

# Mathematik III für Physiker\*

Wintersemester 2014/2015

Stefan Teufel  
Mathematisches Institut  
Uni Tübingen

3. März 2015

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



# Inhaltsverzeichnis

<b>1</b>	<b>Topologische, metrische und normierte Räume</b>	<b>1</b>
<b>2</b>	<b>Stetigkeit</b>	<b>9</b>
<b>3</b>	<b>Kompaktheit</b>	<b>15</b>
<b>4</b>	<b>Differenzierbarkeit</b>	<b>19</b>
<b>5</b>	<b>Taylorformel und lokale Extrema</b>	<b>35</b>
<b>6</b>	<b>Implizite Funktionen</b>	<b>41</b>
<b>7</b>	<b>Differentialrechnung in Banachräumen</b>	<b>49</b>
<b>8</b>	<b>Gewöhnliche Differentialgleichungen und dynamische Systeme</b>	<b>53</b>
<b>9</b>	<b>Lineare Differentialgleichungen</b>	<b>67</b>
<b>10</b>	<b>Funktionentheorie</b>	<b>77</b>
10.1	Differenzierbarkeit in $\mathbb{C}$ . . . . .	77
10.2	Komplexe Wegintegrale . . . . .	80
10.3	Der Cauchy Integralsatz . . . . .	85
10.4	Die Cauchy Integralformel . . . . .	89
10.5	Potenzreihen . . . . .	91
10.6	Laurentreihen und isolierte Singularitäten . . . . .	95
10.7	Der Residuenkalkül . . . . .	97
<b>11</b>	<b>Abstrakte Maß- und Integrationstheorie</b>	<b>103</b>
11.1	Das Inhaltsproblem . . . . .	104
11.2	Grundzüge der Maßtheorie . . . . .	105
11.3	Das Lebesgueintegral . . . . .	111



# 1 Topologische, metrische und normierte Räume

Um Konzepte wie Stetigkeit, Differenzierbarkeit und Integrierbarkeit z.B. für Funktionen  $f(x)$  mehrerer reeller Variable  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  zu definieren, muss man zunächst Begriffe wie Konvergenz von Folgen  $x_n \rightarrow a \in \mathbb{R}$ , offenes Intervall  $(a, b) \in \mathbb{R}$ , Vollständigkeit, etc. auf den  $\mathbb{R}^n$  verallgemeinern.

Da wir schließlich aber auch über den  $\mathbb{R}^n$  hinausgehen werden, machen wir das alles gleich etwas allgemeiner auf sogenannten metrischen Räumen.

## 1.1 Definition. Metrik

Sei  $X$  eine Menge. Eine Abbildung

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

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

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

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

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

Das Paar  $(X, d)$  heißt dann **metrischer Raum**.

## 1.2 Beispiele. Euklidische und diskrete Metrik

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

Die Eigenschaften (i) und (ii) aus Definition 1.1 sind offensichtlich, die Dreiecksungleichung nicht (vgl. Kapitel 6 aus MaPhy2). Falls nicht anders gesagt, sei der  $\mathbb{R}^n$  im Folgenden immer mit der euklidischen Metrik versehen.

(b) Sei  $X$  eine beliebige Menge und

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

Dann ist  $d$  eine Metrik auf  $X$  und heißt **diskrete Metrik**.

**1.3 Bemerkung.** Jede Teilmenge  $Y \subset X$  eines metrischen Raums  $(X, d)$  ist, mit derselben Metrik versehen, selbst wieder ein metrischer Raum  $(Y, d)$ .

## 1.4 Definition. Norm

Sei  $V$  ein Vektorraum über  $\mathbb{K}$ . Eine Abbildung

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

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

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

## 1 Topologische, metrische und normierte Räume

- (ii)  $\|\lambda x\| = |\lambda| \|x\|$ , (Homogenität)  
(iii)  $\|x + y\| \leq \|x\| + \|y\|$ . (Dreiecksungleichung)

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

### 1.5 Bemerkung. Jede Norm induziert eine Metrik

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

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

eine Metrik auf  $V$  definiert.

- Beweis.* (i)  $d(x, y) = 0 \Leftrightarrow \|x - y\| = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y$ ,  
(ii)  $d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |(-1)| \|y - x\| = \|y - x\| = d(y, x)$ ,  
(iii)  $d(x, z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$ .

□

### 1.6 Beispiele. (a) Die euklidische Norm auf $\mathbb{R}^n$ :

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

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

eine Norm definiert, die euklidische Norm. Die von ihr nach Bemerkung 1.5 induzierte Metrik ist die euklidische Metrik.

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

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

Im allgemeinen ist für  $x \in \mathbb{R}^n$  natürlich  $\|x\|_2 \neq \|x\|_\infty$ . Beispielsweise für  $x = (1, 1) \in \mathbb{R}^2$  ist  $\|x\|_2 = \sqrt{2}$  aber  $\|x\|_\infty = 1$ . (In den Übungen sollen Sie die "Einheitskugeln" des  $\mathbb{R}^2$  bezüglich verschiedener Normen skizzieren.)

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

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

Dann ist

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

eine Norm auf  $V$ .

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

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

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

□

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

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

durch

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

eine Norm definiert wird.

### 1.7 Merke. Zusammenhang zwischen Metrik und Norm

Metrik  $\hat{=}$  Abstand zwischen Punkten beliebiger Mengen  
 Norm  $\hat{=}$  Länge eines Vektors

Jede Norm induziert auch eine Metrik, da “Abstand” gleich “Länge des Differenzvektors” gesetzt werden kann.

Auf einem Vektorraum gilt also

Skalarprodukt	$\mapsto$	Norm	$\mapsto$	Metrik
$\langle \cdot, \cdot \rangle$		$\ \cdot\  = \sqrt{\langle \cdot, \cdot \rangle}$		$d(x, y) = \ x - y\ $

### 1.8 Definition. Offene Mengen in metrischen Räumen

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

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

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

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

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

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

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

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

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

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

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

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

(b) Ein offenes Intervall  $(a, b) \subset \mathbb{R}$  ist eine offene Menge bzgl. der euklidischen Metrik  $d(x, y) = |x - y|$ .

(c) Sei  $X$  beliebig und versehen mit der diskreten Metrik (Beispiel 1.2), dann ist jede Teilmenge  $U \subset X$  offen. (Klar:  $B_{\frac{1}{2}}(x) = \{x\} \forall x \in X$ .)

**1.10 Bemerkung.** In einem metrischen Raum  $X$  gilt:

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

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

(c) Sind  $U_i \subset X$  offen ( $i \in \mathcal{I}$ ,  $\mathcal{I}$  eine beliebige Indexmenge), so ist auch  $\bigcup_{i \in \mathcal{I}} U_i$  offen.

## 1 Topologische, metrische und normierte Räume

*Beweis.* (a) Offenbar.

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

□

**1.11 Bemerkung.** Beliebige Durchschnitte offener Mengen sind im Allgemeinen nicht mehr offen, z.B. ist  $I_n = (-\frac{1}{n}, \frac{1}{n}) \subset \mathbb{R}$  für alle  $n \in \mathbb{N}$  offen, aber  $\bigcap_{n \in \mathbb{N}} I_n = \{0\}$  sicher nicht.

### 1.12 Definition. Abgeschlossene Mengen

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

**1.13 Beispiele.** (a)  $[a, b] \subset \mathbb{R}$  ist abgeschlossen.

(b)  $[a, \infty) \subset \mathbb{R}$  ist abgeschlossen.

(c)  $[a, b) \subset \mathbb{R}$  ist für  $-\infty < a < b < \infty$  weder abgeschlossen noch offen.

(d) Sei  $(X, d)$  metrischer Raum. Dann sind  $\emptyset$  und  $X$  sowohl offen als auch abgeschlossen.

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

**1.15 Bemerkung.** In einem metrischen Raum  $X$  gilt:

(a)  $\emptyset$  und  $X$  sind abgeschlossen.

(b) Sind  $U, V \subset X$  abgeschlossen, so ist auch  $U \cup V$  abgeschlossen.

(c) Beliebige Schnitte abgeschlossener Mengen sind abgeschlossen.

*Beweis.* Komplementbildung in Bemerkung 1.10.

□

Man kann nun die folgenden Betrachtungen noch verallgemeinern, indem man vergisst, dass die offenen Mengen mit Hilfe einer Metrik definiert wurden und stattdessen die Eigenschaften in Bemerkung 1.10 zur Definition erhebt.

### 1.16 Definition. Topologische Räume

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

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

(b)  $U, V \in \mathcal{T} \Rightarrow U \cap V \in \mathcal{T}$

(c)  $U_i \in \mathcal{T}$  für  $i \in I \Rightarrow \bigcup_{i \in I} U_i \in \mathcal{T}$

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

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

**Zusammenfassend** sollten Sie sich merken, dass man in einem topologischen Raum weiß, was die offenen Mengen, was Umgebungen und was innere Punkte sind. Ausgehend davon folgen, wie wir sehen werden, Begriffe wie Konvergenz, Stetigkeit, Kompaktheit, etc.

**1.17 Bemerkung.** Topologische Räume in denen die Topologie nicht durch eine Metrik gegeben ist werden im folgenden keine große Rolle spielen. Trotzdem ist es nützlich und in vielen Fällen auch einfacher, den Begriff der Metrik möglichst selten zu verwenden und stattdessen mit Umgebungen und offenen Mengen zu argumentieren. In dieser Vorlesung können Sie aber stets Ihre Intuition über die offenen Mengen im  $\mathbb{R}^n$  bezüglich der euklidischen Metrik verwenden.

**1.18 Definition. Inneres, Abschluss, Rand**

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

- (a)  $\overset{\circ}{Y} := \bigcup_{\substack{U \subset Y \\ U \text{ offen}}} U$  das **Innere** oder der **offene Kern** von  $Y$ ,
- (b)  $\overline{Y} := \bigcap_{\substack{A \supset Y \\ A \text{ abgeschlossen}}} A$  der **Abschluss** oder die **abgeschlossene Hülle** von  $Y$ ,
- (c)  $\partial Y = \overline{Y} \setminus \overset{\circ}{Y}$  der **Rand** von  $Y$ .

Es gilt also jeweils per Definition

- $\overset{\circ}{Y} \subset Y \subset \overline{Y}$
- $\overset{\circ}{Y}$  ist die größte in  $Y$  enthaltene offene Menge, insbesondere ist  $\overset{\circ}{Y}$  offen.
- $Y$  ist offen  $\Leftrightarrow Y = \overset{\circ}{Y}$
- $\overline{Y}$  ist die kleinste abgeschlossene Menge in der  $Y$  enthalten ist, insbesondere ist  $\overline{Y}$  abgeschlossen.
- $Y$  ist abgeschlossen  $\Leftrightarrow Y = \overline{Y}$
- $\overset{\circ}{Y}$  ist das Komplement von  $\overline{X \setminus Y}$
- $\overline{Y}$  ist das Komplement von  $(X \setminus Y)^\circ$

**1.19 Bemerkung. Alternative Charakterisierungen**

In einem topologischen Raum  $X$  gilt:

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

*Beweis.* Übungsaufgabe □

- 1.20 Beispiele.** (a) Sei  $Y = (a, b) \subset \mathbb{R}$ . Dann ist  $\overset{\circ}{Y} = (a, b)$ ,  $\overline{Y} = [a, b]$  und  $\partial Y = \{a, b\}$ .  
 (b) Betrachte  $\mathbb{Q} \subset \mathbb{R}$ : Da in jedem  $\varepsilon$ -Ball  $(x - \varepsilon, x + \varepsilon)$  sowohl eine rationale als auch eine irrationale Zahl liegt, ist  $\partial \mathbb{Q} = \mathbb{R}$ . Somit sind  $\overset{\circ}{\mathbb{Q}} = \mathbb{Q} \setminus \partial \mathbb{Q} = \emptyset$  und  $\overline{\mathbb{Q}} = \mathbb{Q} \cup \partial \mathbb{Q} = \mathbb{R}$ .

**1.21 Definition. Konvergenz von Folgen**

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

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

oder kurz  $x_n \rightarrow a$  für  $n \rightarrow \infty$ , wenn gilt:  
 Für jede Umgebung  $U \subset X$  von  $a$  gibt es ein  $N \in \mathbb{N}$ , so dass  $x_n \in U$  für alle  $n \geq N$ . Eine Folge konvergiert also gegen einen Punkt, wenn jede (noch so kleine) Umgebung des Punktes alle bis auf endlich viele Folgenglieder enthält.

**1.22 Bemerkung. Konvergenz in metrischen Räumen**

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

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

*Beweis.* Übungsaufgabe. □

**1.23 Definition. Häufungspunkt**

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

**1.24 Satz. Folgenkriterium für den Abschluss in metrischen Räumen**

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

$$a \in \bar{A} \iff \exists (x_n) \in A : \lim_{n \rightarrow \infty} x_n = a.$$

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

*Beweis.* „ $\Rightarrow$ “: Sei  $a \in \bar{A}$ . Nach Bemerkung 1.19 (b) gilt für jedes  $n \in \mathbb{N}$ , dass  $B_{\frac{1}{n}}(a) \cap A \neq \emptyset$ . Wir können also für jedes  $n \in \mathbb{N}$  ein  $x_n \in B_{\frac{1}{n}}(a) \cap A$  auswählen. Diese Folge liegt ganz in  $A$  und konvergiert gegen  $a$ .

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

**1.25 Definition. Cauchyfolge**

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

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

**1.26 Bemerkung.** Jede konvergente Folge ist auch eine Cauchy-Folge: Sei  $\lim x_n = a$ . Dann gibt es zu jedem  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  so, dass  $d(x_n, a) < \frac{\varepsilon}{2}$  für alle  $n \geq N$ . Wegen der Dreiecksungleichung ist dann für  $n, m \geq N$  aber  $d(x_n, x_m) \leq d(x_n, a) + d(x_m, a) < \varepsilon$ . Die Umkehrung gilt nur in vollständigen Räumen.

**1.27 Definition. Vollständigkeit und Banachraum**

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

**1.28 Bemerkung.** Konvergenz, Kompaktheit, Stetigkeit, Abschluss, Rand etc. sind topologische Begriffe. Die Konzepte Cauchyfolge und Vollständigkeit benötigen mehr Struktur, z.B. eine Metrik.

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

### 1.29 Definition. Äquivalenz von Normen

Zwei Normen  $\|\cdot\|_1$  und  $\|\cdot\|_2$  auf einem Vektorraum  $V$  heißen **äquivalent**, wenn es Konstanten  $c, C > 0$  gibt, so dass für alle  $x \in V$  gilt:

$$c \|x\|_1 \leq \|x\|_2 \leq C \|x\|_1.$$

(Umgekehrt gilt dann offensichtlich auch

$$\frac{1}{C} \|x\|_2 \leq \|x\|_1 \leq \frac{1}{c} \|x\|_2. \quad )$$

**1.30 Bemerkung.** Seien  $\|\cdot\|_1$  und  $\|\cdot\|_2$  äquivalente Normen auf  $V$ . Dann gilt:

- (a)  $\lim x_n = a$  in  $(V, \|\cdot\|_1)$   $\Leftrightarrow$   $\lim x_n = a$  in  $(V, \|\cdot\|_2)$ .
- (b)  $(x_n)$  ist Cauchy in  $(V, \|\cdot\|_1)$   $\Leftrightarrow$   $(x_n)$  ist Cauchy in  $(V, \|\cdot\|_2)$ .
- (c)  $(V, \|\cdot\|_1)$  ist vollständig  $\Leftrightarrow$   $(V, \|\cdot\|_2)$  ist vollständig.

*Beweis.* (a) Sei  $\lim x_n = a$  in  $(V, \|\cdot\|_1)$ , also  $\lim_{n \rightarrow \infty} \|x_n - a\|_1 = 0$ . Dann ist auch  $\lim_{n \rightarrow \infty} \|x_n - a\|_2 \leq \lim_{n \rightarrow \infty} C \|x_n - a\|_1 = 0$ . (b) Analog. (c) folgt aus (a) und (b). □



## 2 Stetigkeit

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

### 2.1 Definition. Folgenstetigkeit

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

$$\lim_{n \rightarrow \infty} x_n = a \quad \Rightarrow \quad \lim_{n \rightarrow \infty} f(x_n) = f(a) \quad \text{für jede Folge } (x_n) \text{ in } X.$$

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

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

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

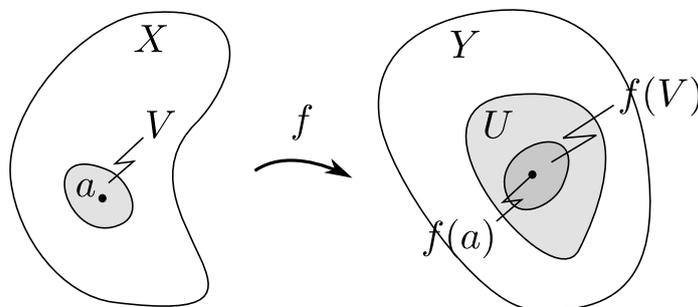
Stetigkeit von  $f$  in  $a$  bedeutet daher in dieser Schreibweise

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

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

### 2.3 Definition. Stetigkeit

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



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

### 2.4 Bemerkung. Stetigkeit impliziert Folgenstetigkeit

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

## 2 Stetigkeit

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

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

### 2.5 Bemerkung. “ $\varepsilon$ - $\delta$ -Stetigkeit” in metrischen Räumen

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

Anders gesagt: Jede  $\varepsilon$ -Kugel um  $f(a)$  enthält das Bild einer  $\delta$ -Kugel um  $a$ .

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

### 2.6 Bemerkung. In metrischen Räumen impliziert Folgenstetigkeit auch Stetigkeit

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

*Beweis.* Angenommen es gibt ein  $\varepsilon > 0$ , so dass für alle  $\delta > 0$  gilt  $f(B_\delta(a)) \not\subset B_\varepsilon(f(a))$ . Für  $\delta = \frac{1}{n}$  wähle dann  $x_n \in B_\delta(a) \setminus f^{-1}(B_\varepsilon(f(a)))$ . Also  $\lim x_n = a$ , aber  $f(x_n) \notin B_\varepsilon(f(a))$  für alle  $n \in \mathbb{N}$  und somit konvergiert  $f(x_n)$  nicht gegen  $f(a)$ .  $\square$

### 2.7 Satz. Charakterisierung von Stetigkeit durch Urbilder offener Mengen

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

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

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

### 2.8 Beispiele. Stetige Abbildungen in metrischen Räumen

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

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

stetig.

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

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

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

- (b) In einem normierten Raum  $(V, \|\cdot\|)$  ist  $\|\cdot\| : V \rightarrow \mathbb{R}$  stetig. *Beweis.*  $\|x\| = d_0(x)$  in (a).
- (c) Die Addition  $\text{add}: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (x, y) \mapsto x + y$  ist stetig, die Multiplikation  $\text{mult}: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (x, y) \mapsto x \cdot y$  ist stetig, die Division  $\text{div}: \mathbb{R} \times \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, (x, y) \mapsto xy^{-1}$  ist stetig.  
*Beweis.* Sei  $(x_n, y_n) \rightarrow (x, y)$  also  $x_n \rightarrow x$  und  $y_n \rightarrow y$  in  $\mathbb{R}$ , dann gilt (MaPhy1)  $x_n + y_n \rightarrow x + y$  usw. □
- (d) Die Hintereinanderschaltung (Komposition) stetiger Abbildungen ist stetig: Seien  $f : Y \rightarrow Z$  und  $g : X \rightarrow Y$  stetig, dann ist  $f \circ g : X \rightarrow Z$  stetig.  
*Beweis.* Sei  $O \subset Z$  offen  $\xrightarrow{f \text{ stetig}} f^{-1}(O) \subset Y$  offen  $\xrightarrow{g \text{ stetig}} g^{-1}(f^{-1}(O)) = (f \circ g)^{-1}(O)$  ist offen in  $X$ . Also ist  $f \circ g$  mit Satz 2.7 stetig. □
- (e) Eine Abbildung  $f : X \rightarrow \mathbb{R}^n, x \mapsto f(x) = (f_1(x), \dots, f_n(x))$ , ist genau dann stetig, wenn alle  $f_i : X \rightarrow \mathbb{R}, i = 1, \dots, n$ , stetig sind.
- (f) Sind  $f, g : X \rightarrow \mathbb{R}$  stetig, so sind auch die Funktionen  $f + g, f \cdot g : X \rightarrow \mathbb{R}$  und, falls  $g(x) \neq 0 \forall x \in X$ ,  $f/g : X \rightarrow \mathbb{R}$  stetig.  
*Beweis.* Z.B. ist  $f + g = \text{add} \circ (f, g)$  mit (c), (d) und (e) stetig. □

## 2.9 Satz. Alle Normen im $\mathbb{R}^n$ sind äquivalent.

*Beweis.* Es genügt zu zeigen, dass jede Norm  $\|\cdot\|$  zu  $\|\cdot\|_1$  mit  $\|x\|_1 = \sum_{j=1}^n |x_j|$  äquivalent ist. Sei  $x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$ . Dann gilt

$$\|x\| = \left\| \sum x_j e_j \right\| \leq \sum |x_j| \|e_j\| \leq n \max\{\|e_j\| \mid j = 1, \dots, n\} \sum |x_j| =: C \|x\|_1.$$

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

$$\frac{1}{k} \|x_k\|_1 > \|x_k\| \quad \text{und o.B.d.A.} \quad \|x_k\|_1 = 1. \quad (2.1)$$

Diese enthält eine bzgl.  $\|\cdot\|_1$  konvergente Teilfolge, wieder mit  $(x_k)$  bezeichnet, da die Komponentenfolgen  $(x_{k,j})$  beschränkte Folgen in  $\mathbb{R}$  sind und somit nach dem Satz von Bolzano-Weierstraß konvergente Teilfolgen enthalten. Wegen (2.1) konvergiert die Folge  $(x_k)$  auch in der  $\|\cdot\|$ -Norm gegen den eindeutigen Grenzwert  $x := \lim_{k \rightarrow \infty} x_k$ . Aufgrund der Stetigkeit der Normen ist dann einerseits  $1 = \lim \|x_k\|_1 = \|\lim x_k\|_1 = \|x\|_1$ , also  $x \neq 0$ . Andererseits ist  $\|x\| = \lim \|x_k\| \leq \lim \frac{1}{k} \|x_k\|_1 = 0$ , also  $x = 0$ . □

## 2.10 Bemerkung. $\mathbb{R}^n$ ist bezüglich jeder Norm vollständig.

*Beweis.* Mit Bemerkung 1.30 (c) und Satz 2.9 genügt es, die Vollständigkeit von  $(\mathbb{R}^n, \|\cdot\|_\infty)$  nachzuweisen. Sei dazu  $(x_k)$  eine Cauchyfolge in  $(\mathbb{R}^n, \|\cdot\|_\infty)$ , dann ist wegen  $|x_{k,j} - x_{\ell,j}| \leq \|x_k - x_\ell\|_\infty$  auch  $(x_{k,j})$  für jedes  $j = 1, \dots, n$  eine Cauchyfolge in  $\mathbb{R}$  und somit wegen der Vollständigkeit von  $\mathbb{R}$  konvergent. Sei  $\lim_{k \rightarrow \infty} x_{k,j} = a_j$  und  $a = (a_1, \dots, a_n)$ , dann ist  $\|x_k - a\|_\infty = \max\{|x_{k,1} - a_1|, \dots, |x_{k,n} - a_n|\} < \varepsilon$  für  $k$  hinreichend groß und somit  $\lim x_k = a$ . □

## 2.11 Definition. Punktweise- und gleichmäßige Konvergenz von Funktionenfolgen

Seien  $X$  eine Menge,  $Y$  ein metrischer Raum und

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

## 2 Stetigkeit

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

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

oder, ausführlich,

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

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

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

oder, ausführlich,

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

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

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

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

### 2.13 Satz. Gleichmäßige Limites stetiger Funktionen sind stetig

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

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

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

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

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

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

Im Schritt (\*) braucht man die gleichmäßige Konvergenz, da  $|f(x_k) - f_N(x_k)| < \frac{\varepsilon}{3}$  für alle  $x_k$  mit  $k \geq k_1$  gelten muss.

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

*Beweis.* Sei  $(f_n)$  eine  $\|\cdot\|_\infty$ -Cauchyfolge in  $C(X, Y)$ . Dann ist  $(f_n(x))$  für jedes  $x \in X$  eine Cauchyfolge in  $Y$ , denn  $\|f_n(x) - f_m(x)\| \leq \|f_n - f_m\|_\infty$ . Da  $Y$  vollständig ist, konvergiert  $(f_n(x))$  in  $Y$  und wir nennen den Grenzwert  $f(x)$ . Es konvergiert also  $(f_n)$  punktweise gegen die so definierte Funktion  $f$ . Wenn wir zeigen, dass  $(f_n) \rightarrow f$  gleichmäßig, dann folgt, mit Satz 2.13, dass  $f$  stetig ist und somit  $f \in C(X, Y)$  und  $\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0$ .

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

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

Also  $(f_n) \rightarrow f$  gleichmäßig. □

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

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

### 2.17 Satz. Banachscher Fixpunktsatz

Sei  $X$  ein vollständiger metrischer Raum,  $A$  eine abgeschlossene Teilmenge und  $f : A \rightarrow A$  eine Abbildung. Falls es eine Konstante  $0 < \theta < 1$  gibt, so dass für alle  $x, y \in A$  gilt

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

also  $f$  eine **Kontraktion** ist, dann hat  $f$  genau einen Fixpunkt  $a \in A$ , d.h. es gilt  $f(a) = a$  für genau ein  $a \in A$ .

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

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

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

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

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

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

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

Aufgrund der Vollständigkeit von  $X$  konvergiert die Cauchyfolge  $(x_n)$  gegen ein  $a \in X$ . Da alle  $x_n$  in  $A$  liegen und  $A$  abgeschlossen ist, muss auch  $a \in A$  sein. Schließlich gilt wegen der Stetigkeit von  $f$  in  $a$  (Übungsaufgabe: Kontraktion  $\Rightarrow$  Stetigkeit), dass

$$f(a) = f(\lim x_n) = \lim f(x_n) = \lim x_{n+1} = a,$$

also ist  $a$  ein Fixpunkt. □



# 3 Kompaktheit

## 3.1 Definition. Offene Überdeckung und kompakte Menge

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

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

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

- (b) Eine Teilmenge  $K \subset X$  heißt **kompakt**, wenn gilt: zu *jeder* offenen Überdeckung  $(U_i)_{i \in \mathcal{I}}$  von  $K$  gibt es eine **endliche Teilüberdeckung** von  $K$ , d.h. es gibt  $i_1, \dots, i_n \in \mathcal{I}$  so, dass bereits

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

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

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

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

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

kompakt.

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

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

## 3.3 Satz. Bolzano-Weierstraß

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

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

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

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

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

Dann liegen aber nur endlich viele Folgenglieder in  $K$  : Widerspruch!  $\square$

### 3 Kompaktheit

**3.4 Bemerkung.** In metrischen Räumen gilt auch die Umkehrung: Sei  $K$  Teilmenge eines metrischen Raumes  $X$  und habe jede Folge in  $K$  auch einen Häufungspunkt in  $K$ , dann ist  $K$  kompakt. (Übungsaufgabe).

### 3.5 Definition. Beschränkte Mengen und ihr Durchmesser

Sei  $X$  ein metrischer Raum.

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

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

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

### 3.6 Satz. Kompakta sind abgeschlossen und beschränkt

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

**Warnung:** Die Umkehrung gilt im Allgemeinen nicht! (vgl. Satz 3.8)

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

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

Damit ist für alle  $x, y \in K$

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

also ist  $K$  beschränkt.

*Abgeschlossenheit:* Sei  $(x_n)$  eine konvergente Folge in  $K$ , also  $\lim x_n = a \in X$  und  $x_n \in K$  für alle  $n \in \mathbb{N}$ . Nach Bolzano-Weierstraß hat  $(x_n)$  einen Häufungspunkt in  $K$ , welcher gleich  $a$  sein muss, denn eine konvergente Folge hat genau einen Häufungspunkt, welcher gleich dem Grenzwert ist. Gemäss Satz 1.24 ist  $K$  also abgeschlossen.  $\square$

### 3.7 Bemerkung. Abgeschlossene Teilmengen kompakter Mengen sind kompakt

Sei  $X$  ein topologischer Raum. Ist  $K \subset X$  eine abgeschlossene Teilmenge und ist  $K$  enthalten in einer kompakten Menge  $L \subset X$ , also  $K \subset L$ , so ist auch  $K$  kompakt.

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

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

und somit auch  $K$  kompakt.  $\square$

### 3.8 Satz. Heine-Borel

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

*Beweis.* “ $\Rightarrow$ ”: Gilt nach Satz 3.6 in beliebig metrischen Räumen.

“ $\Leftarrow$ ”: Sei  $K \subset \mathbb{R}^n$  beschränkt und abgeschlossen. Dann gibt es ein  $R > 0$  so, dass  $K$  im Würfel  $W_R$  mit Kantenlänge  $2R$  enthalten ist,  $K \subset [-R, R]^n =: W_R$ . Wir zeigen in Satz 3.10, dass  $W_R$  kompakt ist, also ist mit Bemerkung 3.7 auch  $K$  kompakt.  $\square$

### 3.9 Proposition. Schachtelungsprinzip

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

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

*Beweis. Eindeutigkeit:* Seien  $a, b \in \bigcap_{n \in \mathbb{N}} A_n$  und  $a \neq b$ . Dann ist  $d(a, b) > 0$ , was im Widerspruch zu  $d(a, b) \leq \text{diam}(A_n) \xrightarrow{n \rightarrow \infty} 0$  steht.

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

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

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

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

*Beweis.* Sei  $(U_i)_{i \in \mathcal{I}}$  eine offene Überdeckung von  $W$  und nehme an, dass es keine endliche Teilüberdeckung gibt. Zerlege  $W_0 := W$  in  $2^n$  abgeschlossene Würfel der Kantenlänge  $R$ . Dann erlaubt auch einer dieser kleineren Würfel keine endliche Teilüberdeckung. Diesen nennen wir  $W_1$  und konstruieren so eine Folge  $W_k$  abgeschlossener Würfel der Kantenlänge  $2^{-k} \cdot 2R$ , die alle keine endliche Teilüberdeckung erlauben. Nach Konstruktion gilt dann  $W_{k+1} \subset W_k$  und  $\lim_{k \rightarrow \infty} \text{diam}(W_k) \leq \lim_{k \rightarrow \infty} 2^{-k} \cdot 2R = 0$ . Mit dem Schachtelungsprinzip 3.10 und der Vollständigkeit von  $\mathbb{R}^n$  folgt  $\{a\} = \bigcap_k W_k \subseteq W$ . Da  $a \in W \subset \bigcup_{i \in \mathcal{I}} U_i$ , gibt es ein  $i_0 \in \mathcal{I}$  mit  $a \in U_{i_0}$ . Da  $U_{i_0}$  offen ist, existiert ein  $\varepsilon > 0$  so, dass  $B_\varepsilon(a) \subset U_{i_0}$ . Sei nun  $k_0 \in \mathbb{N}$  so groß, dass  $\text{diam}(W_{k_0}) < \varepsilon$ , also  $W_{k_0} \subset B_\varepsilon(a) \subset U_{i_0}$ . Aber dann wird  $W_{k_0}$  schon von einem einzigen der  $U_i$  überdeckt, obwohl es nach der Konstruktion nicht mal von endlich vielen überdeckt wird. Widerspruch!  $\square$

### 3.11 Satz. Stetige Bilder kompakter Mengen sind kompakt

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

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

### 3.12 Satz. Weierstraß

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

### 3 Kompaktheit

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

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

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

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

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

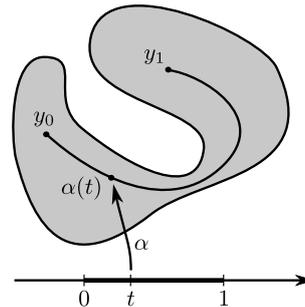
*Beweis. von Satz 3.12:* Sei  $c := \sup_{x \in K} f(x) \in \mathbb{R} \cup \{\infty\}$ . Es gibt dann, nach Definition des Supremums, eine Folge  $(x_n)$  in  $K$  mit  $\lim_{n \rightarrow \infty} f(x_n) = c$ . Da  $K$  kompakt ist, hat  $(x_n)$  einen Häufungspunkt  $a \in K$ . Also existiert eine Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  von  $(x_n)_{n \in \mathbb{N}}$ , sodass  $\lim_{k \rightarrow \infty} x_{n_k} = a$ . Wegen der Stetigkeit von  $f$  gilt aber

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

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

### 3.14 Definition. Wegzusammenhang

Sei  $X$  ein topologischer Raum. Eine Teilmenge  $Y \subset X$  heißt **wegzusammenhängend**, wenn es zu je zwei Punkten  $y_0, y_1 \in Y$  einen **Weg** von  $y_0$  nach  $y_1$  in  $Y$  gibt, d.h. eine stetige Abbildung  $\alpha : [0, 1] \rightarrow Y$  mit  $\alpha(0) = y_0$  und  $\alpha(1) = y_1$ .



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

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

### 3.16 Definition. Gebiet

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

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

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

## 4 Differenzierbarkeit

### 4.1 Definition. Partielle Ableitung

Sei  $n \geq 1$  und  $G \subset \mathbb{R}^n$  ein Gebiet. Für  $x \in G$  und  $j \in \{1, \dots, n\}$  heißt eine Funktion  $f : G \rightarrow \mathbb{R}$  im Punkt  $x$  in die  $j$ -te Koordinatenrichtung partiell differenzierbar, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x + h e_j) - f(x)}{h}$$

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

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

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

**4.2 Bemerkung.** Zur Erinnerung: Dass der  $\lim_{h \rightarrow 0}$  existiert bedeutet, dass für alle Nullfolgen  $(h_k)_{k \in \mathbb{N}}$  in  $\mathbb{R} \setminus \{0\}$  mit  $h_k$  derart, dass  $x + h_k e_j \in G$  ist, der Grenzwert

$$\lim_{k \rightarrow \infty} \frac{f(x + h_k e_j) - f(x)}{h_k}$$

existiert und damit auch für all diese Folgen gleich ist.

**4.3 Bemerkung.** Man führt also in dieser Definition die Differenzierbarkeit einer Funktion mehrerer Veränderlicher auf die Situation einer reellen Veränderlichen zurück. Setzt man  $g(x_j) := f(x_1, x_2, \dots, x_j, \dots, x_n)$  für festes  $x_i$  falls  $i \neq j$ , so ist  $\frac{\partial f}{\partial x_j}(x) = g'(x_j)$ .

### 4.4 Beispiel. Euklidische Norm

Betrachte die Radiusfunktion

$$r : \mathbb{R}^n \rightarrow [0, \infty), \quad x \mapsto r(x) = \|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}.$$

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

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

### 4.5 Definition. Partielle und stetige partielle Differenzierbarkeit

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f$  eine reelle Funktion auf  $G$ , also  $f : G \rightarrow \mathbb{R}$ .

- (a) Für  $x \in G$  heißt  $f$  in  $x$  **partiell differenzierbar**, wenn  $f$  in alle  $n$  Richtungen partiell differenzierbar ist. Man nennt dann den Vektor

$$\text{grad } f(x) = \nabla f(x) = (\partial_1 f(x), \dots, \partial_n f(x)) \in \mathbb{R}^n$$

den **Gradienten von  $f$**  in  $x$ .

#### 4 Differenzierbarkeit

- (b) Es heißt  $f$  **partiell differenzierbar in  $G$** , falls  $f$  in allen Punkten in  $G$  partiell differenzierbar ist.
- (c) Es heißt  $f$  **stetig partiell differenzierbar in  $G$** , falls  $f$  partiell differenzierbar in  $G$  ist und die partiellen Ableitungen  $\partial_j f : G \rightarrow \mathbb{R}$ ,  $j = 1, \dots, n$ , stetig sind.

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

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

Dann ist  $f$  nämlich in  $(x, y) = (0, 0)$  partiell differenzierbar mit  $\text{grad}f(0, 0) = (0, 0)$ , aber offenbar nicht stetig. Wir werden jedoch sehen, dass jede Funktion  $f$  die in einem Gebiet stetig partiell differenzierbar ist, dort auch stetig ist.

#### 4.7 Bemerkung. Rechenregeln für partielle Ableitungen

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

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

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

Insbesondere ist also

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

ein reeller Vektorraum.

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

Für  $j \in \{1, \dots, n\}$  gilt dann

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

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

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

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

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

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

#### 4.9 Definition. Vektorwertige Funktionen

Seien  $m, n \geq 1$ ,  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  eine vektorwertige Funktion, also  $f = (f_1, \dots, f_m)$  mit  $f_i : G \rightarrow \mathbb{R}$  für  $i = 1, \dots, m$ . Wir sagen, dass,  $f$  (**stetig**) **partiell differenzierbar** ist, falls jede Komponente  $f_i$  (stetig) partiell differenzierbar ist.

#### 4.10 Definition. Vektorfeld

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

#### 4.11 Beispiel. Gradient als Vektorfeld

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

#### 4.12 Definition. Divergenz und Laplace

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

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

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

die **Divergenz** von  $f$ .

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

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

**Laplace von  $f$ .**

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

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

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

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

Also ist  $\Delta r = \text{div}(\text{grad} r) = \frac{n-1}{r}$ .

**4.14 Definition.** Die  $r$ -mal stetig partiell differenzierbaren Funktionen  $C^r(G)$

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

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

Wiederholte Anwendung von 4.7 (a) liefert, dass

$$C^r(G) := \{f : G \rightarrow \mathbb{R} \mid f \text{ ist } r\text{-mal stetig partiell differenzierbar}\}$$

ein reeller Vektorraum ist.

**4.15 Notation.** Ist  $f$   $r$ -mal stetig partiell differenzierbar, so schreibt man auch

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

für jedes  $j = (j_1, \dots, j_r) \in \{1, \dots, n\}^r$ .

**4.16 Satz. von Schwarz**

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

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

**4.17 Lemma.** Ist  $f : \mathbb{R}^2 \supset G \rightarrow \mathbb{R}$  stetig partiell differenzierbar und ist das Rechteck  $[a, b] \times [c, d]$  ganz in  $G$  gelegen, so liefert

$$F(x) = \int_c^d f(x, y) dy$$

eine auf ganz  $(a, b)$  differenzierbare Funktion mit

$$F'(x) = \int_c^d \frac{\partial f}{\partial x}(x, y) dy.$$

*Beweis.* Der Beweis ist elementar: man zeigt direkt mit Hilfe des Mittelwertsatzes, dass der Differenzenquotient von  $F$  konvergiert, siehe z.B. Fischer-Kaul §23:2.3. Da wir im vierten Semester ein viel allgemeineres Resultat zu Integralen mit Parametern beweisen werden, verzichten wir an dieser Stelle auf einen Beweis. □

*Beweis. von Satz 4.16.* Für einen festen Punkt  $x \in G$  und  $i \neq j$  sei  $\varphi(u, v) := f(x + u e_i + v e_j)$ . Zu zeigen ist dann  $\partial_u \partial_v \varphi(0, 0) = \partial_v \partial_u \varphi(0, 0)$ . Die Funktion  $\varphi$  ist in einer Umgebung  $U$  des Nullpunktes in  $\mathbb{R}^2$  definiert, welche ein kompaktes Quadrat  $Q_r = \{(u, v) \mid |u| \leq r, |v| \leq r\}$  enthält. Der Hauptsatz der Differential- und Integralrechnung liefert

$$\varphi(u, v) = \varphi(u, 0) + \int_0^v \partial_v \varphi(u, t) dt$$

in  $Q_r$  und nach Lemma 4.17 angewandt auf  $\partial_v \varphi$  gilt

$$\partial_u \varphi(u, v) - \partial_u \varphi(u, 0) = \int_0^v \partial_u \partial_v \varphi(u, t) dt.$$

Leitet man diese Gleichung nach  $v$  ab, liefert der Hauptsatz der Differential- und Integralrechnung

$$\partial_v \partial_u \varphi(u, v) = \partial_u \partial_v \varphi(u, v) \quad \text{in } Q_r.$$

□

#### 4.18 Definition. Hesse-Matrix

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

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

die **Hesse-Matrix von  $f$  in  $x$** . Wegen Satz 4.16 ist  $\text{Hess}f(x)$  symmetrisch.

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

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

#### 4.20 Definition. Rotation eines Vektorfeldes

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

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

#### 4.21 Notation. Nabla-Operator

Führt man  $\nabla$  als “vektorwertigen Operator” ein, den sogenannten **Nabla-Operator**,

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

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

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

**4.22 Korollar. zu Satz 4.16.** Sei  $G \subset \mathbb{R}^3$  ein Gebiet.

(a) Für  $f \in C^2(G)$  gilt  $\text{rot}(\text{grad}f) = 0$ .

#### 4 Differenzierbarkeit

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

*Beweis.* Übungsaufgabe: Man rechne nach, dass man  $\nabla$  formal wie einen Vektor behandeln kann:

$$\begin{aligned}\operatorname{rot}(\operatorname{grad} f) &= \nabla \times \nabla f = 0, \\ \operatorname{div}(\operatorname{rot} v) &= \nabla \cdot (\nabla \times v) = 0.\end{aligned}$$

□

**4.23 Beispiel.** Sei  $F : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$  zweimal stetig partiell differenzierbar und rotationssymmetrisch, also  $F(x) = f(r(x))$  mit  $r(x) = \|x\|$  und  $f \in C^2((0, \infty))$ . Dann ist auch  $\Delta F : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$  rotationssymmetrisch, also  $\Delta F = g \circ r$  für eine stetige Funktion  $g : (0, \infty) \rightarrow \mathbb{R}$  und es gilt

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

*Beweis.* Übungsaufgabe

□

Wir führen nun einen etwas anderen Differenzierbarkeitsbegriff ein, der im Gegensatz zur partiellen Differenzierbarkeit geometrisch motiviert ist und die geometrische Bedeutung der Ableitung als **lineare Approximation** in den Vordergrund stellt.

Für  $f : I \rightarrow \mathbb{R}$ ,  $I \subset \mathbb{R}$  ein offenes Intervall, bedeutet Differenzierbarkeit, dass

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

wobei

$$\varphi(h) = f(x+h) - f(x) - hf'(x).$$

Also kann man  $f$  in einer hinreichend kleinen Umgebung von  $x$  so schreiben:

$$f(x+h) = f(x) + hf'(x) + \varphi(h),$$

wobei der Term  $\varphi(h)$  für  $h \rightarrow 0$  gegenüber den anderen vernachlässigt werden kann, da

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

Es existiert also ein  $\delta > 0$  und eine Funktion  $\varphi : (-\delta, \delta) \rightarrow \mathbb{R}$  und eine lineare Abbildung  $A : \mathbb{R} \rightarrow \mathbb{R}$  so, dass

$$f(x+h) = f(x) + A(h) + \varphi(h)$$

mit

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

#### 4.24 Definition. Totale Differenzierbarkeit

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  eine Abbildung. Es heißt  $f$  in einem Punkt  $x \in G$  **total differenzierbar** (oder einfach nur **differenzierbar**), wenn es eine lineare Abbildung

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

gibt und ein  $\delta > 0$  mit  $B_\delta(x) \subset G$ , so dass für  $h \in B_\delta(0)$  und

$$\varphi(h) := f(x+h) - f(x) - Ah$$

gilt

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

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

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

Mit dieser Schreibweise ist also dann  $f : G \rightarrow \mathbb{R}^m$  differenzierbar in  $x$  wenn sich  $f$  in  $x$  bis auf einen Fehler der Ordnung  $o(\|h\|)$  linear approximieren läßt, also wenn

$$f(x+h) = f(x) + Ah + o(\|h\|).$$

(b) **Zur Erinnerung:** Eine lineare Abbildung  $A : V \rightarrow W$  zwischen Vektorräumen  $V$  und  $W$  wird nach Wahl von Basen in  $V$  und  $W$  durch eine  $(m \times n)$ -Matrix  $(a_{ij})$  beschrieben.

Wir werden im Folgenden die Wahl der Basis nur dann explizit machen, wenn wir **nicht** die kanonische Basis  $(e_1, \dots, e_n)$  von  $\mathbb{R}^n$  bzw.  $(e_1, \dots, e_m)$  von  $\mathbb{R}^m$  zugrunde legen.

(c) Natürlich sagen wir, dass  $f : G \rightarrow \mathbb{R}^m$  **total differenzierbar** ist, wenn  $f$  in allen Punkten  $x \in G$  total differenzierbar ist.

#### 4.26 Beispiel. Quadratische Formen

Sei  $C = (c_{ij}) \in M(n \times n, \mathbb{R})$  eine symmetrische  $n \times n$ -Matrix und

$$\begin{aligned} f : \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto f(x) = \langle x, Cx \rangle = \sum_{i,j=1}^n c_{ij} x_i x_j \end{aligned}$$

die zugehörige quadratische Form. Für  $x, h \in \mathbb{R}^n$  gilt

$$\begin{aligned} f(x+h) &= \langle x+h, Cx+Ch \rangle = \\ &= \langle x, Cx \rangle + \langle h, Cx \rangle + \langle x, Ch \rangle + \langle h, Ch \rangle \\ &= \langle x, Cx \rangle + 2\langle Cx, h \rangle + \langle h, Ch \rangle \\ &= f(x) + Ah + \varphi(h) \end{aligned}$$

mit  $A := 2Cx \in M(1 \times n)$  und  $\varphi(h) = \langle h, Ch \rangle$ .

Da  $|\varphi(h)| \leq \|h\| \cdot \|Ch\| \leq \|C\| \cdot \|h\|^2$ , gilt  $\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|} = 0$ . Also ist  $f$  in  $x$  differenzierbar.

#### 4.27 Definition. Die Norm einer linearen Abbildung

Die **Norm** einer linearen Abbildung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  definiert man durch

$$\|T\| := \sup \{ \|Tx\|_{\mathbb{R}^m} \mid x \in \mathbb{R}^n \text{ mit } \|x\|_{\mathbb{R}^n} \leq 1 \}.$$

**4.28 Bemerkung.** (a) Es gilt  $\|T\| < \infty$ , da  $f : x \mapsto \|Tx\|_{\mathbb{R}^m}$  stetig ist ( $f = \|\cdot\| \circ T$ ) und  $B_1(0) \subset \mathbb{R}^n$  kompakt, also  $f$  als stetige Abbildung auf einem Kompaktum beschränkt ist (vgl. Satz 3.11).

(b) Es gilt für beliebiges  $x \in \mathbb{R}^n$ , dass

$$\|Tx\|_{\mathbb{R}^m} = \|x\|_{\mathbb{R}^n} \|T \frac{x}{\|x\|}\|_{\mathbb{R}^m} \leq \|T\| \|x\|_{\mathbb{R}^n}.$$

**4.29 Satz. und Definition. Stetigkeit diffbarer Funktionen und die Jacobi-Matrix**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  eine Abbildung, die im Punkt  $x \in G$  differenzierbar sei, also

$$f(x+h) = f(x) + Ah + o(\|h\|)$$

mit der Matrix  $A = (a_{ij}) \in M(m \times n, \mathbb{R})$ .

Dann gilt:

(a)  $f$  ist im Punkt  $x$  stetig.

(b) Alle Komponenten  $f_i : G \rightarrow \mathbb{R}$ ,  $1 \leq i \leq m$ , von  $f$  sind in  $x$  partiell differenzierbar mit

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

Aus (b) folgt insbesondere, dass die Matrix  $A$  durch die differenzierbare Abbildung  $f$  eindeutig bestimmt ist. Man nennt  $A$  das **Differential**, die **Jacobi-Matrix** oder die **Funktionalmatrix** von  $f$  im Punkte  $x$  und schreibt:

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

Das Differential  $Df(x)$  ist also die lineare Approximation an  $f$  im Punkt  $x$ :

$$f(x+h) = f(x) + Df(x)h + o(\|h\|).$$

*Beweis.* (a)  $\lim_{h \rightarrow 0} f(x+h) = f(x) + \lim_{h \rightarrow 0} (Ah + o(\|h\|)) = f(x)$ .

(b) Für  $i = 1, \dots, m$  und  $h \in \mathbb{R}^n$  ist

$$f_i(x+h) = f_i(x) + \sum_{j=1}^n a_{ij}h_j + o(\|h\|),$$

also für  $k \in \mathbb{R}$  und  $h = ke_j$

$$f_i(x+ke_j) = f_i(x) + ka_{ij} + o(\|ke_j\|).$$

Damit folgt für die partielle Ableitung von  $f_i$  in Richtung  $e_j$

$$\frac{\partial f_i}{\partial x_j}(x) := \lim_{k \rightarrow 0} \frac{f_i(x+ke_j) - f_i(x)}{k} = a_{ij} + \lim_{k \rightarrow 0} \frac{o(k)}{k} = a_{ij}.$$

□

**4.30 Satz. Stetig partiell differenzierbar  $\Rightarrow$  total differenzierbar**

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

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

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

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

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

wobei

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

Daraus folgt

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

Setzt man  $a_i = \partial_i f(x)$ , so gilt

$$f(x+h) = f(x) + \sum_{i=1}^n a_i h_i + \varphi(h)$$

mit

$$\varphi(h) = \sum_{i=1}^n (\partial_i f(y^{(i)}) - a_i) h_i.$$

Wegen der Stetigkeit von  $\partial_i f$  in  $x$  gilt

$$\lim_{h \rightarrow 0} (\partial_i f(y^{(i)}) - a_i) = \partial_i f(\lim_{h \rightarrow 0} y^{(i)}) - a_i = \partial_i f(x) - a_i = 0,$$

also

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

□

**4.31 Korollar.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  stetig partiell differenzierbar. Dann ist  $f$  total differenzierbar und somit stetig.

**4.32 Merke.** Es gelten also die Implikationen:

$$\begin{array}{c} \text{stetig partiell differenzierbar} \Rightarrow \text{total differenzierbar} \Rightarrow \text{partiell differenzierbar} \\ \downarrow \\ \text{stetig} \end{array}$$

Die Umkehrungen gelten im Allgemeinen nicht!

Im Folgenden werden wir oft “stetig partiell differenzierbar” durch “stetig differenzierbar” abkürzen, da nach obigem eine total differenzierbare Funktion mit stetiger Ableitung ja insbesondere stetige partielle Ableitungen hat.

### 4.33 Satz. Kettenregel

Seien  $G \subset \mathbb{R}^n$  und  $H \subset \mathbb{R}^m$  Gebiete und  $g : G \rightarrow \mathbb{R}^m$  und  $f : H \rightarrow \mathbb{R}^k$  Abbildungen mit  $g(G) \subset H$ . Die Abbildung  $g$  sei im Punkt  $x \in G$  differenzierbar und die Abbildung  $f$  sei im Punkt  $y := g(x)$  differenzierbar. Dann ist die Komposition

$$f \circ g : G \rightarrow \mathbb{R}^k$$

im Punkt  $x$  differenzierbar und für ihr Differential gilt:

$$D(f \circ g)(x) = \underbrace{Df(g(x))}_{k \times m\text{-Matrix}} \cdot \underbrace{Dg(x)}_{m \times n\text{-Matrix}}$$

#### 4 Differenzierbarkeit

*Beweis.* Sei  $A := Dg(x)$  und  $B = Df(y)$ . Es ist zu zeigen, dass  $D(f \circ g)(x) = BA$ . Nach Voraussetzung gelten

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

mit  $\varphi(h) = o(\|h\|)$  und  $\psi(\eta) = o(\|\eta\|)$ . Wählt man

$$\eta := g(x+h) - g(x) = Ah + \varphi(h)$$

so ergibt sich

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

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

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

also

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

□

#### 4.34 Beispiel. Der Laplaceoperator in Polarkoordinaten

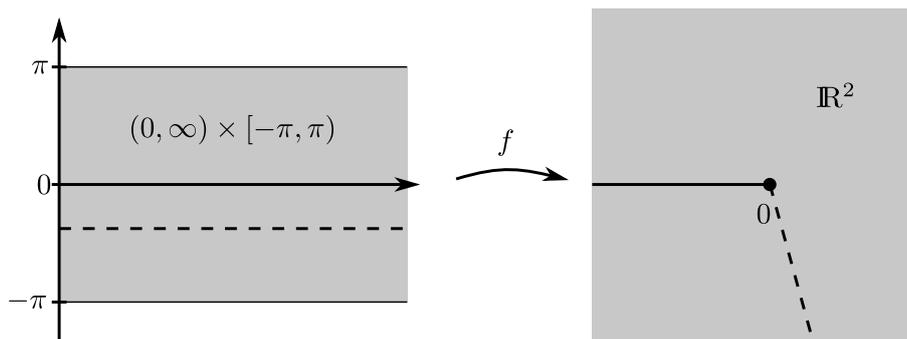
Die Abbildung

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

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

$$(\Delta u) \circ f = \frac{\partial^2(u \circ f)}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2(u \circ f)}{\partial \varphi^2} + \frac{1}{r} \frac{\partial(u \circ f)}{\partial r}. \quad (4.1)$$

*Beweis.* Nachrechnen mit Kettenregel (Übungsaufgabe). □



Aber wie kommt man auf den Ausdruck (4.1)? Dazu betrachten wir zunächst das Differential von  $u$  in Polarkoordinaten. Mit der Kettenregel gilt

$$D(u \circ f) = Du \circ f \cdot Df, \quad \text{also} \quad Du \circ f = D(u \circ f) \cdot (Df)^{-1}, \quad (4.2)$$

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

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

Wir haben also in (4.2) eine allgemeine Formel für das Differential  $Du$  von  $u$  ausgedrückt in Koordinaten geben durch  $f$ . Dabei benötigen wir nur die Invertierbarkeit von  $Df$ . In unserem Beispiel ergibt sich

$$\begin{aligned} Du \circ f &= \left( \frac{\partial(u \circ f)}{\partial r}, \frac{\partial(u \circ f)}{\partial \varphi} \right) \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix} \\ &= \left( \cos \varphi \frac{\partial(u \circ f)}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial(u \circ f)}{\partial \varphi}, \sin \varphi \frac{\partial(u \circ f)}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial(u \circ f)}{\partial \varphi} \right) \\ &=: \underbrace{\left( \cos \varphi \frac{\partial u}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial u}{\partial \varphi} \right)}_{\frac{\partial u}{\partial x} \circ f}, \underbrace{\left( \sin \varphi \frac{\partial u}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial u}{\partial \varphi} \right)}_{\frac{\partial u}{\partial y} \circ f} = \text{grad } u \circ f, \quad (4.3) \end{aligned}$$

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

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

und analog

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

Addiert man die beiden Ausdrücke, so ergibt sich (4.1).

Nochmals zurück zu (4.3): Die Matrix  $Du \circ f$  ist die Linearisierung der Abbildung  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$  im Punkte  $f(r, \varphi)$  bezüglich der kanonischen Basis des  $\mathbb{R}^2$ . Man kann nun aber auch noch versuchen, die lineare Abbildung  $Du \circ f$  bezüglich einer den Koordinaten angepassten Basis

#### 4 Differenzierbarkeit

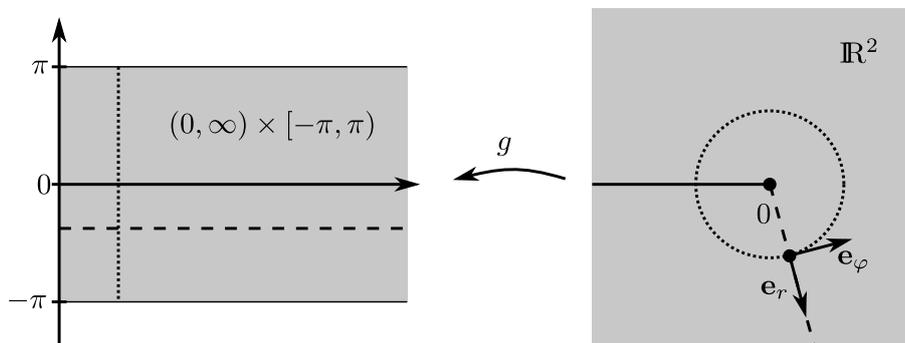
darzustellen, nämlich bezüglich der aus der Physik bekannten Basisvektoren  $\mathbf{e}_r$  und  $\mathbf{e}_\varphi$ . Um diese vernünftig zu definieren, schränken wir zunächst  $f$  auf  $(0, \infty) \times (-\pi, \pi)$  ein und erhalten so einen **Diffeomorphismus**

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

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

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

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



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

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

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

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

wobei  $E$  wie immer die Einheitsmatrix bezeichnet. Also ist

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

was wir für unser Beispiel oben schon berechnet haben. Der Basiswechsel von der neuen Basis in die kanonische ist durch  $(Dg)^T = ((Df)^{-1})^T$  gegeben. Multiplikation von (4.3) durch  $((Df)^{-1})^T$  von rechts (nur im Urbildraum wird die Basis gewechselt) liefert das Differential  $Du$  als Matrix bezüglich der durch  $f$  beschriebenen Basisvektoren

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

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

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

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

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

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

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

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

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

#### 4.35 Definition. Richtungsableitung

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^m$  eine Funktion. Unter der Richtungsableitung von  $f$  im Punkt  $x \in G$  in Richtung  $v \in \mathbb{R}^n$ , versteht man den Differenzialquotienten

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

falls dieser existiert. Für  $v = e_j$  ist  $\partial_v f$  also gleich der  $j$ -ten partiellen Ableitung  $\partial_j f$ .

#### 4.36 Satz. Richtungsableitung und Gradient

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}$  stetig differenzierbar. Dann gilt für jedes  $x \in G$  und jeden Vektor  $v \in \mathbb{R}^n$

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

*Beweis.* Sei  $g : (-\delta, \delta) \rightarrow G$  definiert durch  $g(t) := x + tv$  und  $h := f \circ g : (-\delta, \delta) \rightarrow \mathbb{R}$ . Nach Definition der Richtungsableitung ist

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

Aus der Kettenregel folgt

$$\begin{aligned} \frac{d}{dt} h(t) &= \underbrace{Df(g(t))}_{1 \times n} \cdot \underbrace{Dg(t)}_{n \times 1} = \sum_{j=1}^n \underbrace{\frac{\partial f}{\partial x_j}(g(t))}_{(\text{grad } f)_j} \underbrace{\frac{dg_j}{dt}(t)}_{=v_j} \\ &= \langle v, \text{grad } f(g(t)) \rangle, \end{aligned}$$

also  $\frac{d}{dt} h(0) = \langle v, \text{grad } f(x) \rangle$ . □

**4.37 Bemerkung.** (a) Ist  $\nabla f(x) \neq 0$ , so ist der Winkel  $\theta$  zwischen den Vektoren  $v$  und  $\nabla f(x)$  definiert und es gilt

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

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

(b) Für stetig differenzierbares  $f : G \rightarrow \mathbb{R}^m$  ist die Richtungsableitung somit gegeben durch

$$\partial_v f(x) = \begin{pmatrix} \partial_v f_1(x) \\ \vdots \\ \partial_v f_m(x) \end{pmatrix} = \begin{pmatrix} \nabla f_1(x) \\ \vdots \\ \nabla f_m(x) \end{pmatrix} \cdot v = Df(x) \cdot v,$$

wobei das Produkt das Matrix-Vektor Produkt ist.

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

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

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

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

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

**4.39 Satz. Mittelwertsatz der Differentialrechnung**

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

$$f(x + h) - f(x) = \left( \int_0^1 Df(x + th) \, dt \right) \cdot h$$

**4.40 Bemerkung.** Man beachte, dass man für  $m \geq 2$  das Integral  $\int_0^1 Df(x + th) \cdot h \, dt$  im Allgemeinen **nicht** durch den Wert des Integranden an einer Zwischenstelle  $x + \theta h$  ersetzen kann.

*Beweis. des Mittelwertsatzes.* Setzen wir  $\alpha : [0, 1] \rightarrow G$ ,  $\alpha(t) = x + th$ , so ist nach dem Hauptsatz angewendet auf  $(f_i \circ \alpha) : (0, 1) \rightarrow \mathbb{R}$

$$\begin{aligned} f(x + h) - f(x) &= (f \circ \alpha)(1) - (f \circ \alpha)(0) = \int_0^1 \frac{d}{dt} (f \circ \alpha)(t) \, dt \\ &= \int_0^1 Df(\alpha(t)) \cdot \underbrace{D\alpha(t)}_{=h} \, dt = \int_0^1 Df(x + th) \, dt \cdot h. \end{aligned}$$

□

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

**4.42 Korollar. Schrankensatz**

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

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

gilt

$$\|f(y) - f(x)\| \leq M \underbrace{\int_0^1 \left\| \frac{d\alpha}{dt}(t) \right\| dt}_{\text{Länge des Weges } \alpha}.$$

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

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

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

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

□

*Beweis. des Schrankensatzes.* Mit  $D\alpha(t) = \frac{d\alpha}{dt}(t)$  ist

$$\begin{aligned} \|f(y) - f(x)\| &= \left\| \int_0^1 Df(\alpha(t)) \cdot D\alpha(t) dt \right\| \leq \int_0^1 \|Df(\alpha(t)) \cdot D\alpha(t)\| dt \\ &\leq M \int_0^1 \|D\alpha(t)\| dt = M \int_0^1 \left\| \frac{d\alpha}{dt}(t) \right\| dt. \end{aligned}$$

□

**4.44 Korollar.** Sei  $G$  ein Gebiet,  $f : G \rightarrow \mathbb{R}^m$  partiell differenzierbar und  $Df(x) = 0$  für alle  $x \in G$ . Dann ist  $f$  auf  $G$  konstant.

*Beweis.* Seien  $x, y \in G$  beliebig und  $\alpha : [0, 1] \rightarrow G$  ein Weg mit  $\alpha(0) = x$  und  $\alpha(1) = y$ . Wenn  $\alpha$  differenzierbar ist, so folgt mit dem Schrankensatz sofort

$$\|f(y) - f(x)\| \leq M \int_0^1 \left\| \frac{d\alpha}{dt}(t) \right\| dt = 0,$$

da wegen  $Df \equiv 0$  auch  $M = 0$  ist. Also  $f(x) = f(y)$ . Nun können wir aber im Allgemeinen nur einen stetigen Weg von  $x$  nach  $y$  finden. Das folgende Argument zeigt, dass wir diesen stetigen Weg zumindest lokal durch einen differenzierbaren (nämlich durch eine Strecke) ersetzen können. Setze  $t_0 = \sup\{t \in [0, 1] \mid f(\alpha(\tau)) = f(x) \forall \tau \in [0, t]\}$ . Wegen der Stetigkeit von  $f$  ist dann auch  $f(\alpha(t_0)) = f(x)$ . Angenommen  $t_0 < 1$ , dann gibt es ein  $\delta > 0$  mit  $B_\delta(\alpha(t_0)) \subset G$  und ein  $0 < \tilde{\delta} < \delta$ , so dass auch  $\alpha(t_0 + \tilde{\delta}) \in B_\delta(\alpha(t_0))$  liegt. Da für  $h := \alpha(t_0 + \tilde{\delta}) - \alpha(t_0)$  die ganze Strecke  $[0, 1] \ni s \mapsto \alpha(t_0) + hs \subset B_\delta(\alpha(t_0)) \subset G$  liegt, gilt nach dem Schrankensatz

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

mit

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

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



## 5 Taylorformel und lokale Extrema

Wir haben im ersten Semester gezeigt, dass für eine zweimal stetig differenzierbare Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  die Taylorentwicklung

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

eine lokale quadratische Approximation an die Funktion  $f$  liefert. Für Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  ist die erste Ableitung  $Df(x) = \text{grad}f(x)^T$  ein Zeilenvektor und die zweite Ableitung  $\text{Hess}f(x)$  eine Matrix. Die naheliegende Verallgemeinerung der Taylorschen Formel für solche  $f$  ist also

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

### 5.1 Notation. Multiindices und iterierte Richtungsableitung

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

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

$$\begin{aligned}
 |\alpha| &= \alpha_1 + \dots + \alpha_n = \sum_{j=1}^n \alpha_j \\
 \alpha! &= \alpha_1! \cdot \dots \cdot \alpha_n! = \prod_{j=1}^n \alpha_j! ,
 \end{aligned}$$

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

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

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

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

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

$$(h \cdot \nabla)^k f(x) := \partial_h^k f(x) := \sum_{j_1=1}^n \dots \sum_{j_k=1}^n h_{j_1} \dots h_{j_k} \partial_{j_1} \dots \partial_{j_k} f(x) .$$

## 5 Taylorformel und lokale Extrema

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

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

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

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

*Beweis. von Lemma 5.2.* Nach der Kettenregel ist

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

Nochmals die Kettenregel liefert

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

und  $k$ -malige Anwendung schlieelich

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

Da die Reihenfolge der Differentiationen aber gema Sats 4.16 keine Rolle spielt, fassen wir Terme, in denen  $\alpha_1$ -mal nach der ersten Koordinate,  $\alpha_2$ -mal nach der zweiten Koordinate etc. abgeleitet wird, zusammen. Es gibt nun nach Bemerkung 5.3 gerade  $\frac{k!}{\alpha_1! \dots \alpha_n!}$  solche  $k$ -Tupel  $(i_1, \dots, i_k)$  in denen  $\alpha_1$ -mal der Wert 1,  $\alpha_2$ -mal der Wert 2,  $\dots$ , und  $\alpha_n$ -mal der Wert  $n$  vorkommt. Durch Zusammenfassen der Summanden ergibt sich also

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

□

### 5.4 Satz. von Taylor

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

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

**5.5 Bemerkung.** Man nennt

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

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

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

so ist

$$\begin{aligned} P_0(h) &= f(x) \\ P_1(h) &= \partial_1 f(x) h_1 + \dots + \partial_n f(x) h_n = \langle \text{grad} f(x), h \rangle \\ P_2(h) &= \sum_{\alpha_1 + \alpha_2 = 2} \frac{1}{\alpha_1! \alpha_2!} \partial_1^{\alpha_1} \partial_2^{\alpha_2} f(x) h_1^{\alpha_1} h_2^{\alpha_2} \\ &= \frac{1}{2} \left( \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j \right) = \frac{1}{2} \langle h, \text{Hess} f(x) h \rangle. \end{aligned}$$

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

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

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

Nach Lemma 5.2 erhält man

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

□

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

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

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

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

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

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

## 5 Taylorformel und lokale Extrema

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

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

und deshalb gilt

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

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

$$f(x + h) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \varphi(h)$$

mit  $\varphi(h) = o(\|h\|^k)$ . □

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

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

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

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

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

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

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

## 5.8 Definition. Lokale Extrema

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

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

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

**5.9 Proposition.** Sei  $f : G \rightarrow \mathbb{R}$ ,  $x \in G$  und  $f$  habe in  $x$  ein lokales Extremum. Falls  $f$  in  $x$  partiell differenzierbar ist, so gilt

$$\text{grad} f(x) = 0.$$

*Beweis.* Definiert man für  $\delta > 0$  klein genug die Funktionen  $h_i : (-\delta, \delta) \rightarrow \mathbb{R}$  durch

$$h_i(t) = f(x + te_i),$$

so ist

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

Mit  $f$  hat aber auch  $h_i$  in 0 ein lokales Extremum, also gilt nach Mathematik für Physiker I  $\frac{d}{dt} h_i(0) = 0$  für  $i = 1, \dots, n$ . Somit ist auch  $\text{grad} f(x) = 0$ . □

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

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

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

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

**5.11 Satz.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f \in C^2(G, \mathbb{R})$ . Sei  $x \in G$  eine Nullstelle von  $\text{grad} f$ , also  $\text{grad} f(x) = 0$ .

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

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

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

*Beweis.* Stelle  $x$  in einer Orthonormalbasis aus Eigenvektoren dar. □

*Beweis. von Satz 5.11.* (a) Sei  $\delta > 0$  so klein, dass  $B_\delta(x) \subset G$  ist. Nach Korollar 5.6 gilt dann für  $h \in B_\delta(0)$

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

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

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

Also ist  $f(x)$  ein striktes lokales Minimum.

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

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

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

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

also

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

## 5 Taylorformel und lokale Extrema

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

**5.13 Bemerkung.** Man nennt die Nullstellen von  $\text{grad}f$  auch die **kritischen Punkte** von  $f$ .

### 5.14 Beispiel. Minimum und Sattel

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

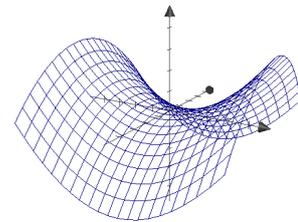
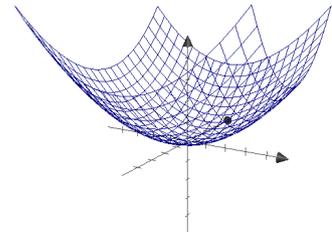
$$\text{Hess}f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \text{Hess}f(0, 0)$$

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

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

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

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



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

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

Wegen Satz 5.11 (c) ist aber die Semidefinitheit ein notwendiges Kriterium für das Vorliegen eines Extremums.

## 6 Implizite Funktionen

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

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

Hierbei wäre  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto F(x, y)$  eine „explizit“ gegebene Funktion. Man möchte dann die implizite Gleichung  $F(x, y) = 0$  explizit machen, d.h. „nach  $y$  auflösen“ und in der Form  $y = g(x)$  schreiben.

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

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

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

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

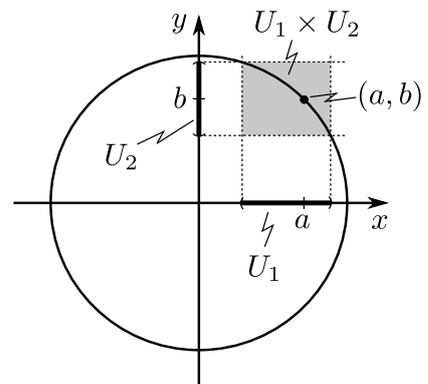
Gibt man allerdings  $(a, b) \in C$  mit  $b \neq 0$  vor, so kann man  $F(x, y)$  lokal um  $(a, b)$  nach  $y$  auflösen, d.h. es gibt eine Umgebung  $U_1 \subset \mathbb{R}$  von  $a$  und eine Umgebung  $U_2$  von  $b$  und eine Funktion  $g : U_1 \rightarrow U_2$  so, dass für alle  $(x, y) \in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

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

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

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

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



Der folgende Satz über implizite Funktionen gibt eine hinreichende Bedingung dafür an, dass man eine implizite Gleichung  $F(x, y) = 0$  lokal um einen Punkt  $(a, b)$  mit  $F(a, b) = 0$  in der Form  $y = g(x)$  explizit machen kann.

### 6.2 Satz. Satz über implizite Funktionen

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

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

## 6 Implizite Funktionen

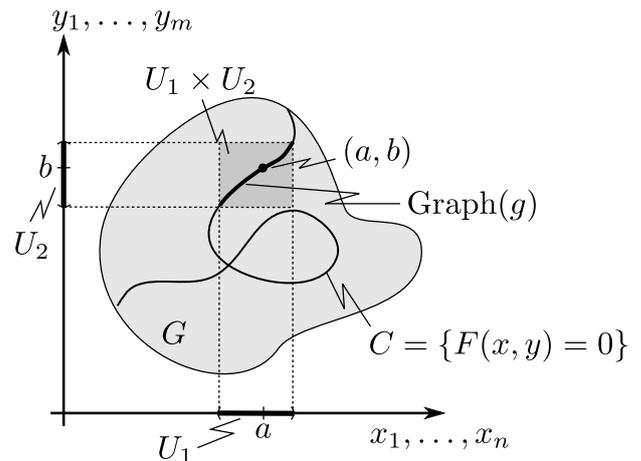
invertierbar ist.

Dann existieren offene Umgebungen  $U_1 \subset \mathbb{R}^n$  von  $a$  und  $U_2 \subset \mathbb{R}^m$  von  $b$  mit  $U_1 \times U_2 \subset G$  und eine stetig partiell differenzierbare Funktion  $g : U_1 \rightarrow U_2$ , sodass für alle  $(x, y) \in U_1 \times U_2$  gilt:

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

Das Differential von  $g$  ist

$$Dg(x) = -[\partial_y F(x, g(x))]^{-1} \partial_x F(x, g(x)).$$



**6.3 Bemerkung.** Die Bedingung  $\det\left(\frac{\partial F}{\partial y}(a, b)\right) \neq 0$  ist gleichbedeutend damit, dass keine Richtungsableitung von  $F$  in „ $y$ -Richtung“ verschwindet, also  $\vec{v} \cdot \frac{\partial F}{\partial y}(a, b) \neq 0$  für alle  $\vec{v} \in \mathbb{R}^m \setminus \{0\}$ . Dadurch wird sichergestellt, dass es eine Umgebung  $U_2$  von  $b$  gibt, so dass für  $y \in U_2$

$$F(a, y) = 0 \quad \Leftrightarrow \quad y = b.$$

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

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

Zu (a): Betrachte die Funktion  $\Phi : G \rightarrow \mathbb{R}^m$  gegeben durch

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

wobei  $B = \frac{\partial F}{\partial y}(0, 0)$  und o.B.d.A  $(a, b) = (0, 0)$  ist. Dann gilt

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

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

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

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

gilt. Daraus folgt nun wieder, dass

$$\Phi(x, \cdot) : \overline{B}_{r_2}(0) \rightarrow \overline{B}_{r_2}(0)$$

für jedes feste  $x$  in einer geeigneten Kugel  $\overline{B}_{r_3}(0) \subset \overline{B}_{r_1}(0)$  eine Kontraktion ist: für  $y \in \overline{B}_{r_2}(0)$  ist

$$\|\Phi(x, y)\| \leq \|\Phi(x, y) - \Phi(x, 0)\| + \|\Phi(x, 0)\| \leq \frac{1}{2} \|y\| + \frac{1}{2} r_2 \leq r_2,$$

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

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

$$A := \{h \in C(\overline{B_{r_3}}, \mathbb{R}^m) \mid \text{Bild}(h) \subset \overline{B_{r_2}}, h(0) = 0\}.$$

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

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

ist dann wieder eine Kontraktion, denn

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

Der Fixpunkt von  $\Psi$  ist

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

Also ist  $g \in A$  und somit stetig.

Zu (c): Da  $F$  stetig differenzierbar ist, gilt mit dem Mittelwertsatz für  $j = 1, \dots, m$

$$\begin{aligned} 0 &= F_j(x, g(x)) - F_j(x_0, g(x_0)) = F_j(x, g(x)) - F_j(x_0, g(x)) + F_j(x_0, g(x)) - F_j(x_0, g(x_0)) \\ &= \left( \int_0^1 \partial_x F_j(x_0 + t(x - x_0), g(x)) \cdot (x - x_0) dt \right) \\ &\quad + \left( \int_0^1 \partial_y F_j(x_0, g(x_0) + t(g(x) - g(x_0))) \cdot (g(x) - g(x_0)) dt \right) \\ &= \partial_x F_j(\tilde{x}_j, g(x)) \cdot (x - x_0) + \partial_y F_j(x_0, \tilde{y}_j) \cdot (g(x) - g(x_0)) \end{aligned}$$

für geeignete Zwischenstellen  $(\tilde{x}_j, \tilde{y}_j) \in G$ . Also ist

$$\begin{aligned} 0 &= \begin{pmatrix} \partial_x F_1(\tilde{x}_1, g(x)) \\ \vdots \\ \partial_x F_m(\tilde{x}_m, g(x)) \end{pmatrix} (x - x_0) + \begin{pmatrix} \partial_y F_1(x_0, \tilde{y}_1) \\ \vdots \\ \partial_y F_m(x_0, \tilde{y}_m) \end{pmatrix} (g(x) - g(x_0)) \\ &=: A(x) \cdot (x - x_0) + B(x) \cdot (g(x) - g(x_0)). \end{aligned}$$

Nun ist  $\partial_y F(0, 0)$  nach Annahme invertierbar, also  $\partial_y F_1(0, 0), \dots, \partial_y F_m(0, 0)$  linear unabhängig. Diese lineare Unabhängigkeit bleibt aber wegen der Stetigkeit der  $\partial_y F_j(x, y)$  in einer hinreichend kleinen Umgebung von  $(0, 0)$  erhalten und für eine hinreichend kleine Umgebung von 0 ist  $B(x)^{-1}$  gleichmäßig beschränkt. Wählen wir  $x_0$  in dieser Umgebung, so gilt

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

Aufgrund der Stetigkeit der partiellen Ableitungen von  $F$  gilt

$$\lim_{x \rightarrow x_0} B^{-1}(x)A(x) = B^{-1}(x_0)A(x_0),$$

also

$$Dg(x_0) = -\partial_y F(x_0, g(x_0))^{-1} \partial_x F(x_0, g(x_0)),$$

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

## 6 Implizite Funktionen

**6.4 Beispiel.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}$  stetig differenzierbar. Wir betrachten die Niveaufäche

$$N_c = \{x \in \mathbb{R}^n \mid f(x) = c\}$$

zu  $c \in \mathbb{R}$  von  $f$ . Ist nun  $\text{grad}(f)(a) \neq 0$  für  $a \in N_c$ , so behaupten wir, dass  $N_c$  bei  $a$  lokal aussieht wie der Graph einer stetig differenzierbaren Funktion  $g$  in  $(n-1)$  Veränderlichen. Sei z.B.  $\frac{\partial f}{\partial x_n}(a) \neq 0$  und  $\bar{x} = (x_1, \dots, x_{n-1})$  und  $y = x_n$ . Dann ist

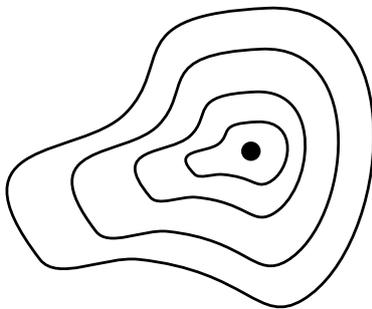
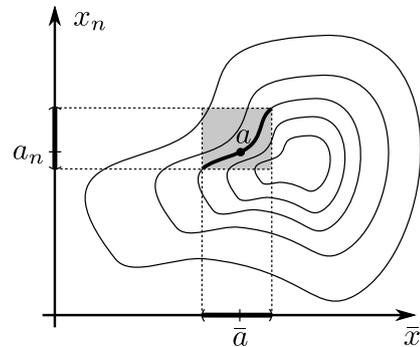
$$x \in N_c \Leftrightarrow \underbrace{f(\bar{x}, y) - c}_{=: F(\bar{x}, y)} = 0$$

und

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

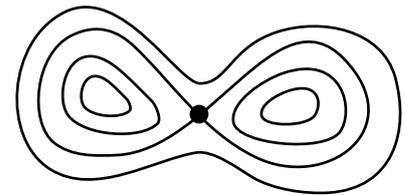
Nach dem Satz über implizite Funktionen existiert eine Umgebung  $\bar{U} \times U_n$  von  $a$  und eine  $C^1$  Funktion  $g : \bar{U} \rightarrow U_n$  mit  $F(\bar{x}, y) = 0 \Leftrightarrow g(\bar{x}) = y$ . Somit ist  $N_c \cap (\bar{U} \times U_n) = \text{Graph}(g) \cap (\bar{U} \times U_n)$ .

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



← striktes Extremum

Sattelpunkt →



Wir kommen nun zu der Frage nach der lokalen Umkehrbarkeit einer stetig differenzierbaren Funktion  $f : \mathbb{R}^n \supset G \rightarrow D \subset \mathbb{R}^n$ . Insbesondere ist es oft wichtig, z.B. bei Koordinationstransformationen, dass auch  $g = f^{-1} : D \rightarrow G$  wieder stetig differenzierbar ist.

### 6.5 Definition. Diffeomorphismus

Seien  $G, D \subset \mathbb{R}^n$  Gebiete. Eine stetig differenzierbare Abbildung  $f : G \rightarrow D$  heißt **Diffeomorphismus**, wenn  $f$  bijektiv und  $f^{-1} : D \rightarrow G$  stetig differenzierbar ist.

**6.6 Bemerkung.** Ist  $f : G \rightarrow D$  ein Diffeomorphismus, so ist für jedes  $x \in G$  die Ableitung  $Df(x)$  invertierbar und es gilt mit  $g = f^{-1}$ , dass  $(Df(x))^{-1} = Dg(f(x))$ .

*Beweis.* Wegen  $g \circ f = \text{id}$  liefert die Kettenregel

$$E = D(\text{id}) = D(g \circ f) = Dg \circ f \cdot Df.$$

□

**6.7 Bemerkung.** Nicht jede stetig differenzierbare Bijektion hat auch eine stetig differenzierbare Umkehrung:  $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3$ , ist bijektiv, stetig differenzierbar, die Umkehrabbildung  $g : \mathbb{R} \rightarrow \mathbb{R}, y \mapsto \sqrt[3]{y}$  ist aber im Nullpunkt nicht differenzierbar. Beachte, dass die nach Bemerkung 6.6 notwendige Bedingung  $f' \neq 0$  für  $x = 0$  nicht erfüllt ist, und somit  $f$  kein Diffeomorphismus sein kann.

**6.8 Satz. Satz über die Umkehrabbildung**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $f : G \rightarrow \mathbb{R}^n$  eine stetig differenzierbare Funktion. Ist nun  $a \in G$  so, dass  $Df(a) \in M(n \times n, \mathbb{R})$  invertierbar ist, so existiert eine offene Umgebung  $U \subset G$  von  $a$  so, dass  $f|_U : U \rightarrow V$  mit  $V = f(U)$  ein Diffeomorphismus ist.

*Beweis.* Sei  $b := f(a)$ . Man möchte die Gleichung  $y = f(x)$  in einer Umgebung von  $(x, y) = (a, b)$  nach  $x$  auflösen, also  $x = g(y)$  schreiben. Dazu betrachtet man die Abbildung

$$F : G \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$F(x, y) = y - f(x).$$

Es ist

$$\frac{\partial F}{\partial x}(a, b) = -Df(a)$$

nach Voraussetzung invertierbar. Deshalb liefert der Satz über implizite Funktionen Umgebungen  $\tilde{U} \subset G$  von  $a$  und  $V \subset \mathbb{R}^n$  von  $b$  und ein stetig differenzierbares  $g : V \rightarrow \tilde{U}$ , so dass für alle  $(x, y) \in \tilde{U} \times V$  gilt

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

Setzt man noch  $U := g(V)$ , so gilt für alle  $(x, y) \in U \times V$

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

Also ist  $f|_U : U \rightarrow V$  bijektiv,  $g = f^{-1}$  und  $g$  stetig differenzierbar und somit  $f|_U$  ein Diffeomorphismus. □

**6.9 Beispiel. Kugelkoordinaten**

Kugelkoordinaten des  $\mathbb{R}^3$  sind gegeben durch

$$f : (0, \infty) \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^3, \quad (r, \vartheta, \varphi) \mapsto (r \sin \vartheta \cos \varphi, r \sin \vartheta \sin \varphi, r \cos \vartheta)$$

Eingeschränkt auf

$$G = (0, \infty) \times (0, \pi) \times (-\pi, \pi)$$

ist  $f$  injektiv und für die Funktionaldeterminante  $\det(Df) : G \rightarrow \mathbb{R}$  gilt

$$\det(Df) = \begin{vmatrix} \sin \vartheta \cos \varphi & r \cos \vartheta \cos \varphi & -r \sin \vartheta \sin \varphi \\ \sin \vartheta \sin \varphi & r \cos \vartheta \sin \varphi & r \sin \vartheta \cos \varphi \\ \cos \vartheta & -r \sin \vartheta & 0 \end{vmatrix} = r^2 \sin \vartheta$$

für alle  $(r, \vartheta, \varphi) \in G$ . Deshalb ist  $f$  ein lokaler Diffeomorphismus. Da  $f$  bijektiv ist, ist es aber auch global ein Diffeomorphismus auf sein Bild  $D := f(G) \subset \mathbb{R}^3$ .

Häufig sucht man Extrema von Funktionen  $f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}$  unter einer Nebenbedingung, die man durch  $h(x) = 0$  für eine geeignete Funktion  $h : G \rightarrow \mathbb{R}$  beschreiben kann. Beispielsweise sucht man das Maximum einer Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  auf der Einheitskreislinie in  $\mathbb{R}^2$ , also unter der Nebenbedingung  $h(x) = \|x\| - 1 = 0$ .

**6.10 Definition. Extrema unter Nebenbedingungen**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und seien  $f, h : G \rightarrow \mathbb{R}$  stetig differenzierbar. Sei  $M = \{x \in G \mid h(x) = 0\}$  und  $a \in M$ . Man sagt,  $f$  habe bei  $a$  ein **lokales Maximum (bzw. Minimum) unter Nebenbedingung**  $h = 0$ , wenn es eine offene Umgebung  $U \subset G$  von  $a$  gibt, sodass für alle  $x \in U \cap M$  gilt

$$f(x) \leq f(a) \quad (\text{bzw. } f(x) \geq f(a)).$$

Der Satz über implizite Funktionen liefert ein notwendiges Kriterium für das Vorliegen lokaler Extrema unter Nebenbedingungen.

**6.11 Satz. Satz über Extrema unter Nebenbedingungen**

Seien  $f, h : G \rightarrow \mathbb{R}$  stetig differenzierbar,  $G \subset \mathbb{R}^n$  ein Gebiet und  $a \in M = \{x \in G \mid h(x) = 0\}$ . Es habe  $f$  ein lokales Extremum unter der Nebenbedingung  $h = 0$  in  $a$  **und** es sei  $\text{grad}h(a) \neq 0$ . Dann gibt es ein  $\lambda \in \mathbb{R}$  mit

$$\text{grad}f(a) = \lambda \text{grad}h(a),$$

**6.12 Bemerkung. Geometrische Bedeutung**

Der Gradient von  $h$  in  $a$  steht senkrecht auf der Niveauläche  $M$  von  $h$ . Damit  $f$  auf  $M$  ein lokales Extremum hat, müssen nur die Richtungsableitungen  $\langle v, \nabla f(a) \rangle$  von  $f$  tangential an  $M$  verschwinden, also

$$\langle v, \nabla f(a) \rangle = 0 \quad \text{falls} \quad \langle v, \nabla h(a) \rangle = 0,$$

d.h. der Gradient von  $f$  muss auf  $M$  senkrecht stehen. Da  $M$  Kodimension 1 hat, folgt daraus

$$\nabla f(a) \parallel \nabla h(a).$$

beachte, dass Satz 6.11 das Analogon zu Proposition 5.9 für den Fall ohne Nebenbedingungen ist. Es ist  $\text{grad}f(a) = \lambda \text{grad}h(a)$  eine notwendige, aber keine hinreichende Bedingung für das Vorliegen eines lokalen Extremums unter der Nebenbedingung  $h = 0$ .

*Beweis.* von Satz 6.11. Da  $\text{grad}h(a) \neq 0$  ist, ist wenigstens eine partielle Ableitung von  $h$  in  $a$  von Null verschieden, sagen wir  $\frac{\partial h}{\partial x_n}(a) \neq 0$ . Sei  $\bar{a} := (a_1, \dots, a_{n-1})$ . Der Satz über implizite Funktionen liefert dann Umgebungen  $V \subset \mathbb{R}^{n-1}$  von  $\bar{a}$  und  $I \subset \mathbb{R}$  von  $a_n$  mit  $V \times I \subset G$  und ein stetig differenzierbares  $g : V \rightarrow I$ , so dass für alle  $(\bar{x}, x_n) \in V \times I$  gilt

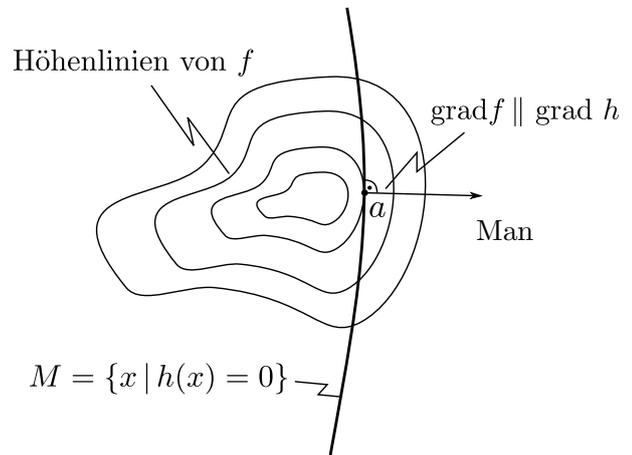
$$h(\bar{x}, x_n) = 0 \quad \Leftrightarrow \quad x_n = g(\bar{x}).$$

Hat nun  $f : G \rightarrow \mathbb{R}$  ein lokales Extremum unter der Nebenbedingung  $h = 0$  in  $a \in M$ , so hat  $f(\bar{x}, g(\bar{x})) =: f \circ \varphi$  ein lokales Extremum in  $\bar{x}$  (ohne Nebenbedingung). Hierbei ist  $\varphi : V \rightarrow M \subset \mathbb{R}^n$ ,  $\varphi(\bar{x}) = (\bar{x}, g(\bar{x}))$ . Es ist also

$$\text{grad}(f \circ \varphi)(\bar{a}) = 0.$$

Für  $i = 1, \dots, n - 1$  ergibt die Kettenregel

$$0 = \frac{\partial (f \circ \varphi)}{\partial x_i}(\bar{a}) = \sum_{j=1}^n \partial_j f(\varphi(\bar{a})) \frac{\partial \varphi_j}{\partial x_i}(\bar{a}) = \partial_i f(a) + \partial_n f(a) \frac{\partial g}{\partial x_i}(\bar{a}). \quad (*)$$



Andererseits gilt wegen  $(h \circ \varphi)(\bar{x}) = 0 \forall \bar{x} \in V$ , dass

$$0 = \frac{\partial(h \circ \varphi)}{\partial x_i}(\bar{a}) = \partial_i h(a) + \partial_n h(a) \frac{\partial g}{\partial x_i}(\bar{a}),$$

also wegen  $\frac{\partial h}{\partial x_n}(a) \neq 0$

$$\frac{\partial g}{\partial x_i}(\bar{a}) = - \left( \frac{\partial h}{\partial x_n}(a) \right)^{-1} \frac{\partial h}{\partial x_i}(a). \quad (**)$$

Setzen wir nun

$$\lambda := \frac{\frac{\partial f}{\partial x_n}(a)}{\frac{\partial h}{\partial x_n}(a)} \in \mathbb{R},$$

so folgt zunächst für  $i = n$

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

aber dann auch für  $i = 1, \dots, n-1$ , wenn man  $(**)$  in  $(*)$  einsetzt, dass

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

also  $\text{grad} f(a) = \lambda \text{grad} h(a)$ . □

**6.13 Beispiel.** Sei  $f : K \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto y^2 - x^2$ , auf der Kreisscheibe  $K := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$  definiert. Da  $f$  stetig ist und  $K$  kompakt, nimmt  $f$  auf  $K$  sein Supremum  $c := \sup f(x, y)$  an. Wir wollen  $c$  und die Stellen  $(x, y) \in K$  wo  $f$  den Wert  $c$  annimmt, berechnen. Dazu suchen wir zunächst lokale Extrema im Inneren und auf dem Rand.

(i) **Innere Punkte:** Da  $\text{grad} f(x, y) = (-2x, 2y) \neq (0, 0)$  für  $(x, y) \neq (0, 0)$  und

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

indefinit ist, hat  $f$  in  $\overset{\circ}{K} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$  kein lokales Extremum. Die Extrema von  $f$  müssen also auf dem Rand

$$M = \{(x, y) \mid x^2 + y^2 = 1\}$$

liegen.

(ii) **Randpunkte:** Betrachte die Nebenbedingung  $h = 0$  für

$$h : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2 - 1.$$

Ein Maximum von  $f$  auf einem Randpunkt  $a \in M$  ist sicherlich auch ein Maximum von  $f$  unter der Nebenbedingung  $h = 0$ . Es muss also gelten, dass

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

für ein  $\lambda \in \mathbb{R}$ . Beachte, dass  $\text{grad} h(x, y) = (2x, 2y) \neq 0$  für  $(x, y) \in M$ . Wir erhalten somit die Gleichungen

$$\begin{aligned} \text{(I.)} \quad & -2x &= & \lambda(2x) \\ \text{(II.)} \quad & 2y &= & \lambda(2y) \\ \text{(III.)} \quad & x^2 + y^2 &= & 1 \end{aligned}$$

## 6 Implizite Funktionen

Ist nun  $x \neq 0$ , so folgt aus I., dass  $\lambda = -1$  und dann aus II. und III., dass  $y = 0$  und  $x = \pm 1$ . Ist dagegen  $y \neq 0$ , so folgt  $x = 0$  und  $y = \pm 1$ . Man erhält also die vier Kandidaten

$$a \in \{(1, 0), (0, 1), (-1, 0), (0, -1)\}.$$

Da  $f(\pm 1, 0) = -1$  und  $f(0, \pm 1) = +1$  ist, gilt  $c = +1$  und das Supremum wird genau in den Punkten  $(0, 1)$  und  $(0, -1)$  angenommen.

**6.14 Bemerkung.** Man leitet die Gleichungen I. - III. aus Beispiel 6.13 oft folgendermaßen ab: statt  $\nabla f = 0$  fordert man bei NB  $h = 0$ , dass für ein  $\lambda \in \mathbb{R}$

$$\nabla(f + \lambda h) = 0 \quad (= \text{I.} + \text{II.})$$

und eben

$$h = 0 \quad (= \text{III.})$$

gelten. Man nennt  $\lambda$  auch den Lagrangeschen Multiplikator.

**6.15 Beispiel.** Wir geben nun einen alternativen Beweis für die Tatsache, dass jede symmetrische Matrix mindestens einen reellen Eigenwert hat. Sei dazu  $A \in M(n \times n, \mathbb{R})$  symmetrisch und

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \langle x, Ax \rangle, \quad \text{sowie} \quad h : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \langle x, x \rangle - 1.$$

Die Fläche  $\{h = 0\}$  ist die Einheitssphäre  $S^{n-1}$  im  $\mathbb{R}^n$  und ist kompakt. Damit nimmt die stetige Funktion  $f$  auf  $S^{n-1}$  ihr Maximum und ihr Minimum an. Sei  $x_0$  ein Punkt an dem  $f$  auf  $S^{n-1}$  maximal wird. Dann gilt nach Satz 6.11, dass

$$\nabla f(x_0) = 2Ax_0 = \lambda \nabla h(x_0) = \lambda 2x_0$$

für ein  $\lambda \in \mathbb{R}$ . Also ist  $\lambda$  Eigenwert zum Eigenvektor  $x_0$ .

### 6.16 Bemerkung. Mehrere Nebenbedingungen

Liegen mehrere Nebenbedingungen  $h_1, \dots, h_k$  vor, sucht man also ein Extremum von  $f$  auf der  $(n - k)$ -dimensionalen "Fläche"

$$\{h_1 = 0\} \cap \dots \cap \{h_k = 0\},$$

so ergibt sich die notwendige Bedingung, dass Richtungsableitungen  $\langle v, \nabla f(a) \rangle$  von  $f$  die tangential an alle Hyperflächen  $M_j := \{h_j = 0\}$  liegen, verschwinden. Also

$$\langle v, \nabla f(a) \rangle = 0 \quad \text{falls} \quad \langle v, \nabla h_j(a) \rangle = 0 \quad \forall j = 1, \dots, k.$$

Damit ist eine notwendige Bedingung für das Vorliegen eines Extremums von  $f|_{\bigcap_{j=1}^k M_j}$ , dass

$$\nabla f(a) \in \text{span}\{\nabla h_1(a), \dots, \nabla h_k(a)\}. \quad (*)$$

Fasst man  $(h_1, \dots, h_k)$  als Vektorwertige Funktion  $h : \mathbb{R}^n \rightarrow \mathbb{R}^k$  auf, so kann man  $(*)$  wie gehabt aus

$$\nabla(f + \lambda \cdot h) = 0$$

herleiten, wobei der Lagrangemultiplikator  $\lambda$  nun ein Vektor im  $\mathbb{R}^k$  ist.

# 7 Differentialrechnung in Banachräumen

Fast alle Resultate der vorhergehenden Kapitel gelten im Wesentlichen unverändert, wenn man den  $\mathbb{R}^n$  durch einen beliebigen Banachraum, also einen vollständigen normierten Raum  $(X, \|\cdot\|)$ , ersetzt. Wir werden das aus Zeitmangel nicht im Detail ausführen können, besprechen aber in diesem kurzen Kapitel die grundlegende Definition und eine Beispielanwendung.

## 7.1 Definition. Fréchet Ableitung

Seien  $X$  und  $Y$  Banachräume und  $G \subset X$  ein Gebiet. Eine Abbildung  $f : G \rightarrow Y$  heißt differenzierbar im Punkt  $x$  in  $G$ , wenn es eine stetige lineare Abbildung  $A : X \rightarrow Y$  gibt so, dass

$$f(x+h) = f(x) + Ah + o(\|h\|)$$

für  $h$  in einer hinreichend kleinen Umgebung der Null.

Wie im endlichdimensionalen Fall ist  $A$  eindeutig bestimmt und heißt **Fréchet Ableitung** von  $f$  bei  $x$  bezeichnet mit  $Df(x)$ .

**7.2 Bemerkung.** (a) Ein wichtiger Punkt in obiger Definition ist, dass die Stetigkeit von  $A$  gefordert wird. In unendlichdimensionalen normierten Räumen impliziert die Linearität einer Abbildung nämlich nicht die Stetigkeit.

(b) Wieder impliziert die Differenzierbarkeit in einem Punkt die Stetigkeit in diesem Punkt, vgl. Satz 4.29.

(c) Der Mittelwertsatz, der Satz über implizite Funktionen, der Satz über die Umkehrabbildung und die Aussagen über lokale Extrema mit und ohne Nebenbedingungen gelten analog auch für differenzierbare Funktionen auf Banachräumen.

## 7.3 Beispiel. Euler-Lagrange-Gleichungen in der klassischen Mechanik

Ohne Beweis stellen wir zunächst fest, dass der Raum der zweimal stetig differenzierbaren Pfade  $X := \{x : [0, t] \rightarrow \mathbb{R}^n \mid x \text{ ist zweimal stetig differenzierbar}\}$  mit der Norm

$$\|x\|_X := \|x\|_\infty + \|\dot{x}\|_\infty + \|\ddot{x}\|_\infty := \sup_{s \in [0, t]} \|x(s)\|_{\mathbb{R}^n} + \sup_{s \in [0, t]} \|\dot{x}(s)\|_{\mathbb{R}^n} + \sup_{s \in [0, t]} \|\ddot{x}(s)\|_{\mathbb{R}^n}$$

ein Banachraum ist, wobei die Punkte wie in der Physik die Zeitableitungen darstellen.

In der Physik ordnet man nun jedem Weg  $x \in X$  eine sogenannte Wirkung zu, d.h. man definiert

$$S : X \rightarrow \mathbb{R}, \quad x \mapsto S(x) := \int_0^t L(x(s), \dot{x}(s)) ds,$$

wobei  $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $(q, v) \mapsto L(q, v)$  die zweimal stetig differenzierbare Lagrangefunktion ist. Für ein nichtrelativistisches Teilchen mit Masse  $m > 0$  in einem differenzierbaren Potential  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  ist beispielsweise

$$L(q, v) = \frac{1}{2}m\|v\|_{\mathbb{R}^n}^2 - V(q).$$

Wir bestimmen zunächst die Ableitung  $DS$  von  $S$ . Dazu stellen wir fest, dass für  $h \in X$

$$\begin{aligned} S(x+h) &= \int_0^t L(x(s) + h(s), \dot{x}(s) + \dot{h}(s)) ds \\ &= \int_0^t \left( L(x(s), \dot{x}(s)) + \left\{ \left\langle \frac{\partial L}{\partial q}(x(s), \dot{x}(s)), h(s) \right\rangle + \left\langle \frac{\partial L}{\partial v}(x(s), \dot{x}(s)), \dot{h}(s) \right\rangle \right\} \right) ds \\ &\quad + O(\|h\|_X^2). \end{aligned}$$

Damit gilt

$$\begin{aligned} S(x+h) - S(x) &= \int_0^t \left( \left\langle \frac{\partial L}{\partial q}(x(s), \dot{x}(s)), h(s) \right\rangle + \left\langle \frac{\partial L}{\partial v}(x(s), \dot{x}(s)), \dot{h}(s) \right\rangle \right) ds + O(\|h\|_X^2) \\ &= \left\langle \frac{\partial L}{\partial v}(x(s), \dot{x}(s)), h(s) \right\rangle \Big|_0^t \\ &\quad + \int_0^t \left\langle \frac{\partial L}{\partial q}(x(s), \dot{x}(s)) - \frac{d}{ds} \frac{\partial L}{\partial v}(x(s), \dot{x}(s)), h(s) \right\rangle ds + O(\|h\|_X^2), \end{aligned}$$

wobei  $O(\|h\|_X^2)$  bedeutet, dass der Restterm betragsmäßig durch eine von  $h$  unabhängige Konstante mal  $\|h\|_X^2$  beschränkt ist. Die Ableitung  $DS(x) : X \rightarrow \mathbb{R}$ ,

$$\begin{aligned} DS(x)h &= \left\langle \frac{\partial L}{\partial v}(x(t), \dot{x}(t)), h(t) \right\rangle - \left\langle \frac{\partial L}{\partial v}(x(0), \dot{x}(0)), h(0) \right\rangle \\ &\quad + \int_0^t \left\langle \frac{\partial L}{\partial q}(x(s), \dot{x}(s)) - \frac{d}{ds} \frac{\partial L}{\partial v}(x(s), \dot{x}(s)), h(s) \right\rangle ds \end{aligned}$$

existiert also an jedem Punkt  $x \in X$  als stetige lineare Abbildung von  $X$  nach  $\mathbb{R}$ .

Die physikalisch realisierten Trajektorien  $x$  bekommt man nun als kritische Punkte der Wirkung bei festgehaltenen Endpunkten  $x(0) = x_0$  und  $x(t) = x_1$ . Wir fordern also  $DS(x)h = 0$  für alle  $h \in X$  mit  $h(0) = h(t) = 0$ , was auf die Euler-Lagrange-Gleichung

$$\frac{\partial L}{\partial q}(x(s), \dot{x}(s)) - \frac{d}{ds} \frac{\partial L}{\partial v}(x(s), \dot{x}(s)) = 0$$

führt.

Im letzten Schritt hätten wir auch unter den jeweils  $n$  Nebenbedingungen

$$H_0(x) = x(0) - x_0 = 0 \quad \text{und} \quad H_1(x) = x(t) - x_1 = 0$$

minimieren können. Wegen

$$DH_0(x)h = h(0) \quad \text{und} \quad DH_1(x)h = h(t)$$

folgt dann aus

$$\begin{aligned} (DS(x) - \lambda_0 DH_0(x) - \lambda_1 DH_1(x))h &= \left\langle \frac{\partial L}{\partial v}(x(t), \dot{x}(t)) - \lambda_1, h(t) \right\rangle - \left\langle \frac{\partial L}{\partial v}(x(0), \dot{x}(0)) + \lambda_0, h(0) \right\rangle \\ &\quad + \int_0^t \left\langle \frac{\partial L}{\partial q}(x(s), \dot{x}(s)) - \frac{d}{ds} \frac{\partial L}{\partial v}(x(s), \dot{x}(s)), h(s) \right\rangle ds \\ &\stackrel{!}{=} 0, \end{aligned}$$

dass  $\lambda_0 = -\frac{\partial L}{\partial v}(x(0), \dot{x}(0))$  und  $\lambda_1 = \frac{\partial L}{\partial v}(x(t), \dot{x}(t))$ , sowie wiederum die Euler-Lagrange Gleichung

$$\frac{\partial L}{\partial q}(x(s), \dot{x}(s)) - \frac{d}{ds} \frac{\partial L}{\partial v}(x(s), \dot{x}(s)) = 0.$$

#### 7.4 Bemerkung. Funktionalableitung

Ist  $f : X \rightarrow \mathbb{R}$  auf einem Raum  $X$ , dessen Elemente  $x$  selbst Funktionen sind, definiert, so schreibt man in der Physik oft  $f[x]$  statt  $f(x)$ . Das Differential einer solchen Abbildung  $f : X \rightarrow \mathbb{R}$  wird oft auch Funktionalableitung genannt. Das Differential  $Df[x]$  ist eine stetige lineare Abbildung von Funktionen in  $X$  nach  $\mathbb{R}$ , und solche Abbildungen heißen (für bestimmte Räume  $X$ ) Distributionen. Weiterhin ist es in der Physik üblich, Distributionen formal wie Funktionen zu notieren, also einer linearen Abbildung  $A : X \rightarrow \mathbb{R}$  den symbolischen Ausdruck  $A(s)$  zuzuordnen und

$$Ah =: \int A(s)h(s)ds$$

zu schreiben. Diese Notation verwendet man auch für  $Df$  und schreibt dann

$$Df[x]h = \int \frac{\delta f}{\delta x}[x](s) h(s) ds.$$

In obigem Beispiel wäre also

$$\frac{\delta S}{\delta x}[x](s) = \frac{\partial L}{\partial v}(x(t), \dot{x}(t)) \delta(s-t) - \frac{\partial L}{\partial v}(x(0), \dot{x}(0)) \delta(s) + \frac{\partial L}{\partial q}(x(s), \dot{x}(s)) - \frac{d}{ds} \frac{\partial L}{\partial v}(x(s), \dot{x}(s)),$$

wobei die Diracsche Delta-Distribution durch

$$\delta : X \rightarrow \mathbb{R}, \quad x \mapsto \delta(x) = x(0) =: \int x(s) \delta(s) ds$$

gegeben ist.

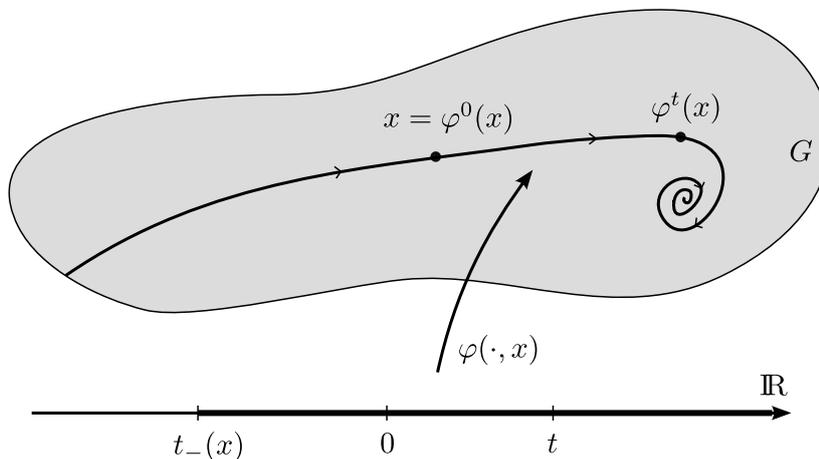


# 8 Gewöhnliche Differentialgleichungen und dynamische Systeme

In einem deterministischen dynamischen System kann man jedem Zustand  $\cong$  Punkt  $x$  im Phasenraum  $G$  eine eindeutige zeitliche Entwicklung zuordnen: Ist das System zum Zeitpunkt  $t_0 = 0$  im Zustand  $x \in G$ , so sei der Zustand zum Zeitpunkt  $t$  gleich  $\varphi(t, x)$ . Dabei kann das "Existenzintervall"  $I(x) \ni 0$ , auf dem die Abbildung  $\varphi(\cdot, x) : I(x) \rightarrow G$  erklärt ist, von  $x$  abhängen. Wir schreiben oft  $\varphi^t(x) := \varphi(t, x)$  und verlangen  $\varphi^0(x) = x$  und die Verträglichkeitsbedingung

$$\varphi^s(\varphi^t(x)) = \varphi^{s+t}(x)$$

für alle  $t \in I(x)$  und  $s \in I(\varphi^t(x))$ .



## 8.1 Definition. Dynamisches System

Ein **dynamisches System** auf einem Gebiet  $G \subset \mathbb{R}^n$  ist gegeben durch eine stetig differenzierbare Abbildung  $\varphi : \Omega \rightarrow G$ ,  $(t, x) \mapsto \varphi^t(x)$  mit folgenden Eigenschaften:

- (a)  $\Omega \subset \mathbb{R} \times G$  ist ein Gebiet, so dass  $\{0\} \times G \subset \Omega$  ist und für jedes  $x \in G$  ist

$$I(x) := \{t \in \mathbb{R} \mid (t, x) \in \Omega\} \subset \mathbb{R}$$

ein Intervall, das sogenannte Existenzintervall zum Startpunkt  $x$ .

- (b) (i) Für alle  $x \in G$  ist  $\varphi^0(x) = x$ ;  
(ii) für alle  $x \in G$  und  $t \in I(x)$  gilt:

Es ist  $s \in I(\varphi^t(x))$  genau wenn  $s + t \in I(x)$ . In diesem Fall gilt

$$\varphi^s(\varphi^t(x)) = \varphi^{s+t}(x).$$

**8.2 Bemerkung.** (a) Man nennt  $G \subset \mathbb{R}^n$  den **Phasenraum** des dynamischen Systems  $\varphi$ .

- (b) Weil  $\Omega \subset \mathbb{R} \times G \subset \mathbb{R}^{n+1}$  offen ist, ist auch  $I(x) \subset \mathbb{R}$  offen und wegen Definition 8.1 (a) ein offenes Intervall,

$$I(x) =: (t_-(x), t_+(x)).$$

Hierbei ist  $t_-(x) \in [-\infty, 0)$  und  $t_+(x) \in (0, +\infty]$ . Wir nennen  $\varphi(x) := \{\varphi^t(x) \mid t \in I(x)\}$  den Orbit von  $x$  und  $I(x)$  sei Existenzintervall.

(c) Setzen wir für jedes  $t \in \mathbb{R}$

$$G_t := \{x \in G \mid (t, x) \in \Omega\}$$

( $G_t$  ist also die Menge derjenigen  $x \in G$ , deren "Zeitentwicklung" mindestens bis zur Zeit  $t$  existiert), so folgt aus Definition 8.1 (b), dass

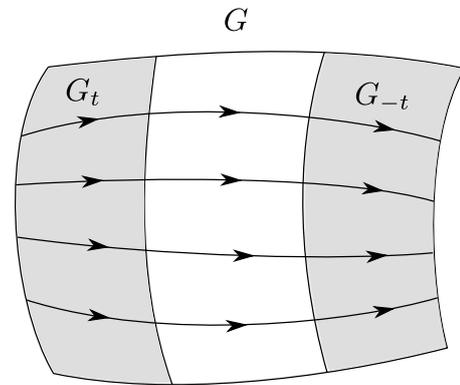
$$\varphi^t(G_t) = \{\varphi^t(x) \mid x \in G_t\} = G_{-t}$$

gilt und, dass

$$\varphi^t : G_t \rightarrow G_{-t}$$

ein Diffeomorphismus ist, denn  $\varphi^{-t} : G_{-t} \rightarrow G_t$  ist sein Inverses:

$$\begin{aligned} \varphi^{-t} \circ \varphi^t(x) &= \varphi^{-t+t}(x) = \varphi^0(x) = x \\ &= \varphi^t \circ \varphi^{-t}(x). \end{aligned}$$



Man nennt die Familie  $(\varphi^t)_{t \in \mathbb{R}}$  von Diffeomorphismen den **Fluss** von  $\varphi$ .

Statt  $\varphi^t(x)$  schreibt man oft auch kurz  $x(t)$ . Nachteil: In dieser Notation ist der Anfangspunkt  $\varphi^0(x) = x = x(0)$  nicht mehr explizit enthalten.

(d) Besonders schön ist die Situation, wenn  $I(x) = \mathbb{R}$  ist, für alle  $x \in G$ , also  $G_t = G$  für alle  $t \in \mathbb{R}$ . In diesem Fall ist  $\Omega = \mathbb{R} \times G$  und man spricht von einem **globalen dynamischen System** bzw. einem globalen Fluss auf  $G$ . Es ist dann also

$$\mathbb{R} \rightarrow \text{Diff}(G), t \mapsto \varphi^t$$

ein Gruppenhomomorphismus von  $(\mathbb{R}, +)$  in die **Diffeomorphismengruppe** von  $G$ ,

$$\text{Diff}(G) = \{\varphi : G \rightarrow G \mid \varphi \text{ ist Diffeomorphismus}\}.$$

Umgekehrt definiert jeder stetig differenzierbare Homomorphismus  $\varphi : \mathbb{R} \rightarrow \text{Diff}(G)$  ein globales dynamisches System auf  $G$ .

### 8.3 Definition. Vektorfeld eines Flusses

Sei  $\varphi$  ein dynamisches System auf einem Gebiet  $G \subset \mathbb{R}^n$ . Dann heißt

$$v : G \rightarrow \mathbb{R}^n, \quad x \mapsto v(x) = \left. \frac{d}{dt} \varphi^t(x) \right|_{t=0} =: \dot{x}(0)$$

das **Vektorfeld** von  $\varphi$  auf  $G$ .

**8.4 Bemerkung.** Sei  $\varphi$  ein dynamisches System auf  $G \subset \mathbb{R}^n$  mit Vektorfeld  $v : G \rightarrow \mathbb{R}^n$ . Dann gilt für alle  $x \in G$  und alle  $t \in I(x)$

$$\frac{d}{dt} \varphi^t(x) = v(\varphi^t(x))$$

oder kurz

$$\dot{x}(t) = v(x(t)).$$

Das Vektorfeld selbst hängt also **nicht** von der Zeit ab.

*Beweis.* Seien  $x \in G$  und  $t \in I(x)$  beliebig. Mit  $y = \varphi^t(x)$  gilt nun:

$$v(\varphi^t(x)) = v(y) = \left. \frac{d}{ds} \varphi^s(y) \right|_{s=0} = \left. \frac{d}{ds} \varphi^s(\varphi^t(x)) \right|_{s=0} = \left. \frac{d}{ds} \varphi^{s+t}(x) \right|_{s=0} = \left. \frac{d}{dt} \varphi^{s+t}(x) \right|_{s=0} = \left. \frac{d}{dt} \varphi^t(x) \right|_{s=0} = \frac{d}{dt} \varphi^t(x).$$

□

**8.5 Bemerkung.** (a) Jede der Kurven  $I(x) \rightarrow \mathbb{R}^n$ ,  $t \mapsto \varphi^t(x) =: x(t)$  erfüllt also das System von  $n$  Gleichungen

$$\begin{aligned} \dot{x}_1(t) &= v_1(x_1(t), \dots, x_n(t)) \\ &\vdots \\ \dot{x}_n(t) &= v_n(x_1(t), \dots, x_n(t)) \end{aligned}$$

oder kurz eben

$$\dot{x} = v(x).$$

Man nennt ein solches System ein **autonomes System gewöhnlicher Differentialgleichungen**. „Differentialgleichung“, weil in den Gleichungen sowohl die Funktionen  $x_1, \dots, x_n$  als auch ihre Ableitungen vorkommen; „gewöhnlich“, weil die Funktionen nur von **einer** reellen Veränderlichen  $t$  abhängen (nicht von mehreren, in welchem Fall man von „partiellen“ Differentialgleichungen spricht); „autonom“, weil das Vektorfeld  $v$  nicht selbst von der Zeit  $t$  abhängt.

(b) In den meisten Fällen ist es nun so, dass man das dynamische System  $\varphi$  auf  $G \subset \mathbb{R}^n$  gerne kennen würde, aber zunächst nur das zugehörige Vektorfeld  $v : G \rightarrow \mathbb{R}^n$  kennt. Es stellt sich also die fundamentale Frage: Inwieweit legt das zugehörige Vektorfeld  $v$  das dynamische System  $\varphi$  fest? Wir werden zeigen, dass jedes stetig differenzierbare Vektorfeld  $v$  auf  $G \subset \mathbb{R}^n$  genau ein (maximales) dynamisches System  $\varphi$  auf  $G$  festlegt.

### 8.6 Beispiel. Klassische Mechanik

Bewegt sich ein Teilchen der Masse  $m > 0$  in einem Gebiet  $D \subset \mathbb{R}^3$  unter dem Einfluss einer Kraft  $K : D \rightarrow \mathbb{R}^3$ , so löst die Bahnkurve des Teilchens  $t \mapsto x(t)$  die Gleichung

$$m \ddot{x}(t) = F(x(t)) \quad (\text{Masse} \times \text{Beschleunigung} = \text{Kraft}).$$

Setzen wir  $G : D \times \mathbb{R}^3 \subset \mathbb{R}^6$  und  $v : G \rightarrow \mathbb{R}^6$ ,

$$v(x, y) = \begin{pmatrix} y \\ \frac{1}{m} F(x) \end{pmatrix}, \quad \text{sowie} \quad z(t) = \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix},$$

so löst  $z(t)$  die Gleichung

$$\dot{z}(t) = v(z(t)),$$

denn

$$\dot{z}(t) = \begin{pmatrix} \dot{x}(t) \\ \dot{\dot{x}}(t) \end{pmatrix} = \begin{pmatrix} \dot{x}(t) \\ \frac{1}{m} F(x(t)) = v(z(t)) \end{pmatrix}.$$

Es ist also  $z$  die Lösung der gewöhnlichen Differentialgleichung

$$\dot{z} = v(z).$$

Bestimmt man also zum Vektorfeld  $v : G \rightarrow \mathbb{R}^6$  das zugehörige dynamische System  $\varphi$  (d.h. alle Lösungskurven  $z : I(z) \rightarrow G$ ), so erhält man durch Projektion auf die ersten 3 Koordinaten alle Teilchenbahnen  $x : I(z) \rightarrow D$  im Kraftfeld  $F$ .

**8.7 Definition. Lösung einer Differentialgleichung**

Sei  $v : G \rightarrow \mathbb{R}^n$  ein stetiges Vektorfeld. Sei  $I \subset \mathbb{R}$  ein offenes Intervall. Man nennt eine stetig differenzierbare Kurve  $x : I \rightarrow G, t \mapsto x(t)$ , eine **Lösung der Differentialgleichung**

$$\dot{x} = v(x),$$

wenn für alle  $t \in I$  gilt

$$\dot{x}(t) = v(x(t)).$$

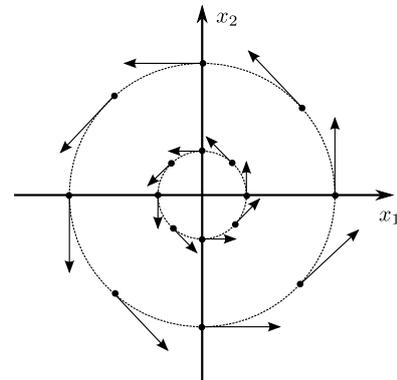
Falls  $0 \in I$  ist, so heißt  $x(t)$  eine **Lösung zum Anfangswert**  $x_0 \in G$ , wobei  $x_0 = x(0)$  ist.

**8.8 Bemerkung. Phasendiagramm**

Man verdeutlicht sich das Vektorfeld  $v : G \rightarrow \mathbb{R}^n$  oft durch das sogenannte **Phasendiagramm**. Sei beispielsweise  $v : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x_1, x_2) \mapsto (-x_2, x_1)$  und

$$\dot{x} = v(x) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

Die Lösungskurven  $x(t)$  sind überall tangential an  $v$  und heißen **Integrialkurven** an das Vektorfeld  $v$ . Beim Lösen einer Differentialgleichung spricht man auch vom „Integrieren der Gleichung“.



**8.9 Beispiele. (a) Der freie Fall**

$$F(x) = -mge_3, \quad g > 0, \quad e_3 = (0, 0, 1).$$

Der zugängliche Ortsraum (= Konfigurationsraum) ist  $D = \mathbb{R} \times \mathbb{R} \times (0, \infty)$  und der Phasenraum ist  $G = D \times \mathbb{R}^3$ . Wir suchen Lösungen von

$$\ddot{x} = \frac{1}{m}F(x) = -ge_3,$$

d.h. wir müssen das System  $\dot{z} = v(z)$  mit  $v(x, y) = (y, -ge_3)$  integrieren:

$$\begin{array}{ll} \dot{x}_1 = y_1 & \dot{y}_1 = 0 \\ \dot{x}_2 = y_2 & \dot{y}_2 = 0 \\ \dot{x}_3 = y_3 & \dot{y}_3 = -g \end{array}$$

Hier bekommt man die Lösungen wirklich unmittelbar durch Integration:

$$\begin{array}{ll} y_1 = b_1 & x_1 = b_1 t + a_1 \\ y_2 = b_2 & x_2 = b_2 t + a_2 \\ y_3 = -gt + b_3 & x_3 = -\frac{g}{2} t^2 + b_3 t + a_3 \end{array}$$

mit Integrationskonstanten  $a_1, \dots, b_3 \in \mathbb{R}$  für alle  $t \in \mathbb{R}$ . Es ist dann  $x(0) = a$  und  $\dot{x}(0) = b$  und das zu  $v : G \rightarrow \mathbb{R}^6, v(x, y) = (y, -ge_3)$  gehörende dynamische System  $\varphi$  auf  $\Omega \subset \mathbb{R} \times G$  ist gegeben durch  $\varphi : \Omega \rightarrow G$ ,

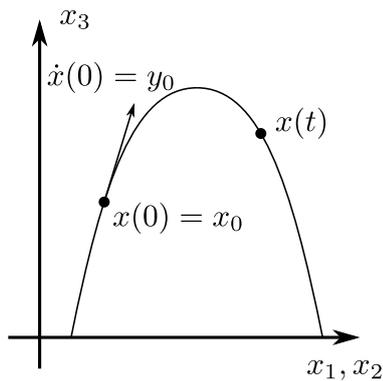
$$\varphi^t(x, y) = \left(-\frac{g}{2} t^2 e_3 + t y + x, -g t e_3 + y\right).$$

Die Existenzintervalle sind durch die Bedingung  $x_3(t) > 0$  eingeschränkt, es gilt also

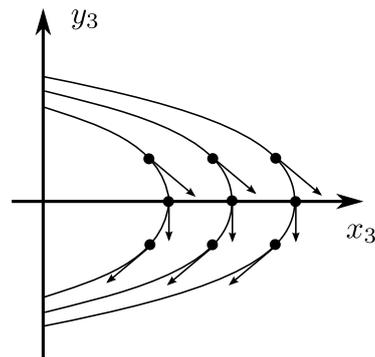
$$I(x, y) = \left( \frac{y_3}{g} + \sqrt{\frac{y_3^2}{g^2} + \frac{2x_3}{g}}, \frac{y_3}{g} - \sqrt{\frac{y_3^2}{g^2} + \frac{2x_3}{g}} \right)$$

und dementsprechend  $\Omega = \{(t, x, y) \in \mathbb{R} \times G \mid t \in I(x, y)\}$ . Die Lösung von  $m\ddot{x} = F(x)$  auf  $D \subset \mathbb{R}^3$  zum Anfangswert  $(x(0), \dot{x}(0)) = (x_0, y_0)$  ist somit

$$x(t) = -\frac{g}{2} t^2 e_3 + t y_0 + x_0. \quad (\text{Parabelbahn})$$



Lösung im Ortsraum



Phasenraumdiagramm

(b) **Das Hookesche Gesetz**

Ein Gewicht der Masse  $m$  hänge an einer Feder der Federkonstante  $k > 0$  und  $x \in \mathbb{R}$  bezeichne die vertikale Auslenkung aus der Ruhelage. Dann wirkt auf das Gewicht die Kraft  $F(x) = -kx$  und die Bewegungsgleichung lautet diesmal

$$m\ddot{x} = -kx \quad \text{bzw.} \quad \ddot{x} = -\omega^2 x$$

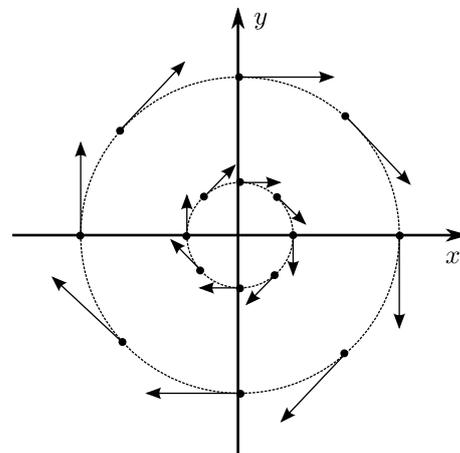
mit  $\omega = \sqrt{\frac{k}{m}}$ . Mit  $G = \mathbb{R} \times \mathbb{R}$  und  $v : G \rightarrow \mathbb{R}^2$ ,  $v(x, y) = (y, -\omega^2 x)$  ist also

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\omega^2 x \end{aligned}$$

zu lösen.

Die qualitative Form der Lösungen liest man direkt aus dem Phasenraumdiagramm ab. Auch die explizite Form erhält man ganz direkt durch Integration,

$$x(t) = x_0 \cos(\omega t) + \frac{y_0}{\omega} \sin \omega t.$$



**8.10 Definition. Systeme gewöhnlicher DGLen  $m$ -ter Ordnung**

Sei  $m \in \mathbb{N}$ . Ein **System gewöhnlicher Differentialgleichungen  $m$ -ter Ordnung** auf einem Gebiet  $D \subset \mathbb{R}^n$  ist gegeben durch eine stetige Funktion

$$f : D \times \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{(m-1)\text{-mal}} \rightarrow \mathbb{R}^n$$

und die Gleichung

$$x^{(m)} = f(x, \dot{x}, \dots, x^{(m-1)}). \quad (*)$$

Unter einer Lösung von  $(*)$  versteht man eine  $m$ -mal stetig differenzierbare Kurve  $x : I \rightarrow D$ ,  $I \subset \mathbb{R}$  ein offenes Intervall, so dass für alle  $t \in I$  gilt

$$x^{(m)}(t) = f(x(t), \dot{x}(t), \dots, x^{(m-1)}(t)).$$

Falls  $0 \in I$  ist, heißt  $x : I \rightarrow D$  Lösung zum Anfangswert  $(x_0, y_1, \dots, y_{m-1})$  wobei  $x(0) = x_0$  und  $x^{(j)} = y_j$  für  $j = 1, \dots, m-1$ .

**8.11 Bemerkung. Reduktion auf ein System erster Ordnung**

Man kann jedes System  $m$ -ter Ordnung auf einem Gebiet  $D \subset \mathbb{R}^n$ ,

$$x^{(m)} = f(x, \dot{x}, \dots, x^{(m-1)}),$$

immer auf ein System erster Ordnung auf dem Gebiet  $G := D \times \mathbb{R}^{n(m-1)} \subset \mathbb{R}^{nm}$  zurückführen. Man definiert dazu einfach das Vektorfeld  $v : G \rightarrow \mathbb{R}^{nm}$  durch

$$v(x, y_1, \dots, y_{m-1}) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ f(x, y_1, \dots, y_{m-1}) \end{pmatrix}$$

und löst dann mit  $z = (x, y_1, \dots, y_{m-1}) \in G$  das System

$$\dot{z} = v(z).$$

*Beweis.* Ist  $x : I \rightarrow D$ ,  $t \mapsto x(t)$ , Lösung von  $x^{(m)} = f(x, \dot{x}, \dots, x^{(m-1)})$ , so ist  $z : I \rightarrow G$

$$z(t) = (x(t), \dot{x}(t), \dots, x^{(m-1)}(t))$$

Lösung von  $\dot{z} = v(z)$ . Ist umgekehrt  $z(t) = (x(t), y_1(t), \dots, y_{m-1}(t))$  Lösung von  $\dot{z} = v(z)$ , so ist  $t \mapsto x(t)$ ,  $I \rightarrow D$  Lösung von  $x^{(m)} = f(x, \dot{x}, \dots, x^{(m-1)})$ .  $\square$

**8.12 Definition. Autonome und nicht-autonome Systeme**

Sei  $I \subset \mathbb{R}$  ein offenes Intervall. Ein stetiges **zeitabhängiges Vektorfeld**  $f$  auf einem Gebiet  $D \subset \mathbb{R}^n$  ist eine stetig differenzierbare Abbildung

$$f : I \times D \rightarrow \mathbb{R}^n, \quad (t, x) \mapsto f(t, x).$$

Man nennt dann das System gewöhnlicher Differentialgleichungen

$$\dot{x} = f(t, x) \quad (*)$$

**nicht-autonom**, wenn  $f$  explizit von  $t \in I$  abhängt. Ist  $f$  (wie bisher) unabhängig von  $t$ , so heißt  $\dot{x} = f(x)$  **autonom**. Sei  $J \subset I$  ein Teilintervall. Eine stetig differenzierbare Kurve  $x : J \rightarrow D$  heißt **Lösung von  $(*)$** , wenn für alle  $t \in J$  gilt

$$\dot{x}(t) = f(t, x(t)).$$

Für  $t_0 \in J$  heißt  $x$  Lösung zum Anfangswert  $x_0 = x(t_0)$  zur Anfangszeit  $t_0$ .

### 8.13 Bemerkung. Reduktion auf ein autonomes System

Man kann jedes nicht-autonome System  $\dot{x} = f(t, x)$  auf einem Gebiet  $D \subset \mathbb{R}^n$  auf ein autonomes System auf  $G := I \times D \subset \mathbb{R}^{n+1}$  wie folgt zurückführen:

Man setzt  $v : G \rightarrow \mathbb{R}^{n+1}$ ,  $v(s, y) = (1, f(s, y))$ ,  $z := (s, y) \in G$ , und löse dann  $\dot{z} = v(z)$  auf  $G$  zum Anfangswert  $(s, y) = (t_0, x_0)$ . Sei  $z(t) = (\underbrace{s(t)}_{=t_0+t}, y(t))$  diese Lösung, dann ist  $x(t) = y(t - t_0)$

Lösung von

$$\dot{x} = f(t, x)$$

zum Anfangswert  $x_0$  zur Anfangszeit  $t_0$ .

*Beweis.* Es ist  $x(t_0) = y(0) = x_0$  und es gilt

$$\dot{x}(t) = \dot{y}(t - t_0) = f(\underbrace{s(t - t_0)}_{=t}, y(t - t_0)) = f(t, x(t)).$$

□

**8.14 Bemerkung.** (a) Im allgemeinen hat man keine Chance, ein System gewöhnlicher Differentialgleichungen

$$\dot{x} = v(x)$$

auf ein Gebiet  $G \subset \mathbb{R}^n$  bei Vorgabe von  $v : G \rightarrow \mathbb{R}^n$  zu integrieren, d.h. den zugehörigen Fluss  $\varphi$  explizit zu bestimmen. Man ist daher schon damit zufrieden, qualitative Aussagen über die Bahn  $t \mapsto x(t)$  zu bekommen, wie z.B. Antworten auf die Fragen:

- (i) Konvergiert  $t \mapsto x(t)$  für  $t \rightarrow \infty$  gegen eine Gleichgewichtslage  $x_* \in G$ , also einen Punkt mit  $v(x_*) = 0$ ?
- (ii) Ist  $t \mapsto x(t)$  periodisch, also  $x(t + T) = x(t)$  für ein  $T > 0$  und alle  $t \in \mathbb{R}$ ?
- (iii) Oder auch nur: für welche  $x \in G$  ist überhaupt  $I(x) = \mathbb{R}$ , existiert also eine Lösung für alle Zeiten?

Man nennt das Studium dieser Fragen die „Qualitative Theorie gewöhnlicher Differentialgleichungen“.

- (b) Hat man spezielle Informationen über  $v$  oder  $G$ , z.B. Symmetrieaussagen oder  $G \subset \mathbb{R}$  so kann man  $\dot{x} = f(x)$  in manchen Fällen doch explizit integrieren.

### 8.15 Proposition. Integration eindimensionaler DGLen

Sei  $G \subset \mathbb{R}$  ein Intervall,  $v : G \rightarrow \mathbb{R}$  stetig und  $x_0 \in G$  mit  $v(x_0) \neq 0$ . Seien weiter  $\delta_1, \delta_2 > 0$  derart, dass  $(x_0 - \delta_1, x_0 + \delta_2) \subset G$  und  $v(x) \neq 0$  für  $x \in (x_0 - \delta_1, x_0 + \delta_2)$  ist. Definiere  $\tau : (x_0 - \delta_1, x_0 + \delta_2) \rightarrow \mathbb{R}$  durch

$$\tau(x) = \int_{x_0}^x \frac{du}{v(u)}.$$

Dann ist  $I := \text{Bild}(\tau) \subset \mathbb{R}$  ein offenes Intervall und  $\tau : (x_0 - \delta_1, x_0 + \delta_2) \rightarrow I$  ist ein Diffeomorphismus. Seine Inverse  $\varphi := \tau^{-1} : I \rightarrow (x_0 - \delta_1, x_0 + \delta_2)$  ist Lösungskurve von  $\dot{x} = v(x)$  auf  $G$  zum Anfangswert  $x_0$ .

**8.16 Bemerkung.** Diese Aussage zeigt die Berechtigung der heuristischen Rechnung

$$\dot{x} = v(x) \Rightarrow \frac{dx}{dt} = v(x) \stackrel{(*)}{\Rightarrow} \frac{dx}{v(x)} = dt \Rightarrow \int_{x_0}^x \frac{du}{v(u)} = \int_0^t dt \Rightarrow \int_{x_0}^x \frac{du}{v(u)} = t.$$

Den Schritt (\*) nennt man oft „Trennung der Variablen“.

*Beweis. von Proposition 8.15.* Es ist  $\tau'(x) = \frac{1}{v(x)} \neq 0$  und nach dem Zwischenwertsatz für alle  $x \in (x_0 - \delta_1, x_0 + \delta_2)$  von gleichem Vorzeichen. Damit ist  $\tau$  streng monoton und deshalb bijektiv auf sein Bild. Nach dem Umkehrsatz ist  $\tau$  ein Diffeomorphismus auf sein offenes Bild  $I$ . Für  $\varphi = \tau^{-1}$  gilt dann  $\dot{\varphi}(t) = \frac{1}{\tau'(\varphi(t))} = v(\varphi(t))$  für alle  $t \in I$ . Wegen  $\tau(x_0) = 0$  ist  $\varphi(0) = x_0$ .  $\square$

- 8.17 Bemerkung.** (a) Man kann also eindimensionale Gleichungen durch „Quadratur“ lösen, d.h. durch die Prozesse „Stammfunktion bilden“ und „Invertieren“.  
 (b) Das Verfahren der Trennung der Variable funktioniert auch für den Fall, dass  $v$  zeitabhängig und von der Form

$$v : I \times G \rightarrow \mathbb{R}, \quad v(t, x) = a(t)b(x)$$

ist, wobei  $I \subset \mathbb{R}$  ein offenes Intervall und  $a$  und  $b$  jeweils stetig sind. Dann liefert Auflösen von

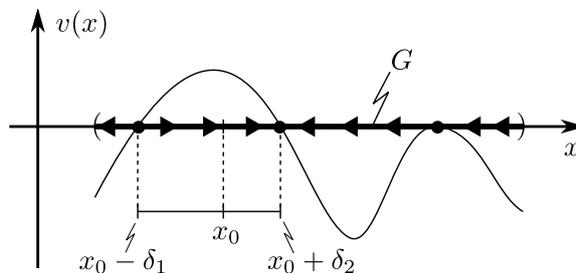
$$\int_{t_0}^{\tau(x)} a(t) dt = \int_{x_0}^x \frac{du}{b(u)}$$

nach  $\tau(x)$  analog zu Proposition 8.15 eine lokale Lösung von

$$\dot{x}(t) = v(t, x(t)) \quad \text{zum Anfangswert } x(t_0) = x_0.$$

- (c) Ganz allgemein sind die Punkte  $x_0 \in G$ , wo ein gegebenes Vektorfeld  $v : G \rightarrow \mathbb{R}^n$  verschwindet,  $v(x_0) = 0$ , für das dynamische System von  $\dot{x} = v(x)$  besonders einfach: Es ist  $I(x_0) = \mathbb{R}$  und  $x_0(t) = \varphi^t(x_0) = x_0$  für alle  $t \in \mathbb{R}$ . Man nennt diese Punkte Gleichgewichtslagen oder stationäre Punkte.  
 (d) Mit Proposition 8.15 ist das dynamische System  $\varphi$  von  $\dot{x} = v(x)$  auf  $G \subset \mathbb{R}$  vollständig bestimmt. Sein Phasendiagramm sieht so aus:

In den Nullstellen von  $v$  ruht das System. Dazwischen läuft es monoton von einer Nullstelle zur nächsten. Ist  $v$  an den Nullstellen differenzierbar, so werden die Gleichgewichtslagen in endlicher Zeit nicht erreicht. An den Rändern des Intervalls  $G$  können die Orbits in endlicher Zeit entweichen.



- 8.18 Beispiele.** (a) Sei  $a \in \mathbb{R}$  und  $v : \mathbb{R} \rightarrow \mathbb{R}$ ,  $v(x) = ax$  ein lineares Vektorfeld. Die Lösung von  $\dot{x} = v(x)$  auf  $G = \mathbb{R}$  zum Anfangswert  $x_0 > 0$  für  $a \neq 0$  ist nach Proposition 8.14 für  $x > 0$  gegeben durch

$$t(x) = \int_{x_0}^x \frac{du}{au} = \frac{1}{a} (\ln x - \ln x_0) = \frac{1}{a} \ln \frac{x}{x_0}$$

also

$$\ln \frac{x}{x_0} = at \Rightarrow x(t) = x_0 e^{at} \quad \forall t \in \mathbb{R}.$$

Diese Formel gilt auch für  $x_0 \leq 0$  und somit ist der Fluss  $\varphi^t : \mathbb{R} \rightarrow \mathbb{R}$ ,  $\varphi^t(x_0) = e^{at} x_0$  global.

- (b) Sei  $v : \mathbb{R} \rightarrow \mathbb{R}$ ,  $v(x) = 1 + x^2$  quadratisch in  $x$ . Die Lösungskurve  $t \rightarrow x(t)$  zum Anfangswert  $x_0 = 0$  ist

$$t(x) = \int_0^x \frac{du}{1+u^2} = \arctan(x)$$

also

$$x(t) = \tan(t).$$

Somit ist  $I(0) = (-\frac{\pi}{2}, \frac{\pi}{2})$  und  $t \mapsto x(t)$  läuft „in endlicher Zeit nach Unendlich“.

- (c) Sei  $v : \mathbb{R} \rightarrow \mathbb{R}$  gegeben durch  $v(x) = \sqrt{|x|}$ . Dann erreichen die Lösungen zu Startwerten  $x(0) < 0$  in endlicher Zeit die Ruhelage  $x = 0$ . Dort können Sie beliebig lange verweilen und dann nach rechts wieder aus der Ruhelage herauslaufen. Die Lösungen sind also nicht eindeutig. (vgl. Übungen).

### 8.19 Definition. Lipschitz-Stetigkeit

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $v : G \rightarrow \mathbb{R}^n$  ein Vektorfeld.

- (a) Es heißt  $v$  **Lipschitz-stetig**, wenn es ein  $L > 0$  gibt, so dass für alle  $x, y \in G$  gilt

$$\|v(x) - v(y)\| \leq L\|x - y\|.$$

Es heißt dann  $L$  eine **Lipschitz-Konstante** für  $v$ .

- (b) Es heißt  $v$  **lokal-Lipschitz-stetig**, wenn jedes  $x \in G$  eine offene Umgebung  $U \subset G$  besitzt, so dass  $v|_U$  Lipschitz-stetig ist.

### 8.20 Bemerkung. Differenzierbar $\Rightarrow$ lokal Lipschitz

Sei  $v : G \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld auf einem Gebiet  $G \subset \mathbb{R}^n$ . Dann ist  $v$  lokal Lipschitz-stetig.

*Beweis.* Übungsaufgabe (leicht). □

**8.21 Proposition.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $v : G \rightarrow \mathbb{R}^n$  ein lokal Lipschitz-stetiges Vektorfeld. Für jedes Kompaktum  $K \subset G$  ist  $v|_K$  Lipschitz-stetig.

*Beweis.* Übungsaufgabe (nicht ganz so leicht). □

### 8.22 Theorem. Existenz- und Eindeutigkeitsatz von Picard-Lindelöf

Sei  $G \subset \mathbb{R}^n$  ein Gebiet,  $v : G \rightarrow \mathbb{R}^n$  ein lokal Lipschitz-stetiges Vektorfeld und  $x_0 \in G$ .

Dann existiert ein  $\delta > 0$ , sodass es genau eine stetig differenzierbare Kurve  $x : (-\delta, \delta) \rightarrow G$  gibt, die Lösung von  $\dot{x} = v(x)$  zum Anfangswert  $x_0$  ist, also

$$\left. \begin{aligned} \dot{x}(t) &= v(x(t)), & \forall t \in (-\delta, \delta) \\ x(0) &= x_0 \end{aligned} \right\} (*)$$

erfüllt.

**8.23 Bemerkung.** Die Grundidee des Beweises ist die Folgende: Ist  $x : (-\delta, \delta) \rightarrow G$  Lösung von (\*), so folgt durch Integration und den Hauptsatz

$$x(t) - \underbrace{x(0)}_{=x_0} = \int_0^t \dot{x}(s) ds = \int_0^t v(x(s)) ds,$$

also

$$x(t) = x_0 + \int_0^t v(x(s)) ds. \quad (**)$$

Jede Lösung von (\*) ist also Fixpunkt der Abbildung  $\Phi$ , die jeder Funktion  $\varphi$  die Funktion  $\Phi[\varphi]$  zuordnet, gegeben durch

$$\Phi[\varphi](t) = x_0 + \int_0^t v(\varphi(s)) ds.$$

Ist umgekehrt  $x$  Fixpunkt von  $\Phi$  und stetig, so liefert Differentiation von (\*\*), dass  $\dot{x}(t) = v(x(t))$  für alle  $t \in (-\delta, \delta)$ , also  $x$  Lösung von  $\dot{x} = v(x)$  zum Anfangswert  $x(0) = x_0$  ist.

Man muss nun lediglich den Funktionenraum auf dem  $\Phi$  operiert so definieren, dass er ein Banachraum stetiger Funktionen ist und  $\Phi$  eine Kontraktion darauf. Dann liefert der Banachsche Fixpunktsatz das Resultat.

*Beweis. von Theorem 8.22.* Weil  $G$  offen ist, existiert ein  $r > 0$ , so dass  $K = \overline{B}_r(x_0) \subset G$  ist. Da  $K$  kompakt und  $v|_K : K \rightarrow \mathbb{R}^n$  stetig ist, existiert ein  $M > 0$  so, dass für alle  $x \in K$  gilt  $\|v(x)\| \leq M$ . Nach Proposition 8.21 existiert auch ein  $L > 0$ , sodass  $v|_K$  Lipschitz-stetig ist mit Lipschitz-Konstante  $L$ .

Setze nun  $\delta := \min\{\frac{1}{L}, \frac{r}{M}\}$  und sei  $0 < \delta_0 < \delta$  beliebig. Wir betrachten den Vektorraum

$$X = \{\varphi : [-\delta_0, \delta_0] \rightarrow \mathbb{R}^n \mid \varphi \text{ ist stetig}\} = C([-\delta_0, \delta_0], \mathbb{R}^n)$$

mit der Supremumsnorm

$$\|\varphi\|_\infty = \sup_{t \in [-\delta_0, \delta_0]} \|\varphi(t)\|.$$

Es ist dann  $(X, \|\cdot\|_\infty)$  ein Banachraum und der Teilraum

$$A = \{\varphi \in X \mid \varphi(0) = x_0 \text{ und } \varphi([-\delta_0, \delta_0]) \subset K\} \subset X$$

ist abgeschlossen. Für  $\varphi \in A$  sei nun  $\Phi[\varphi] : [-\delta_0, \delta_0] \rightarrow \mathbb{R}^n$  gegeben durch

$$\Phi[\varphi](t) := x_0 + \int_0^t v(\varphi(s)) ds.$$

Es ist dann  $\Phi[\varphi]$  wieder in  $A$ , denn  $\Phi[\varphi]$  ist stetig (sogar stetig differenzierbar),  $\Phi[\varphi](0) = x_0$  und  $\Phi[\varphi](t) \in K$  für alle  $t \in [-\delta_0, \delta_0]$ . Letzteres folgt aus

$$\|\Phi[\varphi](t) - x_0\| \leq \left\| \int_0^t v(\varphi(s)) ds \right\| \leq \int_0^t \underbrace{\|v(\varphi(s))\|}_{\substack{\in K \\ \leq M}} ds \leq \delta_0 M < \frac{r}{M} M = r.$$

Damit ist  $\Phi$  eine Abbildung von  $A$  nach  $A$ ,  $\Phi : A \rightarrow A$ . Schließlich ist  $\Phi$  eine Kontraktion mit Kontraktionskonstante  $0 < \theta < 1$ ,  $\theta := \delta_0 L$ . Denn für alle  $\varphi, \psi \in A$  gilt

$$\begin{aligned} \|\Phi[\varphi] - \Phi[\psi]\|_\infty &= \sup_{|t| \leq \delta_0} \left\| \int_0^t (v(\varphi(s)) - v(\psi(s))) ds \right\| \leq \sup_{|t| \leq \delta_0} \int_0^t \|v(\varphi(s)) - v(\psi(s))\| ds \\ &\leq \sup_{|t| \leq \delta_0} \int_0^{\delta_0} L \|\varphi(s) - \psi(s)\| ds \leq L \delta_0 \|\varphi - \psi\|_\infty = \theta \|\varphi - \psi\|_\infty \end{aligned}$$

mit  $\theta = \delta_0 L < 1$ , weil  $\delta_0 < \frac{1}{L}$  ist.

Nach dem Banachschen Fixpunktsatz existiert also genau ein  $\varphi \in A$  mit  $\Phi[\varphi] = \varphi$ .

Setze nun  $x : (-\delta, \delta) \rightarrow G$ ,  $x(t) = \varphi(t)$ , wobei für  $t \in (-\delta, \delta)$  erst ein  $|t| < \delta_0 < \delta$  gewählt sei und  $\varphi$  dann der eindeutige Fixpunkt von  $\Phi = \Phi_{\delta_0}$  sei. Da jeder Fixpunkt von  $\Phi_{\delta_0}$  auch Fixpunkt von  $\Phi_{\delta_1}$  mit  $\delta_1 \leq \delta_0$  ist, hängt diese Definition nicht von der Wahl von  $\delta_0$  ab. Es folgt, dass  $x$  stetig ist und  $\forall t \in (-\delta, \delta)$  gilt:

$$x(t) = x_0 + \int_0^t v(x(s)) ds.$$

Also ist  $x$  sogar stetig differenzierbar mit

$$\dot{x}(t) = v(x(t)) \quad \text{und} \quad x(0) = x_0.$$

Also ist  $x$  Lösung von (\*). Da jede Lösung von (\*) auch (\*\*) erfüllt und die Lösung von (\*\*) ja nach dem Banachschen Fixpunktsatz eindeutig ist, ist  $x$  auch die eindeutige Lösung von (\*).  $\square$

**8.24 Bemerkung.** (a) Man beachte, dass der Existenzsatz nur die Existenz einer Lösung von  $\dot{x} = v(x)$ ,  $x(0) = x_0$  „für kurze Zeiten“ liefert,  $x : (-\delta, \delta) \rightarrow G$ . Man spricht daher von lokaler- bzw. Kurzzeitexistenz.

(b) Der Beweis gibt auch eine untere Schranke für die Lebensdauer der Lösung  $x$ , nämlich

$$\delta(x_0) = \min \left\{ \frac{1}{L}, \frac{r(x_0)}{M} \right\},$$

wobei  $r(x_0)$  so ist, dass  $K = \overline{B_r}(x_0) \subset G$  ist und  $M > 0$  eine Schranke für die „Geschwindigkeit“  $\|v\|$  auf  $K$  und  $L > 0$  eine Lipschitzkonstante (z.B. eine Schranke auf  $\|Dv\|$ , falls  $v$  stetig differenzierbar ist) auf  $K$  ist.

### 8.25 Definition. Maximale Lösung

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $v : G \rightarrow \mathbb{R}^n$  ein stetiges Vektorfeld. Eine Lösung  $x : I \rightarrow G$ ,  $I \subset \mathbb{R}$  ein offenes Intervall, von  $\dot{x} = v(x)$  auf  $G$  heißt **maximal**, wenn gilt: Ist  $\tilde{I} \subset \mathbb{R}$  ein Intervall mit  $\tilde{I} \supset I$  und  $\tilde{x} : \tilde{I} \rightarrow G$  eine Lösung von  $\dot{x} = v(x)$  mit  $\tilde{x}|_I = x$ , so gilt bereits  $\tilde{I} = I$  und somit  $\tilde{x} = x$ . Man sagt auch, eine maximale Lösung kann nicht fortgesetzt werden.

**8.26 Proposition.** Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $v : G \rightarrow \mathbb{R}^n$  ein lokal Lipschitz-stetiges Vektorfeld. Seien  $I, J \in \mathbb{R}$  offene Intervalle mit  $0 \in I \cap J$  und  $x : I \rightarrow G$  und  $y : J \rightarrow G$  Lösungen der Differentialgleichung  $\dot{x} = v(x)$  mit Anfangswert  $x(0) = y(0) = x_0 \in G$ . Dann gilt für alle  $t \in I \cap J : x(t) = y(t)$ .

*Beweis.* Sei  $A = \{t \in I \cap J \mid x(t) = y(t)\}$ . Wir zeigen, dass  $A = I \cap J$  ist. Sei dazu  $I = (t_-, t_+)$  und  $J = (s_-, s_+)$  und o.B.d.A.  $t_+ \leq s_+$  und

$$t_0 := \sup\{\tau \in (0, t_+) \mid x(t) = y(t) \forall 0 \leq t < \tau\}.$$

Nach dem Eindeutigkeitsatz existiert ein  $\delta_{x_0} > 0$ , sodass  $x(t) = y(t)$  für alle  $0 \leq t < \delta_{x_0}$ . Also ist  $0 < \delta \leq t_0 \leq t_+$ . Angenommen  $t_0 < t_+$ . Die Stetigkeit von  $x$  liefert dann  $x(t_0) = y(t_0) =: \tilde{x}_0$ . Wiederum nach Picard-Lindelöf existiert dann aber ein  $\delta_{\tilde{x}_0} > 0$  so, dass  $x(t) = y(t)$  für alle  $t \in (t_0 - \delta_{\tilde{x}_0}, t_0 + \delta_{\tilde{x}_0})$ , was im Widerspruch zur Definition vor  $t_0$  steht. Also ist  $t_0 = t_+$  und mit einem analogen Argument für die andere Intervallgrenze ergibt sich  $A = I \cap J$ .  $\square$

### 8.27 Satz. Existenz und Eindeutigkeit der maximalen Lösung

Sei  $G \subset \mathbb{R}^n$  ein Gebiet und  $v : G \rightarrow \mathbb{R}^n$  ein lokal Lipschitz-stetiges Vektorfeld. Dann existiert zu jedem  $x_0 \in G$  genau eine maximale Lösung  $x : I \rightarrow G$  von  $\dot{x} = v(x)$  mit  $x(0) = x_0$ .

*Beweis.* Ist  $y : J \rightarrow G$  eine Lösung von  $\dot{x} = v(x)$  mit  $y(0) = x_0$  und  $0 \in J$ , so notieren wir diese mit  $(y, J)$ . Sei nun

$$I := \bigcup_{(y, J) \text{ ist Lösung}} J \subset \mathbb{R}$$

Dann ist  $I \subset \mathbb{R}$  ein offenes Intervall. Ist  $t \in I$  beliebig, so wählen wir eine Lösung  $(y, J)$  mit  $t \in J$  und setzen  $x(t) := y(t)$ . Wegen Proposition 8.26 ist dies unabhängig von der Wahl  $(y, J)$ . Es ist dann die so definierte Funktion  $x : I \rightarrow G$  eine Lösung von  $\dot{x} = v(x)$  mit  $x(0) = x_0$ . Nach Konstruktion ist  $x$  maximal und zwar die einzige maximale Lösung.  $\square$

**8.28 Bemerkung.** (a) Wir schreiben für das Definitionsintervall  $I$  der maximalen Lösung von  $\dot{x} = v(x)$ ,  $x(0) = x_0$ ,

$$I(x_0) = (t_-(x_0), t_+(x_0)),$$

wobei  $t_-(x_0) \in [-\infty, 0]$  bzw.  $t_+(x_0) \in [0, \infty]$  die linke bzw. rechte Intervallgrenze bezeichnet.

(b) Wir setzen nun weiter

$$\Omega := \{(t, x) \in \mathbb{R} \times G \mid t \in I(x)\}$$

und  $\varphi : \Omega \rightarrow G$ ,  $(t, x) \mapsto \varphi^t(x) = x(t)$ , wobei  $I(x) \rightarrow G$ ,  $t \mapsto x(t)$  die maximale Lösung von  $\dot{x} = v(x)$  zum Anfangswert  $x \in G$  ist.

- (c) Man kann nun zeigen, dass  $\Omega$  offen ist und  $\varphi : \Omega \rightarrow G$  stetig ist. Man spricht von „stetiger Abhängigkeit von den Anfangsdaten“. Falls  $v : G \rightarrow \mathbb{R}^n$  stetig differenzierbar ist, so ist auch  $\varphi : \Omega \rightarrow G$  stetig differenzierbar.
- (d) Wir zeigen hier noch, dass  $\varphi$  einen Fluss definiert, also die Verträglichkeitsbedingung (b) aus Definition 8.2 erfüllt: Offenbar ist  $\varphi^0(x) = x$  für alle  $x \in G$ . Für  $x \in G$  und  $t \in I(x)$  lösen nun sowohl

$$\phi : (t_-(x) - t, t_+(x) - t) \rightarrow G, \quad \phi(s) = \varphi^{s+t}(x)$$

als auch

$$\psi : (t_-(\varphi^t(x)), t_+(\varphi^t(x))) \rightarrow G, \quad \psi(s) = \varphi^s(\varphi^t(x))$$

die Differentialgleichung  $\dot{x} = v(x)$  zum Anfangswert  $\varphi^t(x)$ . Außerdem sind beide Lösungen maximal und daher gilt:  $s \in I(\varphi^t(x)) \Leftrightarrow s + t \in I(x)$  und für diese  $s$  gilt  $\phi(s) = \psi(s)$  also  $\varphi^s(\varphi^t(x)) = \varphi^{s+t}(x)$ .

- (e) Ähnlich wie zuvor sagen wir, dass ein dynamisches System  $\varphi : \Omega \rightarrow G$  **maximal** ist, wenn gilt: Ist  $\tilde{\varphi} : \tilde{\Omega} \rightarrow G$  ein dynamisches System mit  $\tilde{\Omega} \supset \Omega$  und  $\tilde{\varphi}|_{\Omega} = \varphi$ , so muss bereits  $\tilde{\Omega} = \Omega$  sein und damit  $\tilde{\varphi} = \varphi$ .
- (f) Das dynamische System  $\varphi : \Omega \rightarrow G$ , das wir in (b) und (c) zu einem Vektorfeld  $v : G \rightarrow \mathbb{R}^n$  konstruiert haben, ist offenbar maximal. Sein assoziiertes Vektorfeld ist offenbar gerade  $v$ .

**8.29 Bemerkung.** Insgesamt ergibt sich also, dass jedem maximalen dynamischen System  $\varphi$  auf  $G$  ein stetiges Vektorfeld  $v = \frac{d}{dt}\varphi^t|_{t=0}$  zugeordnet werden kann. Umgekehrt liefert zumindest jedes stetig differenzierbare Vektorfeld  $v$  auf  $G$  ein maximales dynamisches System  $\varphi$  auf  $G$ .

Dabei ist der Übergang  $\varphi \mapsto v$  einfach („Differenzieren kann jeder“) der Übergang  $v \mapsto \varphi$  schwer („Integrieren ist eine Kunst“).

Im allgemeinen weiß man nicht einmal, ob bei gegebenem  $v : G \rightarrow \mathbb{R}^n$  und  $x \in G$  das Ende  $t_+(x) \in (0, \infty]$  endlich oder unendlich ist. Der folgende Satz besagt aber immerhin, dass die Bahnkurve  $t \mapsto x(t)$  eines Punktes  $x \in G$  mit  $t_+(x) < \infty$  jedes Kompaktum  $K$  in  $G$  verlassen muss, wenn  $t \rightarrow t_+(x)$  geht, d.h.  $t \mapsto x(t)$  strebt für  $t \rightarrow t_+(x)$  zum Rand von  $G$  oder nach Unendlich. Anders gesagt,  $t \mapsto x(t)$  kann sich in endlicher Zeit „nicht einfach in Luft auflösen“.

**8.30 Satz. Verhalten für  $t \rightarrow t_+(x)$**

Sei  $G \subset \mathbb{R}^n$  ein Gebiet,  $v : G \rightarrow \mathbb{R}^n$  ein lokal Lipschitz-stetiges Vektorfeld und  $x \in G$  so, dass  $t_+(x) < \infty$ . Ist dann  $K \subset G$  kompakt, so gibt es ein  $0 < \tau < t_+(x)$ , sodass für alle  $\tau < t < t_+(x)$  gilt:  $x(t) \notin K$ .

*Beweis.* Da  $K$  kompakt ist, gibt es ein  $\rho > 0$ , so dass  $B_\rho(x) \subset G$  für alle  $x \in K$  gilt. Denn die Distanzfunktion zum Rand von  $G$ ,  $\text{dist} : \mathbb{R}^n \rightarrow [0, \infty)$ ,

$$\text{dist}(x, \partial G) := \inf \{\|x - y\| \mid y \in \partial G\}$$

ist stetig und nimmt auf  $K$  ihr Minimum  $\rho$  an. Es ist  $\rho > 0$  denn sonst wäre  $K \cap \partial G \neq \emptyset$ , also  $K \not\subset G$ . Seien weiter  $\|v(x)\| \leq M$  für  $x \in \overline{B}_\rho(K)$  und  $L$  eine Lipschitzkonstante für  $v$  auf  $\overline{B}_\rho(K)$ . Dann gilt für  $t \in (0, t_+(x))$  mit  $x(t) \in K$ , dass

$$t_+(x) = t_+(x(t)) + t \geq \delta + t$$

mit  $\delta := \min \{\frac{1}{L}, \frac{\rho}{M}\}$ . Für alle  $t \in (\tau, t_+(x))$  mit  $\tau := t_+(x) - \delta$  muss also gelten, dass  $x(t) \notin K$ . □

**8.31 Bemerkung.** (a) Eine entsprechende Aussage gilt natürlich, wenn  $t_-(x) > -\infty$  ist.

(b) Bleibt eine Lösungskurve  $t \mapsto x(t)$  in einem Kompaktum, z.B. wenn  $\lim_{t \rightarrow t_+} x(t) = p$  für ein  $p \in G$  ist, so muss also  $t_+(x) = \infty$  sein. In diesem Fall muss dann  $p$  eine Gleichgewichtslage sein.

**8.32 Beispiel. Das  $N$ -Körper Problem**

Betrachte  $N$  Punktteilchen im  $\mathbb{R}^3$  die über das Gravitationspotential

$$V(q_i - q_j) = -\frac{1}{|q_i - q_j|}$$

wechselwirken (Masse und Gravitationskonstante = 1). Der Phasenraum des Systems ist der  $\mathbb{R}^{6N}$  und wir schreiben

$$z = (q, p) \in \mathbb{R}^{6N}.$$

Die zugehörige Differentialgleichung ist

$$\dot{z} = v(z)$$

mit

$$v(q, p) = \left( p, \nabla \sum_{i < j} \frac{1}{|q_i - q_j|} \right) = \left( p, -\sum_{i < j} \frac{q_i - q_j}{|q_i - q_j|^3} \right).$$

Es ist  $v$  lokal Lipschitz-stetig auf dem Gebiet

$$G = \{(q, p) \in \mathbb{R}^{6N} \mid q_i \neq q_j \ \forall i \neq j\}.$$

Für  $N = 2$  kann man die Lösungen explizit berechnen (Keplerproblem).

Für  $N = 3$  existieren die Lösungen für „fast alle“ Anfangsdaten global, man kennt sie aber nicht.

Für  $N > 3$  würde man gerne zeigen, dass die Lösungen für „fast alle“ Anfangsdaten existieren.

**Problem:** Für alle  $N \geq 2$  gibt es Lösungen, die nicht global existieren, nämlich die direkte Kollision zweier Teilchen. Für  $N \geq 5$  ist bekannt, dass Lösungen in endlicher Zeit nach unendlich laufen können.

Eine naheliegende mathematische Frage ist also: Wieviel „schlechte“ Anfangsdaten gibt es ?

Das quantenmechanische  $N$ -Körper Problem ist übrigens kein Problem mehr. Die Lösungen der entsprechenden Schrödingergleichung existieren für alle Zeiten, die zugehörigen Teilchentrajektorien in der Bohmschen Mechanik existieren für fast alle Anfangsdaten global.



# 9 Lineare Differentialgleichungen

## 9.1 Definition. Lineare Systeme

Sei  $I \subset \mathbb{R}$  ein offenes Intervall und  $A : I \rightarrow M(n, \mathbb{R})$  eine stetige Abbildung.

(a) Man nennt dann die Differentialgleichung

$$\dot{x} = A(t)x$$

ein **nicht-autonomes, homogenes, lineares System.**

(b) Ist  $b : I \rightarrow \mathbb{R}^n$  stetig, so heißt

$$\dot{x} = A(t)x + b(t)$$

ein **nicht-autonomes, inhomogenes, lineares System.**

**9.2 Motivation.** Aus der linearen Algebra wissen wir, dass im autonom homogenen Fall

$$\dot{x} = Ax$$

die globale Lösung durch  $x(t) = e^{At}x(0)$  gegeben ist. Es liegt nahe, dass auch im nicht-autonomen linearen Fall die Lösungen höchstens exponentiell wachsen können, falls  $\|A(t)\|$  beschränkt bleibt. Da sie dann aber nicht in endlicher Zeit nach Unendlich laufen können, existieren sie gemäß Satz 8.30 für alle Zeiten in  $I$ .

## 9.3 Satz. Globale Existenz für lineare Systeme

Sei  $I \subset \mathbb{R}$  offenes Intervall,  $A : I \rightarrow M(n, \mathbb{R})$  und  $b : I \rightarrow \mathbb{R}^n$  seien stetig. Dann existiert zu jedem  $t_0 \in I$  und  $x_0 \in \mathbb{R}^n$  eine eindeutige maximale Lösung  $x : I \rightarrow \mathbb{R}^n$  von

$$\dot{x} = A(t)x + b(t) \quad \text{mit} \quad x(t_0) = x_0.$$

*Beweis.* Wir führen den Beweis nur für lokal Lipschitz-stetiges  $A(t)$  und  $b(t)$  aus. Sonst müssten wir im Beweis nochmals die Picard-Iteration für dieses spezielle Problem durchgehen, worauf wir aus zeitgründen verzichten. Gemäß Bemerkung 8.13 liefert die Anwendung von Satz 8.27 auf das lokal Lipschitz-stetige Vektorfeld (hier verwenden wir die Zusatzannahme, dass  $A(t)$  und  $b(t)$  auch lokal Lipschitz-stetig sind!)

$$v(t, x) = \begin{pmatrix} 1 \\ A(t)x + b(t) \end{pmatrix}$$

die Existenz einer eindeutigen maximalen Lösung auf dem Zeitintervall  $J := (s_-, s_+) \subset I =: (t_-, t_+)$ . Wir zeigen nun  $s_+ = t_+$  (analog sieht man dann  $s_- = t_-$ , also  $I = J$ ). Angenommen  $s_+ < t_+$ , dann muss nach Satz 8.30 die maximale Lösung  $(s_-, s_+) \rightarrow I \times \mathbb{R}^n$ ,  $t \mapsto (t, x(t))$  jedes Kompaktum von  $I \times \mathbb{R}^n$  verlassen, also  $\|x(t)\| \rightarrow \infty$  für  $t \rightarrow s_+$  gelten. Wir zeigen nun, dass  $\|x(t)\|$  beschränkt bleibt für  $t \rightarrow s_+$  falls  $s_+ < t_+$ . Insgesamt folgt dann  $s_+ = t_+$ .

Setze dazu  $L := \max\{\|A(t)\| \mid t_0 \leq t \leq s_+\}$  und  $M := \max\{\|b(t)\| \mid t_0 \leq t \leq s_+\}$ , so gilt für die

## 9 Lineare Differentialgleichungen

stetige Funktion  $u : [t_0, s_+) \rightarrow [0, \infty)$ ,  $u(t) = \|x(t)\|$ , dass

$$\begin{aligned} u(t) &= \|x(t)\| = \|x(t_0) + \int_{t_0}^t \dot{x}(s) ds\| \leq \|x(t_0)\| + \int_{t_0}^t \|\dot{x}(s)\| ds \\ &= u(t_0) + \int_{t_0}^t \|A(s)x(s) + b(s)\| ds \\ &\leq \left( u(t_0) + \int_{t_0}^{s_+} \|b(s)\| ds \right) + \int_{t_0}^t \underbrace{\|A(s)\|}_{\leq L} \cdot \underbrace{\|x(s)\|}_{=u(s)} ds \\ &\leq C + L \int_{t_0}^t u(s) ds. \end{aligned}$$

### 9.4 Lemma. von Gronwall

Sei  $a < b$  und  $u : [a, b] \rightarrow [0, \infty)$  eine stetige Funktion. Es gebe Konstanten  $L, C \geq 0$  so, dass für alle  $t \in [a, b]$  gilt

$$u(t) \leq C + L \int_a^t u(s) ds.$$

Dann ist

$$u(t) \leq C e^{L(t-a)} \quad \text{für alle } t \in [a, b].$$

*Ende des Beweises von Satz 9.3.* Also ist  $u(t) \leq C e^{L(t-t_0)} \leq C e^{L(s_+-t_0)}$  für alle  $t_0 \leq t < s_+$  und somit  $t \mapsto \|x(t)\|$  beschränkt für  $t \rightarrow s_+$ .  $\square$

*Beweis. des Gronwall Lemmas.* Sei zunächst  $C > 0$ . Setze  $U : [a, b] \rightarrow [0, \infty)$

$$U(t) := C + L \int_a^t u(s) ds.$$

Dann ist  $U$  stetig differenzierbar und  $\dot{U}(t) = L u(t) \geq 0$ , also monoton steigend. Außerdem ist  $U(a) = C > 0$ , also  $U(t) \geq C$  für alle  $t \in [a, b]$ . Es folgt

$$\frac{d}{dt} \ln(U(t)) = \frac{\dot{U}(t)}{U(t)} = \frac{L u(t)}{U(t)} \leq L$$

da  $u \leq U$ . Also ist

$$\ln U(t) - \ln C = \ln U(t) - \ln U(a) = \int_a^t \frac{d}{ds} \ln U(s) ds \leq \int_a^t L ds = L(t-a).$$

und somit

$$u(t) \leq U(t) = \exp(\ln U(t)) \leq \exp(\ln C + L(t-a)) = C e^{L(t-a)}.$$

Falls  $C = 0$ , so liefert unser Argument, dass  $u(t) \leq \tilde{C} e^{Lt}$  für jedes  $\tilde{C} > 0$ , also  $u(t) \equiv 0$ .  $\square$

Die wichtigste Eigenschaft homogener linearer Systeme ist, dass Linearkombinationen von Lösungen wieder Lösungen sind.

### 9.5 Satz. Der Lösungsraum homogener linearer Systeme

Sei  $I \subset \mathbb{R}$  ein offenes Intervall und  $A : I \rightarrow M(n, \mathbb{R})$  stetig. Sei  $L_h \subset C^1(I, \mathbb{R}^n)$  die Menge aller maximalen Lösungen des homogenen linearen Systems

$$\dot{x} = A(t)x \quad (*)$$

auf  $\mathbb{R}^n$ . Dann gilt:

- (a)  $L_h$  ist ein  $n$ -dimensionaler Unterraum von  $C^1(I, \mathbb{R}^n)$ .
- (b) Ist  $t_0 \in I$ ,  $r \in \mathbb{N}$  und  $\xi_1, \dots, \xi_r \in \mathbb{R}^n$ , so seien  $x_1, \dots, x_r \in L_h$  die Lösungen von  $\dot{x} = A(t)x$  mit  $x_j(t_0) = \xi_j$ ,  $j = 1, \dots, r$ . Dann sind äquivalent:
- (i)  $(x_1, \dots, x_r)$  ist linear unabhängig in  $L_h$ .
  - (ii)  $(x_1(t), \dots, x_r(t))$  ist linear unabhängig in  $\mathbb{R}^n$  für alle  $t \in I$ .
  - (iii)  $(\xi_1, \dots, \xi_r)$  ist linear unabhängig in  $\mathbb{R}^n$ .

*Beweis.* Zu (a): Seien  $x, y \in L_h$  und  $\lambda \in \mathbb{R}$ . Dann ist

$$\frac{d}{dt}(x + y) = \dot{x} + \dot{y} = A(t)x + A(t)y = A(t)(x + y)$$

und

$$\frac{d}{dt}(\lambda x) = \lambda \dot{x} = \lambda A(t)x = A(t)(\lambda x).$$

Also sind auch  $x + y$  und  $\lambda x$  Lösungen von (\*) und somit ist  $L_h$  ein Untervektorraum. Die Aussage zur Dimension folgt sofort aus Teil (b).

Zu (b): (i)  $\Rightarrow$  (ii): Seien  $(x_1, \dots, x_r) \subseteq L_h$  linear unabhängig und sei  $\tilde{t} \in I$  beliebig aber fest. Erfüllen dann  $\lambda_1, \dots, \lambda_r \in \mathbb{R}$  die Gleichung

$$\lambda_1 x_1(\tilde{t}) + \dots + \lambda_r x_r(\tilde{t}) = 0 \quad (**)$$

als Gleichung in  $\mathbb{R}^n$ , so ist zu zeigen, dass aus (\*\*) schon  $\lambda_j = 0$  für  $j = 1, \dots, r$  folgt. Setze dazu  $x : I \rightarrow \mathbb{R}^n$ ,  $x(t) := \lambda_1 x_1(t) + \dots + \lambda_r x_r(t)$ . Nach (a) ist  $x(t)$  Lösung von (\*) und wegen (\*\*) ist  $x(\tilde{t}) = 0$  und somit, wegen der Eindeutigkeit der Lösung,  $x(t) = 0$  für alle  $t \in I$ . Also ist

$$\lambda_1 x_1 + \dots + \lambda_r x_r = 0$$

in  $C^1(I, \mathbb{R}^n)$  und nach Voraussetzung  $\lambda_1 = \dots = \lambda_r = 0$ . Es ist also  $(x_1(\tilde{t}), \dots, x_r(\tilde{t})) \subset \mathbb{R}^n$  linear unabhängig für alle  $\tilde{t} \in I$ .

- (ii)  $\Rightarrow$  (iii): klar, weil  $\xi_j = x_j(t_0)$  ist,  
 (iii)  $\Rightarrow$  (i): auch klar, denn ist

$$\lambda_1 x_1 + \dots + \lambda_r x_r = 0$$

so ist insbesondere

$$\lambda_1 \xi_1 + \dots + \lambda_r \xi_r = \lambda_1 x_1(t_0) + \dots + \lambda_r x_r(t_0) = 0.$$

Also muss nach Voraussetzung  $\lambda_1 = \dots = \lambda_r = 0$  sein. □

### 9.6 Korollar. Der Propagator

Sei  $A : I \rightarrow M(n, \mathbb{R})$  stetig,  $t_0 \in I$  und  $\varphi^t : \mathbb{R}^n \rightarrow \mathbb{R}^n$  der zu  $\dot{x} = A(t)x$  gehörige Fluss (vgl. Bemerkung 8.2 (c)) zum Anfangszeitpunkt  $t_0$ : Sei  $x : I \rightarrow \mathbb{R}^n$  die Lösung von  $\dot{x} = A(t)x$  zum Anfangswert  $x(t_0) = x_0$ , so ist also  $\varphi^t(x_0) = x(t)$ .

Dann ist  $\varphi^t$  für jedes feste  $t \in I$  ein linearer Isomorphismus des  $\mathbb{R}^n$ , welcher im Folgenden auch mit  $\Phi(t)$  bezeichnet und **Propagator** genannt wird.

Die Abbildung  $\Phi : \mathbb{R} \rightarrow L(\mathbb{R}^n)$  erfüllt die Matrix Differentialgleichung

$$\dot{\Phi}(t) = A(t)\Phi(t) \quad \text{mit Anfangsbedingung} \quad \Phi(t_0) = \text{Id}.$$

*Beweis.* Übungsaufgabe. □

### 9.7 Definition. Fundamentalsystem

Sei  $A : I \rightarrow M(n, \mathbb{R})$  stetig. Ein  $n$ -Tupel  $(x_1, \dots, x_n) \subset L_h$  von Lösungen der homogenen Gleichung  $\dot{x} = A(t)x$  heißt **Lösungs-Fundamentalsystem**, wenn  $(x_1, \dots, x_n)$  eine Basis von  $L_h$  ist.

**9.8 Satz. Der Lösungsraum inhomogener linearer Systeme**

Sei  $I \subset \mathbb{R}$  ein offenes Intervall und  $A : I \rightarrow M(n, \mathbb{R})$  und  $b : I \rightarrow \mathbb{R}^n$  seien stetig. Mit  $L_i \subset C^1(I, \mathbb{R}^n)$  werde die Teilmenge aller Lösungen der inhomogenen Gleichung

$$\dot{x} = A(t)x + b(t)$$

bezeichnet, also

$$L_i = \{x \in C^1(I, \mathbb{R}^n) \mid \dot{x} = A(t)x + b(t)\}$$

und

$$L_h = \{x \in C^1(I, \mathbb{R}^n) \mid \dot{x} = A(t)x\}.$$

(a) Ist  $x : I \rightarrow \mathbb{R}^n$  eine Lösung der inhomogenen Gleichung, so gilt

$$L_i = x + L_h := \{x + \varphi \in C^1(I, \mathbb{R}^n) \mid \varphi \in L_h\}.$$

(b) **Variation der Konstanten**

Sei  $\Phi(t) : \mathbb{R}^n \rightarrow \mathbb{R}^n$  der Propagator des homogenen Systems zur Anfangszeit  $t_0 \in I$  und sei  $x_0 \in \mathbb{R}^n$ . Dann ist

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

die Lösung der inhomogenen Gleichung  $\dot{x} = A(t)x + b(t)$  zum Anfangswert  $x(t_0) = x_0$ .

*Beweis.* (a) Ist  $x$  eine Lösung von  $\dot{x} = A(t)x + b(t)$ , so ist  $\tilde{x}$  genau dann eine weitere Lösung von  $\dot{x} = A(t)x + b(t)$ , wenn  $\varphi := \tilde{x} - x$  Lösung der homogenen Gleichung ist:

$$\dot{\tilde{x}} = A(t)\tilde{x} + b(t) \quad \Leftrightarrow \quad \dot{\varphi} = \dot{\tilde{x}} - \dot{x} = A(t)\tilde{x} - A(t)x = A(t)\varphi.$$

(b) Es gilt  $x(t_0) = \Phi(t_0)x_0 = x_0$  und

$$\begin{aligned} \dot{x}(t) &= \dot{\Phi}(t) \left( x_0 + \int_{t_0}^t \Phi(s)^{-1} b(s) ds \right) + \Phi(t) \Phi(t)^{-1} b(t) \\ &= A(t)\Phi(t) \left( x_0 + \int_{t_0}^t \Phi(s)^{-1} b(s) ds \right) + b(t) \\ &= A(t)x(t) + b(t). \end{aligned}$$

□

**9.9 Bemerkung.** Die Bezeichnung „Variation der Konstanten“ kommt aus folgendem Ansatz für die Lösung:  $x(t) = \Phi(t)c(t)$  statt  $x(t) = \Phi(t)x_0$  wie für die homogene Gleichung. Dann muss nämlich  $c(t)$  folgende Differentialgleichung erfüllen:

$$\dot{x} = \dot{\Phi}c + \Phi\dot{c} = A\Phi c + \Phi\dot{c} = x + \Phi\dot{c} \stackrel{!}{=} x + b(t)$$

also  $\Phi\dot{c} = b$  oder  $\dot{c}(t) = \Phi^{-1}(t)b(t)$ . Somit folgt

$$c(t) = x_0 + \int_{t_0}^t \Phi^{-1}(s)b(s) ds.$$

**9.10 Bemerkung.** Die Menge aller Lösungen des homogenen Systems  $\dot{x} = A(t)x$  ist ein  $n$ -dimensionaler Unterraum von  $C^1(I, \mathbb{R}^n)$ . Die Menge aller Lösungen des inhomogenen Systems ist ein  $n$ -dimensionaler affiner Unterraum von  $C^1(I, \mathbb{R}^n)$ . Kennt man die vollständige Lösung des homogenen Systems (also zumindest  $n$  linear unabhängige Lösungen), so kann man auch das inhomogene System vollständig lösen.

**9.11 Beispiel.** Betrachte die Differentialgleichung

$$\dot{x} = 2tx + t^3 \quad (*)$$

auf  $\mathbb{R}$ . Mit  $a(t) = 2t$  und  $b(t) = t^3$  wird das zu  $\dot{x} = a(t)x + b(t)$ . Die Lösung der homogenen Gleichung  $\dot{x} = a(t)x$  mit  $x(t_0) = x_0$  ist

$$x(t) = e^{\int_{t_0}^t a(s) ds} x_0,$$

also mit  $t_0 = 0$

$$x(t) = e^{\int_0^t 2(s) ds} x_0 = e^{t^2} x_0.$$

Da  $\dim L_h = 1$  ist, ist  $x(t) = e^{t^2} x_0$  eine Basis für  $L_h$  und der Propagator ist  $\Phi(t) = e^{t^2}$ . Variation der Konstanten liefert nun die Lösung der inhomogenen Gleichung zum Anfangswert  $x_0$  durch

$$x_1(t) = e^{t^2} \left( x_0 + \int_0^t e^{-s^2} s^3 ds \right) = e^{t^2} \left( x_0 + \frac{1}{2} \right) - \frac{1}{2} (1 + t^2),$$

wobei

$$\int_0^t e^{-s^2} s^3 ds = \frac{1}{2} \int_0^{t^2} e^{-\tau} \tau d\tau = -\frac{1}{2} e^{-\tau} \tau \Big|_0^{t^2} + \frac{1}{2} \int_0^{t^2} e^{-\tau} d\tau = -\frac{1}{2} e^{-t^2} t^2 - \frac{1}{2} e^{-t^2} + \frac{1}{2}.$$

**9.12 Bemerkung. Dyson Reihe**

Auch im Fall  $n > 1$  kann man noch eine explizite Formel für die Lösung der homogenen Gleichung angeben, die sogenannte Dyson Reihe. Sei  $A : I \rightarrow M(n, \mathbb{R})$  stetig, dann ist die Lösung  $\Phi : I \rightarrow M(n, \mathbb{R})$  von

$$\dot{\Phi} = A(t)\Phi$$

mit  $\Phi(t_0) = E_n$  gegeben durch die absolut konvergente Reihe

$$\Phi(t) = E_n + \sum_{j=1}^{\infty} \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \cdots \int_{t_0}^{\tau_{j-1}} d\tau_j A(\tau_1) \cdots A(\tau_j).$$

*Beweis.* Übungsaufgabe. □

**9.13 Bemerkung.** Für den autonomen Fall  $A(t) \equiv A$  erhält man die Lösung  $\Phi : \mathbb{R} \rightarrow M(n, \mathbb{R})$  von

$$\dot{\Phi} = A\Phi$$

mit  $\Phi(0) = E_n$  durch exponentieren

$$\Phi(t) = e^{At}.$$

Insbesondere existieren die Lösungen für alle Zeiten. Um  $e^{At} := \sum_{j=0}^{\infty} \frac{(At)^j}{j!}$  explizit auszurechnen, transformiert man  $A$  auf Jordansche Normalform (vgl. Lineare Algebra).

Sei z.B.  $A$  diagonalisierbar, also

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = D$$

eine Diagonalmatrix und  $S \in GL_n(\mathbb{R})$ , so löst  $\Psi = S^{-1}\Phi S$  die Gleichung

$$\dot{\Psi} = S^{-1}\dot{\Phi}S = S^{-1}A\Phi S = S^{-1}AS S^{-1}\Phi S = D\Psi,$$

## 9 Lineare Differentialgleichungen

also

$$\Psi(t) = e^{Dt} = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix}$$

und somit

$$\Phi(t) = S \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} S^{-1}.$$

Um die Normalformtheorie von Matrizen über  $\mathbb{C}$  benutzen zu können (viele Matrizen haben eben keine reellen Eigenwerte, aber komplexe), betrachtet man die Differentialgleichung  $\dot{x} = A(t)x$  statt auf  $\mathbb{R}^n$  auf  $\mathbb{C}^n$  und sucht Lösungen  $z : I \rightarrow \mathbb{C}^n$  von

$$\dot{z} = A(t)z.$$

Die Zeitvariable bleibt aber reell! Die Lösungen bilden dann einen Unterraum der komplexen Dimension  $n$  in  $C^1(I, \mathbb{C}^n)$  und alles in diesem Kapitel gesagte gilt analog.

Beachte, dass für  $A(t) \in M(n, \mathbb{R})$  zwar  $S$  und  $\lambda_1, \dots, \lambda_n$  komplex sein können, der Propagator

$$\Phi(t) = S \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} S^{-1}$$

ist aber immer noch reell.

Die Resultate zu linearen Differentialgleichungssystemen 1. Ordnung lassen sich direkt auf lineare (Systeme)  $n$ -ter Ordnung übertragen:

### 9.14 Definition. Homogene lineare DGL $m$ -ter Ordnung

Sei  $I \subset \mathbb{R}$  ein offenes Intervall und  $a_k : I \rightarrow \mathbb{K}$ ,  $k = 1, \dots, m-1$ , stetige Funktionen. Hier ist  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Dann heißt

$$x^{(m)}(t) + a_{m-1}(t)x^{(m-1)}(t) + \dots + a_1(t)\dot{x}(t) + a_0(t)x(t) = 0$$

homogene lineare Differentialgleichung  $m$ -ter Ordnung.

Ist  $b : I \rightarrow \mathbb{K}$  stetig, so heißt

$$x^{(m)}(t) + a_{m-1}(t)x^{(m-1)}(t) + \dots + a_1(t)\dot{x}(t) + a_0(t)x(t) = b(t)$$

inhomogene lineare Differentialgleichung  $m$ -ter Ordnung.

### 9.15 Satz. Lösungsraum linearer DGLen $m$ -ter Ordnung

(a) Sei  $L_h$  die Menge aller Lösungen  $x : I \rightarrow \mathbb{K}$  der homogenen Gleichung

$$x^{(m)} + a_{m-1}x^{(m-1)} + \dots + a_1\dot{x} + a_0x = 0.$$

Dann ist  $L_h$  ein  $m$ -dimensionaler Unterraum von  $C^m(I, \mathbb{K})$ .

(b) Sei  $L_i$  die Menge aller Lösungen  $x : I \rightarrow \mathbb{K}$  der inhomogenen Differentialgleichung

$$x^{(m)} + a_{m-1}x^{(m-1)} + \dots + a_1\dot{x} + a_0x = b.$$

Dann gilt für beliebiges  $\tilde{x} \in L_i$

$$L_i = \tilde{x} + L_h.$$

(c) Ein  $m$ -Tupel  $(\varphi_1, \dots, \varphi_m)$  in  $L_h$  ist genau dann linear unabhängig, wenn für ein und damit für alle  $t \in I$  die „Wronski-Determinante“

$$W(t) := \det \begin{pmatrix} \varphi_1(t) & \cdots & \varphi_m(t) \\ \dot{\varphi}_1(t) & \cdots & \dot{\varphi}_m(t) \\ \vdots & & \vdots \\ \varphi_1^{(m-1)}(t) & \cdots & \varphi_m^{(m-1)}(t) \end{pmatrix}$$

von Null verschieden ist.

*Beweis.* Die Differentialgleichung

$$x^{(m)} + a_{m-1}x^{(m-1)} + \dots + a_1\dot{x} + a_0x = b \quad (*)$$

ist äquivalent zu dem inhomogenen linearen System 1. Ordnung

$$\begin{aligned} \dot{y}_0 &= y_1 \\ \dot{y}_1 &= y_2 \\ &\vdots \\ \dot{y}_{m-2} &= y_{m-1} \\ \dot{y}_{m-1} &= -a_0y_0 - a_1y_1 - \dots - a_{m-1}y_{m-1} + b \end{aligned} \quad (**)$$

Jeder Lösung  $x : I \rightarrow \mathbb{K}$  von  $(*)$  entspricht eine Lösung

$$\begin{pmatrix} x \\ \dot{x} \\ \vdots \\ x^{(m-1)} \end{pmatrix} : I \rightarrow \mathbb{K}^m$$

von  $(**)$  und umgekehrt. Entsprechendes gilt für die homogenen Gleichungen ( $b = 0$ ). Damit folgen die Behauptungen aus Satz 9.8.  $\square$

**9.16 Beispiel.** Die Differentialgleichung

$$\ddot{x} - \frac{1}{2t}\dot{x} + \frac{1}{2t^2}x = 0$$

auf dem Intervall  $I = (0, \infty)$  besitzt die Lösungen  $\varphi_1(t) := t$  und  $\varphi_2(t) := \sqrt{t}$ , wovon man sich durch Einsetzen überzeugt.

Die Wronski-Determinante von  $(\varphi_1, \varphi_2)$  ist

$$W(t) = \det \begin{pmatrix} \varphi_1(t) & \varphi_2(t) \\ \dot{\varphi}_1(t) & \dot{\varphi}_2(t) \end{pmatrix} = \det \begin{pmatrix} t & \sqrt{t} \\ 1 & \frac{1}{2\sqrt{t}} \end{pmatrix} = -\frac{\sqrt{t}}{2}.$$

Da  $W(t) \neq 0$  für  $t \in I$ , bilden  $\varphi_1$  und  $\varphi_2$  ein Lösungs-Fundamentalsystem. Die allgemeine Lösung der Differentialgleichung ist also

$$\varphi(t) = c_1\varphi_1(t) + c_2\varphi_2(t) = c_1t + c_2\sqrt{t}$$

mit beliebige Konstanten  $c_1, c_2 \in \mathbb{C}$ .

**9.17 Beispiele.** (a) Die Legendresche Differentialgleichung

Die Legendresche Differentialgleichung auf  $I = (-1, 1)$  zu  $n \in \mathbb{N}$  ist

$$(1 - t^2)\ddot{x} - 2t\dot{x} + n(n + 1)x = 0$$

bzw., da  $(1 - t^2) \neq 0$ ,

$$\ddot{x} - \frac{2t}{(1 - t^2)}\dot{x} + \frac{n(n + 1)}{(1 - t^2)}x = 0.$$

Das Legendre Polynom der Ordnung  $n$  ist definiert durch

$$P_n(t) := \frac{1}{2^n n!} \left( \frac{d}{dt} \right)^n (t^2 - 1)^n$$

und löst die Legendresche Differentialgleichung der Ordnung  $n$ .

(b) Die Hermitesche Differentialgleichung

Die Hermitesche Differentialgleichung auf  $I = \mathbb{R}$  zu  $n \in \mathbb{N}$  ist

$$\ddot{x} - 2t\dot{x} + 2nx = 0.$$

Das Hermite Polynom der Ordnung  $n$  ist definiert durch

$$H_n(t) = (-1)^n e^{t^2} \left( \frac{d}{dt} \right)^n e^{-t^2}$$

und löst die Hermitesche Differentialgleichung der Ordnung  $n$ .

(c) Die Laguerresche Differentialgleichung

Die Laguerresche Differentialgleichung auf  $I = (0, \infty)$  zu  $n \in \mathbb{N}$  ist

$$t\ddot{x} + (1 - t)\dot{x} + nx = 0.$$

Das Laguerresche Polynom der Ordnung  $n$  ist definiert durch

$$L_n(t) := e^t \left( \frac{d}{dt} \right)^n (t^n e^{-t})$$

und löst die Laguerresche Differentialgleichung der Ordnung  $n$ .

In allen drei Fällen (a), (b) und (c) sind die Polynome jeweils nur eine spezielle Lösung und gemäß Satz 9.15 gibt es jeweils noch eine weitere linear unabhängige Lösung.

**9.18 Satz.** Sei  $\varphi : I \rightarrow \mathbb{R}$  eine Lösung der homogenen linearen Differentialgleichung 2. Ordnung

$$\ddot{x} + a(t)\dot{x} + b(t)x = 0.$$

Im Intervall  $J \subset I$  gelte  $\varphi(t) \neq 0$ . Dann erhält man über  $J$  eine zweite von  $\varphi$  linear unabhängige Lösung  $\psi : J \rightarrow \mathbb{R}$  durch den Ansatz

$$\psi(t) = \varphi(t)u(t),$$

wobei  $u$  eine nicht-konstante Lösung der Differentialgleichung

$$\ddot{u} + \left( 2 \frac{\dot{\varphi}(t)}{\varphi(t)} + a(t) \right) \dot{u} = 0 \quad (*)$$

ist.

**9.19 Bemerkung.** Die Gleichung (\*) ist eine lineare Differentialgleichung 1. Ordnung für  $\dot{u}$ , welche die Lösung

$$\begin{aligned}\dot{u}(t) &= \dot{u}(t_0) e^{-\int_{t_0}^t \left(2\frac{\dot{\varphi}(s)}{\varphi(s)} + a(s)\right) ds} \\ &= \dot{u}(t_0) e^{-2(\ln\varphi(t) - \ln\varphi(t_0))} e^{-\int_{t_0}^t a(s) ds} \\ &= \dot{u}(t_0) \frac{\varphi(t_0)^2}{\varphi(t)^2} e^{-\int_{t_0}^t a(s) ds}\end{aligned}$$

hat. Man erhält  $u$  dann durch eine weitere Integration.

*Beweis. von Satz 9.18.*

$$\psi = \varphi u \quad \Rightarrow \quad \dot{\psi} = \dot{\varphi} u + \varphi \dot{u} \quad \Rightarrow \quad \ddot{\psi} = \ddot{\varphi} u + 2\dot{\varphi} \dot{u} + \varphi \ddot{u}$$

Also

$$\ddot{\psi} + a\dot{\psi} + b\psi = \ddot{\varphi} u + 2\dot{\varphi} \dot{u} + \varphi \ddot{u} + a\dot{\varphi} u + a\varphi \dot{u} + b\varphi u = 2\dot{\varphi} \dot{u} + \varphi \ddot{u} + a\varphi \dot{u},$$

da  $\ddot{\varphi} + a\dot{\varphi} + b\varphi = 0$ . Somit löst  $\psi$  die Differentialgleichung, wenn  $\varphi \ddot{u} + 2\dot{\varphi} \dot{u} + a\varphi \dot{u} = 0$  bzw. wenn

$$\ddot{u} + \left(2\frac{\dot{\varphi}}{\varphi} + a\right) \dot{u} = 0.$$

Ist  $u$  nicht konstant, so sind  $\psi = u\varphi$  und  $\varphi$  auf  $J$  linear unabhängig. □

**9.20 Beispiel.** Für  $n = 1$  ist die Legendresche Differentialgleichung auf  $I = (-1, 1)$

$$\ddot{x} - \frac{2t}{1-t^2} \dot{x} + \frac{2}{1-t^2} x = 0.$$

Sie hat die Lösung  $P_1(t) = t$ . Also erhält man auf  $(0, 1)$  eine zweite Lösung durch den Ansatz  $\psi(t) = t u(t)$ , wobei

$$\begin{aligned}\dot{u}(t) &= \dot{u}(t_0) \frac{t_0^2}{t^2} e^{\int_{t_0}^t \frac{2s}{1-s^2} ds} = \dot{u}(t_0) \frac{t_0^2}{t^2} e^{-\ln(1-s^2)|_{t_0}^t} = \dot{u}(t_0) \frac{t_0^2}{t^2} \frac{1-t_0^2}{1-t^2} \\ &= \dot{u}(t_0) t_0^2 (1-t_0^2) \left( \frac{1}{t^2} + \frac{1}{2} \left( \frac{1}{1+t} + \frac{1}{1-t} \right) \right)\end{aligned}$$

somit ist

$$\begin{aligned}u(t) &= u(t_0) + \int_{t_0}^t \dot{u}(s) ds = u(t_0) + \dot{u}(t_0) t_0^2 (1-t_0^2) \left( -\frac{1}{s} \Big|_{t_0}^t + \frac{1}{2} (\ln(1+s) - \ln(1-s)) \Big|_{t_0}^t \right) \\ &= u(t_0) + \dot{u}(t_0) t_0^2 (1-t_0^2) \left( \frac{1}{t_0} - \frac{1}{t} + \frac{1}{2} \ln \frac{1+t}{1-t} \right).\end{aligned}$$

Wir wählen nun für beliebiges  $t_0 \in (0, 1)$

$$\dot{u}(t_0) = \frac{1}{t_0^2(1-t_0^2)} \quad \text{und} \quad u(t_0) = -\dot{u}(t_0) t_0 (1-t_0^2)$$

und erhalten so die Lösung

$$u(t) = \frac{1}{2} \ln \frac{1+t}{1-t} - \frac{1}{t}.$$

Es ist also

$$\psi(t) = t u(t) = \frac{t}{2} \ln \frac{1+t}{1-t} - 1$$

eine von  $\varphi(t)$  linear unabhängige Lösung auf dem Intervall  $(0, 1)$ . Man rechnet nun aber direkt nach, dass  $\psi(t)$  die Differentialgleichung auf dem ganzen Intervall  $(-1, 1)$  löst.

**9.21 Bemerkung. Lineare Differentialgleichungen mit konstanten Koeffizienten**

Eine homogene lineare Differentialgleichung  $m$ -ter Ordnung ist von der Form

$$\sum_{j=0}^m a_j x^{(j)}(t) = 0 \quad \text{mit } a_m = 1, \quad (*)$$

für  $x : \mathbb{R} \rightarrow \mathbb{R}$  oder  $x : \mathbb{R} \rightarrow \mathbb{C}$ . Sie ist äquivalent zu dem System erster Ordnung

$$\dot{y}(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ -a_0 & -a_1 & \cdots & -a_{m-2} & -a_{m-1} \end{pmatrix} y(t) =: A y(t)$$

für  $y : \mathbb{R} \rightarrow \mathbb{R}^m$  oder  $y : \mathbb{R} \rightarrow \mathbb{C}^m$ . Jeder Eigenvektor  $v$  von  $A$  zum Eigenwert  $\lambda$  liefert eine Lösung  $y(t) = e^{\lambda t} v$  des Systems erster Ordnung. Das charakteristische Polynom von  $A$  ist

$$P_A(\lambda) = \sum_{j=0}^m a_j \lambda^j,$$

läßt sich also direkt aus der Differentialgleichung  $(*)$  ablesen. Damit haben wir ein Rezept zum Auffinden von Lösungen von  $(*)$  gefunden: zu jeder Nullstelle  $\lambda_0$  von  $P_A(\lambda)$  ist

$$x_{\lambda_0}(t) = e^{\lambda_0 t}$$

eine Lösung von  $(*)$ . Hat  $P_A(\lambda)$  tatsächlich  $m$  verschiedene Nullstellen, so bilden die zugehörigen Lösungen ein Fundamentalsystem. Aber auch beim Vorliegen einer  $\ell$ -fachen Nullstelle kann man direkt  $\ell$  linear unabhängige Lösungen angeben: sei  $\lambda_0$   $\ell$ -fache Nullstelle von  $P_A(\lambda)$ , dann sind

$$x_{\lambda_0,0}(t) = e^{\lambda_0 t}, \quad x_{\lambda_0,1}(t) = t e^{\lambda_0 t}, \quad x_{\lambda_0,2}(t) = t^2 e^{\lambda_0 t}, \quad \dots, \quad x_{\lambda_0,\ell-1}(t) = t^{\ell-1} e^{\lambda_0 t}$$

$\ell$  linear unabhängige Lösungen von  $(*)$ . (Beweis in den Übungen).

**9.22 Beispiel. Der gedämpfte harmonische Oszillator**

Das charakteristische Polynom zur Differentialgleichung

$$\ddot{x} + 2\gamma \dot{x} + \omega^2 x = 0, \quad \gamma, \omega > 0,$$

hat die Form  $P(\lambda) = \lambda^2 + 2\gamma\lambda + \omega^2$  und somit die Nullstellen

$$\lambda_{\pm} = -\gamma \pm \sqrt{\gamma^2 - \omega^2}.$$

Wir unterscheiden wieder drei Fälle:

**Überdämpfte Bewegung:**  $\gamma > \omega$ , also  $\lambda_{\pm} \in \mathbb{R}$  und  $\lambda_+ \neq \lambda_-$ . Zwei linear unabhängige Lösungen sind somit

$$x_1(t) = e^{\lambda_+ t} = e^{(-\gamma + \sqrt{\gamma^2 - \omega^2})t} \quad \text{und} \quad x_2(t) = e^{\lambda_- t} = e^{(-\gamma - \sqrt{\gamma^2 - \omega^2})t}$$

**Kritische Dämpfung:**  $\gamma = \omega$ , also  $\lambda_* = -\gamma$  zweifacher Eigenwert. Zwei linear unabhängige Lösungen sind diesmal

$$x_1(t) = e^{\lambda_* t} = e^{-\gamma t} \quad \text{und} \quad x_2(t) = t e^{\lambda_* t} = t e^{-\gamma t}.$$

**Gedämpfte Schwingung:**  $\gamma < \omega$ , also  $\lambda_{\pm} \in \mathbb{C}$  und  $\lambda_+ \neq \lambda_-$ . Zwei linear unabhängige Lösungen sind

$$x_1(t) = e^{\lambda_+ t} = e^{-\gamma t} e^{i\tilde{\omega}t} \quad \text{und} \quad x_2(t) = e^{\lambda_- t} = e^{-\gamma t} e^{-i\tilde{\omega}t} \quad \text{mit } \tilde{\omega} := \sqrt{\omega^2 - \gamma^2}.$$

In allen drei Fällen läßt sich jede Lösung als Linearkombination von  $x_1(t)$  und  $x_2(t)$  schreiben.

# 10 Funktionentheorie

## 10.1 Differenzierbarkeit in $\mathbb{C}$

Die komplexe Zahlenebene ist mittels

$$j : \mathbb{C} \rightarrow \mathbb{R}^2 \quad , \quad j(z) = (\operatorname{Re} z, \operatorname{Im} z)$$

isomorph zu dem  $\mathbb{R}$ -Vektorraum  $\mathbb{R}^2$ . Da

$$\begin{array}{ccc} & |z| = \|j(z)\| & \\ \nearrow & & \nwarrow \\ \text{Betrag in } \mathbb{C} & & \text{Norm in } \mathbb{R}^2 \end{array}$$

ist die metrische Struktur von  $\mathbb{C}$  mit der von  $\mathbb{R}^2$  identisch. Also ist  $U \subset \mathbb{C}$  genau dann offen, kompakt, abgeschlossen etc., wenn  $j(U) \subset \mathbb{R}^2$  offen, kompakt, abgeschlossen etc. ist. Eine Folge  $(z_n)_{n \in \mathbb{N}}$  konvergiert genau dann in  $\mathbb{C}$ , wenn  $(j(z_n))_{n \in \mathbb{N}}$  in  $\mathbb{R}^2$  konvergiert.

### 10.1 Definition. Reelle Differenzierbarkeit

Sei  $U \subset \mathbb{C}$  offen. Eine Abbildung  $f : U \rightarrow \mathbb{C}$  heißt reell differenzierbar bei  $z_0 \in U$ , falls sie aufgefasst als Abbildung von  $j(U) \subset \mathbb{R}^2$  nach  $\mathbb{R}^2$ , d.h.  $j \circ f \circ j^{-1} : j(U) \rightarrow \mathbb{R}^2$  differenzierbar ist bei  $z_0$ . Es ist also  $f : U \rightarrow \mathbb{C}$  reell differenzierbar bei  $z_0 \in U$  genau dann, wenn es eine reelle Matrix  $A \in M(2 \times 2, \mathbb{R})$  gibt, so dass für alle  $h \in \mathbb{C}$  hinreichend klein

$$j(f(z_0 + h)) = j(f(z_0)) + A j(h) + o(\|j(h)\|)$$

gilt.

Man sagt nun  $f : U \rightarrow \mathbb{C}$  ist komplex differenzierbar bei  $z_0 \in U$ , wenn es eine komplexe Zahl  $w \in \mathbb{C}$  gibt, so dass für  $h \in \mathbb{C}$  klein genug gilt

$$f(z_0 + h) = f(z_0) + wh + o(|h|).$$

### 10.2 Definition. Komplexe Differenzierbarkeit

Sei  $U \subset \mathbb{C}$  offen. Eine Funktion  $f : U \rightarrow \mathbb{C}$  heißt (komplex) differenzierbar bei  $z_0 \in U$ , falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

existiert. Wir definieren in diesem Fall die Ableitung von  $f$  in  $z_0$  durch

$$f'(z_0) = \lim_{h \rightarrow 0} \frac{f(z_0 + h) - f(z_0)}{h}.$$

Ist  $f$  in ganz  $U$  differenzierbar, so heißt  $f$  in  $U$  **holomorph**.

**10.3 Bemerkung.** Die aus dem Reellen bekannten Rechenregeln gelten auch im Komplexen (mit identischen Beweisen): Sind  $f, g : U \rightarrow \mathbb{C}$  holomorph, so auch  $f + g, f \cdot g, \alpha f$  für  $\alpha \in \mathbb{C}$  sowie  $f/g : U \setminus \{g = 0\} \rightarrow \mathbb{C}$  und es gilt

$$(f + g)' = f' + g', (fg)' = f'g + fg', (\alpha f)' = \alpha f'.$$

Außerdem gilt die Kettenregel: Sind  $f : U \rightarrow V$  und  $g : V \rightarrow \mathbb{C}$  holomorph, so ist auch  $g \circ f : U \rightarrow \mathbb{C}$  holomorph und

$$(g \circ f)'(z) = g'(f(z))f'(z) \quad \text{für alle } z \in U.$$

**10.4 Bemerkung.** Jede holomorphe Funktion  $f : U \rightarrow \mathbb{C}$  ist auch reell differenzierbar und somit stetig. Denn für  $w \in \mathbb{C}$  ist

$$M_w : \mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto M_w(z) := wz$$

$\mathbb{R}$ -linear als Abbildung von  $\mathbb{R}^2$  nach  $\mathbb{R}^2$ : Sei  $w = w_1 + iw_2$  und  $z = z_1 + iz_2$  mit  $w_1, w_2, z_1, z_2 \in \mathbb{R}$ . Dann ist

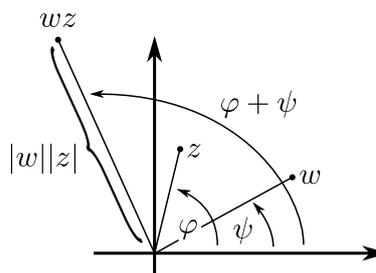
$$wz = (w_1 + iw_2)(z_1 + iz_2) = w_1z_1 - w_2z_2 + i(w_1z_2 + w_2z_1),$$

also

$$j(wz) = \begin{pmatrix} w_1z_1 - w_2z_2 \\ w_1z_2 + w_2z_1 \end{pmatrix} = \underbrace{\begin{pmatrix} w_1 & -w_2 \\ w_2 & w_1 \end{pmatrix}}_{=: A_w} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

**10.5 Erinnerung.** Geometrisch ist die Multiplikation mit einer komplexen Zahl  $w$  eine Drehstreckung,

$$\begin{aligned} w &= w_1 + iw_2 = |w| e^{i\psi} \\ z &= z_1 + iz_2 = |z| e^{i\varphi} \\ wz &= \underbrace{|w||z|}_{\text{Streckung}} \underbrace{e^{i(\psi+\varphi)}}_{\text{Drehung}} \end{aligned}$$



mit  $\psi := \arg(w)$  und  $\varphi := \arg(z)$ .

Die gleiche Drehstreckung in  $\mathbb{R}^2$  ist durch die Matrix  $A_w = \begin{pmatrix} w_1 & -w_2 \\ w_2 & w_1 \end{pmatrix}$  gegeben:

$$\det A_w = w_1^2 + w_2^2 = |w|^2$$

und

$$\frac{1}{(\det A_w)^{1/2}} A_w = \begin{pmatrix} \frac{w_1}{|w|} & -\frac{w_2}{|w|} \\ \frac{w_2}{|w|} & \frac{w_1}{|w|} \end{pmatrix}$$

ist die Drehung um den Winkel  $\cos \varphi = \frac{w_1}{|w|}$  bzw.  $\sin \varphi = \frac{w_2}{|w|}$ . Also gilt der folgende Satz.

**10.6 Satz. Cauchy-Riemannsche Differentialgleichungen**

Sei  $U \subset \mathbb{C}$  offen und  $f : U \rightarrow \mathbb{C}$ . Dann sind äquivalent:

- (i)  $f$  ist holomorph in  $U$ .
- (ii)  $f$  ist reell differenzierbar in  $U$ , und für die Funktionen  $u = \text{Re}f$  und  $v = \text{Im}f$  gilt

$$(*) \quad \begin{cases} \partial_x u(x, y) = \partial_y v(x, y) \\ \partial_y u(x, y) = -\partial_x v(x, y) \end{cases}$$

für alle  $z = x + iy \in U$ .

Die Gleichungen (\*) heißen die Cauchy-Riemannschen Differentialgleichungen und lauten kompakter

$$\partial_x f = -i\partial_y f.$$

*Beweis.* Schreiben wir  $j(f(z)) = \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}$ , also  $z = x + iy$ ,  $u(x, y) = \operatorname{Re} f(z)$  und  $v(x, y) = \operatorname{Im} f(z)$ , so ist

$$Dj(f(z)) = \begin{pmatrix} \partial_x u(x, y) & \partial_y u(x, y) \\ \partial_x v(x, y) & \partial_y v(x, y) \end{pmatrix}.$$

Falls  $f$  reell differenzierbar ist, so gilt

$$j(f(z+h)) = j(f(z)) + \underbrace{Dj(f(z))}_{\in M(2 \times 2, \mathbb{R})} \cdot j(h) + o(\|j(h)\|).$$

Das ist äquivalent zu

$$f(z+h) = f(z) + w \cdot h + o(|h|)$$

genau dann, wenn  $Dj(f(z))$  eine Drehstreckung ist, also

$$Dj(f) = \begin{pmatrix} \partial_x u & \partial_y u \\ \partial_x v & \partial_y v \end{pmatrix} = \begin{pmatrix} w_1 & -w_2 \\ w_2 & w_1 \end{pmatrix}.$$

Es muss also gelten

$$\partial_x u = \partial_y v \quad \text{und} \quad \partial_x v = -\partial_y u.$$

In diesem Fall ist  $w = f'(z)$  also  $\operatorname{Re} f' = \partial_x u = \partial_y v$  und  $\operatorname{Im} f' = \partial_x v = -\partial_y u$ . □

**10.7 Folgerung.** Es ist  $f$  also genau dann holomorph, wenn

$$f'(z) = \frac{\partial}{\partial x} f(z) = -i \frac{\partial}{\partial y} f(z). \quad (*)$$

Wegen  $x = (z + \bar{z})/2$  und  $y = (z - \bar{z})/2i$  ist formal

$$\frac{\partial}{\partial z} f(z) = \frac{\partial}{\partial z} f(x + iy) = \frac{\partial f}{\partial x} \frac{\partial x}{\partial z} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial z} = \frac{1}{2} \left( \frac{\partial f}{\partial x} + \frac{1}{i} \frac{\partial f}{\partial y} \right),$$

und man definiert deshalb die Differentialoperatoren (auch Wirtingerableitungen genannt)

$$\frac{\partial}{\partial z} := \frac{1}{2} \left( \frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \quad \text{und} \quad \frac{\partial}{\partial \bar{z}} := \frac{1}{2} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right).$$

Mit (\*) erfüllt  $f$  also genau dann die Cauchy-Riemannschen Differentialgleichungen, ist also genau dann holomorph, wenn  $\frac{\partial}{\partial \bar{z}} f = 0$ . Andererseits gilt für holomorphes  $f$ , dass  $f'(z) = \frac{\partial}{\partial z} f(z)$ .

### 10.8 Beispiel. Die komplexe Exponentialfunktion

Es ist

$$e^z = e^{x+iy} = e^x e^{iy} = e^x \cos y + i e^x \sin y,$$

also

$$u(x, y) = e^x \cos y \quad \text{und} \quad v(x, y) = e^x \sin y.$$

Damit ergibt sich

$$\partial_x u = e^x \cos y = \partial_y v \quad \text{und} \quad \partial_y u = -e^x \sin y = -\partial_x v.$$

Also ist  $\exp: \mathbb{C} \rightarrow \mathbb{C}$ ,  $z \mapsto e^z$  holomorph und wie erwartet

$$(e^z)' = \partial_x u + i \partial_x v = e^x \cos y + i e^x \sin y = e^z.$$

**10.9 Beispiel. Die Betragsfunktion ist nicht holomorph**

Die Betragsfunktion

$$|\cdot| : \mathbb{C} \rightarrow [0, \infty) \subset \mathbb{C}, \quad z \mapsto |z| = \sqrt{z\bar{z}}$$

ist nicht holomorph, denn

$$\partial_{\bar{z}}|z| = \partial_{\bar{z}}\sqrt{z\bar{z}} = \frac{z}{2\sqrt{z\bar{z}}} \neq 0.$$

Die Betragsfunktion ist aber bekanntermaßen außer bei  $z = 0$  reell differenzierbar.**10.10 Bemerkung. Real- und Imaginärteil holomorpher Funktionen sind harmonisch**Ist  $f : U \rightarrow \mathbb{C}$  holomorph und sind  $u = \operatorname{Re}f$  und  $v = \operatorname{Im}f$  zweimal stetig partiell differenzierbar (was, wie wir sehen werden, aus der Holomorphie folgt), so folgt aus den Cauchy-Riemannschen Differenzialgleichungen

$$\begin{aligned} \Delta u &= \partial_x(\partial_x u) + \partial_y(\partial_y u) = \\ &= \partial_x \partial_y v - \partial_y \partial_x v = 0 \end{aligned}$$

und analog  $\Delta v = 0$ . Real- und Imaginärteil einer holomorphen Funktion sind also immer Lösung der Laplace-Gleichung  $\Delta u = 0$  in  $U$ . Solche Funktionen nennt man harmonische Funktionen.**10.2 Komplexe Wegintegrale****10.11 Definition. Komplexes Wegintegral**Sei  $U \subset \mathbb{C}$  offen und  $\gamma : [\alpha, \beta] \rightarrow U$  eine stetig differenzierbare Kurve. Für jede stetige Funktion  $f : U \rightarrow \mathbb{C}$  sei dann

$$\int_{\gamma} f(z) dz := \int_{\alpha}^{\beta} f(\gamma(t)) \gamma'(t) dt$$

das Integral von  $f$  längs  $\gamma$ . Hierin ist

$$\gamma'(t) := \lim_{s \rightarrow t} \frac{\gamma(s) - \gamma(t)}{s - t} \in \mathbb{C}.$$

Entsprechend definiert man für stückweise glatte Wege, d.h.  $\gamma : [\alpha, \beta] \rightarrow U$  ist stetig und es existieren endlich viele Zwischenpunkte  $\alpha = \alpha_0 < \alpha_1 < \dots < \alpha_r = \beta$ , so dass  $\gamma_k = \gamma|_{[\alpha_{k-1}, \alpha_k]}$  jeweils stetig differenzierbar ist, das komplexe Wegintegral durch

$$\int_{\gamma} f(z) dz = \sum_{k=1}^r \int_{\gamma_k} f(z) dz.$$

Man mache sich klar, dass diese Definition unabhängig von der Wahl der Zwischenpunkte ist.

**10.12 Beispiel.** Sei  $\gamma : [0, 2\pi] \rightarrow \mathbb{C}$ ,  $t \mapsto \gamma(t) = e^{it}$ . Dann ist  $\gamma$  eine glatte geschlossene Kurve, nämlich die Einheitskreislinie, und es gilt  $\gamma'(t) = ie^{it}$ . Also ist z.B. für  $f(z) = z^n$ 

$$\begin{aligned} \int_{\gamma} f(z) dz &= \int_0^{2\pi} f(\gamma(t)) \gamma'(t) dt = \int_0^{2\pi} e^{itn} i e^{it} dt = \\ &= \begin{cases} \frac{e^{it(n+1)}}{n+1} \Big|_0^{2\pi} = \frac{1}{n+1} - \frac{1}{n+1} = 0 & \text{für } n \neq -1, \\ 2\pi i & \text{für } n = -1. \end{cases} \end{aligned}$$

**10.13 Bemerkung.** Wir werden im folgenden stückweise differenzierbare Wege betrachten, ohne dies immer explizit zu sagen. In diesem Zusammenhang ist wichtig, dass in einer offenen Menge  $U \subset \mathbb{C}$  zwei Punkte genau dann durch einen differenzierbaren Weg in  $U$  verbunden sind, wenn sie durch einen stetigen Weg verbunden sind. Denn sei  $\gamma : [0, 1] \rightarrow U$  ein stetiger Weg, so können wir die stetigen Funktionen  $\operatorname{Re}\gamma : [0, 1] \rightarrow \mathbb{R}$  und  $\operatorname{Im}\gamma : [0, 1] \rightarrow \mathbb{R}$  gleichmäßig durch Polynome approximieren (Stone-Weierstraß aus MaPhy I), also zu beliebigem  $\varepsilon > 0$  Polynome  $q$  und  $p$  finden, mit  $q(0) = \operatorname{Re}\gamma(0)$ ,  $q(1) = \operatorname{Re}\gamma(1)$  und

$$\|q - \operatorname{Re}\gamma\|_{\infty} < \varepsilon$$

und analog für  $p$  und  $\operatorname{Im}\gamma$ . Da  $U$  offen ist und  $\gamma([0, 1])$  kompakt, liegt der glatte Weg  $q + ip$  von  $\gamma(0)$  nach  $\gamma(1)$  aber für  $\varepsilon$  hinreichend klein auch ganz in  $U$ .

**Merke:** In einem Gebiet  $U \subset \mathbb{C}$ , also einer offenen und wegzusammenhängenden Menge  $U$ , können je zwei Punkte durch einen differenzierbaren Weg verbunden werden.

#### 10.14 Definition. Reparametrisierung eines Weges

Wir sagen, dass zwei Wege  $\gamma_1$  und  $\gamma_2$  durch Reparametrisierung auseinander hervorgehen, wenn es einen Diffeomorphismus  $\varphi : [\alpha_1, \beta_1] \rightarrow [\alpha_2, \beta_2]$  mit  $\gamma_1 = \gamma_2 \circ \varphi$  gibt. Dann ist entweder

$$\varphi(\alpha_1) = \alpha_2 \text{ und } \varphi' > 0 \quad \text{oder} \quad \varphi(\alpha_1) = \beta_2 \text{ und } \varphi' < 0.$$

Es heißt  $\varepsilon(\varphi) := \operatorname{sgn}\varphi' \in \{-1, 1\}$  die Orientierung der Reparametrisierung.

#### 10.15 Lemma. Reparametrisierung komplexer Wegintegrale

Sei  $f : U \rightarrow \mathbb{C}$  stetig und seien  $\gamma_1$  und  $\gamma_2$  Wege, die durch eine Reparametrisierung  $\varphi$  auseinander hervorgehen. Dann gilt

$$\int_{\gamma_1} f(z) dz = \varepsilon(\varphi) \int_{\gamma_2} f(z) dz.$$

*Beweis.*

$$\begin{aligned} \int_{\gamma_1} f(z) dz &= \int_{\alpha_1}^{\beta_1} f(\gamma_1(t)) \gamma_1'(t) dt = \int_{\alpha_1}^{\beta_1} f(\gamma_2(\varphi(t))) \gamma_2'(\varphi(t)) \varphi'(t) dt \\ &= \int_{\varphi(\alpha_1)}^{\varphi(\beta_1)} f(\gamma_2(s)) \gamma_2'(s) ds = \varepsilon(\varphi) \int_{\alpha_2}^{\beta_2} f(\gamma_2(s)) \gamma_2'(s) ds = \varepsilon(\varphi) \int_{\gamma_2} f(z) dz. \end{aligned}$$

□

**10.16 Bemerkung.** Für einen Weg der Form  $\gamma : [\alpha, \beta] \rightarrow [\alpha, \beta]$ ,  $t \mapsto \gamma(t) = t$  auf der reellen Achse, stimmt das komplexe Wegintegral mit dem üblichen Integral für Funktionen einer reellen Veränderlichen überein:

$$\int_{\gamma} f(z) dz = \int_{\alpha}^{\beta} f(t) dt.$$

Die Unabhängigkeit von Reparametrisierungen ist dann einfach die Substitutionsregel.

#### 10.17 Definition. Länge eines Weges

Die Länge eines Weges  $\gamma : [\alpha, \beta] \rightarrow \mathbb{C}$  ist

$$L(\gamma) := \int_{\alpha}^{\beta} |\gamma'(t)| dt.$$

Die Länge eines Weges ist invariant unter Reparametrisierung. (Übungsaufgabe)

**10.18 Definition. Stammfunktion**

Sei  $f : U \rightarrow \mathbb{C}$  stetig. Dann heißt  $F : U \rightarrow \mathbb{C}$  Stammfunktion von  $f$ , wenn für jede stückweise glatte Kurve  $\gamma : [\alpha, \beta] \rightarrow U$  gilt

$$\int_{\gamma} f(z) dz = F(\gamma(\beta)) - F(\gamma(\alpha)).$$

Es heißt  $f$  **integrabel**, falls  $f$  eine Stammfunktion hat.

**10.19 Lemma.** Sei  $f : U \rightarrow \mathbb{C}$  stetig, dann sind äquivalent:

- (i)  $f$  ist integrabel.
- (ii)  $\int_{\gamma} f(z) dz = 0$  für jede stückweise glatte, geschlossene Kurve  $\gamma$  in  $U$ .
- (iii) Für je zwei stückweise glatte Kurven  $\gamma_1, \gamma_2$  in  $U$  mit gleichem Anfangs- und Endpunkt gilt

$$\int_{\gamma_1} f(z) dz = \int_{\gamma_2} f(z) dz.$$

*Beweis.* (i)  $\Rightarrow$  (ii): klar

(ii)  $\Rightarrow$  (iii):

$$\int_{\gamma_1} f(z) dz - \int_{\gamma_2} f(z) dz = \int_{\gamma} f(z) dz = 0$$

mit dem geschlossenen Weg

$$„\gamma = \gamma_1 - \gamma_2“,$$

in dem erst  $\gamma_1$  vorwärts und dann  $\gamma_2$  rückwärts durchlaufen wird.

(iii)  $\Rightarrow$  (i): O.B.d.A. sei  $U$  ein Gebiet, also wegzusammenhängend. Wähle  $a \in U$  fest und definiere  $F : U \rightarrow \mathbb{C}$  wie folgt: Für jedes  $p \in U$  wähle  $\gamma_p : [0, 1] \rightarrow U$  mit  $\gamma_p(0) = a$  und  $\gamma_p(1) = p$  und setze

$$F(p) := \int_{\gamma_p} f(z) dz.$$

Nach Voraussetzung hängt  $F(p)$  nicht vom gewählten Pfad  $\gamma_p$  ab. Für einen beliebigen Weg  $\gamma : [\alpha, \beta] \rightarrow U$  von  $\gamma(\alpha) = p$  nach  $\gamma(\beta) = q$  gilt dann, dass

$$F(p) + \int_{\gamma} f(z) dz = \int_{\gamma_p} f(z) dz + \int_{\gamma} f(z) dz = \int_{„\gamma_p + \gamma“} f(z) dz = \int_{\gamma_q} f(z) dz = F(q).$$

Also ist  $F$  Stammfunktion von  $f$ . □

**10.20 Definition. Aneinandersetzen von Wegen**

Seien  $\gamma_1 : [a, b] \rightarrow \mathbb{C}$  und  $\gamma_2 : [c, d] \rightarrow \mathbb{C}$  stetige Wege. Dann bezeichne  $-\gamma_1 : [a, b] \rightarrow \mathbb{C}$  den rückwärts durchlaufenen Weg  $-\gamma_1(t) := \gamma_1(a + b - t)$  und  $\gamma_1 + \gamma_2 : [a, b + (d - c)] \rightarrow \mathbb{C}$  die Aneinandersetzung

$$(\gamma_1 + \gamma_2)(t) = \begin{cases} \gamma_1(t) & \text{falls } t \in [a, b] \\ \gamma_2(t + c - b) & \text{falls } t \in [b, b + (d - c)]. \end{cases}$$

Falls  $\gamma_1(b) = \gamma_2(c)$ , so ist  $\gamma_1 + \gamma_2$  auch wieder stückweise glatt.

Wir werden nun in mehreren Schritten einen “Hauptsatz” im Komplexen beweisen. Der folgende Satz sagt zunächst, dass auch im Komplexen  $F$  genau dann Stammfunktion von  $f$  ist, wenn  $F$  differenzierbar ist und  $F' = f$  gilt.

**10.21 Satz.** Sei  $U \subset \mathbb{C}$  offen und  $f : U \rightarrow \mathbb{C}$  stetig. Dann sind äquivalent:

- (i)  $F$  ist Stammfunktion von  $f$
- (ii)  $F$  ist holomorph und  $F' = f$ .

*Beweis.* (i) $\Rightarrow$ (ii): Sei  $z \in U$ , dann ist

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{F(z+h) - F(z)}{h} &= \lim_{h \rightarrow 0} \frac{1}{h} \int_{\gamma} f(z) dz = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^1 f(z+th) h dt \\ &= \lim_{h \rightarrow 0} \int_0^1 f(z+th) dt = f(z). \end{aligned}$$

(ii) $\Rightarrow$ (i): Sei  $\gamma : [\alpha, \beta] \rightarrow U$  ein glatter Weg, dann ist  $h(t) := F(\gamma(t))$  differenzierbar und es gilt  $h'(t) = f(\gamma(t)) \gamma'(t)$ . Mit dem Hauptsatz für Funktionen einer Veränderlichen folgt

$$F(\gamma(\beta)) - F(\gamma(\alpha)) = h(\beta) - h(\alpha) = \int_{\alpha}^{\beta} h'(t) dt = \int_{\alpha}^{\beta} f(\gamma(t)) \gamma'(t) dt = \int_{\gamma} f(z) dz.$$

□

**10.22 Definition. Lokale Integrierbarkeit**

Sei  $U \subset \mathbb{C}$  offen und  $f : U \rightarrow \mathbb{C}$  stetig. Falls zu jedem  $z \in U$  eine Umgebung  $U_z \subset U$  von  $z$  existiert so, dass  $f|_{U_z}$  integrel ist, so heißt  $f$  **lokal integrel**.

Wir kommen nun zur Frage nach der Existenz einer Stammfunktion. Während im Reellen jede stetige Funktion eine Stammfunktion besitzt, ist im Komplexen die komplexe Differenzierbarkeit eine hinreichende (nächster Satz) und, wie wir später sehen werden, auch notwendige Bedingung für lokale Integrierbarkeit.

**10.23 Satz.** Jede holomorphe Funktion ist lokal integrel.

**10.24 Proposition.** Sei  $U \subset \mathbb{C}$  eine offene Kreisscheibe. Dann ist  $f \in C(U)$  bereits dann integrel, wenn

$$\int_{\partial R} f(z) dz = 0$$

für jedes achsenparallele Rechteck  $R \subset U$  gilt.

*Beweis.* O.B.d.A. sei 0 der Mittelpunkt von  $U$ . Für  $a = \alpha + i\beta \in U$  sei  $\gamma$  bzw.  $\sigma$  derjenige Weg, der 0 mit  $a$  längs zweier Geradenstücke über  $\alpha$  bzw.  $i\beta$  verbindet.

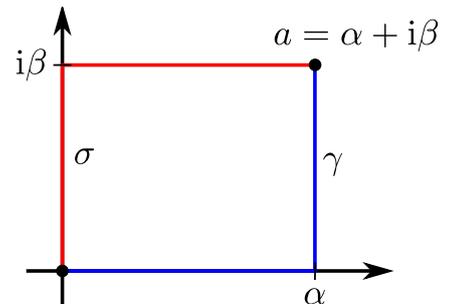
Nach Voraussetzung ist  $F(a) := \int_{\gamma} f(z) dz = \int_{\sigma} f(z) dz$ .

Für  $h \in \mathbb{R}$  gilt nun

$$\frac{\partial F}{\partial x}(a) = \lim_{h \rightarrow 0} \frac{F(a+h) - F(a)}{h} = \lim_{h \rightarrow 0} \int_0^1 f(a+th) dt = f(a)$$

und

$$\frac{\partial F}{\partial y}(a) = \lim_{h \rightarrow 0} \frac{F(a+ih) - F(a)}{h} = \lim_{h \rightarrow 0} i \int_0^1 f(a+ti h) dt = if(a).$$



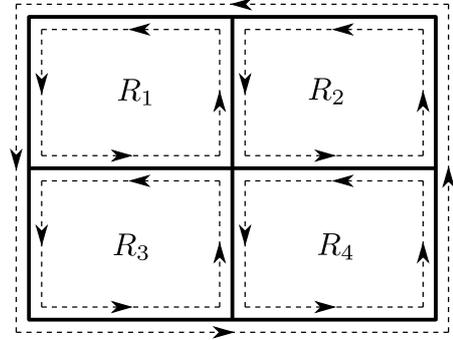
Also gilt in  $U$ , dass  $\partial_x F = f$  und  $\partial_y F = if$  und mit Satz 10.6 ist  $F$  holomorph und  $F' = f$ . □

*Beweis. von Satz 10.23.* Sei  $D \subset \mathbb{C}$  offen und  $f : D \rightarrow \mathbb{C}$  holomorph. Da jedes  $a \in D$  eine Kreisscheibenumgebung in  $D$  hat, genügt es zu zeigen, dass für jedes Rechteck  $R \subset D$  gilt

$$\int_{\partial R} f(z) dz = 0.$$

Sei also  $R \subset D$  fest,  $d := L(\partial R)$  der Umfang von  $R$  und zur Abkürzung  $\omega = f(z) dz$ . Teile nun  $R$  durch die beiden Seitenhalbierenden in vier kongruente Teilrechtecke  $R = R_1 \cup R_2 \cup R_3 \cup R_4$ . Die Wegintegrale über die inneren Seiten heben sich paarweise auf, also

$$\int_{\partial R} \omega = \sum_{k=1}^4 \int_{\partial R_k} \omega.$$



Insbesondere existiert ein  $j \in \{1, 2, 3, 4\}$  so, dass

$$\left| \int_{\partial R} \omega \right| \leq 4 \left| \int_{\partial R_j} \omega \right|.$$

Setze nun  $R^0 = R$  und  $R^1 = R_j$  und iteriere diese Prozedur. Wir erhalten dann eine Folge  $R^0 \supset R^1 \supset \dots \supset R^n \supset \dots$  von Rechtecken mit

$$\left| \int_{\partial R^n} \omega \right| \leq 4 \left| \int_{\partial R^{n+1}} \omega \right| \quad \text{und} \quad L(\partial R^{n+1}) = \frac{1}{2} L(\partial R^n),$$

also

$$L(\partial R^{n+1}) = \frac{d}{2^n} \quad \text{und} \quad \left| \int_{\partial R} \omega \right| \leq 4^n \left| \int_{\partial R^n} \omega \right|.$$

Wegen des Schachtelungsprinzips, Proposition 3.9, existiert ein eindeutiges  $a \in R$  mit  $a \in R^n$  für alle  $n \in \mathbb{N}$ . Da  $f : D \rightarrow \mathbb{C}$  holomorph ist, gilt

$$f(z) = f(a) + f'(a)(z - a) + \varphi(z - a)$$

mit  $\varphi(z - a) = o(|z - a|)$  für  $z$  in einer geeigneten Umgebung von  $a$ . Insbesondere ist  $g(z) := \frac{\varphi(z - a)}{z - a}$  stetig bei  $a$  und  $g(a) = 0$ .

Zu beliebigem  $\varepsilon > 0$  wähle nun  $n \in \mathbb{N}$  mit  $|g(z)| \leq \frac{\varepsilon}{d^2}$  für alle  $z \in R^n$ . Wegen  $|z - a| \leq L(\partial R^n) = \frac{d}{2^n}$  für alle  $z \in R^n$  gilt dann

$$\begin{aligned} \left| \int_{\partial R} \omega \right| &\leq 4^n \left| \int_{\partial R^n} \omega \right| = 4^n \left| \int_{\partial R^n} (f(a) + f'(a)(z - a) + g(z)(z - a)) dz \right| \\ &= 4^n \left| \int_{\partial R^n} g(z)(z - a) dz \right| \leq 4^n L(\partial R^n) \sup(g(z)(z - a)) \\ &\leq 4^n \frac{d}{2^n} \frac{\varepsilon}{d^2} \frac{d}{2^n} = \varepsilon, \end{aligned}$$

da  $z \mapsto f(a) + f'(a)(z - a)$  eine Stammfunktion besitzt (nämlich  $zf(a) + \frac{1}{2}f'(a)(z - a)^2$ ), und somit

$$\int_{\partial R^n} (f(a) + f'(a)(z - a)) dz = 0$$

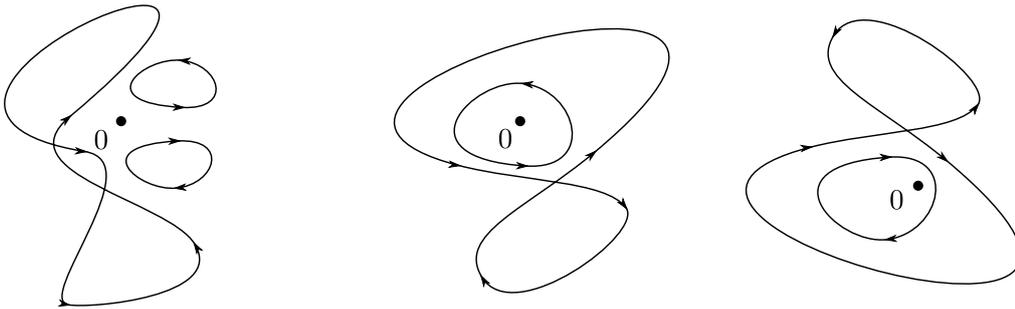
ist. □

### 10.3 Der Cauchy Integralsatz

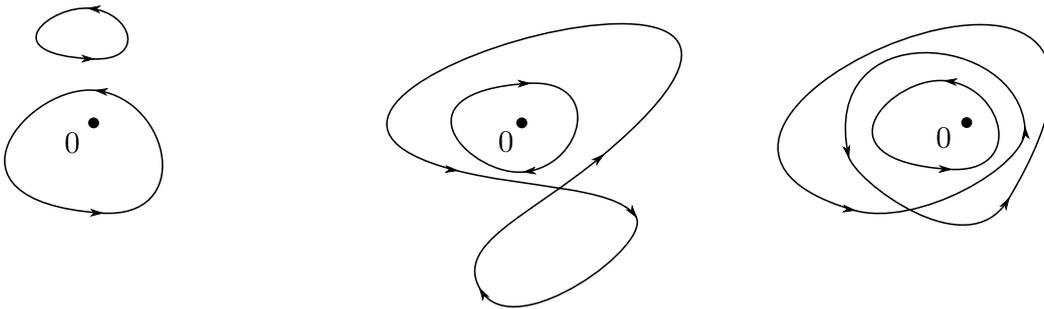
In Worten lautet die Aussage des Cauchyschen Integralsatzes wie folgt: Ist  $f : D \rightarrow \mathbb{C}$  holomorph,  $D$  offen, und kann man zwei geschlossene Wege  $\gamma_1$  und  $\gamma_2$  in  $D$  "stetig ineinander überführen ohne  $D$  zu verlassen", dann gilt

$$\int_{\gamma_1} f(z) dz = \int_{\gamma_2} f(z) dz .$$

**10.25 Beispiel.** Sei  $D = \mathbb{C} \setminus \{0\} =: \mathbb{C}^*$ . Beispiele für jeweils homotope Wege sind



Beispiele für jeweils nicht zueinander homotope Wege sind



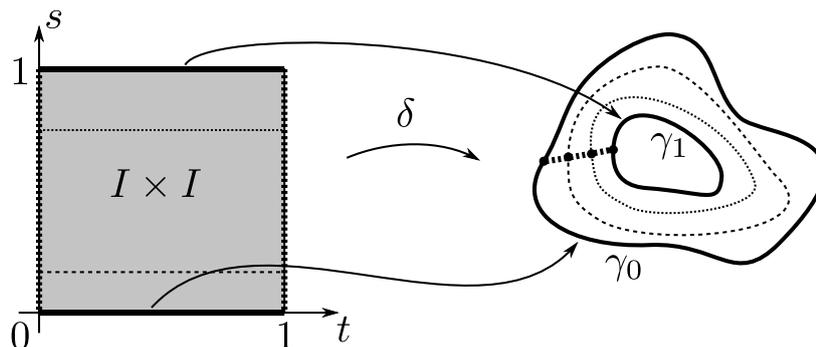
Im folgenden sei  $D \subset \mathbb{C}$  immer offen und  $I := [0, 1]$ .

#### 10.26 Definition. Homotopie geschlossener Kurven

Zwei geschlossenen stetige Kurven  $\gamma_0, \gamma_1 : I \rightarrow D$  heißen **homotop in  $D$** , falls es eine stetige Abbildung  $\delta : I \times I \rightarrow D$  gibt, mit

$$\delta(t, 0) = \gamma_0(t), \quad \delta(t, 1) = \gamma_1(t) \quad \text{und} \quad \delta(0, s) = \delta(1, s)$$

für alle  $t, s \in I$ . Man nennt eine solche Abbildung  $\delta$  dann **Homotopie**.



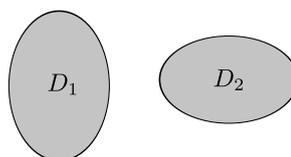
**10.27 Definition. Nullhomotop**

Eine geschlossene stetige Kurve  $\gamma : I \rightarrow D$  heißt **nullhomotop in  $D$** , falls  $\gamma$  homotop in  $D$  ist zu einer konstanten Kurve  $\gamma_0 : I \rightarrow D, \gamma_0(t) = a \in D \forall t \in I$ . Es lässt sich dann  $\gamma$  also “stetig auf einen Punkt zusammenziehen”.

**10.28 Definition. Einfach zusammenhängende Gebiete**

Eine Teilmenge  $D \subset \mathbb{C}$  heißt **einfach zusammenhängend**, falls jede geschlossene Kurve  $\gamma : I \rightarrow D$  in  $D$  nullhomotop ist.

**Achtung:** Mit dieser Definition folgt aus einfach zusammenhängend nicht zusammenhängend. Z.B. ist  $D = D_1 \cup D_2$  im Bild einfach zusammenhängend aber nicht zusammenhängend.



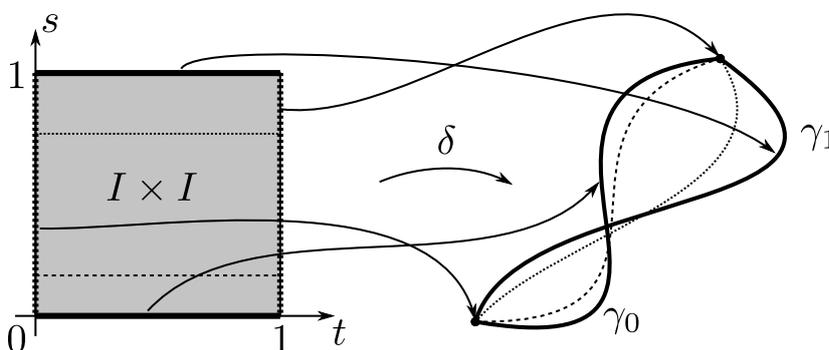
Statt geschlossene Wege stetig zu deformieren, kann man auch Wege mit beliebigen Endpunkten stetig deformieren und dabei die Endpunkte festhalten.

**10.29 Definition. Homotopie bei festen Endpunkten**

Zwei stetige Kurven  $\gamma_0, \gamma_1 : I \rightarrow D$  heißen **homotop in  $D$  bei festen Endpunkten**, falls es eine stetige Abbildung  $\delta : I \times I \rightarrow D$  gibt mit

$$\delta(t, 0) = \gamma_0(t), \quad \delta(t, 1) = \gamma_1(t) \quad \text{und} \quad \delta(0, s) = \delta(0, 0), \quad \delta(1, s) = \delta(1, 0)$$

für alle  $s, t \in I$ .



**10.30 Satz. Cauchy Integralsatz (spezielle Version)**

Sei  $D \subset \mathbb{C}$  offen und  $\gamma$  ein geschlossener und in  $D$  nullhomotoper Weg. Dann gilt für jedes holomorphe  $f : D \rightarrow \mathbb{C}$ , dass

$$\int_{\gamma} f(z) dz = 0.$$

**10.31 Satz. Cauchy Integralsatz**

Sei  $D \subset \mathbb{C}$  offen und seien  $\gamma_0, \gamma_1 : I \rightarrow D$  stückweise glatte Wege. Sei weiterhin eine der folgenden Bedingungen erfüllt:

- (i)  $\gamma_0$  und  $\gamma_1$  sind homotop in  $D$  bei festen Endpunkten.
- (ii)  $\gamma_0$  und  $\gamma_1$  sind homotop in  $D$  als geschlossene Kurven.

Dann gilt für jedes holomorphe  $f : D \rightarrow \mathbb{C}$ , dass

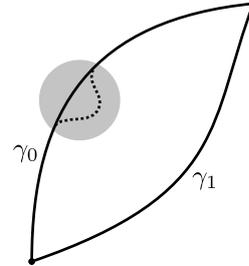
$$\int_{\gamma_0} f(z) dz = \int_{\gamma_1} f(z) dz.$$

**10.32 Bemerkung.** Satz 10.30 folgt aus Satz 10.31, da für jede konstante Kurve  $\gamma$  wegen  $\gamma' = 0$  gilt

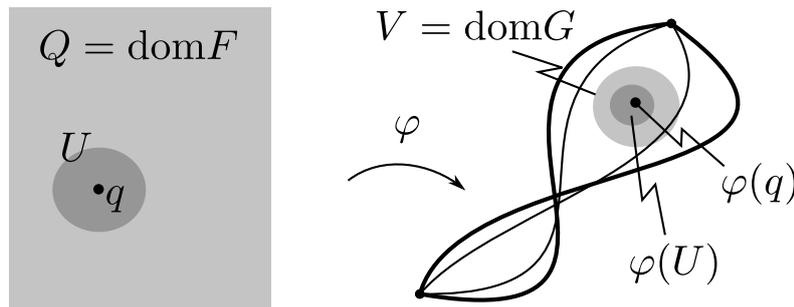
$$\int_{\gamma} f(z) dz = 0.$$

*Beweis. von Satz 10.31.* Die Idee ist einfach: Zumindest (i) würde sofort aus der Existenz einer Stammfunktion für  $f$  folgen.

Die hat man gemäß Satz 10.23 für holomorphes  $f$  zumindest lokal. Setze also “große Änderungen” des Weges aus lauter “kleinen Änderungen” zusammen. Letztere verändern das Integral wegen der lokalen Integrität nicht. Auf einem allgemeinen Gebiet ist das aber schwer präzise zu formulieren, deshalb zieht man alles auf das Quadrat  $I \times I$  zurück.



**10.33 Definition.** Sei  $f : D \rightarrow \mathbb{C}$  holomorph und entweder  $Q \subset \mathbb{R}^2$  ein Rechteck oder  $Q \subset \mathbb{R}$  ein Intervall. Sei  $\varphi : Q \rightarrow D$  stetig (man denke an eine Homotopie oder an eine Kurve). Dann heißt  $F : Q \rightarrow \mathbb{C}$  Stammfunktion von  $f$  längs  $\varphi$ , falls es zu jedem  $q \in Q$  eine Umgebung  $U_q$  von  $q$  in  $Q$ , eine Umgebung  $V$  von  $\varphi(q)$  in  $D$  und eine Stammfunktion  $G$  von  $f|_V$  gibt, sodass  $\varphi(U_q) \subset V$  und  $F(u) = G(\varphi(u))$  für alle  $u \in U$  gilt.



**10.34 Lemma.** Sei  $F : Q \rightarrow \mathbb{C}$  Stammfunktion von  $f$  längs  $\varphi : Q \rightarrow D$ . Dann gilt

- (i)  $F$  ist stetig.
- (ii) Für jede weitere Stammfunktion  $\tilde{F}$  von  $f$  längs  $\varphi$  ist  $\tilde{F} - F$  konstant.

*Beweis.* (i):  $F$  ist lokal Komposition der stetigen Abbildungen  $G$  und  $\varphi$ .

(ii)  $\tilde{F} - F$  ist lokal konstant, da jeweils  $\tilde{G} - G$  lokal konstant ist. □

**10.35 Lemma.** Sei  $\gamma : I \rightarrow D$  eine stückweise glatte Kurve und  $F : I \rightarrow \mathbb{C}$  eine Stammfunktion von  $f : D \rightarrow \mathbb{C}$  längs  $\gamma$ . Dann gilt

$$\int_{\gamma} f(z) dz = F(1) - F(0).$$

*Beweis.* Aus der offenen Überdeckung  $\bigcup_{q \in I} U_q \supset I$  kann man wegen der Kompaktheit von  $I$  eine endliche Teilüberdeckung auswählen. Es gibt also endlich viele Zwischenpunkte  $0 = t_0 < t_1 < \dots < t_n = 1$  und offene Umgebungen  $V_k \subset D$  mit Stammfunktionen  $G_k$  von  $f|_{V_k}$ ,  $k = 1, \dots, n$ , so dass  $\gamma([t_{k-1}, t_k]) \subset V_k$  und  $F(t) = G_k(\gamma(t))$  für  $t \in [t_{k-1}, t_k]$ .

Damit ist

$$\begin{aligned} \int_{\gamma} f(z) dz &= \sum_{k=1}^n \int_{\gamma_k} f(z) dz = \sum_{k=1}^n (G_k(\gamma(t_k)) - G_k(\gamma(t_{k-1}))) = \sum_{k=1}^n (F(t_k) - F(t_{k-1})) \\ &= F(t_n) - F(t_0) = F(1) - F(0). \end{aligned}$$

□

**10.36 Lemma.** Sei  $f : D \rightarrow \mathbb{C}$  stetig und lokal integabel und  $\varphi : I \times I \rightarrow D$  stetig. Dann existiert eine Stammfunktion von  $f$  längs  $\varphi$ .

*Beweis.* Sei  $Q = I \times I$  und sei  $\mathcal{L} \subset \mathbb{R}$  die Menge aller  $\ell > 0$  mit der folgenden Eigenschaft: Für jedes Rechteck  $R \subset Q$  mit Umfang  $L(\partial R) \leq \ell$  existiert eine Stammfunktion von  $f$  längs  $\varphi|_R$ . Zu zeigen: Es gibt ein  $\ell \in \mathcal{L}$  mit  $\ell \geq 4$ , denn  $Q$  hat Umfang vier.

$\mathcal{L} \neq \emptyset$ : Andernfalls gäbe es zu jedem  $n \in \mathbb{N}$  ein Rechteck  $R_n \subset Q$  mit Umfang  $L(\partial R_n) \leq \frac{1}{n}$ , so dass  $f$  längs  $\varphi|_{R_n}$  keine Stammfunktion besitzt. Wähle  $a_n \in R_n$  beliebig, dann hat die Folge  $(a_n)$  einen Häufungspunkt  $a \in Q$ , da  $Q$  kompakt ist. Nun ist aber  $f$  lokal-integabel, also gibt es eine Umgebung  $U$  von  $a \in Q$  und eine Stammfunktion  $F$  von  $f$  längs  $\varphi|_U$ . Da aber  $R_n \subset U$  für  $n$  groß genug, muss die Annahme falsch sein.

$\ell \in \mathcal{L} \Rightarrow \frac{4}{3}\ell \in \mathcal{L}$ : Sei also  $\ell \in \mathcal{L}$  und  $R \subset Q$  ein Rechteck mit  $L(\partial R) \leq \frac{4}{3}\ell$ . Dann ist  $R = R_1 \cup R_2$  wobei  $R_1, R_2$  Rechtecke mit Umfang  $\leq \ell$  sind. Also existieren Stammfunktionen  $F_k$  von  $f$  längs  $\varphi|_{R_k}$  für  $k = 1, 2$ . Die Einschränkungen  $F_k|_{R_1 \cap R_2}$  sind Stammfunktionen längs  $\varphi|_{R_1 \cap R_2}$ , also ist  $\delta = F_1|_{R_1 \cap R_2} - F_2|_{R_1 \cap R_2}$  konstant. Setze

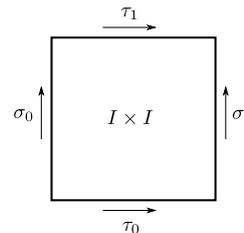
$$F = \begin{cases} F_1 & \text{auf } R_1 \\ F_2 + \delta & \text{auf } R_2 \end{cases},$$

dann ist  $F$  Stammfunktion von  $f$  längs  $\varphi|_R$ , also  $\frac{4}{3}\ell \in \mathcal{L}$ . □

*Ende des Beweises von Satz 10.31.* Sei also  $f : D \rightarrow \mathbb{C}$  stetig und lokal integabel und  $\delta : I \times I \rightarrow D$  eine Homotopie der stückweise glatten Kurven  $\gamma_0, \gamma_1 : I \rightarrow D$ . Wegen Lemma 10.36 existiert eine Stammfunktion  $F$  von  $f$  längs  $\delta$ .

Wir parametrisieren die vier Seiten von  $I \times I$  durch Kurven  $\sigma_k, \tau_j : I \rightarrow I \times I$ ,  $k, j = 0, 1$ , vermöge

$$\sigma_k(s) := (k, s) \quad \text{und} \quad \tau_j(t) = (t, j).$$



Nun ist  $F \circ \tau_j$  Stammfunktion von  $f$  längs  $\delta \circ \tau_j = \gamma_j$  und  $F_k := F \circ \sigma_k$  Stammfunktion von  $f$  längs  $\delta \circ \sigma_k$ . Es gilt also mit Lemma 10.35, dass

$$\int_{\gamma_j} f(z) dz = F(1, j) - F(0, j) = F_1(j) - F_0(j).$$

Es bleibt also zu zeigen, dass  $F_1(s) - F_0(s)$  konstant auf  $I$  ist, falls (i) oder (ii) gilt.

Fall (i): Die Wege  $\gamma_0$  und  $\gamma_1$  sind homotop bei festen Endpunkten. Dann ist  $\delta \circ \sigma_k$  konstant, also z.B.  $h \equiv 0$  Stammfunktion von  $f$  längs  $\delta \circ \sigma_k$ . Also ist auch  $F_k$  konstant und somit  $F_1 - F_0$  ebenfalls.

Fall (ii): Die Wege  $\gamma_0$  und  $\gamma_1$  sind homotop als geschlossene Kurven. Dann ist  $\delta \circ \sigma_0 = \delta \circ \sigma_1$  und somit, da  $F_0, F_1$  Stammfunktionen von  $f$  längs  $\delta \circ \sigma_0$  sind,  $F_1 - F_0$  konstant. □

**10.37 Bemerkung.** Wir haben im Beweis von Satz 10.31 nur die lokale Integribilität verwendet, nirgends direkt die Holomorphie.

## 10.4 Die Cauchy Integralformel

### 10.38 Definition. Umlaufzahl

Sei  $\omega \in \mathbb{C}$  und  $\gamma : [\alpha, \beta] \rightarrow \mathbb{C} \setminus \{\omega\}$  eine stückweise glatte, geschlossene Kurve. Dann heißt

$$I(\gamma, \omega) := \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z - \omega}$$

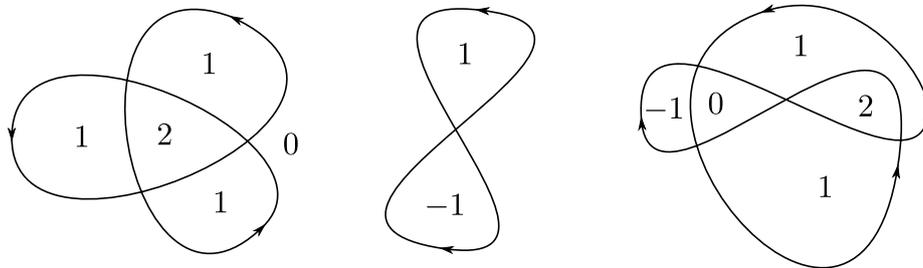
die Umlaufzahl, Windungszahl oder auch kurz der Index von  $\gamma$  in  $\omega$ .

Es gilt  $I(\gamma, \omega) \in \mathbb{Z}$  und  $I(\gamma, \cdot)$  ist lokal konstant auf  $\mathbb{C} \setminus \text{Sp}(\gamma)$ , wobei

$$\text{Sp}(\gamma) := \{\gamma(t) \in \mathbb{C} \mid t \in [\alpha, \beta]\}.$$

*Beweis.* Übungsaufgabe. □

### 10.39 Beispiel. Kurven und ihre Umlaufzahlen:

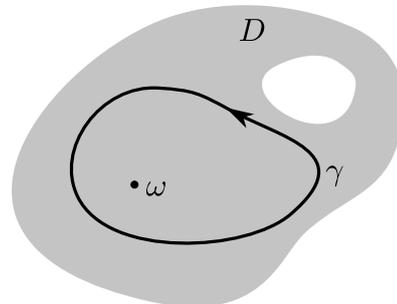


### 10.40 Satz. Cauchy Integralformel

Sei  $D \subset \mathbb{C}$  offen,  $f : D \rightarrow \mathbb{C}$  holomorph und  $\gamma$  eine stückweise glatte, geschlossene Kurve in  $D$ , die in  $D$  nullhomotop ist. Dann gilt

$$I(\gamma, \omega) f(\omega) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - \omega} dz$$

für alle  $\omega \in D \setminus \text{Sp}(\gamma)$ .



*Beweis.* Sei  $\omega \in D \setminus \text{Sp}(\gamma)$  fest. Dann ist

$$\Delta(z) := \frac{f(z) - f(\omega)}{z - \omega}$$

holomorph auf  $D \setminus \{\omega\}$  und stetig auf  $D$  (da  $\lim_{z \rightarrow \omega} \Delta(z) = f'(\omega)$  existiert). Nach Übungsaufgabe 55 ist  $\Delta$  dann aber schon auf ganz  $D$  lokal integrierbar und nach dem Cauchy Integralsatz gilt  $\int_{\gamma} \Delta(z) dz = 0$ . Also gilt wegen  $f(z) = f(\omega) + \Delta(z)(z - \omega)$ , dass

$$\int_{\gamma} \frac{f(z)}{z - \omega} dz = \int_{\gamma} \frac{f(\omega)}{z - \omega} dz + \int_{\gamma} \Delta(z) dz = f(\omega) \int_{\gamma} \frac{1}{z - \omega} dz = f(\omega) 2\pi i I(\gamma, \omega).$$

□

Es ist also  $f$  im Inneren von  $\gamma$ ,  $\text{In}(\gamma) = \{\omega \in \mathbb{C} \setminus \text{Sp}(\gamma) \mid I(\gamma, \omega) \neq 0\}$ , durch die Werte von  $f$  auf  $\text{Sp}(\gamma)$  schon eindeutig bestimmt.

Es kommt noch besser: die rechte Seite der Cauchy Integralformel, also

$$\frac{1}{2\pi i I(\gamma, \omega)} \int_{\gamma} \frac{f(z)}{z - \omega} dz,$$

ist offensichtlich beliebig oft komplex differenzierbar nach  $\omega$  und somit auch  $f(\omega)$ !

**10.41 Korollar. Cauchy-Integralformel für die Ableitung**

Sei  $D \subset \mathbb{C}$  offen,  $f : D \rightarrow \mathbb{C}$  holomorph und  $n \in \mathbb{N}$ . Dann ist  $f$   $n$ -mal komplex differenzierbar und für die  $n$ -te Ableitung  $f^{(n)}(z)$  gilt

$$I(\gamma, z) f^{(n)}(z) = \frac{n!}{2\pi i} \int_{\gamma} \frac{f(\xi)}{(\xi - z)^{n+1}} d\xi,$$

wobei  $\gamma$  eine beliebige nullhomotope geschlossene Kurve in  $D$  und  $z \in D \setminus \text{Sp}(\gamma)$  ist.

*Beweis.* Für  $I(\gamma, z) = 0$  ist  $\gamma$  nullhomotop in  $D \setminus \{z\}$  und die Aussage gilt nach dem Cauchy Integralsatz. Sei also  $I(\gamma, z) \neq 0$ . Induktion nach  $n$ : Für  $n = 0$  ist die Aussage einfach die Cauchy-Integralformel.

$n \Rightarrow n + 1$ : Nach Induktionsannahme ist

$$f^{(n)}(z) = \frac{n!}{2\pi i I(\gamma, z)} \int_{\gamma} \frac{f(\xi)}{(\xi - z)^{n+1}} d\xi$$

und somit

$$\begin{aligned} \frac{d}{dz} f^{(n)}(z) &= \frac{n!}{2\pi i} \underbrace{\left( \frac{d}{dz} \frac{1}{I(\gamma, z)} \right)}_{= 0 \text{ da } I \text{ lokal konstant}} \int_{\gamma} \frac{f(\xi)}{(\xi - z)^{n+1}} d\xi + \frac{n!}{2\pi i I(\gamma, z)} \frac{d}{dz} \int_{\gamma} \frac{f(\xi)}{(\xi - z)^{n+1}} d\xi \\ &= \frac{n!}{2\pi i I(\gamma, z)} \int_{\gamma} \frac{d}{dz} \frac{f(\xi)}{(\xi - z)^{n+1}} d\xi = \frac{(n+1)!}{2\pi i I(\gamma, z)} \int_{\gamma} \frac{f(\xi)}{(\xi - z)^{n+2}} d\xi. \end{aligned}$$

□

**10.42 Folgerung.** (a) Jede einmal komplex differenzierbare Funktion ist beliebig oft komplex differenzierbar.

(b)  $f$  ist holomorph  $\Leftrightarrow f$  ist lokal integrabel.

**10.43 Satz. Cauchy Abschätzung**

Sei  $D \subset \mathbb{C}$  offen,  $f : D \rightarrow \mathbb{C}$  holomorph,  $a \in D$  und  $R > 0$  so, dass  $B_R(a) \subset D$ . Sei  $M = \sup_{z \in B_R(a)} |f(z)|$ , dann gilt für alle  $n \in \mathbb{N}$

$$|f^{(n)}(a)| \leq \frac{n!}{R^n} M.$$

*Beweis.* Wähle  $0 < r < R$  beliebig und sei  $\gamma_r : [0, 2\pi] \rightarrow D$ ,  $\gamma_r(t) = a + re^{it}$ . Dann ist

$$f^{(n)}(a) = \frac{n!}{2\pi i} \int_{\gamma} \frac{f(\xi)}{(\xi - a)^{n+1}} d\xi = \frac{n!}{2\pi i} \frac{1}{r^{n+1}} \int_0^{2\pi} \frac{f(\gamma_r(t)) i r e^{it}}{e^{it(n+1)}} dt = \frac{n!}{2\pi} \frac{1}{r^n} \int_0^{2\pi} f(\gamma_r(t)) e^{-itn} dt$$

also

$$|f^{(n)}(a)| \leq \frac{n!}{r^n} M.$$

□

**10.44 Korollar. Satz von Liouville**

Jede beschränkte, auf ganz  $\mathbb{C}$  holomorphe Funktion ist konstant.

*Beweis.* Sei  $|f(z)| \leq M$ , also  $|f'(z)| \leq \frac{M}{R}$  für  $R > 0$  beliebig. Dann ist  $f'(z) = 0$ .  $\square$

**10.45 Korollar. Fundamentalsatz der Algebra**

Jedes nichtkonstante Polynom mit komplexen Koeffizienten hat wenigstens eine Nullstelle in  $\mathbb{C}$ .

*Beweis.* Sei  $p(z)$  ein komplexes Polynom vom Grad  $n \geq 1$ . Dann existiert ein  $R > 0$  so, dass  $|p(z)| > 1$  für alle  $z$  mit  $|z| > R$ . Angenommen,  $p$  hat keine Nullstelle. Dann ist  $f(z) = \frac{1}{p(z)}$  eine beschränkte holomorphe Funktion auf  $\mathbb{C}$  und somit nach dem Satz von Liouville konstant.  $\square$

**10.46 Definition. Kompakte Konvergenz von Funktionen**

Sei  $D \subset \mathbb{C}$  offen und  $(f_n)$  eine Folge von Funktionen  $f_n : D \rightarrow \mathbb{C}$ . Man sagt  $(f_n)$  **konvergiert kompakt** gegen  $f : D \rightarrow \mathbb{C}$ , falls die Folge  $f_n|_K$  von Funktionen auf jedem Kompaktum  $K \subset D$  gleichmäßig konvergiert, also

$$\lim_{n \rightarrow \infty} \sup_{z \in K} |f_n(z) - f(z)| = 0.$$

Insbesondere impliziert gleichmäßige Konvergenz also die kompakte Konvergenz, aber nicht umgekehrt.

**10.47 Satz. Kompakte Grenzwerte holomorpher Funktionen sind holomorph**

Sei  $(f_n)$  eine Folge holomorpher Funktionen auf  $D \subset \mathbb{C}$ , die kompakt gegen  $f : D \rightarrow \mathbb{C}$  konvergiert. Dann ist  $f$  holomorph und  $(f'_n)$  konvergiert kompakt gegen  $f'$ .

*Beweis.* Als gleichmäßiger Limes stetiger Funktionen ist  $f$  stetig. Für jeden nullhomotopen Weg  $\gamma$  in  $D$  gilt

$$\int_{\gamma} f(z) dz = \lim_{n \rightarrow \infty} \int_{\gamma} f_n(z) dz = 0,$$

also ist  $f$  lokal integrierbar und somit holomorph. Damit gilt auch für jedes Kompaktum  $K \in D$  und für  $\varepsilon$  so klein, dass  $\partial(K + \varepsilon) \subset D$ ,

$$\sup_{z \in K} |f'(z) - f'_n(z)| \leq \frac{1}{2\pi} \sup_{z \in K} \int_{\partial(K+\varepsilon)} \frac{|f(\xi) - f_n(\xi)|}{|\xi - z|^2} d\xi \leq \frac{c_K}{2\pi\varepsilon^2} \sup_{z \in K+\varepsilon} |f(z) - f_n(z)| \xrightarrow{n \rightarrow \infty} 0.$$

$\square$

**10.48 Bemerkung.** Insbesondere sind also gleichmäßige Grenzwerte holomorpher Funktionen wieder holomorph!

**Achtung:** Die analoge Aussage im reellen ist falsch, denn  $f(x) = |x|$  ist gleichmäßiger Limes differenzierbarer Funktionen auf  $\mathbb{R}$ .

**10.5 Potenzreihen**

**10.49 Satz.** Sei  $(c_n)_{n \in \mathbb{N}_0}$  eine Folge in  $\mathbb{C}$  und

$$R := \frac{1}{\limsup \sqrt[n]{|c_n|}}.$$

Dann gilt:

- (i) für jedes  $z \in \mathbb{C}$  mit  $|z| < R$  konvergiert die Reihe  $\sum_{n=0}^{\infty} c_n z^n$  absolut.
- (ii) für jedes  $z \in \mathbb{C}$  mit  $|z| > R$  divergiert die Reihe  $\sum_{n=0}^{\infty} c_n z^n$ .

Die Zahl  $R$  heißt Konvergenzradius der Reihe  $\sum_{n=0}^{\infty} c_n z^n$  und  $\{z \in \mathbb{C} \mid |z| < R\}$  das Konvergenzgebiet.

*Beweis.* (i) Sei  $|z| < R$ , dann ist  $\limsup \sqrt[n]{|c_n|} \cdot |z| < 1$ . Also gibt es ein  $n_0 \in \mathbb{N}$  und ein  $p < 1$  so, dass  $\sqrt[n]{|c_n|} \cdot |z| \leq p < 1$  für alle  $n \geq n_0$ . Somit ist  $|c_n| |z|^n \leq p^n$  und  $\sum |c_n| |z|^n$  wird durch die geometrische Reihe  $\sum p^n$  majorisiert.

- (ii) Angenommen  $\sum c_n z^n$  konvergiert, also  $c_n z^n \rightarrow 0$ . Es gibt dann insbesondere ein  $n_0 \in \mathbb{N}$  so, dass  $|c_n z^n| < 1$ , also  $\sqrt[n]{|c_n|} |z| < 1$  für alle  $n \geq n_0$ . Also ist  $|z| \leq \frac{1}{\limsup \sqrt[n]{|c_n|}}$ .

□

**10.50 Satz.** Es sei  $\sum c_n z^n$  eine Potenzreihe mit Konvergenzradius  $R > 0$ . Dann ist  $f(z) = \sum_{n=0}^{\infty} c_n z^n$  auf  $D = \{z \in \mathbb{C} \mid |z| < R\}$  eine holomorphe Funktion und es gilt  $f'(z) = \sum_{n=1}^{\infty} n c_n z^{n-1}$  sowie  $c_n = \frac{f^{(n)}(0)}{n!}$ . Es ist also  $\sum c_n z^n$  die Taylorreihe von  $f$  in  $z = 0$ .

*Beweis.* Die Partialsummenfolge  $f_N(z) = \sum_{n=0}^N c_n z^n$  konvergiert kompakt auf  $D$  gegen  $f(z) = \sum_{n=0}^{\infty} c_n z^n$ . Denn sei  $K \subset D$  kompakt, also  $K \subset B_r(0)$  mit  $r < R$ , dann ist

$$\left| \sum_{n=N+1}^{\infty} c_n z^n \right| \leq \sum_{n=N+1}^{\infty} |c_n| r^n < \varepsilon$$

für alle  $z \in K$  und  $N$  hinreichend groß. Nach Satz 10.47 ist  $f$  also holomorph auf  $D$  und  $f'(z) = \lim_{N \rightarrow \infty} f'_N(z) = \sum_{n=1}^{\infty} n c_n z^{n-1}$ . Insbesondere ist  $f'(0) = c_1$  und wiederholtes Anwenden liefert  $c_n$ . □

Umgekehrt gilt

**10.51 Satz.** Es sei  $D = \{z \in \mathbb{C} \mid |z| < r\}$  die Kreisscheibe mit Radius  $0 < r \leq \infty$  und  $f : D \rightarrow \mathbb{C}$  holomorph. Dann ist der Konvergenzradius der Taylorreihe  $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n$  von  $f$  mindestens  $r$  und sie stellt auf  $D$  die Funktion  $f$  dar:

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n \quad \forall z \in D.$$

*Beweis.* Sei  $z \in D$  fest und  $|z| < \varrho < r$ . Für  $\xi \in D$  mit  $|\xi| = \varrho$  gilt dann  $|\frac{z}{\xi}| < 1$ , also

$$\frac{1}{\xi - z} = \frac{1}{\xi} \left( \frac{1}{1 - \frac{z}{\xi}} \right) = \frac{1}{\xi} \sum_{n=0}^{\infty} \left( \frac{z}{\xi} \right)^n$$

gleichmäßig für alle  $\xi$  mit  $|\xi| = \varrho$ . Die Cauchy Integralformel liefert dann

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_{|\xi|=\varrho} \frac{f(\xi)}{\xi - z} d\xi = \frac{1}{2\pi i} \int \frac{f(\xi)}{\xi} \sum_{n=0}^{\infty} \left( \frac{z}{\xi} \right)^n d\xi \\ &= \sum_{n=0}^{\infty} \underbrace{\frac{1}{2\pi i} \int \frac{f(\xi)}{\xi^{n+1}} d\xi}_{\frac{f^{(n)}(0)}{n!}} z^n = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n. \end{aligned}$$

□

**10.52 Zusammenfassung. Wichtig!**

- Jede Potenzreihe  $\sum c_n(z-a)^n$  stellt auf ihrem Konvergenzgebiet (welches immer eine Kreisscheibe ist) eine holomorphe Funktion dar.
- Umgekehrt lässt sich jede holomorphe Funktion  $f : D \rightarrow \mathbb{C}$  lokal als konvergente Potenzreihe schreiben. Genauer gilt:  
Die Taylorreihe  $\sum \frac{f^{(n)}(a)}{n!}(z-a)^n$  von  $f$  bei  $a \in D$  konvergiert auf jeder Kreisscheibenumgebung  $B_r(a) = \{z \in \mathbb{C} \mid |z-a| < r\}$  mit  $B_r(a) \subset D$ .

**10.53 Satz. Nullstellen holomorpher Funktionen**

Sei  $f : D \rightarrow \mathbb{C}$  holomorph,  $D$  ein Gebiet und  $f \neq 0$ .

- (a) Ist  $z_0 \in D$  Nullstelle von  $f$ , so gibt es ein  $k \in \mathbb{N}$  und eine holomorphe Funktion  $g : D \rightarrow \mathbb{C}$  mit

$$f(z) = (z - z_0)^k \cdot g(z) \quad \text{und} \quad g(z_0) \neq 0.$$

Die Zahl  $k$  heißt Ordnung der Nullstelle.

- (b) Genau dann hat  $f$  an der Stelle  $z_0$  eine Nullstelle  $k$ -ter Ordnung, wenn

$$f(z_0) = f'(z_0) = \dots = f^{(k-1)}(z_0) = 0 \quad \text{aber} \quad f^{(k)}(z_0) \neq 0.$$

- (c) In jeder kompakten Teilmenge  $K \subset D$  hat  $f$  höchstens endlich viele Nullstellen. Die Nullstellenmenge einer holomorphen Funktion ist also diskret, d.h. sie hat keine Häufungspunkte im Holomorphiegebiet  $D$ .

*Beweis.* (a) Auf einer hinreichend kleinen Kreisscheibe  $B_r(z_0)$  um  $z_0$  gilt

$$f(z) = \sum_{n=1}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n.$$

Dann existiert  $k = \min\{n \geq 1 \mid f^{(n)}(z_0) \neq 0\}$ . (Denn sei  $B = \{z \in D \mid f^{(n)}(z) = 0 \forall n \in \mathbb{N}_0\}$ , dann gilt  $B$  ist abgeschlossen, denn alle Ableitungen  $f^{(n)}(z)$  sind stetig und  $B$  ist offen, da zu jedem  $z \in B$  ein  $\rho$  existiert, so dass die Taylorreihe von  $f$  bei  $z$  auf  $B_\rho(z)$  konvergiert. Da  $D$  zusammenhängend ist, sind die einzigen Teilmengen von  $D$  die sowohl offen als auch abgeschlossen sind  $\emptyset$  und  $D$ . Also  $B = \emptyset$ .) Also ist

$$f(z) = (z - z_0)^k \underbrace{\sum_{n=k}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^{n-k}}_{g(z)},$$

wobei  $g$  auf  $B_r$  holomorph ist und  $g(z_0) = \frac{f^{(k)}(z_0)}{k!} \neq 0$  ist. Setze  $g$  außerhalb von  $B_r$  durch  $g(z) = f(z)/(z - z_0)^k$  fort.

- (b) klar.
- (c) Angenommen  $f$  hat unendlich viele Nullstellen in  $K$ , dann haben die Nullstellen einen Häufungspunkt in  $K$ , da  $K$  kompakt ist. Das kann aber nicht sein, da es wegen (a) zu jeder Nullstelle  $z_0$  eine Umgebung gibt, in der keine weitere Nullstelle liegt. Denn  $f(z) = (z - z_0)^k g(z)$  und  $(z - z_0)^k$  ist nur für  $z = z_0$  Null und  $g(z)$  ist als stetige Funktion auch in einer Umgebung von  $z_0$  ungleich Null.

□

**10.54 Bemerkung.** Die Nullstellen holomorpher Funktionen können sich am Rand des Gebiets häufen. Sei z.B.  $D = \mathbb{C}^* = \mathbb{C} \setminus \{0\}$  und  $f(z) = \sin\left(\frac{\pi}{z}\right)$ . Dann ist  $f\left(\frac{1}{n}\right) = 0$  für  $n \in \mathbb{N}$ .

**10.55 Satz. Identitätssatz**

Sei  $D \subset \mathbb{C}$  ein Gebiet und  $M \subset D$  eine Teilmenge, die mindestens einen Häufungspunkt besitzt. Dann gilt für je zwei holomorphe Funktionen  $f, g : D \rightarrow \mathbb{C}$ , dass aus  $f|_M = g|_M$  folgt  $f = g$  auf  $D$ .

*Beweis.* Die holomorphe Funktion  $f - g$  ist Null auf einer Menge mit Häufungspunkt und muss somit nach Satz 10.53 identisch verschwinden.  $\square$

**10.56 Merke.** Eine holomorphe Funktion  $f : D \rightarrow \mathbb{C}$  ist durch ihre Werte auf einer beliebigen Menge  $M$ , die einen ihrer Häufungspunkte enthält, eindeutig bestimmt. Also z.B. auf einer beliebigen offenen Menge, oder aber auch nur auf einer Folge von Punkten mit Häufungspunkt.

**10.57 Satz.** Sei  $D \subset \mathbb{C}$  offen,  $f : D \rightarrow \mathbb{C}$  holomorph und  $a \in D$ . Dann sind äquivalent

- (i)  $f$  ist lokal injektiv.
- (ii)  $f$  ist lokal biholomorph (d.h. die lokale Inverse ist holomorph).
- (iii)  $f'(a) \neq 0$

*Beweis.* Identifizieren wir  $\mathbb{C}$  mit  $\mathbb{R}^2$ , so ist wegen

$$\det Df = \det \begin{pmatrix} \partial_x u & \partial_y u \\ \partial_x v & \partial_y v \end{pmatrix} = |\partial_x u|^2 + |\partial_x v|^2 = |f'|^2$$

und dem Satz über die Umkehrabbildung, Satz 6.8, (i)  $\Leftrightarrow$  (iii). Weiterhin ist nach Bemerkung 6.6 mit  $Df$  auch die Jacobimatrix der Umkehrabbildung  $D(f^{-1}) = (Df)^{-1}$  eine Drehstreckung und somit  $f^{-1}$  holomorph.  $\square$

**10.58 Satz. Satz von der Gebietstreue**

Sei  $D \subset \mathbb{C}$  ein Gebiet und  $f : D \rightarrow \mathbb{C}$  holomorph und nicht konstant. Dann ist  $f(D)$  auch ein Gebiet.

*Beweis.* Es ist  $f(D)$  offenbar zusammenhängend, da jeder Weg zwischen zwei Urbildpunkten, auf einen Weg zwischen den Bildpunkten abgebildet wird. Dass eine nichtkonstante holomorphe Funktion offene Mengen auf offene Mengen abbildet, folgt mit etwas Mühe aus Satz 10.53 (a).  $\square$

**10.59 Satz. Maximumsprinzip**

Sei  $D \subset \mathbb{C}$  ein Gebiet und  $f : D \rightarrow \mathbb{C}$  holomorph. Falls  $|f|$  ein lokales Maximum besitzt, so ist  $f$  konstant.

*Beweis.* Es habe  $f$  ein lokales Maximum bei  $a$ , d.h. es gibt eine Umgebung  $U$  von  $a$ , sodass  $|f(z)| \leq |f(a)|$  für alle  $z \in U$ . Angenommen,  $f$  ist nicht konstant, dann ist  $f(U)$  Umgebung von  $f(a)$ , was aber im Widerspruch zu  $|f(z)| \leq |f(a)|$  für alle  $z \in U$  steht.  $\square$

**10.60 Folgerung.** Sei  $D \subset \mathbb{C}$  ein Gebiet,  $K \subset D$  kompakt und  $f : D \rightarrow \mathbb{C}$  holomorph und nicht konstant. Dann gilt

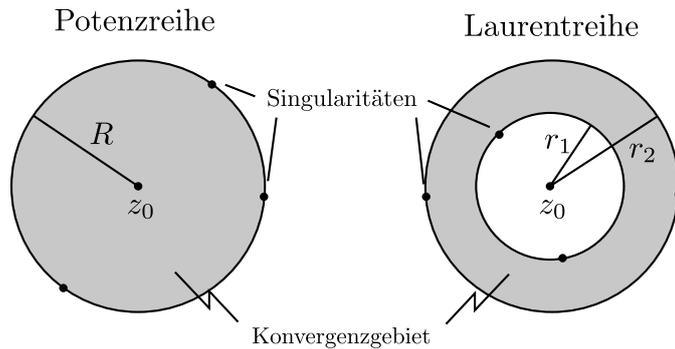
$$\sup_{z \in K} |f(z)| < \sup_{z \notin K} |f(z)|.$$

### 10.6 Laurentreihen und isolierte Singularitäten

Sei  $z_0 \in \mathbb{C}$  und  $0 \leq r_1 < r_2 \leq \infty$ .  
Dann heißt

$$D = \{z \in \mathbb{C} \mid r_1 < |z - z_0| < r_2\}$$

ein Kreisring um  $z_0$  mit Radien  $r_1, r_2$ .



#### 10.61 Satz. Laurentreihe

Jede auf dem Kreisring  $D = \{z \in \mathbb{C} \mid r_1 < |z - z_0| < r_2\}$  holomorphe Funktion  $f$  lässt sich auf  $D$  in eine kompakt konvergente Reihe

$$f(z) = \sum_{n \in \mathbb{Z}} c_n (z - z_0)^n$$

entwickeln, genannt Laurentreihe von  $f$  auf  $D$ . Die Koeffizienten sind durch

$$c_n = \frac{1}{2\pi i} \int_{|z-z_0|=\varrho} \frac{f(z)}{(z - z_0)^{n+1}} dz$$

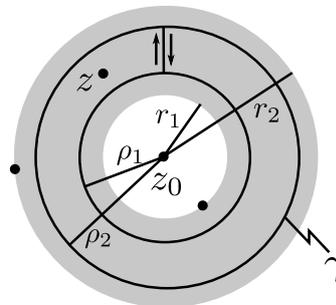
für alle  $n \in \mathbb{Z}$  und ein beliebiges  $r_1 < \varrho < r_2$  eindeutig gegeben.

*Beweis.* Da alle Kreislinien  $|z - z_0| = \varrho$ , die einmal positiv durchlaufen werden, homotop sind, ist die Definition der  $c_n$  unabhängig von  $\varrho$ .

Sei nun o.B.d.A.  $z_0 = 0$ . Sei  $z \in D$  und wähle  $\varrho_1, \varrho_2$  mit

$$r_1 < \varrho_1 < |z| < \varrho_2 < r_2$$

und sei  $\gamma$  der skizzierte Weg.



Dann gilt  $I(\gamma, z) = 1$  und die Cauchy-Integralformel liefert

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\xi)}{\xi - z} d\xi = \underbrace{\frac{1}{2\pi i} \int_{|\xi|=\varrho_2} \frac{f(\xi)}{\xi - z} d\xi}_{=: f_2(z)} - \underbrace{\frac{1}{2\pi i} \int_{|\xi|=\varrho_1} \frac{f(\xi)}{\xi - z} d\xi}_{=: f_1(z)}$$

Für  $f_2$  ergibt sich wieder mit  $|\xi| = \varrho_2 > |z|$ , dass

$$\frac{1}{\xi - z} = \frac{1}{\xi} \cdot \frac{1}{1 - \frac{z}{\xi}} = \frac{1}{\xi} \sum_{n=0}^{\infty} \left(\frac{z}{\xi}\right)^n.$$

Also ist

$$f_2(z) = \sum_{n=0}^{\infty} \left( \frac{1}{2\pi i} \int_{|\xi|=\varrho_2} \frac{f(\xi)}{\xi^{n+1}} d\xi \right) z^n = \sum_{n=0}^{\infty} c_n z^n$$

gleichmäßig konvergent auf jedem Kompaktum  $K \subset \{\varrho_1 < |z| < \varrho_2\}$ .

Für  $f_1$  ergibt sich mit  $|\xi| = \varrho_1 < |z|$ , dass

$$\frac{1}{\xi - z} = -\frac{1}{z} \frac{1}{1 - \frac{\xi}{z}} = -\sum_{n=0}^{\infty} \frac{\xi^n}{z^{n+1}} = -\sum_{n=-\infty}^{-1} \frac{z^n}{\xi^{n+1}}.$$

Also ist

$$f_1(z) = -\sum_{n=-\infty}^{-1} \left( \frac{1}{2\pi i} \int_{|\xi|=\varrho_1} \frac{f(\xi)}{\xi^{n+1}} d\xi \right) z^n = -\sum_{n=-\infty}^{-1} c_n z^n$$

gleichmäßig konvergent auf jedem Kompaktum  $K \subset \{\varrho_1 < |z| < \varrho_2\}$ .

Insgesamt konvergiert also

$$f(z) = f_2(z) - f_1(z) = \sum_{n \in \mathbb{Z}} c_n z^n$$

kompakt auf  $D$  gegen  $f$ .

Diese Darstellung ist eindeutig: Angenommen,  $f(z) = \sum_{n \in \mathbb{Z}} d_n z^n$  ist ebenfalls kompakt konvergent auf  $D$ . Dann gilt für jedes  $n \in \mathbb{Z}$

$$\frac{f(z)}{z^{n+1}} = \frac{d_n}{z} + \sum_{k \neq n} d_k z^{k-n-1},$$

also

$$d_n = \frac{d_n}{2\pi i} \int_{|z|=\varrho} \frac{1}{z} dz = \frac{1}{2\pi i} \int_{|z|=\varrho} \frac{f(z)}{z^{n+1}} dz - \underbrace{\frac{1}{2\pi i} \int \sum_{k \neq n} d_k z^{k-n-1} dz}_{=0} = c_n.$$

□

### 10.62 Definition. Isolierte Singularität

Sei  $D \subset \mathbb{C}$  ein Gebiet  $z_0 \in D$ ,  $f : D \setminus \{z_0\} \rightarrow \mathbb{C}$  holomorph und  $f(z) = \sum_{n \in \mathbb{Z}} c_n (z - z_0)^n$  die Laurentreihe von  $f$  auf  $B_r(z_0) \setminus \{z_0\}$  für  $r$  mit  $B_r(z_0) \subset D$ . Dann heißt  $z_0$  **isolierte Singularität** von  $f$  und  $\omega_f(z_0) = \inf\{n \in \mathbb{Z} \mid c_n \neq 0\} \in \mathbb{Z} \cup \{\pm\infty\}$  ihre **Ordnung**.

Im Allgemeinen ist  $z_0$  hier einfach nur eine isolierter Punkt und die Funktion  $f : D \setminus \{z_0\} \rightarrow \mathbb{C}$  braucht keineswegs singulär zu sein bei  $z_0$ . Vielmehr sagt uns der nächste Satz, wann wir  $f$  zu einer holomorphen Funktion auf ganz  $D$  fortsetzen können.

### 10.63 Satz. Riemannscher Hebbarkeitssatz

Sei  $D \subset \mathbb{C}$  offen und  $z_0 \in D$  isolierte Singularität der holomorphen Funktion  $f : D \setminus \{z_0\} \rightarrow \mathbb{C}$ . Dann sind äquivalent:

- (i)  $\omega_f(z_0) \geq 0$
- (ii)  $f$  ist holomorph fortsetzbar in  $z_0$ , d.h. es gibt eine holomorphe Funktion  $\tilde{f} : D \rightarrow \mathbb{C}$  mit  $\tilde{f}|_{D \setminus \{z_0\}} = f$ .
- (iii) Es gibt eine Umgebung  $U$  von  $z_0$ , so dass  $f|_{U \setminus \{z_0\}}$  beschränkt ist.
- (iv)  $\lim_{z \rightarrow z_0} (z - z_0)f(z) = 0$ .

*Beweis.* (i)  $\Rightarrow$  (ii): Laurentreihe = Potenzreihe

(ii)  $\Rightarrow$  (iii)  $\Rightarrow$  (iv): klar

(iv)  $\Rightarrow$  (i):  $g(z) = (z - z_0)f(z)$  ist stetig bei  $z_0$  und deshalb nach Übungsaufgabe 56 holomorph auf  $D$ . Da  $g(z_0) = 0$  ist, ist  $\omega_g(z_0) \geq 1$  und  $\omega_f(z_0) \geq 0$ . □

**10.64 Satz. Verhalten bei Singularitäten endlicher Ordnung**

Sei  $D \subset \mathbb{C}$  offen und  $z_0 \in D$  isolierte Singularität der holomorphen Funktion  $f : D \setminus \{z_0\} \rightarrow \mathbb{C}$ . Dann sind äquivalent:

- (i)  $-\infty < \omega_f(z_0) < 0$
- (ii) zu jedem  $M > 0$  existiert eine Umgebung  $U \subset \mathbb{C}$  von  $z_0$  mit  $|f(z)| \geq M$  für alle  $z \in U \cap D$   
also  $\lim_{z \rightarrow z_0} |f(z)| = +\infty$ .

*Beweis.* (i)  $\Rightarrow$  (ii): Sei  $k = -\omega_f(z_0)$ . Dann ist nach Satz 10.63  $g(z) = f(z)(z - z_0)^k$  holomorph fortsetzbar auf  $D$  mit  $g(z_0) \neq 0$ . (ii) folgt dann wegen  $f(z) = \frac{g(z)}{(z - z_0)^k}$ .

(ii)  $\Rightarrow$  (i) Wähle eine zu  $M = 1$  entsprechende Umgebung  $U$  und setze auf  $U$

$$g(z) = \begin{cases} 0 & z = z_0 \\ \frac{1}{f(z)} & \text{sonst.} \end{cases}$$

Die Funktion  $g(z)$  ist holomorph auf  $U \setminus \{z_0\}$  und stetig auf  $U$ , also holomorph auf ganz  $U$ . Damit ist  $g(z) = (z - z_0)^k h(z)$  für ein  $k \geq 1$  und  $h$  holomorph auf  $U$  mit  $h(z_0) \neq 0$ , also  $h(z) \neq 0$  auf ganz  $U$ . Damit ergibt sich

$$f(z) = (z - z_0)^{-k} \cdot h(z)^{-1} = (z - z_0)^{-k} \sum_{n=0}^{\infty} h_n (z - z_0)^n,$$

also  $\omega_f(z_0) = -k$ . □

**10.65 Definition. Pole und wesentliche Singularitäten**

Sei  $D \subset \mathbb{C}$  offen und  $z_0 \in D$  eine isolierte Singularität von  $f : D \setminus \{z_0\} \rightarrow \mathbb{C}$ . Dann heißt  $z_0$

- (i) hebbare Singularität, falls  $\omega_f(z_0) \geq 0$ .
- (ii) Pol der Ordnung  $-\omega_f(z_0)$ , falls  $-\infty < \omega_f(z_0) < 0$ .
- (iii) wesentliche Singularität, falls  $\omega_f(z_0) = -\infty$ .

**10.66 Beispiel.** Die Funktion  $\frac{z}{\sin z}$  hat eine hebbare Singularität in  $z = 0$  und einen Pol der Ordnung 1 in  $\pi \cdot n$  für  $n \in \mathbb{Z} \setminus \{0\}$ .

Die Funktion  $e^{\frac{1}{z}}$  hat eine wesentliche Singularität in  $z = 0$ , denn ihre Laurentreihe bei  $z = 0$  ist

$$e^{\frac{1}{z}} = \sum_{n=0}^{\infty} \left(\frac{1}{z}\right)^n \frac{1}{n!} = \sum_{n=-\infty}^0 \frac{1}{(-n)!} z^n.$$

**10.67 Satz. Casorati-Weierstraß**

Sei  $D \subset \mathbb{C}$  offen und  $z_0 \in D$  eine wesentliche Singularität der holomorphen Funktion  $f : D \setminus \{z_0\} \rightarrow \mathbb{C}$ . Dann ist  $f(U \cap (D \setminus \{z_0\}))$  dicht in  $\mathbb{C}$  für jede Umgebung  $U$  von  $z_0$ .

*Beweis.* Übungsaufgabe. □

**10.7 Der Residuenkalkül**

Im Folgenden sei  $D \subset \mathbb{C}$  offen,  $\Delta \subset D$  eine diskrete Teilmenge und  $f : D \setminus \Delta \rightarrow \mathbb{C}$  holomorph. Es ist also  $f$  mit Ausnahme isolierter Singularitäten in  $D$  holomorph. Falls keine der Singularitäten eine wesentliche ist, so nennt man  $f$  **meromorph** in  $D$ .

**10.68 Definition. Residuum**

Sei  $f : D \setminus \Delta \rightarrow \mathbb{C}$  holomorph ( $D$  offen,  $\Delta \subset D$  diskret). Für  $z_0 \in D$  sei  $\sum_{n \in \mathbb{Z}} c_n(z - z_0)^n$  die Laurentreihe um  $z_0$ . Dann heißt

$$\text{Res}_f(z_0) := c_{-1} = \frac{1}{2\pi i} \int_{|z-z_0|=r} f(z) dz \quad (\text{für } r \text{ klein genug})$$

das **Residuum** von  $f$  bei  $z_0$ .

Offensichtlich ist die Menge  $\{z \in D \mid \text{Res}_f(z) \neq 0\}$  in  $\Delta$  enthalten und somit auch diskret in  $D$ .

**10.69 Satz. Residuensatz**

Sei  $D \subset \mathbb{C}$  offen,  $\Delta \subset D$  diskret und  $f : D \setminus \Delta \rightarrow \mathbb{C}$  holomorph. Dann gilt für jede stückweise glatte geschlossene Kurve  $\gamma$  in  $D \setminus \Delta$  die in  $D$  nullhomotop ist:

$$\int_{\gamma} f(z) dz = 2\pi i \sum_{a \in \Delta} I(\gamma, a) \text{Res}_f(a).$$



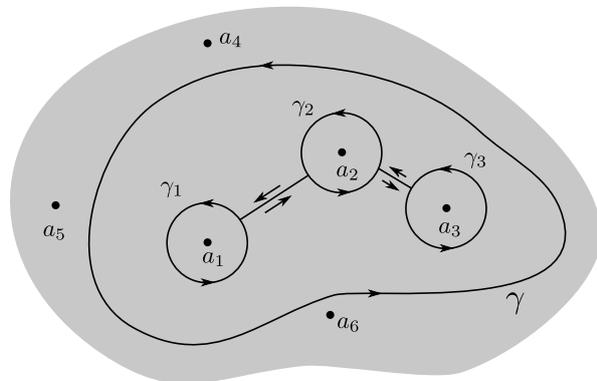
endliche Summe, da  $I(\gamma, a) \neq 0$  nur für endlich viele  $a \in \Delta$ .

*Beweis.* Wir definieren zunächst die „Addition“ von geschlossenen Wegen: Seien  $\gamma_n$  geschlossene Wege und  $d_n \in \mathbb{Z}$ , dann heißt  $c = \sum_{n=1}^N d_n \gamma_n$  ein Zyklus und man setzt

$$\int_c f(z) dz = \sum_{n=1}^N d_n \int_{\gamma_n} f(z) dz.$$

Zu jedem  $a \in \Delta$  sei nun  $\gamma_a$  ein Weg, der  $a$  einmal positiv im Konvergenzgebiet der Laurentreihe um  $a$  umläuft, aber keinen anderen Punkt in  $\Delta$ . Insbesondere ist  $\text{In}(\gamma_a) \cap \Delta = \{a\}$ . Betrachte nun den Zyklus

$$c = \gamma - \sum_{a \in \Delta} I(\gamma, a) \gamma_a.$$



Es gilt  $\text{In}(c) := \{z \in \mathbb{C} \setminus \text{Sp}(c) \mid I(c, z) \neq 0\} \subset D \setminus \Delta$  und man sagt,  $c$  ist nullhomolog in  $D \setminus \Delta$ . Die homologe Version des Cauchy-Integralsatzes sagt dann:

$$\int_c f(z) dz = 0,$$

also

$$\int_{\gamma} f(z) dz = \sum_{a \in \Delta} I(\gamma, a) \int_{\gamma_a} f(z) dz.$$

Nun ist aber

$$\int_{\gamma_a} f(z) dz = \int_{\gamma_a} \sum_{n \in \mathbb{Z}} c_{n,a}(z - a)^n dz = 2\pi i c_{-1,a} = 2\pi i \text{Res}_f(a).$$

□

**10.70 Bemerkung.** Die folgenden Regeln zur Bestimmung des Residuums ergeben sich unmittelbar aus der Definition:

- (a)  $\text{Res}_f(z_0) = \lim_{z \rightarrow z_0} (z - z_0) \cdot f(z)$ , falls der Grenzwert existiert. Dann ist entweder  $z_0$  hebbar und  $\text{Res}_f(z_0) = 0$  oder  $z_0$  ist ein Pol erster Ordnung.
- (b)  $\text{Res}_{\alpha f + \beta g}(z_0) = \alpha \text{Res}_f(z_0) + \beta \text{Res}_g(z_0)$  für  $\alpha, \beta \in \mathbb{C}$ .
- (c) Hat  $f$  an der Stelle  $z_0$  einen Pol erster Ordnung und ist  $g$  holomorph in einer Umgebung von  $z_0$ , so gilt

$$\text{Res}_{fg}(z_0) = g(z_0) \text{Res}_f(z_0).$$

- (d) Hat  $f$  bei  $z_0$  eine einfache Nullstelle, so folgt aus (a), dass

$$\text{Res}_{\frac{1}{f}}(z_0) = \lim_{z \rightarrow z_0} \frac{z - z_0}{f(z) - 0} = \frac{1}{f'(z_0)}.$$

**10.71 Satz. Integrale der Form  $\int_{-\infty}^{\infty} f(x) dx$**

Sei  $\Delta = \{z_1, z_2, \dots, z_n\} \subset \mathbb{C}$  mit  $\Delta \cap \mathbb{R} = \emptyset$  und sei  $f : \mathbb{C} \setminus \Delta \rightarrow \mathbb{C}$  holomorph. Ferner sei

$$|f(z)| \leq \frac{C}{|z|^2} \quad \text{für } |z| \geq r \quad \text{und } \text{Im } z \geq 0.$$

Dann gilt mit  $\Delta^+ = \{z_j \in \Delta \mid \text{Im } z_j > 0\}$ , dass

$$\int_{-\infty}^{\infty} f(x) dx = 2\pi i \sum_{z_j \in \Delta^+} \text{Res}_f(z_j).$$

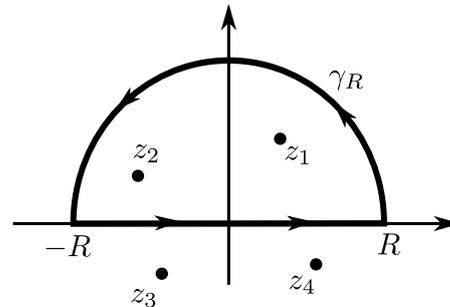
*Beweis.* Sei  $R > r$  und  $\gamma_R$  der skizzierte Weg.

Dann gilt gemäß Residuensatz, dass

$$\int_{\gamma_R} f(z) dz = 2\pi i \sum_{z_j \in \Delta^+} \text{Res}_f(z_j).$$

Andererseits ist

$$\int_{\gamma_R} f(z) dz = \int_{-R}^R f(x) dx + \int_0^{\pi} f(Re^{it}) iR e^{it} dt.$$



Wegen

$$\lim_{R \rightarrow \infty} \left| \int_0^{\pi} f(Re^{it}) iR e^{it} dt \right| \leq \lim_{R \rightarrow \infty} \pi R \frac{C}{R^2} = 0$$

gilt

$$2\pi i \sum_{z_j \in \Delta^+} \text{Res}_f(z_j) = \lim_{R \rightarrow \infty} \int_{\gamma_R} f(z) dz = \lim_{R \rightarrow \infty} \int_{-R}^R f(x) dx = \int_{-\infty}^{\infty} f(x) dx.$$

□

**10.72 Beispiel.** (a) Wir berechnen  $f(x) = \frac{1}{1+x^2}$  zunächst direkt und dann mit Hilfe des Residuenkalküls:

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \arctan x \Big|_{-\infty}^{\infty} = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$$

Andererseits hat

$$\frac{1}{1+z^2} = \frac{1}{(z-i)} \cdot \underbrace{\frac{1}{(z+i)}}_{g(z)}$$

als einzigen Pol in der oberen Halbebene den bei  $+i$  und das Residuum bei  $i$  ist

$$\operatorname{Res}_f(i) = g(i) \cdot \underbrace{\operatorname{Res}_{\frac{1}{z-i}}(i)}_{=1} = \frac{1}{2i},$$

also

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = 2\pi i \operatorname{Res}_f(i) = \pi.$$

- (b) Eine Stammfunktion für  $f(x) = \frac{1}{1+x^4}$  zu finden, ist schon deutlich aufwendiger. Die Rechnung mit Hilfe des Residuenkalküls ist hingegen immer noch ganz einfach: Die vierten Wurzeln aus  $-1$  in der oberen Halbebene sind  $z_1 = e^{i\frac{\pi}{4}}$  und  $z_2 = e^{i\frac{3}{4}\pi}$ . Sei  $g(z) = 1+z^4$ , also  $f(z) = \frac{1}{g(z)}$ . Dann ist

$$\operatorname{Res}_f(z_1) = \frac{1}{g'(z_1)} = \frac{1}{4z_1^3} = \frac{1}{4}z_3 = \frac{1}{4}e^{-\frac{3}{4}i\pi}$$

und

$$\operatorname{Res}_f(z_2) = \frac{1}{g'(z_2)} = \frac{1}{4z_2^3} = \frac{1}{4}z_4 = \frac{1}{4}e^{-\frac{i\pi}{4}}.$$

Also

$$\int_{-\infty}^{\infty} \frac{1}{1+x^4} dx = \frac{2\pi i}{4} \left( e^{-\frac{3}{4}i\pi} + e^{-\frac{i\pi}{4}} \right) = \frac{\pi}{2} \left( e^{-\frac{i\pi}{4}} + e^{+\frac{i\pi}{4}} \right) = \pi \cos\left(\frac{\pi}{4}\right) = \frac{\pi}{\sqrt{2}}.$$

- (c) Man muss ein gegebenes reelles Integral nicht notwendigerweise als ein komplexes Wegintegral entlang der reellen Achse auffassen. Insbesondere bei trigonometrischen Funktionen im Integranden lassen sich reelle Integrale auch als geschlossene Wegintegrale im Komplexen auffassen: Zur Berechnung von

$$\int_0^{2\pi} \frac{1}{4+3\cos t} dt$$

setzen wir  $z(t) = e^{it} = \cos t + i \sin t$  also  $\cos t = \frac{1}{2}(z + \frac{1}{z})$ , also

$$\frac{1}{4+3\cos t} = \frac{1}{4+\frac{3}{2}(z+\frac{1}{z})} = \frac{z}{\frac{3}{2}(z^2+\frac{8}{3}z+1)} = \frac{2z}{3(z-z_1)(z-z_2)},$$

mit  $z_1 = -\frac{4}{3} + \sqrt{\frac{7}{9}}$  und  $z_2 = -\frac{4}{3} - \sqrt{\frac{7}{9}}$ . Setzen wir  $\gamma(t) = z(t) = e^{it}$ , also  $\dot{\gamma}(t) = ie^{it} = iz$ , so ergibt sich

$$\int_{\gamma} \frac{2z}{3(z-z_1)(z-z_2)} \cdot \frac{1}{iz} dz = \int_0^{2\pi} \frac{1}{4+3\cos t} \cdot \frac{ie^{it}}{ie^{it}} dt = \int_0^{2\pi} \frac{1}{4+3\cos t} dt.$$

Andererseits ist

$$\int_{\gamma} \frac{-2i}{3(z-z_1)(z-z_2)} dz = -\frac{2}{3}i \cdot 2\pi i \cdot \operatorname{Res}_{\frac{1}{(z-z_1)(z-z_2)}}(z_1) = \frac{4}{3}\pi \cdot \frac{3}{2\sqrt{7}} = \frac{2}{\sqrt{7}}\pi.$$

**10.73 Satz. Berechnung von Fouriertransformierten**  $\int_{-\infty}^{\infty} f(x) e^{-ikx} dx$

Sei  $\Delta = \{z_1, z_2, \dots, z_n\} \subset \mathbb{C}$  mit  $\Delta \cap \mathbb{R} = \emptyset$  und sei  $f : \mathbb{C} \setminus \Delta \rightarrow \mathbb{C}$  holomorph. Ferner sei

$$|f(z)| \leq \frac{C}{|z|} \quad \text{für } |z| \geq r \text{ und } \text{Im } z \geq 0.$$

Dann gilt mit  $\Delta^+ = \{z_j \in \Delta \mid \text{Im } z_j > 0\}$  und  $\Delta^- = \{z_j \in \Delta \mid \text{Im } z_j < 0\}$ , dass

$$\text{P.V.} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx := \lim_{R \rightarrow \infty} \int_{-R}^R f(x) e^{-ikx} dx = \begin{cases} 2\pi i \sum_{z_j \in \Delta^+} \text{Res}_g(z_j) & \text{für } k < 0, \\ -2\pi i \sum_{z_j \in \Delta^-} \text{Res}_g(z_j) & \text{für } k > 0, \end{cases}$$

wobei  $g(z) := f(z)e^{-ikz}$  ist. Hier steht P.V. für den Cauchy Hauptwert (Cauchy principle value) des Integrals, welches unter den gegebenen Voraussetzungen weder als uneigentliches Riemannintegral noch als Lebesgueintegral zu existieren braucht.

*Beweis.* Sei  $k < 0$ . Analog zum Beweis von Satz 10.71 ist nur zu zeigen, dass

$$\begin{aligned} \lim_{R \rightarrow \infty} \left| \int_0^\pi f(Re^{it}) e^{-ikRe^{it}} iRe^{it} dt \right| &\leq \lim_{R \rightarrow \infty} R \int_0^\pi |f(Re^{it})| e^{kR \sin t} dt \\ &\leq C \lim_{R \rightarrow \infty} \int_0^\pi e^{kR \sin t} dt \\ &\leq 2C \lim_{R \rightarrow \infty} \int_0^{\pi/2} e^{kR 2t/\pi} dt \\ &\leq 2C \lim_{R \rightarrow \infty} \left( \frac{\pi e^{kR 2t/\pi}}{2kR} \right) \Big|_0^{\pi/2} = 0. \end{aligned}$$

Für den Fall  $k > 0$  mache man sich klar, dass für  $h(z) := f(-z)$  gilt  $\text{Res}_h(z) = -\text{Res}_f(-z)$ .  $\square$

**10.74 Beispiel.** Die Fouriertransformierte von  $f(x) = \frac{1}{1+x^4}$  ist somit für  $k < 0$

$$\begin{aligned} \hat{f}(k) &= \int_{-\infty}^{\infty} \frac{1}{1+x^4} e^{-ikx} dx = 2\pi i \left( \text{Res}_f(z_1) e^{-ikz_1} + \text{Res}_f(z_2) e^{-ikz_2} \right) \\ &= \frac{2\pi i}{4} \left( \bar{z}_2 e^{-ikz_1} + \bar{z}_1 e^{-ikz_2} \right) = \frac{\pi i}{2} \left( -z_1 e^{-ikz_1} + \bar{z}_1 e^{ik\bar{z}_1} \right) \\ &= \frac{\pi}{2i} \left( z_1 e^{-ikz_1} - \bar{z}_1 e^{ik\bar{z}_1} \right) = \pi \text{Im} \left( z_1 e^{-ikz_1} \right) = \pi \text{Im} \left( e^{\frac{i\pi}{4}} e^{-\frac{ik(1+i)}{\sqrt{2}}} \right) \\ &= \pi e^{-\frac{|k|}{\sqrt{2}}} \text{Im} \left( e^{i\left(\frac{\pi}{4} + \frac{|k|}{\sqrt{2}}\right)} \right) = \pi e^{-\frac{|k|}{\sqrt{2}}} \sin \left( \frac{\pi}{4} + \frac{|k|}{\sqrt{2}} \right). \end{aligned}$$

Bei der Berechnung der Residuen haben wir verwendet, dass  $f$  nur einfache Pole hat. Da für reellwertige Funktionen  $f(x)$  gilt, dass  $\hat{f}(-k) = \overline{\hat{f}(k)}$ , stimmt die obige Formel auch für  $k > 0$ .



# 11 Abstrakte Maß- und Integrationstheorie

## Grundideen:

### (a) Das Regel- bzw. Riemannintegral

Zerlege die Urbildmenge („ $x$ -Achse“) in endlich viele Intervalle  $[a_i, a_{i+1})$  und approximiere die zu integrierende Funktion  $f$  durch **Treppenfunktionen**

$$g(x) = \sum_{i=1}^n \alpha_i \chi_{[a_i, a_{i+1})}(x),$$

wobei für  $A \subset \mathbb{R}$  die charakteristische Funktion der Menge  $A$  gegeben ist durch

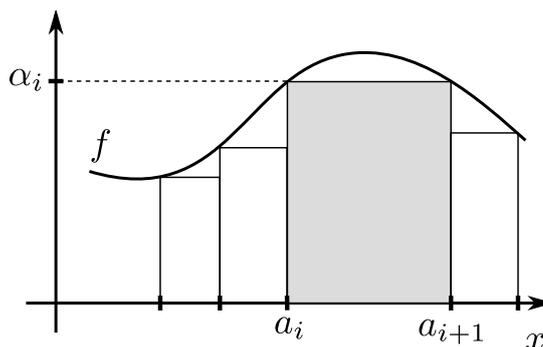
$$\chi_A(x) = \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{sonst.} \end{cases}$$

Die Werte  $\alpha_i$  werden durch die Funktion  $f$  bestimmt, z.B.  $\alpha_i = \inf\{f(x) \mid x \in [a_i, a_{i+1})\}$ . Das Integral über eine Treppenfunktion ist dann elementar definiert durch

$$\int g(x) dx := \sum_{i=1}^n \alpha_i (a_{i+1} - a_i)$$

und man erhält das Integral über geeignete Limesbildung.

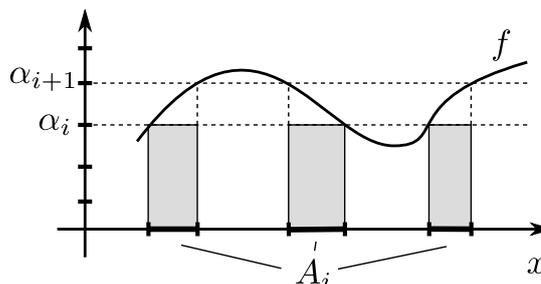
**Prinzip:** Einteilung des Urbildes in endlich viele Intervalle a priori und Gewichtung der Länge eines Intervalls mit Funktionswerten bei der Integration.



### (b) Das Lebesgue-Integral

Zerlege den Bildbereich („ $y$ -Achse“) in endlich viele Intervalle  $[\alpha_i, \alpha_{i+1})$  und approximiere die zu integrierende Funktion  $f$  durch sogenannte **einfache Funktionen**

$$g(x) = \sum_{i=1}^n \alpha_i \chi_{A_i}(x),$$



also Funktionen die auf Mengen  $A_i \subset \mathbb{R}$  konstant sind, welche durch die Funktion  $f$  definiert werden, z.B.  $A_i = f^{-1}([\alpha_i, \alpha_{i+1}))$ . Hier bezeichnet  $f^{-1}$  nicht die Umkehrabbildung sondern das Urbild. Die Mengen  $A_i$  sind dann im Allgemeinen keine Intervalle.

Das Integral für einfache Funktionen ist wieder elementar,

$$\int g(x) dx = \sum_{i=1}^n \alpha_i \lambda(A_i),$$

vorausgesetzt, wir können der „Länge“ bzw. dem Maß  $\lambda(A_i)$  jeder Menge  $A_i$  einen präzisen Sinn geben.

**11.1 Beispiel.** Die Funktion  $f(x) = \chi_{\mathbb{Q} \cap [0,1]}(x)$  ist weder Regel- noch Riemannintegrierbar, aber wegen

$$\lambda(\mathbb{Q} \cap [0, 1]) = 0 \quad (\text{intuitiv klar})$$

sollte

$$\int_0^1 f(x) dx = 1 \cdot \lambda(\mathbb{Q} \cap [0, 1]) + 0 \cdot \underbrace{\lambda([0, 1] \setminus \mathbb{Q})}_{=1} = 0$$

gelten.

**Prinzip:** Beim Lebesgueintegral zerlegt man statt des Urbildbereichs zunächst den Bildbereich. Es wird sich zeigen, dass dieser Ansatz zwei große Vorteile hat:

- a) Es gibt „mehr“ integrierbare Funktionen  $\Rightarrow$  Räume Lebesgue-integrierbarer Funktionen sind vollständig.
- b) Das Prinzip ist leicht auf andere Urbildmengen als  $\mathbb{R}$  oder  $\mathbb{R}^n$  verallgemeinerbar, da keine Zerlegung des Urbildbereichs in Intervalle oder Würfel notwendig ist.

**Aber:** Alles hängt an der Definition geeigneter Maße  $\lambda$ .

## 11.1 Das Inhaltsproblem

Der Inhalt eines Intervalls  $[a, b] \subset \mathbb{R}$  ist seine Länge,  $\lambda([a, b]) = b - a$ . Dieser Inhalt ist

- (a) **translationsinvariant**, d.h.

$$\lambda([a + c, b + c]) = \lambda([a, b]) \quad \forall c \in \mathbb{R},$$

- (b) **additiv** für disjunkte Intervalle,

$$\lambda\left(\bigcup_i [a_i, b_i]\right) = \sum_i (b_i - a_i),$$

- (c) **subadditiv** für überlappende Intervalle,

$$\lambda\left(\bigcup_i [a_i, b_i]\right) \leq \sum_i (b_i - a_i).$$

**Frage:** Kann man diesen primitiven Inhaltsbegriff unter Beibehaltung der Eigenschaften (a), (b) und (c) auf beliebige Teilmengen von  $\mathbb{R}$  erweitern?

**Antwort:** Nein! Es gibt keine Abbildung

$$\lambda : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty] \quad (\mathcal{P} = \text{Potenzmenge} = \text{Menge aller Teilmengen})$$

mit den Eigenschaften

- (a)  $\lambda([0, 1]) = 1$
- (b)  $\lambda(A + c) = \lambda(A) \quad \forall A \subset \mathbb{R}, c \in \mathbb{R}$ ,
- (c)  $\lambda(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \lambda(A_i)$  falls  $A_i \cap A_j = \emptyset$  für  $i \neq j$  ( $\sigma$ -Additivität).

Um das einzusehen, betrachten wir das Intervall  $[0, 1]$  als Kreisring und Verschiebungen modulo 1, also  $\mathbb{R}/\mathbb{Z}$ . Wir konstruieren gleich eine disjunkte Zerlegung des  $[0, 1]$ -Rings in abzählbar unendlich

viele Mengen  $V_j$ ,  $j \in \mathbb{N}$ , die alle durch Translation auseinander hervorgehen,  $V_j = V + r_j$ . Dann folgt mit Normierung,  $\sigma$ -Additivität und Translationsinvarianz

$$\begin{aligned} 1 &\stackrel{(a)}{=} \lambda([0, 1]) = \lambda\left(\bigcup_{j=1}^{\infty} V_j\right) \stackrel{(c)}{=} \sum_{j=1}^{\infty} \lambda(V_j) \\ &= \sum_{j=1}^{\infty} \lambda(V + r_j) \stackrel{(b)}{=} \sum_{j=1}^{\infty} \lambda(V) = \lambda(V) \cdot \sum_{j=1}^{\infty} 1 = \begin{cases} 0 & \text{falls } \lambda(V) = 0 \\ \infty & \text{sonst.} \end{cases} \end{aligned}$$

Der Menge  $V$  kann man also keinen solchen Inhalt zuweisen, man sagt sie ist nicht „messbar“.

Um so ein  $V$  zu konstruieren, betrachten wir die Äquivalenzrelation  $x \sim y \Leftrightarrow x - y \in \mathbb{Q}$  auf  $\mathbb{R}/\mathbb{Z}$ . Die Äquivalenzklassen  $[x]$  bilden eine disjunkte Zerlegung von  $\mathbb{R}/\mathbb{Z}$  und wir können aus jeder einen Vertreter  $x$  wählen. Die Menge dieser Vertreter nennen wir  $V$  und setzen  $V_r := V + r$ ,  $r \in \mathbb{Q} \cap [0, 1]$ . Dann gilt

$$V_r \cap V_{r'} = \emptyset \quad \text{für } r \neq r'$$

und

$$\bigcup_{r \in \mathbb{Q} \cap [0, 1]} V_r = [0, 1].$$

**Fazit:** Es existieren Mengen, denen man keinen Inhalt zuweisen kann, falls man die Eigenschaften (a), (b) und (c) von dem Inhaltsbegriff fordert. Das ist nicht weiter schlimm und soll an dieser Stelle nur die Tatsache motivieren, dass wir im Folgenden Maße auf Teilmengensystemen betrachten, die nicht notwendigerweise die Potenzmenge sind.

**11.2 Bemerkung.** Bei der Konstruktion von  $V$  haben wir das **Auswahlaxiom** verwendet, da wir aus jeder Äquivalenzklasse  $[x]$  einen Vertreter ausgewählt haben, ohne diesen genau zu spezifizieren.

### 11.3 Bemerkung. Der Satz von Banach-Tarski

Man könnte hoffen, dass alle Mengen messbar werden, wenn man statt  $\sigma$ -Additivität nur endliche Additivität fordert. Auf  $\mathbb{R}$  ist das tatsächlich der Fall: Es gibt einen endlich additiven und translationsinvarianten Inhalt auf  $\mathbb{R}$ . In  $\mathbb{R}^d$  mit  $d \geq 3$  gibt es aber schon nicht messbare Mengen, wenn man nur endliche Additivität und Invarianz unter Bewegungen, also Translationen und Rotationen, fordert:

Eine Kugel im  $\mathbb{R}^3$  kann so in endlich viele disjunkte Teilmengen zerlegt werden, dass aus diesen Teilen allein durch starres Verschieben und Rotieren zwei Kugeln der ursprünglichen Größe gebildet werden können.

Diese Konstruktion geht auf Banach und Tarski zurück und heißt deshalb auch das Banach-Tarski “Paradoxon”. Es handelt sich hierbei um ein Paradoxon im Sinne eines scheinbaren Widerspruchs bzw. eines Widerspruchs gegen die Intuition. Mathematisch gesehen liegt hier kein Widerspruch vor. Der Beweis wird für Interessierte in einer Zusatzstunde besprochen.

## 11.2 Grundzüge der Maßtheorie

Ein Maß  $\mu$  ordnet Teilmengen  $A \subset M$  einer Obermenge  $M$  eine Maßzahl  $\mu(A) \in [0, \infty]$  zu. Für Teilmengen von  $\mathbb{R}$ ,  $\mathbb{R}^2$ ,  $\mathbb{R}^3$  spricht man dann beispielsweise von Länge, Fläche und Volumen. Ein anderes Beispiel sind Wahrscheinlichkeitsmaße, die Mengen von “Ereignissen” Wahrscheinlichkeiten zuordnen.

Wie wir am Beispiel  $\mathbb{R}$  bereits gezeigt haben, ist es im Allgemeinen nicht möglich, allen Teilmengen einer Menge eine Maßzahl zuzuordnen, wenn man zusätzliche Eigenschaften (wie z.B.

Translationsinvarianz) fordert. Man muss sich also auf geeignete Systeme von Teilmengen beschränken:

**11.4 Definition.  $\sigma$ -Algebra**

Eine Familie  $\mathcal{A} \subseteq \mathcal{P}(X)$  von Teilmengen einer Menge  $X$  heißt  **$\sigma$ -Algebra** auf  $X$ , falls

- (i)  $\emptyset \in \mathcal{A}$ ,
- (ii)  $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ ,
- (iii)  $A_k \in \mathcal{A}$  für  $k \in \mathbb{N} \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$ .

Eine  $\sigma$ -Algebra ist also ein System von Teilmengen, welches die leere Menge enthält und abgeschlossen ist unter Komplementbildung und abzählbaren Vereinigungen.

**11.5 Proposition.** Sei  $\mathcal{A}$  eine  $\sigma$ -Algebra auf  $X$ . Dann gilt

- (a)  $X \in \mathcal{A}$
- (b)  $A_k \in \mathcal{A}$  für  $k \in \mathbb{N} \Rightarrow \bigcap_{k=1}^{\infty} A_k \in \mathcal{A}$
- (c)  $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}, A \cap B \in \mathcal{A}$  und  $A \setminus B \in \mathcal{A}$ .

*Beweis.* Übungsaufgabe. □

**11.6 Beispiel.** (a)  $\mathcal{P}(X)$  und  $\{\emptyset, X\}$  sind  $\sigma$ -Algebren auf  $X$ .

(b) Sind  $\mathcal{A}_j, j \in I$ ,  $\sigma$ -Algebren auf  $X$ , so ist auch  $\bigcap_{j \in I} \mathcal{A}_j$  eine  $\sigma$ -Algebra.

**11.7 Definition. Erzeugendensystem**

Aus Beispiel 11.6 (b) folgt insbesondere, dass jede Familie  $\mathcal{F} \subset \mathcal{P}(X)$  eine kleinste  $\sigma$ -Algebra **erzeugt**, welche  $\mathcal{F}$  enthält:

$$\mathcal{A}_{\mathcal{F}} := \bigcap_{\substack{\mathcal{B} \text{ ist } \sigma\text{-Algebra} \\ \text{mit } \mathcal{F} \subset \mathcal{B}}} \mathcal{B}.$$

Jede Familie  $\mathcal{F} \subset \mathcal{P}(X)$  die  $\mathcal{A}$  erzeugt, heißt **Erzeugendensystem** für  $\mathcal{A}$ .

**11.8 Definition. Maß, Messraum, Maßraum, messbare Mengen**

Sei  $\mathcal{A} \subset \mathcal{P}(X)$  eine  $\sigma$ -Algebra auf  $X$ .

- (a) Eine Abbildung  $\mu : \mathcal{A} \rightarrow [0, \infty]$  heißt **Maß**, falls
  - (i)  $\mu(\emptyset) = 0$
  - (ii) Für paarweise disjunkte  $A_k \in \mathcal{A}, k \in \mathbb{N}$ , gilt

$$\mu \left( \bigcup_{k=1}^{\infty} A_k \right) = \sum_{k=1}^{\infty} \mu(A_k) \quad (\sigma\text{-Additivität}).$$

Falls  $\mu(X) < \infty$ , so heißt  $\mu$  ein **endliches Maß**.

Falls  $X = \bigcup_{k=1}^{\infty} A_k$  mit  $\mu(A_k) < \infty \forall k \in \mathbb{N}$ , so heißt  $\mu$  ein  **$\sigma$ -endliches Maß**.

- (b) Das Tupel  $(X, \mathcal{A})$  heißt **Messraum**, das Tripel  $(X, \mathcal{A}, \mu)$  heißt **Maßraum**. Die Elemente von  $\mathcal{A}$  heißen die  **$\mathcal{A}$ -messbaren Mengen**.

**11.9 Beispiel. Zählmaß und Diracmaß**

(a) Das **Zählmaß**  $\nu$  ist auf der Potenzmenge jeder Menge  $X$  definiert: Für  $A \subset X$  ist

$$\nu(A) := \begin{cases} |A| = \text{Anzahl der Elemente von } A & \text{falls } A \text{ endlich ist,} \\ \infty & \text{sonst.} \end{cases}$$

(b) Ebenfalls auf  $\mathcal{P}(X)$  ist das **Diracmaß**  $\delta_{x_0}$  bei  $x_0 \in X$  definiert durch

$$\delta_{x_0}(A) = \begin{cases} 1 & \text{falls } x_0 \in A, \\ 0 & \text{sonst.} \end{cases}$$

Aus der Definition ergeben sich direkt die folgenden elementaren Eigenschaften von Maßen.

**11.10 Proposition.** Sei  $\mu$  ein Maß auf  $(X, \mathcal{A})$ .

(a) Seien  $A, B \in \mathcal{A}$  mit  $A \subset B$ . Dann gilt

$$\mu(B) = \mu(A) + \mu(B \setminus A).$$

Insbesondere ist also  $\mu(A) \leq \mu(B)$ .

(b) Seien  $A, B \in \mathcal{A}$ . Dann gilt

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

(c) Seien  $A_j \in \mathcal{A}$  für  $j \in \mathbb{N}$ . Dann gilt

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mu(A_j) \quad (\text{Subadditivität}).$$

(d) Sei  $(A_j)$  eine aufsteigende Folge messbarer Mengen, also  $A_j \in \mathcal{A}$  und  $A_j \subset A_{j+1}$  für alle  $j \in \mathbb{N}$ , dann gilt

$$\lim_{j \rightarrow \infty} \mu(A_j) = \mu\left(\bigcup_{j=1}^{\infty} A_j\right).$$

*Beweis.* Übungsaufgabe. □

Wie können wir nun unseren translationsinvarianten Inhaltsbegriff für Intervalle in  $\mathbb{R}$  auf Mengen aus einer geeigneten  $\sigma$ -Algebra ausdehnen? Da gibt es im Wesentlichen zwei Möglichkeiten.

**1. Möglichkeit: Erweitern**

Sei  $\mathcal{B}$  die kleinste  $\sigma$ -Algebra welche die offenen Intervalle enthält (also die von den offenen Intervallen erzeugte  $\sigma$ -Algebra) und zeige, dass sich  $\lambda$  in eindeutiger Weise zu einem Maß auf  $\mathcal{B}$  fortsetzen läßt. Diese Strategie führt zu den sogenannten Maß-Erweiterungssätzen, auf die wir aus Zeitgründen nicht weiter eingehen können.

**11.11 Bemerkung. Borel- $\sigma$ -Algebra**

In einem metrischen Raum  $X$  heißt die von den offenen Mengen erzeugte  $\sigma$ -Algebra  $\mathcal{B}$  die **Borel- $\sigma$ -Algebra**. Auf  $\mathbb{R}$  stimmt sie mit der von den offenen Intervallen erzeugten  $\sigma$ -Algebra überein:

**11.12 Lemma.** Sei  $G \subset \mathbb{R}$  offen, dann ist  $G$  abzählbare Vereinigung offener Intervalle

$$G = \bigcup_{j=1}^{\infty} (a_j, b_j)$$

*Beweis.* Wähle zu jedem Punkt  $x \in \mathbb{Q} \cap G$  das größte offene Intervall  $I_x$  mit  $x \in I_x \subset G$ . Dann ist  $G = \bigcup_{x \in \mathbb{Q} \cap G} I_x$ .  $\square$

## 2. Möglichkeit: Einschränken

Definiere ein sog. äußeres Maß  $\lambda^*$  auf allen Teilmengen von  $\mathbb{R}$ , welches auf den Intervallen mit  $\lambda$  übereinstimmt. Dieses kann mit dem zuvor gesagten nicht  $\sigma$ -additiv sein. Finde dann die größte  $\sigma$ -Algebra  $\mathcal{L}$ , so dass  $\lambda^*$  eingeschränkt auf  $\mathcal{L}$  ein Maß ist.

### 11.13 Definition. Äußeres Maß

Eine Abbildung  $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$  heißt äußeres Maß auf der Menge  $X$ , falls

- (i)  $\mu^*(\emptyset) = 0$
- (ii) Falls  $A \subset \bigcup_{j=1}^{\infty} A_j$ , so ist

$$\mu^*(A) \leq \sum_{j=1}^{\infty} \mu^*(A_j) \quad (\text{Subadditivität})$$

Für  $A \subset B$  gilt dann insbesondere  $\mu^*(A) \leq \mu^*(B)$  (Monotonie).

### 11.14 Beispiel. Äußeres Lebesguemaß

Offenbar wird nun für  $A \subset \mathbb{R}$  durch

$$\lambda^*(A) := \inf \left\{ \sum_{j=1}^{\infty} (b_j - a_j) \mid A \subset \bigcup_{j=1}^{\infty} (a_j, b_j), a_j < b_j \in \mathbb{R} \right\}$$

ein äußeres Maß auf  $\mathbb{R}$  definiert und es gilt  $\lambda^*((a, b)) = (b - a)$  (Übungsaufgabe). Es heißt  $\lambda^*$  das äußere Lebesguemaß.

Wie kommt man nun an die messbaren Mengen?

### Lebesgues ursprüngliche Idee:

Definiere auch ein inneres Maß  $\lambda_*$  und zeige, dass  $\{A \subset \mathbb{R} \mid \lambda^*(A) = \lambda_*(A)\}$  eine  $\sigma$ -Algebra ist und auf solchen Mengen durch  $\lambda := \lambda^* = \lambda_*$  ein Maß definiert wird.

### Carathéodorys Trick:

Für Teilmengen  $A \subset [0, 1]$  ist

$$\lambda_*(A) := \lambda^*([0, 1]) - \lambda^*([0, 1] \setminus A) = 1 - \lambda^*([0, 1] \setminus A).$$

Es ist also  $A$  messbar, d.h.  $\lambda_*(A) = \lambda^*(A)$ , genau dann wenn  $\lambda^*(A) + \lambda^*([0, 1] \setminus A) = 1$ , oder

$$\lambda^*(A \cap [0, 1]) + \lambda^*(A^c \cap [0, 1]) = \lambda^*([0, 1]).$$

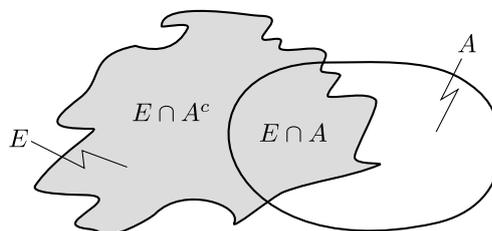
Carathéodorys Idee ist es nun, hier  $[0, 1]$  durch eine beliebige Menge  $E$  zu ersetzen und die Gleichung zur Definition zu erheben.

### 11.15 Definition. $\mu^*$ -Messbarkeit

Sei  $\mu^*$  ein äußeres Maß auf einer Menge  $X$ . Eine Menge  $A \subset X$  heißt  $\mu^*$ -messbar, falls

$$\mu^*(E) = \mu^*(A \cap E) + \mu^*(A^c \cap E) \quad \text{für jedes } E \subset X.$$

$A$  ist also  $\mu^*$  messbar, falls jede Menge  $E$  von  $A$  „sauber“ zerlegt wird, also  $\mu^*$  bzgl. der Zerlegung durch  $A$  additiv ist.



**11.16 Bemerkung.** Wegen der Subadditivität gilt immer

$$\mu^*(E) \leq \mu^*(A \cap E) + \mu^*(A^c \cap E).$$

Man muss also zur Messbarkeit immer nur

$$\mu^*(E) \geq \mu^*(A \cap E) + \mu^*(A^c \cap E)$$

nachweisen.

**11.17 Satz. Carathéodory**

Sei  $\mu^*$  ein äußeres Maß auf einer Menge  $X$ , dann bilden die  $\mu^*$ -messbaren Mengen eine  $\sigma$ -Algebra  $\mathcal{A}$  und  $\mu^*|_{\mathcal{A}}$  ist ein Maß.

*Beweis.* Sei  $\mathcal{A}$  die Menge der  $\mu^*$ -messbaren Mengen.

- (1) Falls  $\mu^*(A) = 0$  ist, so ist  $A \in \mathcal{A}$ , denn die Subadditivität impliziert dann

$$\mu^*(A \cap E) + \mu^*(A^c \cap E) \leq \mu^*(A) + \mu^*(E) = \mu^*(E).$$

Es ist also insbesondere  $\emptyset \in \mathcal{A}$ .

- (2)  $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$  folgt sofort aus der Symmetrie der Definition.  
 (3)  $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cup A_2 \in \mathcal{A}$ , denn

$$\begin{aligned} \mu^*(E) &= \mu^*(A_1 \cap E) + \mu^*(A_1^c \cap E) \\ &= \mu^*(A_1 \cap E) + \mu^*(A_1^c \cap E \cap A_2) + \mu^*(A_1^c \cap E \cap A_2^c) \\ &\geq \mu^*(E \cap (A_1 \cup A_2)) + \mu^*(E \cap (A_1 \cup A_2)^c). \end{aligned}$$

Die erste Gleichheit folgt aus der Messbarkeit von  $A_1$ , die zweite aus der Messbarkeit von  $A_2$ . Die Ungleichung folgt schließlich aus der Subadditivität und den Identitäten

$$\begin{aligned} E \cap (A_1 \cup A_2) &= (E \cap A_1) \cup (E \cap A_1^c \cap A_2) \\ E \cap (A_1 \cup A_2)^c &= E \cap A_1^c \cap A_2^c. \end{aligned}$$

- (4) Für  $A_j \subset X$ ,  $j \in \mathbb{N}$ , paarweise disjunkt und messbar,  $E \subset X$  beliebig,  $B_n := \bigcup_{j=1}^n A_j$  und  $n \in \mathbb{N}$  gilt

$$\mu^*(E \cap B_n) = \sum_{j=1}^n \mu^*(E \cap A_j).$$

Beweis durch Induktion: Der Fall  $n = 1$  ist klar. Wegen (3) ist  $B_n$  messbar, also gilt

$$\begin{aligned} \mu^*(E \cap B_{n+1}) &= \mu^*(E \cap B_{n+1} \cap B_n) + \mu^*(E \cap B_{n+1} \cap B_n^c) \\ &= \mu^*(E \cap B_n) + \mu^*(E \cap A_{n+1}) = \sum_{j=1}^{n+1} \mu^*(E \cap A_{n+1}). \end{aligned}$$

Insbesondere folgt also auch

$$\mu^*(E \cap (\bigcup_{j=1}^{\infty} A_j)) = \sum_{j=1}^{\infty} \mu^*(E \cap A_j),$$

da  $\mu^*(E \cap (\bigcup_{j=1}^{\infty} A_j)) \geq \mu^*(E \cap B_n) = \sum_{j=1}^n \mu^*(E \cap A_j)$  für alle  $n \in \mathbb{N}$ .

(5) Für messbare  $A_j \subset X$ ,  $j \in \mathbb{N}$ , ist auch  $\bigcup_{j=1}^{\infty} A_j$  messbar.

Um das zu sehen, schreiben wir zunächst  $\bigcup_{j=1}^{\infty} A_j$  als disjunkte Vereinigung. Sei

$$\tilde{A}_1 = A_1, \quad \tilde{A}_2 = A_2 \setminus A_1, \quad \tilde{A}_3 = A_3 \setminus (A_1 \cup A_2) \quad \dots$$

dann sind die  $\tilde{A}_j$  messbar und  $\bigcup_{j=1}^n \tilde{A}_j = \bigcup_{j=1}^n A_j$ . Also gilt für beliebiges  $E \subset X$

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap (\bigcup_{j=1}^n A_j)) + \mu^*(E \cap (\bigcup_{j=1}^n A_j)^c) \\ &\geq \mu^*(E \cap (\bigcup_{j=1}^n \tilde{A}_j)) + \mu^*(E \cap (\bigcup_{j=1}^{\infty} A_j)^c) \\ &= \sum_{j=1}^n \mu^*(E \cap \tilde{A}_j) + \mu^*(E \cap (\bigcup_{j=1}^{\infty} A_j)^c) \end{aligned}$$

und für  $n \rightarrow \infty$

$$\begin{aligned} \mu^*(E) &\geq \sum_{j=1}^{\infty} \mu^*(E \cap \tilde{A}_j) + \mu^*(E \cap (\bigcup_{j=1}^{\infty} A_j)^c) \\ &= \mu^*(E \cap (\bigcup_{j=1}^{\infty} A_j)) + \mu^*(E \cap (\bigcup_{j=1}^{\infty} A_j)^c). \end{aligned}$$

□

### 11.18 Definition. Lebesguemaß und Lebesgue- $\sigma$ -Algebra auf $\mathbb{R}^n$

Das äußere Lebesguemaß  $\lambda^*(A)$  einer Menge  $A \subset \mathbb{R}^n$  ist definiert durch das Infimum der Inhalte aller abzählbaren Überdeckungen durch offene Quader,

$$\lambda^*(A) := \inf \left\{ \sum_{j=1}^{\infty} \prod_{k=1}^n (b_{jk} - a_{jk}) \mid A \subset \bigcup_{j=1}^{\infty} \prod_{k=1}^n (a_{jk}, b_{jk}), a_{jk} < b_{jk} \in \mathbb{R} \right\}.$$

Die Restriktion  $\lambda$  von  $\lambda^*$  auf die  $\lambda^*$ -messbaren Mengen heißt das  $n$ -dimensionale **Lebesguemaß** und die  $\sigma$ -Algebra  $\mathcal{L}^n$  der  $\lambda^*$ -messbaren Mengen heißt **Lebesgue- $\sigma$ -Algebra**.

### 11.19 Bemerkung. Translationsinvarianz des Lebesguemaßes

Da offenbar  $\lambda^*$  translationsinvariant ist, ist auch  $\lambda = \lambda^*|_{\mathcal{L}^n}$  translationsinvariant, d.h.

$$\lambda(A + x) = \lambda(A) \quad \forall A \in \mathcal{L}^n, x \in \mathbb{R}^n.$$

### 11.20 Proposition. Borelmengen sind Lebesguemessbar, also $\mathcal{B}^n \subset \mathcal{L}^n$ .

*Beweis.* Es genügt zu zeigen, dass offene Quader  $\prod_{k=1}^n (a_k, b_k)$   $\lambda^*$ -messbar sind, denn diese erzeugen  $\mathcal{B}^n$ : wäre  $\mathcal{B}^n \not\subset \mathcal{L}^n$ , so wäre  $\mathcal{B}^n \cap \mathcal{L}^n \subsetneq \mathcal{B}^n$  eine kleinere  $\sigma$ -Algebra welche alle offenen Quader enthält, im Widerspruch zur Definition von  $\mathcal{B}^n$ . Dass offene Quader messbar sind, ist für den Fall  $n = 1$  eine Übungsaufgabe. □

**11.21 Beispiel.** Es ist  $\mathbb{Q}$  als abzählbare Vereinigung von abgeschlossenen Mengen (Punkten) eine Borel-Menge, also Lebesguemessbar. In den Übungen wurde gezeigt, dass  $\lambda^*(\mathbb{Q}) = 0$ , also ist auch  $\lambda(\mathbb{Q}) = 0$ . Die rationalen Zahlen bilden somit eine Lebesgue-Nullmenge.

**11.22 Bemerkung.** Die  $\sigma$ -Algebra  $\mathcal{B}$  der Borelmengen ist echt kleiner als die  $\sigma$ -Algebra der Lebesguemengen. Man nennt  $\lambda^*|_{\mathcal{B}^n}$  deshalb auch das Lebesgue-Borel-Maß, um es von dem Lebesgue-Maß  $\lambda = \lambda^*|_{\mathcal{L}^n}$  zu unterscheiden.  $\lambda$  erhält man aus  $\lambda^*|_{\mathcal{B}^n}$  durch „Vervollständigung“: Ein Maß  $\mu$  auf einer  $\sigma$ -Algebra  $\mathcal{A}$  heißt vollständig, falls jede Teilmenge einer Nullmenge messbar ist. (siehe Übungsaufgabe).

**11.23 Satz. Eindeutigkeitsatz**

Sei  $(X, \mathcal{A})$  ein Messraum und  $\mathcal{E} \subset \mathcal{A}$  ein durchschnittstabiles Erzeugendensystem für  $\mathcal{A}$ , d.h.  $A, B \in \mathcal{E} \Rightarrow A \cap B \in \mathcal{E}$ . Weiter gebe es eine aufsteigende Folge  $(E_n)$  in  $\mathcal{E}$  mit  $E_1 \subset E_2 \subset E_3 \dots$  und  $\cup_{n=1}^{\infty} E_n = X$ . Sind nun  $\mu_1$  und  $\mu_2$  Maße auf  $(X, \mathcal{A})$  mit

- (i)  $\mu_1|_{\mathcal{E}} = \mu_2|_{\mathcal{E}}$
- (ii)  $\mu_1(E_n) = \mu_2(E_n) < \infty \quad \forall n \in \mathbb{N}$ ,

so gilt bereits  $\mu_1 = \mu_2$ .

*Beweis.* Siehe z.B. Bauer, *Maß- und Integrationstheorie*, Satz 5.4. □

**11.24 Korollar. Eindeutigkeit des Lebesguemaßes**

Das Lebesgue- und das Lebesgue-Borel-Maß sind jeweils eindeutig bestimmt durch die Forderung nach Translationsinvarianz und der Normierung  $\lambda((0, 1)^n) = 1$ .

*Beweis.* Skizze: Translationsvarianz und Normierung fixieren  $\lambda$  auf beliebigen Quadern,

$$\lambda \left( \prod_{j=1}^n (a_j, b_j) \right) = \prod_{j=1}^n (b_j - a_j),$$

welche ein  $\cap$ -stabiles Erzeugendensystem von  $\mathcal{B}^n$  bilden. Also ist  $\lambda$  auf  $\mathcal{B}^n$  eindeutig bestimmt. Auf  $\mathcal{L}^n$  ist  $\lambda$  als Vervollständigung von  $\lambda$  auf  $\mathcal{B}^n$  somit ebenfalls eindeutig bestimmt. □

## 11.3 Das Lebesgueintegral

Wir entwickeln nun die Integrationstheorie zunächst für reelwertige Funktionen. Um auch singuläre Funktionen integrieren zu können, müssen wir die Funktionswerte  $\pm\infty$  explizit zulassen.

**11.25 Definition.** Es sei

$$\bar{\mathbb{R}} = [-\infty, \infty] := \mathbb{R} \cup \{\pm\infty\},$$

wobei  $\infty = +\infty \neq -\infty \notin \mathbb{R}$  zwei Symbole sind. Wir erweitern die Ordnung und die Rechenregeln von  $\mathbb{R}$  auf  $[-\infty, \infty]$  durch

$$\begin{aligned} -\infty < a < \infty & \quad \forall a \in \mathbb{R} \\ a + \infty = \infty + a & := \infty & \quad \forall -\infty < a \leq \infty \\ a - \infty = -\infty + a & := -\infty & \quad \forall -\infty \leq a < \infty \\ (\pm\infty)(\pm a) = (\pm a)(\pm\infty) & := +\infty & \quad \forall 0 < a \leq \infty \\ (\pm\infty)(\mp a) = (\mp a)(\pm\infty) & := -\infty & \quad \forall 0 < a \leq \infty \\ \pm\infty \cdot 0 = 0 \cdot (\pm\infty) & := 0 \end{aligned}$$

Nicht definiert sind die Ausdrücke  $\infty - \infty$  und  $-\infty + \infty$ .

**11.26 Definition. Numerische Funktionen und die Borel- $\sigma$ -Algebra auf  $\bar{\mathbb{R}}$**

Eine Funktion  $f : X \rightarrow \bar{\mathbb{R}}$  heißt **numerische Funktion**. Wir versehen  $\bar{\mathbb{R}}$  im Folgenden immer mit der Borel- $\sigma$ -Algebra

$$\bar{\mathcal{B}} := \{B \subset \bar{\mathbb{R}} \mid B \cap \mathbb{R} \in \mathcal{B}\}$$

und erweitern das Lebesgue-Borel-Maß gemäß

$$\lambda(B) := \lambda(B \cap \mathbb{R}) \quad \forall B \in \bar{\mathcal{B}}.$$

Wie in der Einleitung skizziert, beruht das Lebesgueintegral für reellwertige Funktionen auf der Idee, den Wertebereich in Intervalle zu zerlegen und die Urbildmengen von Intervallen zu "messen". Wir werden also von den zu integrierenden Funktionen verlangen müssen, dass Urbilder von Intervallen oder (wie wir sehen werden äquivalent) Urbilder von Borel-messbaren Mengen wieder messbar sind. Allgemein definiert man deshalb folgenden Begriff der messbaren Funktion:

**11.27 Definition. Messbare Funktionen**

Seien  $(X, \mathcal{A})$  und  $(Y, \mathcal{C})$  Messräume. Eine Abbildung

$$f : X \rightarrow Y \quad \text{heißt } \mathcal{A}\text{-}\mathcal{C}\text{-messbar,}$$

falls für alle  $C \in \mathcal{C}$  gilt, dass  $f^{-1}(C) \in \mathcal{A}$ , also falls  $f$ -Urbilder messbarer Mengen messbar sind.

Für numerische Funktionen ergibt sich sofort:

**11.28 Bemerkung. Messbare numerische Funktionen**

Sei  $(X, \mathcal{A})$  ein Messraum. Für eine Abbildung  $f : X \rightarrow \bar{\mathbb{R}}$  sind äquivalent

- (a)  $f$  ist  $\mathcal{A}$ - $\bar{\mathcal{B}}$ -messbar
- (b)  $\forall a \in \mathbb{R}$  ist  $\{f \geq a\} := \{x \in X \mid f(x) \geq a\} = f^{-1}([a, \infty]) \in \mathcal{A}$
- (c)  $\forall a \in \mathbb{R}$  ist  $\{f > a\} = f^{-1}((a, \infty]) \in \mathcal{A}$
- (d)  $\forall a \in \mathbb{R}$  ist  $\{f \leq a\} = f^{-1}([-\infty, a]) \in \mathcal{A}$
- (e)  $\forall a \in \mathbb{R}$  ist  $\{f < a\} = f^{-1}([-\infty, a)) \in \mathcal{A}$

*Beweis.* Die Behauptung folgt aus folgendem Lemma,

**11.29 Lemma.** Sei  $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{C})$  und  $\mathcal{E} \subset \mathcal{C}$  ein Erzeugendensystem von  $\mathcal{C}$ . Falls  $f^{-1}(E) \in \mathcal{A}$  für alle  $E \in \mathcal{E}$ , so ist  $f$  messbar.

*Beweis.* Sei  $\mathcal{C}_f := \{C \subset Y \mid f^{-1}(C) \in \mathcal{A}\}$  die größte  $\sigma$ -Algebra auf  $Y$  so, dass  $f$   $\mathcal{A}$ - $\mathcal{C}_f$ -messbar ist. Dann gilt nach Voraussetzung  $\mathcal{E} \subset \mathcal{C}_f$  und somit  $\mathcal{C} \subset \mathcal{C}_f$ . □

und der Tatsache, dass die Familien  $\{[a, \infty] \mid a \in \mathbb{R}\}$ ,  $\{(a, \infty] \mid a \in \mathbb{R}\}$  usw. jeweils die Borel- $\sigma$ -Algebra  $\bar{\mathcal{B}}$  erzeugen. □

**11.30 Beispiel.** Alle stetigen Funktionen  $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  sind  $\mathcal{B}^n$ - $\bar{\mathcal{B}}$ -messbar, da die offenen Mengen in  $\bar{\mathbb{R}}$  die  $\sigma$ -Algebra  $\bar{\mathcal{B}}$  erzeugen und deren Urbilder als offene Mengen  $\mathcal{B}^n$ -messbar sind. Und jede  $\mathcal{B}^n$ - $\bar{\mathcal{B}}$ -messbare Funktion ist auch  $\mathcal{L}^n$ - $\bar{\mathcal{B}}$ -messbar, da  $\mathcal{B}^n \subset \mathcal{L}^n$ .

**11.31 Proposition. Verknüpfungen messbarer numerischer Funktionen**

Sei  $(X, \mathcal{A})$  ein Messraum,  $f, g, f_j : X \rightarrow \bar{\mathbb{R}}$ ,  $j \in \mathbb{N}$ , seien  $\mathcal{A}$ -messbare Funktionen und  $\varphi : \bar{\mathbb{R}} \rightarrow \bar{\mathbb{R}}$  sei messbar. Dann sind die Funktionen

$$f + g, \quad f \cdot g, \quad \frac{f}{g}, \quad |f|, \quad \varphi \circ f, \quad \min(f, g), \quad \max(f, g),$$

$$\sup_{j \in \mathbb{N}} f_j, \quad \inf_{j \in \mathbb{N}} f_j, \quad \limsup_{j \rightarrow \infty} f_j, \quad \liminf_{j \rightarrow \infty} f_j$$

alle wieder  $\mathcal{A}$ -messbar (falls sie wohldefiniert sind).

*Beweis.* Zunächst bemerken wir, dass Kompositionen messbarer Abbildungen messbar sind. Also ist  $\varphi \circ f$  messbar.

Für  $f + g, f \cdot g, |f|, \min(f, g)$  und  $\max(f, g)$  beachte, dass

$$(f + g)^{-1}(\pm\infty) = f^{-1}(\pm\infty) \cup g^{-1}(\pm\infty) \in \mathcal{A}$$

bzw.

$$(f \cdot g)^{-1}(+\infty) = ((f^{-1}(\infty) \cap g^{-1}((0, \infty])) \cup \dots) \in \mathcal{A}$$

usw. Es genügt also,  $f, g : X \rightarrow \mathbb{R}$  zu betrachten.

Nun sind  $+: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \cdot : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, |\cdot| : \mathbb{R} \rightarrow [0, \infty), \min : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  und  $\max : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  alle stetig und somit  $\mathcal{B}^2$ - $\mathcal{B}$ -messbar.

Ebenso ist die Abbildung  $(f, g) : X \rightarrow \mathbb{R} \times \mathbb{R}, x \mapsto (f(x), g(x))$ ,  $\mathcal{A}$ - $\mathcal{B}^2$ -messbar, denn  $(f, g)^{-1}((a, b) \times (c, d)) = f^{-1}((a, b)) \cap g^{-1}((c, d)) \in \mathcal{A}$ .

Also sind  $f + g = +(f, g)$  usw. als Kompositionen messbarer Abbildungen alle wieder messbar.

Da

$$\left\{ x \in X \mid \sup_{j \in \mathbb{N}} f_j(x) \leq a \right\} = \bigcap_{j \in \mathbb{N}} \left\{ x \in X \mid f_j(x) \leq a \right\}$$

abzählbarer Schnitt messbarer Mengen ist, bzw.

$$\left\{ x \in X \mid \inf_{j \in \mathbb{N}} f_j(x) < a \right\} = \bigcup_{j \in \mathbb{N}} \left\{ x \in X \mid f_j(x) < a \right\},$$

abzählbare Vereinigung messbarer Mengen ist, sind  $\sup f_j$  und  $\inf f_j$   $\mathcal{A}$ -messbar. Ebenso sind  $\limsup f_j$  und  $\liminf f_j$   $\mathcal{A}$ -messbar, da

$$\limsup_{j \rightarrow \infty} f_j = \inf_{n \in \mathbb{N}} \sup_{j \geq n} f_j \quad \text{bzw.} \quad \liminf_{j \rightarrow \infty} f_j = \sup_{n \in \mathbb{N}} \inf_{j \geq n} f_j. \quad \square$$

### 11.32 Korollar. Punktweise Limites messbarer Funktionen sind messbar

Sei  $(f_n)$  eine Folge messbarer numerischer Funktionen die punktweise gegen eine Funktion  $f$  konvergiert. Dann ist  $f$  messbar.

*Beweis.* Das folgt wegen  $\lim_{n \rightarrow \infty} f_n = \limsup_{n \rightarrow \infty} f_n = \liminf_{n \rightarrow \infty} f_n$ . □

### 11.33 Definition. Einfache Funktionen

Eine Funktion  $g : X \rightarrow \bar{\mathbb{R}}$  heißt **einfach** (oder Treppenfunktion), wenn  $g(X)$  endlich ist,  $g$  also nur endlich viele verschiedene Werte annimmt.

Eine einfache Funktion kann in eindeutiger Weise in der **kanonischen Form**

$$g(x) = \sum_{j=1}^k \alpha_j \chi_{A_j}(x)$$

mit  $\alpha_j \in \bar{\mathbb{R}}, A_j \subset X, A_j \cap A_i = \emptyset$  und  $\alpha_j \neq \alpha_i$  für  $i \neq j$  geschrieben werden.

### 11.34 Definition. Das Integral für einfache Funktionen

Sei  $(X, \mathcal{A}, \mu)$  ein Maßraum und  $g : X \rightarrow [0, \infty]$  einfach und messbar (also  $A_j \in \mathcal{A}$  für alle  $j$ ), dann definieren wir das Integral von  $g$  bzgl.  $\mu$  durch

$$\int_X g \, d\mu := \sum_{j=1}^k \alpha_j \mu(A_j),$$

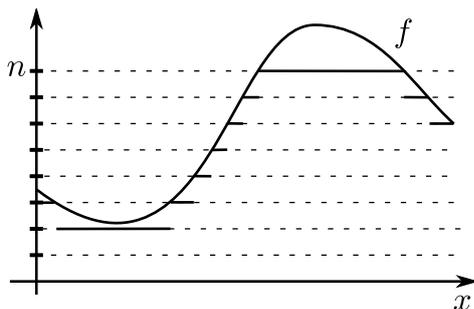
wobei  $g = \sum_{j=1}^k \alpha_j \chi_{A_j}$  in der kanonischen Form gegeben ist.

Aus dem nächsten Lemma werden wir folgern, dass man diese Integraldefinition in natürlicher Weise auf messbare Funktionen ausdehnen kann.

**11.35 Lemma.** Sei  $f : X \rightarrow [0, \infty]$  messbar. Dann gibt es eine monoton steigende Folge  $(g_n)$  (also  $g_n \leq g_{n+1}$  für alle  $n \in \mathbb{N}$ ) einfacher messbarer Funktionen, die punktweise gegen  $f$  konvergiert.

*Beweis.* Definiere  $g_n : X \rightarrow [0, \infty)$  durch

$$g_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{falls } \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \\ n & \text{falls } f(x) \geq n. \end{cases} \quad \text{für ein } k \in \mathbb{N} \text{ mit } k \leq n \cdot 2^n$$



Da  $f$  messbar ist, sind die Mengen  $\{\frac{k-1}{2^n} \leq f < \frac{k}{2^n}\}$  messbar und somit ist auch  $g_n$  messbar.

□

**11.36 Definition. Das Integral für positive messbare Funktionen**

Sei  $(X, \mathcal{A}, \mu)$  ein Maßraum und  $f : X \rightarrow [0, \infty]$  messbar, dann definieren wir das Integral von  $f$  bzgl.  $\mu$  durch

$$\int f \, d\mu = \sup \left\{ \int g \, d\mu \mid g : X \rightarrow [0, \infty] \text{ einfach und messbar mit } g \leq f \right\}.$$

**11.37 Bemerkung.** Das Integral über eine positive messbare Funktion ist also immer wohldefiniert, kann aber den Wert  $+\infty$  annehmen: Z.B. ist für  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{x} \chi_{[1, \infty]}$  und  $g_n(x) = \frac{1}{\lfloor x \rfloor} \chi_{(1, n+1]}(x)$

$$\int_{\mathbb{R}} f \, d\mu \geq \sup_{n \in \mathbb{N}} \int_{\mathbb{R}} g_n \, d\mu = \sup_{n \in \mathbb{N}} \sum_{j=1}^n \frac{1}{j+1} = \infty.$$

Um nun die üblichen Eigenschaften des Integrals (Linearität und Monotonie) zu zeigen, weisen wir diese zunächst für einfache Funktionen nach (was einfach ist, daher der Name) und zeigen dann, dass sie sich durch monotone Approximation gemäß Lemma 11.35 auf das Integral messbarer Funktionen übertragen lassen.

Im Folgenden bezeichne  $\mathcal{T}(X)$  die Menge der nicht negativen, einfachen und messbaren Funktionen auf  $(X, \mathcal{A})$ .

**11.38 Lemma. Eigenschaften des Integrals über Treppenfunktionen**

(i) Sei  $g = \sum_{i=1}^n \beta_i \chi_{B_i}$  einfach, messbar und nicht negativ und die Mengen  $B_i$  seien paarweise disjunkt, dann gilt

$$\int_X g \, d\mu = \sum_{i=1}^n \beta_i \mu(B_i).$$

(ii) Seien  $g, h \in \mathcal{T}(X)$  und  $\alpha \geq 0$ . Dann gilt

a)  $\alpha g \in \mathcal{T}(X)$  und  $\int \alpha g \, d\mu = \alpha \int g \, d\mu$

- b)  $g + h \in \mathcal{T}(X)$  und  $\int (g + h) d\mu = \int g d\mu + \int h d\mu$   
 c)  $g \leq h \Rightarrow \int g d\mu \leq \int h d\mu$ .

*Beweis.* (i) Folgt sofort aus der Additivität des Maßes  $\mu$ : Es sei  $g = \sum_{j=1}^k \alpha_j \chi_{A_j}$  die kanonische Darstellung von  $g$  und  $I_j := \{i \mid \beta_i = \alpha_j\}$ . Dann ist  $A_j = \bigcup_{i \in I_j} B_i$  und  $\mu(A_j) = \sum_{i \in I_j} \mu(B_i)$ , also

$$\sum_{i=1}^n \beta_i \mu(B_i) = \sum_{j=1}^k \sum_{i \in I_j} \alpha_j \mu(B_i) = \sum_{j=1}^k \alpha_j \mu(A_j) = \int g d\mu.$$

- (ii) a) Offensichtlich.  
 b) Es genügt  $g = \alpha \chi_A$  und  $h = \beta \chi_B$  zu betrachten. Dann ist

$$g + h = (\alpha + \beta) \chi_{A \cap B} + \alpha \chi_{A \setminus B} + \beta \chi_{B \setminus A} \in \mathcal{T}(X)$$

und mit Teil (i) und der Additivität des Maßes gilt

$$\int (g+h) d\mu \stackrel{(i)}{=} (\alpha+\beta)\mu(A \cap B) + \alpha\mu(A \setminus B) + \beta\mu(B \setminus A) = \alpha\mu(A) + \beta\mu(B) = \int g d\mu + \int h d\mu.$$

- c) Mit (b) ist  $h - g \in \mathcal{T}(X)$  und  $\int h d\mu = \int g d\mu + \underbrace{\int (h - g) d\mu}_{\geq 0} \geq \int g d\mu$ . □

**11.39 Bemerkung.** Mit Teil (ii) (b) von Lemma 11.38 folgt dann, dass für jede beliebige Darstellung einer Treppenfunktion  $g = \sum_{i=1}^n \beta_i \chi_{B_i}$  gilt

$$\int g d\mu = \sum_{i=1}^n \beta_i \mu(B_i).$$

**11.40 Lemma.** Sei  $f : X \rightarrow [0, \infty]$  messbar und  $(g_n)$  eine monoton wachsende Folge in  $\mathcal{T}(X)$  die punktweise gegen  $f$  konvergiert. Dann gilt

$$\int f d\mu = \lim_{n \rightarrow \infty} \int g_n d\mu.$$

*Beweis.* Da  $g_n \leq f$ , ist nach der Definition von  $\int f d\mu$  offensichtlich  $\int g_n d\mu \leq \int f d\mu$  für alle  $n \in \mathbb{N}$ , also auch  $\lim_{n \rightarrow \infty} \int g_n d\mu \leq \int f d\mu$ .

Für die umgekehrte Ungleichung zeigen wir:

$$(*) \quad \text{Sei } h \in \mathcal{T}(X) \text{ mit } h \leq f \Rightarrow \int h d\mu \leq \lim_{n \rightarrow \infty} \int g_n d\mu.$$

Denn dann gilt für jede Folge  $(h_m)$  in  $\mathcal{T}(X)$  mit  $h_m \leq f$  und  $\int h_m d\mu \nearrow \int f d\mu$  (solche Folgen gibt es aufgrund der Definition von  $\int f d\mu$  als Supremum), dass

$$\int f d\mu = \lim_{m \rightarrow \infty} \int h_m d\mu \leq \lim_{n \rightarrow \infty} \int g_n d\mu.$$

Sei für (\*) also  $h = \sum_{j=1}^k \alpha_j \chi_{A_j} \in \mathcal{T}(X)$  und  $\delta \in (0, 1)$ . Setze  $C_n := \{x \mid g_n(x) \geq \delta h(x)\}$ . Die Mengen  $C_n$  sind messbar und wegen  $g_n \geq \delta h \chi_{C_n}$  ist

$$\int g_n d\mu \geq \delta \int h \chi_{C_n} d\mu.$$

Wegen der Monotonie von  $g_n$  ist  $C_n \subset C_{n+1}$  und da  $\delta < 1$ , ist  $\bigcup_{n=1}^{\infty} C_n = X$  (kurz:  $C_n \nearrow X$ ), also insbesondere auch  $C_n \cap A_j \nearrow A_j$ . Also ergibt sich für jedes  $\delta < 1$ , dass

$$\begin{aligned} \int h \, d\mu &= \sum_{j=1}^k \alpha_j \mu(A_j) = \lim_{n \rightarrow \infty} \sum_{j=1}^k \alpha_j \mu(C_n \cap A_j) \\ &= \lim_{n \rightarrow \infty} \int h \chi_{C_n} \, d\mu \leq \frac{1}{\delta} \lim_{n \rightarrow \infty} \int g_n \, d\mu, \end{aligned}$$

was (\*) impliziert. □

**11.41 Folgerung.** Seien  $f, g : X \rightarrow [0, \infty]$  messbar und  $\alpha \geq 0$ . Dann gilt

- (i)  $\int \alpha f \, d\mu = \alpha \int f \, d\mu$
- (ii)  $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$
- (iii)  $f \leq g \Rightarrow \int f \, d\mu \leq \int g \, d\mu$ .

*Beweis.* Seien  $(f_n)$  und  $(g_n)$  Folgen in  $\mathcal{T}(X)$  mit  $f_n \nearrow f$  und  $g_n \nearrow g$ .

- (i)  $\alpha f_n \nearrow \alpha f \Rightarrow \int \alpha f \, d\mu = \lim_{n \rightarrow \infty} \int \alpha f_n \, d\mu = \lim_{n \rightarrow \infty} \alpha \int f_n \, d\mu = \alpha \int f \, d\mu$ .
- (ii)  $(f_n + g_n) \nearrow (f + g) \Rightarrow \dots$  wie in (i).
- (iii)  $f_n \leq g \stackrel{\text{nach Def.}}{\Rightarrow} \int f_n \, d\mu \leq \int g \, d\mu$  für alle  $n \in \mathbb{N} \Rightarrow$  Behauptung. □

Der folgende Satz von Beppo Levi ist einer der zentralen Konvergenzsätze für die Integrationstheorie.

**11.42 Satz. Satz von der monotonen Konvergenz (von Beppo Levi)**

Seien  $f_n : X \rightarrow [0, \infty]$  messbar und  $f_n \leq f_{n+1}$  für alle  $n \in \mathbb{N}$ . Setze  $f := \lim_{n \rightarrow \infty} f_n$  (punktweise), dann gilt

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \int f \, d\mu.$$

*Beweis.* Da  $f_n \leq f$  für alle  $n \in \mathbb{N}$ , gilt  $\int f_n \, d\mu \leq \int f \, d\mu$  und somit auch  $\lim_{n \rightarrow \infty} \int f_n \, d\mu \leq \int f \, d\mu$ . Wir konstruieren nun eine Folge  $(g_n)$  in  $\mathcal{T}(X)$  mit  $g_n \nearrow f$  und  $g_n \leq f_n$ . Denn dann folgt auch

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int g_n \, d\mu \leq \lim_{n \rightarrow \infty} \int f_n \, d\mu.$$

Dazu wähle zu jedem  $n \in \mathbb{N}$  einer Folge  $(h_{nj})_{j \in \mathbb{N}}$  in  $\mathcal{T}(X)$  mit  $(h_{nj})_{j \in \mathbb{N}} \nearrow f_n$  und setze

$$g_n := \max(h_{1n}, \dots, h_{nn}).$$

Dann ist

$$g_{n+1} = \max(h_{1,n+1}, \dots, h_{n,n+1}, h_{n+1,n+1}) \geq \max(h_{1n}, \dots, h_{nn}) = g_n,$$

da jeweils  $h_{j,n+1} \geq h_{jn}$ . Weiterhin gilt  $g_n \leq f_n$  da  $h_{jn} \leq f_j \leq f_n$  für  $j \leq n$ . Schließlich folgt  $\lim_{n \rightarrow \infty} g_n = f$  aus

$$f = \sup_j (f_j) = \sup_j \sup_{n \geq j} \underbrace{(h_{jn})}_{\leq g_n} \leq \sup_n g_n \leq \sup_n f_n = f. \quad \square$$

**11.43 Korollar. Lemma von Fatou**

Seien  $f_n : X \rightarrow [0, \infty]$  messbar. Dann gilt

$$\int \liminf_{n \rightarrow \infty} f_n \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu.$$

*Beweis.* Mit Proposition 11.31 sind die Funktionen  $g_k := \inf_{n \geq k} f_n$  messbar und  $g_k \leq f_n$  für  $n \geq k$ . Somit ist

$$\int g_k \, d\mu \leq \inf_{n \geq k} \int f_n \, d\mu.$$

Nun ist  $g_k \leq g_{k+1}$  und  $\lim_{k \rightarrow \infty} g_k = \liminf_{n \rightarrow \infty} f_n$  und der Satz der monotonen Konvergenz liefert

$$\int \liminf_{n \rightarrow \infty} f_n \, d\mu = \int \lim_{k \rightarrow \infty} g_k \, d\mu = \lim_{k \rightarrow \infty} \int g_k \, d\mu \leq \lim_{k \rightarrow \infty} \inf_{n \geq k} \int f_n \, d\mu = \liminf_{n \rightarrow \infty} \int f_n \, d\mu. \quad \square$$

Das Integral über Funktionen ohne festes Vorzeichen definiert man dadurch, dass man die Bereiche mit positivem und negativem Vorzeichen separat integriert. Um die Situation  $\infty - \infty = ?$  zu vermeiden, fordert man, dass beide Teile jeweils endlich sind.

#### 11.44 Definition. Integrierbare Funktionen und ihr Integral

Eine messbare Funktion  $f : X \rightarrow \overline{\mathbb{R}}$  heißt **integrierbar**, falls für

$$f_+ := \max(f, 0) \quad \text{und} \quad f_- := \max(-f, 0)$$

gilt

$$\int f_+ \, d\mu < \infty \quad \text{und} \quad \int f_- \, d\mu < \infty.$$

Man setzt dann

$$\int f \, d\mu := \int f_+ \, d\mu - \int f_- \, d\mu.$$

#### 11.45 Bemerkung. Kriterien für Integrierbarkeit

Für eine messbare Funktion  $f : X \rightarrow \overline{\mathbb{R}}$  sind äquivalent:

- (a)  $f$  ist integrierbar.
- (b)  $|f|$  ist integrierbar.
- (c) Es gibt eine integrierbare Funktion  $g : X \rightarrow [0, \infty]$  mit  $|f| \leq g$ .

*Beweis.* Übungsaufgabe □

#### 11.46 Satz. Eigenschaften des Integrals über integrierbare Funktionen

Seien  $f, g : X \rightarrow \overline{\mathbb{R}}$  integrierbar und  $\alpha \in \mathbb{R}$ . Dann gilt

- (a)  $\alpha f$  ist integrierbar und  $\int \alpha f \, d\mu = \alpha \int f \, d\mu$ .
- (b) Falls  $f + g$  definiert ist ( $\infty - \infty$  ist nicht definiert, vgl. dazu aber Folgerung 11.50), so ist  $f + g$  integrierbar und

$$\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

- (c)  $f \leq g \Rightarrow \int f \, d\mu \leq \int g \, d\mu$ .
- (d)  $|\int f \, d\mu| \leq \int |f| \, d\mu$ .

*Beweis.* (a) Für  $\alpha \geq 0$  ist

$$\begin{aligned} \int \alpha f \, d\mu &= \int (\alpha f)_+ \, d\mu - \int (\alpha f)_- \, d\mu = \int \alpha f_+ \, d\mu - \int \alpha f_- \, d\mu \\ &= \alpha \left( \int f_+ \, d\mu - \int f_- \, d\mu \right) = \alpha \int f \, d\mu \end{aligned}$$

und  $\alpha < 0$  geht analog.

- (b) Da  $|f + g| \leq |f| + |g|$  ist nach Bemerkung 11.45 (c)  $f + g$  integrierbar und mit  $h := f + g$  gilt

$$h_+ - h_- = h = f + g = f_+ + g_+ - (f_- + g_-),$$

also

$$h_+ + f_- + g_- = h_- + f_+ + g_+.$$

Folgerung 11.41 liefert

$$\int h_+ \, d\mu + \int f_- \, d\mu + \int g_- \, d\mu = \int h_- \, d\mu + \int f_+ \, d\mu + \int g_+ \, d\mu,$$

also

$$\int h \, d\mu = \int h_+ \, d\mu - \int h_- \, d\mu = \int f_+ \, d\mu - \int f_- \, d\mu + \int g_+ \, d\mu - \int g_- \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

- (c) Da  $f \leq g$  impliziert, dass  $f_+ \leq g_+$  und  $g_- \leq f_-$ , liefert Folgerung 11.41, dass

$$\int f \, d\mu = \int f_+ \, d\mu - \int f_- \, d\mu \leq \int g_+ \, d\mu - \int g_- \, d\mu = \int g \, d\mu.$$

- (d) Da  $f \leq |f|$  und  $-f \leq |f|$ , folgt  $\int f \, d\mu \leq \int |f| \, d\mu$  und  $-\int f \, d\mu \leq \int |f| \, d\mu$ . □

### 11.47 Definition. Fast überall oder fast sicher

Wir sagen, eine Eigenschaft von Punkten  $x$  gilt fast überall (oder fast sicher) bezüglich eines Maßes  $\mu$  auf  $X$ , falls diese Eigenschaft für  $x \in A \subset X$  gilt und

$$\mu(X \setminus A) = 0.$$

Eine Eigenschaft gilt also fast überall, falls sie nur auf einer Nullmenge nicht gilt.

**11.48 Beispiel.** (a) Da  $\lambda(\mathbb{Q}) = 0$ , ist eine reelle Zahl fast sicher irrational, also  $\chi_{\mathbb{Q}}(x) = 0$   $\lambda$ -fast überall auf  $\mathbb{R}$ .

- (b) Ist  $f : X \rightarrow \overline{\mathbb{R}}$  integrierbar, so ist  $f(x) \in \mathbb{R}$  fast überall. Denn sei  $N = f^{-1}(\{-\infty, \infty\})$ , dann ist  $\alpha \chi_N \leq |f|$  für alle  $\alpha \in \mathbb{R}$ , also  $\alpha \mu(N) = \alpha \int \chi_N \, d\mu \leq \int |f| \, d\mu < \infty$  für alle  $\alpha \in \mathbb{R}$ . Dann muss aber  $\mu(N) = 0$  sein.

Eine integrierbare Funktion kann also nur auf einer Nullmenge die Werte  $\pm\infty$  annehmen.

**11.49 Bemerkung.** Sei  $f : X \rightarrow [0, \infty]$  messbar. Dann gilt:

$$\int_X f \, d\mu = 0 \quad \Leftrightarrow \quad f = 0 \quad \text{f.ü.}$$

*Beweis.* Übungsaufgabe. □

**11.50 Folgerung.** (a) Man kann also eine integrierbare Funktion  $f : X \rightarrow \overline{\mathbb{R}}$  auf jeder Nullmenge  $N \subset X$  beliebig abändern ohne ihr Integral dabei zu ändern.

- (b) Insbesondere darf man das für  $N = f^{-1}(\{-\infty, \infty\})$ . D.h. für integrierbare Funktionen kann man zumindest das Integral betreffend annehmen, dass  $f$  reellwertig ist.
- (c) Somit ist  $f + g$  für integrierbare Funktionen immer f.ü. definiert und die Aussage von Folgerung 11.46 (b) gilt in diesem Sinne immer.

**11.51 Definition. Das Integral komplexwertiger Funktionen**

- (a) Eine Funktion  $f : X \rightarrow \mathbb{C}$  auf einem Messraum  $(X, \mathcal{A})$  heißt **messbar**, falls  $\operatorname{Re}(f)$  und  $\operatorname{Im}(f) : X \rightarrow \mathbb{R}$  messbar sind.
- (b) Eine Funktion  $f : X \rightarrow \mathbb{C}$  heißt **integrierbar**, falls  $|f| : X \rightarrow [0, \infty)$  integrierbar ist und ihr Lebesgue-Integral definieren wir durch

$$\int f \, d\mu := \int \operatorname{Re}(f) \, d\mu + i \int \operatorname{Im}(f) \, d\mu.$$

**11.52 Bemerkung.** Satz 11.46 und Folgerungen 11.50 gelten analog für  $\mathbb{C}$ -wertige Funktionen. Es ist deshalb nicht nötig, zu  $\mathbb{C}$  wieder einen unendlich fernen Punkt hinzuzunehmen. Singuläre integrierbare Funktionen können eben nur auf einer Nullmenge singulär sein und dann spielt der Wert an der Singularität keine Rolle, solange wir die Funktion nur integrieren wollen.

Schließlich stellen wir noch fest, dass das eindimensionale Lebesgueintegral für Regelfunktionen mit dem im ersten Semester eingeführten Regelintegral übereinstimmt.

**11.53 Proposition. Vergleich zum Regelintegral**

Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Regelfunktion, so ist  $f$   $\mathcal{B}$ - $\overline{\mathcal{B}}$ -messbar und integrierbar und die Integrale stimmen überein:

$$\int_a^b f(x) \, dx = \int_{[a,b]} f \, d\lambda := \int f \chi_{[a,b]} \, d\lambda.$$

*Beweis.* Übung. □