(s)\, ds\right).$$

*Dann ist*

$$\begin{array}{rcl} \varphi : I & \to & \mathbb{R} \\ t & \mapsto & \psi(t) \cdot \left(c + \int_{t_0}^{t} \psi(s)^{-1} b(s)\, ds\right) \end{array}$$

*die einzige Lösung der Differentialgleichung $y' = a(t)y + b(t)$, welche die Anfangsbedingung $\varphi(t_0) = c$ erfüllt.*

*Beweis.* Wir haben vor dem Satz schon bewiesen, dass $\varphi$ eine Lösung der Differentialgleichung ist (und man kann das auch noch mal leicht durch einsetzen nachweisen), es ist zudem $\varphi(t_0) = 1$ und daher $\psi(t_0) = c$. Wir müssen nun noch zeigen, dass $\psi$ die einzige Lösung der Differentialgleichung $y' = a(t)y+b(t)$ ist, welche die Anfangsbedingung $\varphi(t_0) = c$ erfüllt. Sei dazu $\phi(t)$ eine weitere Lösung. Nachdem $\psi(t) \neq 0$ für alle $t$ existiert eine Funktion $v$ (nämlich $v(t) = \frac{\phi(t)}{\psi(t)}$), so dass $\phi(t) = \psi(t)v(t)$. Aber vor dem Satz haben wir schon gesehen, dass es genau eine solche Funktion $v$ gibt, also $v = u$ und $\phi = \varphi$. $\square$

*Beispiel.* Betrachten wir die Differentialgleichung

$$y' = 2ty + t^3.$$

Wir suchen eine Lösung welche die Anfangsbedingung $\varphi(0) = c$ erfüllt. Die zugehörige homogene Differentialgleichung $y' = 2xy$ besitzt die Lösung

$$\psi(t) = \exp\left(\int_0^t 2s\, ds\right) = \exp(t^2),$$

welche die Anfangsbedingung $\psi(0) = 1$ erfüllt. Aus dem obigen Satz folgt, dass

$$\varphi(t) = e^{t^2} \cdot \left(c + \int_0^t s^3 e^{-s^2}\, ds\right)$$

die Lösung der inhomogenen Differentialgleichung mit der Anfangsbedingung $\varphi(0) = c$ ist. Mithilfe der Substitution $u = -s^2$ und durch partielle Integration bestimmt man, dass

$$\int_0^t s^3 e^{-s^2}\, ds = \frac{1}{2} - \frac{1}{2}(t^2+1)e^{-t^2},$$

also

$$\varphi(t) = (c + \frac{1}{2})e^{t^2} - \frac{1}{2}(t^2+1).$$

## 11. Der Existenz– und Eindeutigkeitssatz

Sei
$$y' = f(t, y)$$
eine Differentialgleichung. Es stellen sich folgende Fragen:

(1) Besitzt die Differentialgleichung immer eine Lösung?
(2) Ist eine Lösung, zu einer gegebenen Anfangsbedingung, eindeutig bestimmt?

### 11.1. **Der Eindeutigkeitssatz.**

*Beispiel.* Wir betrachten die Differentialgleichung
$$y' = y^{2/3}.$$
Offensichtlich ist $\varphi(t) := 0$ eine Lösung mit $\varphi(0) = 0$. Andererseits gilt für
$$\psi(t) := \frac{1}{27}t^3$$
dass $\psi(0) = 0$ und
$$\psi(t)' = \frac{1}{9}t^2 = \left(\frac{1}{27}t^3\right)^{2/3} = \psi(t)^{2/3}.$$
Wir sehen also, dass $\psi(t)$ ebenfalls eine Lösung für die gleiche Anfangsbedingung ist. Die Differentialgleichung $y' = y^{2/3}$ besitzt also keine eindeutig bestimmte Lösung.

Man kann das obige Beispiel noch wie folgt verallgemeinern: es sei $a < 0$ und $b > 0$, dann kann man leicht nachweisen, dass
$$\phi(t) := \begin{cases} \frac{1}{27}(t-a)^3, & \text{für } t < a, \\ 0, & \text{für } t \in [a, b], \\ \frac{1}{27}(t-b)^3, & \text{für } t > b. \end{cases}$$
ebenfalls eine Lösung ist, welche die Anfangsbedingung $\phi(0) = 0$ erfüllt.

Wir sehen also, dass eine Differentialgleichung i.a. keine eindeutig bestimmte Lösung besitzt. Das obige Beispiel beinhaltet die Funktion $y^{2/3}$, welche bekanntermaßen nicht differenzierbar ist, und es liegt nahe zu vermuten, dass das unerwartete Verhalten der Differentialgleichung etwas damit zu tun hat, dass $y^{2/3}$ nicht differenzierbar ist und am Punkt $y = 0$ ‘unendliche Steigung’ besitzt.

*Definition.* Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und es sei
$$\begin{array}{rcl} f : G & \to & \mathbb{R}^n \\ (t, y) & \mapsto & f(t, y) \end{array}$$
eine Funktion.

(1) Wir sagen *$f$ ist Lipschitz–stetig bezüglich der Variablen $y$ mit Lipschitzkonstante* $L \geq 0$, wenn

$$\|f(t,y) - f(t,y')\| \leq L\|y-y'\| \text{ für alle } (t,y),(t,y') \in G.$$

(2) Wir sagen $f$ ist *lokal Lipschitz–stetig bezüglich der Variablen $y$*, falls jeder Punkt $(a,b) \in G$ eine Umgebung $U$ besitzt, so dass $f: G \cap U \to \mathbb{R}^n$ Lipschitz–stetig bezüglich der Variablen $y$ ist.

*Beispiel.* (1) Die Funktion

$$\begin{array}{rcl} f: \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (t,y) & \mapsto & y^2 + t^{\frac{2}{3}} \end{array}$$

ist nicht Lipschitz–stetig bezüglich der Variablen $y$ aber $f$ ist lokal Lipschitz–stetig bezüglich der Variablen $y$. Beides ist nicht schwierig zu zeigen, und ist Teil von Übungsblatt 12.

(2) Die Funktion

$$\begin{array}{rcl} f: \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (t,y) & \mapsto & y^{2/3} \end{array}$$

ist nicht lokal Lipschitz–stetig bezüglich der Variablen $y$ am Punkt $(0,0$, insbesondere ist $f$ nicht Lipschitz–stetig bezüglich der Variablen $y$.

Im folgenden werden wir kurz schreiben '$f$ ist (lokal) Lipschitz–stetig', und nicht mehr erwähnen, dass wir die Lipschitzstetigkeit bezüglich der Variablen $y$ betrachten.

**Satz 11.1.** *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und es sei*

$$\begin{array}{rcl} f: G & \to & \mathbb{R}^n \\ (t,y) & \mapsto & f(t,y) \end{array}$$

*eine Funktion, welche bezüglich der Variablen $y = (y_1, \ldots, y_n)$ stetig partiell differenzierbar ist. Dann ist $f$ lokal Lipschitz–stetig.*

*Beweisskizze.* Wir betrachten den Fall $G = \mathbb{R} \times \mathbb{R}^n$, der allgemeine Fall wird ganz ähnlich bewiesen. Sei also $(a,b) \in \mathbb{R} \times \mathbb{R}^n$.

Sei erst einmal $v \in \mathbb{R}^n$ beliebig. Dann besagt Korollar 6.10, dass

$$\|f(a,b+v) - f(a,b)\| \leq M\|v\|,$$

wobei $M$ das Maximum der Normen von den Matrizen

$$\left(\frac{\partial f_i}{\partial y_j}\right)_{i,j=1,\ldots,n}$$

auf dem Geradenstück zwischen $(a,b)$ und $(a,b+v)$ ist. Nach der Bemerkung auf Seite 49, gilt für eine beliebige $n \times n$–Matrix $A$, dass

$$\|A\| \leq n \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}| \leq n^3 \max\{|a_{ij}| \mid i,j \in \{1,\ldots,n\}\}.$$

In unserem Falle gibt uns also eine Schränke für die partiellen Ableitungen $\frac{\partial f_i}{\partial y_j}$ eine Schranke für $M$.

Um solch eine Schranke zu finden, wählen wir ein beliebiges $r > 0$ und betrachten

$$V := \{(t,y) \in \mathbb{R} \mid |t-a| \leq r \text{ und } \|y-b\| \leq r\}.$$

Nachdem $V$ kompakt und nachdem die partiellen Ableitungen $\frac{\partial f}{\partial y_i}$ nach Voraussetzung stetig sind, existiert ein $C \geq 0$, so dass

$$\left|\frac{\partial f}{\partial y_i}\right| \leq C \text{ für alle } (t,y) \in V.$$

Für alle $v \in \mathbb{R}^n$ mit $\|v\| \leq r$ folgt dann aus obiger Diskussion, dass

$$\|f(a,b+v) - f(a,b)\| \leq M\|v\| \leq n^3 C\|v\|.$$

Wir sehen nun, dass die offene Menge

$$V := \{(t,y) \in \mathbb{R} \mid |t-a| < r \text{ und } \|y-b\| < r\}$$

und die Konstante $L = Cn^3$ die gewünschten Eigenschaften besitzen.

$\square$

**Satz 11.2. (Eindeutigkeitssatz)** *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ eine Funktion, welche lokal Lipschitz–stetig ist. Wenn die Differentialgleichung*

$$y' = f(t,y)$$

*zwei Lösungen*

$$\varphi, \psi : I \to \mathbb{R}^n \ \ (I \subset \mathbb{R} \text{ ein Intervall})$$

*besitzt, so dass $\varphi(t_0) = \psi(t_0)$ für ein $t_0 \in I$, dann gilt*

$$\varphi(t) = \psi(t) \text{ für alle } t \in I.$$

Um den Eindeutigkeitssatz zu beweisen benötigen wir folgenden Hilfssatz:

**Satz 11.3.** *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ eine stetige Abbildung, sei $(a,c) \in G$, und sei $I \subset \mathbb{R}$ ein Intervall mit $a \in I$. Es sei $\varphi : I \to \mathbb{R}^n$ eine Kurve. Dann sind folgende beiden Aussagen äquivalent:*

(1) *$\varphi$ ist eine Lösung der Differentialgleichung $y' = f(t,y)$ bezüglich der Anfangsbedingung $\varphi(a) = c$,*

(2) *es gilt*

$$\varphi(t) = c + \int_a^t f(s, \varphi(s))\, ds \text{ für alle } t \in I.$$

Der Satz besagt also, dass eine Funktion $\varphi$ ein Differentialgleichung erfüllt, genau dann, wenn $\varphi$ eine gewisse Integralgleichung erfüllt. Der Beweis des Satzes ist eine Anwendung des Hauptsatzes der Differential– und Integralrechnung:

*Beweis.* Nehmen wir zuerst an, dass $\varphi$ eine Lösung der Differentialgleichung bezüglich der Anfangsbedingung $\varphi(a) = c$. Dann folgt, dass

$$\int_a^t f(s, \varphi(s))\, ds = \int_a^t \varphi'(s)\, ds = \varphi(t) - \varphi(a) = \varphi(t) - c.$$

Nehmen wir nun an, dass $\varphi$ die Integralgleichung erfüllt. Dann gilt $\varphi(a) = c$ und aus dem Hauptsatz der Differential– und Integralrechnung folgt, dass

$$\varphi(t) = \frac{d}{dt} \int_a^t f(s, \varphi(s))\, ds = f(t, \varphi(t)).$$

□

Wir wenden uns jetzt dem Beweis des Eindeutigkeitssatzes zu:

*Beweis von Satz 11.2.* Wir behandeln nur den Fall, dass $I = \mathbb{R}$. Der allgemeine Fall wird ganz ähnlich bewiesen. Wir zeigen zuerst folgende Behauptung:

*Behauptung.* Es sei $a \in I$, so dass $\varphi(a) = \psi(a)$. Dann existiert ein $\varepsilon > 0$, so dass

$$\varphi(t) = \psi(t) \text{ für alle } t \in (a - \varepsilon, a + \varepsilon).$$

Aus Satz 11.3 folgt, dass

$$\varphi(t) - \psi(t) = \int_a^t f(s, \varphi(s)) - f(s, \psi(s))\, ds.$$

Da $f$ lokal Lipschitz–stetig ist, gibt es ein $L \geq 0$ und ein $\delta > 0$, so dass

$$\|f(s, \varphi(s)) - f(s, \psi(s))\| \leq L\|\varphi(s) - \psi(s)\|$$

für alle $s \in (a-\delta, a+\delta)$. Es folgt also aus Lemma 6.8, dass für alle $t \in (a-\delta, a+\delta)$ gilt:

$$\begin{aligned}
\|\varphi(t)-\psi(t)\| &= \left\|\int_a^t f(s,\varphi(s)) - f(s,\psi(s))\,ds\right\| \\
&\leq \int_a^t \|f(s,\varphi(s)) - f(s,\psi(s))\|\,ds \\
&\leq \left|\int_a^t L\|\varphi(s)-\psi(s)\|\,ds\right| \\
&= L\left|\int_a^t \|\varphi(s)-\psi(s)\|\,ds\right|.
\end{aligned}$$

Die Differenzen $\|\varphi(t)-\psi(t)\|$ tauchen nun also auf beiden Seiten der Ungleichung auf. Wir wollen wirklich zeigen, dass diese Differenzen gleich null sind.

Wir setzen

$$\varepsilon := \min\{\frac{\delta}{2}, \frac{1}{2L}\}.$$

Nachdem $\varphi$ und $\psi$ stetig sind existiert ein $t \in [a-\varepsilon, a+\varepsilon]$, so dass

$$\|\varphi(s)-\psi(s)\| \leq M := \|\varphi(t)-\psi(t)\| \quad \text{für alle } s \in [a-\varepsilon, a+\varepsilon].$$

Es folgt nun aus obiger Abschätzung[69], dass

$$\begin{aligned}
M &= \|\varphi(t)-\psi(t)\| \\
&\leq L\left|\int_a^t \|\varphi(s)-\psi(s)\|\,ds\right| \\
&\leq L\left|\int_a^t M\,ds\right| \\
&= L|t-a|M \leq L\varepsilon M \\
&\leq L\tfrac{1}{2L}M = \tfrac{1}{2}M.
\end{aligned}$$

Wir sehen also, dass $0 \leq M \leq \frac{1}{2}M$, aber dies ist nur möglich wenn $M=0$. Wir sehen also, dass $\varphi(t)=\psi(t)$ für alle $t \in [a-\varepsilon, a+\varepsilon]$.

Der Satz folgt nun aus folgender Behauptung, welche in Übungsblatt 12 bewiesen wird:

*Behauptung.* Es seien $f,g : \mathbb{R} \to \mathbb{R}^n$ zwei stetige Funktionen, so dass $f(a)=g(a)$ für ein $a \in \mathbb{R}$. Nehmen wir an, dass es für jedes $b \in \mathbb{R}$ mit $f(b)=g(b)$ ein $\varepsilon > 0$ gibt, so dass $f(t)=g(t)$ für alle $t \in (b-\varepsilon, b+\varepsilon)$. Dann gilt $f(t)=g(t)$ für alle $t \in \mathbb{R}$.

□

[69] Um die obige Abschätzung verwenden zu können, benötigen wir die Bedingung, dass $\varepsilon < \frac{\delta}{2}$.

### 11.2. **Der Existenzsatz.**

**Satz 11.4. (Existenzsatz von Picard–Lindelöf)** *Es sei* $G \subset \mathbb{R}\times\mathbb{R}^n$ *und* $f : G \to \mathbb{R}^n$ *eine Funktion, welche lokal Lipschitz–stetig bezüglich der Variablen* $y$ *ist. Dann gibt es zu jedem* $(a,c) \in G$ *ein* $\varepsilon > 0$ *und eine Lösung*

$$\varphi : [a-\varepsilon, a+\varepsilon] \to \mathbb{R}^n$$

*der Differentialgleichung*

$$y' = f(t,y),$$

*welche die Anfangsbedingung* $\varphi(a) = c$ *erfüllt.*

Man beachte, dass nach dem Eindeutigkeitssatz die Lösung $\varphi$ zudem eindeutig ist.

*Bemerkung.* Selbst wenn $G = \mathbb{R}\times\mathbb{R}^n$, d.h. wenn wir für alle Zeiten und alle Punkte ein Vektorfeld gegeben haben, und wenn $f$ Lipschitz–stetig ist, ist die Lösung nicht notwendigerweise auf ganz $\mathbb{R}$ definiert, wie wir schon im Beispiel nach Satz 10.1 gesehen haben.

Wir werden die Existenz der Lösung mit Hilfe des Banachschen Fixpunktsatzes zeigen. Zur Erinnerung:

(1) Ein Banachraum ist ein vollständiger normierter Vektorraum, d.h. ein normierter Vektorraum, in dem jede Cauchyfolge konvergiert.
(2) Für alle $a \le b \in \mathbb{R}$ ist
$$C([a,b],\mathbb{R}^n) := \left\{ \begin{array}{c} \text{Vektorraum aller stetigen} \\ \text{Abbildungen } f\colon [a,b] \to \mathbb{R}^n \end{array} \right\}$$
mit der Supremumsnorm, d.h. mit
$$\|f\| := \sup\{\|f(t)\| \,|\, t \in [a,b]\}$$
ein Banachraum.
(3) Es sei $V$ ein Banachraum. Eine Abbildung $T : V \to V$ heißt *Kontraktion*, wenn es eine Konstante $\Theta \in [0,1)$ (genannt Kontraktionsfaktor) gibt, so dass
$$\|T(x) - T(y)\| \le \Theta\|x-y\| \text{ für alle } x,y \in V.$$
Ein Punkt $x \in V$ heißt *Fixpunkt von* $T$, wenn $T(x) = x$.
(4) Der Banachsche Fixpunktsatz besagt folgendes: Es sei $A$ eine abgeschlossene Teilmenge eines Banachraumes $(V, \|-\|)$ und es sei $T : A \to A$ eine Kontraktion. Dann besitzt $T$ genau einen Fixpunkt.

*Beweis des Existenzsatz von Picard–Lindelöf.* Wir beweisen den Existenzsatz für den Spezialfall $G = \mathbb{R} \times \mathbb{R}^n$. Der allgemeine Fall wird ganz ähnlich bewiesen. Wir nehmen zudem o.B.d.A. an, dass $a = 0$ und $c = 0$.

Es sei nun erstmal $\varepsilon > 0$ beliebig. Nach Satz 11.3 ist eine differenzierbare Kurve $\varphi : [-\varepsilon, \varepsilon] \to \mathbb{R}^n$ eine Lösung der Differentialgleichung $y' = f(t, y)$ mit Anfangsbedingung $\varphi(a) = c$ genau dann, wenn

$$\varphi(t) = \int_0^t f(s, \varphi(s))\, ds \text{ für alle } s \in I. \tag{11.1}$$

Wir betrachten die Abbildung

$$\begin{array}{rcl} T : C([-\varepsilon, \varepsilon], \mathbb{R}^n) & \to & C([-\varepsilon, \varepsilon], \mathbb{R}^n) \\ \psi & \mapsto & \left( \begin{array}{rcl} T(\psi) : [-\varepsilon, \varepsilon] & \to & \mathbb{R}^n \\ t & \mapsto & \int_0^t f(s, \psi(s))\, ds. \end{array} \right) \end{array}$$

Dann folgt sofort aus der Definition von $T$, dass $\varphi$ die Integralgleichung (11.1) genau dann erfüllt, wenn $T(\varphi) = \varphi$, d.h. genau dann, wenn $\varphi$ ein Fixpunkt unter $T$ ist.

Um den Banachschen Fixpunktsatz anwenden zu können, müssen wir nun zeigen, dass es ein geeignetes $\varepsilon > 0$ und eine abgeschlossene Teilmenge $A \subset C([-\varepsilon, \varepsilon], \mathbb{R}^n)$ gibt, so dass die Einschränkung von $T$ auf $A$ eine Kontraktion ist. [70]

Nach Voraussetzung ist $f$ lokal Lipschitz–stetig, es existiert also ein $\delta > 0$ und ein $r > 0$, so dass die Einschränkung von $f$ auf

$$Q_{\delta,r} := \{(t, y) \in \mathbb{R} \times \mathbb{R}^n \,|\, |t| \le \delta, \|y\| \le r\}$$

Lipschitz–stetig bezüglich einer Lipschitzkonstanten $L > 0$ ist. Nachdem $f$ stetig und $Q_{\delta,r}$ kompakt ist, gibt es eine Konstante $M \ge 0$, so dass

$$\|f(t, y)\| \le M \text{ für alle } (t, y) \in Q_{\delta,r}.$$

Wir setzen jetzt

$$\begin{array}{rcl} \varepsilon & := & \min\left(\delta, \frac{r}{M}, \frac{1}{2L}\right), \\ A & := & \{\psi \in C([-\varepsilon, \varepsilon], \mathbb{R}^n) \,|\, \|\psi\| \le r\}. \end{array}$$

[70] Der Banachsche Fixpunktsatz, gibt uns dann also einen eindeutig bestimmten Fixpunkt von $T : A \to A$. Dieser Fixpunkt ist eine Lösung der Differentialgleichung. Es folgt aber nicht die Eindeutigkeit der Lösung, nachdem eine weitere Lösung $\psi : [-\varepsilon, \varepsilon] \to \mathbb{R}^n$ der Differentialgleichung nicht notwendigerweise in $A$ liegen muss. Man kann den Eindeutigkeitssatz deswegen nicht auf die Eindeutigkeit des Fixpunktes zurückführen.

Man beachte, dass $A \subset C([-\varepsilon,\varepsilon],\mathbb{R}^n)$ eine abgeschlossene Teilmenge ist. [71]

*Behauptung.* Die obige Abbildung

$$T : C([-\varepsilon,\varepsilon],\mathbb{R}^n) \to C([-\varepsilon,\varepsilon],\mathbb{R}^n)$$

hat folgende Eigenschaften:

(1) $T(A) \subset A$,
(2) die Einschränkung von $T$ auf $T : A \to A$ ist kontrahierend.

Wie wir oben schon diskutiert haben, folgt aus der Behauptung und dem Banachschen Fixpunktsatz sofort der Satz von Picard–Lindelöf.

Wir wenden uns also jetzt dem Beweis der Behauptung zu:

(1) Sei also $\psi \in A$. Wir müssen zeigen, das $T(\psi) \in A$, d.h. wir müssen zeigen, dass für alle $t \in [-\varepsilon,\varepsilon]$ gilt:

$$\|T(\psi)(t)\| = \left\| \int_0^t f(s,\psi(s)\,ds \right\| \le r.$$

Wir beachten zuerst, dass für alle $s \in [-\varepsilon,\varepsilon]$ gilt $|s| \le \varepsilon \le \delta$ und nachdem $\psi \in A$ gilt, dass $||\psi(s)|| \le r$, d.h. für alle $s \in [-\varepsilon,\varepsilon]$ gilt $(s,\psi(s)) \in A$. Wir daher erhalten für alle $t \in [-\varepsilon,\varepsilon]$, dass

$$\begin{aligned} \|T(\psi)(t)\| = \left\| \int_0^t f(s,\psi(s)\,ds \right\| &\le \int_0^t ||f(s,\psi(s)||\,ds \\ &\le \int_0^t M\,ds = |t|M \le \varepsilon M \le \tfrac{r}{M}M = r. \end{aligned}$$

Wir haben also gezeigt, dass $T(\psi)$ in $A$ liegt.

(2) Es seien $\varphi_1,\varphi_2 \in A$. Wir werden nun zeigen, dass

$$||T(\varphi_1) - T(\varphi_2)|| \le \frac{1}{2}||\varphi_1 - \varphi_2||.$$

In der Tat. Sei $t \in [-\varepsilon,\varepsilon]$. Dann gilt

$$T(\varphi_1)(t) - T(\varphi_2)(t) = \int_0^t \left(f(s,\varphi_1(s)) - f(s,\varphi_2(s))\right)\,ds.$$

[71] Dies ist eine ganz allgemeine Aussage: für jeden metrischen Raum $(X,d)$, jedes $x \in X$ und jedes $r \ge 0$, ist $A := \{y \in X \mid d(x,y) \le r\} \subset X$ abgeschlossen.

Nachdem $(s, \varphi_1(s)), (s, \varphi_2(s)) \in Q$ (siehe (1)), können wir die Lipschitzbedingung anwenden, es folgt für alle $t \in [-\varepsilon, \varepsilon]$, dass

$$\begin{array}{rcl}
\|T(\varphi_1)(t) - T(\varphi_2)(t)\| & = & \left\| \int_0^t \left(f(s,\varphi_1(s)) - f(s,\varphi_2(s))\right) ds \right\| \\
& \le & \int_0^t \|f(s,\varphi_1(s)) - f(s,\varphi_2(s))\| \, ds \\
& \le & \left| \int_0^t L\|\varphi_1(s) - \varphi_2(s)\| \, ds \right| \\
& \le & L|t| \cdot \|\varphi_1 - \varphi_2\| \\
& \le & L\varepsilon \cdot \|\varphi_1 - \varphi_2\| \\
& \le & L\frac{1}{2L} \cdot \|\varphi_1 - \varphi_2\| = \frac{1}{2}\|\varphi_1 - \varphi_2\|.
\end{array}$$

Nachdem dies für alle $t \in [-\varepsilon, \varepsilon]$ gilt, folgt dass

$$\begin{array}{rcl}
\|T(\varphi_1) - T(\varphi_2)\| & = & \sup\{\|T(\varphi_1)(t) - T(\varphi_2)(t)\| \mid t \in [-\varepsilon, \varepsilon]\} \\
& \le & \frac{1}{2}\|\varphi_1 - \varphi_2\|.
\end{array}$$

□

In dem Beweis vom Banachscher Fixpunktsatz hatten wir folgende, etwas stärkere Aussage bewiesen: Es sei $A$ eine abgeschlossene Teilmenge eines Banachraumes $(V, \| - \|)$ und es sei $T : A \to A$ eine Kontraktion. Es sei $v_0 \in A$. Dann konvergiert die durch $v_k := \Phi(v_{k-1})$ rekursiv definiert Folge gegen ein $v \in A$ mit $T(v) = v$.

Es sei $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ eine Funktion, welche lokal Lipschitz–stetig ist und es sei $(a, c) \in G$. Es sei $\varphi_0 : [-\varepsilon, \varepsilon] \to \mathbb{R}^n$ eine differenzierbare Abbildung mit $\varphi(a) = c$. Wir betrachten die Funktionenfolge

$$\begin{array}{rcl}
\varphi_0(t) & & \\
\varphi_1(t) & := & c + \int_0^t f(s, \varphi_0(s)) \, ds \\
\varphi_2(t) & := & c + \int_0^t f(s, \varphi_1(s)) \, ds \\
\vdots & &
\end{array}$$

Für geeignetes $\varepsilon > 0$ folgt aus der obigen Formulierung des Banachschen Fixpunktsatzes, dass diese Folge von Funktionen im Banachraum $C([a - \varepsilon, a + \varepsilon], \mathbb{R}^n)$ gegen eine Lösung der Differentialgleichung $y' = f(t, y)$ konvergiert. Wir verweisen auf Forster, Analysis II Seite 154 für ein explizites Beispiel.

## 12. Differentialgleichungen höherer Ordnung

*Definition.* Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und $f : G \to \mathbb{R}$ eine stetige Funktion. Wir nennen

$$y^{(n)} = f(t, y, y', \ldots, y^{(n-1)})$$

eine *Differentialgleichung $n$–ter Ordnung*, eine Lösung davon ist eine $n$–mal differenzierbare Funktion

$$\varphi : I \to \mathbb{R},$$

welche folgende Bedingung erfüllt:

$$\varphi'(t) = f(t, \varphi(t), \varphi'(t), \dots, \varphi^{(n-1)}(t)) \text{ für alle } t \in I.$$

(Damit die rechte Seite definiert ist, muss gelten, dass

$$(t, \varphi(t), \varphi'(t), \dots, \varphi^{(n-1)}(t)) \in G$$

für alle $t \in I$.)

*Beispiel.* Die Funktionen $\varphi(t) = -\sin(t) + \cos(t)$ und $\psi(t) = e^t$ sind Lösungen der Differentialgleichung 4. Ordnung

$$y^{(4)} = \frac{1}{2}(y + y').$$

Man beachte, dass $\varphi(0) = 1 = \psi(0)$. Der Eindeutigkeitssatz ist allerdings nicht verletzt, weil für Differentialgleichungen $n$–ter Ordnung die Werte für die ersten $(n-1)$–Ableitungen vorgegeben sein müssen, um eine eindeutige Lösung zu erhalten. Siehe Satz 12.3.

Wir werden jetzt sehen, dass man das Studium von Differentialgleichung höherer Ordnung auf das Studium von Differentialgleichungen 1. Ordnung zurückführen kann.

Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und $f : G \to \mathbb{R}$ eine stetige Funktion. Wir betrachten die Differentialgleichung $n$–ter Ordnung

$$y^{(n)} = f(t, y, y', \dots, y^{(n-1)}).$$

Wir können auch folgende Differentialgleichung erster Ordnung betrachten:

$$\begin{array}{rcl} y_0' & = & y_1 \\ y_1' & = & y_2 \\ & \vdots & \\ y_{n-2}' & = & y_{n-1} \\ y_{n-1}' & = & f(t, y_0, \dots, y_{n-1}). \end{array}$$

Wir setzen

$$Y := \begin{pmatrix} y_0 \\ \vdots \\ y_{n-2} \\ y_{n-1} \end{pmatrix} \text{ und } F(t, Y) := \begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \\ f(t, Y) \end{pmatrix}.$$

Dann können wir obige Differentialgleichung erster Ordnung auch schreiben als Differentialgleichung $Y' = F(t, Y)$.

Die folgenden beiden Lemmas besagen nun, dass Lösungen der Differentialgleichung $n$–ter Ordnung $y^{(n)} = f(t, y, y', \dots, y^{(n-1)})$ in ein–eindeutiger Beziehung zu den Lösungen der Differentialgleichung erster Ordnung $Y' = F(t, Y)$ in $n$–Variablen stehen.

**Lemma 12.1.** *Es sei $\varphi : I \to \mathbb{R}$ eine Lösung der Differentialgleichung $y^{(n)} = f(t, y, y', \dots, y^{(n-1)})$. Dann ist*

$$\begin{array}{rcl} \Phi : I & \to & \mathbb{R}^n \\ t & \mapsto & \Phi(t) := \begin{pmatrix} \varphi(t) \\ \varphi'(t) \\ \vdots \\ \varphi^{(n-1)}(t) \end{pmatrix} \end{array}$$

*eine Lösung der Differentialgleichung $Y' = F(t, Y)$.*

*Beweis.* Dieses Lemma folgt sofort durch einsetzen. In der Tat, es ist

$$\Phi'(t) = \begin{pmatrix} \varphi'(t) \\ \varphi''(t) \\ \vdots \\ \varphi^{(n)}(t) \end{pmatrix} = \begin{pmatrix} \varphi'(t) \\ \varphi''(t) \\ \vdots \\ f(t, \varphi(t), \dots, \varphi^{(n-1)}(t)) \end{pmatrix} = F(t, \Phi(t)).$$

□

**Lemma 12.2.** *Es sei $\Phi = (\varphi_1, \dots, \varphi_n) : I \to \mathbb{R}^n$ eine Lösung der Differentialgleichung $Y' = F(t, Y)$, dann ist $\varphi_1$ eine Lösung der Differentialgleichung $y^{(n)} = f(t, y, y', \dots, y^{(n-1)})$.*

*Beweis.* Es sei $\Phi = (\varphi_1, \dots, \varphi_n) : I \to \mathbb{R}^n$ eine Lösung der Differentialgleichung $Y' = F(t, Y)$. Ausgeschrieben bedeutet dies, dass

$$\begin{array}{rcl} \varphi_1'(t) & = & \varphi_2(t) \\ \varphi_2'(t) & = & \varphi_3(t) \\ & \vdots & \\ \varphi_n'(t) & = & f(t, \Phi(t)) = f(t, \varphi_1(t), \dots, \varphi_n(t)). \end{array}$$

Die ersten $n-1$ Gleichung implizieren, dass $\varphi_{k+1}(t) = \varphi_1^{(k)}(t)$, $k = 1, \dots, n-1$. Die letzte Gleichung besagt nun, dass

$$\varphi_1^{(n)}(t) = f(t, \varphi_1(t), \varphi_1'(t), \dots, \varphi_1^{(n-1)}(t)),$$

d.h. $\varphi_1$ ist eine Lösung der Differentialgleichung

$$y^{(n)} = f(t, y, y', \dots, y^{(n-1)}).$$

□

Nachdem Lösungen der Differentialgleichung

$$y^{(n)} = f(t, y, y', \dots, y^{(n-1)})$$

in ein–eindeutiger Beziehung zu den Lösungen der Differentialgleichung $Y' = F(t, Y)$ stehen, übertragen sich der Eindeutigkeitssatz und der Existenzsatz für Differentialgleichungen erster Ordnung zu Aussagen über Differentialgleichungen beliebiger Ordnung. Genauer gesprochen erhalten wir folgenden Satz:

**Satz 12.3.** *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und $f : G \to \mathbb{R}$ eine stetige Funktion, welche lokal Lipschitz–stetig* [72] *ist. Wir betrachten die Differentialgleichung*

$$y^{(n)} = f(t, y, y', \dots, y^{(n-1)}).$$

(1) *(Eindeutigkeit) Es seien $\varphi, \psi : I \to \mathbb{R}$ zwei Lösungen, so dass*
$$\varphi(a) = \psi(a),\ \varphi'(a) = \psi'(a),\ \dots,\ \varphi^{(n-1)}(a) = \psi^{(n-1)}(a),$$
*für ein $a \in I$. Dann gilt $\varphi(t) = \psi(t)$ für alle $t \in I$.*
(2) *(Existenz) Für jeden Punkt $(a, c_0, \dots, c_{n-1}) \in G$ gibt es ein $\varepsilon > 0$ und eine Lösung*
$$\varphi : [a - \varepsilon, a + \varepsilon] \to \mathbb{R}$$
*der Differentialgleichung, welche die Anfangsbedingung*
$$\varphi(a) = c_0,\ \varphi'(a) = c_1, \dots, \varphi^{(n-1)}(a) = c_{n-1}$$
*erfüllt.*

*Beweisskizze.* Es sei $f : G \to \mathbb{R}$ eine stetige Funktion, welche lokal Lipschitz–stetig ist. Dann kann man leicht sehen, dass die Funktion

$$\begin{array}{rcl} \mathbb{R} \times \mathbb{R}^n & \to & \mathbb{R}^n \\ (t, y_1, \dots, y_n) & \mapsto & F(t, Y) := \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \\ f(t, y_0, \dots, y_{n-1}) \end{pmatrix} \end{array}$$

ebenfalls lokal Lipschitz–stetig ist. Der Satz folgt nun sofort aus Lemmas 12.1 und 12.2, sowie aus dem Eindeutigkeitssatz 11.2 und dem Existenzsatz 11.4. □

[72] In diesem Fall soll das Folgendes heißen: für alle $(a, c) = (a, c_0, \dots, c_{n-1}) \in G$ existiert eine offene Umgebung $U \in R \times \mathbb{R}^n$ und ein $L \geq 0$, so dass für alle $(b, c')$ und $(b, c'')$ in $G \cap U$ gilt:

$$|f(b, c') - f(b, c'')| \leq L||c' - c''||.$$

## 13. Lineare Differentialgleichungen

*Definition.* Es sei $I \subset \mathbb{R}$ ein Intervall und

$$\begin{array}{rcl} I & \to & M(n,\mathbb{R}) \\ t & \mapsto & A(t) := (a_{ij}(t)) \end{array}$$

eine stetige Abbildung, dann nennen wir

$$y' = A(t)y$$

ein homogene lineare Differentialgleichung. [73] Es sei

$$\begin{array}{rcl} I & \to & \mathbb{R}^n \\ t & \mapsto & b(t) = (b_1(t), \dots, b_n(t)) \end{array}$$

eine stetige Abbildung, dann heißt

$$y' = A(t)y + b(t)$$

inhomogene lineare Differentialgleichung.

**Satz 13.1.** *Es sei $I \subset \mathbb{R}$ ein Intervall und es seien $A : I \to M(n,\mathbb{R})$ und $b : I \to \mathbb{R}^n$ stetige Abbildungen. Es sei $t_0 \in I$ und $c \in \mathbb{R}^n$. Dann besitzt die Differentialgleichung*

$$y' = A(t)y + b(t)$$

*genau eine Lösung $\varphi : I \to \mathbb{R}^n$, welche die Anfangsbedingung $\varphi(t_0) = c$ erfüllt.*

Für den Beweis, siehe Forster, Analysis II, Seite 162.

*Bemerkung.*
1. Wir betrachten die Differentialgleichung $y' = f(t,y)$ mit $f(t,y) = A(t)y + b(t)$. Man kann leicht zeigen, dass $f$ lokal Lipschitz–stetig ist. Aus Satz 11.2 folgt also die Eindeutigkeit der Lösung $\varphi$. Der Existenzsatz von Picard–Lindelöf (Satz 11.4) hingegen besagt nur, dass es eine Lösung gibt, welche auf einem Intervall $[t_0-\varepsilon, t_0+\varepsilon]$ definiert ist. Die Aussage, dass $\varphi$ auf ganz $I$ definiert ist, ist eine stärkere Aussage.
2. Die Funktion $b$ kann natürlich die Nullfunktion sein, d.h. der Satz gilt insbesondere für homogene lineare Differentialgleichungen.

[73] Diese Differentialgleichung wird oft auch als *homogenes lineares Differentialgleichungssystem* bezeichnet.

13.1. **Homogene lineare Differentialgleichungen.** Wir studieren nun zuerst die homogenen linearen Differentialgleichungen.

**Lemma 13.2.** *Es sei $I \subset \mathbb{R}$ ein Intervall und*

$$\begin{array}{rcl} I & \to & M(n,\mathbb{R}) \\ t & \mapsto & A(t) \end{array}$$

*eine stetige Abbildung. Die Menge aller Lösungen $\mathcal{L}$ der Differentialgleichung $y' = A(t)y$ ist ein Untervektorraum des reellen Vektorraums*

$$V := \{ \text{ alle differenzierbaren Abbildungen } I \to \mathbb{R}^n \}.$$

*Beweis.* Es seien $\varphi, \psi : I \to \mathbb{R}^n$ Lösungen der Differentialgleichung $y' = A(t)y$. Wir müssen zeigen, dass

$$\varphi + \psi : t \mapsto \varphi(t) + \psi(t)$$

ebenfalls eine Lösung ist. Dies kann man leicht verifizieren:

$$\begin{array}{rcl} (\varphi+\psi)(t)' & = & \varphi(t)' + \psi(t)' \\ & = & A(t)\varphi(t) + A(t)\psi(t) = A(t)(\varphi(t) + \psi(t)) = A(t)(\varphi+\psi)(t). \end{array}$$

Ganz analog zeigt man, dass wenn $\varphi$ eine Lösung ist, und $\lambda \in \mathbb{R}$, dann ist $t \mapsto \lambda\varphi(t)$ ebenfalls eine Lösung der Differentialgleichung. $\square$

Nachdem die Menge der Lösungen $\mathcal{L}$ der Differentialgleichung $y' = A(t)y$ ein Vektorraum ist, genügt es eine Basis zu finden, um $\mathcal{L}$ zu beschreiben. Zur Erinnerung: Funktionen $f_1, \dots, f_k : I \to \mathbb{R}^n$ sind linear unabhängig genau dann, wenn gilt:

$$\lambda_1 f_1 + \dots + \lambda_k f_k = 0 \ \text{ (d.h. die Nullabbildung) } \Rightarrow \lambda_1 = \dots = \lambda_k = 0.$$

anders ausgedrückt, $f_1, \dots, f_k : I \to \mathbb{R}^n$ sind linear unabhängig genau dann, wenn gilt:

$$\lambda_1 f_1(t) + \dots + \lambda_k f_k(t) = 0 \text{ für alle } t \in I \Rightarrow \lambda_1 = \dots = \lambda_k = 0.$$

Dies bedeutet nicht, dass für jedes einzelne $t$ die Vektoren $f_1(t), \dots, f_k(t)$ linear unabhängig sind.

**Satz 13.3.** *Es sei $I \subset \mathbb{R}$ ein Intervall und*

$$\begin{array}{rcl} I & \to & M(n,\mathbb{R}) \\ t & \mapsto & A(t) \end{array}$$

*eine stetige Abbildung. Es seien $\varphi_1, \dots, \varphi_k : I \to \mathbb{R}^n$ Lösungen von $y' = A(t)y$. Dann sind folgende Aussagen äquivalent:*

(1) *$\varphi_1, \dots, \varphi_k$ sind linear unabhängig in $\mathcal{L}$,*
(2) *es existiert ein $t_0 \in I$, so dass die Vektoren*

$$\varphi_1(t_0), \dots, \varphi_k(t_0) \in \mathbb{R}^n$$

*linear unabhängig sind,*

(3) *für jedes $t_0 \in I$, sind die Vektoren*

$$\varphi_1(t_0), \ldots, \varphi_k(t_0) \in \mathbb{R}^n$$

*linear unabhängig.*

*Beweis.* Es ist offensichtlich, dass (3) $\Rightarrow$ (2) und (2) $\Rightarrow$ (1). Wir müssen also nur noch zeigen, dass (1) $\Rightarrow$ (3). Seien also $\varphi_1, \ldots, \varphi_k$ linear unabhängige Lösungen von $y' = A(t)y$ und es sei $t_0 \in I$. Wir müssen zeigen, dass

$$\lambda_1\varphi_1(t_0) + \cdots + \lambda_k\varphi_k(t_0) = 0 \Rightarrow \lambda_1 = \cdots = \lambda_k = 0.$$

Es seien also $\lambda_1, \ldots, \lambda_k$ gegeben, so dass $\lambda_1\varphi_1(t_0) + \cdots + \lambda_k\varphi_k(t_0) = 0$. Dann ist $\varphi := \lambda_1\varphi_1 + \cdots + \lambda_k\varphi_k$ eine Lösung der Differentialgleichung $y' = A(t)y$ bezüglich der Anfangsbedingung $\varphi(t_0) = 0$. Andererseits ist $\psi(t) = 0$ ebenfalls eine Lösung der Differentialgleichung $y' = A(t)y$ bezüglich der Anfangsbedingung $\psi(t_0) = 0$. Aus Satz 13.1 folgt nun, dass $\varphi(t) = \psi(t) = 0$ für alle $t \in I$. Insbesondere ist $\lambda_1\varphi + \cdots + \lambda_k\varphi_k$ die Nullfunktion, aus der linearen Unabhängigkeit von den Funktionen $\varphi_1, \ldots, \varphi_k$ folgt nun, dass $\lambda_1 = \cdots = \lambda_k = 0$. $\square$

**Satz 13.4.** *Es sei $I \subset \mathbb{R}$ ein Intervall und*

$$\begin{array}{rcl} I & \to & M(n, \mathbb{R}) \\ t & \mapsto & A(t) \end{array}$$

*eine stetige Abbildung. Dann ist die Menge aller Lösungen $\mathcal{L}$ der Differentialgleichung $y' = A(t)y$ ein $n$–dimensionaler Vektorraum.*

*Beweis.* Es sei $t_0 \in I$. Aus Satz 13.1 folgt, dass es für jedes $i \in \{1, \ldots, n\}$ eine Lösung $\varphi_i$ der Differentialgleichung $y' = A(t)y$ gibt, welche die Anfangsbedingung $\varphi_i(t_0) = e_i$ erfüllt.[74] Wir werden nun zeigen, dass $\varphi_1, \ldots, \varphi_n$ eine Basis von $\mathcal{L}$ ist.

Aus Satz 13.3 folgt, dass die Funktionen $\varphi_1, \ldots, \varphi_n$ linear unabhängig sind. Sei nun $\psi$ eine beliebige Lösung der Differentialgleichung $y' = A(t)y$. Wir schreiben $\psi(t_0) =: (\lambda_1, \ldots, \lambda_n)$. Dann gilt

$$\psi(t_0) = (\lambda_1, \ldots, \lambda_n) = \lambda_1\varphi(t_0) + \cdots + \lambda_n\varphi(t_0).$$

Die Funktionen $\psi$ und $\lambda_1\varphi + \cdots + \lambda_n\varphi$ sind also Lösungen der Differentialgleichung und haben den gleichen Wert zur Zeit $t_0$, aus Satz 13.1 folgt nun, dass $\psi = \lambda_1\varphi + \cdots + \lambda_n\varphi$. $\square$

*Definition.* Es sei $y' = A(t)y$ eine lineare homogene Differentialgleichung, eine Basis des Lösungsraumes der Differentialgleichung wird *Lösungsfundamentalsystem* genannt.

[74] Wie üblich bezeichnet $e_i$ den $i$–ten Vektor der Standardbasis des $\mathbb{R}^n$, d.h. $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ mit der 1 in der $i$–ten Koordinate.

*Beispiel.* Es sei $c \in \mathbb{R}$. Wir betrachten die lineare homogene Differentialgleichung

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \underbrace{\begin{pmatrix} 0 & -c \\ c & 0 \end{pmatrix}}_{=A} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

(d.h. $y' = Ay$ wobei $A$ eine konstante $2 \times 2$–Matrix ist). Dann sind

$$\varphi_1(t) = \begin{pmatrix} \cos(ct) \\ \sin(ct) \end{pmatrix} \text{ und } \varphi_2(t) = \begin{pmatrix} -\sin(ct) \\ \cos(ct) \end{pmatrix}$$

Lösungen der Differentialgleichung (wie man leicht durch einsetzen sieht). Nachdem die Vektoren

$$\varphi_1(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ und } \varphi_2(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

linear unabhängig sind, folgt aus Satz 13.3 und Satz 13.4, dass $\varphi_1$ und $\varphi_2$ ein Lösungsfundamentalsystem der Differentialgleichung $y' = Ay$ sind. Insbesondere kann jede andere Lösung geschrieben werden als

$$\psi(t) = \lambda_1 \varphi_1(t) + \lambda_2 \varphi_2(t).$$

### 13.2. **Die Wronski–Matrix.**

*Definition.* Es sei $I \subset \mathbb{R}$ ein Intervall und

$$\begin{array}{ccc} I & \to & M(n, \mathbb{R}) \\ t & \mapsto & A(t) \end{array}$$

eine stetige Abbildung. Es seien $\varphi_1, \ldots, \varphi_n : I \to \mathbb{R}^n$ Lösungen der Differentialgleichung $y' = A(t)y$. Dann heißt

$$\Phi(t) := \begin{pmatrix} \varphi_1(t) & \ldots & \varphi_n(t) \end{pmatrix}$$

die *Wronski–Matrix* und

$$W(t) := \det(\Phi(t))$$

die *Wronski–Determinante* der Lösungen $\varphi_1, \ldots, \varphi_n$.

Wir können nun Satz 13.3 mithilfe von Satz 13.4 wie folgt umformulieren: Es seien

$$\varphi_1, \ldots, \varphi_n : I \to \mathbb{R}^n$$

Lösungen der Differentialgleichung $y' = A(t)y$, dann sind die folgenden Aussagen äquivalent:

(1) $\varphi_1, \ldots, \varphi_n$ bilden ein Lösungsfundamentalsystem,
(2) es existiert ein $t_0 \in I$, so dass $W(t_0) \neq 0$,
(3) $W(t) \neq 0$ für alle $t \in I$.

Die Wronski–Determinante hat folgende geometrische Bedeutung: Es seien $\varphi_1, \ldots, \varphi_n : I \to \mathbb{R}^n$ Lösungen der Differentialgleichung $y' = A(t)y$. Dann ist

$$|W(t)| = |\det\begin{pmatrix}\varphi_1(t) & \ldots & \varphi_n(t)\end{pmatrix}|$$

das Volumen des durch $\varphi_1(t), \ldots, \varphi_n(t)$ aufgespannten Parallelotops. [75] In vielen Fällen studiert man volumenerhaltende Differentialgleichungen, d.h. Differentialgleichung, so dass $|W(t)|$ konstant ist für ein Lösungsfundamentalsystem. [76] Beispielsweise ist jede Differentialgleichung, welche den Fluss von Wasser beschreibt notwendigerweise volumenerhaltend.

**Satz 13.5.** *(Satz von Liouville) Es sei $I \subset \mathbb{R}$ ein Intervall und*

$$\begin{array}{rcl} I & \to & M(n, \mathbb{R}) \\ t & \mapsto & A(t) \end{array}$$

*eine stetige Abbildung. Es seien $\varphi_1, \ldots, \varphi_n : I \to \mathbb{R}^n$ Lösungen der Differentialgleichung $y' = A(t)y$. Die dazugehörige Wronski–Determinante erfüllt die Differentialgleichung*

$$W'(t) = \operatorname{spur}(A(t)) \cdot W(t).$$

*Bemerkung.* Sei $t_0 \in I$. Aus Satz 10.2 folgt, dass

$$W(t) = W(t_0) \cdot e^{\int_{t_0}^{t} \operatorname{spur}(A(s))\, ds}.$$

Insbesondere folgt, dass $W(t_0) \neq 0$, genau dann, wenn $W(t) \neq 0$ für alle $t$. Der Satz von Liouville ergibt damit insbesondere einen neuen Beweis der Äquivalenz der Aussagen (2) und (3) in Satz 13.3.

*Beweisskizze.* Wir beweisen zuerst folgende Behauptung:

---

[75] Zur Erinnerung: es seien $v_1, \ldots, v_n \in \mathbb{R}^n$ gegeben, dann heißt

$$P := \{\lambda_1 v_1 + \cdots + \lambda_n v_n \,|\, \lambda_1, \ldots, \lambda_n \in [0,1]\}$$

das durch $v_1, \ldots, v_n$ aufgespannte *Parallelotop* (oder auch *Parallelepiped* beziehungsweise *Spat*). Es gilt ganz allgemein, dass

$$|\det\begin{pmatrix}v_1 & \ldots & v_n\end{pmatrix}| = \text{Volumen von } P.$$

Dieser Zusammenhang wird in Analysis III erläutert, in Analysis III wird auch das Volumen von geeigneten Teilmengen von $\mathbb{R}^n$ definiert.

[76] Man beachte, dass die Wronski–Determinante konstant ist für ein Lösungsfundamentalsystem, genau dann, wenn die Wronski–Determinante konstant ist für jedes Lösungsfundamentalsystem. Dies folgt daraus, dass es für zwei Wronski–Matrizen $\Phi$ und $\Psi$ eine Matrix $P \in GL(n, \mathbb{R})$ gibt mit $\Phi(t) = P\Psi(t)$ für alle $t$.
In der Tat: sei $t_0 \in I$, wir setzen $P := \Phi(t_0) \cdot \Psi(t_0)^{-1}$. Dann folgt aus dem Eindeutigkeitssatz, dass $\Phi(t) = P\Psi(t)$ für alle $t$.

*Behauptung.* Es sei $B : I \to M(n, \mathbb{R})$ eine differenzierbare Abbildung[77]. Wir bezeichnen mit $b_i(t)$ den $i$–ten Zeilenvektor von $B(t)$. Dann gilt

$$\frac{d}{dt}\det(B(t)) = \sum_{i=1}^{n} \det \begin{pmatrix} b_1(t) \\ \vdots \\ b_i(t)' \\ \vdots \\ b_n(t) \end{pmatrix}$$

Wir wenden uns dem Beweis der Behauptung zu. Wir schreiben $B(t+h) = B(t) + hB'(t) + R(h)$ wobei $R(h) = o(h)$ für $h \to 0$. [78] Wir bezeichnen zudem mit $r_i$ die $i$–Zeile von $R$. Dann gilt

$$\begin{aligned}
&\tfrac{d}{dt}\det(B(t)) \\
= \ &\lim_{h\to 0}\tfrac{1}{h}(\det(B(t) + hB'(t) + R(h)) - \det(B(t))) \\
= \ &\lim_{h\to 0}\tfrac{1}{h}\left(\det\begin{pmatrix} b_1 + hb_1' + r_1 \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix} - \det(B(t))\right) \\
= \ &\lim_{h\to 0}\tfrac{1}{h}\left(\det\begin{pmatrix} b_1 \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix} + h\det\begin{pmatrix} b_1' \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix} + r_1\det\begin{pmatrix} 1 \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix}\right. \\
&\left. - \det(B(t))\right) \\
= \ &\lim_{h\to 0}\tfrac{1}{h}\left(\det\begin{pmatrix} b_1 \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix} - \det(B(t))\right) \\
&+\lim_{h\to 0}\det\begin{pmatrix} b_1' \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix} + \lim_{h\to 0}\tfrac{r_1(h)}{h}\det\begin{pmatrix} 1 \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix}.
\end{aligned}$$

(Die dritte Gleichheit folgt in dem wir die Linearität der Determinante auf die erste Zeile anwenden.) Die beiden letzteren Grenzwerte kann

[77] D.h. alle $n \times n$–Einträge sind differenzierbar

[78] Zur Erinnerung, $R(h) = o(h)$ für $h \to 0$ bedeutet, dass für alle $i, j \in \{1, \ldots, n\}$ gilt:

$$\lim_{h\to 0}\frac{r_{ij}(h)}{h} = 0.$$

man leicht bestimmen, in der Tat durch einsetzen sieht man, dass

$$\lim_{h\to 0} \det \begin{pmatrix} b_1' \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix} = \det \begin{pmatrix} b_1' \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

und aus $r_1(h) = o(h)$, folgt, dass

$$\lim_{h\to 0} \frac{r_1(h)}{h} \det \begin{pmatrix} 1 \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix} = 0.$$

Wir erhalten also, dass

$$\frac{d}{dt} \det(B(t)) = \lim_{h\to 0} \frac{1}{h} \left( \det \begin{pmatrix} b_1 \\ b_2 + hb_2' + r_2 \\ \vdots \\ b_n + hb_n' + r_n \end{pmatrix} - \det(B(t)) \right) + \det \begin{pmatrix} b_1' \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Wir wenden nun das obige Verfahren für Zeilen $2, 3, \ldots, n$ an, dann erhalten wir, dass

$$\begin{aligned} \frac{d}{dt} \det(B(t)) &= \lim_{h\to 0} \frac{1}{h} \left( \det \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} - \det(B) \right) + \sum_{i=1}^n \det \begin{pmatrix} b_1 \\ \vdots \\ b_i' \\ \vdots \\ b_n \end{pmatrix} \\ &= \sum_{i=1}^n \det \begin{pmatrix} b_1 \\ \vdots \\ b_i' \\ \vdots \\ b_n \end{pmatrix}. \end{aligned}$$

Wir haben damit die Behauptung bewiesen.

Wir bezeichnen nun mit $B$ die Wronski–Matrix der Lösungen $\varphi_1, \ldots, \varphi_n$. Nachdem $\varphi_i' = A\varphi_i$ folgt sofort, dass $B' = AB$, d.h.

$$b_i' = \sum_{j=1}^n a_{ij} \cdot b_j.$$

Setzen wir dies in die obige Gleichung ein, dann erhalten wir, dass

$$\begin{aligned}\frac{d}{dt}\det(B(t)) &= \sum_{i=1}^n \det\begin{pmatrix} b_1 \\ \vdots \\ \sum_{j=1}^n a_{ij}\cdot b_j \\ \vdots \\ b_n\end{pmatrix} \leftarrow i\text{–te Zeile} \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} \det\begin{pmatrix} b_1 \\ \vdots \\ b_j \\ \vdots \\ b_n\end{pmatrix} \leftarrow i\text{–te Zeile.}\end{aligned}$$

Die Matrixen haben jeweils zwei gleiche Zeilen, außer wenn $i = j$. Wir erhalten also die gewünschte Aussage, dass

$$\frac{d}{dt}\det(B(t)) = \sum_{i=1}^n a_{ii} \det\begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n\end{pmatrix} \leftarrow i\text{–te Zeile} = \operatorname{spur}(A)\cdot\det(B).$$

□

## 13.3. Inhomogene lineare Differentialgleichungen.

**Satz 13.6.** *Es sei $I \subset \mathbb{R}$ ein Intervall und es seien $A : I \to M(n, \mathbb{R})$ und $b : I \to \mathbb{R}^n$ stetige Abbildungen. Wir bezeichnen mit $\mathcal{L}$ den Vektorraum der Lösungen der linearen homogenen Differentialgleichung $y' = A(t)y$. Es sei nun $\psi$ eine Lösung der linearen inhomogenen Differentialgleichung $y' = A(t)y + b(t)$. Dann gilt:*

$$\phi \text{ ist Lösung von } y' = A(t)y + b(t) \Leftrightarrow \phi = \psi + \varphi \text{ für ein } \varphi \in \mathcal{L}.$$

Der Satz besagt also, dass es genügt *alle* Lösungen der linearen homogenen Differentialgleichung $y' = A(t)y$ und *eine* Lösung der linearen inhomogenen Differentialgleichung $y' = A(t)y + b(t)$ zu finden, um alle Lösungen der linearen inhomogenen Differentialgleichung $y' = A(t)y + b(t)$ zu beschreiben.

*Beweis.* Sei zuerst $\phi = \psi + \varphi$ für ein $\varphi \in \mathcal{L}$. Dann gilt

$$\phi' = \psi' + \varphi' = A(t)\psi + b(t) + A(t)\varphi = A(t)(\psi + \varphi) + b(t) = A(t)\phi + b(t),$$

d.h. $\phi$ ist in der Tat ein Lösung der Differentialgleichung $y' = A(t)y + b(t)$.

Sei nun $\phi$ eine Lösung von $y' = A(t)y + b(t)$. Wir setzen $\varphi = \phi - \psi$. Dann gilt

$$\varphi' = \phi' - \psi' = (A(t)\phi + b(t)) - (A(t)\psi + b(t)) = A(t)(\phi - \psi) = A(t)\varphi,$$

d.h. $\varphi$ ist Lösung der linearen homogenen Differentialgleichung $y' = A(t)y$, d.h. $\varphi \in \mathcal{L}$.

□

Der folgende Satz ist nun eine Verallgemeinerung von Satz 10.3.

**Satz 13.7.** *Es sei $I \subset \mathbb{R}$ ein Intervall und es seien $A : I \to M(n, \mathbb{R})$ und $b : I \to \mathbb{R}^n$ stetige Abbildungen. Es sei $\varphi_1, \ldots, \varphi_n$ ein Lösungsfundamentalsystem der linearen homogenen Differentialgleichung $y' = A(t)y$. Für jedes $t \in I$ bezeichne*

$$\Phi(t) := \begin{pmatrix} \varphi_1(t) & \ldots & \varphi_n(t) \end{pmatrix} \in M(n, \mathbb{R})$$

*die Wronski–Matrix der Lösungen $\varphi_1(t), \ldots, \varphi_n(t)$. Dann ist eine Lösung der linearen inhomogenen Differentialgleichung $y' = A(t)y + b(t)$ gegeben durch*[79]

$$\psi(t) = \Phi(t)u(t) \textit{ mit } u(t) = \int_{t_0}^{t} \Phi(s)^{-1} b(s)\, dt + \textit{const.}^{80}$$

*Beweis.* Wir zeigen, dass $\psi$ eine Lösung der linearen inhomogenen Differentialgleichung $y' = A(t)y + b(t)$ durch einsetzen: Es gilt

$$\begin{aligned} \psi' &= (\Phi u)' \\ &= \Phi' u + \Phi u' \\ &= \begin{pmatrix} \varphi' & \ldots & \varphi_n' \end{pmatrix} u + \Phi \cdot \Phi^{-1} b(t) \\ &= \begin{pmatrix} A(t)\varphi_1 & \ldots & A(t)\varphi_n \end{pmatrix} u + b(t) \\ &= A(t) \begin{pmatrix} \varphi_1 & \ldots & \varphi_n \end{pmatrix} u + b(t) \\ &= A(t)\Phi u + b(t) = A(t)\psi + b(t). \end{aligned}$$

Die zweite Gleichheit ist die Produktregel für Ableitungen komponentenweise angewandt. Die dritte Gleichheit folgt aus dem Hauptsatz der Differential– und Integralrechnung. Die vierte Gleichheit folgt daraus, dass $\varphi_1, \ldots, \varphi_n$ Lösungen der Differentialgleichung $y' = A(t)y$ sind.

□

[79] Es folgt aus Satz 13.3, dass für jedes $t \in I$ die Vektoren $\varphi_1(t), \ldots, \varphi_n(t)$ linear unabhängig sind, d.h. für jedes $t$ ist die Matrix $\Phi(t)$ invertierbar.

[80] Der Integrant $\Phi(t)^{-1}b(t)$ ist eine Funktion $I \to \mathbb{R}^n$, dass Integral einer solchen Funktion wird komponentenweise bestimmt.

# 14. Homogene lineare autonome Differentialgleichungen

## 14.1. Der Fluss eines autonomen Differentialgleichungssystems. [81]

*Definition.* Es sei $M \subset \mathbb{R}^n$ eine Teilmenge. Ein *Fluss* ist eine stetige Abbildung

$$\Phi : \mathbb{R} \times M \to M,$$

welche die folgenden beiden Eigenschaften erfüllt:

(1) $\Phi(0, x) = x$ für alle $x \in M$,
(2) für alle $s, t \in \mathbb{R}$ und $x \in M$ gilt

$$\Phi(s, \Phi(t, x)) = \Phi(s + t, x).$$

*Bemerkung.* Ein Fluss wird manchmal auch als *dynamisches System* bezeichnet.

*Beispiel.* Die Abbildung

$$\begin{array}{rcl} \Phi : \mathbb{R} \times \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (t, \begin{pmatrix} x \\ y \end{pmatrix}) & \mapsto & \underbrace{\begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}}_{\substack{\text{Rotation um} \\ \text{Winkel t}}} \begin{pmatrix} x \\ y \end{pmatrix} \end{array}$$

definiert einen Fluss auf $\mathbb{R}^2$. (Siehe Übungsblatt 13).

*Definition.* (1) Es sei $\Phi : \mathbb{R} \times M \to M$ ein Fluss und $x \in M$. Dann heißt die Kurve

$$\begin{array}{rcl} \varphi_x : \mathbb{R} & \to & M \\ t & \mapsto & \Phi(t, x) \end{array}$$

die *Flusslinie zum Anfangspunkt $x$ bei* $t = 0$.
(2) Ein Punkt $x$ heißt *Fixpunkt* oder auch *singulärer Punkt*, wenn $\Phi(x, t) = x$ für alle $t \in \mathbb{R}$.
(3) Eine Flusslinie $\varphi_x$ heißt *periodisch*, wenn diese nicht konstant ist, und wenn es ein $p > 0$ gibt, so dass

$$\varphi_x(t + p) = \varphi_x(t)$$

[81] Der Stoff dieses Kapitels basiert nicht auf Forster: Analysis II, sondern basiert lose auf Bröcker: Analysis III, Kapitel I.

für alle $t \in \mathbb{R}$. [82] Das kleinste $p > 0$ mit der Eigenschaft, dass $\varphi_x(t+p) = \varphi_x(t)$ für alle $t \in \mathbb{R}$, nennt man die *Periode* der Flusslinie. [83]

Im obigen Beispiel ist $(0,0)$ der einzige Fixpunkt des Flusses. Für jedes $(x,y) \neq (0,0)$ ist die dazugehörige Flusslinie periodisch mit Periode $2\pi$.

*Definition.* Es sei $M \subset \mathbb{R}^n$ eine Teilmenge und $f : M \to \mathbb{R}^n$ eine stetige Abbildung. Eine Differentialgleichung vom Typ

$$y' = f(y)$$

heißt *autonome Differentialgleichung* (wenn $n > 1$ nennt man es auch autonomes Differentialgleichungssystem). [84]

**Satz 14.1.** *Es sei $M \subset \mathbb{R}^n$ eine Teilmenge und $f : M \to \mathbb{R}^n$ eine Lipschitz stetige Abbildung. Wir nehmen an, dass für jedes $x \in M$ die Lösung $\varphi_x$ der Differentialgleichung $y' = f(y)$, welche die Anfangsbedingung $\varphi_x(0) = x$ erfüllt, auf ganz $\mathbb{R}$ definiert ist. Dann ist*

$$\begin{array}{rcl} \Phi : \mathbb{R} \times M & \to & M \\ (t,x) & \mapsto & \varphi_x(t) \end{array}$$

*ein Fluss auf $M$.*

*Beispiel.* Wir betrachten die Differentialgleichung

$$\begin{array}{rcl} y_1' & = & -y_2 \\ y_2' & = & y_1 \end{array} \quad \text{d.h} \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Dann ist Dann ist der dazugehörige Fluss gerade der Fluss im obigen Beispiel.

*Beweisskizze.* Für jedes $x \in M$ ist die Abbildung

$$\begin{array}{rcl} \mathbb{R} & \to & M \\ t & \mapsto & \varphi_x(t) \end{array}$$

[82] Wenn $\varphi_x(t+p) = \varphi_x(t)$ für *ein* $t \in \mathbb{R}$, dann folgt aus der zweiten Eigenschaft eines Flusses, dass $\varphi_x(t+p) = \varphi_x(t)$ für *alle* $t \in \mathbb{R}$.

[83] Es sei $\varphi_x$ eine periodische Flusslinie. Wir betrachten dann

$$Q := \{q > 0 \,|\, \varphi_x(t+q) = \varphi_x(t) \text{ für alle } t \in \mathbb{R}\}.$$

Dies ist eine nichtleere Menge. Zudem kann man zeigen, dass $\inf(Q) > 0$, und dass $p := \inf(Q)$ ebenfalls die Eigenschaft hat, dass $\varphi_x(t+p) = \varphi_x(t)$ für alle $t \in \mathbb{R}$. Dies zeigt, dass es wirklich Sinn macht, von dem 'kleinsten $p > 0$ mit der Eigenschaft, dass $\varphi_x(t+p) = \varphi_x(t)$ für alle $t \in \mathbb{R}$' zu sprechen.

[84] Autonom heißt also, dass das Vektorfeld nicht von der Zeit $t$ abhängt.

differenzierbar und insbesondere stetig. Dies impliziert allerdings nicht, dass die Abbildung

$$\begin{array}{rcl}\Phi : \mathbb{R} \times M & \to & M \\ (t, x) & \mapsto & \varphi_x(t)\end{array}$$

stetig ist. Der Beweis der Stetigkeit von $\Phi$ ist nicht trivial. Einen Beweis kann man z.B. in Bröcker: Analysis III, Satz 2.3 finden. [85]

Wir müssen nun noch zeigen, dass

(1) $\Phi(0, x) = x$ für alle $x \in M$,
(2) für alle $s, t \in \mathbb{R}$ und $x \in M$ gilt

$$\Phi(s, \Phi(t, x)) = \Phi(s + t, x).$$

Die erste Aussage folgt sofort daraus, dass $\varphi_x$ die Anfangsbedingung $\varphi_x(0) = x$ erfüllt. Sei nun $x \in M$ und $t \in \mathbb{R}$. Wir betrachten die Abbildungen

$$\begin{array}{rcl}\mathbb{R} & \to & M \\ s & \mapsto & \Phi(s, \Phi(t, x)) \text{ und} \\ s & \mapsto & \Phi(s + t, x).\end{array}$$

Beide Kurven nehmen denn Wert $\Phi(t, x)$ zum Zeitpunkt $s = 0$ an. Nachdem $\varphi_x$ eine Lösung der Differentialgleichung $y' = f(y)$ ist, folgt

$$\frac{d}{ds}\Phi(s + t, x) = \frac{d}{ds}\varphi_x(s + t) = f(\varphi_x(s + t)) = f(\Phi(s + t, x))$$

[86] und nachdem $\varphi_{\Phi(t,x)}$ eine Lösung der Differentialgleichung $y' = f(y)$ ist, folgt

$$\frac{d}{ds}\Phi(s, \Phi(t, x)) = \frac{d}{ds}\varphi_{\Phi(t,x)}(s) = f(\varphi_{\Phi(t,x)}(s)) = f(\Phi(s, \Phi(t, x))).$$

D.h. beide Kurven erfüllen die Differentialgleichung $y' = f(y)$ bezüglich der gleichen Anfangsbedingung. Es folgt nun aus dem Eindeutigkeitssatz, dass

$$\Phi(s, \Phi(t, x)) = \Phi(s + t, x) \text{ für alle } s \in \mathbb{R}.$$

$\square$

[85] `http://www.mathematik.uni-regensburg.de/broecker/index.htm`

[86] Wenn man genau hinschaut, sieht man, dass man für die zweite Gleichheit die Kettenregel auf $s \mapsto \alpha_x(s)$ und $s \mapsto s + t$ anwenden muss.

14.2. **Homogene lineare autonome Differentialgleichungen I.** Zur Erinnerung: eine Differentialgleichung $y' = f(t, y)$ heißt linear, wenn $f(t, y) = A(t)y + b(t)$, sie heißt homogen linear, wenn $b(t) = 0$. Zudem heißt $y' = f(t, y)$ autonom, wenn $f$ nicht von $t$ abhängt. Eine homogene lineare autonome Differentialgleichungen ist also von der form

$$y' = Ay$$

für eine festgewählte $n \times n$–Matrix $A$.

Wir erinnern uns zuerst an die Lösung im Fall $n = 1$: Es sei $a \in \mathbb{R}$ und $c \in \mathbb{R}$. Die Lösung $\varphi$ der Differentialgleichung $y' = ay$ bezüglich der Anfangsbedingung $\varphi(0) = c$ ist gegeben durch $\varphi(t) = e^{at}c$, wobei

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Wir kehren zurück zum allgemeinen Fall, es sei also $A$ eine beliebige $n \times n$–Matrix. Wir betrachten die homogene lineare autonome Differentialgleichung

$$y' = Ay.$$

Unser Ziel ist alle Lösungen zu beschreiben, oder anders ausgedrückt, den dazugehörigen Fluss zu beschreiben.

**Lemma 14.2.** *Es sei $A$ eine beliebige reelle $n \times n$–Matrix. Dann konvergiert*

$$e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

*im Vektorraum der reellen $n \times n$–Matrizen bezüglich der Norm*

$$\|A\| := \max\{|a_{ij}| \mid i, j = 1, \ldots, n\}.$$

Hier verwenden wir die Konvention, dass

$$\text{Nullmatrix}^0 = id_n,$$

es folgt, dass

$$e^{\text{Nullmatrix}} = \mathrm{id}_n .$$

*Beispiel.* Es sei $D$ eine Diagonalmatrix mit den Einträgen $\lambda_1,\dots,\lambda_n$, dann gilt

$$\begin{aligned} e^D = \sum_{k=0}^{\infty}\tfrac{1}{k!}D^k &= \sum_{k=0}^{\infty}\tfrac{1}{k!}\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \lambda_n\end{pmatrix}^k \\ &= \sum_{k=0}^{\infty}\tfrac{1}{k!}\begin{pmatrix}\lambda_1^k & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \lambda_n^k\end{pmatrix} \\ &= \begin{pmatrix}\sum_{k=0}^{\infty}\tfrac{1}{k!}\lambda_1^k & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \sum_{k=0}^{\infty}\tfrac{1}{k!}\lambda_n^k\end{pmatrix} = \begin{pmatrix}e^{\lambda_1} & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & e^{\lambda_n}\end{pmatrix}. \end{aligned}$$

*Beweis von Lemma 14.2.* Wir beweisen zuerst folgende Behauptung:

*Behauptung.* Es seien $A,B\in M(n,\mathbb{R})$. Dann gilt

$$\|A\cdot B\|\le n\|A\|\cdot\|B\|.$$

Es seien also $A,B\in M(n,\mathbb{R})$. Wir schreiben $C=A\cdot B$. Dann gilt

$$\begin{aligned} |c_{ij}| &= |\textstyle\sum_{k=1}^n a_{ik}b_{kj}| \\ &\le \textstyle\sum_{k=1}^n |a_{ik}b_{kj}| \\ &\le \textstyle\sum_{k=1}^n \max\{|a_{ij}|\,|\,i,j=1,\dots,n\}\cdot\max\{|b_{ij}|\,|\,i,j=1,\dots,n\} \\ &= n\|A\|\cdot\|B\|. \end{aligned}$$

Es folgt also, dass

$$||C||=\max\{|c_{ij}|,\,|\,i,j=1,\dots,n\}\le n||A||\cdot||B||.$$

Wir haben also die Behauptung bewiesen.

Es folgt aus der obigen Behauptung durch Induktion, dass

$$\left\|\frac{A^k}{k!}\right\|\le\frac{1}{k!}n^{k-1}\|A\|^k=\frac{1}{k!}\frac{1}{n}(n\|A\|)^k.$$

Wir wählen ein $l\in\mathbb{N}$ mit $l\ge 2n\|A\|$. Dann gilt für $k\ge l$, dass

$$\begin{aligned} \tfrac{1}{k!}\tfrac{1}{n}(n\|A\|)^k &= \tfrac{1}{n}\tfrac{1}{l!}(n\|A\|)^l\prod_{i=1}^{k-l}\underbrace{\frac{n\|A\|}{l+i}}_{<\frac{1}{2}} \\ &\le \tfrac{1}{n}\tfrac{1}{l!}(n\|A\|^l)\tfrac{1}{2^{k-l}} \\ &= \underbrace{\frac{1}{n}\frac{1}{l!}(n\|A\|)^l2^l}_{\text{Konstante}}\tfrac{1}{2^k}. \end{aligned}$$

Wir setzen nun

$$C:=\frac{1}{n}\frac{1}{l!}(n\|A\|)^l2^l.$$

Es sei $K \geq l$ und $k_2 \geq k_1 \geq K$, dann gilt

$$\begin{aligned}
\left\| \sum_{k=k_1}^{k_2} \frac{A^k}{k!} \right\| &\leq \sum_{k=k_1}^{k_2} \left\| \frac{A^k}{k!} \right\| \\
&\leq C \sum_{k=k_1}^{k_2} \frac{1}{2^k} \\
&< C \sum_{k=k_1}^{\infty} \frac{1}{2^k} \\
&= C \frac{1}{2^{k-1}} \leq C \frac{1}{2^{K-1}}.
\end{aligned}$$

Wir sehen also, dass die Partialsummen der Reihe eine Cauchyfolge bilden. Der endlich dimensionale Vektorraum $M(n, \mathbb{R})$ ist ein Banachraum, jede Cauchyfolge konvergiert also, insbesondere konvergiert die Reihe $\sum_{k=0}^{\infty} \frac{A^k}{k!}$. □

Das folgende Lemma wird ähnlich wie Analysis I: Satz 5.12 bewiesen.

**Lemma 14.3.** *Für alle $A, B \in M(n, \mathbb{R})$ gilt:*

$$e^{A+B} = e^A \cdot e^B.$$

*Insbesondere ist*

$$(e^A)^{-1} = e^{-A},$$

*d.h. die Matrix $e^A$ ist invertierbar.*

**Satz 14.4.** *Es sei $A$ eine beliebige $n \times n$–Matrix und $v \in \mathbb{R}^n$. Die Lösung $\varphi$ der homogenen linearen autonomen Differentialgleichung*

$$y' = Ay$$

*bezüglich der Anfangsbedingung $\varphi(0) = v$ ist gegeben durch*

$$\begin{aligned}
\varphi : \mathbb{R} &\to \mathbb{R}^n \\
t &\mapsto e^{At} v.
\end{aligned}$$

Anders ausgedrückt, der Fluss, welcher durch die Differentialgleichung $y' = Ay$ definiert ist, ist gegeben durch

$$\begin{aligned}
\Phi : \mathbb{R} \times \mathbb{R}^n &\to \mathbb{R}^n \\
(t, v) &\mapsto e^{At} v.
\end{aligned}$$

*Beweis.* Es sei also $\varphi(t) = e^{At}v$. Dann gilt

$$\begin{array}{rcl}
\varphi'(t) & = & \lim_{h\to 0} \frac{e^{A(t+h)}v - e^{At}v}{h} \\
& = & \lim_{h\to 0} \frac{(e^{At}e^{Ah}) - e^{At})v}{h} \\
& = & e^{At}\left(\lim_{h\to 0} \frac{e^{Ah}-\mathrm{id}}{h}\right) v \\
& = & e^{At}\left(\lim_{h\to 0} \frac{\mathrm{id} + Ah + \sum_{k=2}^{\infty} \frac{A^k}{k!}h^k\ - \mathrm{id}}{h}\right) v \\
& = & e^{At}\left(\lim_{h\to 0} A + \sum_{k=2}^{\infty} \underbrace{A^k h^{k-1}}_{\lim_{h\to 0}=0}\right) v \\
& = & e^{At}Av = A(e^{At}v) = A\cdot\varphi(t).
\end{array}$$

(In der vorletzten Gleichheit haben wir verwendet, dass $e^{At}A = Ae^{At}$, dies folgt daraus, dass $A^n$ und $A$ kommutieren.) $\square$

Wir haben damit, zumindest theoretisch, eine Lösung der homogenen linearen autonomen Differentialgleichung

$$y' = Ay$$

gefunden. Aber was soll $t \mapsto e^{At}v$ schon heißen?

14.3. **Homogene lineare autonome Differentialgleichungen II.** Zur Erinnerung:

(1) jede symmetrische Matrix $A$ über $\mathbb{R}$ kann diagonalisiert werden, genauer gesagt, es gibt eine orthogonale Matrix $P$, so dass $PAP^{-1}$ eine diagonale Matrix ist.
(2) zu jeder quadratischen Matrix $A$ gibt es eine invertierbare (i.a. komplexe) Matrix $P$, so dass $PAP^{-1}$ in Jordan–Form ist, d.h. $PAP^{-1}$ besteht aus Jordan–Blöcken von der Form

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda & 1 \\ 0 & \dots & 0 & 0 & \lambda \end{pmatrix}.$$

Die Jordan–Matrix $J$ ist natürlich nicht notwendigerweise eine reelle Matrix, sondern ist i.a. eine komplexe Matrix. Für eine komplexe Matrix $A$ kann man genau wie oben $e^A$ wie folgt definieren:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \in \mathrm{GL}(n, \mathbb{C}).$$

Es gelten zudem die gleichen Aussagen wie in Lemma 14.3 für komplexe Matrizen $A$ und $B$.

Der folgende Satz erlaubt es uns nun $e^{At}$ für eine beliebige Matrix $A$ schnell zu berechnen:

**Satz 14.5.** (1) *Es sei $A$ eine quadratische Matrix und $P$ eine invertierbare Matrix, dann gilt*

$$e^{PAP^{-1}t} = Pe^{At}P^{-1}.$$

(2) *Für zwei beliebige quadratische Matrizen $A, B$ gilt*

$$e^{\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} t} = \begin{pmatrix} e^{At} & 0 \\ 0 & e^{Bt} \end{pmatrix}.$$

(3) *Für*

$$J = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda & 1 \\ 0 & \dots & 0 & 0 & \lambda \end{pmatrix} \text{ ist } e^{Jt} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \dots & \frac{t^{n-2}}{(n-2)!} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & t \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

*Beweis.* (1) Es ist

$$\begin{aligned} (PAP^{-1})^k &= PAP^{-1}PAP^{-1} \cdot \dots \cdot PAP^{-1} \\ &= PA \cdot \dots \cdot AP^{-1} = PA^kP^{-1}. \end{aligned}$$

Es folgt, dass

$$\begin{aligned} e^{PAP^{-1}t} &= \textstyle\sum_{k=0}^{\infty} \frac{1}{k!}(PAP^{-1}t)^k \\ &= \textstyle\sum_{n=0}^{\infty} \frac{1}{k!}P(At)^kP^{-1} \\ &= P\left(\textstyle\sum_{k=0}^{\infty} \frac{1}{k!}(At)^n\right)P^{-1} = Pe^{At}P^{-1}. \end{aligned}$$

(2) Diese Aussage folgt leicht aus der Beobachtung, dass

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}^k = \begin{pmatrix} A^k & 0 \\ 0 & B^k \end{pmatrix}.$$

(3) Wir setzen

$$U := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & \vdots \\ & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix}.$$

Dann gilt

$$e^{Jt} = e^{(\lambda\,\mathrm{id}+U)t} = e^{\lambda\,\mathrm{id}\,t}e^{Ut} = e^{\lambda t}\sum_{k=0}^{\infty}\frac{1}{k!}(Ut)^k.$$

Nachdem $U^n$ die Nullmatrix ist, folgt, dass

$$\sum_{k=0}^{\infty}\frac{1}{k!}(Ut)^k = \sum_{k=0}^{n-1}\frac{1}{k!}(Ut)^k,$$

und diese Summe kann man leicht explizit berechnen, und man sieht, dass

$$\sum_{k=0}^{n-1}\frac{1}{k!}(Ut)^k = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \dots & \frac{t^{n-2}}{(n-2)!} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & t \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

□

*Beispiel.* Wir betrachten die Differentialgleichung

$$y' = Ay \text{ mit } A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Die Matrix $A$ hat Eigenwerte $\pm i$ mit Eigenvektoren $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ und $\begin{pmatrix} 1 \\ i \end{pmatrix}$. Wir setzen

$$P := \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}, \text{ dann gilt } P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Es folgt also, dass

$$\begin{aligned} e^{At} &= e^{P\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}P^{-1}t} \\ &= \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2}(e^{it}+e^{-it}) & \frac{1}{2}(ie^{it}-ie^{-it}) & \frac{1}{2}(-ie^{it}+ie^{-it}) \\ \frac{1}{2}(e^{it}+e^{-it}) & & \end{pmatrix}. \end{aligned}$$

Aus $e^{it} = \cos(t) + i\sin(t)$ folgt nun, dass

$$e^{At} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

Diese Lösung hatten weiter oben schon 'erraten' (siehe das Beispiel nach Satz 14.1), aber in diesem Beispiel sehen wir nun, dass wir die Lösung auf lineare Algebra zurückführen können.

14.4. **Homogene lineare autonome Differentialgleichungen in Dimension 2.** Sei nun $A$ eine reelle $2 \times 2$–Matrix. Wir können dann folgende Fälle unterscheiden:

(1) $A$ ist diagonalisierbar mit reellen Eigenwerten $\lambda_1 < \lambda_2 \in \mathbb{R}$. Wir unterscheiden:
   (a) $\lambda_1, \lambda_2 > 0$,
   (b) $\lambda_1, \lambda_2 < 0$,
   (c) $\lambda_1 < 0 < \lambda_2$, oder
   (d) $\lambda_1 = 0$ oder $\lambda_2 = 0$.
(2) $A$ ist diagonalisierbar mit komplexen Eigenwerten $\lambda$ und $\overline{\lambda}$,
(3) $A$ ist nicht diagonalisierbar.

In allen drei Fällen können wir den Fluss explizit beschreiben.

(1) Seien $v_1, v_2$ die Eigenvektoren zu den reellen Eigenwerten $\lambda_1$ und $\lambda_2$. Es folgt aus der Diskussion in dem vorherigen Kapitel (oder durch explizites nachrechnen), dass die Kurven $t \mapsto e^{\lambda_1 t} v_1$ und $t \mapsto e^{\lambda_2 t} v_2$ sind also Lösungskurven. In den vier obigen Fällen erhalten wir dann folgende Bilder für den Fluss. (Bilder gibt es nur in der Vorlesung)
(2) Es sei nun $\lambda = a + ib$ ein komplexer Eigenwert (d.h. $b \neq 0$), mit komplexen Eigenvektor $u+iv$ (wobei $u, v \in \mathbb{R}^2$). Dann kann man zeigen (wie im letzten Beispiel im vorherigen Kapitel), dass die Flusslinien Ellipsen um den Ursprung beschreiben, dabei sind die Hauptachsen durch $u$ und $v$ gegeben.
(3) Wenn $A$ nicht diagonalisierbar ist, dann besitzt $A$ einen reellen Eigenwert. [87] Der Fluss ist dann wie folgt gegeben:

[87] Jede Matrix besitzt einen Eigenwert, wenn eine reelle Matrix einen Eigenwert $\lambda$ besitzt, dann ist auch $\overline{\lambda}$ ein Eigenwert. Insbesondere, wenn eine reelle Matrix einen komplexen Eigenwert $\lambda = a + bi$ mit $b \neq 0$ besitzt, dann besitzt die Matrix zwei verschiedene Eigenwerte $\lambda$ und $\overline{\lambda}$. Eine $2 \times 2$–Matrix mit zwei verschiedenen Eigenwerten ist aber diagonalisierbar.