

---

# Grundlagen der Analysis

---

**Frank Klinker**

---

Ausarbeitung einer dreisemestrigen Vorlesung Analysis I-III  
Technische Universität Dortmund, Fakultät für Mathematik  
*Email:* [mail@frank-klinker.de](mailto:mail@frank-klinker.de)



---

## Inhaltsverzeichnis

---

|       |  |    |
|-------|--|----|
| 1     | Analysis I   | 1  |
| 1.1   | Mengensprache, Induktion, Relationen, Zahlbereiche   | 1  |
| 1.1.1 | Mengen und Mengenverknüpfungen   | 1  |
| 1.1.2 | Die natürlichen Zahlen und das Induktionsprinzip   | 5  |
| 1.1.3 | Relationen, ganze und rationale Zahlen   | 10 |
| 1.2   | Zahlenfolgen und reelle Zahlen   | 18 |
| 1.2.1 | Zahlenfolgen in $\mathbb{Q}$   | 18 |
| 1.2.2 | Die reellen Zahlen   | 26 |
| 1.2.3 | Cauchyfolgen in $\mathbb{R}$ : die Vollständigkeit von $\mathbb{R}$ und das Supremumsprinzip | 29 |
| 1.2.4 | Häufungspunkte und weitere Eigenschaften von $\mathbb{R}$                                    | 33 |
| 1.2.5 | Rechenregeln, Intervalle   | 36 |
| 1.3   | Reelle Funktionen  | 38 |
| 1.3.1 | Grundlagen und erste Beispiele   | 38 |
| 1.3.2 | Wurzelfunktionen   | 43 |
| 1.3.3 | Polynome und rationale Funktionen  | 46 |

|       |   |     |
|-------|---|-----|
| 1.3.4 | Die allgemeine Exponentialfunktion und Logarithmusfunktion.....   | 51  |
| 1.4   | Stetige Funktionen und ihre Eigenschaften .....   | 55  |
| 1.4.1 | Stetigkeit.....   | 55  |
| 1.4.2 | Eigenschaften stetiger Funktionen auf Intervallen.....  | 58  |
| 1.4.3 | Grenzwerte von Funktionen .....   | 60  |
| 1.4.4 | Weitere Stetigkeitsbegriffe .....   | 64  |
| 1.5   | Differenzierbarkeit .....   | 67  |
| 1.5.1 | Ableitung und Differenzierbarkeit .....   | 67  |
| 1.5.2 | Eigenschaften differenzierbarer Funktionen .....  | 73  |
| 1.5.3 | Höhere Ableitungen und der Satz von Taylor .....  | 76  |
| 1.6   | Zahlenreihen und Potenzreihen .....   | 86  |
| 1.6.1 | Zahlenreihen.....   | 86  |
| 1.6.2 | Konvergenzkriterien für Zahlenreihen.....   | 91  |
| 1.6.3 | Potenzreihen.....   | 92  |
| 1.7   | Die Winkelfunktionen und die Arkusfunktionen .....  | 97  |
| 1.7.1 | Die Winkelfunktionen .....  | 97  |
| 1.7.2 | Die Arkusfunktionen .....   | 104 |
| 1.8   | Anhang Analysis I: Grundbegriffe der formalen Logik .....   | 106 |
| 1.9   | Anhang Analysis I: Zahlbereichserweiterungen $\mathbb{N} \rightarrow \mathbb{Z} \rightarrow \mathbb{Q}$ ..... | 111 |
| 1.9.1 | Die ganzen Zahlen .....   | 111 |
| 1.9.2 | Die rationalen Zahlen .....   | 115 |
| 2     | Analysis II.....  | 119 |
| 2.1   | Funktionenfolgen .....  | 119 |
| 2.1.1 | Konvergenzbegriffe .....  | 119 |
| 2.1.2 | Eigenschaften konvergenter Funktionenfolgen .....   | 122 |
| 2.1.3 | Funktionenreihen .....  | 124 |
| 2.2   | Integration .....   | 127 |
| 2.2.1 | Stammfunktion und Integral .....  | 127 |
| 2.2.2 | Integrationsregeln .....  | 129 |
| 2.2.3 | Das Riemann-Integral .....  | 135 |

---

|       |  |     |
|-------|--|-----|
| 2.2.4 | Geometrische Anwendungen der Integration .....               | 142 |
| 2.3   | Abbildungen, Stetigkeit und Topologie .....                  | 149 |
| 2.3.1 | Grundbegriffe des Zahlenraums .....                          | 149 |
| 2.3.2 | Abbildungen in mehreren Variablen und Stetigkeit .....       | 150 |
| 2.3.3 | Folgen und Grenzwerte im Zahlenraum .....                    | 157 |
| 2.3.4 | Topologische Grundbegriffe des Zahlenraums .....             | 160 |
| 2.3.5 | Exkurs: Topologie .....                                      | 165 |
| 2.4   | Topologie und Eigenschaften stetiger Abbildungen .....       | 167 |
| 2.4.1 | Zusammenhang .....   | 167 |
| 2.4.2 | Kompaktheit und der Satz von Heine-Borel .....               | 170 |
| 2.4.3 | Der Banachsche Fixpunktsatz .....                            | 174 |
| 2.5   | Differenzierbare Abbildungen des Zahlenraums .....           | 176 |
| 2.5.1 | Richtungsableitung und partielle Differentiation .....       | 176 |
| 2.5.2 | Das totale Differential .....                                | 181 |
| 2.5.3 | Der Satz über implizite Funktionen und der Umkehrsatz .....  | 189 |
| 2.6   | Höhere Ableitungen und Extremwerte .....                     | 198 |
| 2.6.1 | Höhere Ableitungen und der Satz von Taylor .....             | 198 |
| 2.6.2 | Extremwerte ohne und mit Nebenbedingungen .....              | 202 |
| 2.7   | Grundlagen gewöhnlicher Differentialgleichungen .....        | 214 |
| 2.7.1 | Problemstellung und elementare Beispiele .....               | 214 |
| 2.7.2 | Existenz und Eindeutigkeit von Lösungen .....                | 226 |
| 2.8   | Anhang Analysis II: Werkzeuge aus der linearen Algebra ..... | 230 |
| 2.8.1 | Normen und Matrixnormen .....                                | 230 |
| 2.8.2 | Symmetrische Multilinearformen .....                         | 235 |
| 2.8.3 | Matrixexponential .....                                      | 238 |
| 2.8.4 | Einige Eigenschaften der Determinante .....                  | 240 |
| 3     | Analysis III .....   | 242 |
| 3.1   | Lineare DGLn und lineare DGL-Systeme .....                   | 242 |
| 3.1.1 | Die lineare DGL $n$ -ter Ordnung .....                       | 242 |
| 3.1.2 | Die Schwingungsgleichung .....                               | 246 |
| 3.1.3 | Lineare DGL-Systeme .....                                    | 250 |

|       |  |     |
|-------|--|-----|
| 3.2   | Das Jordan-Maß und messbare Mengen . . . . .                                     | 262 |
| 3.2.1 | Quaderapproximation . . . . .  | 263 |
| 3.2.2 | Messbare Teilmengen des $\mathbb{R}^n$ und das Jordan-Maß . . . . .              | 266 |
| 3.2.3 | Das Verhalten des Jordan-Maßes unter Abbildungen . . . . .                       | 268 |
| 3.3   | Integration im $\mathbb{R}^n$ . . . . .  | 269 |
| 3.3.1 | Das Darboux-Integral in $\mathbb{R}$ . . . . .                                   | 270 |
| 3.3.2 | Integration im $\mathbb{R}^n$ . . . . .  | 272 |
| 3.3.3 | Die Transformationsformel . . . . .  | 279 |
| 3.4   | Reguläre Teilmengen des $\mathbb{R}^n$ . . . . .                                 | 284 |
| 3.4.1 | Vorbereitungen . . . . .   | 284 |
| 3.4.2 | Untermannigfaltigkeiten des $\mathbb{R}^n$ . . . . .                             | 287 |
| 3.4.3 | Kurven und Flächen . . . . .   | 292 |
| 3.4.4 | Der Tangentialraum . . . . .   | 295 |
| 3.5   | Integralsätze und die Existenz von Potentialen . . . . .                         | 298 |
| 3.5.1 | Integration über Mannigfaltigkeiten im $\mathbb{R}^n$ . . . . .                  | 298 |
| 3.5.2 | Vektorfelder und Potentiale . . . . .  | 303 |
| 3.5.3 | Die geometrische Interpretation der Divergenz und der<br>Satz von Gauß . . . . . | 310 |
| 3.5.4 | Der Satz von Stokes und die Existenz von Potentialen . . . . .                   | 314 |
| 3.6   | Grundzüge der Funktionentheorie . . . . .  | 319 |
| 3.6.1 | Komplex versus reell . . . . .   | 319 |
| 3.6.2 | Komplexe Differentiation . . . . .   | 322 |
| 3.6.3 | Komplexe Stammfunktion und komplexes Wegintegral . . . . .                       | 325 |
| 3.6.4 | Der Cauchy-Integralsatz und die Cauchy-Integralformel . . . . .                  | 328 |
| 3.7   | Anhang Analysis III: Komplexe Zahlen . . . . .                                   | 331 |
| 3.7.1 | Beschreibung der komplexen Zahlen . . . . .                                      | 331 |
| 3.7.2 | Polarkoordinaten, Wurzeln und komplexe Exponential-<br>funktion . . . . .        | 335 |
|       | Abbildungsverzeichnis . . . . .  | 342 |
|       | Literaturverzeichnis . . . . .   | 343 |

# KAPITEL 1

---

## Analysis I

---

### 1.1 Mengensprache, Induktion, Relationen, Zahlbereiche

#### 1.1.1 Mengen und Mengenverknüpfungen

**Definition 1.1.1** (Menge). Eine Menge  $M$  ist eine Zusammenfassung von Objekten zu einem Ganzen. Diese Objekte heißen dann *Elemente der Menge*. Ist  $x$  ein Objekt der Menge  $M$ , dann schreiben wir  $x \in M$ . Ist  $x$  kein Objekt der Menge  $M$ , dann schreiben wir  $x \notin M$ .

**Bemerkung 1.1.2.** Mengen lassen sich beschreiben durch

- Aufzählen der Elemente mit Mengenklammern  $\{\dots\}$ , oder
- Angabe einer Eigenschaft  $E$ , die die Elemente beschreibt

$\{x \mid x \text{ hat die Eigenschaft } E\}$ .

**Beispiel 1.1.3.** •  $M = \{\text{⚽}, \text{⚾}, \text{🚲}, \text{⚔}\}$

- $M = \{\text{🚲} \mid \text{🚲 ist grün}\}$

- $M = \{\text{Student} \mid \text{Student sitzt in diesem Hörsaal}\}$
- $M = \emptyset = \{\} \longleftrightarrow \text{Leere Menge}$

Man kann mit Mengen operieren

**Definition 1.1.4** (Mengenverknüpfungen). Es seien  $M, N, M_1, M_2, \dots, M_k$  Mengen.

1. Die *Vereinigung* von  $M$  und  $N$  ist die Menge

$$M \cup N = \{x \mid x \in M \text{ oder } x \in N\}.$$

Es ist also  $x \in M \cup N \iff x \in M \vee x \in N$ .

2. Der *Schnitt* von  $M$  und  $N$  ist die Menge

$$M \cap N = \{x \mid x \in M \text{ und } x \in N\}.$$

Es gilt also  $x \in M \cap N \iff x \in M \wedge x \in N$ .

3.  $M$  heißt *Teilmenge* von  $N$ , wenn jedes Element aus  $M$  auch ein Element in  $N$  ist. Wir schreiben dann  $M \subset N$  oder auch  $N \supset M$ . Es ist

$$M \subset N \iff \forall x : (x \in M \implies x \in N) \iff \forall x \in M : x \in N.$$

4. Die *Differenz* von  $M$  und  $N$  ist die Menge

$$M \setminus N = \{x \mid x \in M \text{ und } x \notin N\}.$$

Damit ist  $x \in M \setminus N \iff x \in M \wedge \neg(x \in N)$

5. Ist  $M \subset N$ , so ist das *Komplement von  $M$  (bezüglich  $N$ )* die Menge

$$M^c = \{x \mid x \in N \text{ und } x \notin M\},$$

also  $x \in M^c \iff x \in N \wedge \neg(x \in M)$ .

6. Die *symmetrische Differenz* von  $M$  und  $N$  ist die Menge

$$M \diamond N = (M \setminus N) \cup (N \setminus M).$$

7. Das *Kreuzprodukt* von  $M$  und  $N$  ist die Menge

$$M \times N = \{(m, n) \mid m \in M \text{ und } n \in N\}$$

und die Elemente von  $M \times N$  nennt man auch *Paare*.

Analog definieren wir das Kreuzprodukt mehrerer Mengen als

$$M_1 \times M_2 \times \dots \times M_k = \{(m_1, m_2, \dots, m_k) \mid \forall i : m_i \in M_i\}$$

und ihre Elemente heißen *k-Tupel*.

**Bemerkung 1.1.5.** • Ist in 4.  $M \cap N = \emptyset$ , dann ist  $N \setminus M = N$  und  $M \setminus N = M$ .

- Ist in 4.  $M \subset N$ , dann ist  $N \setminus M = M^c$  und  $M \setminus N = \emptyset$ .
- Es ist<sup>(i)</sup>

$$\begin{aligned} M = N &\iff M \subset N \wedge N \subset M \\ &\iff \forall x : (x \in M \implies x \in N) \wedge \forall x : (x \in N \implies x \in M) \\ &\iff \forall x : ((x \in M \implies x \in N) \wedge (x \in N \implies x \in M)) \\ &\iff \forall x : (x \in M \iff x \in N). \end{aligned}$$

- Es ist  $M \times N \times P \neq (M \times N) \times P \neq M \times (N \times P)$ .

**Satz 1.1.6** (Rechenregeln für Mengenverknüpfungen). *Es seien  $M, N$  und  $P$  Mengen. Dann gilt*

1.  $M \cup N = N \cup M$   
 $M \cap N = N \cap M$
2.  $(M \cup N) \cup P = M \cup (N \cup P)$   
 $(M \cap N) \cap P = M \cap (N \cap P)$
3.  $(M \cup N) \cap P = (M \cap P) \cup (N \cap P)$   
 $(M \cap N) \cup P = (M \cup P) \cap (N \cup P)$
4.  $(M^c)^c = M$

---

<sup>(i)</sup>Die letzte Äquivalenz ist genau der Ansatz für den Beweis einer Mengengleichheit.

5.  $(M \cap N)^c = M^c \cup N^c$   
 $(M \cup N)^c = M^c \cap N^c$  *De Morgansche Regeln*
6.  $M \cup \emptyset = M$  und  $M \cap \emptyset = \emptyset$
7.  $M \diamond N = (M \cup N) \setminus (M \cap N)$

Die Aussagen aus Satz 1.1.6 beweist man, indem man zu der logischen Formulierung übergeht. Wir erläutern das am Beispiel  $M \cup N = N \cup M$ . Der entsprechende logische Ausdruck, der zu zeigen ist, ist  $x \in M \cup N \iff x \in N \cup M$ . Mit der Definition der Vereinigungsmenge ist das<sup>(i)</sup>

$$(x \in M \vee x \in N) \iff (x \in N \vee x \in M)$$

Dabei kann man im Prinzip auf zwei Arten vorgehen

- Wegen Bemerkung 1.8.6 sind die folgenden beiden Beweise gleichwertig:
  - Zeige, dass die gesamte Aussage eine Tautologie ist. Hierbei kann man eine WWT benutzen oder durch Umformung mit den Rechenregeln aus dem Anhang das Ergebnis bekommen.
  - Zeige, dass die WWT beider Seiten der Äquivalenz, das gleiche Ergebnis liefern.
- Zeige, durch Umformung und mit Hilfe bekannter Äquivalenzen, dass aus der linken Seite der Äquivalenz die rechte Seite folgt, und umgekehrt.

**Bemerkung 1.1.7.** Graphisch kann man sich die Verknüpfungen sehr gut mit *Venn-Diagrammen* veranschaulichen. Mit ihnen lassen sich Mengengleichheiten sehr gut auf Plausibilität überprüfen — Achtung: Ihre Verwendung ersetzt keinen formalen Beweis.

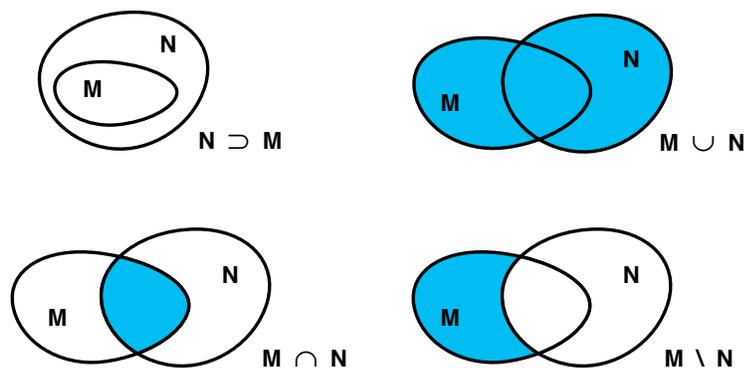
**Bezeichnung 1.1.8.** Ist  $I$  eine beliebige Indexmenge und  $\mathcal{M} = \{M_\alpha \mid \alpha \in I\}$  eine Menge von Mengen<sup>(ii)</sup>. Dann schreiben wir

$$\begin{aligned} \bigcap \mathcal{M} &:= \{x \mid \forall \alpha \in I : x \in M_\alpha\}, \\ \bigcup \mathcal{M} &:= \{x \mid \exists \alpha \in I : x \in M_\alpha\}. \end{aligned}$$

<sup>(i)</sup>Die hier zu zeigende Aussage ist gerade die dritte Äquivalenz aus Beispiel A.8.

<sup>(ii)</sup>Man sagt auch  $\mathcal{M}$  ist eine Familie von Mengen.

Abbildung 1.1.1: Venn-Diagramme



**Beispiel 1.1.9.** • Es sei die Indexmenge  $I = \{a, b, c, d, e\}$  gegeben und die Mengen  $M_a = \{1, 2, 3, 4, 5\}$ ,  $M_b = \{4, 5, 8, 9\}$ ,  $M_c = \{1, 3, 4, 5, 7, 9\}$ ,  $M_d = \{2, 4, 5, 6, 8, 10\}$ ,  $M_e = \{1, 4, 5\}$ . Dann gilt für die Familie  $\mathcal{M} = \{M_\alpha \mid \alpha \in I\}$

$$\cap \mathcal{M} = \{4, 5\}, \quad \cup \mathcal{M} = \{1, 2, \dots, 10\}.$$

- Es sei  $M$  eine Menge und  $I = M$ . Definiere nun  $M_x = \{x\} \subset M$  und  $\mathcal{M} = \{M_x \mid x \in M\}$ . Dann ist

$$\cap \mathcal{M} = \emptyset, \quad \cup \mathcal{M} = M.$$

- Es sei  $\mathcal{M} = \{M_i \mid i \in \mathbb{N}\}$  mit  $M_i = \{k \in \mathbb{N} \mid k \leq i\}$ . Dann ist

$$\cap \mathcal{M} = \{0\}, \quad \cup \mathcal{M} = \mathbb{N}.$$

**Definition 1.1.10.** Es sei  $M$  eine Menge. Die *Potenzmenge*  $\mathcal{P}(M)$  von  $M$  ist die Menge der Teilmengen von  $M$ , also

$$\mathcal{P}(M) = \{N \mid N \subset M\}.$$

## 1.1.2 Die natürlichen Zahlen und das Induktionsprinzip

### 1.1.2.1 $\mathbb{N}$ und seine Verknüpfungen

Wir möchten mit den Elementen einer Menge "rechnen". Wir werden die folgende Definition später noch etwas präzisieren, siehe Definition [1.1.34](#).

**Definition 1.1.11.** 1. Eine *Verknüpfung*  $\circ$  auf einer Menge  $M$  ordnet jedem Paar  $(m_1, m_2) \in M \times M$  ein Element in  $M$  zu, dieses nennen wir  $m_1 \circ m_2$ .

(a)  $\circ$  heißt *kommutativ*, wenn

$$\forall m_1, m_2 \in M : m_1 \circ m_2 = m_2 \circ m_1 .$$

(b)  $\circ$  heißt *assoziativ*, wenn

$$\forall m_1, m_2, m_3 \in M : (m_1 \circ m_2) \circ m_3 = m_1 \circ (m_2 \circ m_3) .$$

(c)  $e \in M$  heißt *neutrales Element* bezüglich  $\circ$ , wenn

$$\forall m \in M : m \circ e = e \circ m = m .$$

(d)  $\tilde{m} \in M$  heißt *inverses Element* zu  $m \in M$  bezüglich  $\circ$ , wenn

$$\tilde{m} \circ m = m \circ \tilde{m} = e .$$

2. Das Paar  $(M, \circ)$  heißt *Gruppe*, wenn  $\circ$  assoziativ ist, ein neutrales Element besitzt und zu jedem  $m \in M$  das inverse Element  $\tilde{m} \in M$  existiert.

**Beispiel 1.1.12.** Für eine Menge  $M$  sind  $\cap, \cup, \setminus$  und  $\diamond$  Verknüpfungen auf der Potenzmenge  $\mathcal{P}(M)$ .

Die Menge der natürlichen Zahlen mit ihren Rechenregeln setzen wir hier als bekannt voraus, werden aber im nächsten Abschnitt noch einige Details nachliefern. Die folgende Definition ist insofern eher eine Konvention.

**Definition 1.1.13.** (Natürliche Zahlen) Die Menge der natürlichen Zahlen ist die Menge

$$\mathbb{N} = \{0, 1, 2, 3, 4, \dots\} .$$

Wir schreiben

$$\mathbb{N}^* = \mathbb{N} \setminus \{0\} \quad \text{und} \quad \mathbb{N}^{\geq k} = \mathbb{N} \setminus \{0, 1, 2, \dots, k-1\} .$$

**Beispiel 1.1.14.** • Die Addition  $+$  auf  $\mathbb{N}$  ist eine Verknüpfung auf den natürlichen Zahlen mit den Eigenschaften (a), (b) und (c). Das neutrale Element ist 0 und dies ist auch das einzige Element, das ein Inverses besitzt.

- Die Multiplikation  $\cdot$  auf  $\mathbb{N}$  oder auf  $\mathbb{N}^*$  ist eine Verknüpfung mit den Eigenschaften (a), (b) und (c). Das neutrale Element ist 1, und dies ist auch das einzige Element, das ein Inverses besitzt.
- Auf der Menge  $\mathbb{N}$  gilt zusätzlich noch das *Distributivgesetz*, das die beiden Verknüpfungen verbindet

$$\forall a, b, c \in \mathbb{N} : a \cdot (b + c) = a \cdot b + a \cdot c.$$

### 1.1.2.2 Axiomatik von $\mathbb{N}$ und vollständige Induktion

In diesem Abschnitt soll einen kurzen Einblick in die Axiomatik der Natürlichen Zahlen gewähren. Dies geschieht ohne Anspruch an Vollständigkeit sondern vor allem zur Begründung des Induktionsprinzips.

Eine schöne Einführung der verschiedenen Zahlbereiche  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$  und  $\mathbb{C}$  findet man etwa in dem Buch [La]. Auch wenn das Vorgehen dort von unserem hier und in Abschnitt 1.9 vorgestellten abweicht, werden einige Dinge auch ähnlich behandelt.

**Definition 1.1.15** (Axiomatik der natürlichen Zahlen).

1. Die Peano-Axiome zur Einführung der natürlichen Zahlen sind

$$(PA0) \quad 0 \in \mathbb{N}$$

Es gibt eine Zuordnung  $\nu$ , die für jede natürlichen Zahl  $n$  eine neue natürliche Zahl  $\nu(n)$  liefert. Diese Zahl nennen wir den *Nachfolger* von  $n$ . Die *Nachfolgezurordnung*  $\nu$  erfülle

$$(PA1) \quad \forall k \in \mathbb{N} : \nu(k) \neq 0$$

$$(PA2) \quad \forall k, \ell \in \mathbb{N} : (\nu(k) = \nu(\ell) \implies k = \ell)$$

$$(PA3) \quad \text{Für jede beliebige Menge } M \text{ gilt}$$

$$\left( 0 \in M \wedge \forall k \in \mathbb{N} : (k \in M \implies \nu(k) \in M) \right) \implies \mathbb{N} \subset M.$$

Zusatz: Setzt man in (PA3)  $M \subset \mathbb{N}$  voraus, so folgt schon  $M = \mathbb{N}$ .

2. Auf  $\mathbb{N}$  sind durch die folgenden Festlegungen zwei Verknüpfungen  $+$  und  $\cdot$  erklärt:

$$(VA1) \quad k + 0 := k$$

$$(VA2) \quad k + \nu(\ell) := \nu(k + \ell)$$

$$(VA3) \quad k \cdot 0 = 0$$

$$(VA4) \quad k \cdot \nu(\ell) := (k \cdot \ell) + k$$

Setzt man  $\nu(0) := 1$ , so folgt  $\nu(k) = k + 1$ .

Mit Hilfe der obigen Axiome lassen sich nun die bekannten Rechenregeln für natürliche Zahlen und weitere nützliche Eigenschaften beweisen.

**Bemerkung 1.1.16.** 1.  $+$ ,  $\cdot$  sind kommutativ, assoziativ und distributiv.

2. Für alle  $k, \ell, n \in \mathbb{N}$  gilt  $n + k = n + \ell \implies k = \ell$ .
3. Es ist  $\mathbb{N} = \{0\} \cup \nu(\mathbb{N})$ . Damit gilt: für alle  $n \neq 0$  existiert ein  $n' \in \mathbb{N}$  mit  $\nu(n') = n$ .
4. Für alle  $n, m \in \mathbb{N}$  gilt  $n + m = 0 \implies (n = 0 \wedge m = 0)$ .
5. Für alle  $m, n \in \mathbb{N}$  gilt  $n \cdot m = 0 \implies (m = 0 \vee n = 0)$ .
6. Für alle  $k, \ell, n \in \mathbb{N}$  mit  $n \neq 0$  gilt  $k \cdot n = \ell \cdot n \implies k = \ell$ .

Das Axiom (PA3) liefert direkt die Lösung des folgenden Problems:

Es sei  $A$  eine Aussageform über der Grundmenge  $\mathbb{N}$ , siehe Definition 1.8.15. Man kann  $\{A(n)\}_{n \in \mathbb{N}}$  als eine Familie von Aussagen interpretieren. Gesucht ist ein effizientes Beweisverfahren für die Aussage

$$\forall n \in \mathbb{N} : A(n) \text{ ist wahr}$$

**Satz 1.1.17** (Induktionsprinzip). Sei  $\{A(n)\}_{n \in \mathbb{N}^{\geq n_0}}$  eine Familie von Aussagen, die für alle  $n \in \mathbb{N}^{\geq n_0}$  definiert ist. Weiter gelte

(IA)  $A(n_0)$  ist wahr, und

(IS) Für alle  $\ell \in \mathbb{N}^{\geq n_0}$  gilt: Ist  $A(\ell)$  wahr, so ist auch  $A(\ell + 1)$  wahr.

Dann ist  $A(n)$  wahr für alle  $n \in \mathbb{N}^{\geq n_0}$ .

Etwas formaler formuliert lautet das Induktionsprinzip wie folgt:

$$\left( A(n_0) \wedge \left( \forall \ell \in \mathbb{N}^{\geq n_0} : A(\ell) \implies A(\ell + 1) \right) \right) \implies \forall n \in \mathbb{N}^{\geq n_0} : A(n)$$

**Beweisskizze.** Es sei  $\{B(n)\}_{n \in \mathbb{N}}$  die Familie von Aussagen mit  $B(\ell) : \iff A(\ell + n_0)$ . In Termen von  $B$  lauten (IA) und (IS)

(IA)  $B(0)$  ist wahr, und

(IS) Für alle  $\ell \in \mathbb{N}$  gilt: Ist  $B(\ell)$  wahr, so ist auch  $B(\ell + 1)$  wahr.

Sei nun  $M := \{m \in \mathbb{N} \mid B(m) \text{ ist wahr}\} \subset \mathbb{N}$ . Dann sind (IA) und (IS) gerade die Voraussetzungen aus (PA3), sodass mit dem Zusatz nach (PA3) schon  $M = \mathbb{N}$  folgt. Das bedeutet:  $B(n)$  gilt für alle  $n \in \mathbb{N}$ , bzw.  $A(n)$  gilt für alle  $n \in \mathbb{N}^{\geq n_0}$   $\square$

**Beispiel 1.1.18.** Im Folgenden sind immer  $A(n)$  und die Gültigkeitsgrenze  $n_0$  angegeben (machen Sie sich auch klar, dass diese Grenze jeweils optimal ist)

$$1. \sum_{k=0}^n k^3 = \frac{n^2(n+1)^2}{4}, n_0 = 0$$

$$2. \sum_{k=0}^n q^k = \frac{q^{n+1} - 1}{q - 1}, n_0 = 0$$

$$3. \prod_{k=0}^n \left(1 + \frac{1}{2^{2^k}}\right) = \frac{2^{2^{n+1}} - 1}{2^{2^{n+1} - 1}}, n_0 = 0$$

$$4. 3 \text{ teilt } 13^n + 2, n_0 = 0$$

$$5. 47 \text{ teilt } 7^{2^n} - 2^n, n_0 = 0$$

$$6. 2^n > n^2, n_0 = 5$$

$$7. \text{ Besitzt die Menge } M \text{ genau } n \text{ Elemente, dann hat } \mathcal{P}(M) \text{ genau } 2^n \text{ Elemente, } n_0 = 0.$$

## 1.1.3 Relationen, ganze und rationale Zahlen

## 1.1.3.1 Grundbegriffe

**Definition 1.1.19** (Relation). Es seien  $M$  und  $N$  Mengen.

1. Eine *Relation zwischen  $M$  und  $N$*  ist eine Teilmenge  $R \subset M \times N$ .
2. Ist  $M = N$ , also  $R \subset M \times M$ , dann heißt  $R$  eine *Relation auf  $M$* .

Ist  $(a, b) \in R \subset M \times N$  so sagen wir  *$a$  steht in Relation zu  $b$  bezüglich  $R$*  und wir schreiben

$$a \sim_R b$$

statt  $(a, b) \in R$ . Ist klar, welche Relation gemeint ist, so schreiben wir noch kürzer  $a \sim b$ . Für spezielle Relationen benutzen wir später auch eigenständige Symbole.

**Beispiel 1.1.20.** •  $S$  sei die Menge aller Schüler einer Klasse und  $N = \{1, 2, 3, 4, 5, 6\}$ . Die Relation

$$R = \left\{ (s, n) \in S \times N \mid \begin{array}{l} \text{der Schüler } s \text{ hat in der} \\ \text{Klassenarbeit die Note } n \end{array} \right\}$$

ist eine Relation zwischen  $S$  und  $N$ . Diese nennt sich Notenspiegel.

- $S$  sei wie oben. Die Relation

$$R = \left\{ (s_1, s_2) \in S \times S \mid \begin{array}{l} \text{der Schüler } s_1 \text{ hat in der Klassen-} \\ \text{arbeit die gleiche Note wie } s_2 \end{array} \right\}$$

Ist eine Relation auf der Menge der Schüler  $S$ .

## 1.1.3.2 Ordnungs- und Äquivalenzrelationen

**Definition 1.1.21** (Eigenschaften von Relationen). Eine Relation  $R \subset M \times M$  auf einer Menge  $M$  heißt

1. *reflexiv*  $\iff \forall a \in M : a \sim_R a$

$$2. \textit{ transitiv} \iff \forall a, b, c \in M : (a \sim_R b \wedge b \sim_R c \implies a \sim_R c)$$

$$3. \textit{ symmetrisch} \iff \forall a, b \in M : (a \sim_R b \implies b \sim_R a)$$

$$4. \textit{ antisymmetrisch} \iff \forall a, b \in M : (a \sim_R b \wedge b \sim_R a \implies a = b)$$

$$5. \textit{ total} \iff \forall a, b \in M : (a \sim_R b \vee b \sim_R a)$$

**Definition 1.1.22** (Spezielle Relationen). Eine Relation  $R \subset M \times M$  auf einer Menge  $M$  heißt

1. *Halbordnung*, wenn sie 1, 2 und 4 aus Definition 1.1.21 erfüllt.
2. *Totalordnung* oder *Ordnung*, wenn sie 1, 2, 4 und 5 aus Definition 1.1.21 erfüllt.
3. *Äquivalenzrelation*, wenn sie 1, 2 und 3 aus Definition 1.1.21 erfüllt.

**Beispiel 1.1.23.** 1. Die *Teilbarkeitsrelation* auf  $\mathbb{N}^*$ ,  $T \subset \mathbb{N}^* \times \mathbb{N}^*$ , ist definiert durch

$$a \sim_T b \iff a|b \iff \exists k \in \mathbb{N}^* : b = k \cdot a$$

und ist eine Halbordnung.

2. Die *Gleichheit* auf einer Menge  $M$ ,  $G \subset M \times M$ , ist definiert durch

$$a \sim_G b \iff a = b$$

und ist eine Äquivalenzrelation.

3. Die *Ordnungsrelation* auf  $\mathbb{N}$ ,  $O \subset \mathbb{N} \times \mathbb{N}$ , ist definiert durch

$$a \sim_O b \iff a \leq b \iff \exists k \in \mathbb{N} : b = a + k$$

und ist eine Totalordnung.

4. Es sei  $M$  eine Menge und  $\mathcal{P}(M)$  ihre Potenzmenge. Wir definieren die Relation  $\tau$  auf  $\mathcal{P}(M)$  für  $U, V \in \mathcal{P}(M)$  durch

$$U \sim_\tau V \iff U \subset V.$$

Dies ist eine Halbordnung.

5. Die zweite Relation aus Beispiel 1.1.20 ist eine Äquivalenzrelation

**Bezeichnung 1.1.24.** Bezüglich der Ordnungsrelation auf  $\mathbb{N}$  schreiben wir auch

- $a \geq b \iff b \leq a,$
- $a < b \iff a \leq b \wedge a \neq b,$
- $a > b \iff b < a.$

**Bemerkung 1.1.25.** Mit der obigen Bezeichnung gilt insbesondere für alle  $a, b \in \mathbb{N}$  entweder(!)  $a < b$  oder  $a = b$  oder  $b < a$ .

**Satz 1.1.26.** Die Ordnungsrelation auf  $\mathbb{N}$  ist im folgenden Sinne mit der Addition und der Multiplikation verträglich. Sind  $a, b \in \mathbb{N}$  mit  $a \leq b$  so gilt für alle  $c \in \mathbb{N}$

$$a + c \leq b + c \quad \text{und} \quad a \cdot c \leq b \cdot c.$$

**Beweisskizze.** Wegen  $a \leq b$  gibt es ein  $k \in \mathbb{N}$  mit  $b = a + k$ . Damit ist

$$a + c \leq (a + k) + c \iff a + c \leq (a + c) + k \stackrel{\tilde{a}:=a+c}{\iff} \tilde{a} \leq \tilde{a} + k$$

und ebenso

$$a \cdot c \leq b \cdot c \iff a \cdot c \leq (a + k) \cdot c \iff a \cdot c \leq a \cdot c + k \cdot c \stackrel{\tilde{a}:=a \cdot c, \tilde{k}=k \cdot c}{\iff} \tilde{a} \leq \tilde{a} + \tilde{k}.$$

Die rechten Seiten der Äquivalenzen gelten nun per Definition und somit auch die linken.  $\square$

**Definition 1.1.27.** Ist  $\sim$  eine Äquivalenzrelation auf der Menge  $M$  so nennt man die Teilmenge

$$[a]_{\sim} := \{b \in M \mid b \sim a\}$$

die Äquivalenzklasse zu  $a$  (bezüglich  $\sim$ ). Wenn die Relation klar ist, dann schreiben wir auch  $[a]$ .

**Satz 1.1.28.** Ist  $\sim$  eine Äquivalenzrelation auf der Menge  $M$  so gilt

$$[a] = [b] \iff a \sim b.$$

Es ist entweder  $[a] \cap [b] = \emptyset$  oder  $[a] = [b]$

**Definition 1.1.29** (Klasseneinteilung). Es sei  $M$  eine Menge. Eine Teilmenge  $\mathcal{K} \subset \mathcal{P}(M)$  heißt *Klasseneinteilung von  $M$* , wenn die folgenden drei Bedingungen gelten

1.  $\forall x \in M \exists U \in \mathcal{K} : x \in U$ ,
2.  $\forall U \in \mathcal{K} : U \neq \emptyset$ ,
3.  $\forall U, V \in \mathcal{K} : (U \cap V \neq \emptyset \implies U = V)$ .

Gelten nur 1. und 2. so spricht man auch von einer *Überdeckung von  $M$* .

Dass die Begriffe Klasseneinteilung und Äquivalenzrelation im Wesentlichen die gleichen Strukturen beschreiben, zeigt der folgende Satz.

**Satz 1.1.30.** 1. Es sei  $\sim$  eine Äquivalenzrelation auf der Menge  $M$ . Dann ist

$$M / \sim := \{ [x] \mid x \in M \}$$

eine Klasseneinteilung von  $M$ . Genauer nennt man  $M / \sim$  auch eine *Äquivalenzklasseneinteilung*.

2. Es sei  $\mathcal{K}$  eine Klasseneinteilung auf  $M$ . Dann definiert diese eine Äquivalenzrelation  $\sim$  durch

$$a \sim b : \iff \exists U \in \mathcal{K} : a, b \in U.$$

Liegt die Betonung bei einer Klasseneinteilung auf der sie definierenden Äquivalenzrelation, dann spricht man auch von einer *Äquivalenzklasseneinteilung*.

**Definition 1.1.31.** Es sei  $\mathcal{K}$  eine Klasseneinteilung der Menge  $M$ .

1. Für  $U \in \mathcal{K}$  heißt ein Element  $x \in U$  ein *Repräsentant von  $U$* .
2. Wählen wir aus jeder Menge  $U \in \mathcal{K}$  genau einen Repräsentanten  $x_U \in U$ , dann heißt die Menge

$$\{x_U \mid U \in \mathcal{K}\} \subset M$$

ein *Repräsentantensystem* von  $\mathcal{K}$ .

**Beispiel 1.1.32** (Quotientenvektorraum). Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum und  $U \subset V$  ein Untervektorraum. Definiere

$$x \sim y \iff x - y \in U$$

Dann definiert das eine Äquivalenzrelation auf  $V$  und die Klasseneinteilung wird mit  $V/U$  bezeichnet. Insbesondere ist das nach geeigneter Definition der Strukturen selbst wieder ein Vektorraum.

**Beispiel 1.1.33** (Zahlenbereichserweiterungen).

1. Die Menge  $\mathbb{Z}$  der ganzen Zahlen ist eine Erweiterung der natürlichen Zahlen zu einem kommutativen Ring mit Eins. Es wird erreicht, dass alle Elemente ein additives Inverses besitzen, d.h. die Subtraktion ist uneingeschränkt durchführbar. Die Erweiterung geschieht wie folgt:

Wir definieren auf der Menge  $\mathbb{N} \times \mathbb{N}$  die Äquivalenzrelation

$$(a, b) \sim_Z (a', b') : \iff a + b' = a' + b.$$

Auf der Äquivalenzklasseneinteilung  $(\mathbb{N} \times \mathbb{N}) / \sim_Z$  definieren

$$\begin{aligned} [(a, b)] +_Z [(a', b')] &:= [(a + a', b + b')] \\ [(a, b)] \cdot_Z [(a', b')] &:= [(a \cdot a' + b \cdot b', a \cdot b' + a' \cdot b)] \end{aligned}$$

Verknüpfungen, die diese zu einem kommutativen Ring mit Eins machen. Die natürlichen Zahlen findet man darin als Elemente der Form  $[(a, 0)]$  wieder. Die Restklasseneinteilung bezeichnet man nun mit

$$\mathbb{Z} = (\mathbb{N} \times \mathbb{N}) / \sim_Z$$

Jede Restklasse hat einen ausgezeichneten Repräsentanten und für diesen schreibt man

$$a := [(a, 0)], \quad 0 := [(0, 0)], \quad -a := [(0, a)].$$

Weitere Details – insbesondere zur Einbettung von  $\mathbb{N}$  in  $\mathbb{Z}$  und zur Ordnungsrelation auf  $\mathbb{Z}$  – findet man in Abschnitt 1.9.1.

2. Ähnlich wie im vorangegangenen Beispiel ist die Menge  $\mathbb{Q}$  der ganzen Zahlen eine Erweiterung der ganzen Zahlen zu einem Körper. Es wird

erreicht, dass alle Elemente ungleich 0 ebenfalls ein Multiplikatives Inverses besitze, d.h. auch die Division ist uneingeschränkt durchführbar. Diese Zahlbereichserweiterung geschieht wie folgt:

Wir definieren auf der Menge  $\mathbb{Z} \times \mathbb{Z}^*$  die Äquivalenzrelation

$$(a, b) \sim_Q (a', b') : \iff a \cdot b' = a' \cdot b.$$

Auf der Äquivalenzklasseneinteilung  $(\mathbb{Z} \times \mathbb{Z}^*) / \sim_Q$  definieren

$$\begin{aligned} [(a, b)] +_Q [(a', b')] &:= [(a \cdot b' + a' \cdot b, b \cdot b')] \\ [(a, b)] \cdot_Q [(a', b')] &:= [(a \cdot a', b \cdot b')] \end{aligned}$$

Verknüpfungen, die diese zu einem Körper machen. Die ganzen Zahlen findet man darin als Elemente der Form  $[(a, 1)]$  wieder.

Die Restklasseneinteilung bezeichnet man nun mit

$$\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z}^*) / \sim_Q$$

Ohne sich um ausgezeichnete Restklassen zu kümmern schreibt man die Restklassen in Form von (nicht gekürzten) Brüchen

$$\frac{a}{b} := [(a, b)].$$

Weitere Details – insbesondere zur Einbettung von  $\mathbb{Z}$  in  $\mathbb{Q}$  und zur Ordnungsrelation auf  $\mathbb{Q}$  – findet man in Abschnitt [1.9.2](#).

### 1.1.3.3 Abbildungen

In diesem Abschnitt finden wir eine konkrete Definition des Begriffs der Abbildung. Wir haben ihn schon vorher benutzt im Sinne von Zuordnung, siehe Definitionen [1.1.4](#), [1.1.11](#), [1.8.15](#) und auch Beispiel [1.1.14](#).

**Definition 1.1.34.** Es seien  $M$  und  $N$  Mengen.

1. Eine Relation  $f \subset M \times N$  heißt *Abbildung* von  $M$  nach  $N$ , wenn

$$\forall x \in M \exists! y \in N : (x, y) \in f.$$

Die Menge  $M$  nennt man die *Definitionsmenge* und  $N$  den *Wertebereich* der Abbildung  $f$ .

Betonen wir  $f \subset M \times N$  als Teilmenge des Kreuzproduktes, so schreiben wir  $G_f \subset M \times N$  und nennen  $G_f$  den *Graphen der Abbildung*.

Da in der Regel der Zuordnungscharakter einer Abbildung  $f$  im Vordergrund steht, schreiben wir

$$f : M \rightarrow N \quad \text{und} \quad y = f(x)$$

statt  $f \subset M \times N$  und  $(x, y) \in f$ .

2. Ist  $f(x) = y$ , dann nennen wir  $y \in N$  das *Bild von  $x \in M$  unter  $f$*  und  $x \in M$  das *Urbild von  $y \in N$  unter  $f$* .
3. Ist  $U \subset M$ , so nennen wir die Teilmenge

$$f(U) = \{y \in N \mid \exists x \in U : f(x) = y\} \subset N$$

das *Bild von  $U$  unter  $f$* . Ist hier  $U = M$ , dann nennen wir  $\text{Bild}(f) = f(M)$  das *Bild von  $f$* .

4. Ist  $V \subset N$ , so nennen wir die Menge

$$f^{-1}(V) = \{x \in M \mid f(x) \in V\} \subset M$$

das *Urbild von  $V$  unter  $f$* . Es gilt stets  $f^{-1}(N) = M$ .

**Beispiel 1.1.35.** 1. Eine Verknüpfung  $\circ$  auf einer Menge  $M$  ist eine Abbildung  $\circ : M \times M \rightarrow M$ . Als Relation ist damit  $\circ \subset (M \times M) \times M$ .

2. Es sei  $M$  eine Menge, dann ist die Gleichheit auf  $M$  eine Abbildung. Diese wird *Identität* auf  $M$  genannt und mit  $\text{id}_M : M \rightarrow M$  und es gilt  $\text{id}_M(m) = m$  für alle  $m \in M$ .
3. Ist  $\sim$  eine Äquivalenzrelation auf  $M$ , so ist die Zuordnung  $\pi : M \rightarrow M/\sim$  mit  $\pi(x) := [x]$  eine Abbildung. Sie wird *kanonische Projektion der Äquivalenzrelation* genannt.

**Definition/Bemerkung 1.1.36.** Es sei  $f : M \rightarrow N$  eine Abbildung.

1.  $f$  heißt *injektiv*, wenn

$$\forall x, x' \in M : (f(x) = f(x') \implies x = x')$$

oder äquivalent

$$\forall x, x' \in M : (x \neq x' \implies f(x) \neq f(x')).$$

2.  $f$  heißt *surjektiv*, wenn

$$\forall y \in N \exists x \in M : f(x) = y.$$

3.  $f : M \rightarrow N$  heißt *bijektiv*, wenn  $f$  injektiv und surjektiv ist, d.h. genau dann, wenn

$$\forall y \in N \exists! x \in M : f(x) = y.$$

4. Ist  $U \subset M$ , dann nennt man die Abbildung  $f|_U : U \rightarrow N$  mit  $f|_U(x) := f(x)$  die *Einschränkung von  $f$  auf  $U$* .

5. Man kann jede Abbildung  $f : M \rightarrow N$  surjektiv machen, indem man sie als Abbildung  $f : M \rightarrow \text{Bild}(f)$  betrachtet.

**Beispiel 1.1.37.** Die Identität auf einer Menge  $M$  ist bijektiv und die kanonische Projektion einer Äquivalenzrelation ist surjektiv.

**Definition 1.1.38.** Sind  $f : M \rightarrow N$  und  $g : N \rightarrow P$  Abbildungen, dann nennt man die Abbildung

$$g \circ f : M \rightarrow P \quad \text{mit} \quad (g \circ f)(m) := g(f(m)).$$

Die *Verkettung* oder *Hintereinanderausführung* von  $f$  und  $g$ .

**Definition/Bemerkung 1.1.39.** Es sei  $f : M \rightarrow N$  eine bijektive Abbildung. Dann gilt:

- Es gibt eine eindeutige Abbildung  $f^{-1} : N \rightarrow M$  mit

$$f \circ f^{-1} = \text{id}_N \quad \text{und} \quad f^{-1} \circ f = \text{id}_M.$$

Diese heißt *Umkehrabbildung* von  $f$ .<sup>(i)</sup>

---

<sup>(i)</sup>Achtung: Die Abbildung  $f^{-1}$  gibt es nur, wenn  $f$  bijektiv ist. Verwechseln Sie dies nicht der Bezeichnung für das Urbild gemäß Definition/Bemerkung 1.1.34.4.

- Ist  $n \in N$ , so gilt für die Umkehrabbildung  $f^{-1}(n) = m$  für dasjenige  $m \in M$  mit  $f(m) = n$ .
- $f^{-1}$  besitzt selbst eine Umkehrabbildung, und es gilt  $(f^{-1})^{-1} = f$ .

## 1.2 Zahlenfolgen und reelle Zahlen

### 1.2.1 Zahlenfolgen in $\mathbb{Q}$

**Definition 1.2.1.** Eine *Zahlenfolge* in  $\mathbb{Q}$  ist eine Abbildung

$$x : \mathbb{N} \rightarrow \mathbb{Q}.$$

Die Bilder der Folge  $x$  bezeichnen wir mit

$$x_k := x(k)$$

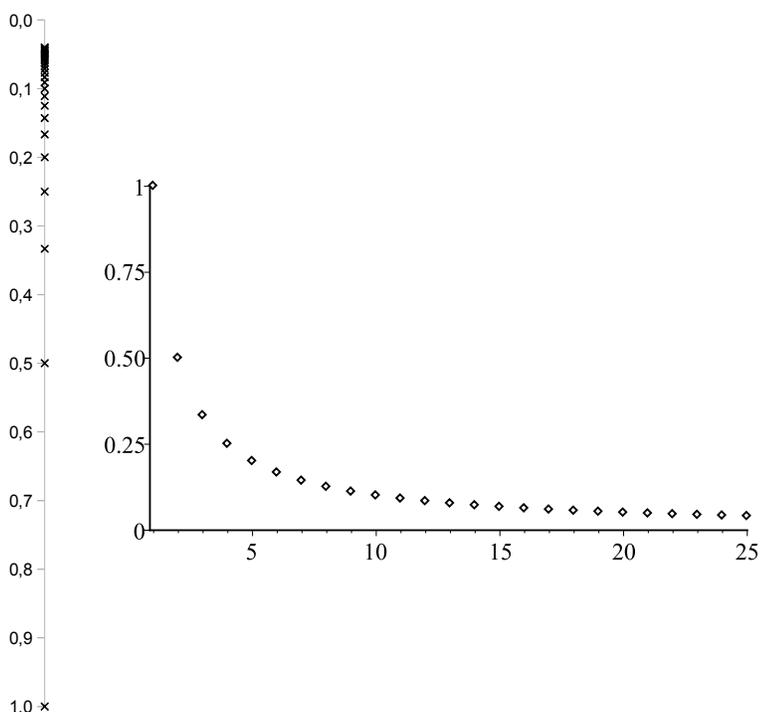
und nennen dieses Element als  $k$ -tes *Folglied*. Für eine Folge  $x$  schreiben wir auch  $(x_n)_{n \in \mathbb{N}}$  oder kürzer  $(x_n)$ .

**Beispiel 1.2.2.** •  $x : \mathbb{N} \rightarrow \mathbb{Q}$  mit  $x_n = a$  für alle  $n \in \mathbb{N}$  heißt *konstante Folge* und wir schreiben  $(a)$  für diese Folge.

- $x : \mathbb{N}^* \rightarrow \mathbb{Q}$  mit  $x_n := \frac{1}{n}$ .
- $x : \mathbb{N} \rightarrow \mathbb{Q}$  mit  $x_n := \frac{4(n+1)(n-4)}{(n^2+10)(n+5)}$ .
- $x : \mathbb{N} \rightarrow \mathbb{Q}$  mit  $x_n := (-1)^n$ .
- $x : \mathbb{N} \rightarrow \mathbb{Q}$  mit  $x_n := a^n$  für ein festes  $a \in \mathbb{Q}$ .

Man kann Zahlenfolgen auf zwei Arten veranschaulichen: Man kann Ihre Bilder auf einer "Zahlengerade" einzeichnen, oder man kann den Graphen der Folge in der "Ebene" betrachten, siehe Abbildung 1.2.1. Wir werden die zweite Art der Darstellung bevorzugen.

Wenn von jetzt an über Folgen gesprochen wird, dann ist immer eine Zahlenfolge im Sinne von Definition 1.2.1 gemeint.

Abbildung 1.2.1: Die Zahlenfolge  $(\frac{1}{n})$ 

**Definition 1.2.3.** Eine Folge  $(x_n)$  heißt

1. (a) *nach oben beschränkt*, wenn es eine Zahl  $M \in \mathbb{Q}$  gibt, sodass  $x_n < M$  für alle  $n \in \mathbb{N}$ .
- (b) *nach unten beschränkt*, wenn es eine Zahl  $M' \in \mathbb{Q}$  gibt, sodass  $x_n > M'$  für alle  $n \in \mathbb{N}$ .
- (c) *beschränkt*, wenn es eine Zahl  $C \in \mathbb{Q}$  gibt, sodass  $|x_n| < C$  für alle  $n \in \mathbb{N}$ .
2. (a) *monoton steigend*, wenn  $x_n \leq x_m$  für alle  $n, m \in \mathbb{N}$  mit  $n < m$ .
- (b) *streng monoton steigend*, wenn  $x_n < x_m$  für alle  $n, m \in \mathbb{N}$  mit  $n < m$ .
- (c) *monoton fallend*, wenn  $x_n \geq x_m$  für alle  $n, m \in \mathbb{N}$  mit  $n < m$ .
- (d) *streng monoton fallend*, wenn  $x_n > x_m$  für alle  $n, m \in \mathbb{N}$  mit  $n < m$ .

**Beispiel 1.2.4.** In der folgenden Tabelle listen wir einige Beispielfolgen mit ihren Eigenschaften.

| $x_n$                               | Eigenschaften aus Definition 1.2.3  |
|-------------------------------------|---|
| $a$                                 | konstante Folge ( $\underline{a}$ ) (beschränkt mit gleicher oberer und unterer Schranke) |
| $\frac{1}{n}$                       | beschränkt, streng monoton fallend  |
| $n^2$                               | nach unten beschränkt, streng monoton steigend  |
| $(-1)^n$                            | beschränkt  |
| $a^n$ (mit $a > 1$ )                | streng monoton steigend, nach unten beschränkt  |
| $a^n$ (mit $0 < a < 1$ )            | streng monoton fallend, beschränkt  |
| $a^n$ (mit $0 > a > -1$ )           | beschränkt  |
| $a^n$ (mit $a < -1$ )               | —   |
| $\frac{4(n+1)(n-4)}{(n^2+10)(n+5)}$ | beschränkt  |
| $20n^2 - 50n - n^3$                 | nach oben beschränkt  |
| $(-1)^n n^2$                        | —   |

**Definition 1.2.5.** • Eine Folge  $(x_n)$  heißt *konvergent*, wenn es eine Zahl  $a \in \mathbb{Q}$  gibt, sodass für alle rationalen Zahlen  $\epsilon > 0$  eine natürliche Zahl  $n_0 \in \mathbb{N}$  existiert, mit  $|x_n - a| < \epsilon$  für alle  $n > n_0$ .

Etwas formaler:

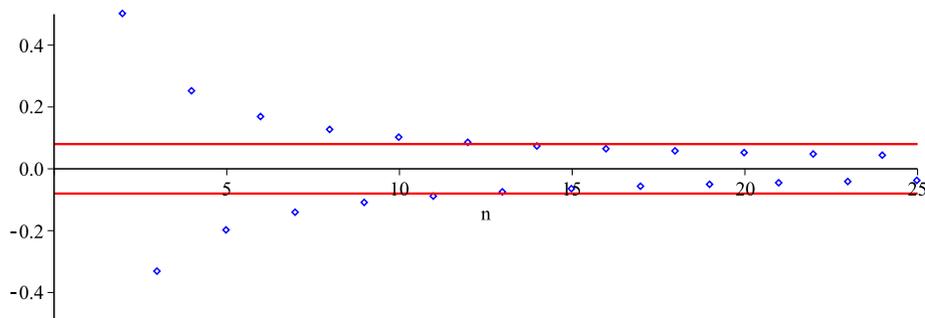
$$\forall \epsilon \in \mathbb{Q}^{>0} \exists n_0 \in \mathbb{N} \forall n \in \mathbb{N}^{>n_0} : |x_n - a| < \epsilon.$$

- $a$  nennt man *Grenzwert* der Folge  $(x_n)$  und wir schreiben

$$a = \lim_{n \rightarrow \infty} x_n \text{ oder } x_n \rightarrow a \text{ (für } n \rightarrow \infty \text{)}.$$

- Eine gegen  $a = 0$  konvergente Folge heißt auch *Nullfolge*
- Ist eine Folge nicht konvergent so heißt sie *divergent*.

Abbildung 1.2.2: Konvergenz von Zahlenfolgen



**Bemerkung 1.2.6.** Konvergiert eine Folge  $(x_n)$  gegen  $a$ , so kann es für jedes  $n_0$  nur endlich viele Zahlen  $m \in \mathbb{N}$  geben, mit  $|x_m - a| \geq \epsilon$ . Für diese  $m$  muss nämlich  $m \leq n_0$  gelten.

**Satz 1.2.7.** *Der Grenzwert einer konvergenten Folge ist eindeutig.*

**Beispiel 1.2.8.** • Die Folge  $(x_n)$  mit  $x_n = \frac{1}{n}$  konvergiert gegen  $a = 0$ . Ist  $\epsilon > 0$ , so wähle  $n_0 > \frac{1}{\epsilon}$ .<sup>(i)</sup> Dann ist für alle  $n > n_0$

$$|x_n - a| = \left| \frac{1}{n} - 0 \right| = \frac{1}{n} < \frac{1}{n_0} < \epsilon.$$

- Die Folge  $(x_n)$  mit  $x_n = \frac{(-1)^n}{n}$  konvergiert gegen  $a = 0$ , siehe Abbildung 1.2.2. Ist  $\epsilon > 0$ , so wähle wie im vorigen Beispiel  $n_0 > \frac{1}{\epsilon}$ . Dann ist für alle  $n > n_0$

$$|x_n - a| = \left| \frac{(-1)^n}{n} - 0 \right| = \frac{1}{n} < \frac{1}{n_0} < \epsilon.$$

- Die Folgen  $((-1)^n)$  und  $(n^2)$  sind divergent. Für die erste Folge zeigt man das mit der vorigen Bemerkung und für die zweite folgt das aus Satz 1.2.16.

Es folgen noch einige nützliche Konvergenzkriterien.

<sup>(i)</sup>Ist  $a > 0$  eine rationale Zahl, dann gibt es stets eine Zahl  $k \in \mathbb{N}$ , sodass  $k > a$ . eine solche Zahl erhält man etwa als den um Eins vergrößerten Zähler, in der Darstellung von  $a$  als gekürzten Bruch.

- Satz 1.2.9.** 1. Es seien  $(x_n)$  und  $(y_n)$  Zahlenfolgen mit  $\lim_{n \rightarrow \infty} (x_n - y_n) = 0$  und die Folge  $(y_n)$  konvergiere gegen  $a$ . Dann konvergiert auch  $(x_n)$  gegen  $a$ .
2. (Einschlusskriterium). Es sei  $(x_n)$  eine Folge und  $(y_n), (z_n)$  konvergente Folgen mit gleichem Grenzwert  $a$ . Weiter gebe es ein  $n_0 \in \mathbb{N}$ , sodass für alle  $n \geq n_0$  die Ungleichung  $y_n \leq x_n \leq z_n$  gelte. Dann konvergiert  $(x_n)$  ebenfalls gegen  $a$ .
3. Es sei  $(y_n)$  eine Nullfolge und  $(x_n)$  eine Folge für die ein  $a \in \mathbb{Q}$  und  $n_0 \in \mathbb{N}$  existiert mit  $|x_n - a| < |y_n|$  für alle  $n \geq n_0$ . Dann konvergiert  $(x_n)$  gegen  $a$ .
4. (Minorantenkriterium). Es seien  $(x_n)$  eine Folge und  $(y_n)$  eine unbeschränkte Folge. Weiter gebe es ein  $n_0 \in \mathbb{N}$ , sodass für alle  $n \geq n_0$  die Ungleichung  $|x_n| \geq |y_n|$  gelte. Dann ist auch  $(x_n)$  unbeschränkt.
5. (Majorantenkriterium). Es seien  $(x_n)$  eine Folge und  $(y_n)$  eine Nullfolge. Weiter gebe es ein  $n_0 \in \mathbb{N}$ , sodass für alle  $n \geq n_0$  die Ungleichung  $0 < x_n \leq y_n$  gelte. Dann ist auch  $(x_n)$  eine Nullfolge.

**Beispiel 1.2.10.** Die Folge  $((\frac{1}{k})^n)$  ist für  $k \in \mathbb{N}^{\geq 2}$  eine Nullfolge. Zur Begründung zeigt man zunächst per Induktion, dass  $2^n \geq n$  für alle  $n \geq 1$ . Dann gilt auch  $k^n > n$  für alle natürlichen Zahlen  $n \geq 1$  und  $k \geq 2$ . Damit ist  $(\frac{1}{k})^n < \frac{1}{n}$  und wir wenden Satz 1.2.9.3 an.

**Folgerung 1.2.11.** Es sei  $(a_n)$  eine Folge mit  $a_n > 0$  für alle  $n \in \mathbb{N}$  eine monoton steigende, unbeschränkte Folge. Dann ist die Folge  $(\frac{1}{a_n})$  eine Nullfolge.

**Beweisskizze.** Da  $(a_n)$  monoton steigend ist, gilt  $a_m \geq a_n$  für alle  $m > n$ . Da  $(a_n)$  nicht beschränkt ist gibt es zu jedem  $k \in \mathbb{N}$  ein  $f(k) \in \mathbb{N}$  so dass  $a_{f(k)} > k$ . Damit ist dann auch  $a_m > k$  für alle  $m > f(k)$ . Damit gilt

$$\begin{aligned} & \forall k \in \mathbb{N} \exists f(k) \in \mathbb{N} \forall m \geq f(k) : |a_m| > k \\ \implies & \forall k \in \mathbb{N} \exists f(k) \in \mathbb{N} \forall m \geq \max\{k, f(k)\} : |a_m| > k \geq m \\ \implies & \exists n_0 \in \mathbb{N} \forall m \geq n_0 : |a_m| > m \\ \implies & \exists n_0 \in \mathbb{N} \forall n \geq n_0 : \left| \frac{1}{a_n} \right| < \frac{1}{n} \end{aligned}$$

Mit Punkt 5. des vorigen Satzes und Beispiel 1.2.10 folgt nun die Behauptung.  $\square$

**Definition 1.2.12.** Es sei  $(x_n)$  eine Folge und  $f : \mathbb{N} \rightarrow \mathbb{N}$  eine Abbildung mit  $f(k) > f(\ell)$  für  $k > \ell$ . Dann nennt man die Folge  $(x_n^f)$  mit  $x_n^f := x_{f(n)}$  eine *Teilfolge* von  $(x_n)$ .

**Bemerkung 1.2.13.** Eine Abbildung  $f : \mathbb{N} \rightarrow \mathbb{N}$  mit  $f(k) > f(\ell)$  für  $k > \ell$  ist insbesondere injektiv und es gilt  $f(k) \geq k$  für alle  $k \in \mathbb{N}$ .

**Beispiel 1.2.14.** •  $(\frac{1}{2n})$  ist eine Teilfolge von  $(\frac{1}{n})$  – Es ist  $f(n) = 2n$ .

- Die konstante Folge  $(\underline{1})$  ist eine Teilfolge von  $((-1)^n)$  —  $f$  ist nicht eindeutig, es liefern z.B.  $f(n) = 2n$  und  $f(n) = 10n$  beide die konstante Teilfolge  $(\underline{1})$ .
- Es sei  $(x_n)$  eine Folge für die unendlich viele Folgenglieder nicht verschwinden. Dann ist die Folge  $(\hat{x}_n)$ , die durch Streichen der verschwindenden Folgenglieder entsteht eine Teilfolge von  $(x_n)$ .
- Der *Shift* einer Folge ist durch  $f(n) = n + k$  für  $k > 0$  definiert:  $(x_n^f)$  ist dann durch  $x_n^f = x_{n+k}$  erklärt.

**Satz 1.2.15.** *Es sei  $(x_n)$  eine konvergente Folge mit Grenzwert  $a$ . Ist  $(x_n^f)$  eine Teilfolge von  $(x_n)$ , dann ist  $(x_n^f)$  konvergent mit Grenzwert  $a$ .*

**Satz 1.2.16.** *Jede konvergente Folge in  $\mathbb{Q}$  ist beschränkt.*

**Satz 1.2.17** (Rechenregeln). *Es seien  $(x_n)$  und  $(y_n)$  konvergent mit  $\lim_{n \rightarrow \infty} x_n = a$  und  $\lim_{n \rightarrow \infty} y_n = b$ . Dann gilt*

1. Die Folge  $(x_n \pm y_n)$  konvergiert und es gilt

$$\lim_{n \rightarrow \infty} (x_n \pm y_n) = a \pm b.$$

2. Die Folge  $(x_n y_n)$  konvergiert und es gilt

$$\lim_{n \rightarrow \infty} (x_n y_n) = ab.$$

3. Ist  $y_n \neq 0$  für alle  $n \in \mathbb{N}$  und  $b \neq 0$ , so konvergiert die Folge  $\left(\frac{1}{y_n}\right)$  und es gilt

$$\lim_{n \rightarrow \infty} \frac{1}{y_n} = \frac{1}{b}.$$

4. Ist  $y_n \neq 0$  für alle  $n \in \mathbb{N}$  und  $b \neq 0$ , so konvergiert die Folge  $\left(\frac{x_n}{y_n}\right)$  und es gilt

$$\lim_{n \rightarrow \infty} \frac{x_n}{y_n} = \frac{a}{b}.$$

**Bemerkung 1.2.18.** Die Umkehrung von Satz 1.2.16 ist in der Regel nicht wahr: Ist eine Folge in  $\mathbb{Q}$  beschränkt so muss sie nicht konvergent sein, wie das Beispiel  $((-1)^n)$  zeigt.

Ist eine Folge jedoch beschränkt und monoton, so drängt sich die Frage auf, ob sie auch schon konvergiert. Im Fall der Folge  $\left(\frac{1}{n}\right)$  ist das genau der Fall. Im allgemeinen ist das jedoch nicht der Fall, wie wir im Folgenden diskutieren werden.

**Bezeichnung 1.2.19.** Eine Folge  $(x_n)$ , wie wir sie bisher betrachtet haben, war durch eine Bildungsvorschrift gegeben, die einem erlaubt jedes Folgenglied unabhängig voneinander zu berechnen. Solch eine Form nennt man auch *explizit* oder *geschlossen*. Es gibt jedoch auch Folgen, die nicht in geschlossener Form gegeben sind. Sie benötigen zur Berechnung eines Folgengliedes Informationen über eines oder mehrerer vorangehender Folgenglieder. Solche Folgen nennt man *rekursiv definierte Folgen*.

- Die sogenannte *Fibonacci Folge* ist definiert als die Folge  $(f_n)$  mit

$$f_1 = 1, \quad f_2 = 1 \quad \text{und} \quad f_n = f_{n-1} + f_{n-2} \quad \text{für } n \geq 2.$$

Für diese Folge gibt es "keine" geschlossene Form in  $\mathbb{Q}$ .<sup>(i)</sup>

- Ein weiteres Beispiel einer rekursiv definierten Folge ist  $(x_n)$  mit

$$x_0 = 1 \quad \text{und} \quad x_{n+1} = \frac{1}{2}x_n + \frac{1}{x_n}.$$

---

<sup>(i)</sup>Eine geschlossene Form der Fibonacci-Folge ist  $f_n = \frac{1}{\sqrt{5}} \left( \left( \frac{1+\sqrt{5}}{2} \right)^n - \left( \frac{1-\sqrt{5}}{2} \right)^n \right)$ .

**Beispiel 1.2.20.** Die rekursive Folge  $(x_n)$  mit  $x_0 = 1$  und  $x_{n+1} = \frac{1}{2}x_n + \frac{1}{x_n}$ . Sie ist beschränkt und monoton fallend, denn sie erfüllt

$$x_n^2 > 2 \quad \text{und} \quad x_{n+1} < x_n$$

für alle  $n \geq 1$ . Diese Folge ist nicht konvergent.

Diese „besondere“ Art der Nicht-Konvergenz wollen wir nun näher untersuchen.

**Definition 1.2.21.** Eine Folge  $(x_n)$  heißt *Cauchy-Folge*, wenn

$$\forall \epsilon \in \mathbb{Q}^{>0} \exists n_0 \in \mathbb{N} \forall n, m \in \mathbb{N}^{\geq n_0} : |x_n - x_m| < \epsilon.$$

**Bemerkung 1.2.22.** Zur Überprüfung, ob eine Folge eine Cauchyfolge ist, sind die folgenden zwei Punkte sehr nützlich:

1.  $(x_n)$  ist genau dann eine Cauchy-Folge, wenn

$$\forall \epsilon \in \mathbb{Q}^{>0} \exists n_0 \in \mathbb{N} \forall n \in \mathbb{N}^{\geq n_0} \forall k \in \mathbb{N} : |x_{n+k} - x_n| < \epsilon.$$

2. Es seien  $(x_n)$  eine Folge und  $(y_n)$  eine Nullfolge und es gelte

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 \forall k : |x_{n+k} - x_n| < |y_n|$$

so ist  $(x_n)$  eine Cauchy-Folge.

**Beispiel 1.2.23.** Die rekursiv definierte Folge aus Beispiel 1.2.20 ist eine Cauchyfolge. Wir nehmen zur Begründung an, dass als Startwert  $x_0 = 2$  gewählt ist, so dass die Folge insgesamt streng monoton fallend wird. Insbesondere ist  $x_m x_n < x_0^2 = 4$  und  $x_m x_n > 2$  für alle  $m > n > 1$ . Zunächst ist

$$\begin{aligned} |x_{n+1} - x_{m+1}| &= \left| \frac{x_n}{2} + \frac{1}{x_n} - \frac{x_m}{2} - \frac{1}{x_m} \right| \\ &= \left| \frac{1}{2}(x_n - x_m) - \frac{1}{x_m x_n}(x_n - x_m) \right| \\ &= \frac{x_m x_n - 2}{2x_m x_n} |x_n - x_m| \leq \frac{4-2}{2 \cdot 2} |x_n - x_m| = \frac{1}{2} |x_n - x_m| \end{aligned}$$

und damit

$$|x_{n+1} - x_n| \leq \frac{1}{2} |x_n - x_{n-1}| \leq \dots \leq \left(\frac{1}{2}\right)^n |x_1 - x_0|.$$

Das liefert nun

$$\begin{aligned} |x_{n+k} - x_n| &= |x_{n+k} - x_{n+k-1} + x_{n+k-1} - \cdots - x_{n+1} + x_{n+1} - x_n| \\ &\leq \sum_{\ell=0}^{k-1} |x_{n+\ell+1} - x_{n+\ell}| \leq \sum_{\ell=0}^{k-1} \left(\frac{1}{2}\right)^{n+\ell} |x_1 - x_0| \\ &= \left(\frac{1}{2}\right)^n |x_1 - x_0| \frac{1 - \left(\frac{1}{2}\right)^k}{1 - \frac{1}{2}} < |x_1 - x_0| \left(\frac{1}{2}\right)^{n-1} \end{aligned}$$

Wegen Beispiel 1.2.10 ist  $\left(\left(\frac{1}{2}\right)^n\right)_{n \in \mathbb{N}}$  eine Nullfolge und wir können die vorige Bemerkung anwenden.

**Satz 1.2.24.** 1. Jede Teilfolge  $(x_n^f)$  einer Cauchy-Folge  $(x_n)$  ist wieder eine solche.

2. Jede Cauchy-Folge in  $\mathbb{Q}$  ist beschränkt.

3. Jede konvergente Folge in  $\mathbb{Q}$  ist eine Cauchy-Folge in  $\mathbb{Q}$ .

Bisher haben wir lediglich handliche Kriterien dafür, herauszufinden, ob eine Folge nicht konvergiert. Nachzuprüfen, ob eine Folge konvergiert, können wir bisher nur, indem wir einen potentiellen Kandidaten hernehmen und testen, ob dieser tatsächlich ein Grenzwert ist.

### 1.2.2 Die reellen Zahlen

Man kann auch mit Folgen rechnen und im Prinzip haben wir das schon in 1.2.17 gemacht: Wir schreiben

$$\begin{aligned} (x_n) \pm (y_n) &:= (x_n \pm y_n) \\ (x_n) \cdot (y_n) &:= (x_n y_n) \\ (x_n) : (y_n) &:= \left(\frac{x_n}{y_n}\right), \end{aligned}$$

wobei die letzte Operation möglich ist, wenn  $y_n \neq 0$  für alle  $n \in \mathbb{N}$ .

**Bezeichnung 1.2.25.** Die Menge aller Folgen in  $\mathbb{Q}$  bezeichnen wir mit  $\mathcal{F}(\mathbb{Q})$ , die Teilmenge aller beschränkten Folgen mit  $\mathcal{BF}(\mathbb{Q})$  und die Teilmenge aller konvergenten Folgen in  $\mathbb{Q}$  mit  $\mathcal{KF}(\mathbb{Q})$ . Eine wichtige Teilmenge davon ist die

Menge  $\mathcal{NF}(\mathbb{Q})$  aller *Nullfolgen* in  $\mathbb{Q}$ , das sind die konvergenten Folgen mit Grenzwert 0. Die Menge aller Cauchy-Folgen in  $\mathbb{Q}$  bezeichnen wir entsprechend mit  $\mathcal{CF}(\mathbb{Q})$ . Es gilt

$$\mathcal{NF}(\mathbb{Q}) \subset \mathcal{KF}(\mathbb{Q}) \subset \mathcal{CF}(\mathbb{Q}) \subset \mathcal{BF}(\mathbb{Q}) \subset \mathcal{F}(\mathbb{Q}).$$

**Bemerkung 1.2.26.** 1. Dass die Menge  $\mathcal{F}(\mathbb{Q})$  unter den obigen Verknüpfungen abgeschlossen ist, sieht man sofort.

2. Satz 1.2.17 sagt gerade, dass die Menge  $\mathcal{KF}(\mathbb{Q})$  ebenfalls bezüglich der obigen Operationen abgeschlossen ist. Daraus folgt die gleiche Eigenschaft auch für  $\mathcal{NF}(\mathbb{Q})$ , wobei die Division allerdings keinen Sinn macht.
3. Ähnlich wie in Satz 1.2.17 zeigt man nun, dass die Mengen  $\mathcal{BF}(\mathbb{Q})$  und  $\mathcal{CF}(\mathbb{Q})$  ebenfalls unter Addition und Multiplikation abgeschlossen sind.
4. Weiter zeigt man durch Überprüfen der Rechenregeln, dass  $\mathcal{F}(\mathbb{Q})$ ,  $\mathcal{BF}(\mathbb{Q})$ ,  $\mathcal{KF}(\mathbb{Q})$  und  $\mathcal{CF}(\mathbb{Q})$  jeweils kommutative Ringe mit Einselement sind und  $\mathcal{NF}(\mathbb{Q})$  ein kommutativer Ring (ohne Einselement!) ist. Null und eins sind durch  $(\underline{0})$  und  $(\underline{1})$  gegeben.

All die Ringe aus der obigen Bemerkung haben Nullteiler!

**Bemerkung 1.2.27.** Ist  $(x_n) \cdot (y_n) = (\underline{0})$  so hat mindestens einer der beteiligten Folgen  $(\underline{0})$  als Teilfolge, also unendlich viele verschwindende Folgenglieder.

Wir wissen wegen Satz 1.2.24.1, dass für  $(x_n) \in \mathcal{CF}(\mathbb{Q})$  jede Teilfolge  $(x_n^f)$  von  $(x_n)$  selbst in  $\mathcal{CF}(\mathbb{Q})$  liegt. Damit folgt nun:

**Satz 1.2.28.** Ist  $(x_n) \in \mathcal{CF}(\mathbb{Q})$  und  $(x_n^f)$  eine Teilfolge, so ist  $(x_n) - (x_n^f) \in \mathcal{NF}(\mathbb{Q})$ .

**Folgerung 1.2.29.** Die Nullteiler in  $\mathcal{CF}(\mathbb{Q})$  sind Nullfolgen.

**Definition/Satz 1.2.30.** Auf dem Ring  $\mathcal{CF}(\mathbb{Q})$  der Cauchy-Folgen in  $\mathbb{Q}$  definieren wir eine Relation durch

$$(x_n) \sim (y_n) : \iff (x_n - y_n) \in \mathcal{NF}(\mathbb{Q}).$$

Dies ist eine Äquivalenzrelation und die Äquivalenzklasseneinteilung bezeichnen wir mit

$$\widehat{\mathbb{R}} := \mathcal{CF}(\mathbb{Q}) / \sim .$$

**Satz 1.2.31.** 1. Die Addition und die Multiplikation auf  $\mathcal{CF}(\mathbb{Q})$  induzieren auf natürliche Weise eine Addition und Multiplikation auf  $\widehat{\mathbb{R}}$ . Es ist

$$[(x_n)] \oplus [(y_n)] := [(x_n + y_n)], \quad [(x_n)] \odot [(y_n)] := [(x_n y_n)].$$

2. Diese Verknüpfungen machen  $\widehat{\mathbb{R}}$  zu einem Körper.

**Beweisskizze.** (Satz 1.2.31). Zunächst zeigt man für 1. die Repräsentantenunabhängigkeit der Verknüpfungen. Für 2. bemerkt man, dass die neutralen Elemente in  $\mathbb{R}$  die Klassen der neutralen Elemente in  $\mathcal{CF}(\mathbb{Q})$  sind, also  $[(\underline{0})]$  und  $[(\underline{1})]$ .

Ist  $\alpha \in \widehat{\mathbb{R}} \setminus \{[(\underline{0})]\}$ , so besitzt jeder Repräsentant  $(x_n) \in \alpha$  nur endlich viele verschwindende Folgenglieder. Betrachte dazu dann die Teilfolge  $(\hat{x}_n)$ , die aus  $(x_n)$  durch streichen eben dieser Folgenglieder entsteht. Schließlich zeigt man, dass  $\left[\left(\frac{1}{\hat{x}_n}\right)\right]$  das multiplikative Inverse zu  $[(x_n)]$  ist.  $\square$

Die rationalen Zahlen lassen sich so in die reellen Zahlen einbetten, dass alle Strukturen, die wir in  $\mathbb{Q}\mathbb{Q}$  kennengelernt haben weiterhin gültig und mit den analogen Strukturen in  $\mathbb{R}$  verträglich sind.

**Satz 1.2.32.** Die Abbildung  $\Phi : \mathbb{Q} \rightarrow \widehat{\mathbb{R}}$ , die jedem  $q \in \mathbb{Q}$  die Äquivalenzklasse der konstanten Folge  $(q)$  zuordnet, ist ein Körperhomomorphismus. Da ein Körperhomomorphismus stets injektiv ist, ist  $\Phi : \mathbb{Q} \rightarrow \Phi(\mathbb{Q})$  bijektiv.

**Bezeichnung 1.2.33.** Ist eine Cauchy-Folge  $(x_n)$  in  $\mathbb{Q}$  konvergent mit Grenzwert  $a \in \mathbb{Q}$ , dann ist  $[(x_n)] = [(\underline{a})]$  und wir identifizieren  $[(x_n)] \in \widehat{\mathbb{R}}$  und  $a \in \mathbb{Q}$ .

**Definition 1.2.34.** Ein Element  $\alpha \in \widehat{\mathbb{R}}$  nennen wir *positiv*, wenn es ein  $\epsilon \in \mathbb{Q}^{>0}$  und eine Cauchy-Folge  $(x_n) \in \alpha$  gibt, sodass  $x_n \geq \epsilon$  für alle  $n \in \mathbb{N}$ . Etwas formaler

$$\exists \epsilon \in \mathbb{Q}^{>0} \exists (x_n) \in \alpha \forall n \in \mathbb{N} : x_n \geq \epsilon.$$

Weiter definieren wir

$$\alpha \otimes \beta : \iff \alpha - \beta \text{ ist positiv,}$$

$$\alpha \supseteq \beta : \iff \alpha \otimes \beta \vee \alpha = \beta,$$

$$\alpha \ominus \beta : \iff \beta \supseteq \alpha,$$

$$\alpha \cong \beta : \iff \beta \supseteq \alpha.$$

Wir nennen  $\alpha$  *negativ*, wenn  $-\alpha$  positiv ist. Es ist  $\alpha$  positiv  $\iff \alpha \otimes [(\mathbb{Q})]$  und  $\alpha$  negativ  $\iff \alpha \ominus [(\mathbb{Q})]$ .

**Bemerkung 1.2.35.** Ist  $\alpha$  positiv und  $\epsilon \in \mathbb{Q}^{>0}$  wie oben, dann gibt es für alle  $(x_n) \in \alpha$  ein  $n_0 \in \mathbb{N}$ , sodass  $x_k > \frac{\epsilon}{2}$  für alle  $k > n_0$ .

**Satz 1.2.36.** •  $\otimes$  ist eine Totalordnung auf  $\widehat{\mathbb{R}}$  und erfüllt die gewöhnlichen Rechenregeln, siehe Kapitel 1.2.5.

- Für alle  $q, r \in \mathbb{Q}$  gilt

$$\Phi(q) \otimes \Phi(r) \iff q \leq r.$$

**Bezeichnung 1.2.37.** Wir schreiben

$$\begin{array}{ll} \mathbb{R} & \text{statt } \widehat{\mathbb{R}} \\ + & \text{statt } \oplus \\ \cdot & \text{statt } \odot \\ \alpha - \beta & \text{statt } \alpha \oplus (-\beta) \\ \frac{\alpha}{\beta} & \text{statt } \alpha \odot \beta^{-1} \\ \leq & \text{statt } \otimes \\ \mathbb{Q} & \text{statt } \Phi(\mathbb{Q}) \\ a & \text{statt } [(\underline{a})] \end{array}$$

Die Zahlen  $\mathbb{R} \setminus \mathbb{Q}$  heißen *irrationale Zahlen*.

### 1.2.3 Cauchyfolgen in $\mathbb{R}$ : die Vollständigkeit von $\mathbb{R}$ und das Supremumsprinzip

Bevor wir uns den speziellen Eigenschaften reeller Zahlen zuwenden, wollen wir den Folgenbegriff auf die Menge der reellen Zahlen ausweiten.

**Definition 1.2.38.** Eine *reelle Zahlenfolge* ist eine Abbildung  $\alpha : \mathbb{N} \rightarrow \mathbb{R}$ . Für solch eine Folge  $\alpha$  mit Bildern  $\alpha_n := \alpha(n)$  schreiben wir wieder  $(\alpha_n)_{n \in \mathbb{N}}$  oder  $(\alpha_n)$ .

Die Definitionen aus Kapitel 1.2.1 werden gleichlautend übernommen, wobei stets  $\mathbb{Q}$  durch  $\mathbb{R}$  ersetzt wird.

Aus der Verträglichkeit der Einbettung  $\Phi : \mathbb{Q} \rightarrow \mathbb{R}$  und der Bemerkung 1.2.35 folgt nun die wichtige Eigenschaft

**Satz 1.2.39.** *a) Für eine Cauchyfolge  $(x_n)$  in  $\mathbb{Q}$  ist die Folge  $(\Phi(x_n))$  eine Cauchyfolge in  $\mathbb{R}$  ist.*

*b) Für eine Cauchyfolge  $(\alpha_n)$  in  $\mathbb{R}$  mit  $\alpha_n \in \Phi(\mathbb{Q})$  für alle  $n$ , ist  $(\Phi^{-1}(\alpha_n))$  eine Cauchyfolge in  $\mathbb{Q}$ .*

Weiter besitzt jede Cauchyfolge in  $\mathbb{Q}$  einen Grenzwert in  $\mathbb{R}$ , oder in Termen von  $\mathbb{R}$ :

**Satz 1.2.40.** *Ist  $(\alpha_n)$  eine Cauchyfolge in  $\mathbb{R}$  mit  $\alpha_n \in \Phi(\mathbb{Q})$  für alle  $n$ , dann gibt es ein  $\alpha \in \mathbb{R}$  mit  $\lim_{n \rightarrow \infty} \alpha_n = \alpha$*

Weiter lässt sich jede reelle Zahl beliebig genau durch eine rationale Zahl annähern:

**Satz 1.2.41.** *Ist  $\alpha \in \mathbb{R}$  und  $\epsilon \in \mathbb{R}^{>0}$ , dann gibt es eine rationale Zahl  $x \in \mathbb{Q}$  mit  $|\alpha - \Phi(x)| < \epsilon$ .*

Die drei vorigen Sätze, die Aussagen über die Beziehung zwischen reellen Zahlen und rationalen Zahlen, sowie deren Cauchyfolgen machen, liefern nun die wichtige Aussage

**Theorem 1.2.42.** *Der Körper  $\mathbb{R}$  der reellen Zahlen ist vollständig, das heißt jede Cauchyfolge in  $\mathbb{R}$  konvergiert in  $\mathbb{R}$ .*

**Satz 1.2.43.** *Alle Sätze und Bemerkungen aus Kapitel 1.2.1 bleiben nach Ersetzen von  $\mathbb{Q}$  durch  $\mathbb{R}$  wahr mit Ausnahme von Bemerkung 1.2.18, wie wir gleich sehen werden. Es gilt insbesondere*

- *Jede konvergente reelle Folge ist beschränkt.*
- *Jede reelle Cauchy-Folge ist beschränkt.*
- *Jede konvergente reelle Folge ist eine Cauchy-Folge.*

Als wichtigstes Argument zur Einführung der reellen Zahlen wird nun die Bemerkung 1.2.18 durch folgenden zentralen Satz ersetzt.

**Satz 1.2.44** (Cauchy-Konvergenzkriterium). *Eine reelle Zahlenfolge ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.*

Bevor wir uns weiteren Eigenschaften reeller Zahlenfolgen ansehen, wenden wir uns zunächst Eigenschaften allgemeinerer Teilmengen zu. Dazu benötigen wir die folgende Definition.

**Definition 1.2.45.** Es sei  $\mathcal{M}$  eine Menge mit einer Ordnungsrelation  $\leq$  und  $\mathcal{M}' \subset \mathcal{M}$  eine Teilmenge.

1. (a) Ein Element  $M \in \mathcal{M}'$  heißt *Maximum von  $\mathcal{M}'$* , wenn es das größte Element der Menge  $\mathcal{M}'$  ist; d. h.

$$\forall x \in \mathcal{M}' : x \leq M .$$

Wir schreiben dann  $M = \max(\mathcal{M}')$ .

- (b) Ein Element  $m \in \mathcal{M}'$  heißt *Minimum von  $\mathcal{M}'$* , wenn es das kleinste Element der Menge  $\mathcal{M}'$  ist; d. h.

$$\forall x \in \mathcal{M}' : x \geq m .$$

Wir schreiben dann  $m = \min(\mathcal{M}')$ .

2. (a) Ein Element  $C \in \mathcal{M}$  heißt *obere Schranke von  $\mathcal{M}'$* , wenn es größer/gleich allen Elementen der Menge  $\mathcal{M}'$  ist; d. h..

$$\forall x \in \mathcal{M}' : x \leq C .$$

- (b) Ein Element  $c \in \mathcal{M}$  heißt *untere Schranke von  $\mathcal{M}'$* , wenn es kleiner/gleich allen Elementen der Menge  $\mathcal{M}'$ ; d. h.

$$\forall x \in \mathcal{M}' : x \geq c .$$

3. (a) Ein Element  $S \in \mathcal{M}$  heißt *Supremum* der Menge  $\mathcal{M}'$ , wenn es die kleinste ober Schranke von  $\mathcal{M}'$  ist; d. h.

$$\forall S' \in \mathcal{M} (\forall x \in \mathcal{M}' : x \leq S \wedge x \leq S') \implies S \leq S' .$$

Wir schreiben dann  $S = \sup(\mathcal{M}')$ .

- (b) Ein Element  $s \in \mathcal{M}$  heißt *Infimum* der Menge  $\mathcal{M}'$ , wenn es die größte untere Schranke von  $\mathcal{M}'$  ist; d. h.

$$\forall s' \in \mathcal{M} (\forall x \in \mathcal{M}' : x \geq s \wedge x \geq s') \implies s \geq s'.$$

Wir schreiben dann  $s = \inf(\mathcal{M}')$ .

Insbesondere ist ein Maximum (Minimum) einer Menge auch ein Supremum (Infimum).

**Bemerkung 1.2.46.** In einer geordneten Menge müssen für eine Teilmenge weder Maximum, Minimum, Supremum oder Infimum existieren:

Betrachte dazu die Potenzmenge  $\mathcal{P}(M)$  einer Menge  $M$ . Diese ist durch die Teilmengenbeziehung geordnet und besitzt ein Minimum ( $\emptyset$ ) und ein Maximum ( $M$ ) als untere und obere Schranken.

Entfernt man nun diese beiden Elemente, so ist die Menge  $\mathcal{P}(M) \setminus \{\emptyset, M\}$  immer noch durch  $\subset$  geordnet. Sie besitzt weder Maximum noch Minimum. Als Teilmenge von  $\mathcal{P}(M)$  ist sie jedoch weiter durch  $\emptyset$  und  $M$  beschränkt und sie bilden hier das Infimum und das Supremum.

Betrachten wir eine beliebige Teilmenge  $\mathcal{A} \subset \mathcal{P}(M)$  so gilt allgemeiner  $\inf(\mathcal{A}) = \cap \mathcal{A}$  und  $\sup(\mathcal{A}) = \cup \mathcal{A}$ .

In den reellen Zahlen gilt jedoch die folgende Eigenschaft.

**Satz 1.2.47** (Supremumseigenschaft).

- *Jede nach oben beschränkte Teilmenge  $M \subset \mathbb{R}$  hat ein Supremum.*
- *Jede nach unten beschränkte Teilmenge  $M \subset \mathbb{R}$  hat ein Infimum.*

Wir erhalten nun die "fehlenden" Eigenschaften von Zahlenfolgen, die uns als Motivation zur Einführung der reellen Zahlen dienten.

**Satz 1.2.48.** • *Jede nach oben beschränkte, monoton steigende reelle Folge ist konvergent.*

- *Jede nach unten beschränkte, monoton fallende reelle Folge ist konvergent.*
- *Jede beschränkte, monotone reelle Folge ist konvergent.*

1.2.4 Häufungspunkte und weitere Eigenschaften von  $\mathbb{R}$ 

Auch wenn eine Zahlenfolge keinen Grenzwert besitzt, kann es Werte geben, in deren Nähe sich die Folgenglieder bevorzugt aufhalten. Das Standardbeispiel ist sich die Zahlenfolge  $((-1)^n)_{n \in \mathbb{N}}$ , die nicht konvergiert, aber deren Werte sich stets 'in der Nähe' von  $+1$  oder  $-1$  aufhalten.

**Definition 1.2.49.** Ein Punkt  $a \in \mathbb{R}$  heißt *Häufungspunkt* der reellen Folge<sup>(i)</sup>  $(x_n)$ , wenn es zu jeder reellen Zahl  $\epsilon > 0$  eine unendliche Menge  $N \subset \mathbb{N}$  gibt, sodass

$$|x_n - a| < \epsilon \text{ für alle } n \in N.$$

**Satz 1.2.50.**  $a \in \mathbb{R}$  ist genau dann ein Häufungspunkt der reellen Folge  $(x_n)$ , wenn es eine Teilfolge  $(x_n^f)$  von  $(x_n)$  gibt, die gegen  $a$  konvergiert.

**Beweisskizze.** Für die Richtung " $\Rightarrow$ " konstruieren wir eine Teilfolge über ihre definierende Funktion  $f : \mathbb{N} \rightarrow \mathbb{N}$  wie folgt:

Sei  $B_n := \{k \in \mathbb{N} \mid |x_k - a| < \frac{1}{n+1}\}$  und damit  $f(0) := \min(B_0)$  und rekursiv  $f(k) := \min(B_k \setminus \mathbb{N}^{\leq f(k-1)})$  für  $k \geq 1$  □

**Satz 1.2.51.** 1. Jede reelle Folge besitzt eine monotone Teilfolge.

2. (Bolzano-Weierstraß). Jede beschränkte reelle Folge hat einen Häufungspunkt.

3. Jede beschränkte Folge mit genau einem Häufungspunkt ist konvergent.

**Beweisskizze.** (zu 1) Wir konstruieren eine Teilfolge über Ihre definierende Funktion  $f : \mathbb{N} \rightarrow \mathbb{N}$  wie folgt:

Sei  $B := \{k \in \mathbb{N} \mid \forall n \geq k : x_k \leq x_n\}$ . Dann unterscheiden wir zwei Fälle: Ist  $B$  unendlich, dann sei  $f(0) := \min(B)$  und für  $k \geq 1$  induktiv  $f(k) := \min(B \setminus \mathbb{N}^{\leq f(k-1)})$ . Das liefert eine monoton steigende Teilfolge.

Ist  $B$  endlich, so gibt es ein  $N \in \mathbb{N} \setminus B$ , das eine obere Schranke für  $B$  ist. Dazu gibt es ein  $f(0) \in \mathbb{N} \setminus B$  mit  $f(0) > N$  und  $x_{f(0)} < x_N$ . Jetzt wählen wir induktiv  $f(k) \in \mathbb{N} \setminus B$  mit  $f(k) > f(k-1)$  und  $x_{f(k)} < x_{f(k-1)}$ . Das definiert nun eine monoton fallende Teilfolge.

---

<sup>(i)</sup>Ab jetzt werden wir Elemente aus  $\mathbb{R}$  und  $\mathbb{Q}$  nicht mehr unterscheiden und nutzen die gleichen Bezeichnungen.

(Zu  $\beta$ ) Wir zeigen, dass der Häufungspunkt auch der Grenzwert der Folge ist. Wir betrachten zu  $\epsilon > 0$  die Menge  $B_\epsilon := \{n \in \mathbb{N} \mid |x_n - a| < \epsilon\}$  und zeigen, dass  $\mathbb{N} \setminus B_\epsilon$  endlich ist.

Dazu nehmen wir an, dass  $\mathbb{N} \setminus B_\epsilon$  nicht endlich ist. Sei nun  $f(0) = \min(\mathbb{N} \setminus B_\epsilon)$  und induktiv  $f(k) := \min((\mathbb{N} \setminus B_\epsilon) \setminus \mathbb{N}^{\leq f(k-1)})$  für  $k \geq 1$ . Mit  $(x_n)$  ist auch  $(x_n^f)$  beschränkt. Da  $(x_n^f)$  monoton ist, hat diese Folge einen Häufungspunkt. Wegen der Definition von  $B_\epsilon$  kann dieser dann aber nicht  $a$  sein.  $\square$

Aus dem Supremumsprinzip folgt, dass  $\mathbb{R}$  im folgenden Sinne zusammenhängend ist:

**Satz 1.2.52** (Zusammenhang von  $\mathbb{R}$ ). *Es seien  $M, M' \subset \mathbb{R}$  Teilmengen mit  $x \leq x'$  für alle  $x \in M$  und  $x' \in M'$ . Dann existiert ein  $y \in \mathbb{R}$  mit  $x \leq y \leq x'$  für alle  $x \in M, x' \in M'$ .*

**Bemerkung 1.2.53.** • Dass Satz 1.2.52 über  $\mathbb{Q}$  nicht gilt, zeigen die Mengen  $M = \{x \in \mathbb{Q} \mid x^2 < 2\}$  und  $M' = \{x \in \mathbb{Q} \mid x^2 > 2\}$ .

- Ganz allgemein sind auf einer total geordneten Menge  $(M, \leq)$  die folgenden Aussagen äquivalent:
  - Zu zwei Teilmengen  $N, N' \subset M$  mit  $n \leq n'$  für alle  $n \in N, n' \in N'$  existiert ein  $m \in M$ , sodass  $n \leq m \leq n'$  für alle  $n \in N, n' \in N'$ .
  - Jede nach oben beschränkte Menge hat ein Supremum.
  - Jede nach unten beschränkte Menge hat ein Infimum.

**Satz 1.2.54.** *Für die reellen Zahlen gelten die folgenden Aussagen:*

1. (a)  $\mathbb{N} \subset \mathbb{R}$  ist nicht beschränkt.  
 (b) Zu  $a \in \mathbb{R}$  gibt es zu jeder Zahl  $\epsilon \in \mathbb{R}^+$  ein  $n \in \mathbb{N}$ , sodass  $n\epsilon > a$ .  
 (c) Zu jedem  $\epsilon \in \mathbb{R}^+$  gibt es ein  $n \in \mathbb{N}$ , sodass  $\frac{1}{n} < \epsilon$ .
2. (a) Zwischen zwei reellen Zahlen liegt immer eine rationale Zahl.  
 (b) Zwischen zwei reellen Zahlen liegt immer eine irrationale Zahl.  
 (c) Zwischen zwei rationalen Zahlen liegt immer eine irrationale Zahl.
3. Zwischen zwei reellen/rationalen Zahlen liegen sowohl unendlich viele rationale Zahlen als auch irrationale Zahlen.

4. Zu jeder reellen Zahl  $\alpha \in \mathbb{R}$  und zu jedem  $\epsilon \in \mathbb{R}^+$  gibt es ein  $a \in \mathbb{Q}$  mit  $|a - \alpha| < \epsilon$ .

**Bemerkung 1.2.55.** Die Eigenschaften 1.(a)-(c) machen  $\mathbb{R}$  zu einem *archimedischen Körper*. Sie gelten auch, wenn man  $\mathbb{R}$  durch  $\mathbb{Q}$  ersetzt, sodass auch die rationalen Zahlen einen archimedischen Körper bilden. Dies haben wir bei der Untersuchung der rationalen Zahlenfolge  $(\frac{1}{n})$  verwendet, siehe Fußnote (i).

Damit haben wir auch das folgende elementare Beispiel als Erweiterung zu Beispiel 1.2.10.

**Beispiel 1.2.56.** • Die rationale Folge  $(q^n)_{n \in \mathbb{N}}$  ist für  $0 < q < 1$  eine streng monoton fallende Nullfolge.

- Die rationale Folge  $(q^n)_{n \in \mathbb{N}}$  ist für  $q > 1$  streng monoton steigend und unbeschränkt.

Zur Begründung der beiden äquivalenten Aussagen schreiben wir im zweiten Fall  $q = (1 + r)$  mit  $r \in \mathbb{Q}^{>0}$  und nutzen die Bernoulli-Ungleichung<sup>(i)</sup>

$$(1 + r)^n \geq 1 + nr$$

sowie, dass die Folge  $(rn)_{n \in \mathbb{N}}$  unbeschränkt ist.

Wir können Infimum und Supremum einer Menge auch wie folgt charakterisieren

**Satz 1.2.57.** *Es sei  $M \subset \mathbb{R}$  eine Teilmenge und  $a \in \mathbb{R}$ . Dann gilt*

$$1. \ a = \inf(M) \iff \begin{cases} (i) & a \text{ ist untere Schranke von } M \\ (ii) & \forall \epsilon > 0 \exists x \in M : x - a < \epsilon \end{cases}$$

$$2. \ a = \sup(M) \iff \begin{cases} (i) & a \text{ ist obere Schranke von } M \\ (ii) & \forall \epsilon > 0 \exists x \in M : a - x < \epsilon \end{cases}$$

<sup>(i)</sup>Diese zeigt man durch Induktion über  $n$  und sie gilt für  $x > -1$ .

## 1.2.5 Rechenregeln, Intervalle

Betrag, Vorzeichen und Ordnungsrelation auf  $\mathbb{R}$  erfüllen die üblichen Rechenregeln, so wie wir sie zum Teil im vorigen Abschnitt schon benutzt haben.

**Bemerkung 1.2.58** (Rechenregeln). Es seien im Folgenden  $a, b, c \in \mathbb{R}$  und  $m, n \in \mathbb{N}$ . Dann gilt der *Binomische Lehrsatz*

$$a^{n+1} - b^{n+1} = (a - b) \sum_{k=0}^n a^k b^{n-k} \quad \text{und} \quad (a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$$

mit  $\binom{n}{k} := \frac{n!}{k!(n-k)!}$  für  $n \geq k$ .<sup>(i)</sup> Außerdem gelten die Rechenregeln aus Tabelle 1.

**Definition/Bemerkung 1.2.59.** 1. Für  $a, b \in \mathbb{R}$  heißen

$$]a, b[ := \{x \in \mathbb{R} \mid a < x < b\} \quad \text{und} \quad [a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

das *offene Intervall* und das *abgeschlossene Intervall* mit den *Intervallgrenzen*  $a$  und  $b$ .

2. Die Zahlen  $m = m(a, b) := \frac{a+b}{2}$  und  $r = r(a, b) := \frac{|a-b|}{2}$  erfüllen

$$]a, b[ = \{x \in \mathbb{R} \mid |x - m| < r\} \quad \text{und} \quad [a, b] = \{x \in \mathbb{R} \mid |x - m| \leq r\}$$

und heißen *Mittelpunkt*  $x_0$  und *Radius* des Intervalls. Wir schreiben manchmal

$$B_r(x_0) := \{x \in \mathbb{R} \mid |x - x_0| < r\} \quad \text{und} \quad \bar{B}_r(x_0) := \{x \in \mathbb{R} \mid |x - x_0| \leq r\}.$$

3. Ein *Intervall*  $I \subset \mathbb{R}$  ist eine Menge mit der Eigenschaft, dass mit je zwei Punkten  $a, b$  die Punkte dazwischen ebenfalls in  $I$  liegen, d.h.

$$\forall a, b \in I : [a, b] \subset I \vee ]a, b[ \subset I.$$

Zu den Intervallen gehören somit auch die *halboffenen Intervalle*

$$]a, b] := \{x \in \mathbb{R} \mid a < x \leq b\} \quad \text{und} \quad [a, b[ := \{x \in \mathbb{R} \mid a \leq x < b\}$$

---

<sup>(i)</sup>Um auch die Fälle  $a + b = 0$ ,  $n = 0$  oder  $ab = 0$  einzuschließen verabreden wir hier  $0^0 = 1$ .

Tabelle 1: Rechenregeln

|                         |  |                                    |                                     |
|-------------------------|--|------------------------------------|-------------------------------------|
| $a < b$                 | $\implies a + c < b + c$                 | $ a $                              | $= \max\{a, -a\}$                   |
| $a < b$                 | $\implies -a > -b$                       | $- a $                             | $= \min\{a, -a\}$                   |
| $a > 0, b > 0$          | $\implies a + b > 0$                     | $ -a $                             | $=  a $                             |
| $a < b, c > 0$          | $\implies ac < bc$                       | $  a  $                            | $=  a $                             |
| $a < b, c < 0$          | $\implies ac > ab$                       | $ a - b $                          | $=  b - a $                         |
| $a > 0, b > 0$          | $\implies ab > 0$                        | $\max\{a, b\}$                     | $= \frac{a+b}{2} + \frac{ a-b }{2}$ |
| $0 < a < b$             | $\implies 0 < \frac{1}{b} < \frac{1}{a}$ | $\min\{a, b\}$                     | $= \frac{a+b}{2} - \frac{ a-b }{2}$ |
| $a < b < 0$             | $\implies \frac{1}{b} < \frac{1}{a} < 0$ | $ ab $                             | $=  a   b $                         |
| $a < b$                 | $\implies a < \frac{a+b}{2} < b$         | $ \frac{a}{b} $                    | $= \frac{ a }{ b }, b \neq 0$       |
| $a \neq 0, m > 0$       | $\implies a^m \neq 0$                    | $ a + b $                          | $\leq  a  +  b $                    |
| $a > 1, m > 0$          | $\implies a^m > 1$                       | $ a - b $                          | $\geq  a  -  b $                    |
| $a > 1, m < n$          | $\implies a^m < a^n$                     | $ a - b $                          | $\geq   a  -  b  $                  |
| $0 < a < 1, m < n$      | $\implies a^n < a^m$                     | $a^m a^n$                          | $= a^{m+n}$                         |
| $0 \leq a < b, m > 0$   | $\implies a^m < b^m$                     | $a^m b^m$                          | $= (ab)^m$                          |
| $a \geq -1$             | $\implies (1 + a)^m \geq 1 + ma$         | $(a^n)^m$                          | $= a^{mn}$                          |
| $ a  > c \wedge a > -c$ | $\implies  a  >  c $                     | $\binom{n}{k}$                     | $= \binom{n}{n-k}$                  |
| $ a  < c \wedge a < -c$ | $\implies  a  <  c $                     | $\binom{n+1}{k}$                   | $= \binom{n}{k-1} + \binom{n}{k}$   |
| $ a  = 0$               | $\iff a = 0$                             | $\sum_{k=0}^n (-1)^k \binom{n}{k}$ | $= 0$ für $n \geq 1$                |
| $ a - b  < c$           | $\iff b - c < a < b + c$                 | $\sum_{k=0}^n \binom{n}{k}$        | $= 2^n$                             |

und die *unbeschränkten Intervalle*

$$[a, \infty[ := \mathbb{R}^{\geq a}, ]a, \infty[ := \mathbb{R}^{> a}, ]\infty, a] := \mathbb{R}^{\leq a}, ]\infty, a[ := \mathbb{R}^{< a}.$$

Der Schnitt zweier Intervalle ist ein Intervall und der Schnitt abzählbar

vieler Intervalle ist ein Intervall.

4. Die Vereinigung zweier nichtleerer Intervalle ist genau dann ein Intervall, wenn der Schnitt nicht leer ist, oder, wenn die Vereinigungen von der Form<sup>(ii)</sup>  $]a, b[ \cup ]b, c[$  oder  $]a, b[ \cup [b, c[$  sind.

Mit diesen Bezeichnungen lässt sich der folgende Satz elegant beweisen.

**Satz 1.2.60** (Überabzählbarkeit von  $\mathbb{R}$ ). *Die Menge der reellen Zahlen  $\mathbb{R}$  und die Menge der irrationalen Zahlen  $\mathbb{R} \setminus \mathbb{Q}$  sind überabzählbar, d.h. nicht abzählbar.*

**Bemerkung 1.2.61.** Zur Abzählbarkeit siehe auch Satz 1.9.9, Bemerkung 1.9.10 und Bemerkung 1.9.19.

## 1.3 Reelle Funktionen

### 1.3.1 Grundlagen und erste Beispiele

**Definition 1.3.1.** Es seien  $U, V \subset \mathbb{R}$  Teilmengen der reellen Zahlen.

1. Eine (*reelle*) *Funktion* auf  $U$  ist eine Abbildung

$$f : U \rightarrow \mathbb{R}.$$

Eine besondere Rolle werden Funktionen spielen, deren Definitionsbereich Intervalle  $I \subset \mathbb{R}$  ist.

2. Ein Wert  $x \in U$  heißt *Nullstelle* der Funktion  $f : U \rightarrow \mathbb{R}$ , wenn  $f(x) = 0$ . Die Menge der Nullstellen von  $f$  bezeichnen wir mit  $\mathcal{N}_f := \{x \in U \mid f(x) = 0\}$ .
3. Sind  $f : U \rightarrow \mathbb{R}$  und  $g : V \rightarrow \mathbb{R}$  Funktionen und gilt  $f(x) = g(x)$  für einen Wert  $x \in U \cap V$ , so nennt man den Punkt  $(x, f(x)) = (x, g(x))$  *Schnittpunkt* der Funktionen  $f$  und  $g$ .

---

<sup>(ii)</sup>Mit dem Symbol  $] \dots [$  wollen wir ausdrücken, dass an diesem Rand des Intervalls die Art des Randes unerheblich ist.

**Definition 1.3.2.** Es sei  $f : U \rightarrow \mathbb{R}$  eine Funktion.  $f$  ist

- *monoton steigend*  $\iff \forall x, x' \in U : (x \leq x' \implies f(x) \leq f(x'))$ ,
- *monoton fallend*  $\iff \forall x, x' \in U : (x \leq x' \implies f(x) \geq f(x'))$ ,
- *streng monoton steigend*  $\iff \forall x, x' \in U : (x < x' \implies f(x) < f(x'))$ ,
- *streng monoton fallend*  $\iff \forall x, x' \in U : (x < x' \implies f(x) > f(x'))$ ,

**Beispiel 1.3.3.** • Jede Zahlenfolge  $(x_n)_{n \in \mathbb{N}}$  definiert eine Funktion  $f : \mathbb{N} \rightarrow \mathbb{R}$  durch  $f(n) := x(n)$ .

- Es sei  $a \in \mathbb{R}$ . Die Funktion, die für jedes  $x \in \mathbb{R}$  den Wert  $a$  annimmt, bezeichnen wir mit  $\underline{a} : \mathbb{R} \rightarrow \mathbb{R}$ , d.h.  $\underline{a}(x) = a$  für alle  $x \in \mathbb{R}$ . Ihr Graph in der Koordinatenebene ist eine Parallele zur  $x$ -Achse durch den Punkt  $(0, a)$ .

Die konstanten Funktionen sind sowohl monoton fallend als auch monoton steigend.

- Die Identität  $\text{id}$  auf  $\mathbb{R}$  ist die Funktion mit

$$\text{id}(x) = x$$

Ihr Graph ist die Winkelhalbierende in der Koordinatenebene. Die Identität ist streng monoton steigend.

**Definition 1.3.4.** Es sei  $U$  eine Menge mit  $a - x \in U \implies a + x \in U$  für ein  $a \in \mathbb{R}$ . Eine Funktion  $f : U \rightarrow \mathbb{R}$  heißt

- *achsensymmetrisch zur Achse  $x = a$* , wenn  $f(a - x) = f(a + x)$  für alle  $x \in U$ , und
- *punktsymmetrisch zum Punkt  $(a, b)$* , wenn  $f(a + x) - b = b - f(a - x)$  für alle  $x \in U$ .
- Eine Funktion die achsensymmetrisch zur Achse  $x = 0$  ist, also  $f(x) = f(-x)$ , heißt *gerade*. Eine Funktion, die punktsymmetrisch zum Ursprung ist, also  $f(x) = -f(-x)$ , heißt *ungerade*.

**Beispiel 1.3.3 (Fortsetzung).** Wie man an den bisherigen Beispielen schon sieht, werden Funktionen in der Regel beschrieben, indem man eine *Funktionsvorschrift* angibt. Wir schreiben das dann auch in der Form

$$f : x \mapsto f(x).$$

- Für  $n \in \mathbb{N}^*$  ist die  $n$ -te Potenzfunktion die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f(x) = x^n$$

Sie hat genau eine Nullstelle bei  $x = 0$ . Sie ist für gerades  $n$  auf  $\mathbb{R}^{\leq 0}$  streng monoton fallend und auf  $\mathbb{R}^{\geq 0}$  streng monoton steigend. Für ungerades  $n$  ist die Potenzfunktion streng monoton steigend auf ihrem gesamten Definitionsbereich.

- Als *Normalhyperbel* bezeichnen wir die Funktion  $f : \mathbb{R}^* \rightarrow \mathbb{R}$  mit

$$f : x \mapsto \frac{1}{x}.$$

Diese Funktion ist streng monoton fallend.

- Die Betragsfunktion ist die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit

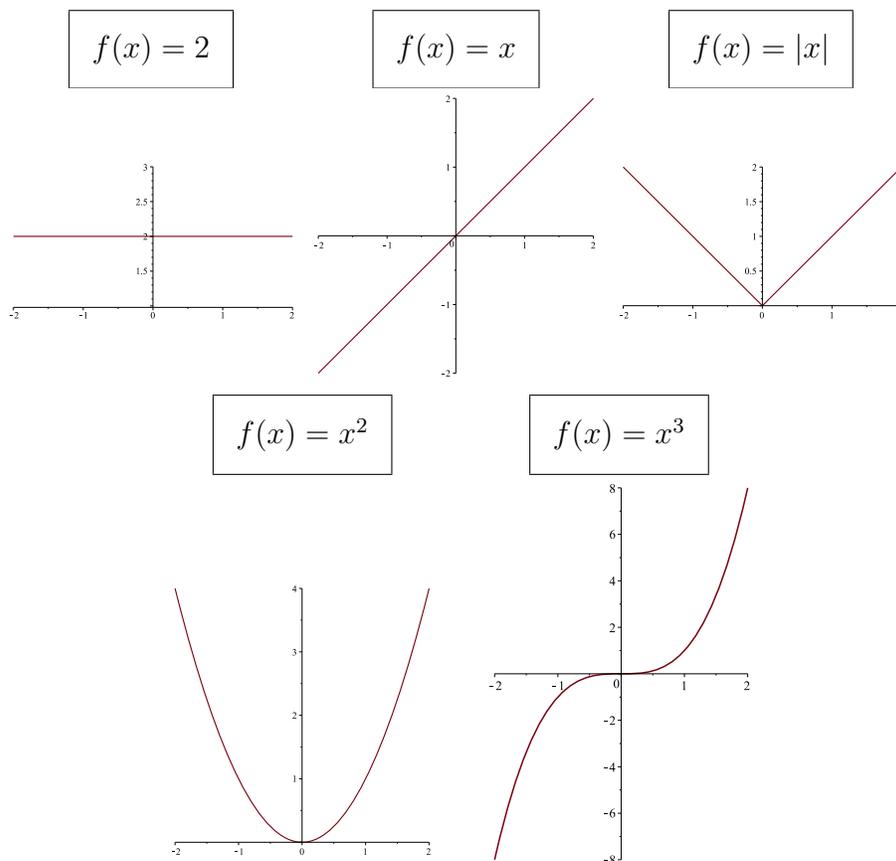
$$f : x \mapsto |x| = \begin{cases} x & \text{für } x \geq 0 \\ -x & \text{für } x < 0 \end{cases}$$

**Bemerkung 1.3.5.** Es seien  $U, V \subset \mathbb{R}$  und  $f : U \rightarrow V$  eine Funktion.

1.  $f : U \rightarrow V$  ist injektiv  $\iff$  der Graph von  $f$  schneidet für alle  $a \in V$  den Graphen der Funktion  $\underline{a}$  höchstens einmal.
2.  $f : U \rightarrow V$  ist surjektiv  $\iff$  der Graph von  $f$  schneidet für alle  $a \in V$  den Graphen der Funktion  $\underline{a}$  mindestens einmal.
3.  $f : U \rightarrow V$  ist bijektiv  $\iff$  der Graph von  $f$  schneidet für alle  $a \in V$  den Graphen der Funktion  $\underline{a}$  genau einmal.

Man kann Funktionen auf verschiedene Arten verknüpfen.

Abbildung 1.3.1: Einige Funktionen und ihre Graphen



**Definition 1.3.6.** 1. Es seien  $f : U \rightarrow \mathbb{R}$  und  $g : U \rightarrow \mathbb{R}$  Funktionen mit dem selben Definitionsbereich. Dann sind die Funktionen  $f + g$ ,  $f - g$ ,  $f \cdot g$  und  $\frac{f}{g}$  über die folgenden Funktionsvorschriften definiert

$$f + g : U \rightarrow \mathbb{R} \quad \text{mit} \quad (f + g)(x) := f(x) + g(x)$$

$$f - g : U \rightarrow \mathbb{R} \quad \text{mit} \quad (f - g)(x) := f(x) - g(x)$$

$$f \cdot g : U \rightarrow \mathbb{R} \quad \text{mit} \quad (f \cdot g)(x) := f(x)g(x)$$

$$\frac{f}{g} : U \setminus \mathcal{N}_g \rightarrow \mathbb{R} \quad \text{mit} \quad \left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}$$

2. Es seien  $f : U \rightarrow \mathbb{R}$  und  $g : V \rightarrow \mathbb{R}$  Funktionen mit  $f(U) \subset V$ . Dann

ist die Verkettung  $g \circ f : U \rightarrow \mathbb{R}$  ebenfalls eine Funktion und es gilt

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

**Beispiel 1.3.7.** Es sei  $f : U \rightarrow \mathbb{R}$  eine Funktion und  $b : \mathbb{R} \rightarrow \mathbb{R}$  mit  $b(x) = |x|$  die Betragsfunktion. Die Verkettung liefert eine neue Funktion, die wir mit  $|f| : U \rightarrow \mathbb{R}$  bezeichnen. Es gilt

$$|f| = b \circ f \quad \text{also} \quad |f|(x) = |f(x)|.$$

**Definition 1.3.8.** Eine Funktion  $f : U \rightarrow \mathbb{R}$  heißt

- nach unten beschränkt, wenn es eine Zahl  $C \in \mathbb{R}$  gibt, sodass  $f(x) \geq C$  für alle  $x \in U$ ,
- nach oben beschränkt, wenn es eine Zahl  $C \in \mathbb{R}$  gibt, sodass  $f(x) \leq C$  für alle  $x \in U$ ,
- beschränkt, wenn es eine Zahl  $C \in \mathbb{R}$  gibt, sodass  $|f(x)| < C$  für alle  $x \in U$ .

Insbesondere ist eine Funktion genau dann beschränkt, wenn sie nach oben und nach unten beschränkt ist.

**Bezeichnung 1.3.9.** Es seien  $f, g : U \rightarrow \mathbb{R}$  zwei Funktionen. Dann schreiben wir  $f < g$ , wenn  $f(x) < g(x)$  für alle  $x \in U$ . Ist  $f < \underline{a}$  für ein  $a \in \mathbb{R}$  so schreiben wir auch kürzer  $f < a$ . Genauso gilt das auch für  $\leq, >$  und  $\geq$ .

Eine Funktion ist in dieser Notation nach oben beschränkt, nach unten beschränkt oder beschränkt, wenn es eine Zahl  $C \in \mathbb{R}$  gibt, sodass  $f \leq C$ ,  $f \geq C$  oder  $|f| \leq C$ .

**Satz 1.3.10.** *Es sei  $f : U \rightarrow \mathbb{R}$  eine streng monoton Funktion. Dann gilt:*

1.  $f$  ist bijektiv, wenn man den Wertebereich auf  $f(U) \subset \mathbb{R}$  einschränkt.
2. Die Umkehrfunktion  $f^{-1} : f(U) \rightarrow \mathbb{R}$  hat das gleiche Monotonieverhalten wie  $f$ .

**Beispiel 1.3.11.** Es seien  $a \in \mathbb{R}^*, b \in \mathbb{R}$ . Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = ax + b$  ist für  $a > 0$  streng monoton steigend und für  $a < 0$  streng monoton fallend. Die Umkehrfunktion zu  $f$  ist die Abbildung  $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f^{-1}(x) = \frac{1}{a}x - \frac{b}{a}$  und hat das gleiche Monotonieverhalten.

**Bemerkung 1.3.12.** Ist  $f : U \rightarrow \mathbb{R}$  eine invertierbare reelle Funktion und  $f^{-1} : f(U) \rightarrow \mathbb{R}$  ihre Umkehrfunktion, so gilt  $G_f = \{(x, f(x)) \mid x \in U\}$  und  $G_{f^{-1}} = \{(y, f^{-1}(y)) \mid y \in f(U)\} = \{(f(x), x) \mid x \in U\}$ . Deshalb ergeben sich die Graphen jeweils aus dem anderen durch Spiegelung an der Winkelhalbierenden, also an  $G_{\text{id}}$ .

### 1.3.2 Wurzelfunktionen

Die  $n$ -te Potenzfunktion  $f : x \mapsto x^n$  ist für gerade  $n$  auf  $\mathbb{R}^{\geq 0}$  und für ungerades  $n$  auf  $\mathbb{R}$  streng monoton steigend, siehe Beispiel 1.3.3. Damit ist die Potenzfunktion wegen Satz 1.3.10 invertierbar.

Wie sieht jedoch das Bild der Potenzfunktion aus - und damit der Definitionsbereich der Umkehrfunktion? Die Antwort liefert der folgende Satz:

- Satz 1.3.13.**
1. Ist  $n \in \mathbb{N}^*$ , dann gibt es zu jeder Zahl  $b \geq 0$  genau ein Zahl  $a > 0$  mit  $a^n = b$ .
  2. Ist  $n \in \mathbb{N}^*$  gerade, dann gibt es zu jeder Zahl  $b > 0$  - und nur zu diesen - genau zwei Zahlen  $a_1 < 0 < a_2$  mit  $a_1^n = a_2^n = b$ .
  3. Ist  $n \in \mathbb{N}$  ungerade, dann gibt es zu jeder Zahl  $b \in \mathbb{R}$  genau eine Zahl  $a \in \mathbb{R}$  mit  $a^n = b$ .

**Beweisskizze.** Zum Beweis der Existenz von  $a$  in 1. für  $n \geq 2$  definiere  $M := \{x \in \mathbb{R} \mid x^n < b\}$  und setze  $a := \sup(M)$ . Dann zeige  $a^n = b$  durch Widerspruch der Annahmen  $a^n < b$  und  $a^n > b$ .

Sei  $a^n > b$ . Es ist  $a(1 - \epsilon) \in M$  für alle  $0 < \epsilon < 1$ , da sonst  $a$  nicht die kleinste obere Schranke für  $M$  wäre. Damit ist  $a^n(1 - \epsilon)^n < b$  und weiter  $b > a^n(1 - \epsilon)^n > a^n(1 - n\epsilon)$ . Das gilt insbesondere für  $0 < \epsilon = \frac{a^n - b}{na^n} < \frac{1}{n} < 1$  und damit folgt  $b > b$ .

Sei  $a^n < b$ . Es ist  $\frac{a}{1 - \epsilon} > a$  für alle  $0 < \epsilon < 1$ , und daher  $\frac{a}{1 - \epsilon} \notin M$ . Damit ist  $\frac{a^n}{(1 - \epsilon)^n} \geq b$  und weiter  $a^n \geq b(1 - \epsilon)^n > b(1 - n\epsilon)$ . Das gilt insbesondere für  $0 < \epsilon = \frac{b - a^n}{nb} < \frac{1}{n} < 1$  und damit folgt  $a^n > a^n$ .  $\square$

**Folgerung 1.3.14.** Schränkt man den Definitionsbereich der Potenzfunktion  $x \mapsto x^n$  geeignet ein, so ist sie bijektiv auf das Bild:

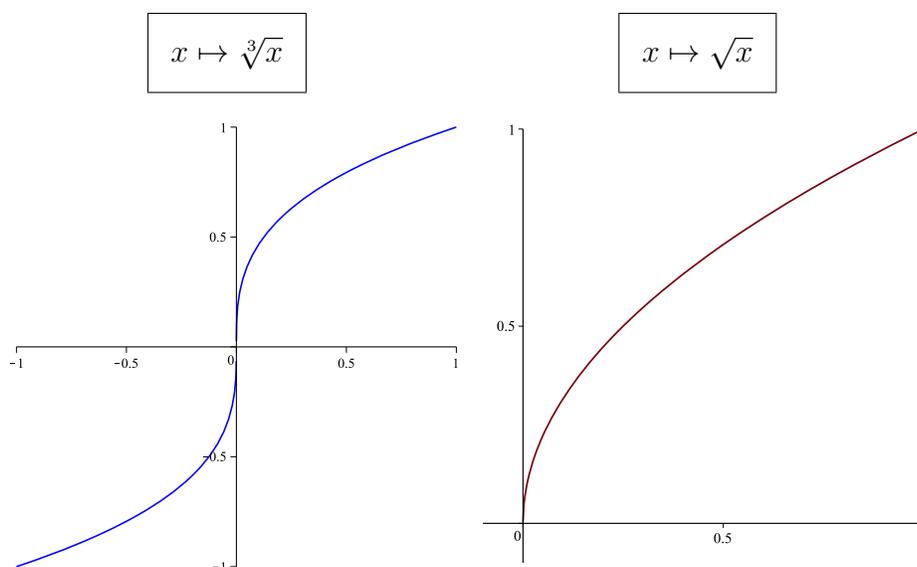
$$\begin{aligned} \mathbb{R}^{\geq 0} \ni x \mapsto x^n \in \mathbb{R}^{\geq 0}, & \quad \text{falls } n \text{ gerade,} \\ \mathbb{R} \ni x \mapsto x^n \in \mathbb{R}, & \quad \text{falls } n \text{ ungerade.} \end{aligned}$$

**Definition 1.3.15.** Die Umkehrfunktion der (eingeschränkten) Potenzfunktion heißt *n-te Wurzelfunktion* und wir schreiben

$$x \mapsto \sqrt[n]{x}.$$

Für ungerades  $n$  ist der Definitions- und der Wertebereich  $\mathbb{R}$ , für  $n$  gerade ist der Definitions- und der Wertebereich  $\mathbb{R}^{\geq 0}$ , siehe Abbildung 1.3.2.

Abbildung 1.3.2: Wurzelfunktionen



**Definition/Bemerkung 1.3.16** (allgemeine Potenzfunktion 1). Wir führen die *allgemeine Potenzfunktion* für rationale Potenzen ein, indem wir sie zunächst für die Potenz Null, für Stammbruchpotenzen und negative ganze Potenzen definieren:

- Für  $x \neq 0$  ist  $x^n = x^{n+0} = x^n x^0$  also setzen wir

$$x^0 = 1.$$

- Wir schreiben für alle  $x \geq 0$  wegen  $(\sqrt[n]{x})^n = \sqrt[n]{x^n} = x$  für  $n > 0$

$$x^{\frac{1}{n}} := \sqrt[n]{x}.$$

- Wegen  $1 = x^0 = x^{n-n}$  für alle  $x \neq 0$  definieren wir für eine negative ganze Zahl  $-n$

$$x^{-n} := \frac{1}{x^n}.$$

Ist  $r = \frac{m}{n} \in \mathbb{Q}$  eine rationale Zahl so definieren wir mit diesen Vorüberlegungen durch

$$x \mapsto x^r := \sqrt[n]{x^m}$$

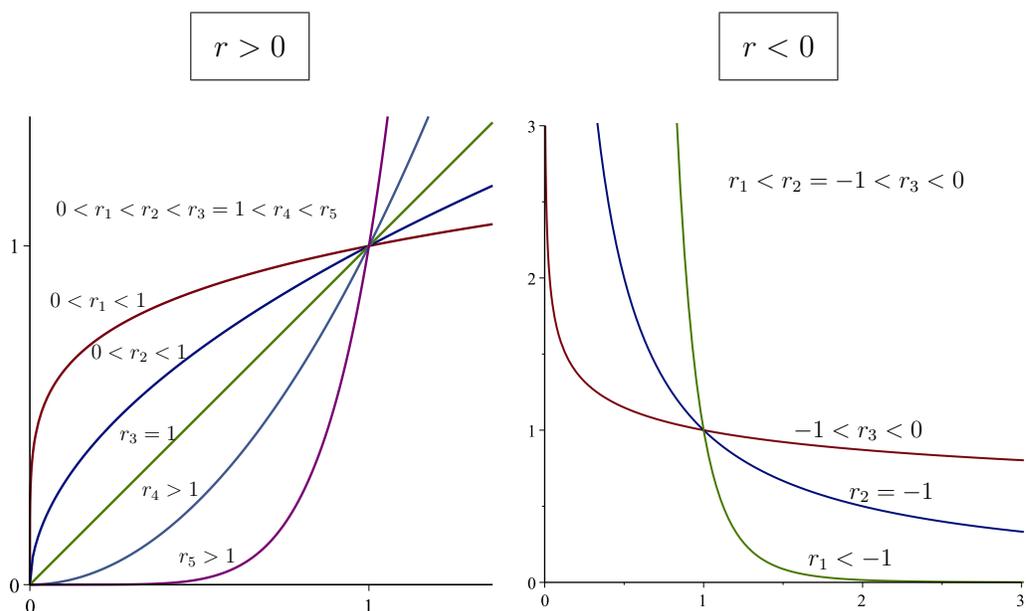
eine Funktion mit Definitionsbereich  $\mathbb{R}^{\geq 0}$  falls  $r > 0$  bzw.  $\mathbb{R}^{> 0}$  falls  $r \leq 0$ , siehe Abbildung 1.3.3.

Die so definierte allgemeine Potenzfunktion erfüllt die gängigen Potenzgesetze:

$$x^r y^r = (xy)^r, \quad x^r x^s = x^{r+s}, \quad (x^r)^s = x^{rs}.$$

für alle  $x, y$  im jeweiligen Definitionsbereich.

Abbildung 1.3.3: Die allgemeine Potenzfunktion  $x \mapsto x^r$



## 1.3.3 Polynome und rationale Funktionen

**Definition 1.3.17.** Ein Polynom ist eine Funktion  $p : \mathbb{R} \rightarrow \mathbb{R}$ , die sich als Summe und Produkten von Potenzfunktionen und konstanten Funktionen ergibt. Sie hat die Form

$$p(x) = \sum_{k=0}^n a_k x^k = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0.$$

Ist  $a_n \neq 0$ , dann heißt die Zahl  $n$  der *Grad* des Polynoms und die  $a_i \in \mathbb{R}$  heißen die *Koeffizienten*. Speziell der höchste Koeffizient  $a_n$  heißt *Leitkoeffizient*.

**Satz 1.3.18.** *Ein Polynom  $p$  vom Grad  $\geq 1$  ist unbeschränkt. Es gilt:*

- *Ist der Grad von  $p$  ungerade, so ist  $p$  nach unten und oben unbeschränkt.*
- *Ist der Grad von  $p$  gerade, so ist  $p$  nach unten beschränkt und nach oben unbeschränkt, wenn der Leitkoeffizient positiv ist, und nach oben beschränkt und nach unten unbeschränkt, wenn der Leitkoeffizient negativ ist.*

Zu diesem Funktionsverhalten siehe auch Abbildung 1.3.4.

**Beweisskizze.** Es sei  $p(x) = \sum_{k=0}^n a_k x^k$ . Dann gilt mit  $C := \frac{2}{|a_n|} \sum_{k=0}^{n-1} |a_k|$  für alle  $x$  mit  $|x| > C$

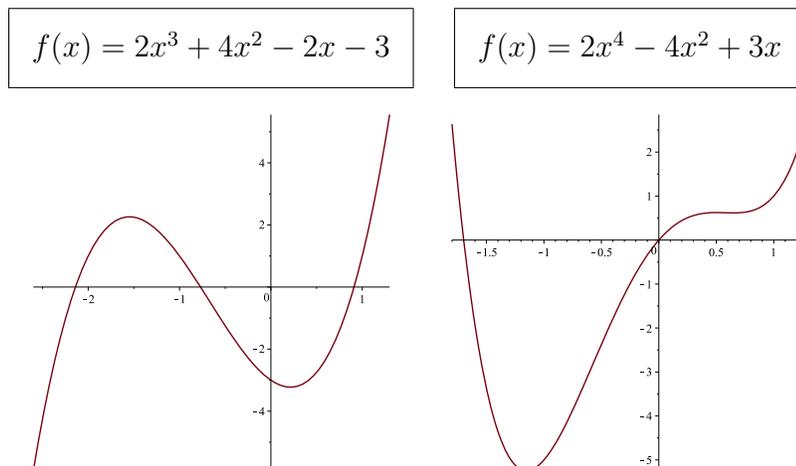
$$a_n x^n - \frac{1}{2} |a_n| |x|^n \leq p(x) \leq a_n x^n + \frac{1}{2} |a_n| |x|^n.$$

Eine Fallunterscheidung für  $a_n$  positiv/negativ und  $n$  gerade/ungerade liefert zusammen mit  $|p(x)| \leq \sum_{k=0}^n |a_k| C^k$  für alle  $x$  mit  $|x| \leq C$  das Resultat.  $\square$

**Bemerkung 1.3.19.** • Ein Polynom  $p(x) = a_1 x + a_0$  vom Grad 1 hat genau eine Nullstelle, nämlich  $x_0 = -\frac{a_0}{a_1}$ .

- Ein Polynom  $p(x) = a_2 x^2 + a_1 x + a_0$  vom Grad 2 hat keine Nullstelle, wenn  $\Delta := \frac{a_1^2}{4} - a_0 a_2 < 0$ , zwei Nullstellen, wenn  $\Delta > 0$  und eine

Abbildung 1.3.4: Polynome



Nullstelle, wenn  $\Delta = 0$ . Es gilt nämlich

$$\begin{aligned} p(x) &= \frac{1}{a_2} \left( a_2^2 x^2 + a_1 a_2 x + \left( \frac{a_1}{2} \right)^2 - \left( \frac{a_1}{2} \right)^2 + a_0 a_2 \right) \\ &= \frac{1}{a_2} \left( a_2 x + \frac{a_1}{2} \right)^2 - \frac{\Delta}{a_2}. \end{aligned}$$

**Satz 1.3.20.** *Es sei  $p$  ein Polynom  $n$ -ten Grades mit einer Nullstelle  $x = x_0$ . Dann gibt es ein Polynom  $\hat{p}$  vom Grad  $n - 1$ , sodass für alle  $x \in \mathbb{R}$*

$$p(x) = (x - x_0)\hat{p}(x).$$

**Beweisskizze.** Wir zeigen eine etwas allgemeinere Aussage: Für alle  $x_0 \in \mathbb{R}$  existiert ein Polynom  $\hat{p}$  wie angegeben, sodass  $p(x) = p(x_0) + (x - x_0)\hat{p}(x)$  für alle  $x \in \mathbb{R}$ .  $\square$

**Folgerung 1.3.21.** Es sei  $p$  ein Polynom  $n$ -ten Grades mit den Nullstellen  $x_1, x_2, \dots, x_m$ . Dann gibt es Zahlen  $n_1, \dots, n_r \in \mathbb{N}^*$  und ein Polynom  $\hat{p}$  vom Grad  $n - \ell$ , wobei  $\ell = n_1 + \dots + n_r$ , mit den folgenden Eigenschaften

- $\hat{p}$  hat keine Nullstellen
- für alle  $x \in \mathbb{R}$  gilt  $p(x) = (x - x_1)^{n_1} (x - x_2)^{n_2} \cdot \dots \cdot (x - x_r)^{n_r} \hat{p}(x)$ .

Die natürliche Zahl  $n_k$  nennt man die *Vielfachheit* der Nullstelle  $x_k$

**Folgerung 1.3.22** (Fundamentalsatz der Algebra). Ein Polynom  $n$ -ten Grades, das nicht das Nullpolynom ist, hat höchstens  $n$  Nullstellen.

**Folgerung 1.3.23** (Identitätssatz für Polynome). Haben zwei Polynome  $n$ -ten Grades  $n + 1$  Schnittpunkte, so sind sie gleich.

**Bemerkung 1.3.24** (Koeffizientenvergleich). Die letzte Folgerung sagt uns, dass zwei Polynome  $p, q$  genau dann die gleiche Funktion liefern, wenn ihre Grade und die Koeffizienten gemäß Definition 1.3.17 übereinstimmen.

**Folgerung 1.3.25.** 1. Es sei  $p$  ein Polynom  $n$ -ten Grades und  $x_0 \in \mathbb{R}$ . Dann gibt es eindeutige Zahlen  $b_0, \dots, b_n \in \mathbb{R}$ , sodass für alle  $x \in \mathbb{R}$

$$p(x) = \sum_{i=0}^n b_i (x - x_0)^i.$$

2. Es sei  $q$  ein Polynom  $n$ -ten Grades,  $x_0 \in \mathbb{R}$  mit  $q(x_0) \neq 0$  und  $m \in \mathbb{N}^*$ . Dann gibt es Polynome  $Q, R$  mit  $\text{grad}(Q) < m$ ,  $\text{grad}(R) < n$ , sodass für alle  $x \in \mathbb{R}$

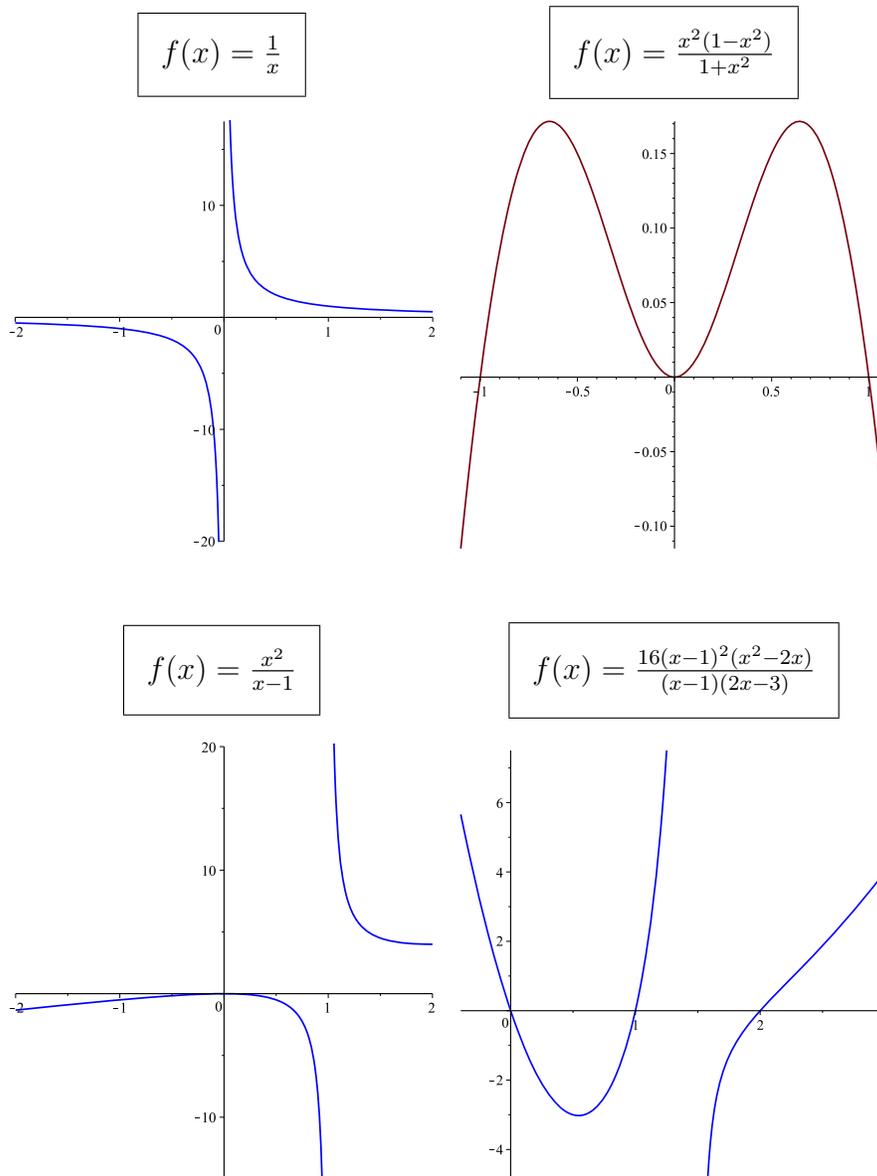
$$Q(x)q(x) + R(x)x^m = 1.$$

**Definition 1.3.26.** Es seien  $p, q : \mathbb{R} \rightarrow \mathbb{R}$  zwei Polynome. Die Funktion  $\frac{p}{q} : \mathbb{R} \setminus \mathcal{N}_q \rightarrow \mathbb{R}$  heißt *rationale Funktion*. Die Menge  $\mathcal{N}_q$  nennt man auch die Menge der *Definitionslücken* der rationalen Funktion  $\frac{p}{q}$ , siehe Abbildung 1.3.5.

**Bemerkung 1.3.27.** Die Zähler  $p$  und Nenner  $q$  einer rationalen Funktion  $f = \frac{p}{q}$  haben eine gemeinsame Nullstelle  $x = x_0$  mit der Vielfachheit  $r$  bzw  $s$ , d.h. es gibt Polynome  $\tilde{p}$  und  $\tilde{q}$  mit  $\tilde{q}(x_0)\tilde{p}(x_0) \neq 0$  und  $p(x) = (x - x_0)^r \tilde{p}(x)$  sowie  $q(x) = (x - x_0)^s \tilde{q}(x)$ . Dann gilt:

- Ist  $r \geq s$ , so hat die gekürzte rationale Funktion  $\tilde{f}$  mit  $\tilde{f}(x) = (x - x_0)^{r-s} \frac{\tilde{p}(x)}{\tilde{q}(x)}$  keine Definitionslücke in  $x = x_0$  und der Verlauf der Graphen von  $f$  und  $\tilde{f}$  stimmen an allen Stellen  $x \neq x_0$  überein. In diesem Fall spricht man bei  $x_0$  von einer *hebbaren Definitionslücke* von  $f = \frac{p}{q}$ .
- Ist  $r < s$  so hat  $\tilde{f}$  mit  $\tilde{f}(x) = \frac{\tilde{p}(x)}{(x-x_0)^{s-r}\tilde{q}(x)}$  die gleichen Definitionslücken wie  $\frac{p}{q}$  und ihre Graphen stimmen überall überein. In diesem Fall nennt man  $x_0$  einen *Pol* von  $f = \frac{p}{q}$  der *Ordnung*  $s - r$ .

Abbildung 1.3.5: Rationale Funktionen



**Satz 1.3.28** (Partialbruchzerlegung). *Das Polynom  $q$  vom Grad  $n$  besitze die Darstellung  $q(x) = (x-x_1)^{n_1}(x-x_2)^{n_2} \cdots (x-x_r)^{n_r} \hat{q}(x)$  wobei  $\hat{q}$  nullstellenfrei ist. Weiter sei  $p$  ein Polynom vom Grad  $m$ , dass keine gemeinsame Nullstelle mit  $q$  hat. Dann gibt es für  $1 \leq i \leq r$ ,  $1 \leq j_i \leq n_i$  Zahlen  $b_{ij_i}$  sowie Polynome*

$p_0$  vom Grad  $m - n$  und  $R_0$  mit  $\text{grad}(R_0) < \text{grad}(\hat{q})$ , sodass für alle  $x \in \mathbb{R}$

$$\frac{p(x)}{q(x)} = p_0(x) + \sum_{i=1}^r \sum_{j_i=1}^{n_i} \frac{b_{ij_i}}{(x - x_i)^{j_i}} + \frac{R_0(x)}{\hat{q}(x)}.$$

Ist  $m < n$ , so ist  $p_0 = \underline{0}$ .

**Beweisskizze.** Das Polynom  $p_0$  erhalten wir mittels Polynomdivision aus  $\frac{p(x)}{q(x)} = p_0(x) + \frac{\tilde{p}(x)}{q(x)}$  und  $\text{grad}(\tilde{p}) < \text{grad}(q)$ . Zeige mit Hilfe von Folgerung 1.3.25.2 zunächst, dass es ein Polynom  $R_0$  wie oben und Polynome  $P_1, \dots, P_r$  mit  $\text{grad}(P_i) < n_i$  gibt, sodass

$$\frac{p(x)}{q(x)} = p_0(x) + \frac{P_1(x)}{(x - x_1)^{n_1}} + \dots + \frac{P_r(x)}{(x - x_r)^{n_r}} + \frac{R_0(x)}{\hat{q}(x)}.$$

Wende dann Folgerung 1.3.25.1 an. □

**Bemerkung/Beispiel 1.3.29.** Wenn man eine explizite Zerlegung zu machen hat, dann wird man nicht den Weg des Beweises gehen, sondern einen direkten Ansatz machen und Koeffizientenvergleich durchführen.

Wir betrachten als Beispiel die Polynome  $p(x) = x^6$  und  $q(x) = x^5 - x^4 - x + 1$ . Man sieht, dass  $q$  die Nullstelle  $x = 1$  hat. Hier liefert Polynomdivision  $q(x) = (x - 1)(x^4 - 1)$  und die Binomischen Formeln weiter  $q(x) = (x - 1)(x^2 - 1)(x^2 + 1) = (x - 1)(x - 1)(x + 1)(x^2 + 1)$ . Das heißt, es ist  $\hat{q}(x) = x^2 + 1$  und insgesamt

$$q(x) = (x - 1)^2(x + 1)(x^2 + 1).$$

Polynomdivision liefert zunächst  $p_0(x) = x + 1$ :

$$\frac{p(x)}{q(x)} = (x + 1) + \frac{x^4 + x^2 - 1}{(x - 1)^2(x + 1)(x^2 + 1)}.$$

Der Satz 1.3.28 sagt uns nun, dass die Restfunktion in der Form

$$\frac{x^4 + x^2 - 1}{(x - 1)^2(x + 1)(x^2 + 1)} = \frac{b_{11}}{(x - 1)} + \frac{b_{12}}{(x - 1)^2} + \frac{b_{21}}{(x + 1)} + \frac{Ax + B}{x^2 + 1}$$

geschrieben werden kann. Multiplizieren wir das mit dem Hauptnenner, nämlich  $q(x)$ , so gibt das

$$\begin{aligned} x^4 + x^2 - 1 &= b_{11}(x - 1)(x + 1)(x^2 + 1) + b_{12}(x + 1)(x^2 + 1) \\ &\quad + b_{21}(x - 1)^2(x^2 + 1) + (Ax + B)(x - 1)^2(x + 1). \end{aligned}$$

Durch einsetzen von  $x = 1$  erhält man sofort  $1 = 4b_{12}$ , also  $b_{12} = \frac{1}{4}$  und damit

$$\begin{aligned} x^4 + x^2 - 1 &= x^4(A + b_{11} + b_{21}) + x^3(-A + B - 2b_{21} + \frac{1}{4}) \\ &\quad + x^2(-A - B + 2b_{21} + \frac{1}{4}) + x(A - B - 2b_{21} + \frac{1}{4}) \\ &\quad + (B - b_{11} + b_{21} + \frac{1}{4}). \end{aligned}$$

Durch Koeffizientenvergleich erhalten wir das LGS

$$\left. \begin{array}{r} A \quad +b_{11} \quad +b_{21} = 1 \\ -A \quad +B \quad \quad -2b_{21} = -\frac{1}{4} \\ -A \quad -B \quad \quad +2b_{21} = \frac{3}{4} \\ A \quad -B \quad \quad -2b_{21} = -\frac{1}{4} \\ \quad \quad B \quad -b_{11} \quad +b_{21} = -\frac{5}{4} \end{array} \right\} \iff \left\{ \begin{array}{l} A = -\frac{1}{4} \\ B = -\frac{1}{4} \\ b_{11} = \frac{9}{8} \\ b_{21} = \frac{1}{8} \end{array} \right.$$

Damit ergibt sich die Zerlegung

$$\frac{x^6}{x^5 - x^4 - x + 1} = x + 1 + \frac{9}{8(x-1)} + \frac{1}{4(x-1)^2} + \frac{1}{8(x+1)} - \frac{x+1}{4(x^2+1)}$$

### 1.3.4 Die allgemeine Exponentialfunktion und Logarithmusfunktion

**Satz 1.3.30.** Für alle reellen Zahlen  $a \in \mathbb{R}^{>0}$  gibt es genau eine Funktion  $E_a : \mathbb{R} \rightarrow \mathbb{R}$  mit den Eigenschaften

- (i)  $E_a(1) = a$ ,
- (ii)  $E_a(x+y) = E_a(x)E_a(y)$  für alle  $x, y \in \mathbb{R}$ ,
- (iii)  $E_a$  ist monoton.

Für diese Funktion gilt dann:  $E_a$  ist streng monoton steigend/fallend, wenn  $a > 1/a < 1$  und das Bild ist  $E_a(\mathbb{R}) = \mathbb{R}^+$ . Ist  $a = 1$  so ist  $E_1 = \underline{1}$  konstant.

**Beweisskizze.** Ein Beweis von Satz 1.3.30 verläuft in mehreren Schritten.

- Zeige zunächst mit (i) und (ii), dass  $E_a(q) = a^q$  für alle  $q \in \mathbb{Q}^{\geq 0}$ . Insbesondere ist die Funktion  $E_a$  dann monoton steigend, falls  $a \geq 1$  und sie ist monoton fallend falls  $a < 1$ .

Wir betrachten zunächst den Fall  $a > 1$ .

- Definiere für  $x \in \mathbb{R}^+$  die Mengen  $M_-^x := \{a^r \mid r \in \mathbb{Q}^+, r \leq x\}$  und  $M_+^x := \{a^r \mid r \in \mathbb{Q}^+, r \geq x\}$ . Wegen der Monotonie (iii) gilt dann  $a^{r-} \leq E_a(x) \leq a^{r+}$  für alle  $a^{r\pm} \in M_{\pm}^x$ . Zeige, dass es kein Intervall  $I$  mit positiver Länge gibt, für das  $M_-^x \leq I \leq M_+^x$  gilt.
- Definiere damit eindeutig(!)  $E_a(x) = \sup(M_-^x) = \inf(M_+^x)$ .
- Zeige, dass die so definierte Funktion  $E_a : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}$  streng monoton steigt, also (iii).
- Zeige, dass die so definierte Funktion  $E_a : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}$  die Funktionalgleichung (ii) erfüllt.
- Erweitere  $E_a$  auf die negativen Zahlen und zeige, dass das Resultat weiter (ii) und (iii) erfüllt.

Betrachte nun analog den Fall  $a < 1$ . □

**Satz 1.3.31.** *Die Funktion  $E_a$  aus Satz 1.3.30 erfüllt neben der Funktionalgleichung (ii) noch die folgende wichtige Identität:*

$$E_a(xy) = E_{E_a(x)}(y) \quad \text{für alle } x, y \in \mathbb{R}.$$

**Definition 1.3.32** (Exponentialfunktion). Die Funktion  $E_a : \mathbb{R} \rightarrow \mathbb{R}^+$  heißt *Exponentialfunktion zur Basis  $a$*  und wir schreiben auch für irrationale  $x \in \mathbb{R}$

$$E_a(x) = a^x.$$

Da die Exponentialfunktion zur Basis  $a \neq 1$  streng monoton ist, existiert ihre Umkehrfunktion.

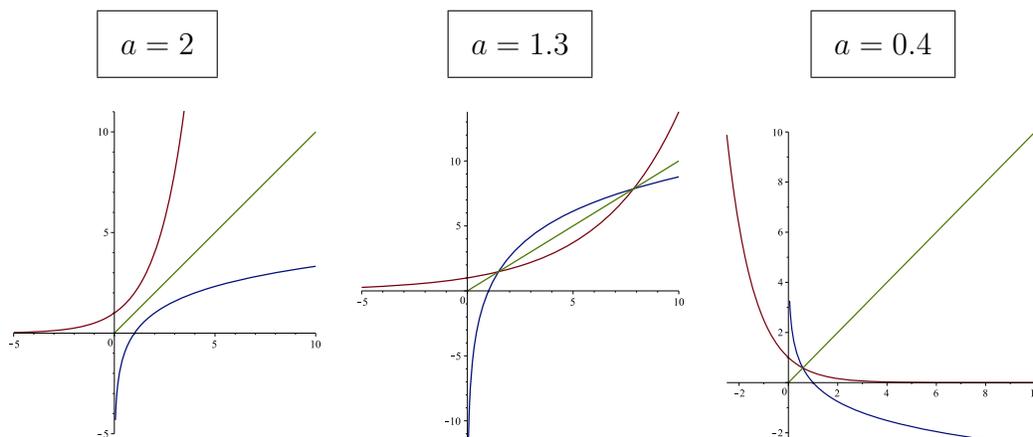
**Definition 1.3.33** (Logarithmusfunktion). Die Umkehrfunktion der Exponentialfunktion zur Basis  $a > 0, a \neq 1$  heißt *Logarithmusfunktion zur Basis  $a$* . Wir schreiben dafür

$$\log_a : \mathbb{R}^+ \rightarrow \mathbb{R}$$

mit

$$\log_a(x) = y \iff x = a^y.$$

Abbildung 1.3.6: Die Graphen der Exponentialfunktion  $x \mapsto a^x$  (rot) und der Logarithmusfunktion  $x \mapsto \log_a(x)$  (blau)



Die folgenden Eigenschaften der Logarithmusfunktion beweist man, indem man sie auf die Eigenschaften der Exponentialfunktion zurückführt.

**Satz 1.3.34** (Logarithmengesetze). *Es gelten die folgenden Rechenregeln beim Umgang mit dem Logarithmus. Für alle  $x, y \in \mathbb{R}^{>0}$  ist*

1.  $\log_a(1) = 0$  und  $\log_a(a) = 1$
- 2.a.  $\log_a(xy) = \log_a(x) + \log_a(y)$
- 2.b.  $\log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y)$
3.  $\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$
4.  $\log_a(x^r) = r \log_a(x)$ ,  $r \in \mathbb{Q}$

Die Eigenschaft 4. nutzen wir zur Definition der allgemeinen Potenzfunktion für reelle Exponenten.

**Definition/Bemerkung 1.3.35** (allgemeine Potenzfunktion 2). Es sei  $r \in \mathbb{R}$ . Dann ist die Potenzfunktion

$$x \mapsto x^r$$

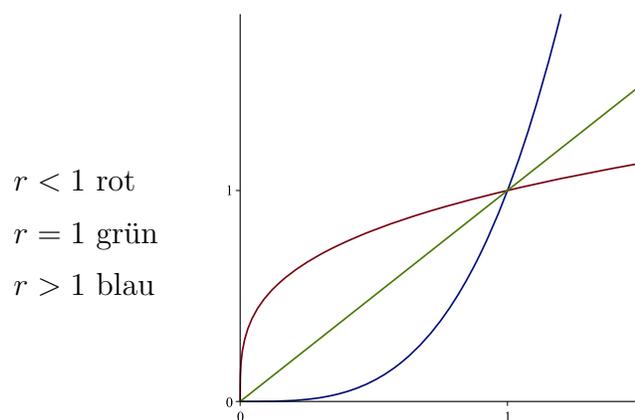
mit Definitionsbereich  $\mathbb{R}^+$  definiert durch

$$x^r := a^{r \log_a(x)}.$$

Diese Definition ist unabhängig von der Wahl einer beliebigen positiven Basis  $a \neq 1$ .

Die Graphen der Potenzfunktionen  $f(x) = x^r$  haben qualitativ die Gestalt aus Abbildung 1.3.7 (vergleiche auch mit Abb. 1.3.3).

Abbildung 1.3.7: Allgemeine Potenzfunktionen

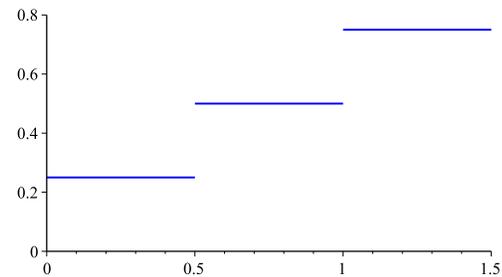


## 1.4 Stetige Funktionen und ihre Eigenschaften

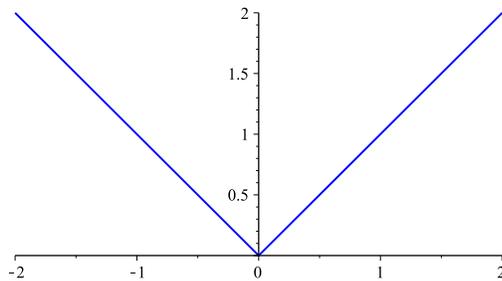
### 1.4.1 Stetigkeit

Abbildung 1.4.1: Einige Funktionsgraphen

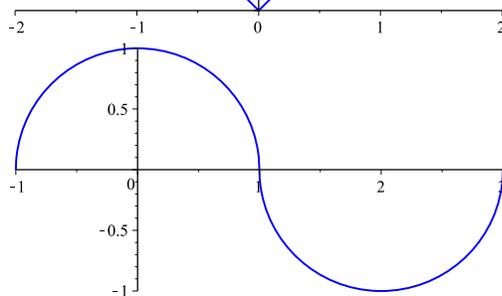
$$x \mapsto \begin{cases} \frac{1}{4}, & x \in [0, \frac{1}{2}] \\ \frac{1}{2}, & x \in ]\frac{1}{2}, 1] \\ \frac{3}{4}, & x \in ]1, \frac{3}{2}] \end{cases}$$



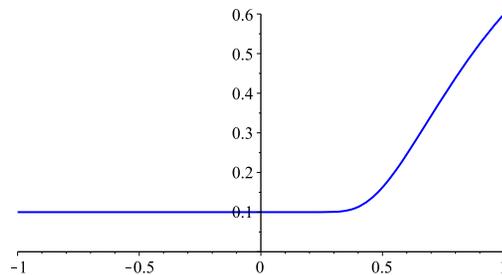
$$x \mapsto |x|, \quad x \in [-2, 2]$$



$$x \mapsto \begin{cases} \sqrt{1-x^2}, & x \in [-1, 1] \\ -\sqrt{1-(2-x)^2}, & x \in ]1, 3] \end{cases}$$



$$x \mapsto \begin{cases} \frac{1}{10}, & x \in [-1, 0] \\ \frac{1}{10} + 2^{-\frac{1}{x^2}}, & x \in ]0, 1] \end{cases}$$



Bei den Funktionen aus Abb. 1.4.1 sehen wir uns das verschiedene Verhalten an den "Verklebestellen" an. Dabei erkennen wir optisch keinen Unterschied

zwischen der dritten und vierten. Ein deutlicher Unterschied besteht jedoch zwischen der ersten einerseits und den weiteren drei andererseits: "Läuft man von beiden Seiten auf eine Verklebestelle zu", so gelangt man bei der zweiten bis vierten Funktion zum gleichen Wert, jedoch nicht bei der ersten.

**Definition 1.4.1** (Stetigkeit). Es sei  $f : M \rightarrow \mathbb{R}$  eine auf der Menge  $M \subset \mathbb{R}$  definierte Funktion.

1. Die Funktion  $f$  heißt *stetig im Punkt*  $x_0 \in M$ , wenn man für jede vorgegebene Zahl  $\epsilon > 0$  eine Zahl  $\delta > 0$  finden kann, sodass  $|f(x) - f(x_0)| < \epsilon$  für alle  $x \in M$  mit  $|x - x_0| < \delta$ .

Etwas formaler:

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in M : (|x - x_0| < \delta \implies |f(x) - f(x_0)| < \epsilon)$$

oder

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in M \cap B_\delta(x_0) : f(x) \in B_\epsilon(f(x_0))$$

2. Die Funktion  $f$  heißt *stetig auf*  $M$ , wenn  $f$  in jedem Punkt  $x_0 \in M$  stetig ist.

**Beispiel 1.4.2.** 1. Jede durch eine Zahlenfolge definierte Funktion ist stetig.

2. Die konstanten Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = a$  für ein festes  $a \in \mathbb{R}$  sind stetig.
3. Die Identität auf  $\mathbb{R}$ , also  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = x$ , ist stetig.
4. Die Quadratfunktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = x^2$  ist stetig.
5. Die Funktion  $f : \mathbb{R}^* \rightarrow \mathbb{R}$  mit  $f(x) = \frac{1}{x}$  ist stetig.

6. Es sei  $M \subset \mathbb{R}$  und  $\text{ind}_M : \mathbb{R} \rightarrow \mathbb{R}$  mit  $\text{ind}_M(x) = \begin{cases} 1, & x \in M \\ 0, & x \notin M \end{cases}$ .

$\text{ind}_M$  ist genau dann stetig auf  $\mathbb{R}$ , wenn  $M = \emptyset$  oder  $M = \mathbb{R}$ .<sup>(i)</sup>

<sup>(i)</sup>Die erste Funktion aus Abbildung 1.4.1 lässt sich als  $\frac{1}{4}\text{ind}_{[0, \frac{1}{2}]} + \frac{1}{2}\text{ind}_{[\frac{1}{2}, 1]} + \frac{3}{4}\text{ind}_{[1, \frac{3}{2}]}$  schreiben.

- Die Funktion  $\text{ind}_{\mathbb{Q}} : \mathbb{R} \rightarrow \mathbb{R}$  ist in keinem Punkt ihres Definitionsbereiches stetig.
- Ist  $I \subset \mathbb{R}$  ein Intervall, dann ist  $\text{ind}_I : \mathbb{R} \rightarrow \mathbb{R}$  in allen Punkten stetig mit Ausnahme der Randpunkte des Intervalls.

**Satz 1.4.3.** *Es seien  $M, N \subset \mathbb{R}$  und  $f : M \rightarrow \mathbb{R}$  und  $g : N \rightarrow \mathbb{R}$  Funktionen mit  $f(M) \subset N$ , sodass die Verkettung  $g \circ f : M \rightarrow \mathbb{R}$  definiert ist. Dann gilt:*

- Ist  $f$  stetig in  $x_0 \in M$  und  $g$  stetig in  $y_0 = f(x_0) \in N$ , so ist  $g \circ f$  stetig in  $x_0 \in M$ .
- Sind  $f$  und  $g$  stetig auf ihren Definitionsbereichen, so ist  $g \circ f : M \rightarrow \mathbb{R}$  stetig auf  $M$ .

**Satz 1.4.4.** *Es seien  $f : M \rightarrow \mathbb{R}$  und  $g : M \rightarrow \mathbb{R}$  Funktionen. Weiter seien  $f$  und  $g$  beide stetig in  $x_0 \in M$ . Dann gilt:*

1. Die Summe  $f + g : M \rightarrow \mathbb{R}$  und die Differenz  $f - g : M \rightarrow \mathbb{R}$  von  $f$  und  $g$  sind stetig in  $x_0 \in M$ .
2. Das Produkt  $fg : M \rightarrow \mathbb{R}$  ist stetig in  $x_0 \in M$ .
3. Ist  $g(x_0) \neq 0$  so ist der Quotient  $\frac{f}{g} : M \setminus \mathcal{N}_g \rightarrow \mathbb{R}$  stetig in  $x_0 \in M$ .

*Sind  $f, g : M \rightarrow \mathbb{R}$  stetig auf  $M$ , so gilt dies ebenfalls für  $f + g$ ,  $f - g$ ,  $fg$  und  $\frac{f}{g}$  auf ihren jeweiligen Definitionsbereichen.*

**Beispiel 1.4.5.** 1. Polynome und rationale Funktionen sind auf ihren Definitionsbereichen stetig.

2. Es seien  $f : [a, b] \rightarrow \mathbb{R}$  und  $g : [b, c] \rightarrow \mathbb{R}$  stetige Funktionen mit  $f(b) = g(b)$ . Dann ist die Funktion  $h : [a, c] \rightarrow \mathbb{R}$  mit

$$h(x) := \begin{cases} f(x), & \text{falls } x \in [a, b] \\ g(x), & \text{falls } x \in ]b, c] \end{cases}$$

stetig. Hierbei ist es unerheblich, ob  $a$  und  $c$  zum jeweiligen Intervall dazugehören, die Intervalle können an diesen Seiten auch unbeschränkt sein.

3. Es seien  $f, g : M \rightarrow \mathbb{R}$  stetige Funktionen auf der Teilmenge  $M \subset \mathbb{R}$ . Dann sind die Funktionen  $h_{\uparrow}, h_{\downarrow} : M \rightarrow \mathbb{R}$  mit

$$h_{\uparrow}(x) := \max\{f(x), g(x)\} \quad \text{und} \quad h_{\downarrow}(x) := \min\{f(x), g(x)\}$$

ebenfalls stetig.

**Satz 1.4.6.** *Es sei  $f : M \rightarrow \mathbb{R}$  eine monotone Funktion. Dann gilt*

1. *Ist das Bild  $f(M)$  ein Intervall, dann ist  $f$  stetig auf  $M$ .*
2. *Ist  $f$  sogar streng monoton und  $M$  selbst ein Intervall, dann ist die Umkehrfunktion  $f^{-1} : f(M) \rightarrow \mathbb{R}$  stetig auf  $f(M)$ .*

**Folgerung 1.4.7.** *Wurzelfunktionen, Exponentialfunktionen, Logarithmusfunktionen und allgemeine Potenzfunktionen sind stetig.*

#### 1.4.2 Eigenschaften stetiger Funktionen auf Intervallen

Stetige Funktionen haben Eigenschaften, die anschaulich oft recht klar sind, in einer allgemeinen Situation aber in der Regel nicht gegeben sind.

**Satz 1.4.8** (Nullstellensatz). *Eine stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$ , die sowohl positive als auch negative Werte annimmt, hat eine Nullstelle.*

**Folgerung 1.4.9.** *Es sei  $f : [a, b] \rightarrow [a, b]$  eine stetige Funktion. Dann gibt es ein  $c \in [a, b]$  mit  $f(c) = c$*

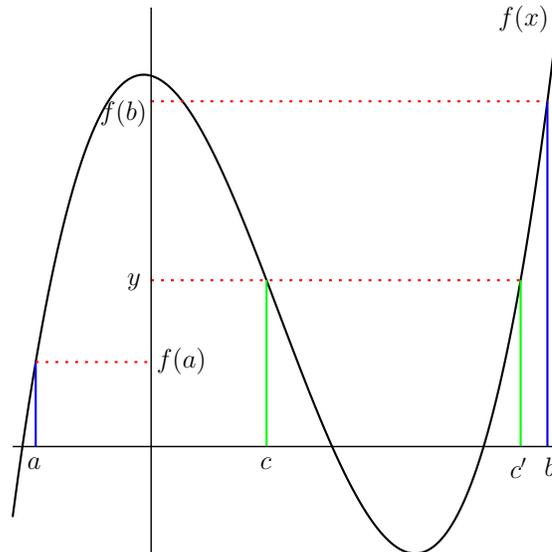
Der Nullstellensatz besitzt die folgende Verallgemeinerung.

**Satz 1.4.10** (Zwischenwertsatz). *Es sei  $f : I \rightarrow \mathbb{R}$  eine stetige Funktion auf dem Intervall  $I$ . Weiter seien  $a, b \in I$  mit  $f(a) \neq f(b)$ . Dann gibt es zu jedem  $y \in [f(a), f(b)] \cup [f(b), f(a)]$  ein  $c \in [a, b]$  mit  $f(c) = y$ .*

**Folgerung 1.4.11.** *Das Bild einer auf einem Intervall definierten stetigen Funktion ist selbst ein Intervall.*

Der Nullstellensatz hat als Konsequenz ebenfalls die folgende Aussage

Abbildung 1.4.2: Zwischenwertsatz



**Satz 1.4.12.** *Es sei  $I \subset \mathbb{R}$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  injektiv und stetig. Dann gilt:*

1.  *$f$  ist streng monoton, und*
2.  *$f^{-1} : f(I) \rightarrow \mathbb{R}$  ist stetig.*

Ein weiterer recht anschauliche Aussage ist im folgenden Satz formuliert – zumindest in Teil 1.

**Satz 1.4.13** (Existenz von Extremwerten). *Es sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion auf einem abgeschlossenen Intervall. Dann gilt:*

1.  *$f$  ist beschränkt,*
2. *und es gibt Punkte  $x_{\min}, x_{\max} \in [a, b]$  mit  $f(x_{\min}) \leq f(x) \leq f(x_{\max})$  für alle  $x \in [a, b]$ .*

*Man sagt auch: Stetige Funktionen auf abgeschlossenen Intervallen nehmen Maximum und Minimum an.*

### 1.4.3 Grenzwerte von Funktionen

Wir haben zu Beginn des Abschnitts 1.4.1 davon gesprochen, dass man "von zwei Seiten gegen einen Punkt läuft". Um gegen einen Punkt laufen zu können muss man sich "dem Punkt auch nähern können". Das zu formalisieren und einige Eigenschaften zu folgern ist der Inhalt dieses Abschnitts.

**Definition 1.4.14.** Sei  $M \subset \mathbb{R}$  eine Menge und  $x_0 \in \mathbb{R}$  ein weiterer Punkt. Dann heißt  $x_0$  *Häufungspunkt* der Menge  $M$ , wenn in jedem Intervall, das  $x_0$  enthält, mindestens ein Punkt  $x \neq x_0$  aus  $M$  liegt.

Etwas formaler:

$$\forall \epsilon > 0 \exists x \in M \setminus \{x_0\} : |x - x_0| < \epsilon.$$

**Beispiel 1.4.15.** 1. Die Häufungspunkte eines beschränkten Intervalls  $I$  (offen, abgeschlossen oder halboffen) sind die Punkte aus  $I$  sowie die beiden Randpunkte von  $I$ .

2. Es sei  $(x_n)$  eine Folge in  $\mathbb{R}$  und  $M = \{x_n \mid n \in \mathbb{N}\}$ . Dann sind die Häufungspunkte der Menge  $M$  gerade die Häufungspunkte der Folge. Ist  $(x_n)$  konvergent, dann hat  $M$  nur einen Häufungspunkt, nämlich den Grenzwert.

**Achtung:** Häufungspunkte einer Menge müssen nicht selbst zu ihr gehören.

**Satz 1.4.16.** *Es sei  $M \subset \mathbb{R}$  eine Teilmenge. Dann gelten die folgenden Aussagen:*

1.  $x_0 \in M$  ist genau dann Häufungspunkt von  $M$ , wenn für alle  $\delta > 0$  die Menge  $M \cap B_\delta(x_0)$  unendlich ist.
2.  $x_0 \in M$  ist genau dann kein Häufungspunkt von  $M$ , wenn es ein  $\delta > 0$  gibt, sodass  $M \cap B_\delta(x_0) = \{x_0\}$ .

**Definition/Bemerkung 1.4.17.** • Eine Teilmenge  $M \subset \mathbb{R}$  heißt *diskrete Menge*, wenn keine ihrer Elemente ein Häufungspunkt ist.

- Als Erweiterung von Beispiel 1.4.2.1 erhält man nun die Aussage, dass jede Funktion auf einer diskreten Menge stetig ist.

**Definition 1.4.18.** Es sei  $f : M \rightarrow \mathbb{R}$  eine Funktion,  $x_0 \in \mathbb{R}$  ein Häufungspunkt von  $M$  und  $y_0 \in \mathbb{R}$ . Dann heißt  $y_0$  *Grenzwert der Funktion  $f$  in  $x_0$* , wenn es für jede Zahl  $\epsilon > 0$  eine Zahl  $\delta > 0$  gibt, sodass  $|f(x) - y_0| < \epsilon$  für alle  $x \in M \setminus \{x_0\}$  mit  $|x - x_0| < \delta$ .

Etwas formaler:

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in M \setminus \{x_0\} : (|x - x_0| < \delta \implies |f(x) - y_0| < \epsilon).$$

Wir schreiben in diesem Fall  $y_0 = \lim_{x \rightarrow x_0} f(x)$ .

Vergleicht man das mit der Stetigkeit gemäß Definition 1.4.1 so folgt direkt der Satz 1.4.21. Davor brauchen wir allerdings noch eine Definition.

**Definition 1.4.19.** Es sei  $f : M \rightarrow \mathbb{R}$  eine Funktion und  $x_0 \in \mathbb{R}$  ein Häufungspunkt von  $M$ . Eine Funktion  $\hat{f} : M \cup \{x_0\} \rightarrow \mathbb{R}$  heißt *stetige Ergänzung von  $f$  in  $x_0$* , wenn  $\hat{f}$  in  $x_0$  stetig ist und  $f(x) = \hat{f}(x)$  für alle  $x \in M \setminus \{x_0\}$ .

**Bemerkung/Beispiel 1.4.20.** 1. In praktischen Anwendungen wird  $f$  in der Regel auf  $M \setminus \{x_0\}$  stetig sein, sodass die Funktion  $\hat{f}$  dann auf dem gesamten erweiterten Definitionsbereich  $M \cup \{x_0\}$  stetig ist.

2. Die rationale Funktion  $\hat{f}(x) = \frac{x^2 - 2}{x^3 - 4x^2 + x - 4}$  ist eine stetige Ergänzung der rationalen Funktion  $f(x) = \frac{x^3 - x^2 - 2x + 2}{x^4 - 5x^3 + 5x^2 - 5x + 4}$  im Punkt  $x_0 = 1$ .

**Satz 1.4.21.**  $y_0$  ist genau dann Grenzwert der Funktion  $f : M \rightarrow \mathbb{R}$  im Häufungspunkt  $x_0 \in \mathbb{R}$  von  $M$ , wenn  $f$  eine stetige Ergänzung  $\hat{f}$  in  $x_0$  besitzt mit  $\hat{f}(x_0) = y_0$ .

**Satz 1.4.22** (Eindeutigkeit des Grenzwertes). Ist  $x_0 \in \mathbb{R}$  ein Häufungspunkt von  $M$ , so hat  $f : M \rightarrow \mathbb{R}$  höchstens einen Grenzwert in  $x_0$ .

**Folgerung 1.4.23.**  $f : M \rightarrow \mathbb{R}$  ist genau dann im Häufungspunkt  $x_0 \in M$  von  $M$  stetig, wenn  $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ .

Es folgen einige Rechenregeln, deren Beweise im wesentlichen analog zu den Beweisen der Sätze 1.4.3 und 1.4.4 verlaufen.

**Satz 1.4.24.** 1. Für die Funktion  $f : M \rightarrow \mathbb{R}$  gelte  $\lim_{x \rightarrow x_0} f(x) = y_0$  im Häufungspunkt  $x_0 \in \mathbb{R}$  von  $M$ . Weiter sei  $N \subset \mathbb{R}$  mit  $f(M) \cup \{y_0\} \subset N$  und  $g : N \rightarrow \mathbb{R}$  sei stetig in  $y_0$ . Dann hat  $g \circ f : M \rightarrow \mathbb{R}$  in  $x_0$  den Grenzwert  $g(y_0)$ , d.h.

$$\lim_{x \rightarrow x_0} g(f(x)) = g\left(\lim_{x \rightarrow x_0} f(x)\right).$$

2. Besitzen die Funktionen  $f, g : M \rightarrow \mathbb{R}$  im Häufungspunkt  $x_0 \in \mathbb{R}$  von  $M$  einen Grenzwert, so besitzen  $f + g, f - g, fg : M \rightarrow \mathbb{R}$  und, falls  $\lim_{x \rightarrow x_0} g(x) \neq 0$ , auch  $\frac{f}{g} : M \setminus \mathcal{N}_g \rightarrow \mathbb{R}$  Grenzwerte und es gilt

$$\lim_{x \rightarrow x_0} (f(x) + g(x)) = \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x),$$

$$\lim_{x \rightarrow x_0} (f(x) - g(x)) = \lim_{x \rightarrow x_0} f(x) - \lim_{x \rightarrow x_0} g(x),$$

$$\lim_{x \rightarrow x_0} (f(x)g(x)) = \lim_{x \rightarrow x_0} f(x) \lim_{x \rightarrow x_0} g(x),$$

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} g(x)}.$$

Wir möchten auch untersuchen, ob eine Funktion einen Grenzwert besitzt, wenn wir  $x$  sehr groß werden lassen, also  $x \rightarrow \infty$ .

**Definition/Bemerkung 1.4.25.** Es sei  $f : M \rightarrow \mathbb{R}$  und  $y_0 \in \mathbb{R}$ .

- Ist  $M$  nach oben unbeschränkt, dann ist

$$\lim_{x \rightarrow \infty} f(x) = y_0 : \iff \forall \epsilon > 0 \exists C > 0 \forall x \in M : (x > C \implies |f(x) - y_0| < \epsilon)$$

- Ist  $M$  nach unten unbeschränkt, dann ist

$$\lim_{x \rightarrow -\infty} f(x) = y_0 : \iff \forall \epsilon > 0 \exists C < 0 \forall x \in M : (x < C \implies |f(x) - y_0| < \epsilon)$$

Die Rechenregeln aus Satz 1.4.24 gelten für diese Grenzwerte analog.

**Bemerkung 1.4.26.** Ist  $M$  nach oben unbeschränkt, dann ist 0 ein Häufungspunkt der Menge  $M_+ := \{\frac{1}{x} \mid x \in M \cap \mathbb{R}^{>0}\}$ . Ist  $f : M \rightarrow \mathbb{R}$  eine Funktion, dann ist  $f_+ : M_+ \rightarrow \mathbb{R}$  mit  $f_+(x) := f(\frac{1}{x})$  definiert. Damit gilt:

$$\lim_{x \rightarrow \infty} f(x) \text{ existiert} \iff \lim_{x \rightarrow 0} f_+(x) \text{ existiert}$$

und beide Grenzwerte stimmen überein.

Genauso wollen wir untersuchen ob der Grenzwert einer Funktion „unendlich“ ist.

**Definition/Bemerkung 1.4.27.** Es sei  $f : M \rightarrow \mathbb{R}$  und  $x_0 \in \mathbb{R}$  ein Häufungspunkt von  $M$ . Wir definieren

$$\begin{aligned} \lim_{x \rightarrow x_0} f(x) = \infty & \\ & : \iff \forall C > 0 \exists \delta > 0 \forall x \in M \setminus \{x_0\} : (|x - x_0| < \delta \implies f(x) > C) \\ \lim_{x \rightarrow x_0} f(x) = -\infty & \\ & : \iff \forall C < 0 \exists \delta > 0 \forall x \in M \setminus \{x_0\} : (|x - x_0| < \delta \implies f(x) < C) \end{aligned}$$

- Die Rechenregeln aus Satz 1.4.24 gelten für diese Grenzwerte analog, wenn man die folgenden formalen Regeln beachtet:

$$\infty + \infty = \infty, \quad -\infty - \infty = -\infty, \quad \pm\infty + c = \pm\infty$$

für  $c \in \mathbb{R}$ , und weiter

$$\infty \cdot (\pm\infty) = \pm\infty, \quad c \cdot \infty = \infty$$

für alle  $c \in \mathbb{R}^*$ .

- Man beachte, dass die Ausdrücke  $\infty - \infty$  und  $0 \cdot \infty$  explizit ausgenommen sind. Wie man an einfachen Beispielen erkennt, können diese nicht sinnvoll definiert werden.

Zum konkreten Berechnen von Grenzwerten ist die folgende Aussage hilfreich

**Definition/Bemerkung 1.4.28.** 1. Es sei  $f : M \rightarrow \mathbb{R}$  eine Funktion,  $x_0 \in \mathbb{R}$  ein Häufungspunkt von  $M$  und  $y_0 \in \mathbb{R}$ .

- $y_0$  heißt *oberer Grenzwert der Funktion  $f$  in  $x_0$*  und wir schreiben  $y_0 = \lim_{x \searrow x_0} f(x)$ , wenn

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in M \setminus \{x_0\} : (0 \leq x - x_0 < \delta \implies |f(x) - y_0| < \epsilon).$$

- $y_0$  heißt *unterer Grenzwert der Funktion  $f$  in  $x_0$*  und wir schreiben  $y_0 = \lim_{x \nearrow x_0} f(x)$ , wenn

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in M \setminus \{x_0\} : (0 \leq x_0 - x < \delta \implies |f(x) - y_0| < \epsilon).$$

2. Es gilt  $y_0 = \lim_{x \rightarrow x_0} f(x) \iff y_0 = \lim_{x \nearrow x_0} f(x) \wedge y_0 = \lim_{x \searrow x_0} f(x)$ .

## 1.4.4 Weitere Stetigkeitsbegriffe

In Abschnitt 1.4.3 haben wir "in der Nähe eines Punktes liegen" mit Hilfe des Häufungspunktes einer Menge formalisiert und schließlich den Begriff des Grenzwertes einer Funktion eingeführt. Hier wollen wir uns noch einen weiteren Aspekt anschauen, der das von "von zwei Seiten gegen einen Punkt laufen" vom Beginn des Abschnitts 1.4.1 konkretisiert.

**Definition 1.4.29** (Folgenstetigkeit). Eine Funktion  $f : M \rightarrow \mathbb{R}$  heißt *folgenstetig in*  $x_0 \in M$ , wenn für alle Folgen  $(y_n)$  in  $M$  mit  $\lim_{n \rightarrow \infty} y_n = x_0$  stets  $\lim_{n \rightarrow \infty} f(y_n) = f(x_0)$  gilt.

Die obige Definition sagt, dass eine Funktion folgenstetig ist, wenn die Grenzwertbildung und die Anwendung der Funktionsvorschrift vertauschbar sind. Der folgende Satz sagt uns jedoch, dass wir damit nichts Neues definiert haben.

**Satz 1.4.30.** *Eine Funktion ist genau dann folgenstetig, wenn sie stetig ist.*

**Bemerkung 1.4.31.** 1. Der vorige Satz liefert einem eine Möglichkeit auf "einfache" Weise zu zeigen, dass eine Funktion an einer Stelle nicht stetig ist.

2. Eine kleine Umformulierung des Satzes liefert eine Beschreibung der Grenzwertbegriffe für eine Funktion  $f : M \rightarrow \mathbb{R}$  mit Hilfe von Zahlenfolgen, siehe Definition 1.4.18 und Definition/Bemerkung 1.4.28. Es gilt:

- $\lim_{x \rightarrow x_0} f(x) = \xi \iff$  Für alle  $(y_n)_{n \in \mathbb{N}}$  in  $M \setminus \{x_0\}$  gilt

$$\lim_{n \rightarrow \infty} y_n = x_0 \implies \lim_{n \rightarrow \infty} f(y_n) = \xi$$

- $\lim_{x \nearrow x_0} f(x) = \xi \iff$  Für alle  $(y_n)_{n \in \mathbb{N}}$  in  $M \setminus \{x_0\}$  gilt

$$(\forall n : y_n \leq x_0) \wedge \lim_{n \rightarrow \infty} y_n = x_0 \implies \lim_{n \rightarrow \infty} f(y_n) = \xi$$

- $\lim_{x \searrow x_0} f(x) = \xi \iff$  Für alle  $(y_n)_{n \in \mathbb{N}}$  in  $M \setminus \{x_0\}$  gilt

$$(\forall n : y_n \geq x_0) \wedge \lim_{n \rightarrow \infty} y_n = x_0 \implies \lim_{n \rightarrow \infty} f(y_n) = \xi$$

Bei der Überprüfung der Stetigkeit zu einer vorgegebenen Schranke  $\epsilon$ , kann die Wahl der Zahl  $\delta$  stark von dem Punkt  $x_0$  abhängen kann – wir haben das bei der Funktion  $x \mapsto \frac{1}{x}$  gesehen. ”Schöner” ist die Situation, wenn für jedes  $\epsilon$  dieses  $\delta$  universell – d.h. unabhängig von  $x_0$  – gewählt werden kann. Diese Eigenschaft macht natürlich ”punktuell” keinen Sinn.

Funktionen mit dieser Eigenschaft sind Gegenstand der folgenden Definition.

**Definition 1.4.32** (Gleichmäßige Stetigkeit). Eine Funktion  $f : M \rightarrow \mathbb{R}$  heißt *gleichmäßig stetig*, wenn es zu jeder Zahl  $\epsilon > 0$  eine Zahl  $\delta > 0$  gibt, sodass  $|f(x) - f(y)| < \epsilon$  für alle  $x, y \in M$  mit  $|x - y| < \delta$ .

Etwas formaler:

$$\forall \epsilon > 0 \exists \delta > 0 \forall x, y \in M : (|x - y| < \delta \implies |f(x) - f(y)| < \epsilon).$$

**Beispiel 1.4.33.** • Die stetige Funktion  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  mit  $f(x) = \frac{1}{x}$  ist nicht gleichmäßig stetig.

- Die Funktion  $f : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}$  mit  $f(x) = \sqrt{x}$  ist gleichmäßig stetig.

**Bemerkung 1.4.34.** Eine gleichmäßig stetige Funktion ist auch stetig.

**Satz 1.4.35.** *Eine stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  auf einem abgeschlossenen Intervall ist gleichmäßig stetig.*

Die gleichmäßige Stetigkeit lässt sich noch etwas verschärfen:

**Definition 1.4.36** (Lipschitz-Stetigkeit). Eine Funktion  $f : M \rightarrow \mathbb{R}$  heißt *Lipschitz-stetig*, wenn es eine Zahl  $L > 0$  gibt – die Lipschitz-Konstante –, sodass

$$\forall x, y \in M : |f(x) - f(y)| \leq L|x - y|.$$

**Bemerkung 1.4.37.** Eine Lipschitz-stetige Funktion ist auch gleichmäßig stetig.

**Beispiel 1.4.38.** • Die gleichmäßig stetige Funktion  $f : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}$  mit  $f(x) = \sqrt{x}$  ist nicht Lipschitz-stetig.

- Die Einschränkung der vorigen Funktion auf  $\mathbb{R}^{\geq \frac{1}{2}}$  ist Lipschitz-stetig.

Wir wissen wegen Folgerung 1.4.9 schon, dass jede stetige Funktion auf einem abgeschlossenen Intervall, deren Bild ebenfalls in diesem Intervall liegt, einen Fixpunkt hat. Da jede Lipschitz-stetige Funktion auch stetig ist, gilt das auch für diese.

Die Lipschitz-Stetigkeit wird im kommenden Semester noch eine wesentliche Rolle spielen. Deshalb formulieren wir hier schon eine Variante von Folgerung 1.4.9, die zusätzlich noch ein Verfahren liefert, wie man diesen existierenden Fixpunkt – zumindest näherungsweise – berechnen kann.

**Satz 1.4.39** (kleiner Fixpunktsatz). *Es sei  $[a, b] \subset \mathbb{R}$  ein abgeschlossenes Intervall und  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion mit  $f([a, b]) \subset [a, b]$ . Weiter sei  $f$  Lipschitz-stetig mit Lipschitz-Konstante  $L < 1$ . Dann gilt*

1. Die Funktion hat einen eindeutigen Fixpunkt  $x^* \in [a, b]$ , das heißt  $f(x^*) = x^*$ , siehe auch Folgerung 1.4.9.
2. Dieser lässt sich als Grenzwert der rekursiv definierten Folge  $(x_n)$  berechnen mit  $x_k := f(x_{k-1})$  für  $k \geq 1$  und  $x_0 \in [a, b]$  beliebig. Es gilt dabei

$$(a) \quad |x_k - x^*| \leq L |x_{k-1} - x^*|,$$

$$(b) \quad |x_k - x^*| \leq \frac{L^k}{1-L} |x_1 - x_0|,$$

$$(c) \quad |x_k - x^*| \leq \frac{L}{1-L} |x_k - x_{k-1}|.$$

**Bemerkung 1.4.40.** 1. Eine Funktion  $f : M \rightarrow \mathbb{R}$  mit  $f(M) \subset M$  heißt *Selbstabbildung*.

2. Man nennt eine Lipschitz-stetige Funktion  $f : M \rightarrow M$  mit Lipschitzkonstante  $L < 1$  auch *Kontraktion* oder *kontrahierende Funktion*.
3. Die Funktionen, die in Satz 1.4.39 betrachtet werden, sind somit kontrahierende Selbstabbildungen.
4. Dass man auf keine der Voraussetzungen in Satz 1.4.39 verzichten kann, zeigen die folgenden Beispiele:

$$(a) \quad \text{Das Intervall ist nicht abgeschlossen: } f : ]0, 1] \rightarrow ]0, 1] \text{ mit } f(x) = \frac{1}{2}x.$$

- (b) Die Abbildung ist keine Selbstabbildung:  $f : [0, 1] \rightarrow \mathbb{R}^+$  mit  $f(x) = \frac{1}{2}x + \frac{2}{3}$ .
- (c) Die Abbildung ist keine Kontraktion:  $f : [0, 1] \rightarrow [0, 1]$  mit  $f(x) = x$ .
5. Dass die Bedingungen jedoch nicht notwendig sind für die Folgerungen 1. und 2. aus Satz 1.4.39, zeigt das Beispiel der Lipschitz-stetige Selbstabbildung  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = 2x$ . Sie hat einen eindeutigen Fixpunkt,  $x^* = 0$  obwohl Intervall offen ist und die Lipschitz-Konstante  $L = 2 > 1$ . Man kann diesen Fixpunkt jedoch nicht mit Hilfe der im Satz angegebenen rekursiven Folge berechnen.

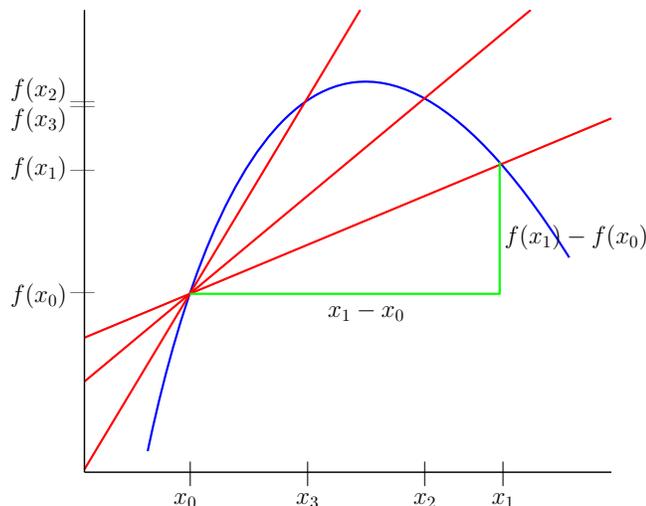
## 1.5 Differenzierbarkeit

Wir widmen uns in diesem Abschnitt dem Unterschied der "Verklebestellen" der zweiten bis vierten Funktion zu Beginn von Abschnitt 1.4.1. All diese Funktionen waren stetig, dennoch unterscheidet sich optisch die zweite von der dritten und vierten: sie besitzt eine "Ecke". Das wollen wir im folgenden konkretisieren und auch sehen, dass sich die Art der "Verklebestellen" der dritten und vierten Funktion unterscheiden.

### 1.5.1 Ableitung und Differenzierbarkeit

Es ist hinlänglich bekannt, was die Steigung einer linearen Funktion ist: Man berechnet sie z.B. mit Hilfe eines Steigungsdreiecks. Im Fall einer nicht-linearen Funktion macht das Steigungsdreieck so keinen Sinn, da man für unterschiedliche  $x$ -Werte in der Regel unterschiedliche Dreiecke erhält – das sogar, wenn man einen der beteiligten Punkte festhält.

Abbildung 1.5.1: Differenzierbarkeit



**Definition 1.5.1.** 1. Die reelle Funktion  $f : M \rightarrow \mathbb{R}$  heißt *differenzierbar im Häufungspunkt*  $x_0 \in M$ , wenn der Grenzwert<sup>(i)</sup>

$$f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert.  $f'(x_0)$  heißt dann die *Ableitung von  $f$  an der Stelle  $x_0$* . Der Ausdruck  $\frac{f(x) - f(x_0)}{x - x_0}$  heißt *Differenzenquotient*.

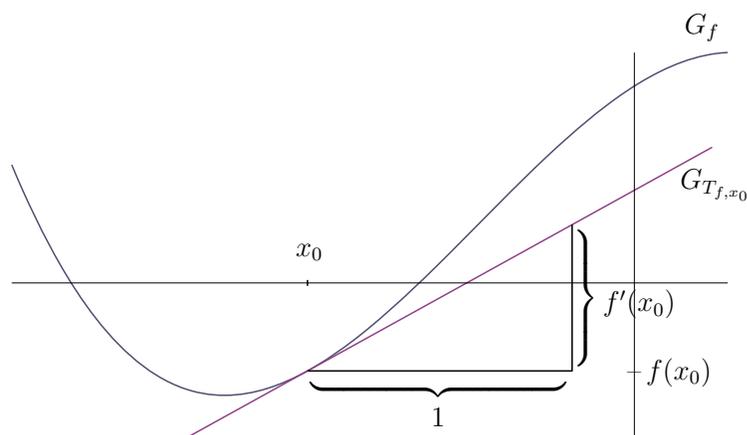
2. Ist  $f$  in  $x_0$  differenzierbar mit Ableitung  $f'(x_0)$ , dann nennt man die Gerade bzw. ihre beschreibende Funktion

$$T_{f,x_0}(x) = f(x_0) + f'(x_0)(x - x_0)$$

die *Tangente an  $f$  in  $x_0$* . Sie entsteht als Grenzobjekt der Sekanten mit den jeweiligen Differenzenquotienten als Steigung, siehe Abbildung 1.5.2

<sup>(i)</sup>Wenn wir im Folgenden von der Differenzierbarkeit einer Funktion  $f : M \rightarrow \mathbb{R}$  in einem Punkt  $x_0 \in M$  sprechen, dann setzen wir implizit voraus, dass dieser Punkt ein Häufungspunkt von  $M$  ist.

Abbildung 1.5.2: Tangente



**Bemerkung 1.5.2.** Eine Funktion  $f : M \rightarrow \mathbb{R}$  ist genau dann in  $x_0 \in M$  differenzierbar, wenn

- $f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0+h) - f(x_0)}{h}$  existiert.
- $f'(x_0) = \lim_{n \rightarrow \infty} \frac{f(x_0) - f(a_n)}{x_0 - a_n}$  für alle Folgen  $(a_n)$  in  $M$  mit  $x_0 = \lim_{n \rightarrow \infty} a_n$  existiert.
- $f'(x_0) = \lim_{n \rightarrow \infty} \frac{f(x_0+a_n) - f(x_0)}{a_n}$  für alle Nullfolgen  $(a_n)$  existiert.

**Bemerkung 1.5.3.** In der Regel handelt es sich beim Definitionsbereich der Funktion  $f$  um ein Intervall (oder um die Vereinigung von Intervallen). Unter der Differenzierbarkeit in einem Randpunkt ist dann der einseitige Grenzwert zu verstehen.

**Beispiel 1.5.4.** 1. Für alle  $n \in \mathbb{N}$  ist die Funktion  $f(x) = x^n$  in jedem Punkt  $x_0$  ihres Definitionsbereiches differenzierbar mit  $f'(x_0) = nx_0^{n-1}$ .

2. Die Funktion  $E_2(x) = 2^x$  ist in  $x = 0$  differenzierbar und es gilt

$$E_2'(0) = \lim_{n \rightarrow \infty} n(\sqrt[n]{2} - 1).$$

**Satz 1.5.5** (Lineare Approximierbarkeit). Die Funktion  $f : M \rightarrow \mathbb{R}$  ist genau dann in  $x_0 \in M$  von  $M$  differenzierbar,

1. wenn es eine in  $x_0$  stetige Funktion  $\hat{f} : M \rightarrow \mathbb{R}$  gibt mit

$$f(x) = f(x_0) + \hat{f}(x)(x - x_0)$$

für alle  $x \in M$ . Es gilt dann  $\hat{f}(x_0) = f'(x_0)$ .

2. wenn es eine in  $x_0$  stetige Funktion  $r : M \rightarrow \mathbb{R}$  mit  $r(x_0) = 0$  und eine Zahl  $c \in \mathbb{R}$  gibt, sodass

$$f(x) = f(x_0) + c(x - x_0) + r(x)(x - x_0)$$

für alle  $x \in M$ . Dann ist  $c = f'(x_0)$ .

**Beweisskizze.** Insbesondere sind 1. und 2. äquivalent. Direkt sieht man das wie folgt: Für 1.  $\Rightarrow$  2. definiere  $c := \hat{f}(x_0)$  und  $r(x) := \hat{f}(x) - \hat{f}(x_0)$  und für 2.  $\Rightarrow$  1. definiere  $\hat{f}(x) := c + r(x)$ .  $\square$

**Bemerkung 1.5.6.** Die Formulierung 2. des vorigen Satzes begründet auch den Namen "Lineare Approximierbarkeit": Die Tangentenfunktion  $T_{f,x_0} : M \rightarrow \mathbb{R}$  approximiert die Funktion  $f$  in  $x_0$  von erster Ordnung<sup>(i)</sup>, d.h.

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

**Folgerung 1.5.7.** Ist  $f : M \rightarrow \mathbb{R}$  in  $x_0 \in M$  differenzierbar, so ist  $f$  in  $x_0$  auch stetig.

**Satz 1.5.8.** 1. (Kettenregel). Ist  $f : M \rightarrow \mathbb{R}$  differenzierbar im Punkt  $x_0 \in M$  und ist  $g : N \rightarrow \mathbb{R}$  mit  $f(M) \subset N$  im Punkt  $f(x_0)$  differenzierbar, dann ist auch die Verkettung  $g \circ f : M \rightarrow \mathbb{R}$  im Punkt  $x_0$  differenzierbar und es gilt

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

2. (Summenregel). Sind  $f, g : M \rightarrow \mathbb{R}$  beide in  $x_0 \in M$  differenzierbar, so auch die Differenz  $f - g$  und die Summe  $f + g$  und es gilt

$$(f \pm g)'(x_0) = f'(x_0) \pm g'(x_0).$$

---

<sup>(i)</sup>Man sagt, eine Funktion  $g : M \rightarrow \mathbb{R}$  approximiert die Funktion  $f : M \rightarrow \mathbb{R}$  in  $x_0 \in M$  von  $k$ -ter Ordnung, wenn  $\lim_{x \rightarrow x_0} \frac{f(x) - g(x)}{(x - x_0)^k} = 0$ .

3. (Produktregel). Sind  $f, g : M \rightarrow \mathbb{R}$  beide in  $x_0 \in M$  differenzierbar, so auch das Produkt  $fg$  und es gilt

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

**Folgerung 1.5.9.** Es ergeben sich sofort die folgenden Spezialfälle: Sind  $f, g : M \rightarrow \mathbb{R}$  in  $x_0$  differenzierbar mit  $g(x_0) \neq 0$  und ist  $c \in \mathbb{R}$  so gilt

$$\begin{aligned} (cf)'(x_0) &= cf'(x_0), \\ \left(\frac{1}{g}\right)'(x_0) &= -\frac{g'(x_0)}{g(x_0)^2}, \\ \left(\frac{f}{g}\right)'(x_0) &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}. \end{aligned}$$

Die letzte Rechenregel heißt auch *Quotientenregel*.

Eine weitere wichtige Konsequenz aus den Rechenregeln ist der folgende Satz.

**Satz 1.5.10** (Ableitung der Umkehrfunktion). *Es sei  $f : M \rightarrow \mathbb{R}$  injektiv und in  $x_0 \in M$  differenzierbar. Dann sind folgende Aussagen äquivalent:*

- $f^{-1}$  ist in  $f(x_0)$  stetig und es ist  $f'(x_0) \neq 0$ .
- $f(x_0)$  ist ein Häufungspunkt von  $f(M)$  und die Umkehrfunktion  $f^{-1} : f(M) \rightarrow \mathbb{R}$  ist in  $f(x_0)$  differenzierbar.

Für die Ableitung gilt in diesem Fall

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

**Beispiel 1.5.11.** 1. Aus der Kettenregel folgt mit  $E_a(x) = E_2(x \log_2(a))$ , dass die allgemeine Exponentialfunktion  $E_a(x) = a^x$  in  $x = 0$  differenzierbar ist. Es gilt  $E'_a(0) = \lim_{n \rightarrow \infty} n(\sqrt[n]{2} - 1) \cdot \log_2 a$ .

2. Mit Satz 1.5.10 und der Ableitung der Potenzfunktion folgt, dass die Wurzelfunktion  $x \mapsto \sqrt[n]{x}$  ist in jedem Punkt ihres Definitionsbereiches differenzierbar außer in  $x_0 = 0$ .

**Definition 1.5.12.** Eine Funktion  $f : M \rightarrow \mathbb{R}$  heißt *differenzierbar*, wenn sie an jeder Stelle  $x_0 \in M$  differenzierbar ist. Die Funktion

$$f' : M \rightarrow \mathbb{R}$$

die jeder Stelle  $x \in M$  die Ableitung  $f'(x)$  an dieser Stelle zuordnet, heißt *Ableitung von  $f$* .

**Beispiel 1.5.13.** 1. Aus Produkt und Quotientenregel folgt, dass alle Polynome und rationalen Funktionen auf ihren Definitionsbereichen differenzierbar sind.

2. Beispiel 1.5.11 sagt, dass alle Wurzelfunktionen auf  $\mathbb{R}^+$  differenzierbar sind.

3. Aus  $E_a(x) = E_a(x_0)E_a(x - x_0)$  folgt, dass die allgemeine Exponentialfunktion  $E_a(x) = a^x$  differenzierbar ist mit

$$E'_a(x) = E'_a(0) \cdot E_a(x),$$

wobei  $E'_a(0) = \lim_{n \rightarrow \infty} n(\sqrt[n]{a} - 1) \cdot \log_2 a$  von oben.

**Definition/Bemerkung 1.5.14** (Exponentialfunktion). 1. Die *Exponentialfunktion* ist die allgemeine Exponentialfunktion mit der Eigenschaft  $E'_e(x) = E_e(x)$ . Die Basis heißt *Eulersche Zahl  $e$*  und ist gegeben durch<sup>(i)</sup>

$$e = E_2\left(\frac{1}{\lim_{n \rightarrow \infty} n(\sqrt[n]{2} - 1)}\right) = 2^{\frac{1}{\lim_{n \rightarrow \infty} n(\sqrt[n]{2} - 1)}}.$$

2. Die Umkehrfunktion der Exponentialfunktion heißt *natürlicher Logarithmus* und wir schreiben

$$\ln := \log_e : \mathbb{R}^+ \rightarrow \mathbb{R}.$$

**Beispiel 1.5.15.** 1. Mit den Bezeichnungen von oben gilt

$$E'_a(x) = \ln(a) E_a(x).$$

<sup>(i)</sup>Wir werden später etwas weniger sperrige Eigenschaften der Eulerschen Zahl kennenlernen, siehe Bemerkung 1.6.37.

2. Die Ableitung des natürlichen Logarithmus ist die Normalhyperbel, d.h. für alle  $x \in \mathbb{R}^+$  ist

$$\ln'(x) = \frac{1}{x}.$$

Wegen der Symmetrie folgt sogar  $(\ln|x|)' = \frac{1}{x}$  für alle  $x \in \mathbb{R}^*$ .

3. Die Ableitung der allgemeinen Potenzfunktion mit  $f(x) = x^r := e^{r \ln(x)}$  ist

$$f'(x) = e^{r \ln(x)} \cdot \frac{r}{x} = r e^{r \ln(x) - \ln(x)} = r e^{(r-1) \ln(x)} = r x^{r-1}.$$

### 1.5.2 Eigenschaften differenzierbarer Funktionen

So wie die stetigen Funktionen spezielle Eigenschaften haben, gilt dieses auch für die differenzierbaren Funktionen.

**Satz 1.5.16.** *Ist  $f : M \rightarrow \mathbb{R}$  differenzierbar und monoton steigend (fallend), so ist  $f'(x) \geq 0$  ( $f'(x) \leq 0$ ) für alle  $x \in M$ .*

**Definition 1.5.17.** • Eine Funktion  $f : M \rightarrow \mathbb{R}$  hat in  $x_0 \in M$  ein *lokales Maximum (lokales Minimum)*, wenn es eine Zahl  $\epsilon > 0$  gibt mit  $f(x) \leq f(x_0)$  ( $f(x) \geq f(x_0)$ ) für alle  $x \in M \cap U_\epsilon(x_0)$ .

- Gilt  $f(x) \leq f(x_0)$  ( $f(x) \geq f(x_0)$ ) sogar für alle  $x \in M$ , dann nennt man  $x_0$  ein *globales Maximum (globales Minimum)*.
- Unterscheidet man nicht zwischen Maximum oder Minimum, so spricht man von *lokalen bzw. globalen Extrema*.

**Satz 1.5.18.** *Besitzt die differenzierbare Funktion  $f : M \rightarrow \mathbb{R}$  im Häufungspunkt  $x_0 \in M$  ein lokales Extremum so gilt  $f'(x_0) = 0$ .*

**Bemerkung 1.5.19.** Satz 1.5.18 muss für globale Extrema nicht richtig sein.

**Bemerkung 1.5.20.** Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig mit  $f(a) = f(b)$ , dann hat  $f$  auf  $]a, b[$  ein Extremum.

**Satz 1.5.21** (Satz von Rolle). *Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig und auf  $]a, b[$  differenzierbar und gilt  $f(a) = f(b)$ , so gibt es mindestens ein  $x_0 \in ]a, b[$  mit  $f'(x_0) = 0$ .*

Tabelle 2: Einige Ableitungen

| $f(x)$            | $f'(x)$                         | Einschränkungen   |
|-------------------|---------------------------------|---|
| $c$               | $0$                             | $x \in \mathbb{R}$  |
| $x$               | $1$                             | $x \in \mathbb{R}$  |
| $x^n$             | $nx^{n-1}$                      | $n \in \mathbb{N}^*, x \in \mathbb{R}$  |
| $\frac{1}{x^n}$   | $-\frac{n}{x^{n+1}}$            | $n \in \mathbb{N}^*, x \in \mathbb{R}^*$  |
| $\sqrt{x}$        | $\frac{1}{2\sqrt{x}}$           | $x \in \mathbb{R}^+$  |
| $\sqrt[n]{x}$     | $\frac{\sqrt[n]{x}}{nx}$        | $n \in \mathbb{N}^{\geq 2}, \begin{cases} x \in \mathbb{R}^+, & \text{falls } n \text{ gerade} \\ x \in \mathbb{R}^*, & \text{falls } n \text{ ungerade} \end{cases}$ |
| $x^{\frac{m}{n}}$ | $\frac{m x^{\frac{m}{n}}}{n x}$ | $m, n \in \mathbb{N}^*, x \in \mathbb{R}^+$   |
| $e^x$             | $e^x$                           | $x \in \mathbb{R}$  |
| $\ln(x)$          | $\frac{1}{x}$                   | $x \in \mathbb{R}^+$  |
| $a^x$             | $\ln(a)a^x$                     | $a > 0, x \in \mathbb{R}$   |
| $\log_a(x)$       | $\frac{1}{x \ln(a)}$            | $a > 0, x \in \mathbb{R}^+$   |
| $x^r$             | $r x^{r-1}$                     | $r \in \mathbb{R}, x \in \mathbb{R}^+$  |

**Satz 1.5.22** (Mittelwertsatz der Differentialrechnung). 1. Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig und auf  $]a, b[$  differenzierbar. Dann gibt es mindestens ein  $x_0 \in ]a, b[$  mit

$$f'(x_0) = \frac{f(a) - f(b)}{a - b}.$$

2. (verallgemeinerte Variante). Sind  $f, g : [a, b] \rightarrow \mathbb{R}$  stetig und auf  $]a, b[$  differenzierbar, so gibt es mindestens ein  $x_0 \in ]a, b[$ , sodass

$$(f(a) - f(b))g'(x_0) = (g(a) - g(b))f'(x_0).$$

3. Sei  $f : I \rightarrow \mathbb{R}$  differenzierbar. Dann gibt es zu je zwei Punkten  $x, x_0 \in I$  ein  $y$  zwischen  $x$  und  $x_0$  mit

$$f(x) = f(x_0) + f'(y)(x - x_0).$$

**Beweisskizze.** In 1. setzen wir  $h(x) = f(x)(a - b) - x(f(a) - f(b))$  und in 2.  $h(x) = f(x)(g(a) - g(b)) - g(x)(f(a) - f(b))$ . Auf diese Funktionen wenden wir dann den Satz von Rolle an.  $\square$

**Folgerung 1.5.23.** 1. Ist  $I$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  differenzierbar mit  $f'(x) = 0$  für alle  $x \in I$  so ist  $f$  konstant.

2. Ist  $I$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  differenzierbar mit  $f'(x) \neq 0$  für alle  $x \in I$ , so ist  $f$  injektiv.
3. Ist der Definitionsbereich der differenzierbaren Funktion  $f : I \rightarrow \mathbb{R}$  ein Intervall, so gilt auch die Umkehrung von Satz 1.5.16: Ist  $f'(x) \geq 0$  ( $f'(x) \leq 0$ ) für alle  $x \in I$ , so ist  $f$  monoton steigend (fallend).
4. Ist  $I$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  differenzierbar mit  $f'(x) \neq 0$  für alle  $x \in I$ , so ist  $f^{-1} : f(I) \rightarrow \mathbb{R}$  ebenfalls differenzierbar auf dem gesamten Bildintervall  $f(I)$  und es gilt

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}.$$

5. Ist  $I$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  differenzierbar mit beschränkter Ableitung, dann ist  $f$  Lipschitz-stetig.
6. Ist  $f' : [a, b] \rightarrow \mathbb{R}$  stetig, dann ist  $f$  Lipschitz-stetig.
7. Ist in den beiden vorigen Punkten insbesondere  $|f'(x)| \leq L < 1$  für alle  $x \in I$ , dann hat die Funktion  $f$  einen Fixpunkt.

Eine weitere wichtige Folgerung formulieren wir hier als Satz:

**Folgerung 1.5.24.** Gegeben seien zwei differenzierbare Funktionen  $f, g : I \rightarrow \mathbb{R}$  auf dem endlichen Intervall  $I$ . Für den Randpunkt  $x_0$  von  $I$  gelte  $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$  und weiter sei  $g'(x) \neq 0$  für alle  $x \in I$ .

Wenn nun  $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} \in \mathbb{R}$  existiert, dann existiert auch  $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$  und beide Grenzwerte stimmen überein.

Dieser Satz besitzt die folgende Verallgemeinerung:

**Satz 1.5.25** (Regel von L'Hospital). *Gegeben seien zwei differenzierbare Funktionen  $f, g : I \rightarrow \mathbb{R}$  auf dem Intervall  $I = ]a, b[$ , wobei  $a = -\infty$  oder  $b = \infty$  zugelassen sind. Weiter gelte  $g'(x) \neq 0$  für alle  $x \in ]a, b[$  und es existiere*

$$\lim_{x \rightarrow b} \frac{f'(x)}{g'(x)} \in \mathbb{R} \cup \{\pm\infty\}.$$

*Gilt dann noch eine der beiden Aussagen*

$$\lim_{x \rightarrow b} f(x) = \lim_{x \rightarrow b} g(x) = 0 \quad \text{oder} \quad \lim_{x \rightarrow b} f(x) = \lim_{x \rightarrow b} g(x) = \pm\infty.$$

*so folgt*

$$\lim_{x \rightarrow b} \frac{f(x)}{g(x)} = \lim_{x \rightarrow b} \frac{f'(x)}{g'(x)}.$$

*Analog gilt die Aussage für den anderen Randpunkt des Intervalls auf dem die Funktionen definiert sind.*

**Beispiel 1.5.26.** •  $\lim_{x \rightarrow \infty} \frac{e^{ax}}{x^b} = \infty$ ,  $\lim_{x \rightarrow \infty} x^b e^{-ax} = 0$  für alle  $a, b > 0$

•  $\lim_{x \rightarrow \infty} \frac{(\ln x)^a}{x^b} = 0$ ,  $\lim_{x \rightarrow 0} x^b (\ln x)^a = 0$  für alle  $a, b > 0$

### 1.5.3 Höhere Ableitungen und der Satz von Taylor

Ist eine Funktion  $f : M \rightarrow \mathbb{R}$  in jedem Punkt differenzierbar, dann existiert die Ableitungsfunktion  $f' : M \rightarrow \mathbb{R}$ . Es stellt sich da die Frage, was diese Funktion nun für Eigenschaften hat. Z.B: Ist diese Funktion gegebenenfalls selbst differenzierbar? Wir wissen aus Folgerung 1.5.7, dass eine in einem Punkt differenzierbare Funktion dort auch stetig ist. Wenn also  $f'$  differenzierbar sein soll, dann muss  $f'$  auch stetig sein.

**Bezeichnung 1.5.27.** Eine Funktion  $f : M \rightarrow \mathbb{R}$  deren Ableitungsfunktion  $f' : M \rightarrow \mathbb{R}$  existiert und stetig ist, heißt *stetig differenzierbar*.

**Definition 1.5.28.** 1. Eine Funktion  $f : M \rightarrow \mathbb{R}$  heißt *n-mal differenzierbar* für  $n \in \mathbb{N}^*$ , wenn die Ableitungen  $f', f'' := (f')', \dots, f^{(n)} := (f^{(n-1)})' : M \rightarrow \mathbb{R}$  existieren.  $f^{(k)}$  heißt dann die *k-te Ableitung von f*.

2. Ist die  $k$ -te Ableitung  $f^{(k)}$  einer  $k$ -mal differenzierbaren Funktion stetig, so heißt  $f$   *$k$ -mal stetig differenzierbar*.
3. Existiert  $f^{(n)}$  für jedes  $n \in \mathbb{N}^*$  so heißt  $f$  *beliebig oft differenzierbar*.
4. Ist jede Ableitung einer beliebig oft differenzierbaren Funktion  $f$  stetig, so heißt  $f$  *glatt*.

**Beispiel 1.5.29.** Die Funktion  $x \mapsto x^k|x|$  ist  $k$ -mal differenzierbar, aber nicht  $(k+1)$ -mal.

**Bemerkung 1.5.30.** Alle Rechenregeln bleiben auch für die höheren Ableitungen gültig, bzw. lassen sich erweitern. Für das Produkt zweier Funktionen gilt

$$(fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} \cdot g^{(n-k)}.$$

Für die Verkettung zweier Funktionen gilt

$$\begin{aligned} (f \circ g)'' &= (f'' \circ g)(g')^2 + (f' \circ g)g'', \\ (f \circ g)''' &= (f''' \circ g)(g')^3 + 3(f'' \circ g)g'g'' + (f' \circ g)g''', \end{aligned}$$

oder allgemein die *Formel von Faà di Bruno*

$$(f \circ g)^{(n)} = \sum_{(i_1, \dots, i_n) \in \mathcal{P}_n} \frac{n!}{i_1! \cdot \dots \cdot i_n!} (f^{(i_1 + \dots + i_n)} \circ g) \prod_{j=1}^n \left( \frac{1}{j!} g^{(j)} \right)^{i_j}$$

wobei  $\mathcal{P}_n$  alle  $n$ -Tupel  $(i_1, \dots, i_n)$  enthält mit  $i_1 + 2i_2 + \dots + ni_n = n$ .

**Satz 1.5.31** (Satz von Taylor). *Es sei  $I$  ein Intervall, das mehr als einen Punkt enthält. Weiter sei  $f : I \rightarrow \mathbb{R}$  eine  $(n+1)$ -mal differenzierbare Abbildung. Dann gelten die folgenden äquivalenten Aussagen*

- Für alle  $x_0, x \in I$  gibt es eine Zahl  $y \in [x_0, x] \cup [x, x_0]$ , sodass

$$\begin{aligned} f(x) &= f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \dots \\ &\quad + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \frac{f^{(n+1)}(y)}{(n+1)!}(x-x_0)^{n+1}. \end{aligned}$$

- Für alle  $x_0, x \in I$  gibt es eine Zahl  $t \in [0, 1]$ , sodass

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots \\ + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(tx + (1-t)x_0)}{(n+1)!}(x - x_0)^{n+1}.$$

- Für alle  $x_0 \in I$  und  $t \in \mathbb{R}$  mit  $x_0 + t \in I$  gibt es ein  $s \in [0, t] \cup [t, 0]$ , sodass

$$f(x_0 + t) = f(x_0) + \frac{f'(x_0)}{1!}t + \frac{f''(x_0)}{2!}t^2 + \dots + \frac{f^{(n)}(x_0)}{n!}t^n \\ + \frac{f^{(n+1)}(x_0 + s)}{(n+1)!}t^{n+1}.$$

**Beweisskizze.** (Per Induktion)

$n = 0$ : Mittelwertsatz [1.5.22](#)

$n - 1 \rightarrow n$ : Sei  $x_1 \in I$ . Zu zeigen: Es gibt ein  $y$  zwischen  $x_1$  und  $x_0$ , sodass

$$f(x_1) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x_1 - x_0)^k + \frac{f^{(n+1)}(y)}{(n+1)!}(x_1 - x_0)^{n+1}.$$

Für  $x_1 = x_0$  ist nichts zu zeigen und wir setzen  $x_1 \neq x_0$  voraus.

Idee: Wir definieren eine Konstante  $c$ , die von  $x_1$  abhängt, durch  $f(x_1) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x_1 - x_0)^k - c(x_1 - x_0)^{n+1} = 0$ , und wir berechnen  $c$ .

Sei dazu  $g(x) := f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k - c(x - x_0)^{n+1}$ , sodass

$$g'(x) = f'(x) - \sum_{k=0}^{n-1} \frac{f^{(k+1)}(x_0)}{k!}(x - x_0)^k - c(n+1)(x - x_0)^n.$$

Wegen  $g(x_0) = g(x_1) = 0$  gibt es mit Satz [1.5.21](#) ein  $\tilde{y}$  zwischen  $x_1$  und  $x_0$  mit  $g'(\tilde{y}) = 0$  oder

$$f'(\tilde{y}) = \sum_{k=0}^{n-1} \frac{f^{(k+1)}(x_0)}{k!}(\tilde{y} - x_0)^k + c(n+1)(\tilde{y} - x_0)^n.$$

Nun lässt sich die Induktionsannahme auf  $f'$  anwenden und liefert

$$f'(\tilde{y}) = \sum_{k=0}^{n-1} \frac{(f')^{(k)}(x_0)}{k!} (\tilde{y} - x_0)^k + \frac{(f')^{(n)}(y)}{n!} (\tilde{y} - x_0)^n$$

für ein  $y$  zwischen  $\tilde{y}$  und  $x_0$ . Ein Vergleich der beiden Ausdrücke gibt schließlich  $c = \frac{f^{(n+1)}(y)}{(n+1)!}$ .  $\square$

**Definition 1.5.32.** Es gelten die Bezeichnungen aus dem obigen Satz. Dann heißt das Polynom

$$T_{n,f,x_0}(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

das  $n$ -te *Taylorpolynom* von  $f$  an der Stelle  $x_0$ . Der Term  $R_{n,f,x_0}(x) = f(x) - T_{n,f,x_0}(x)$  heißt *Restglied* und die Form

$$R_{n,f,x_0}(x) = \frac{f^{(n+1)}(y)}{(n+1)!} (x - x_0)^{n+1}$$

heißt *Lagrangedarstellung* des Restgliedes.

**Bemerkung 1.5.33.** • Das  $n$ -te Taylorpolynom hat maximal den Grad  $n$ .

- Das lineare Taylorpolynom  $T_{1,f,x_0}$  entspricht der Tangente von  $f$  in  $x_0$ .
- Für eine Funktion  $f$  hängt die Zwischenstelle  $y$  aus dem Satz von Taylor von allen eingehenden Größen ab, also von  $x_0, x$  und  $n$ . Weiter liefert der Satz von Taylor nur die Existenz der Zwischenstelle  $y$ , er macht jedoch keine Aussage über dessen genaue Lage. Daher können wir im Allgemeinen keine Aussage darüber machen, wie gut das Taylorpolynom die Funktion bei wachsendem  $n$  "approximiert".
- Im Fall guter Approximation, kann das Taylorpolynom zur Berechnung von Näherungswerten von Funktionswerten genutzt werden. Ist etwa die  $(n+1)$ -te Ableitung von  $f : [a, b] \rightarrow \mathbb{R}$  stetig, und  $C = \max_{x \in [a, b]} |f^{(n+1)}(x)|$ , so ist der Fehler beschränkt durch

$$|f(x) - T_{n,f,x_0}(x)| \leq \frac{C(b-a)}{(n+1)!}.$$

**Beispiel 1.5.34.** Wir betrachten  $f(x) = e^{-x}$  und wollen die Zahl  $\frac{1}{e}$  auf vier Nachkommastellen genau berechnen.

Wir haben in diesem Fall

$$T_{n,f,0}(x) = \sum_{k=0}^n \frac{(-1)^k}{k!} x^k \quad \text{und} \quad R_{n+1,f,0}(1) = (-1)^{n+1} \frac{e^{-1}}{(n+1)!}$$

für ein  $y \in [0, 1]$ . Damit ist der maximale Fehler den wir machen, wenn wir  $f$  durch sein Taylorpolynom ersetzen,

$$\left| \frac{1}{e} - T_{n,f,0}(1) \right| \leq \frac{1}{(n+1)!}.$$

Wir sind mit unserer Näherung sicher auf 4 Nachkommastellen genau, wenn der Fehler kleiner als  $10^{-5}$  ist. Das erreichen wir, wenn wir  $n$  so wählen, dass  $(n+1)! > 10^5$ , also  $n+1 > 8$ . Mit  $n = 8$  erhalten wir

$$T_{8,f,0}(1) = 0.3678|5714\dots$$

Wenn wir das mit  $\frac{1}{e} = 0.3678|7944\dots$  und zusätzlich mit  $T_{7,f,0}(1) = 0.3678|5714\dots$  und  $T_{6,f,0}(1) = 0.3680|5555\dots$  vergleichen, so sehen wir, dass das gewählte  $n$  nicht optimal ist: bereits  $n = 7$  liefert die gewünschte Genauigkeit.

**Satz 1.5.35.** 1. Es sei  $f : I \rightarrow \mathbb{R}$  eine  $n$ -mal differenzierbare Funktion und  $p$  ein Polynom höchstens  $n$ -ten Grades mit  $f^{(k)}(x_0) = p^{(k)}(x_0)$  für alle  $k = 0 \dots, n$ . Dann gilt

$$T_{n,f,x_0}(x) = p(x).$$

2. Es sei  $f : I \rightarrow \mathbb{R}$  eine  $n$ -mal differenzierbare Funktion und  $p$  ein Polynom höchstens  $n$ -ten Grades. Weiter gebe es eine Funktion  $g$  mit  $g(x_0) = 0$ , sodass  $f$  die Darstellung

$$f(x) = p(x) + g(x)(x - x_0)^n.$$

hat. Dann gilt

$$T_{n,f,x_0}(x) = p(x).$$

Wegen des Satzes von Taylor 1.5.31 gilt in 2. die Rückrichtung auch: setze  $g(x) = \frac{x \cdot f^{(n+1)}(y(x))}{(n+1)!}$ .

**Folgerung 1.5.36.** 1. Aus der Verträglichkeit der Differentiation mit der Summenbildung und der Multiplikation mit Skalaren folgt, dass die Zuordnung  $f \mapsto T_{n,f,x_0}$  linear ist. Das heißt für Funktionen  $f, g$  und Zahlen  $\alpha \in \mathbb{R}$  gilt

$$\begin{aligned} T_{n,f+g,x_0} &= T_{n,f,x_0} + T_{n,g,x_0}, \\ T_{n,\alpha f,x_0} &= \alpha T_{n,f,x_0}. \end{aligned}$$

2. Es seien für  $k \in \mathbb{N}$  und  $\alpha \in \mathbb{R}$  die Funktionen  $q_k : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$q_k(x) = x^k \quad \text{und} \quad \rho_\alpha(x) := \alpha x$$

gegeben. Nun sei  $T_{n,f,0}$  das Taylorpolynom von  $f$  im Entwicklungspunkt  $x_0 = 0$ . Dann gilt für die Funktionen  $f \circ q_k$  und  $f \circ \rho_\alpha$

$$\begin{aligned} T_{n,f \circ \rho_\alpha,0} &= T_{n,f,0} \circ \rho_\alpha, \\ T_{kn,f \circ q_k,0} &= T_{n,f,0} \circ q_k. \end{aligned}$$

Insbesondere ist im zweiten Fall zusätzlich

$$T_{kn+\ell,f \circ q_k,0} = T_{kn,f \circ q_k,0} \quad \text{für } \ell = 1, \dots, k-1.$$

Diese Eigenschaften folgen aus dem zweiten Teil von Satz 1.5.35. Insbesondere ist hier die Einschränkung auf den Entwicklungspunkt  $x_0 = 0$  wichtig.

Ein Nachweis dieser Eigenschaften ist auch direkt möglich: die erste Aussage folgt aus einer einfachen Rechnung, die zweite Aussage ist wesentlich aufwändiger und die Formel von Faá di Bruno kann hier hilfreich sein.

**Beispiel 1.5.37.** 1. Das  $n$ -te Taylorpolynom eines Polynoms  $n$ -ten Grades ist das Polynom selbst und das Restglied verschwindet.

2. Das  $n$ -te Taylorpolynom und das Restglied der Exponentialfunktion  $x \mapsto e^x$  an der Stelle  $x_0 = 0$  ist

$$T_{n,f,0}(x) = \sum_{k=0}^n \frac{x^k}{k!} \quad \text{und} \quad R_{n,f,0}(x) = \frac{e^y}{(n+1)!} x^{n+1}.$$

Man sieht, dass dieses Restglied für beliebige  $x$  immer kleiner wird (unabhängig davon, wie  $y$  aussieht). Damit ist das Taylorpolynom eine gute Näherung für die Funktion selbst. Insbesondere gilt

$$e = \lim_{n \rightarrow \infty} T_{n,f,0}(1) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!}.$$

3. Die Funktion  $f(x)$  mit

$$x \mapsto \begin{cases} 0; & x \in \mathbb{R}^{\leq 0} \\ e^{-\frac{1}{x^2}}; & x \in \mathbb{R}^{> 0} \end{cases}$$

ist in  $x_0 = 0$  beliebig oft differenzierbar mit  $f^{(k)}(x_0) = 0$ . Damit ist das Taylorpolynom  $n$ -ter Ordnung durch die Nullfunktion gegeben,  $T_{n,f,0}(x) = 0$ . Da  $\lim_{x \rightarrow \infty} f(x) = 1$  ist, muss für große  $x$  das Restglied ebenfalls näherungsweise 1 sein.

Für das Restglied gilt

$$R_{n,f,0}(x) = e^{-\frac{1}{y^2}} \frac{p_{2n}(y)}{y^{3(n+1)}} x^{n+1}$$

für ein Polynom  $p_{2n}$  vom Grad  $2n$ . Bei genauerer Betrachtung des Restgliedes sieht man, dass  $p_{2n}$  ein gerades Polynom ist, seine Koeffizienten sehr stark mit  $n$  wachsen und die Vorzeichen der Koeffizienten alternieren. Genauer gilt z.B. für den Leitkoeffizienten und das absolute Glied

$$p_{2n}(y) = (-1)^n (n+2)y^{2n} \mp \dots + \frac{2^{n+1}}{(n+1)!}$$

Diese Polynome haben genau  $n$  positive und damit auch  $n$  negative Nullstellen.

Da der Term  $x^{n+1}$  für  $x > 1$  bei  $n \rightarrow \infty$  gegen  $\infty$  strebt, muss  $y$  sich mit steigendem  $n$  einer Nullstelle von  $p_{2n}$  nähern, damit  $R_{n,f,0}(x)$  gegen einen endlichen Wert streben kann, siehe Tabelle 3.

Mit Hilfe des Satzes von Taylor können wir die Aussage aus Satz 1.5.18 verbessern

Tabelle 3: Nullstellen von  $p_{2n}$ 

| $n$ | Lösung $y$ von $f(x) = R_{n,f,0}(x)$ | Nullstelle von $p_{2n}$ |
|-----|--------------------------------------|-------------------------|
| 1   | 0.7703946416                         | 0.8164965809            |
| 2   | 0.5181291883                         | 0.5208516546            |
| 3   | 0.4030427309                         | 0.4032669285            |
| 4   | 0.4874825731                         | 0.4874306605            |
| 5   | 0.5626559372                         | 0.5626757648            |
| 6   | 0.4506953519                         | 0.4506960946            |
| 7   | 0.5004738241                         | 0.5004735728            |
| 8   | 0.5483631114                         | 0.5483632168            |
| 9   | 0.4551422528                         | 0.4551422554            |
| 10  | 0.4901515813                         | 0.4901515804            |
| 11  | 0.5244074428                         | 0.5244074432            |
| 12  | 0.4472533403                         | 0.4472533403            |

**Satz 1.5.38.** *Es sei  $f : I \rightarrow \mathbb{R}$  eine mehrfach stetig differenzierbare Funktion und es sei  $x_0 \in I$  ein Extremum von  $f$ , d.h. es gilt  $f'(x_0) = 0$ , wobei  $x_0$  kein Randpunkt ist.*

*Ist nun  $f''(x_0) = \dots = f^{(k-1)}(x_0) = 0$  und  $f^{(k)}(x_0) \neq 0$ , dann gilt Folgendes:*

- *Ist  $k$  ungerade, so hat  $f$  in  $x_0$  kein lokales Extremum.*
- *Ist  $k$  gerade, so hat  $f$  in  $x_0$  ein Extremum; und zwar*
  - *ein lokales Maximum, wenn  $f^{(k)}(x_0) < 0$ , und*
  - *ein lokales Minimum, wenn  $f^{(k)}(x_0) > 0$ .*

**Beweisskizze.** Der Satz von Taylor liefert in diesem Fall  $f(x) - f(x_0) = \frac{1}{k!} f^{(k)}(y)(x - x_0)^k$  für ein  $y$  zwischen  $x$  und  $x_0$ . Wir benutzen nun, dass aus Stetigkeitsgründen  $f^{(k)}(x)$  und  $f^{(k)}(x_0)$  und damit auch  $f^{(k)}(y)$  das gleiche Vorzeichen haben. Ein Durchgehen der Fälle liefert nun das Gewünschte.  $\square$

**Bemerkung 1.5.39** (Funktionsdiskussion). Die Extrema einer Funktion  $f$  sind nützliche Objekte, wenn man den Verlauf des Graphen  $G_f$  untersuchen möchte.

- $f$  hat in  $x_0$  ein Extremum  $\leftrightarrow (x_0, f(x_0))$  Extrempunkt  $\leftrightarrow f'(x_0) = \dots = f^{(2k-1)}(x_0) = 0$  und  $f^{(2k)}(x_0) \neq 0$  für ein  $k > 0$ .

Weiter nützliche Punkte sind die Schnittpunkte des Graphen mit den Koordinatenachsen:

- Schnittpunkte mit der  $x$ -Achse  $\leftrightarrow$  Nullstellen von  $f$
- Schnittpunkt mit der  $y$ -Achse  $\leftrightarrow (0, f(0))$ .

Daneben ist das "Krümmungsverhalten" des Graphen interessant. Dabei heißt der Graph *linksgekrümmt* (*rechtsgekrümmt*), wenn man beim Durchlaufen des Graphen in positive  $x$ -Richtung eine Linkskurve (Rechtskurve) macht. Ein Punkt des Graphen, in dem sich das Krümmungsverhalten ändert heißt, *Wendepunkt* und der zugehörige  $x$ -Wert *Wendestelle*.

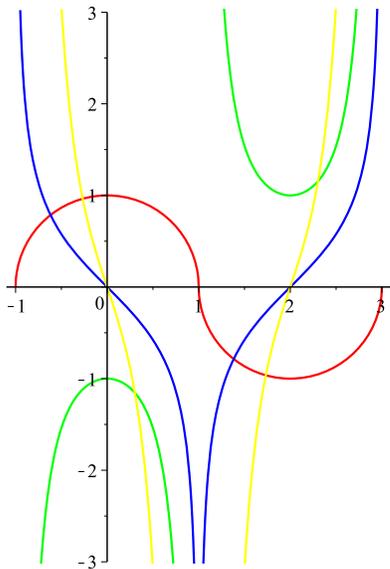
- $G_f$  in  $x_0$  linksgekrümmt (rechtsgekrümmt)  $\leftrightarrow$  Die Steigung der Kurve steigt (fällt)  $\leftrightarrow f''(x_0) > 0$  ( $f''(x_0) < 0$ ).
- $(x_0, f(x_0))$  ist Wendepunkt  $\leftrightarrow f'$  hat in  $x_0$  ein Extremum  $\leftrightarrow f''(x_0) = \dots = f^{(2k)}(x_0) = 0$  und  $f^{(2k+1)}(x_0) \neq 0$  für ein  $k > 0$ .

Ebenfalls hilfreich für die Untersuchung des Verlaufs des Graphen ist die Bestimmung von möglicherweise vorhandenen Asymptoten.

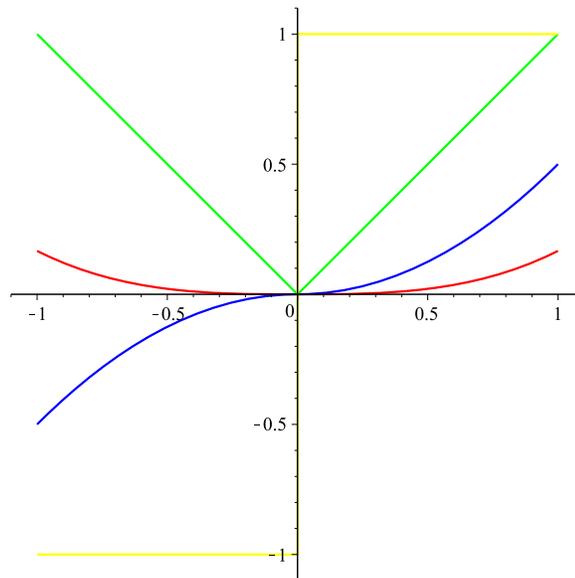
**Beispiel 1.5.40.** In Abbildung 1.5.3 haben wir die Graphen dreier Abbildungen zusammengestellt, die eine "Verklebestelle" haben. Obwohl alle Klebestellen optisch "glatt" sind, verhalten sie sich bezüglich Differentiation gänzlich unterschiedlich; eingezeichnet sind in der Abbildung jeweils  $f(x)$  rot,  $f'(x)$  blau,  $f''(x)$  grün und  $f'''(x)$  gelb.

Abbildung 1.5.3: Typen von "Glattheit"

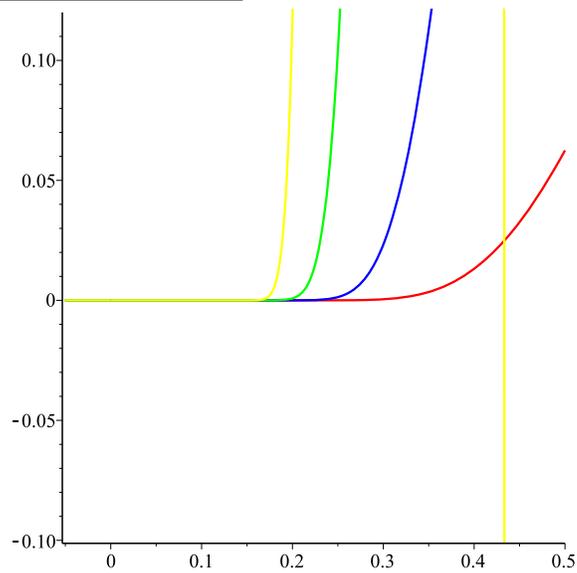
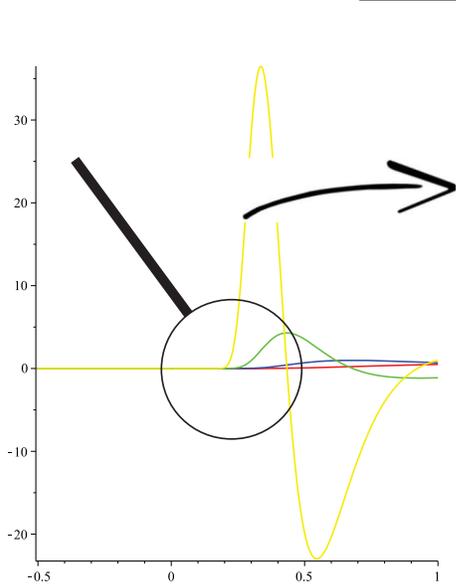
$$f(x) = \begin{cases} \sqrt{1-x^2} & x \in [-1, 1] \\ -\sqrt{1-(x-2)^2} & x \in ]1, 3] \end{cases}$$



$$f(x) = |x^3|$$



$$f(x) = \begin{cases} 0 & x \leq 0 \\ 2^{-\frac{1}{x^2}} & x > 0 \end{cases}$$



## 1.6 Zahlenreihen und Potenzreihen

### 1.6.1 Zahlenreihen

Wir hatten als elementares Beispiel zur Induktion die geometrische Summenformel bewiesen:

$$1 + q + q^2 + \dots + q^n = \frac{1 - q^{n+1}}{1 - q}.$$

Außerdem haben wir uns die Zahlenfolge  $(q^n)_{n \in \mathbb{N}}$  angeschaut, und gesehen, dass diese Folge für  $|q| < 1$  konvergiert mit Grenzwert  $\lim_{n \rightarrow \infty} q^n = 0$ . Wendet man die Regeln im Umgang mit Grenzwerten an, so sehen wir dass für die rechte Seite der geometrischen Summenformel

$$\lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}$$

gilt. Wir interpretieren nun die linke Seite ebenfalls als Zahlenfolge  $(s_n)_{n \in \mathbb{N}}$  mit

$$s_n := \sum_{k=0}^n q^k.$$

Nutzen wir die obige Diskussion, so sehen wir, dass die Folge  $(s_n)$  konvergent mit Grenzwert

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{1 - q}.$$

Auf ähnliche Weise haben wir die Summe

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots$$

untersucht und gesehen, dass diese über alle Grenzen wächst. Schreiben wir ähnlich wie gerade

$$s_n := \sum_{k=1}^n \frac{1}{k}$$

so bedeutet das, dass die Folge  $(s_n)_{n \in \mathbb{N}^*}$  unbeschränkt ist.

**Definition 1.6.1.** 1. Es sei  $(a_n)_{n \in \mathbb{N}}$  eine Zahlenfolge und es sei  $(s_n)_{n \in \mathbb{N}}$  die Folge, die durch Aufsummieren der einzelnen Folgenglieder entsteht:

$$s_n := \sum_{k=0}^n a_k.$$

Dann nennen wir die Zahlenfolge  $(s_n)_{n \in \mathbb{N}}$  eine *Zahlenreihe* und wir schreiben auch

$$(s_n)_{n \in \mathbb{N}} =: \sum a_k.$$

Das Folgenglied  $s_n$  heißt auch die *n-te Teilsumme der Reihe*.

2. Wenn der Grenzwert der Folge  $(s_n)$  existiert, und gilt  $\lim_{n \rightarrow \infty} s_n = r$ , dann heißt die Reihe *konvergent* und wir schreiben

$$\sum_{k=0}^{\infty} a_k = r.$$

Einige einfache Folgerungen aus dem Umgang mit Zahlenfolgen und ihren Eigenschaften sammeln wir in der folgenden Bemerkung.

**Bemerkung 1.6.2.** 1. Eine Reihe  $\sum a_k$  ist genau dann (streng) monoton wachsend, wenn  $a_n \geq 0$  ( $a_n > 0$ ) für alle  $n \in \mathbb{N}$

2. (Cauchy-Konvergenzkriterium). Eine Reihe ist genau dann konvergent, wenn die Teilsummen eine Cauchyfolge bilden. Das heißt:

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n > n_0 \forall \ell \geq 0 : |s_{n+\ell} - s_n| = \left| \sum_{k=n}^{n+\ell} a_k \right| < \epsilon$$

3. Man kann endlich viele der Folgenglieder der Folge  $(a_n)$  ändern ohne die Konvergenzeigenschaft der Reihe  $\sum a_k$  zu beeinflussen.

4. Existiert  $\sum_{k=0}^{\infty} a_k$ , so existiert auch  $\sum_{k=m}^{\infty} a_k$  für jedes  $m \in \mathbb{N}$ .

5. (Monotoniekriterium). Ist  $a_n > 0$  für alle  $n$ , so konvergiert die Reihe  $\sum a_n$  genau dann, wenn die Folge  $(s_n)$  der Teilsummen beschränkt ist.

**Satz 1.6.3.** Eine Reihe  $\sum a_k$  ist höchstens dann konvergent, wenn  $(a_n)$  eine Nullfolge bildet.

**Beispiel 1.6.4.** 1. Die geometrische Reihe  $\sum q^k$  ist für  $|q| < 1$  konvergent und für  $|q| \geq 1$  divergent.

2. Die harmonische Reihe  $\sum \frac{1}{k}$  divergiert.

3. Die Reihe  $\sum \frac{1}{k(k+1)}$  konvergiert mit  $\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1$ .

4. Die Reihe  $\sum \frac{1}{k!}$  konvergiert und es ist

$$\sum_{k=0}^{\infty} \frac{1}{k!} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

Für die Teilsummen von  $\sum \frac{1}{k!}$  gilt

$$s_n = \sum_{k=0}^n \frac{1}{k!} \leq 1 + \sum_{k=1}^n \left(\frac{1}{2}\right)^{k-1} < 1 + \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = 3,$$

wegen  $k! \geq 2^{k-1}$  für  $k \geq 1$ . Damit erfüllt  $\sum \frac{1}{k!}$  die Voraussetzungen aus Bemerkung 1.6.2.2 und es sei  $s = \lim_{n \rightarrow \infty} s_n = \sum_{k=0}^{\infty} \frac{1}{k!}$ .

Weiter wissen wir aus den Übungen, dass die Folge  $(a_n)$  mit  $a_n = \left(1 + \frac{1}{n}\right)^n$  konvergiert, etwa  $\lim_{n \rightarrow \infty} a_n = a$ .

Nun gilt für alle  $n < m$

$$a_m = \sum_{k=0}^m \binom{m}{k} \frac{1}{m^k} \leq s_m, \quad a_m \geq 2 + \sum_{k=2}^n \frac{1}{k!} \prod_{\ell=1}^{k-1} \left(1 - \frac{\ell}{m}\right)$$

damit ist  $a \leq s$  und  $a \geq s_n$  also  $a \geq s$ . Zusammen erhalten wir  $A = s$ .

**Folgerung 1.6.5.** Mit Definition/Bemerkung 1.5.14 und Beispiel 1.5.37.2 haben wir damit die folgenden Darstellungen für die Eulersche Zahl  $e$ :

$$e = 2^{\lim_{n \rightarrow \infty} \frac{1}{n(\sqrt[2]{2}-1)}} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^{\infty} \frac{1}{k!}.$$

**Definition 1.6.6.** Eine Zahlenreihe  $\sum a_n$  heißt *absolut konvergent*, wenn die Reihe  $\sum |a_k|$  konvergiert.

**Folgerung 1.6.7.** 1. Eine absolut konvergente Reihe ist sicher konvergent. Die Dreiecksungleichung für die jeweiligen Teilsummen überträgt sich auch auf den Grenzwert:

$$\left| \sum_{k=0}^{\infty} a_k \right| \leq \sum_{k=0}^{\infty} |a_k|.$$

2. Eine Reihe  $\sum a_k$  konvergiert genau dann absolut, wenn die Folge der Summen  $\sum_{k=0}^n |a_k|$  beschränkt ist.

Ähnlich wie mit konvergenten Folgen, kann man auch mit konvergenten Reihen rechnen und der Beweis verläuft völlig analog. Die Grenzwerte verhalten sich dann so wie ihre Teilsummen:

**Satz 1.6.8.** *Es seien  $\sum a_k$  und  $\sum b_k$  konvergente Reihen. Dann sind auch die Reihen  $\sum(a_k + b_k)$ ,  $\sum(a_k - b_k)$  und für  $\alpha \in \mathbb{R}$  auch die Reihe  $\sum(\alpha a_k)$  konvergent mit*

$$\sum_{k=0}^{\infty} (a_k \pm b_k) = \sum_{k=0}^{\infty} a_k \pm \sum_{k=0}^{\infty} b_k, \quad \sum_{k=0}^{\infty} (\alpha a_k) = \alpha \sum_{k=0}^{\infty} a_k.$$

**Satz 1.6.9.** *Ist  $\sum a_k$  eine konvergente Reihe, so darf man die  $a_k$  beliebig zusammenfassen, wenn man ihre Reihenfolge beibehält.*

*Das heißt, ist  $0 \leq n_0 < n_1 < \dots$  und die Folge  $(b_n)$  gegeben durch*

$$b_0 := a_0 + \dots + a_{n_0} \quad \text{und} \quad b_k = a_{n_{k-1}+1} + \dots + a_{n_k} \quad \text{für} \quad k \geq 1,$$

*so ist  $\sum b_k$  konvergent mit*

$$\sum_{k=0}^{\infty} b_k = \sum_{k=0}^{\infty} a_k.$$

**Bemerkung 1.6.10.** Umgekehrt darf man vorhandene Klammerungen in einer konvergenten Reihe nicht unbedingt weglassen; Man darf das genau dann tun, wenn die daraus resultierende Reihe selbst konvergiert.

Dass man in einer Reihe nicht beliebig klammern oder ausklammern darf, führt direkt zur Frage, wie sich eine Reihe bei einer "Umsortierung" der Glieder verhält.

**Definition 1.6.11.** Es sei  $f : \mathbb{N} \rightarrow \mathbb{N}$  eine bijektive Abbildung und  $\sum a_k$  sei eine Zahlenreihe. Dann nennt man die Reihe  $\sum a_{f(k)}$  eine *Umsortierung* der Ausgangsreihe.

**Satz 1.6.12.** 1. Ist  $\sum a_k$  eine absolut konvergente Reihe, so konvergiert jede Umsortierung  $\sum a_{f(k)}$  und es gilt  $\sum_{k=0}^{\infty} a_{f(k)} = \sum_{k=0}^{\infty} a_k$ .

2. (Riemannscher Umordnungssatz). Ist  $\sum a_k$  eine konvergente aber nicht absolut konvergente Reihe, dann gibt es zu jeder Zahl  $r \in \mathbb{R}$  eine Umsortierung  $\sum a_{f(k)}$ , sodass  $\sum_{k=0}^{\infty} a_{f(k)} = r$ .

Wir haben uns oben schon für Summen und Differenzen von Reihen interessiert. Daher ist die Frage nach Produkten ebenfalls natürlich.

**Definition 1.6.13.** Es seien  $(a_k)$  und  $(b_k)$  Folgen und es sei  $\mathcal{K} = \{N_0, N_1, \dots\}$  eine beliebige Klasseneinteilung der Menge  $\{a_i b_j \mid i, j \in \mathbb{N}\}$  wobei jedes  $N_k$  endlich ist. Weiter sei  $(p_n)$  die Folge mit  $p_n = \sum_{z \in N_n} z$ . Dann nennt man  $\sum p_k$  eine *Produktreihe* der Reihen  $\sum a_k$  und  $\sum b_k$ .

**Satz 1.6.14.** Die Reihen  $\sum a_k$  und  $\sum b_k$  seien absolut konvergent und es sei  $\sum p_k$  eine ihrer Produktreihen. Dann ist  $\sum p_k$  absolut konvergent und es ist

$$\sum_{k=0}^{\infty} p_k = \left( \sum_{k=0}^{\infty} a_k \right) \left( \sum_{k=0}^{\infty} b_k \right).$$

**Beweisskizze.** Es reicht wegen Satz 1.6.12 eine spezielle Produktreihe zu betrachten. Wir wählen diejenige aus der abschließenden Folgerung 1.6.15,

$$\text{also } p_k = \sum_{j=0}^k a_j b_{k-j}.$$

Wir setzen  $s_n^a := \sum_{k=0}^n |a_k|$  und  $s_n^b := \sum_{k=0}^n |b_k|$ , sowie  $s_a := \lim_{n \rightarrow \infty} s_n^a = \sum_{k=0}^{\infty} |a_k|$  und

$s_b := \lim_{n \rightarrow \infty} s_n^b = \sum_{k=0}^{\infty} |b_k|$ . Dann ist  $\sum |p_k|$  absolut konvergent, denn

$$\sum_{k=n}^{n+\ell} |p_k| = \sum_{k=n}^{n+\ell} \left| \sum_{j=0}^k a_j b_{k-j} \right| \leq \sum_{k=n}^{n+\ell} \sum_{j=0}^k |a_j| |b_{k-j}| = \left( \sum_{k=n}^{n+\ell} |a_k| \right) \left( \sum_{k=n}^{n+\ell} |b_k| \right).$$

Es gilt zusätzlich  $\sum_{k=0}^{\infty} p_k = s_a s_b$ , denn

$$\left| \sum_{k=0}^n p_k - s_a s_b \right| = \left| \sum_{j=0}^k a_j b_{k-j} - s_a s_b \right| = \left| \sum_{k=0}^n a_k \sum_{k=0}^n b_k - s_a s_b \right| = |s_n^a s_n^b - s_a s_b|.$$

□

**Folgerung 1.6.15.** Sind  $\sum a_k$  und  $\sum b_k$  absolut konvergent, so gilt

$$\sum_{k=0}^{\infty} \left( \sum_{\ell=0}^k a_{\ell} b_{k-\ell} \right) = \left( \sum_{k=0}^{\infty} a_k \right) \left( \sum_{k=0}^{\infty} b_k \right).$$

### 1.6.2 Konvergenzkriterien für Zahlenreihen

Aus den Eigenschaften der Folge  $(x_n)$  bzw. aus ihr gewonnener Folgen, lassen sich oft Rückschlüsse auf die Konvergenz einer zugeordneten Reihe ziehen. Das liefert dann aber in der Regel keine Aussage über den Grenzwert der Reihe.

**Satz 1.6.16** (Leibniz-Kriterium). *Ist  $(a_n)$  eine monoton fallende Nullfolge, so konvergiert die alternierende Reihe  $\sum (-1)^k a_k$ .*

**Folgerung 1.6.17.** Zu jeder Zahl  $r \in \mathbb{R}$  gibt es eine Umsortierung der alternierenden harmonischen Reihe  $\sum \frac{(-1)^k}{k}$ , die gegen  $r$  konvergiert.

**Satz 1.6.18** (Majorantenkriterium). *Es sei  $\sum c_k$  eine konvergente Reihe mit  $c_n \geq 0$  für alle  $n \in \mathbb{N}$ . Weiter sei  $(a_n)$  eine Folge und es geben ein  $n_0 \in \mathbb{N}$ , mit  $|a_n| \leq c_n$  für alle  $n \geq n_0$ . Dann ist die Reihe  $\sum a_k$  ebenfalls konvergent.*

**Satz 1.6.19.** *Es sei  $\sum a_k$  eine Reihe.*

1. (Wurzelkriterium) *Gibt es  $n_0 \in \mathbb{N}$  und eine Zahl  $0 < \eta < 1$  mit*

$$\sqrt[n]{|a_n|} \leq \eta$$

*für alle  $n \geq n_0$ , so konvergiert die Reihe  $\sum a_k$  absolut.*

2. (Quotientenkriterium) *Gibt es  $n_0 \in \mathbb{N}$  und eine Zahl  $0 < \eta < 1$ , sodass  $a_n \neq 0$  und*

$$\left| \frac{a_{n+1}}{a_n} \right| \leq \eta$$

*für alle  $n \geq n_0$ , so konvergiert  $\sum a_k$  absolut.*

3. Gilt in beiden Fällen jedoch unendlich oft " $\geq 1$ ", so ist  $\sum x_k$  divergent.

**Folgerung 1.6.20.** Es sei  $\sum a_k$  eine Zahlenreihe.

1. (Minorantenkriterium) Ist  $\sum c_k$  eine divergente Reihe mit  $c_n \geq 0$  für alle  $n \in \mathbb{N}$ , und ist ab einem Index  $n_0$  stets  $x_n \geq c_n$ , so ist  $\sum a_k$  ebenfalls divergent
2. Gilt  $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1$  oder  $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1$ , so konvergiert  $\sum a_n$  absolut.
3. Gilt  $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1$  oder  $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1$ , so divergiert  $\sum a_n$ .
4. Ist der Limes = 1, so kann man keine Entscheidung treffen

**Satz 1.6.21** (Verdichtungskriterium). Die Folge  $(a_n)$  sei nicht-negativ und monoton fallend. Dann ist  $\sum a_k$  genau dann konvergent, wenn die Reihe  $\sum 2^k a_{2^k}$  konvergent ist.

**Beispiel 1.6.22.** •  $\sum \frac{1}{k^\alpha}$  und  $\sum \frac{1}{k(\log_{10} k)^\alpha}$  sind genau dann konvergent, wenn  $\alpha > 1$  ist.

- $\sum \frac{k!}{k^k}$  ist konvergent.
- $\sum \frac{2^k}{k!}$  ist konvergent.

### 1.6.3 Potenzreihen

Wir haben in Satz 1.5.31 gesehen, dass die Taylorentwicklung einer reellen Funktion  $f$  im Punkt  $x_0$  die Form  $f(x) = \sum_{k=0}^n a_k (x - x_0)^k + R_n(x)$  hat, wobei

$a_k = \frac{f^{(k)}(x_0)}{k!}$ . Wir nehmen jetzt an, dass  $f$  beliebig oft differenzierbar und  $x_0 = 0$  ist<sup>(i)</sup>. Es ist also für jedes  $n \in \mathbb{N}$

$$f(x) = \sum_{k=0}^n a_k x^k + R_n(x).$$

<sup>(i)</sup>Das dürfen wir ohne Einschränkung annehmen, indem wir zur Funktion  $\tilde{f}$  mit  $\tilde{f}(x) = f(x - x_0)$  übergehen.

Das  $n$ -te Taylorpolynom an einer Stelle  $x \in \mathbb{R}$  lässt sich somit als  $n$ -te Teilsumme der Reihe

$$\sum a_k x^k$$

auffassen.

**Definition 1.6.23.** Es sei  $(a_n)$  eine Zahlenfolge und  $x_0 \in \mathbb{R}$ .

- Die Zuordnung, die jedem Punkt  $x \in \mathbb{R}$  die Zahlenreihe  $\sum a_k(x - x_0)^k$  zuordnet, heißt eine *Potenzreihe* mit Mittelpunkt  $x_0$ .
- Eine Potenzreihe heißt *konvergent in  $x = x_1$* , wenn die Zahlenreihe  $\sum a_k(x_1 - x_0)^k$  konvergiert. Sie heißt *absolut konvergent in  $x = x_1$* , wenn die Zahlenreihe  $\sum |a_k||x_1 - x_0|^k$  konvergiert.

Da jede Potenzreihe in ihrem Mittelpunkt konvergiert, nutzen wir noch folgende Differenzierung

- Eine Potenzreihe heißt *konvergent* wenn sie nicht nur in  $x = x_0$  konvergiert.
- Eine Potenzreihe heißt *divergent*, wenn sie nur in  $x = x_0$  konvergiert.

**Bemerkung 1.6.24.** Zwei Potenzreihen<sup>(i)</sup>  $\sum a_k x^k$  und  $\sum b_k x^k$  sind genau dann gleich, wenn  $a_n = b_n$  für alle  $n \in \mathbb{N}$ . Hierbei heißen zwei Potenzreihen gleich, wenn alle ihre Teilsummen gleich sind.

Aus den Konvergenzkriterien für Reihen erhalten wir sofort die folgende Aussage.

**Satz 1.6.25.** Es sei  $\sum a_k x^k$  eine Potenzreihe.

1. Konvergiert die Potenzreihe in einem Punkt  $x_0 \neq 0$ , dann konvergiert sie in jedem Punkt  $x \in ] -|x_0|, |x_0|[$  absolut.
2. Die Vereinigung  $K \subset \mathbb{R}$  aller offenen Intervalle, in denen die Potenzreihe konvergiert, ist ein um 0 symmetrisches offenes Intervall. Dort konvergiert die Potenzreihe absolut.

---

<sup>(i)</sup>Wir nehmen im Folgenden auch für Potenzreihen an, dass  $x_0 = 0$  ist.

3. Existiert

$$R = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{|a_n|}},$$

dann konvergiert die Potenzreihe in jedem Punkt  $x \in ]-R, R[$  absolut.

4. Existiert

$$R = \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right|,$$

dann konvergiert die Potenzreihe in jedem Punkt  $x \in ]-R, R[$  absolut.

**Bezeichnung 1.6.26.** Das Intervall  $K$  aus 2. heißt *Konvergenzbereich* und die Hälfte seiner Länge heißt *Konvergenzradius*. Der Konvergenzradius lässt sich mit 3. oder 4. berechnen. Hier lassen wir auch  $R = 0, K = \emptyset$  und  $R = \infty, K = \mathbb{R}$  zu.

**Bemerkung 1.6.27.** 1. Konvergiert  $\sum a_k x^k$  in  $\hat{x} \in \mathbb{R}$  nicht, dann gilt  $K \subset ]-\hat{x}, \hat{x}[$ .

2. Es kann sein, dass eine Potenzreihe in keinem, in einem oder in beiden Randpunkten von  $K$  konvergiert.

Z.B. konvergiert  $\sum \frac{x^k}{k}$  nicht in  $x = 1$ , sodass  $K \subset ]-1, 1[$ . Andererseits konvergiert sie in  $x = -1$ , sodass  $] -1, 1[ \subset K$ . Damit ist  $K = ]-1, 1[$  mit Konvergenz in einem Randpunkt.

**Bemerkung 1.6.28.** Betrachtet man eine Potenzreihe der Form  $\sum a_k (x - x_0)^k$ , so bleibt das obige alles richtig, nur dass das Konvergenzbereich um  $x_0$  zentriert ist, also  $K = ]x_0 - R, x_0 + R[$ , wobei  $R$  etwa wie in Satz 1.6.25.3 oder 4 berechnet wird.

**Satz 1.6.29.** Die Potenzreihe  $\sum a_k x^k$  habe den Konvergenzbereich  $K$  mit Konvergenzradius  $R > 0$  und es sei  $x_0 \in K$ . Dann gilt für alle  $x \in K$  mit  $|x - x_0| < R - |x_0|$

$$\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} b_k (x - x_0)^k$$

$$\text{mit } b_k = \sum_{\ell=k}^{\infty} \binom{\ell}{k} a_{\ell} x_0^{\ell-k}.$$

**Beweisskizze.** Die obige Formel wird plausibel, wenn man endliche Summen betrachtet. Ein formaler Beweis der obigen Aussage benutzt Aussagen über die Umsortierung von Doppelreihen, das sind Reihen der Form  $\sum_{i,j=1}^{\infty} a_{ij}$ .  $\square$

**Bezeichnung 1.6.30.** Ist die Potenzreihe  $\sum a_k x^k$  auf  $K \neq \emptyset$  konvergent, so definiert sie eine Funktion  $f : K \rightarrow \mathbb{R}$  mit  $f(x) = \sum_{k=0}^{\infty} a_k x^k$  für alle  $x \in K$ , die so genannte *Grenzfunktion*. Diese bezeichnen wir ebenfalls mit  $\sum a_k x^k$ .

Zum Abschluss dieses Abschnitts wollen wir uns noch mit einigen Eigenschaften der Grenzfunktionen von Potenzreihen beschäftigen.

**Satz 1.6.31.** *Hat die Potenzreihe  $\sum a_k x^k$  einen Konvergenzbereich  $K \neq \emptyset$ , so ist ihre Grenzfunktion dort stetig.*

**Satz 1.6.32.** *Hat die Potenzreihe  $\sum a_k x^k$  einen Konvergenzbereich  $K \neq \emptyset$ , so ist die Grenzfunktion dort beliebig oft differenzierbar, und die Ableitung erhält man durch summandenweises differenzieren, also*

$$\left( \sum a_k x^k \right)' = \sum (k+1) a_{k+1} x^k.$$

*Insbesondere hat die Ableitungsreihe den gleichen Konvergenzradius.*

Wir erhalten die folgende nützliche Aussage.

**Folgerung 1.6.33.** *Jede Potenzreihe ist die Taylorentwicklung ihrer Grenzfunktion.*

**Satz 1.6.34** (Identitätssatz für Potenzreihen). *Es seien  $\sum a_k x^k$  und  $\sum b_k x^k$  zwei Potenzreihen mit dem gleichen Konvergenzbereich  $K \neq \emptyset$ . Stimmen die beiden Grenzfunktionen auf einem noch so kleinen Intervall  $B_r(0) \subset K$  überein, so stimmen sie schon auf ganz  $K$  überein.*

Da Folgerung 1.6.33 in Anwendungen sehr zentral ist, wollen wir auch die „Rückrichtung“ der dortigen Aussage diskutieren.

**Bemerkung 1.6.35.** Es sei  $I \subset \mathbb{R}$  ein offenes Intervall und  $f : I \rightarrow \mathbb{R}$  beliebig oft differenzierbar. Weiter sei  $T_{n,f,x_0}(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$  das Taylorpolynom von  $f$  im Entwicklungsmittelpunkt  $x_0$ . Ihre Taylorentwicklung – also die Potenzreihe, die genau die Taylorpolynome von  $f$  als Teilsummen hat – heißt *Taylorreihe von  $f$* . Dann gilt

- Die Restglieder der Taylorentwicklung konvergieren auf einem Intervall  $J \subset I$  genau dann gegen Null, wenn  $f|_J$  mit der Grenzfunktion der Taylorreihe auf  $J$  übereinstimmt.
- Die Grenzfunktion der Taylorreihe muss an keiner Stelle mit der Ausgangsfunktion  $f$  übereinstimmen (außer im Entwicklungsmittelpunkt  $x_0$ ).

**Beispiel 1.6.36.** 1.  $x \mapsto e^x$  ist auf ganz  $\mathbb{R}$  die Grenzfunktion ihrer Taylorreihe  $\sum \frac{x^k}{k!}$ .

2.  $f : x \mapsto \ln\left(\frac{1}{1+x}\right)$  eingeschränkt auf  $K = ]-1, 1[$  ist die Grenzfunktion ihrer Taylorreihe  $\sum \frac{(-1)^k}{k} x^k$ .

3. Die Taylorreihe der Funktion  $f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$  aus Beispiel

1.5.37 ist die „Nullreihe“. Diese konvergiert offensichtlich auf ganz  $\mathbb{R}$  mit Nullfunktion als Grenzfunktion. Für  $x \leq 0$  stimmt die Grenzfunktion mit  $f$  überein. Genau dort konvergiert auch das Restglied gegen 0.

4. Die leicht abgewandelte Funktion  $f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$  ist in  $x_0 = 0$

ebenfalls beliebig oft differenzierbar mit verschwindenden Ableitungen. Damit ist die Taylorreihe wieder die Nullreihe. Ihre Grenzfunktion stimmt (außer in 0) an keiner Stelle mit  $f$  überein.

**Bemerkung 1.6.37.** In Ergänzung zu Folgerung 1.6.5 erhalten wir die folgenden Darstellungen der Exponentialfunktion:

$$e^x = 2^{\frac{x}{\lim_{n \rightarrow \infty} n(\sqrt[n]{2}-1)}} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

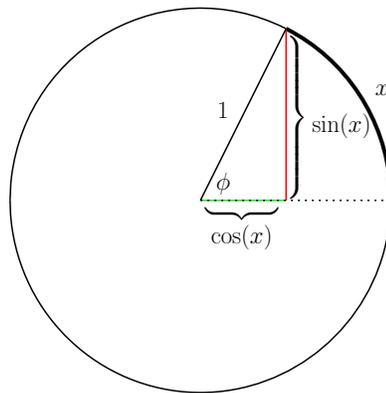
Hier sind die ersten beiden Aussagen leicht zu sehen, für die letzte zeigt man, dass der Ausdruck die Funktionalgleichung und die Normierungsbedingung aus Satz 1.3.30 erfüllt.

## 1.7 Die Winkelfunktionen und die Arkusfunktionen

### 1.7.1 Die Winkelfunktionen

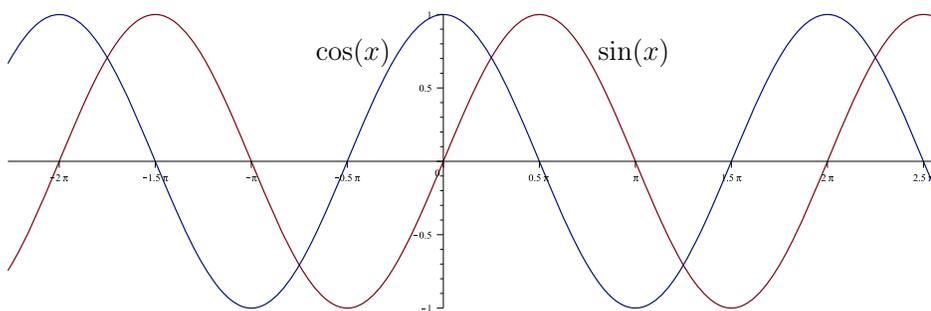
Wir kennen die Definition des Sinus und des Kosinus als Größen, die einem Winkel  $\phi$  in einem rechtwinkligen Dreieck zugeordnet sind,  $\sin(\phi)$ ,  $\cos(\phi)$ . Statt des Winkels  $\phi$  kann man auch die Bogenlänge  $x = \pi \frac{\phi}{180^\circ}$  eines Segments des Einheitskreises betrachten, siehe Abbildung 1.7.1.

Abbildung 1.7.1: Der Einheitskreis



Die  $2\pi$ -periodische Erweiterung<sup>(i)</sup> des Bogenlängenbereichs  $x \in [0, 2\pi)$  auf ganz  $\mathbb{R}$  liefert zwei Funktionen  $\sin : \mathbb{R} \rightarrow \mathbb{R}$  und  $\cos : \mathbb{R} \rightarrow \mathbb{R}$  den Graphen gemäß Abbildung 1.7.2

Abbildung 1.7.2: Sinus und Cosinus



<sup>(i)</sup>Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt  $T$ -periodisch oder periodisch mit Periode  $T$ , wenn  $f(x) = f(x + T)$  für alle  $x \in \mathbb{R}$ .

Mit dieser (rein anschaulichen) "Definition" können wir einige (ebenfalls anschauliche) Eigenschaften ablesen:

[W1] Variiert man die Bogenlänge  $x$  nur ein wenig, so ändern sich die Werte für Sinus und Kosinus ebenfalls nur wenig:

Sinus und Kosinus sind stetige auf ganz  $\mathbb{R}$  definierte Funktionen.

[W2]  $\cos$  ist eine gerade und  $\sin$  ist eine ungerade Funktion.

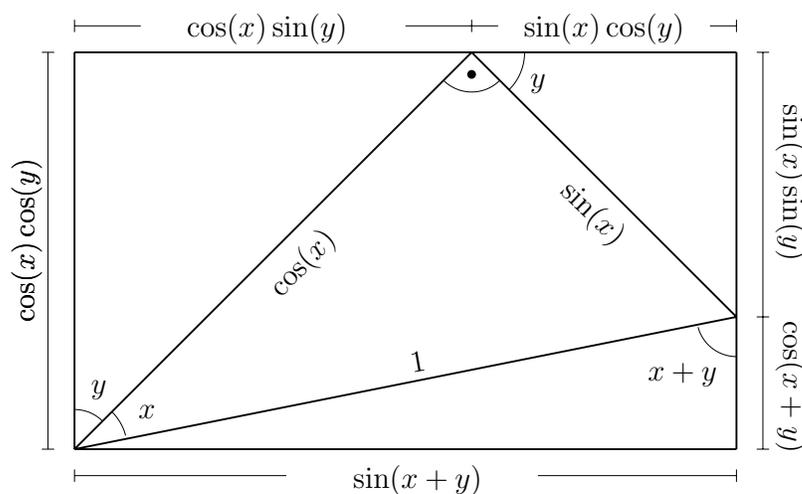
[W3] Es gilt  $\cos(0) = 1$ .

[W4] Wird  $x$  sehr klein, dann weicht die Länge des Bogens nicht weit von dem Lot  $\sin(x)$  ab:

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = 1$$

Wir gehen zurück zu den Betrachtungen an rechtwinkligen Dreiecken und betrachten die folgenden Konstellationen:

Abbildung 1.7.3: Additionstheoreme



Dann sehen wir, dass (zumindest für die Größen in der Grafik)

[W5] die folgenden Additionstheoreme gelten

$$\sin(x+y) = \sin(x) \cos(y) + \cos(x) \sin(y),$$

$$\cos(x + y) = \cos(x) \cos(y) - \sin(x) \sin(y).$$

Wir wissen hier nicht, ob es tatsächlich Funktionen gibt, die die obigen (anschaulichen) Eigenschaften [W1]-[W5] haben. Unabhängig von ihrer Existenz definieren wir dennoch für zwei hypothetische Funktionen wie folgt.

**Definition 1.7.1** (Winkelfunktionen 1). Als Sinus und Kosinus bezeichnen wir zwei Funktionen  $\sin : \mathbb{R} \rightarrow \mathbb{R}$  und  $\cos : \mathbb{R} \rightarrow \mathbb{R}$  mit den Eigenschaften [W1] bis [W5].

**Folgerung 1.7.2** (Eigenschaften der Winkelfunktionen). Es gelten die folgenden Rechenregeln:

1.  $\sin(x - y) = \sin(x) \cos(y) - \cos(x) \sin(y)$   
 $\cos(x - y) = \cos(x) \cos(y) + \sin(x) \sin(y)$
2.  $\sin(2x) = 2 \sin(x) \cos(x)$   
 $\cos(2x) = \cos^2(x) - \sin^2(x)$
3.  $\sin(x) - \sin(y) = 2 \cos\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$   
 $\cos(x) - \cos(y) = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$   
 $\cos(x) + \cos(y) = 2 \cos\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right)$
4.  $\sin^2(x) + \cos^2(x) = 1$
5.  $2 \sin^2\left(\frac{x}{2}\right) = 1 - \cos(x)$   
 $2 \cos^2\left(\frac{x}{2}\right) = 1 + \cos(x)$
6. Der Wertebereich von Sinus und Kosinus ist im Intervall  $[-1, 1]$  enthalten.

**Beweisskizze.** 1 folgt aus [W5] und [W2]. 2 folgt aus [W5] mit  $x = y$ . 3 folgt aus [W5] und 1 für die Argumente  $\frac{x \pm y}{2}$ . 4 folgt aus [W3] und 1 mit  $x = y$ . 5 folgt aus 4 und [W5] mit  $x = \frac{x}{2} + \frac{x}{2}$ . 6 folgt aus 5.  $\square$

**Satz 1.7.3.** *Es ist  $\sin(0) = 0$  und die Winkelfunktionen sind differenzierbar mit Ableitungen*

$$\sin'(x) = \cos(x) \quad \text{und} \quad \cos'(x) = -\sin(x).$$

**Beweisskizze.**  $\sin(0) = 0$  folgt aus 4 und [W3]. Die Ableitungen folgen aus [W1], [W4] und 1.7.2.3.  $\square$

**Definition/Satz 1.7.4.** • *cos besitzt positive Nullstellen.*

- *Wir definieren nun  $\pi$  als das Doppelte der kleinsten positiven Nullstelle von cos.*

**Beweisskizze.** Die Annahme, dass es keine positive Nullstelle gibt führt mit Satz 1.7.3 und Folgerung 1.7.2.2 zum Widerspruch. Wegen [W3] und [W1] muss das Infimum der positiven Nullstellenmenge von cos echt positiv und selbst eine Nullstelle sein, also das Minimum dieser Menge.  $\square$

**Folgerung 1.7.2 (cont.).** Mit Hilfe der kleinsten Nullstelle von cos erhalten wir aus [W3], [W5] und 1.7.2.1-5

$$7. \sin\left(\frac{\pi}{2}\right) = 1, \quad \sin(\pi) = 0, \quad \cos(\pi) = -1$$

$$8. \sin(x + \pi) = -\sin(x) \\ \cos(x + \pi) = -\cos(x)$$

$$9. \sin\left(x + \frac{\pi}{2}\right) = \cos(x) \\ \cos\left(x + \frac{\pi}{2}\right) = -\sin(x)$$

$$10. \sin(\pi + x) = -\sin(\pi - x) \\ \cos(\pi + x) = \cos(\pi - x)$$

$$11. \sin\left(\frac{\pi}{2} + x\right) = \sin\left(\frac{\pi}{2} - x\right) \\ \cos\left(\frac{\pi}{2} + x\right) = -\cos\left(\frac{\pi}{2} - x\right)$$

**Folgerung 1.7.6.** 1. Das Intervall  $[-1, 1]$  ist der Wertebereich von Sinus und Kosinus.

2. Sinus und Kosinus besitzen die folgenden Nullstellenmengen:

$$\mathcal{N}_{\sin} = \{k\pi \mid k \in \mathbb{Z}\}, \quad \mathcal{N}_{\cos} = \left\{ \frac{2k+1}{2}\pi \mid k \in \mathbb{Z} \right\}.$$

**Beweisskizze.** 1 folgt aus [W1] und dem Zwischenwertsatz 1.4.10. Für cos folgt in 2 die Richtung "⊃" aus Folgerung 1.7.2.7-11 und "⊂" mit den gleichen Formeln durch Widerspruch aus der Annahme es gäbe in weiteres zwischen  $\frac{\pi}{2}$  und  $\frac{3\pi}{2}$ .  $\square$

**Satz 1.7.7.** *sin und cos sind die Grenzfunktionen der folgenden Potenzreihen*

$$\sum \frac{(-1)^k}{(2k+1)!} x^{2k+1} \quad \text{und} \quad \sum \frac{(-1)^k}{(2k)!} x^{2k}.$$

Das bedeutet, beide Potenzreihen haben Konvergenzbereich  $\mathbb{R}$  und ihre Grenzfunktionen erfüllen die Punkte [W1]-[W5].

**Beweisskizze.** Dass der Konvergenzbereich der Reihen  $\mathbb{R}$  ist folgt aus dem Quotientenkriterium. Es seien  $s(x)$  und  $c(x)$  die Grenzfunktionen der obigen Reihen. Dass  $s$  und  $c$  [W1]-[W4] erfüllen folgt direkt durch Nachrechnen und den Ergebnissen aus dem Abschnitt über Potenzreihen. [W5] kann man zeigen, indem für die Teilsummen  $s_n(x)$  und  $c_n(x)$  nachweist, dass  $\lim_{n \rightarrow \infty} s_{2n}(x+y) = \lim_{n \rightarrow \infty} (s_n(x)c_n(y) - s_n(y)c_n(x))$  gilt.  $\square$

**Definition 1.7.8** (Winkelfunktionen 2).

1. Der Sinus,  $\sin : \mathbb{R} \rightarrow \mathbb{R}$ , und der Kosinus,  $\cos : \mathbb{R} \rightarrow \mathbb{R}$ , sind durch ihre Potenzreihen gemäß Satz 1.7.7 definiert. Sie haben qualitativ den in der Einführung zu diesem Abschnitt skizzierten Verlauf.
2. Tangens,  $\tan : \mathbb{R} \setminus \mathcal{N}_{\cos} \rightarrow \mathbb{R}$ , und Kotangens,  $\cot : \mathbb{R} \setminus \mathcal{N}_{\sin} \rightarrow \mathbb{R}$ , sind definiert durch

$$\tan(x) := \frac{\sin(x)}{\cos(x)}, \quad \cot(x) := \frac{\cos(x)}{\sin(x)} = \frac{1}{\tan(x)}.$$

3. Sekans,  $\sec : \mathbb{R} \setminus \mathcal{N}_{\sin} \rightarrow \mathbb{R}$ , und Kosekans,  $\csc : \mathbb{R} \setminus \mathcal{N}_{\cos} \rightarrow \mathbb{R}$ , sind definiert durch

$$\sec(x) := \frac{1}{\cos(x)}, \quad \csc(x) := \frac{1}{\sin(x)}.$$

**Bemerkung 1.7.9.** 1. Tangens und Kotangens sind  $\pi$ -periodisch, Sekans und Kosekans sind  $2\pi$ -periodisch und haben Graphen gemäß Abbildung 1.7.4-1.7.5.

2. Die Ableitungen der weiteren Winkelfunktionen sind

$$\tan'(x) = \frac{1}{\cos^2(x)} = \sec^2(x) = 1 + \tan^2(x),$$

$$\cot'(x) = -\frac{1}{\sin^2(x)} = -\csc^2(x) = -(1 + \cot^2(x)),$$

$$\sec'(x) = -\tan(x) \sec(x) = -\sin(x) \sec^2(x),$$

$$\csc'(x) = \cot(x) \csc(x) = \cos(x) \csc^2(x).$$

3. Folgerung 1.7.2 liefert nun weiter Identitäten:

$$\tan\left(\frac{x}{2}\right) = \pm \sqrt{\frac{1 - \cos(x)}{1 + \cos(x)}} = \frac{\sin(x)}{1 + \cos(x)} = \frac{1 - \cos(x)}{\sin(x)},$$

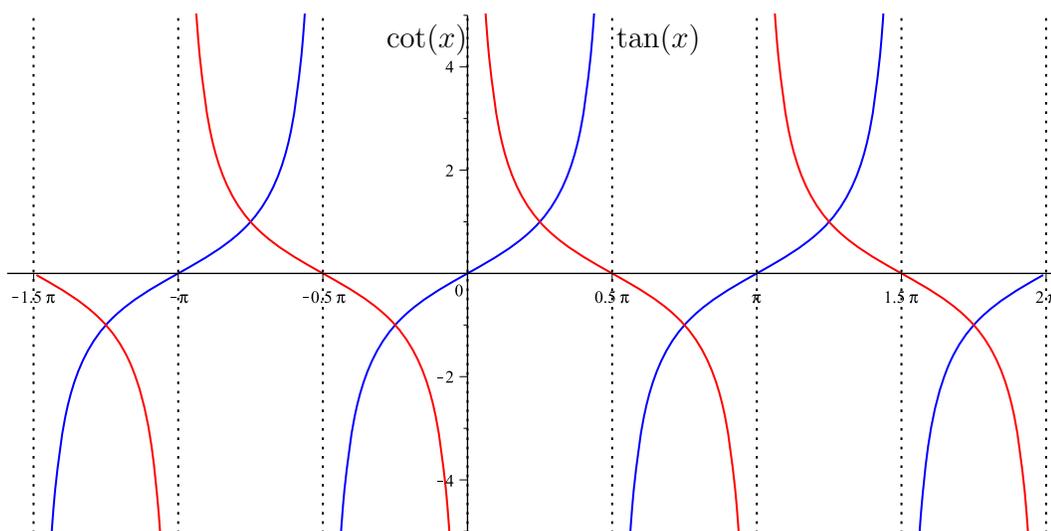
$$\tan(x \pm y) = \frac{\tan(x) \pm \tan(y)}{1 \mp \tan(x) \tan(y)},$$

$$\cot(x \pm y) = \frac{\cot(x) \cot(y) \mp 1}{\cot(x) \pm \cot(y)},$$

$$\tan(x) \tan(y) = \frac{\tan(x) + \tan(y)}{\cot(x) + \cot(y)},$$

$$\tan(x) \cot(y) = \frac{\tan(x) + \cot(y)}{\cot(x) + \tan(y)}.$$

Abbildung 1.7.4: Tangens, Cotangens



**Bemerkung 1.7.10.** Nehmen wir an, dass die hier gefundenen Funktionen tatsächlich die eingangs diskutierte Situation am Einheitskreis widerspiegelt, dann findet man die Größen dort wie in Abbildung 1.7.6.

Abbildung 1.7.5: Sekans und Kosekans

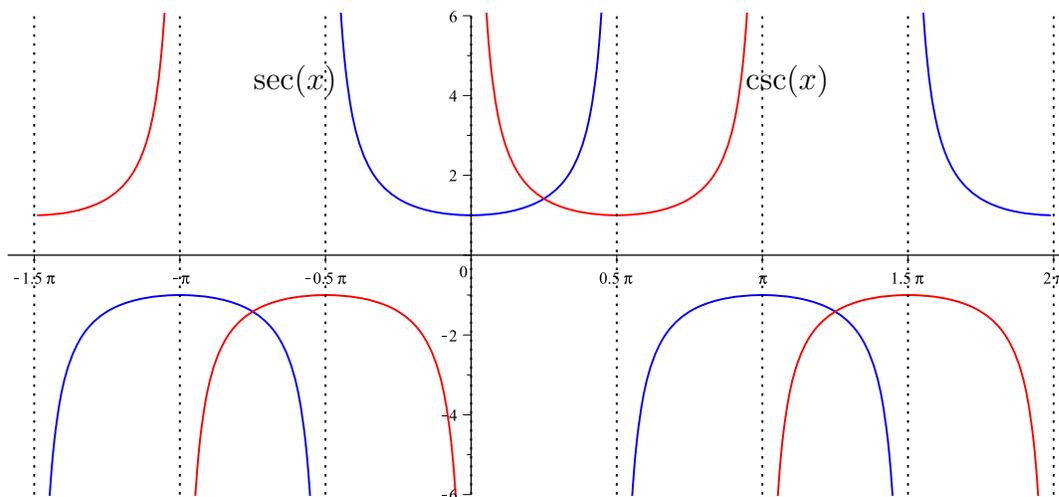
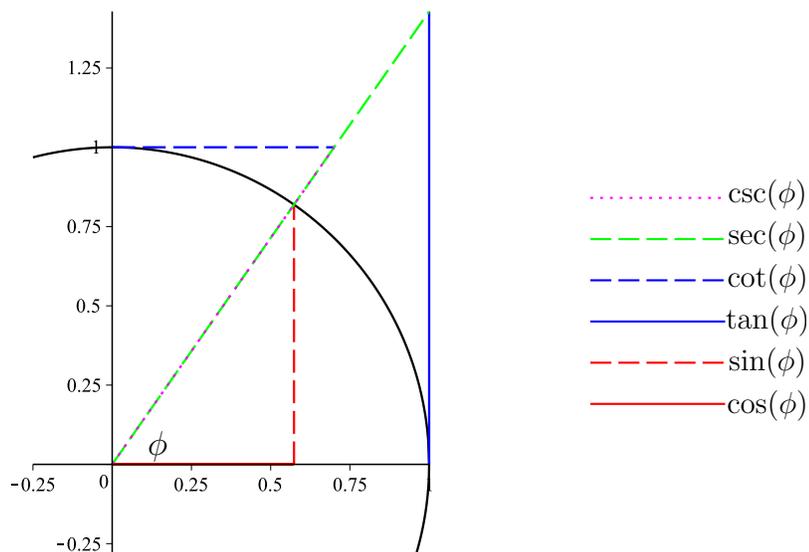


Abbildung 1.7.6: Winkelfunktionen am Kreis



**Satz 1.7.11.** Es sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  definiert durch

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

Dann gelten folgende Aussagen:

- $f$  ist auf  $\mathbb{R}^*$  beliebig oft stetig differenzierbar.
- $x \mapsto f(x)$  ist in  $x = 0$  nicht stetig.<sup>(i)</sup>
- $x \mapsto xf(x)$  ist stetig in  $x = 0$ , aber dort nicht differenzierbar.
- $x \mapsto x^2f(x)$  ist in  $x = 0$  differenzierbar, aber die Ableitung ist dort nicht stetig.
- $x \mapsto x^3f(x)$  ist in  $x = 0$  stetig differenzierbar, aber dort nicht zweimal differenzierbar.

Allgemein gilt für  $k > 0$ :

- $x \mapsto x^{2k}f(x)$  ist in  $x = 0$   $(k - 1)$ -mal stetig differenzierbar und  $k$ -mal differenzierbar, aber die  $k$ -te Ableitung ist dort nicht stetig.
- $x \mapsto x^{2k+1}f(x)$  ist in  $x = 0$   $k$ -mal stetig differenzierbar, aber dort nicht  $(k + 1)$ -mal differenzierbar.

### 1.7.2 Die Arkusfunktionen

Die Winkelfunktionen sind wegen ihrer Periodizität selbstverständlich nicht injektiv und daher nicht invertierbar. Schränkt man sie jedoch auf geeignete Intervalle ein, so erhält man sogar bijektive Funktionen. Ihre Umkehrfunktionen nennt man Arkusfunktionen.

**Definition 1.7.12.** 1. Der *Arkussinus*,

$$\arcsin : [-1, 1] \rightarrow \mathbb{R},$$

ist die Umkehrfunktion des Sinus nachdem man diesen auf das Intervall  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  eingeschränkt hat.

2. Der *Arkuskosinus*,

$$\arccos : [-1, 1] \rightarrow \mathbb{R},$$

ist die Umkehrfunktion des Kosinus, nachdem man diesen auf das Intervall  $[0, \pi]$  eingeschränkt hat.

---

<sup>(i)</sup>Das ist sogar unabhängig von der Wahl der Ergänzung in  $x = 0$

3. Der *Arkustangens*,

$$\arctan : \mathbb{R} \rightarrow \mathbb{R},$$

ist die Umkehrfunktion des Tangens, nachdem man diesen auf das Intervall  $]-\frac{\pi}{2}, \frac{\pi}{2}[$  eingeschränkt hat.

4. Der *Arkuskotangens*,

$$\operatorname{arccot} : \mathbb{R} \rightarrow \mathbb{R},$$

ist die Umkehrfunktion des Kotangens, nachdem man diesen auf das Intervall  $]0, \pi[$  eingeschränkt hat.

5. Der *Arkussekans*,

$$\operatorname{arcsec} : \mathbb{R} \setminus ]-1, 1[ \rightarrow \mathbb{R},$$

ist die Umkehrfunktion des Sekans, nachdem man diesen auf die Menge  $[0, \pi] \setminus \{\frac{\pi}{2}\}$  eingeschränkt hat.

6. Der *Arkuskosekans*,

$$\operatorname{arccsc} : \mathbb{R} \setminus ]-1, 1[ \rightarrow \mathbb{R},$$

ist die Umkehrfunktion des Kosekans, nachdem man diesen auf die Menge  $[-\frac{\pi}{2}, \frac{\pi}{2}] \setminus \{0\}$  eingeschränkt hat.

Die Graphen der Arkusfunktionen sind in Abbildung 1.7.7 skizziert.

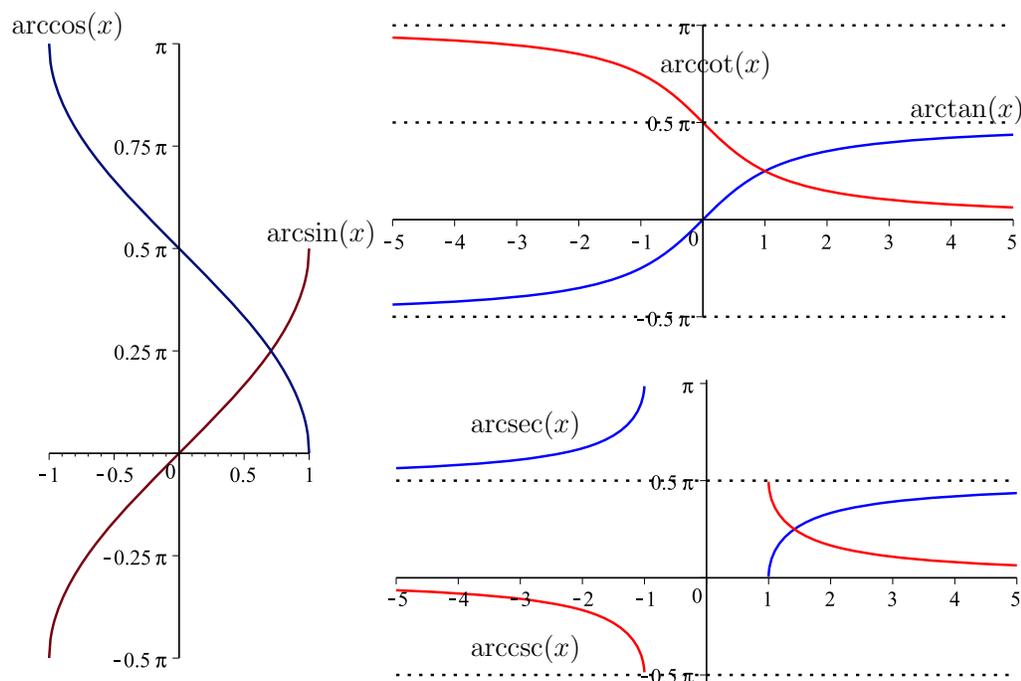
**Bemerkung 1.7.13.** 1. Die Ableitungen der Arkusfunktionen sind

$$\begin{aligned} \arcsin'(x) &= \frac{1}{\sqrt{1-x^2}}, & \arccos'(x) &= -\frac{1}{\sqrt{1-x^2}}, \\ \arctan'(x) &= \frac{1}{1+x^2}, & \operatorname{arccot}'(x) &= -\frac{1}{1+x^2}, \\ \operatorname{arcsec}'(x) &= -\frac{1}{x\sqrt{x^2-1}}, & \operatorname{arccsc}'(x) &= \frac{1}{x\sqrt{x^2-1}}. \end{aligned}$$

Die Ableitungen von arcsin, arccos, arcsec und arccsc existieren in den Punkten  $\pm 1$  nicht.

2. Der Arkustangens bzw. der Arkuskotangens bildet die reelle Achse bijektiv auf das offene Intervall  $]-\frac{\pi}{2}, \frac{\pi}{2}[$  bzw.  $]0, \pi[$  ab.

Abbildung 1.7.7: Die Arkusfunktionen



## 1.8 Anhang Analysis I: Grundbegriffe der formalen Logik

**Definition 1.8.1** (Wahrheitswerte, Aussagen). Eine (*logische*) *Aussage*  $A$  ist eine Behauptung über einen Sachverhalt, der genau einer der beiden *Wahrheitswerte* "wahr" ( $w$ ) oder "falsch" ( $f$ ) zugeordnet werden kann.

**Definition 1.8.2** (Negation, NICHT). Ist  $A$  eine Aussage so nennt man  $\neg A$  die *NEGATION* von  $A$  (man sagt auch "nicht  $A$ "). Sie ist definiert über ihren Wahrheitsgehalt:

Wenn  $A$  falsch ist, dann ist  $\neg A$  wahr.

Wenn  $A$  wahr ist, dann ist  $\neg A$  falsch.

**Definition 1.8.3** (Konjunktion, UND). Sind  $A$  und  $B$  Aussagen, so bezeichnet  $A \wedge B$  die *KONJUNKTION* und man sagt " $A$  und  $B$ ". Sie ist definiert über ihren Wahrheitsgehalt:

Wenn  $A$  und  $B$  wahr sind, dann ist  $A \wedge B$  wahr.

Wenn  $A$  wahr und  $B$  falsch ist, dann ist  $A \wedge B$  falsch.

Wenn  $A$  falsch und  $B$  wahr ist, dann ist  $A \wedge B$  falsch.

Wenn  $A$  und  $B$  falsch sind, dann ist  $A \wedge B$  falsch.

**Definition 1.8.4** (Disjunktion, ODER). Sind  $A$  und  $B$  Aussagen, so bezeichnet  $A \vee B$  die *Disjunktion* und man sagt “ $A$  oder  $B$ ”. Sie ist definiert über ihren Wahrheitsgehalt:

Wenn  $A$  und  $B$  wahr sind, dann ist  $A \vee B$  wahr.

Wenn  $A$  wahr und  $B$  falsch ist, dann ist  $A \vee B$  wahr.

Wenn  $A$  falsch und  $B$  wahr ist, dann ist  $A \vee B$  wahr.

Wenn  $A$  und  $B$  falsch sind, dann ist  $A \vee B$  falsch.

**Definition 1.8.5** (Tautologie, Kontradiktion). Es sei  $A$  eine beliebige Aussage.

1. Die *Tautologie*  $\mathbb{W}$  ist die Aussage mit dem Wahrheitswert der Aussage  $(\neg A) \vee A$ .
2. Die *Kontradiktion*  $\mathbb{F}$  ist die Aussage mit dem Wahrheitswert der Aussage  $(\neg A) \wedge A$ .

Das heißt,  $\mathbb{W}$  ist immer wahr und  $\mathbb{F}$  ist immer falsch.

Die Definitionen 1.8.2-1.8.5 kann man gut mit Hilfe von *Wahrheitstabelle* (WWT) beschreiben:

| $A$ | $B$ | $\neg A$ | $A \wedge B$ | $A \vee B$ | $\mathbb{W}, A \vee \neg A$ | $\mathbb{F}, A \wedge \neg A$ |
|-----|-----|----------|--------------|------------|-----------------------------|-------------------------------|
| $w$ | $w$ | $f$      | $w$          | $w$        | $w$                         | $f$                           |
| $w$ | $f$ | $f$      | $f$          | $w$        | $w$                         | $f$                           |
| $f$ | $w$ | $w$      | $f$          | $w$        | $w$                         | $f$                           |
| $f$ | $f$ | $w$      | $f$          | $f$        | $w$                         | $f$                           |

**Definition 1.8.6** (Logische Äquivalenz). Es seien  $A, B, \dots$  Aussagen und weiter seien  $F(A, B, \dots)$  und  $G(A, B, \dots)$  Ausdrücke die durch Verknüpfung der Aussagen entstehen. Dann heißen  $F(A, B, \dots)$  und  $G(A, B, \dots)$  *logisch äquivalent*, wenn für alle Kombinationen von Wahrheitswerten der Aussagen  $A, B, \dots$  die Aussagen  $F(A, B, \dots)$  und  $G(A, B, \dots)$  den gleichen Wahrheitswert haben.

Wir schreiben dann  $F(A, B, \dots) \iff G(A, B, \dots)$ .

**Bemerkung 1.8.7.** Um die logische Äquivalenz von Aussagenverknüpfungen zu prüfen, in die  $k$  Aussagen  $A_1, A_2, \dots, A_k$  eingehen, braucht man eine WWT mit  $2^k$  Zeilen.

**Beispiel 1.8.8.** Es seien  $A$  und  $B$  Aussagen. Dann gelten folgende logische Äquivalenzen:

$$\begin{array}{ll} \neg(\neg A) \iff A & \\ A \wedge B \iff B \wedge A & A \vee B \iff B \vee A \\ A \vee A \iff A & A \wedge A \iff A \\ A \vee \mathbb{W} \iff \mathbb{W} & A \wedge \mathbb{W} \iff A \\ A \vee \mathbb{F} \iff A & A \wedge \mathbb{F} \iff \mathbb{F} \end{array}$$

**Satz 1.8.9** (Rechenregeln). *Es seien  $A, B$  und  $C$  Aussagen. Dann gelten folgende logischen Äquivalenzen:*

1.  $(A \wedge B) \wedge C \iff A \wedge (B \wedge C)$   
 $(A \vee B) \vee C \iff A \vee (B \vee C)$
2.  $A \wedge (B \vee C) \iff (A \wedge B) \vee (A \wedge C)$   
 $A \vee (B \wedge C) \iff (A \vee B) \wedge (A \vee C)$
3.  $\neg(A \wedge B) \iff \neg A \vee \neg B$  (De Morgansche Regeln)  
 $\neg(A \vee B) \iff \neg A \wedge \neg B$

**Definition 1.8.10** (Folgerung, Äquivalenz). Es seien  $A$  und  $B$  Aussagen. Dann ist die *Folgerung*  $A \rightarrow B$  durch

$$A \rightarrow B := \neg A \vee B$$

und die *Äquivalenz*  $A \leftrightarrow B$  durch

$$A \leftrightarrow B := (A \rightarrow B) \wedge (B \rightarrow A)$$

definiert.

**Bemerkung 1.8.11.** Die WWT dieser Verknüpfungen lauten

| $A$ | $B$ | $A \rightarrow B$ | $A \leftrightarrow B$ |
|-----|-----|-------------------|-----------------------|
| $w$ | $w$ | $w$               | $w$                   |
| $w$ | $f$ | $f$               | $f$                   |
| $f$ | $w$ | $w$               | $f$                   |
| $f$ | $f$ | $w$               | $w$                   |

Mit Hilfe der Äquivalenz kann man die logische Äquivalenz zweier Aussageverknüpfungen auch wie folgt charakterisieren:

**Bemerkung 1.8.12.** Es seien  $F(A, \dots)$  und  $G(A, \dots)$  zwei Aussageverknüpfungen. Dann gilt:

$$F(A, \dots) \iff G(A, \dots)$$

ist gleichbedeutend mit

$$F(A, \dots) \leftrightarrow G(A, \dots) \iff \mathbb{W}$$

**Definition 1.8.13** (Logische Folgerung). Mit den Bezeichnungen aus Definition 1.8.6 definieren wir die *logische Folgerung* analog zu Bemerkung 1.8.12

$$F(A, \dots) \implies G(A, \dots)$$

ist gleichbedeutend mit

$$F(A, \dots) \rightarrow G(A, \dots) \iff \mathbb{W}$$

**Bezeichnung 1.8.14.** Wir werden im Text die Unterscheidung zwischen logischer Folgerung und Folgerung sowie logischer Äquivalenz und Äquivalenz nicht machen und schreiben in beiden Fällen  $\implies$  und  $\iff$ .

**Definition 1.8.15.** Eine *Aussageform*  $\mathcal{A}$  ist eine Zuordnung, die jedem Wert  $x$  aus einer Grundmenge  $G$  einen Aussage  $\mathcal{A}(x)$  zuordnet.<sup>(i)</sup>

**Definition 1.8.16** (Allquantor, Existenzquantor). Es sei  $\mathcal{A}$  eine Aussageform über der Grundmenge  $G$ .

- $\forall x \in G : \mathcal{A}(x)$  bedeutet, dass die Aussageform  $\mathcal{A}$  allgemeingültig ist. Man sagt "Für jedes  $x \in G$  gilt  $\mathcal{A}(x)$ ".
- $\exists x \in G : \mathcal{A}(x)$  bedeutet, dass die Aussageform  $\mathcal{A}$  erfüllbar ist. Man sagt "Es gibt ein  $x \in G$ , sodass  $\mathcal{A}(x)$  gilt".

**Bemerkung 1.8.17.** Es sei  $\mathcal{A}$  eine Aussageform über  $G$ . Dann gilt

$$1. \mathcal{A} \text{ ist erfüllbar} \iff \exists x \in G : \mathcal{A}(x).$$

<sup>(i)</sup>siehe z.B. Satz 1.1.17, wo jedem  $n \in \mathbb{N}$  die Aussage  $A(n)$  zugeordnet wird.

2.  $\mathcal{A}$  ist *nicht erfüllbar*  $\iff \forall x \in G : \neg \mathcal{A}(x)$ .
3.  $\mathcal{A}$  ist *allgemeingültig*  $\iff \forall x \in G : \mathcal{A}(x)$ .
4.  $\mathcal{A}$  ist *nicht allgemeingültig*  $\iff \exists x \in G : \neg \mathcal{A}(x)$ .

In der letzten Bemerkung haben wir schon von folgendem Sachverhalt Gebrauch gemacht:

**Satz 1.8.18** (Negation von Quantoren). *Es sei  $\mathcal{A}$  eine Aussageform über der Grundmenge  $G$ . Dann gilt:*

1.  $\neg(\exists x \in G : \mathcal{A}(x)) \iff \forall x \in G : \neg \mathcal{A}(x)$
2.  $\neg(\forall x \in G : \mathcal{A}(x)) \iff \exists x \in G : \neg \mathcal{A}(x)$

**Bezeichnung 1.8.19.** Ist die Aussageform  $\mathcal{A}$  für genau ein  $x \in G$  erfüllbar, d.h.

$$\exists x \in G : \left( \mathcal{A}(x) \wedge \forall y \in G : (y \neq x \rightarrow \neg \mathcal{A}(y)) \right)$$

so kürzen wir das ab mit

$$\exists! x \in G : \mathcal{A}(x).$$

**Bemerkung 1.8.20** (Rechenregeln).

1. Es seien  $\mathcal{A}$  und  $\mathcal{B}$  Aussageformen auf der Grundmenge  $G$ . Dann gilt:

$$\begin{aligned} \forall x \in G : (\mathcal{A}(x) \wedge \mathcal{B}(x)) &\iff (\forall x \in G : \mathcal{A}(x)) \wedge (\forall x \in G : \mathcal{B}(x)) \\ \exists x \in G : (\mathcal{A}(x) \vee \mathcal{B}(x)) &\iff (\exists x \in G : \mathcal{A}(x)) \vee (\exists x \in G : \mathcal{B}(x)) \\ \forall x \in G : (\mathcal{A}(x) \rightarrow \mathcal{B}(x)) &\iff (\exists x \in G : \mathcal{A}(x)) \rightarrow (\forall x \in G : \mathcal{B}(x)) \\ \exists x \in G : (\mathcal{A}(x) \rightarrow \mathcal{B}(x)) &\iff (\forall x \in G : \mathcal{A}(x)) \rightarrow (\exists x \in G : \mathcal{B}(x)) \end{aligned}$$

2. Es seien  $\mathcal{A}$  und  $\mathcal{B}$  Aussageformen auf der Grundmenge  $G$  und  $D$  eine Aussage, die nicht von  $x$  abhängt, dann gilt

$$\begin{aligned} \forall x \in G : D &\iff \exists x \in G : D \iff D \\ \forall x \in G : (\mathcal{A}(x) \wedge D) &\iff (\forall x \in G : \mathcal{A}(x)) \wedge D \end{aligned}$$

$$\begin{aligned} \forall x \in G : (\mathcal{A}(x) \vee D) &\iff (\forall x \in G : \mathcal{A}(x)) \vee D \\ \forall x \in G : (\mathcal{A}(x) \rightarrow D) &\iff (\exists x \in G : \mathcal{A}(x)) \rightarrow D \\ \exists x \in G : (\mathcal{A}(x) \vee D) &\iff (\exists x \in G : \mathcal{A}(x)) \vee D \\ \exists x \in G : (\mathcal{A}(x) \wedge D) &\iff (\exists x \in G : \mathcal{A}(x)) \wedge D \\ \exists x \in G : (\mathcal{A}(x) \rightarrow D) &\iff (\forall x \in G : \mathcal{A}(x)) \rightarrow D \end{aligned}$$

3. Es sei  $\mathcal{C}$  eine Aussageform auf der Menge  $G_1 \times G_2$ . Dann gilt<sup>(i)</sup>

$$\begin{aligned} \forall x \in G_1 \forall y \in G_2 : \mathcal{C}(x, y) &\iff \forall y \in G_2 \forall x \in G_1 : \mathcal{C}(x, y) \\ \exists x \in G_1 \exists y \in G_2 : \mathcal{C}(x, y) &\iff \exists y \in G_2 \exists x \in G_1 : \mathcal{C}(x, y) \end{aligned}$$

**Beispiel 1.8.21.** • Ist in 3.  $\mathcal{C}(x, y) = \mathcal{A}(x) \wedge \mathcal{B}(y)$  so gilt

$$\forall x \in G_1 \forall y \in G_2 : (\mathcal{A}(x) \wedge \mathcal{B}(y)) \iff (\forall x \in G_1 : \mathcal{A}(x)) \wedge (\forall y \in G_2 : \mathcal{B}(y))$$

• Ist in 2.  $\mathcal{C}(x, y) = \mathcal{A}(x) \vee \mathcal{B}(y)$  so gilt

$$\exists x \in G_1 \exists y \in G_2 : (\mathcal{A}(x) \vee \mathcal{B}(y)) \iff (\exists x \in G_1 : \mathcal{A}(x)) \vee (\exists y \in G_2 : \mathcal{B}(y))$$

## 1.9 Anhang Analysis I: Zahlbereichserweiterungen $\mathbb{N} \rightarrow \mathbb{Z} \rightarrow \mathbb{Q}$

Hier präsentieren wir die Details zur Zahlenbereichserweiterung  $\mathbb{N} \rightarrow \mathbb{Z} \rightarrow \mathbb{Q}$  mittels Äquivalenzrelationen gemäß Beispiel 1.1.33.

### 1.9.1 Die ganzen Zahlen

Die ganzen Zahlen erhält man, indem man eine kleinere natürliche Zahl von einer größeren Zahl "abzieht". Aber was soll das heißen, wenn man lediglich die Addition auf  $\mathbb{N}$  kennt. Für einen Spezialfall kann man die Subtraktion definieren, nämlich wenn  $a \geq b$  ist. In diesem Fall gibt es genau ein  $k \in \mathbb{N}$ , sodass  $a = b + k$  ist. Wir setzen in diesem Fall  $a - b := k$

Wir konstruieren die ganzen Zahlen mit Hilfe von Paaren natürlicher Zahlen.

<sup>(i)</sup>Wenn zwei Quantoren aufeinanderfolgen, so hat man das geklammert zu verstehen, z.B.  $\forall x \in G_1 \forall y \in G_2 : \mathcal{C}(x, y) \iff \forall x \in G_1 : (\forall y \in G_2 : \mathcal{C}(x, y))$

**Definition/Satz 1.9.1.** Wir betrachten die Menge  $\mathbb{N} \times \mathbb{N}$  und auf ihr eine Relation  $\sim_Z$ , die wie folgt definiert ist

$$(a, b) \sim_Z (a', b') \iff a + b' = a' + b.$$

Dann ist  $\sim_Z$  eine Äquivalenzrelation auf  $\mathbb{N} \times \mathbb{N}$ . Die Äquivalenzklasseneinteilung bezeichnen wir mit

$$\widehat{\mathbb{Z}} = \{[(a, b)]_Z \mid (a, b) \in \mathbb{N} \times \mathbb{N}\}.$$

**Bemerkung 1.9.2.** In der Menge  $\widehat{\mathbb{Z}}$  gibt es lediglich drei verschiedene Typen von Klassen:  $[(a, b)]$  mit  $a > b$ ,  $[(a, a)]$  und  $[(a, b)]$  mit  $a < b$ . Dann gilt

$$[(a, b)] = \begin{cases} [(k, 0)] & \text{falls } a > b \text{ und } a = b + k \\ [(0, 0)] & \text{falls } a = b \\ [(0, k')] & \text{falls } a < b \text{ und } b = a + k' \end{cases}$$

Ein "schönes" Repräsentantensystem ist also

$$\{(0, k), (0, 0), (k, 0) \mid k \in \mathbb{N}^*\}.$$

**Definition/Satz 1.9.3.** Auf der Menge  $\widehat{\mathbb{Z}}$  definieren wir eine Addition und eine Multiplikation wie folgt

$$\begin{aligned} [(a, b)] +_Z [(a', b')] &:= [(a + a', b + b')], \\ [(a, b)] \cdot_Z [(a', b')] &:= [(a \cdot a' + b \cdot b', a \cdot b' + b \cdot a')]. \end{aligned}$$

Diese ist wohldefiniert und es gilt

- $(\widehat{\mathbb{Z}}, +_Z)$  ist eine kommutative Gruppe, dessen neutrales Element mit  $\widehat{0}$  bezeichnet wird.
- $(\widehat{\mathbb{Z}} \setminus \{\widehat{0}\}, \cdot_Z)$  ist assoziativ, kommutativ und besitzt ein Einselement. Das neutrale Element bezeichnen wir mit  $\widehat{1}$ .
- $(\widehat{\mathbb{Z}}, +_Z, \cdot_Z)$  ist distributiv.

Diese drei Punkte machen  $\widehat{\mathbb{Z}}$  zu einem kommutativen Ring mit Einselement.

*Proof.* Nachdem man die Wohldefiniertheit gezeigt hat, sieht man, dass  $\widehat{0} = [(0, 0)]$  und  $[(a, b)] +_Z [(b, a)] = [(0, 0)]$ . Außerdem ist  $\widehat{1} = [(1, 0)]$ . Alles andere rechnet man nach.  $\square$

**Definition/Satz 1.9.4.** Auf der Menge  $\widehat{\mathbb{Z}}$  definieren wir eine Ordnungsrelation durch

$$(a, b) \leq_Z (a', b') \iff a + b' \leq a' + b.$$

**Satz 1.9.5.** Die Abbildung  $\phi : \mathbb{N} \rightarrow \widehat{\mathbb{Z}}$  mit

$$\phi(a) = [(a, 0)]$$

ist mit den Strukturen verträglich, d.h.

$$\phi(a + b) = \phi(a) +_Z \phi(b),$$

$$\phi(a \cdot b) = \phi(a) \cdot_Z \phi(b),$$

$$a \leq b \iff \phi(a) \leq_Z \phi(b).$$

**Bezeichnung 1.9.6.** Wir schreiben für  $a > 0$

|                |       |  |
|----------------|-------|--|
| $\mathbb{Z}$   | statt | $\widehat{\mathbb{Z}}$                           |
| $\mathbb{Z}^*$ | statt | $\widehat{\mathbb{Z}} \setminus \{\widehat{0}\}$ |
| $k$            | statt | $[(a + k, a)]$                                   |
| $-k$           | statt | $[(a, a + k)]$                                   |
| $0$            | statt | $[(a, a)]$                                       |
| $+$            | statt | $+_Z$  |
| $\cdot$        | statt | $\cdot_Z$  |
| $k - \ell$     | statt | $k + (-\ell)$                                    |
| $\leq$         | statt | $\leq_Z$   |
| $\mathbb{N}$   | statt | $\phi(\mathbb{N})$                               |

**Definition 1.9.7.** Wir benutzen die Schreibweisen

$$\text{sign}(a) := \begin{cases} 1 & \text{falls } a \geq 0 \\ -1 & \text{falls } a < 0 \end{cases}$$

$$|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases}$$

und nennen  $\text{sign}(a)$  das *Vorzeichen* und  $|a|$  den *Betrag* von  $a \in \mathbb{Z}$ . Es gilt stets  $a = \text{sign}(a) \cdot |a|$ .

**Definition/Satz 1.9.8** (Restklassenringe). 1. Es sei  $n \in \mathbb{N}^{\geq 2}$ . Auf  $\mathbb{Z}$  definieren wir die Äquivalenzrelation

$$k \sim \ell \iff n \mid (k - \ell).$$

Die zugehörige Äquivalenzklasseneinteilung wird mit  $\mathbb{Z}_n$  und besitzt genau  $n$ -Elemente. Ein Repräsentantensystem ist durch die Zahlen  $\{0, 1, \dots, n-1\}$  gegeben. Die Klasse zu  $k \in \mathbb{Z}$  wird wieder mit  $[k]$  bezeichnet.

2. Die folgenden Definition liefern wohldefinierte Verknüpfungen auf  $\mathbb{Z}_n$ :

$$[k] + [\ell] := [k + \ell], \quad [k] \cdot [\ell] := [k \cdot \ell].$$

3.  $\mathbb{Z}_n$  wird mit diesen Verknüpfungen zu einem kommutativen Ring mit Einselement.

4.  $\mathbb{Z}_n$  ist genau dann nullteilerfrei, wenn  $n$  eine Primzahl ist. Insbesondere ist  $\mathbb{Z}_n$  in diesem Fall ein Körper.

**Satz 1.9.9.** Die Menge  $\mathbb{Z}$  der ganzen Zahlen ist abzählbar, d.h. es gibt eine surjektive Abbildung  $\gamma : \mathbb{N} \rightarrow \mathbb{Z}$ , etwa

$$\gamma(k) = \begin{cases} \ell & \text{falls } k = 2\ell \text{ mit } \ell \geq 0 \\ -\ell & \text{falls } k = 2\ell - 1 \text{ mit } \ell \geq 1 \end{cases}$$

(diese Abbildung ist sogar bijektiv).

**Bemerkung 1.9.10.** Die Mengen  $\mathbb{N}^2$  und damit auch  $\mathbb{Z}^2$  sind abzählbar.

- Für die Abbildung, die die Abzählbarkeit von  $\mathbb{N}^2$  zeigt, betrachten wir zwei Primzahlen  $p, q \in \mathbb{N}$ .<sup>(i)</sup> Damit definieren wir die Abbildung  $\delta_{pq} : \mathbb{N}^2 \rightarrow \mathbb{N}$  mit  $\delta_{pq}(n, m) = p^n q^m$ . Diese ist injektiv und, wenn wir die Bildmenge auf  $M := \text{Bild}(\delta_{pq})$  einschränken, ist  $\delta_{pq} : \mathbb{N}^2 \rightarrow M$  sogar bijektiv. Damit existiert die Umkehrabbildung  $\delta_{pq}^{-1} : M \rightarrow \mathbb{N}^2$  und wir definieren damit die surjektive Abbildung

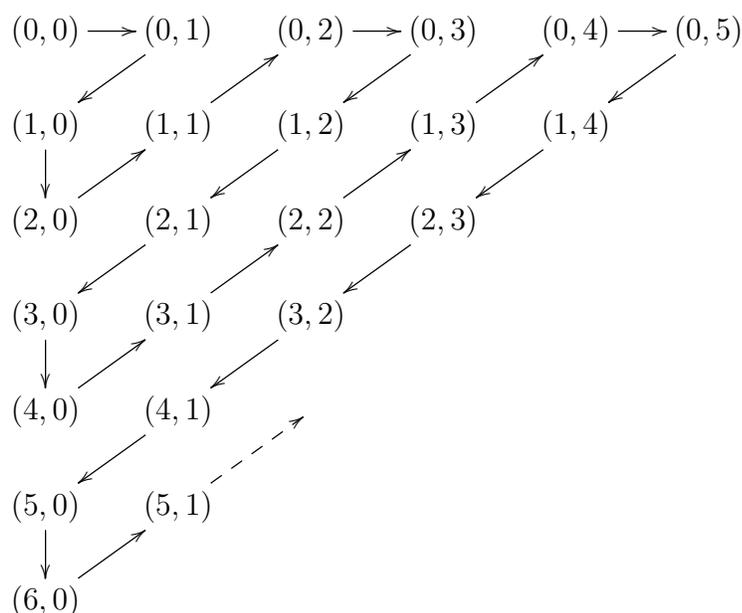
$$\hat{\gamma} : \mathbb{N} \rightarrow \mathbb{N}^2, \quad \hat{\gamma}(n) = \begin{cases} \delta_{pq}^{-1}(n) & \text{falls } n \in M \\ (0, 0) & \text{falls } n \in \mathbb{N} \setminus M \end{cases}$$

<sup>(i)</sup> $p \in \mathbb{N}$  heißt Primzahl, wenn  $p > 1$  und  $p$  nur 1 und  $p$  selbst als Teiler besitzt.

- Für die Abzählbarkeit von  $\mathbb{N}^2$  kann man auch das *Cantorsche Diagonalverfahren* heranziehen, siehe Abb. 1.9.1.
- Die Abzählbarkeit von  $\mathbb{Z}^2$  folgt damit leicht: Schreiben wir für die Bilder der Abbildung  $\hat{\gamma}$  jetzt  $\hat{\gamma}(n) = (\hat{\gamma}_1(n), \hat{\gamma}_2(n))$ , so erhalten wir eine surjektive Abbildung  $\tilde{\gamma} : \mathbb{N} \rightarrow \mathbb{Z}^2$  durch

$$\tilde{\gamma}(n) = \left( \gamma(\hat{\gamma}_1(n)), \gamma(\hat{\gamma}_2(n)) \right).$$

Abbildung 1.9.1: Das Cantorsche Diagonalverfahren



### 1.9.2 Die rationalen Zahlen

Wir haben bei den natürlichen Zahlen schon von Teilbarkeit gesprochen und wir können diese Diskussion direkt auf die ganzen Zahlen übertragen. Wir sehen, dass 4 durch 2 und  $-6$  durch  $-2$  oder 3 teilbar ist und natürlich dass alle Zahlen durch sich selbst teilbar sind. Aber was ist die allgemeine Division?

Wir konstruieren die rationalen Zahlen mit Hilfe von Paaren ganzer Zahlen und gehen analog zum vorhergehenden Abschnitt vor.

**Definition/Satz 1.9.11.** Wir betrachten die Menge  $\mathbb{Z} \times \mathbb{Z}^*$  und auf ihr eine Relation  $\sim_Q$ , die wie folgt definiert ist

$$(a, b) \sim_Q (a', b') \iff ab' = a'b.$$

Dann ist  $\sim_Q$  eine Äquivalenzrelation auf  $\mathbb{Z} \times \mathbb{Z}^*$ . Die Äquivalenzklasseneinteilung bezeichnen wir mit

$$\widehat{\mathbb{Q}} = \{[(a, b)]_Z \mid (a, b) \in \mathbb{Z} \times \mathbb{Z}^*\}.$$

**Bemerkung 1.9.12.** Es seien  $a, b \neq 0$  und es gelte  $b|a$  oder  $a|b$ . Dann ist im ersten Fall  $a = b \cdot k$  und  $[(a, b)] = [(k, 1)] \in \widehat{\mathbb{Q}}$ . Im zweiten Fall ist  $b = a \cdot k$  und es gilt  $[(a, b)] = [(\text{sign}(a \cdot b), |k|)] \in \widehat{\mathbb{Q}}$ . Ist  $a = 0$  so ist  $[(a, b)] = [(0, 1)] \in \widehat{\mathbb{Q}}$ . Etwas allgemeiner kann man sagen: ist  $a = ka', b = kb'$  so ist  $[(a, b)] = [(a', b')]$ . Man kann also stets einen Repräsentanten  $(a', b') \in [(a, b)]$  mit  $\text{ggT}(a', b') = 1$  und  $b > 0$  wählen. Ein "schönes" Repräsentantensystem ist also

$$\{(a, b), (0, 1) \mid a \in \mathbb{Z}^*, b \in \mathbb{N}^*, \text{ggT}(|a|, b) = 1\}.$$

**Definition/Satz 1.9.13.** Auf der Menge  $\widehat{\mathbb{Q}}$  definieren wir eine Addition und eine Multiplikation wie folgt

$$\begin{aligned} [(a, b)] +_Q [(a', b')] &:= [(a \cdot b' + a' \cdot b, b \cdot b')], \\ [(a, b)] \cdot_Q [(a', b')] &:= [(a \cdot a', b \cdot b')]. \end{aligned}$$

Diese ist wohldefiniert und es gilt

- $(\widehat{\mathbb{Q}}, +_Q)$  ist eine kommutative Gruppe, dessen neutrales Element mit  $\widehat{0}$  bezeichnet wird.
- $(\widehat{\mathbb{Q}} \setminus \{\widehat{0}\}, \cdot_Q)$  ist eine kommutative Gruppe, dessen neutrales Element mit  $\widehat{1}$ .
- $(\widehat{\mathbb{Q}}, +_Q, \cdot_Q)$  ist distributiv.

Diese drei Punkte machen  $\widehat{\mathbb{Q}}$  zu einem Körper.

*Proof.* Nachdem man die Wohldefiniertheit gezeigt hat, sieht man  $\widehat{0} = [(0, 1)]$  und  $[(a, b)] +_Q [(-a, b)] = [(0, 1)]$  sowie  $\widehat{1} = [(1, 1)]$  und  $[(a, b)] \cdot_Q [(b, a)] = [(1, 1)]$  für  $a \neq 0$ . Alles andere rechnet man nach.  $\square$

**Definition/Satz 1.9.14.** Auf der Menge  $\widehat{\mathbb{Q}}$  definieren wir eine Ordnungsrelation durch

$$(a, b) \leq_Q (a', b') \iff \text{sign}(b) \cdot a \cdot |b'| \leq \text{sign}(b') \cdot a' \cdot |b|.$$

**Satz 1.9.15.** Die Abbildung  $\psi : \mathbb{Z} \rightarrow \widehat{\mathbb{Q}}$  mit

$$\psi(a) = [(a, 1)]$$

ist mit den Strukturen verträglich, d.h.

$$\begin{aligned} \psi(a + b) &= \psi(a) +_Q \psi(b), \\ \psi(a \cdot b) &= \psi(a) \cdot_Q \psi(b), \\ a \leq b &\iff \psi(a) \leq_Q \psi(b). \end{aligned}$$

**Bezeichnung 1.9.16.** Wir schreiben

$$\begin{array}{ll} \mathbb{Q} & \text{statt } \widehat{\mathbb{Q}} \\ \frac{a}{b} & \text{statt } [(a, b)] \\ a & \text{statt } [(a, 1)] \\ + & \text{statt } +_Q \\ \cdot & \text{statt } \cdot_Q \\ a : b & \text{statt } a \cdot \frac{1}{b} \\ \leq & \text{statt } \leq_Q \\ \mathbb{Z} & \text{statt } \psi(\mathbb{Z}) \end{array}$$

**Bemerkung 1.9.17.** Zwischen zwei verschiedenen rationalen Zahlen liegt stets eine weitere rationale Zahl und damit unendlich viele. Wählt man zwei rationale Zahlen  $a, b \in \mathbb{Q}$  mit  $a < b$ , so gilt stets  $a < \frac{a+b}{2} < b$ .

Dass es trotzdem noch "Lücken" gibt, die es zu stopfen gilt, sieht man zum Beispiel wie folgt: Zeichne ein Quadrat mit der Seitenlänge 1. Dann ist das Quadrat der Diagonale des Dreiecks wegen des Satzes von Pythagoras gegeben durch  $\ell^2 = 1^2 + 1^2 = 2$ , wobei  $\ell$  die Länge der Diagonale ist. Andererseits ist  $\ell$  jedoch nicht rational.

Die Behandlung dieser "Lücken" ist Thema von Abschnitt 1.2.

**Bemerkung 1.9.18.** Der Betrag  $|\cdot|$  und das Vorzeichen  $\text{sign}$  sind genauso wie in Definition 1.9.7 auch auf  $\mathbb{Q}$  definiert. Auf die Rechenregeln gehen wir näher in Abschnitt 1.2.5 ein.

**Bemerkung 1.9.19.** Die Menge  $\mathbb{Q}$  der rationalen Zahlen ist abzählbar. Dazu definiere eine surjektive Abbildung  $\delta : \mathbb{Z}^2 \rightarrow \mathbb{Q}$  durch

$$\delta(r, s) = \frac{r}{s}.$$

Diese liefert uns zusammen mit der Abbildung  $\tilde{\gamma} : \mathbb{N} \rightarrow \mathbb{Z}^2$  eine surjektive Abbildung  $\check{\gamma} : \mathbb{N} \rightarrow \mathbb{Q}$  durch

$$\check{\gamma}(n) = \delta(\tilde{\gamma}(n)).$$

# KAPITEL 2

---

## Analysis II

---

### 2.1 Funktionenfolgen

#### 2.1.1 Konvergenzbegriffe

In Kapitel 1.6.3 haben wir Potenzreihen kennengelernt. Dort hat man Werten  $x \in \mathbb{R}$  jeweils Zahlenreihe  $\sum a_k x^k$  zugeordnet. Dabei haben wir dann für festes  $x$  jeweils diese Zahlenreihen untersucht. Im dem Fall, dass die Potenzreihe konvergiert, haben wir gesehen, dass sie dies für Werte aus einem offenen Intervall tut. Damit haben wir dann die Grenzfunktion der Portenzreihe definiert.

Wir wechseln jetzt den Standpunkt. Gegeben sei die Potenzreihe  $\sum a_k x^k$ . Betrachtet man  $x$  als Variable, kann die  $n$ -te Teilsumme  $\sum_{k=0}^n a_k x^k$  als Polynomfunktion interpretiert werden. Man erhält somit eine Folge von Polynomen

$$(p_n)_{n \in \mathbb{N}} \quad \text{mit} \quad p_n(x) = \sum_{k=0}^n a_k x^k.$$

Und die Grenzfunktion der Potenzreihe ist in diesem Sinne der Grenzwert der ‘Funktionsfolge’  $(p_n)$ .

Wir werden in diesem Kapitel insbesondere die Beweise der Sätze 1.6.31 und 1.6.32 nachholen.

**Definition 2.1.1.** Es seien für alle  $n \in \mathbb{N}$  Funktionen  $f_n : M \rightarrow \mathbb{R}$  auf der selben Definitionsmenge  $M$  definiert. Dann nennt man die Folge  $(f_n)_{n \in \mathbb{N}}$  eine *Funktionsfolge auf  $M$* .

**Definition 2.1.2.** Es sei  $(f_n)$  eine Funktionsfolge auf  $M \subset \mathbb{R}$ .

1.  $(f_n)$  heißt *punktweise konvergent gegen die Funktion  $f : M \rightarrow \mathbb{R}$* , wenn es für alle  $x \in M$  und für alle  $\epsilon > 0$  ein  $n_0 \in \mathbb{N}$  gibt, sodass für alle  $n \geq n_0$  die Abschätzung  $|f_n(x) - f(x)| < \epsilon$  gilt, d.h.

$$\forall x \in M \forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |f_n(x) - f(x)| < \epsilon$$

In diesem Fall heißt  $f$  auch der *punktweise Limes* der Funktionsfolge  $(f_n)$  und wir schreiben auch  $f_n \xrightarrow{\text{pktw}} f$ .

2.  $(f_n)$  heißt *gleichmäßig konvergent gegen die Funktion  $f : M \rightarrow \mathbb{R}$* , wenn es für alle  $\epsilon > 0$  ein  $n_0 \in \mathbb{N}$  gibt, sodass für alle  $n \geq n_0$  und für alle  $x \in M$  die Abschätzung  $|f_n(x) - f(x)| < \epsilon$  gilt, d.h.

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 \forall x \in M : |f_n(x) - f(x)| < \epsilon$$

In diesem Fall heißt  $f$  auch der *gleichmäßige Limes* der Funktionsfolge  $(f_n)$  und wir schreiben auch  $f_n \xrightarrow{\text{glm}} f$  oder  $\lim_{n \rightarrow \infty} f_n = f$ .

**Folgerung 2.1.3.** • Eine gleichmäßig konvergente Funktionsfolge ist ebenfalls punktweise konvergent. Insbesondere sind die Grenzfunktionen in beiden Fällen die gleichen.

- Um eine Funktionsfolge auf gleichmäßige Konvergenz zu untersuchen, benötigt man in der Regel die Grenzfunktion. Diese (potentielle) Grenzfunktion findet man daher zunächst über eine punktweise Betrachtung.

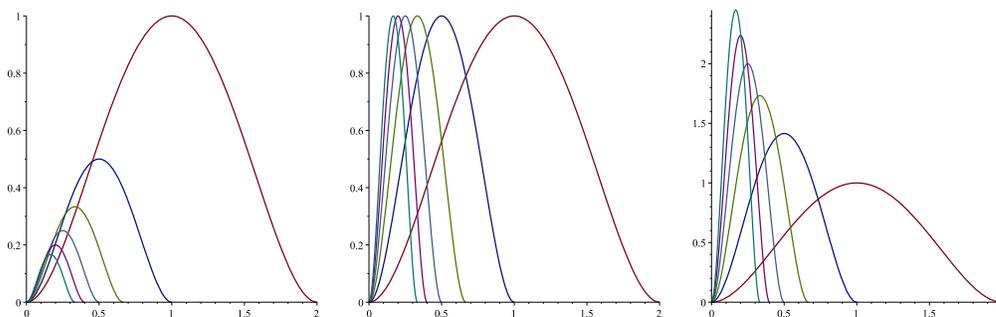
**Beispiel 2.1.4.** 1. Die Funktionsfolge  $(f_n)$  mit  $f_n(x) = x^n$  ist auf dem Intervall  $[0, 1]$  punktweise, aber nicht gleichmäßig konvergent.

2. Die Funktionenfolge  $(f_n)$  mit  $f_n(x) = x^n$  auf dem Intervall  $[0, 1[$  ist nicht gleichmäßig konvergent.
3. Die Funktionenfolge  $(f_n)$  mit  $f_n(x) = x^n$  auf dem Intervall  $[0, r]$  mit  $0 < r < 1$  ist gleichmäßig konvergent.
4. Es sei  $(\alpha_n)$  eine positive reelle Zahlenfolge. Die Funktionenfolge  $(f_n)$  auf  $\mathbb{R}$  mit

$$f_n(x) = \begin{cases} \alpha_n ((nx - 1)^2 - 1)^2 & \text{falls } x \in [0, \frac{2}{n}] \\ 0 & \text{sonst} \end{cases}$$

ist punktweise konvergent. Sie ist genau dann gleichmäßig konvergent, wenn  $(\alpha_n)$  eine Nullfolge ist, siehe Abbildung 2.1.1.

Abbildung 2.1.1:  $f_n(x)$  aus Bsp. 2.1.4.4. für  $\alpha_n = \frac{1}{n}$ ,  $\alpha_n = 1$  und  $\alpha_n = \sqrt{n}$



**Definition/Bemerkung 2.1.5.** • Für eine Funktion  $f : M \rightarrow \mathbb{R}$  definieren wir die Zahl

$$\|f\| := \sup_{x \in M} |f(x)| \in \mathbb{R}^{\geq 0} \cup \{\infty\}.$$

- Eine Funktionenfolge  $(f_n)$  ist genau dann gleichmäßig konvergent gegen die Grenzfunktion  $f$ , wenn

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : \|f_n - f\| < \epsilon$$

bzw.

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

- Die Menge

$$B(M) = \{f : M \rightarrow \mathbb{R} \mid \|f\| < \infty\}$$

der beschränkten Funktionen auf  $M$  ist ein Untervektorraum aller reellen Funktionen auf  $M$ . Damit ist

$$\|\cdot\| : B(M) \rightarrow \mathbb{R}^{\geq 0}$$

eine Norm, die *Supremumsnorm*.<sup>(i)</sup>

Ein wichtiger Spezialfall ist hierbei der Fall  $M = [a, b]$ . In diesem Fall ist die Menge  $C([a, b])$  der auf  $[a, b]$  stetigen reellen Funktionen ein Untervektorraum von  $B([a, b])$  und es gilt dort  $\|f\| = \max_{x \in [a, b]} |f(x)|$ .

### 2.1.2 Eigenschaften konvergenter Funktionenfolgen

**Satz 2.1.6.** *Es sei  $(f_n)$  eine gleichmäßig konvergente Funktionenfolge auf  $M$  bei der jedes Folgenglied in  $x_0 \in M$  stetig ist. Dann ist auch die Grenzfunktion in  $x_0$  stetig. Es gilt also*

$$\lim_{n \rightarrow \infty} \left( \lim_{x \rightarrow x_0} f_n(x) \right) = \lim_{x \rightarrow x_0} \left( \lim_{n \rightarrow \infty} f_n(x) \right).$$

**Folgerung 2.1.7.** Ist  $(f_n)$  eine gleichmäßig konvergente Folge stetiger Funktionen auf  $M$ , dann ist die Grenzfunktion ebenfalls stetig auf  $M$ .

Prinzipiell kann man auch ohne Kenntnis einer potentiellen Grenzfunktion eine Funktionenfolge auf punktweise bzw. gleichmäßige Konvergenz untersuchen. Für die punktweise Konvergenz zieht man sich auf das Wissen über Zahlenfolgen zurück und kann das Cauchy-Kriterium anwenden. Dieses besitzt aber auch eine Variante im Fall der gleichmäßigen Konvergenz:

---

<sup>(i)</sup>Es sei  $V$  ein reeller Vektorraum. Eine Norm ist eine Abbildung  $n : V \rightarrow \mathbb{R}^{\geq 0}$  mit den folgenden drei Eigenschaften:

1.  $n(v) = 0 \iff v = 0$
2.  $n(\alpha v) = |\alpha| n(v)$
3.  $n(v + w) \leq n(v) + n(w)$  für alle  $v, w \in V$  und  $\alpha \in \mathbb{R}$

**Satz 2.1.8.** *Eine Funktionenfolge  $(f_n)$  ist genau dann gleichmäßig konvergent, wenn*

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n, m \geq n_0 : \|f_n - f_m\| < \epsilon.$$

**Definition/Bemerkung 2.1.9.** • Eine Funktionenfolge  $(f_n)$  auf  $M$  heißt

- *beschränkt*, wenn jedes einzelne Folgenglied  $f_n$  beschränkt ist.
- *gleichgradig beschränkt*, wenn es eine Zahl  $C > 0$  gibt mit  $|f_n(x)| < C$  für alle  $x \in M$  und alle  $n \in \mathbb{N}$ .
- Ist eine Funktionenfolge  $(f_n)$  gleichgradig beschränkt, so ist sie beschränkt.
- Ist eine Funktionenfolge allerdings beschränkt, so muss sie nicht unbedingt gleichgradig beschränkt sein. Es gilt vielmehr:

Eine Folge  $(f_n)$  beschränkter Funktionen ist genau dann gleichgradig beschränkt, wenn jede Folge  $(C_n)$  von Schranken beschränkt ist. Dabei sei  $C_n$  jeweils eine Schranke von  $f_n$ .

**Satz 2.1.10.** 1. *Es seien  $(f_n)$  und  $(g_n)$  gleichmäßig konvergente Funktionenfolgen und es sei  $c \in \mathbb{R}$ . Dann sind die Funktionenfolgen  $(f_n + g_n)$  und  $(cf_n)$  ebenfalls gleichmäßig konvergent.*

*Ist hier  $f_n \xrightarrow{glm} f$  und  $g_n \xrightarrow{glm} g$ , so gilt  $f_n + g_n \xrightarrow{glm} f + g$  und  $cf_n \xrightarrow{glm} cf$ .*

2. *Es sei  $(f_n)$  eine gleichmäßig konvergente Funktionenfolgen mit beschränkter Grenzfunktion und es sei  $(c_n)$  eine konvergente Zahlenfolge. Dann ist die Funktionenfolge  $(c_n f_n)$  ebenfalls gleichmäßig konvergent.*

*Ist hier  $f_n \xrightarrow{glm} f$  und  $c = \lim_{n \rightarrow \infty} c_n$ , so gilt  $c_n f_n \xrightarrow{glm} cf$ .*

3. *Es seien  $(f_n)$  und  $(g_n)$  gleichmäßig konvergente Funktionenfolgen mit beschränkten Grenzfunktionen. Dann ist auch die Funktionenfolge  $(f_n g_n)$  gleichmäßig konvergent.*

*Ist in diesem Fall  $f_n \xrightarrow{glm} f$  und  $g_n \xrightarrow{glm} g$ , so gilt  $f_n g_n \xrightarrow{glm} fg$ .*

4. *Es seien  $(f_n)$  gleichmäßig konvergent gegen die Nullfunktion und  $(g_n)$  gleichgradig beschränkt. Dann ist die Funktionenfolge  $(g_n f_n)$  auch gleichmäßig konvergent.*

5. Es seien  $(f_n)$  gleichmäßig konvergent gegen die Nullfunktion und  $(g_n)$  eine gleichgradig beschränkte Funktionenfolge. Dann konvergiert die Funktionenfolge  $(f_n g_n)$  ebenfalls gleichmäßig.

**Bemerkung/Beispiel 2.1.11.** Dass in dem vorigen Satz die Aussage 2. und damit die Aussage 3. ohne den Zusatz der beschränkten Grenzfunktion in der Regel nicht richtig ist, zeigt das folgende Beispiel:

Wir betrachten die Funktionenfolge  $(f_n)$  mit  $f_n(x) = x^2 + \frac{1}{n}$  auf  $M = \mathbb{R}$  und die Zahlenfolge  $(c_n)$  mit  $c_n = 1 + \frac{1}{n}$ . Dann ist  $(f_n)$  punktweise konvergent mit punktwiser Grenzfunktion  $f(x) = x^2$ , denn für alle  $x \in \mathbb{R}$  ist  $\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} (x^2 + \frac{1}{n}) = x^2$ . Weiter ist  $(f_n)$  gleichmäßig konvergent, denn es ist  $\|f_n - f\| = \frac{1}{n}$  konvergent. Weiter ist  $(c_n)$  konvergent mit Grenzwert  $c = 1$ .

Für jedes  $x \in \mathbb{R}$  gilt damit  $\lim_{n \rightarrow \infty} (c_n f_n(x)) = c f(x) = x^2$ , sodass  $c_n f_n \xrightarrow{\text{pktw}} c f$ .

Allerdings ist  $(c_n f_n)$  nicht gleichmäßig konvergent. Denn sei  $\epsilon > 0$  und  $n \in \mathbb{N}$  gegeben, dann findet man stets ein  $x$  sodass  $|c_n f_n(x) - x^2| > \epsilon$ . Sei etwa  $x > \sqrt{n\epsilon}$ , dann ist

$$|c_n f_n(x) - x^2| = \frac{1}{n} \left( x^2 + 1 + \frac{1}{n} \right) > \frac{1}{n} x^2 > \epsilon.$$

### 2.1.3 Funktionenreihen

Analog wie wir bei Zahlenfolgen auch Zahlenreihen untersucht haben, können wir das auch bei Funktionen tun.

**Definition/Bemerkung 2.1.12.** • Ist  $(f_n)$  eine Funktionenfolge auf  $M \subset \mathbb{R}$ , so nennt man die Funktionenfolge  $(F_n)$  mit  $F_n := \sum_{k=0}^n f_k$  eine *Funktionenreihe auf  $M$* . Wir schreiben für eine solche Funktionenreihe statt  $(F_n)$  wie im Falle von Zahlenreihen  $\sum f_k$ .

- Eine Funktionenreihe  $\sum f_k$  heißt punktweise/gleichmäßig konvergent, wenn das für die Folge  $(F_n)$  der Teilsummen gilt.
- Im Falle der punktweisen Konvergenz bezeichnen wir die Grenzfunktion mit  $\sum_{k=0}^{\infty} f_k$ , also

$$\left( \sum_{k=0}^{\infty} f_k \right) (x) := \sum_{k=0}^{\infty} f_k(x).$$

In diesem Fall konvergieren alle Zahlenfolgen  $f_n(x)$  punktweise gegen Null.

**Beispiel 2.1.13** (Zentrales Beispiel). Wir haben bereits eine Potenzreihe  $\sum a_k x^k$  als Funktionenfolge interpretiert, nämlich als die Folge ihrer Teilsummen  $p_n$  mit  $p_n(x) = \sum_{k=0}^n a_k x^k$ , die allesamt Polynome sind.

Im Sinne der obigen Definition ist eine Potenzreihe auch eine Funktionenreihe, nämlich die Reihe über die Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  mit  $f_n(x) = a_n x^n$ .

**Folgerung 2.1.14.** 1. Sind  $\sum f_k$  und  $\sum g_k$  punktweise konvergente Funktionenreihen und ist  $c \in \mathbb{R}$ , so sind  $\sum (f_k + g_k)$  und  $\sum (c f_k)$  ebenfalls punktweise konvergent und es gilt

$$\sum_{k=0}^{\infty} (f_k + g_k) = \sum_{k=0}^{\infty} f_k + \sum_{k=0}^{\infty} g_k, \quad \sum_{k=0}^{\infty} (c f_k) = c \sum_{k=0}^{\infty} f_k.$$

2. Ist die Funktionenreihe  $\sum f_k$  gleichmäßig konvergent, dann konvergiert die Funktionenfolge  $(f_n)$  gleichmäßig gegen die Nullfunktion.

**Bemerkung 2.1.15.** Im Punkt 2 von Folgerung 2.1.14 reicht die punktweise Konvergenz der Reihe nicht aus, um auf die gleichmäßige Konvergenz von  $(f_n)$  gegen Null zu schließen.

Als Beispiel nehme man die Reihe über die Funktionenfolge aus Beispiel 2.1.4.4 mit beliebiger Folge  $(\alpha_n)$ . Hier gilt für ein festes  $x_0 \in \mathbb{R}$

$$\sum_{k=0}^{\infty} f_k(x_0) = \begin{cases} \sum_{k=0}^{n_0} f_k(x_0) & \text{falls } x_0 \in ]0, 2[ \\ 0 & \text{sonst} \end{cases}$$

für  $n_0 = \max \left\{ m \mid m < \frac{2}{x_0} \right\}$ , also punktweise Konvergenz. Aber  $(f_n)$  ist nicht für alle Folgen  $(\alpha_n)$  gleichmäßig konvergent.

**Satz 2.1.16** (Majorantenkriterium für Funktionenreihen). *Gegeben sei die Funktionenreihe  $\sum f_k$  auf  $M$  und es gelte eine der beiden folgenden äquivalenten Bedingungen:*

A. *Es gibt eine positive Nullfolge  $(c_n)$  derart, dass  $\sum c_k$  konvergiert, und ein  $n_0 \in \mathbb{N}$ , sodass  $|f_n(x)| \leq c_n$  für alle  $n \geq n_0$  und für alle  $x \in M$ .*

B. Die Zahlenreihe  $\sum \|f_k\|$  konvergiert, d.h.  $\sum_{k=0}^{\infty} \|f_k\| < \infty$ .

Dann ist  $\sum f_k$  punktweise absolut konvergent und gleichmäßig konvergent.

**Beweisskizze.** [ $A. \implies B.$ ] Es ist  $\sum_{k=m}^n c_k < \epsilon$  für  $m, n$  groß genug. Damit ist für  $x \in M$  und diese  $m, n$

$$\left| \sum_{k=m}^n f_k(x) \right| \leq \sum_{k=m}^n |f_k(x)| \leq \sum_{k=m}^n c_k < \epsilon.$$

Das gilt auch für das Supremum. □

**Folgerung 2.1.17** ( $\hat{=}$  Satz 1.6.31). Es sei  $\sum a_k x^k$  eine Potenzreihe mit Konvergenzradius  $R$ . Dann gilt

1. Für jedes  $r < R$  konvergiert die Potenzreihe auf dem Intervall  $[-r, r]$  gleichmäßig.
2. Die Grenzfunktion der Potenzreihe ist auf dem gesamten Konvergenzreich stetig.

**Satz 2.1.18** (Konvergenz und Ableitung). *Es sei  $(f_n)$  eine Folge differenzierbarer Funktionen auf einem beschränkten Intervall  $I$ . Weiter gebe es ein  $x_0 \in I$  für das die Zahlenfolge  $(f_n(x_0))$  konvergiert. Außerdem sei die Folge der Ableitungsfunktionen  $(f'_n)$  auf  $I$  gleichmäßig konvergent. Dann gilt*

1. Die Funktionenfolge  $(f_n)$  konvergiert ebenfalls gleichmäßig.
2. Die Grenzfunktion  $f$  von  $(f_n)$  ist differenzierbar.
3.  $f'$  ist die Grenzfunktion von  $(f'_n)$ , d.h.

$$\left( \lim_{n \rightarrow \infty} f_n \right)' = \lim_{n \rightarrow \infty} f'_n.$$

**Bemerkung 2.1.19.** Man kann auf keine der beiden Voraussetzungen im Satz verzichten.

- Die gleichmäßige Konvergenz von  $(f'_n)$  ist notwendig für die obige Aussage; das heißt, die gleichmäßige Konvergenz von  $(f_n)$  und die Differenzierbarkeit von  $f$  reicht nicht aus.  
Dies zeigt das Beispiel  $f_n(x) = \frac{1}{n} \sin(nx)$ . Hier ist  $(f_n)$  gleichmäßig konvergent mit differenzierbarer Grenzfunktion  $f = 0$ . Aber die Folge  $(f'_n)$  der Ableitungen konvergiert nicht.
- Auch die Konvergenz von  $(f_n)$  in einem Punkt kann man nicht weglassen. Konvergiert etwa die Ableitungsfolge  $(f'_n)$  gleichmäßig, so divergiert in der Regel die Folge  $(n + f_n)$  in jedem Punkt, sie hat jedoch die gleiche Ableitungsfolge wie  $(f_n)$ .

**Folgerung 2.1.20** ( $\hat{=}$  **Satz 1.6.32**). Es sei  $\sum a_k x^k$  eine Potenzreihe mit Grenzfunktion  $f$ . Dann ist die Reihe der Ableitungen  $\sum (k+1)a_{k+1}x^k$  konvergent mit gleichem Konvergenzbereich und Grenzfunktion  $f'$ .

## 2.2 Integration

### 2.2.1 Stammfunktion und Integral

**Definition 2.2.1.** Es sei  $f : M \rightarrow \mathbb{R}$  eine Funktion auf der Menge  $M$ , die nur aus Häufungspunkten besteht.<sup>(i)</sup> Eine differenzierbare Funktion  $F : M \rightarrow \mathbb{R}$  heißt *Stammfunktion von  $f$* , wenn für alle  $x \in M$

$$F'(x) = f(x).$$

Existiert zu der Funktion  $f$  eine Stammfunktion, so nennen wir  $f$  *integrierbar*.

**Beispiel 2.2.2.** Es folgte eine kleine Liste von Funktionen und ihren Stammfunktionen. Die Liste ist nicht neu, denn man muss nur die Liste von Funktionen

---

<sup>(i)</sup>Das heißt  $M$  besitzt keine isolierten Punkte.

und ihren Ableitungen rückwärts lesen.

| Funktion $f$     | Stammfunktion $F$     | Funktion $f$             | Stammfunktion $F$ |
|------------------|-----------------------|--------------------------|-------------------|
| $x^c, c \neq -1$ | $\frac{x^{c+1}}{c+1}$ | $\frac{1}{1+x^2}$        | $\arctan(x)$      |
| $\frac{1}{x}$    | $\ln(x)$              | $\frac{1}{\sqrt{1-x^2}}$ | $\arcsin(x)$      |
| $e^x$            | $e^x$                 | $\cos(x)$                | $\sin(x)$         |
| $a^x$            | $\frac{a^x}{\ln(a)}$  | $\sin(x)$                | $-\cos(x)$        |

**Bemerkung 2.2.3.** • Die Stammfunktion einer Funktion  $f$  ist nicht eindeutig, denn für jede Konstante  $c \in \mathbb{R}$  ist mit  $F$  auch  $F + \underline{c}$  eine Stammfunktion zu  $f$ .

- Ist der Definitionsbereich ein Intervall, so gibt es umgekehrt für je zwei Stammfunktionen  $F_1$  und  $F_2$  der Funktion  $f$  eine Konstante  $c \in \mathbb{R}$  mit  $F_1 - F_2 = \underline{c}$ .

**Definition/Bemerkung 2.2.4.** Es sei  $f : I \rightarrow \mathbb{R}$  eine Funktion auf dem Intervall  $I$  mit Stammfunktion  $F : I \rightarrow \mathbb{R}$ . Sind  $a, b \in I$ , dann heißt der Ausdruck

$$\int_a^b f(t)dt := F(b) - F(a) =: F(x) \Big|_a^b$$

das *bestimmte Integral von  $f$  von  $a$  bis  $b$* . Der Wert  $a$  heißt dabei *untere* und  $b$  *obere Integralgrenze*.

Diese Definition ist unabhängig von der Wahl der Stammfunktion.

**Definition 2.2.5.** Als *unbestimmtes Integral* von  $f$  bezeichnen wir jede nicht näher bezeichnete Stammfunktion von  $f$  und schreiben dafür

$$\int f(x)dx.$$

**Bemerkung 2.2.6.** Ist  $f : I \rightarrow \mathbb{R}$  integrierbar, so ist für jedes  $a \in I$  die Funktion  $F_a : I \rightarrow \mathbb{R}$  mit

$$F_a(x) := \int_a^x f(t)dt$$

diejenige Stammfunktion von  $f$  mit  $F_a(a) = 0$ . Diese Definition ist wieder unabhängig von der Wahl der Stammfunktion zur Berechnung der rechten Seite. Damit gilt insbesondere

$$F'_a(x) = \left( \int_a^x f(t) dt \right)' = f(x).$$

**Bemerkung 2.2.7** (Zwischenstand). • Um die Integrierbarkeit einer Funktion nachzuweisen, bleibt uns zurzeit lediglich die Möglichkeit eine Stammfunktion explizit anzugeben. D.h. wir haben kein "Verfahren" eine Stammfunktion zu suchen.

- Schön wäre es ja zumindest ein hinreichendes Kriterium für die Integrierbarkeit einer Funktion zu benennen. Solch ein "schönes" Kriterium wäre z.B. die Stetigkeit von  $f$  als hinreichend für die Existenz einer Stammfunktion, auch wenn man diese gegebenenfalls gar nicht explizit angeben kann.

- Unabhängig davon, sieht man jedoch hier schon, dass es auch integrierbare Funktionen gibt, die nicht stetig sind: z.B. ist die Funktion

$$F(x) := \begin{cases} \sqrt{x^3} \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases} \text{ differenzierbar, aber die Ableitung}$$

$$F'(x) = \begin{cases} \frac{3}{2}\sqrt{x} \sin\left(\frac{1}{x}\right) - \frac{1}{\sqrt{x}} \cos\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases} \text{ ist nicht stetig in } x = 0,$$

aber natürlich integrierbar.

- Wir werden in einem späteren Abschnitt einen weiteren Begriff der Integrierbarkeit kennenlernen, der es uns in gewissen Fällen ermöglicht eine Stammfunktion (bzw. ihren Wert an einer vorgegebenen Stelle) konstruktiv zu berechnen.
- Wir werden dabei auch diskutieren, wann dieser neue Integrationsbegriff mit unserem bisherigen Integrationsbegriff über Stammfunktionen übereinstimmt und wann es Unterschiede gibt.

### 2.2.2 Integrationsregeln

Mit Hilfe von Bemerkung 2.2.6 lassen sich bereits sämtliche Rechenregeln der Integration herleiten und angeben. Die meisten von ihnen folgen direkt

aus den Definitionen des vorigen Abschnitts oder sind Folgerungen aus den Differentiationsregeln aus Abschnitt 1.5.

**Satz 2.2.8.** *Es seien  $f, f_1, f_2 : I \rightarrow \mathbb{R}$  auf dem Intervall  $I \subset \mathbb{R}$  definierte, integrierbare Funktionen und  $a, b \in I$ . Dann gilt: <sup>(i)</sup>*

$$1. \int_a^b f(t)dt = - \int_b^a f(t)dt.$$

$$2. \text{ F\u00fcr alle } c \in I \text{ ist } \int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

$$3. \int_a^b (f_1(t) \pm f_2(t))dt = \int_a^b f_1(t)dt \pm \int_a^b f_2(t)dt.$$

$$4. \text{ F\u00fcr alle } c \in \mathbb{R} \text{ gilt } \int_a^b (cf(t))dt = c \int_a^b f(t)dt.$$

$$5. \text{ Ist } f(x) \geq 0 \text{ f\u00fcr alle } x \in I \text{ und ist } a \leq b, \text{ so gilt } \int_a^b f(t)dt \geq 0.$$

$$6. \text{ Ist } f_1(x) \geq f_2(x) \text{ f\u00fcr alle } x \in I \text{ und ist } a \leq b, \text{ so gilt}$$

$$\int_a^b f_1(t)dt \geq \int_a^b f_2(t)dt.$$

$$7. \text{ Ist auch } |f| \text{ integrierbar und ist } a \leq b, \text{ so gilt}$$

$$\left| \int_a^b f(t)dt \right| \leq \int_a^b |f(t)|dt.$$

**Satz 2.2.9** (Partielle Integration). *Sind  $f_1, f_2 : I \rightarrow \mathbb{R}$  differenzierbar und ist das Produkt  $f_1 f_2' : I \rightarrow \mathbb{R}$  integrierbar, dann ist auch das Produkt  $f_1' f_2 : I \rightarrow \mathbb{R}$  integrierbar und es gilt f\u00fcr alle  $a, b \in I$*

$$\int_a^b f_1'(t)f_2(t)dt = f_1(x)f_2(x) \Big|_a^b - \int_a^b f_1(t)f_2'(t)dt.$$

<sup>(i)</sup>Die Rechenregeln 3. und 4. gelten analog auch f\u00fcr unbestimmte Integrale.

**Satz 2.2.10** (Substitutionsregel). *Es sei  $g : I \rightarrow J$  eine differenzierbare Funktion dessen Bild im Intervall  $J \subset \mathbb{R}$  enthalten ist. Ist nun  $f : J \rightarrow \mathbb{R}$  integrierbar, so auch  $(f \circ g) \cdot g' : I \rightarrow \mathbb{R}$  und es gilt für alle  $a, b \in I$*

$$\int_a^b f(g(t)) g'(t) dt = \int_{g(a)}^{g(b)} f(t) dt.$$

**Bemerkung 2.2.11.** Angewendet wird der Satz auch wie folgt:

Gesucht sei die Stammfunktion  $F$  der Funktion  $f : [a, b] \rightarrow \mathbb{R}$ , bzw. das bestimmte Integral  $\int_a^b f(t) dt$ . Dazu suche man eine bijektive, differenzierbare Funktion  $g : [c, d] \rightarrow [a, b]$  mit  $\{c, d\} = \{g^{-1}(a), g^{-1}(b)\}$ , sodass die Funktion  $(f \circ g) \cdot g' : [c, d] \rightarrow \mathbb{R}$  eine Stammfunktion hat. Dann gilt wegen des vorigen Satzes

$$\int_a^b f(t) dt = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(t)) g'(t) dt.$$

Ist nun  $\tilde{F} : [c, d] \rightarrow \mathbb{R}$  eine Stammfunktion zu  $(f \circ g) \cdot g' : [c, d] \rightarrow \mathbb{R}$ , so ist  $F = \tilde{F} \circ g^{-1} : [a, b] \rightarrow \mathbb{R}$  eine Stammfunktion zu  $f$ .

**Beispiel 2.2.12.** • Es sei  $p$  ein Polynom  $n$ -ten Grades und  $a \in \mathbb{R}$ . Die folgenden Integrale berechnet man mit  $n + 1$ -maliger partieller Integration:

$$\int p(x) e^{ax} dx, \quad \int p(x) \sin(ax) dx, \quad \int p(x) \cos(ax) dx, \\ \int p(x) \sinh(ax) dx, \quad \int p(x) \cosh(ax) dx.$$

- Wir suchen eine Stammfunktion von  $\frac{\sqrt{\tan(x)}}{\cos^2(x)}$ . Es ist

$$\frac{\sqrt{\tan(x)}}{\cos^2(x)} = \sqrt{\tan(x)} \tan'(x),$$

und wir wählen in der Substitution  $f(x) = \sqrt{x}$  und  $g(x) = \tan(x)$ , sodass

$$\int_a^b \frac{\sqrt{\tan(x)}}{\cos^2(x)} dx = \int_a^b \sqrt{\tan(x)} \tan'(x) dx$$

$$\begin{aligned}
&= \int_{\tan(a)}^{\tan(b)} \sqrt{t} \, dt \\
&= \frac{2}{3} \sqrt{t^3} \Big|_{\tan(a)}^{\tan(b)} \\
&= \frac{2}{3} \sqrt{\tan(x)} \Big|_a^b
\end{aligned}$$

- Eine formale Durchführung der Substitution verläuft für das vorige Beispiel etwa wie folgt

$$\int_a^b \frac{\sqrt{\tan(x)}}{\cos^2(x)} dx = \int_{\tan(a)}^{\tan(b)} \frac{\sqrt{t}}{\cos^2(x)} \cos^2(x) dt$$

↑

setze  $t := \tan(x)$

$\implies t' = \frac{dt}{dx} = \frac{1}{\cos^2(x)} \iff dx = \cos^2(x) dt$

$t(a) = \tan(a), t(b) = \tan(b)$

das wird nun in das Ausgangsintegral  
"substituiert", wobei alle  $x$  durch  
 $t$  ersetzt werden müssen.

$$= \int_{\tan(a)}^{\tan(b)} \sqrt{t} \, dt = \dots$$

- Eine Stammfunktion von  $f(x) = \frac{\cos(x)}{\sin^2(x)}$  findet man etwa durch die Substitution  $u = \tan(\frac{x}{2})$ . Dazu nutzen wir

$$\cos(x) = \frac{1 - u^2}{1 + u^2} \quad \text{und} \quad \sin(x) = \frac{2u}{1 + u^2}.$$

Das gibt

$$\int_a^b \frac{\cos(x)}{\sin^2(x)} dx = \int_{\tan(\frac{a}{2})}^{\tan(\frac{b}{2})} \frac{(1 + u^2)^2}{4u^2} \cdot \frac{1 - u^2}{1 + u^2} \cdot \frac{2 du}{1 + u^2}$$

$\uparrow$  setze  $u := \tan(\frac{x}{2})$   
 $\implies u' = \frac{du}{dx} = \frac{1}{2}(1 + \tan^2(\frac{x}{2})) \iff dx = \frac{2du}{1+u^2}$   
 $u(a) = \tan(\frac{a}{2}), u(b) = \tan(\frac{b}{2})$   
 das wird nun in das Ausgangsintegral  
 "substituiert", wobei alle  $x$  durch  
 $u$  ersetzt werden müssen.

$$\begin{aligned}
 &= \frac{1}{2} \int_{\tan(\frac{a}{2})}^{\tan(\frac{b}{2})} \left( \frac{1}{u^2} - 1 \right) du \\
 &= \frac{1}{2} \left( -\frac{1}{u} - u \right) \Big|_{\tan(\frac{a}{2})}^{\tan(\frac{b}{2})} \\
 &= -\frac{1}{2} \left( \frac{1}{\tan(\frac{x}{2})} + \tan(\frac{x}{2}) \right) \Big|_a^b \\
 &= -\frac{1}{2} \left( \frac{1 + \tan^2(\frac{x}{2})}{\tan(\frac{x}{2})} \right) \Big|_a^b \\
 &= -\left( \frac{1}{\sin(x)} \right) \Big|_a^b
 \end{aligned}$$

- Schneller hätte man das vorige Ergebnis mit der Substitution  $v = \sin(x)$  erhalten.

⚠ Aber: Für Integrale vom Typ  $\int R(\cos(x), \sin(x)) dx$  wobei  $R(u, v)$  eine rationale Funktion in den beiden Variablen  $(u, v)$  ist, führt die Substitution  $t = \tan(\frac{x}{2})$  immer auf das Integral einer rationalen Funktion  $R(t)$ .

Wenden wir uns – auch wegen des letzten Beispiels – nun den rationalen Funktionen zu.

Die Aussage aus Folgerung 1.3.21 kann man noch wie folgt verfeinern.

**Satz 2.2.13** (Faktorisierungssatz für reelle Polynome). *Es sei  $p$  ein Polynom  $n$ -ten Grades mit Leitkoeffizient  $a_n = 1$  und den Nullstellen  $x_1, \dots, x_m$  mit Vielfachheiten  $r_1, \dots, r_m$ . Weiter sei  $r := \sum_{i=1}^m r_i$ . Dann ist  $n - r = 2s$  gerade*

*und es gibt Zahlen  $b_j, c_j$  für  $j = 1, \dots, p$  und Zahlen  $s_1, \dots, s_p$  mit  $s = \sum_{j=1}^p s_j$ ,*

sodass Folgendes gilt:

$$p(x) = (x - x_1)^{r_1} \cdot \dots \cdot (x - x_m)^{r_m} (x^2 + b_1x + c_1)^{s_1} \cdot \dots \cdot (x^2 + b_px + c_p)^{s_p}$$

und die quadratischen Polynome  $x^2 + b_jx + c_j$  haben keine Nullstellen.

**Satz 2.2.14** (Partialbruchzerlegung). *Das Polynom  $p$  habe die Darstellung aus dem obigen Satz und es sei  $\text{grad}(q) < \text{grad}(p)$ . Dann gibt es  $r$  Zahlen  $A_1^1, \dots, A_{r_1}^1, \dots, A_1^m, \dots, A_{r_m}^m$  und  $2s$  Zahlen  $B_1^1, C_1^1, \dots, B_{s_1}^1, C_{s_1}^1, \dots, B_1^p, C_1^p, \dots, B_{s_p}^p, C_{s_p}^p$  mit*

$$\begin{aligned} \frac{q(x)}{p(x)} &= \frac{A_1^1}{x - x_1} + \frac{A_2^1}{(x - x_1)^2} + \dots + \frac{A_{r_1}^1}{(x - x_1)^{r_1}} + \\ &+ \dots + \\ &+ \frac{A_1^m}{x - x_m} + \frac{A_2^m}{(x - x_m)^2} + \dots + \frac{A_{r_m}^m}{(x - x_m)^{r_m}} + \\ &+ \frac{B_1^1x + C_1^1}{x^2 + b_1x + c_1} + \frac{B_2^1x + C_2^1}{(x^2 + b_1x + c_1)^2} + \dots + \frac{B_{s_1}^1x + C_{s_1}^1}{(x^2 + b_1x + c_1)^{s_1}} + \\ &+ \dots + \\ &+ \frac{B_1^px + C_1^p}{x^2 + b_px + c_p} + \frac{B_2^px + C_2^p}{(x^2 + b_px + c_p)^2} + \dots + \frac{B_{s_p}^px + C_{s_p}^p}{(x^2 + b_px + c_p)^{s_p}}. \end{aligned}$$

**Folgerung 2.2.15** (Integration rationaler Funktionen). Die Integration der rationalen Funktion  $f(x) = \frac{p(x)}{q(x)}$  lässt sich auf die Integration der folgenden Typen elementarer rationaler Funktionen zurückführen:

1.  $\int \frac{dx}{x - x_0} = \ln |x - x_0| + \alpha$
2.  $\int \frac{dx}{(x - x_0)^{n+1}} = -\frac{1}{n(x - x_0)^n} + \alpha$  für  $n > 0$ .
3.  $\int \frac{dx}{x^2 + bx + c} = \frac{2}{\sqrt{-D}} \arctan \left( \frac{2x + b}{\sqrt{-D}} \right) + \alpha$
4.  $\int \frac{dx}{(x^2 + bx + c)^{n+1}} = -\frac{2x + b}{nD(x^2 + bx + c)^n} - \frac{2(2n - 1)}{nD} \int \frac{dx}{(x^2 + bx + c)^n}$   
für  $n > 0$ .

wobei 3. im Fall  $D = b^2 - 4c < 0$  gilt und 4. auch speziell in diesem Fall verwendet wird. Außer 3. und 4. brauchen wir keine weiteren Integrale, denn es ist

$$\frac{Bx + C}{(x^2 + bx + c)^{n+1}} = \frac{B}{2} \frac{2x + b}{(x^2 + bx + c)^{n+1}} + \frac{2C - Bb}{2} \frac{1}{(x^2 + bx + c)^{n+1}}.$$

Der zweite Summand ist nun vom Typ 3. oder 4. und der erste Summand lässt sich direkt integrieren, denn es gilt

$$\int \frac{2x + b}{x^2 + bx + c} dx = \ln(x^2 + bx + c) + \alpha,$$

$$\int \frac{2x + b}{(x^2 + bx + c)^{n+1}} dx = -\frac{1}{n} \frac{1}{(x^2 + bx + c)^n} + \alpha \quad \text{für } n > 0.$$

### 2.2.3 Das Riemann-Integral

Wir nehmen zunächst an, dass wir von einer Funktion  $F : [a, b] \rightarrow \mathbb{R}$  die Ableitung  $F' = f : [a, b] \rightarrow \mathbb{R}$  sowie den Anfangspunkt  $F(a)$  kennen. Können wir nun Aussagen über den Wert  $F(b)$  machen?

Ein "Auflösen" der Gleichung aus Definition/Bemerkung 2.2.4 ist sicher nicht zielführend, da das bestimmte Integral gerade durch die Kenntnis der beiden Werte  $F(a)$  und  $F(b)$  definiert ist.

Wir kennen jedoch den Mittelwertsatz 1.5.22, der uns sagt, dass es einen Wert  $\xi \in ]a, b[$  gibt, mit  $F(b) = F(a) + f(\xi)(b - a)$ . Nun hilft uns das auch nicht weiter, da wir den Wert  $\xi$  nicht kennen.

Im Fall, dass die Ableitungsfunktion  $f = F'$  positiv ist, können wir diese Aussage allerdings geometrisch interpretieren: Der Wert des bestimmten Integrals  $F(b) - F(a)$  entspricht der Fläche eines Rechtecks mit der Grundseite  $b - a$  und der Höhe  $f(\xi)$ .

Steigt die Funktion  $f$  nun nicht zu stark, ist die Wahl eines beliebigen  $\xi$  schon eine gute Näherung. Diese Annahme an  $f$  kann man aber in der Regel nicht erwarten. Ist  $f$  stattdessen stetig, so wissen wir, dass die Funktion auch nur ein wenig steigt, wenn man sich nur wenig von  $a$  fortbewegt. Dies wollen wir als Grundidee zur Berechnung eines Näherungswertes von  $F(b)$  heranziehen.

**Definition 2.2.16** (Zerlegung eines Intervalls). • Es sei  $[a, b] \subset \mathbb{R}$  ein beschränktes Intervall. Eine Teilmenge  $Z = \{x_0, x_1, \dots, x_k\} \subset \mathbb{R}$  mit  $a = x_0 < x_1 < \dots < x_{k-1} < x_k = b$  heißt *Zerlegung von  $[a, b]$  der Länge*

$k$ . Wir schreiben  $h_j := x_j - x_{j-1}$  für  $1 \leq j \leq k$ , sodass  $x_j = a + \sum_{i=1}^j h_i$ .

Weiter schreiben wir  $|Z| := \max_{1 \leq j \leq k} h_j$ .

- Eine Zerlegung heißt *äquidistant*, wenn  $h_j = x_j - x_{j-1}$  für alle  $1 \leq j \leq k$  konstant ist. In diesem Fall ist  $h_j = h = \frac{b-a}{k}$  und es ist  $x_j = a + jh$  für  $0 \leq j \leq k$ .
- Eine Zerlegung  $Z'$  heißt *Verfeinerung* der Zerlegung  $Z$ , wenn  $Z \subset Z'$  und  $|Z'| \leq |Z|$ . Gilt sogar  $|Z'| < |Z|$  so heißt  $Z'$  eine *echte* Verfeinerung von  $Z$ .

**Beispiel 2.2.17.** • Für zwei Zerlegungen  $Z_1$  und  $Z_2$  ist die Zerlegung  $Z_1 \cup Z_2$  eine gemeinsame Verfeinerung. Diese ist aber nicht notwendig echt.

- Sind  $Z_1$  und  $Z_2$  äquidistante Zerlegungen der Länge  $k_1$  und  $k_2$  mit  $k_1 > k_2$  und ist  $Z_2 \subset Z_1$ , dann ist  $Z_1$  eine echte Verfeinerung von  $Z_2$ . Das ist zum Beispiel der Fall, wenn  $k_1$  und  $k_2$  beide Potenzen von 2 sind.

Es sei nun  $Z$  eine Zerlegung von  $[a, b]$  und wie oben  $h_j := x_j - x_{j-1}$ . Es ist dann

$$F(b) - F(a) = \sum_{j=1}^k F(x_j) - F(x_{j-1})$$

und wir finden wegen des Mittelwertsatzes 1.5.22 jeweils  $\xi_j \in ]x_{j-1}, x_j[$ , sodass

$$F(b) = F(a) + \sum_{j=1}^k f(\xi_j)h_j.$$

Wählt man nun die  $\xi_i$  wieder beliebig in den Teilintervallen, so erhält man durch die obige Formel eine Näherung für den Wert  $F(b)$ . Diese Näherung wird hoffentlich besser, wenn wir die Zerlegung feiner wählen, d.h. wenn  $k$  größer und  $h_j$  kleiner wird.

**Definition 2.2.18.** Es sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion. Weiter sei  $(Z_n)_{n \in \mathbb{N}} = (Z_1, Z_2, \dots)$  eine Folge von Zerlegungen  $Z_n = \{x_0^n, x_1^n, \dots, x_{k_n}^n\}$  von  $[a, b]$  mit  $\lim_{n \rightarrow \infty} |Z_n| = 0$ . Dazu sei  $(\vec{\xi}_n)_{n \in \mathbb{N}}$  eine Folge von Vektoren  $\vec{\xi}_n = (\xi_1^n, \dots, \xi_{k_n}^n)$  mit  $\xi_j^n \in ]x_{j-1}^n, x_j^n[$  und wir schreiben wie oben  $h_j^n = x_j^n - x_{j-1}^n$  für  $1 \leq j \leq k_n$ .

Dann heißt

$$R_{f;a,b}(Z_n, \vec{\xi}_n) := \sum_{j=1}^{k_n} f(\xi_j^n) h_j^n$$

Riemann-Zwischensumme zu  $f$  und  $[a, b]$  (zur Zerlegungsfolge  $(Z_n)$  und zur Zwischenwertfolge  $(\vec{\xi}_n)$ ). Die Folge

$$\left( R_{f;a,b}(Z_n, \vec{\xi}_n) \right)_{n \in \mathbb{N}}$$

heißt *Riemann-Folge zu  $f$  und  $[a, b]$*  (zur Zerlegungsfolge  $(Z_n)$  und zur Zwischenwertfolge  $(\vec{\xi}_n)$ ).

**Bemerkung 2.2.19.** 1. Wenn alle Riemann-Folgen konvergieren, dann gegen den selben Grenzwert.

2. Wenn alle Riemann-Folgen konvergieren, dann kann man zur Zerlegung  $Z_n$  als Einträge des Zwischenwertvektors  $\vec{\xi}_n$  stets die Zwischenstellen gemäß des Mittelwertsatzes wählen. Damit gilt

$$F(b) = F(a) + \lim_{n \rightarrow \infty} R_{f;a,b}(Z_n, \vec{\xi}_n).$$

und dies muss auch für alle anderen Wahlen  $(\vec{\xi}_n)$  richtig sein.

Wir verallgemeinern die Vorbetrachtungen zur folgenden Definition.

**Definition 2.2.20.** • Eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  heißt *Riemann-integrierbar (R-integrierbar)*, wenn alle Riemann-Folgen konvergieren. Für den Grenzwert schreiben wir

$$\int_a^b f(t) dt$$

und nennen diesen das *Riemann-Integral (R-Integral) von  $f$  in den Grenzen von  $a$  bis  $b$* .

- Wir definieren zusätzlich für  $b < a$

$$\int_a^b f(t) dt := - \int_b^a f(t) dt$$

und wenn  $f$  in  $a$  definiert ist

$$\int_a^a f(t) dt := 0.$$

**Satz 2.2.21.** 1. Eine  $R$ -integrierbare Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist beschränkt.

2. Sind  $f, g : [a, b] \rightarrow \mathbb{R}$   $R$ -integrierbar und gilt  $f \leq g$  auf  $[a, b]$ , dann ist

$$\int_a^b f(t) dt \leq \int_a^b g(t) dt.$$

3. Ist  $f : [a, b] \rightarrow \mathbb{R}$   $R$ -integrierbar, dann gilt

$$\int_a^b f(t) dt \leq \|f\|(b-a).$$

Die Diskussion bisher liefert nun den folgenden ersten Teil des Hauptsatzes der Differential- und Integralrechnung

**Satz 2.2.22** (Hauptsatz der Differential- und Integralrechnung – Teil 1).  
Besitzt  $F : [a, b] \rightarrow \mathbb{R}$  eine  $R$ -integrierbare Ableitung  $F'$ , dann gilt

$$\int_a^b F'(t) dt = F(b) - F(a).$$

**Bemerkung 2.2.23.** 1. Zur Berechnung von  $\int_a^b f(t) dt$  muss man den obigen Grenzprozess durchführen, da  $f$  in der Regel nicht als Ableitung gegeben ist.

2. Wissen wir jedoch, dass  $\int_a^b f(t) dt$  existiert und kennen darüberhinaus eine Stammfunktion  $F$  von  $f$ , so gilt

$$\int_a^b f(t) dt = \int_a^b f(t) dt = F(b) - F(a).$$

3. Ist  $F'$  auf  $[a, b]$  stetig, so auch auf  $[a, x]$   $a \leq x \leq b$ . Wegen  $F(x) = F(a) + \int_a^x F'(t) dt$  haben wir damit die Funktion  $F$  aus der Kenntnis seiner Änderung  $F'$  zurückerhalten.

4. Es sind nicht alle Ableitungen R-integrierbar, obwohl sie selbstverständlich Stammfunktionen besitzen, wie das Beispiel

$$F(x) = \begin{cases} x^{\frac{3}{2}} \sin\left(\frac{1}{x}\right) & \text{falls } x \neq 0 \\ 0 & \text{falls } x = 0 \end{cases}$$

mit

$$F'(x) = \begin{cases} \frac{3}{2}x^{\frac{1}{2}} \sin\left(\frac{1}{x}\right) - x^{-\frac{1}{2}} \cos\left(\frac{1}{x}\right) & \text{falls } x \neq 0 \\ 0 & \text{falls } x = 0 \end{cases}$$

zeigt.

5. Nicht alle R-integrierbaren Funktionen sind stetig, wie das Beispiel der Funktion  $H : [-1, 1] \rightarrow \mathbb{R}$  mit  $H(x) = \begin{cases} 0 & \text{falls } x \in [-1, 0] \\ 1 & \text{falls } x \in ]0, 1] \end{cases}$ .

Im nächsten Satz formulieren wir die Konvergenz der Riemann-Folgen mit Hilfe eines  $\epsilon$ - $\delta$ -Arguments

**Satz 2.2.24.**  $f : [a, b] \rightarrow \mathbb{R}$  ist genau dann R-integrierbar, wenn

$$\forall \epsilon > 0 \exists \delta > 0 \forall |Z| < \delta \forall \vec{\xi} : \left| R_{f;a,b}(Z, \vec{\xi}) - \int_a^b f(t) dt \right| < \epsilon.$$

Damit lässt sich ähnlich wie schon bei Zahlenfolgen und Funktionenfolgen die Frage nach der R-Integrierbarkeit auch ohne Kenntnis des Grenzwertes beantworten. Es gilt das folgende *Cauchy-Konvergenzkriterium*:

**Satz 2.2.25.**  $f : [a, b] \rightarrow \mathbb{R}$  ist genau dann R-integrierbar, wenn

$$\forall \epsilon > 0 \exists \delta > 0 \forall |Z|, |Z'| < \delta \forall \vec{\xi}, \vec{\xi}' : \left| R_{f;a,b}(Z, \vec{\xi}) - R_{f;a,b}(Z', \vec{\xi}') \right| < \epsilon.$$

Damit lässt sich bequem die folgende, in praktischen Anwendungen verwendete Eigenschaft begründen

**Bemerkung 2.2.26.** 1. Ist  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion und  $c \in [a, b]$  derart, dass  $f|_{[a,c]}$  und  $f|_{[c,b]}$  R-integrierbar sind, so ist auch  $f$  R-integrierbar, und es gilt

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt.$$

In 1. gilt sogar die Äquivalenz:

- 1'. Eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist genau dann R-integrierbar, wenn für alle  $c \in [a, b]$  die Funktionen  $f|_{[a,c]}$  und  $f|_{[c,b]}$  R-integrierbar sind. Dann gilt die Gleichheit aus 1.

Den Beweis für Aussage 1' werden wir später nachholen. Für eine sehr spezielle Situation werden wir das Ergebnis jedoch gleich als direkte Folgerung erhalten.

**Satz 2.2.27.** *Jede stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist R-integrierbar.*

**Folgerung 2.2.28.** Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig, und sind  $y_1, y_2, y_3 \in [a, b]$ , so gilt

$$\int_{y_1}^{y_3} f(t) dt = \int_{y_1}^{y_2} f(t) dt + \int_{y_2}^{y_3} f(t) dt.$$

**Folgerung 2.2.29.** Jede beschränkte Funktion  $f : [a, b] \rightarrow \mathbb{R}$  mit nur endlich vielen Unstetigkeitsstellen ist R-integrierbar.

**Satz 2.2.30.** 1. *Es seien  $f, g : [a, b] \rightarrow \mathbb{R}$  R-integrierbar und  $c \in \mathbb{R}$ . Dann sind auf  $f + g$  und  $cf$  R-integrierbar mit*

$$\begin{aligned} \int_a^b (f + g)(t) dt &= \int_a^b f(t) dt + \int_a^b g(t) dt \\ \int_a^b (cf)(t) dt &= c \int_a^b f(t) dt. \end{aligned}$$

2. *Es sei  $R([a, b]) = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ ist R-integrierbar}\}$ . Dann gilt die folgende Inklusion bezüglich der Supremumsnorm normierter Vektorräume:*

$$C([a, b]) \subsetneq R([a, b]) \subsetneq B([a, b]).$$

*Insbesondere ist die Abbildung  $f \mapsto \int_a^b f(t) dt$  eine Linearform auf  $R([a, b])$  und auf  $C([a, b])$ .*

**Beweisskizze.** (zu 2.) Die Inklusionen folgen aus den Sätzen 2.2.21 und 2.2.27. Dass es sich hier um echte Inklusionen handelt, folgt einerseits aus

Folgerung 2.2.29 und andererseits mit Hilfe der beschränkten, jedoch nicht  $\mathbb{R}$ -integrierbaren Funktion

$$f : [a, b] \rightarrow \mathbb{R}, \quad f(x) = \text{ind}_{\mathbb{Q} \cap [a, b]}(x).$$

□

**Satz 2.2.31** (Hauptsatz der Differential- und Integralrechnung – Teil 2). *Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig, dann besitzt  $f$  eine Stammfunktion  $F : [a, b] \rightarrow \mathbb{R}$ , etwa*

$$F(x) = \int_a^x f(t) dt.$$

**Bemerkung 2.2.32.** Der Hauptsatz der Differential- und Integralrechnung erlaubt es uns

- wegen Teil 1 bei der Kenntnis einer Stammfunktion  $F$  der  $\mathbb{R}$ -integrierbaren Funktion  $f$  den Wert  $\int_a^b f(t) dt$  des Konvergenzprozesses explizit anzugeben.
- wegen Teil 2 zu vorgegebener stetiger Funktion  $f : [a, b] \rightarrow \mathbb{R}$  durch einen Konvergenzprozess eine Stammfunktion  $F : [a, b] \rightarrow \mathbb{R}$  von  $f$  anzugeben, d.h.  $F' = f$ .

**Bemerkung 2.2.33.** Sind  $f, g : [a, b] \rightarrow \mathbb{R}$   $\mathbb{R}$ -integrierbar, und stimmen  $f$  und  $g$  auf einer *dichten* Teilmenge von  $[a, b]$  überein, dann gilt

$$\int_a^b f(t) dt = \int_a^b g(t) dt.$$

**Bezeichnung 2.2.34.** Wegend des Hauptsatzes der Differential- und Integralrechnung werden wir, weil wir uns in der Regel mit mindestens stückweise stetigen Funktionen beschäftigen, nicht mehr zwischen den Symbolen  $\int$  und  $\int$  unterscheiden.

**Folgerung 2.2.35.** Wegen der starken Voraussetzungen der Sätze 2.2.9 und 2.2.10 gelten die partielle Integration und die Substitutionsregel für die  $\mathbb{R}$ -Integration analog.

Abschließend untersuchen wir das Verhalten der Integration im Zusammenhang mit Funktionenfolgen.

**Satz 2.2.36.** *Es sei  $(f_n)$  eine gleichmäßig konvergente Folge integrierbarer Funktionen auf dem Intervall  $[a, b]$ . Dann gilt*

1. Die Zahlenfolge  $\left(\int_a^b f_n(t) dt\right)$  ist konvergent.
2. Die Grenzfunktion  $f = \lim_{n \rightarrow \infty} f_n$  von  $(f_n)$  ist integrierbar mit

$$\int_a^b f(t) dt = \lim_{n \rightarrow \infty} \int_a^b f_n(t) dt.$$

Man kann also sagen: "Unter den gegebenen Voraussetzungen lassen sich Integration und Grenzwertbildung vertauschen".

Übertragen wir das auf Potenzreihen erhalten wir

**Folgerung 2.2.37.** Hat die Potenzreihe  $\sum a_k x^k$  einen Konvergenzbereich  $K \neq \emptyset$ , so ist die Grenzfunktion dort integrierbar, und eine Stammfunktion erhält man durch summandenweises integrieren, also

$$\int_a^b \left(\sum a_k x^k\right) dx = \sum \frac{a_{k+1}}{k+1} x^{k+1} \Big|_a^b.$$

für  $a, b \in K$ .

## 2.2.4 Geometrische Anwendungen der Integration

### 2.2.4.1 Flächen- und Volumenberechnung

Der Mittelwertsatz der Differentialrechnung [1.5.22](#) liefert die folgende Aussage für Integrale:

**Satz 2.2.38** (Mittelwertsatz der Integralrechnung). *Es sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig, dann gibt es ein  $\xi \in [a, b]$ , sodass*

$$\int_a^b f(t) dt = f(\xi)(b - a).$$

**Bemerkung 2.2.39.** Die obige Aussage gilt in der Regel nicht mehr, wenn man die Stetigkeit fallen lässt.

Dazu betrachten wir die Funktion  $f : [-1, 1] \rightarrow \mathbb{R}$  mit  $f(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$ .

Es ist hier  $\int_a^b f(t) dt = 1$  und  $f(\xi)(b-a) \in \{0, 2\}$  für alle  $\xi \in [a, b]$ .

Betrachten wir nun eine stetige Funktion  $f$ , die nur positive Werte annimmt. Dann entspricht der rechte Term der Gleichung aus Satz 2.2.38 dem Flächeninhalt eines Rechtecks mit Grundseitenlänge  $b-a$  und Höhe  $f(\xi)$ . Schauen wir nun die Form der Riemannsumme in Definition 2.2.18 genauer an, so sehen wir, dass  $R_{f;a,b}(Z_n, \vec{\xi}_n)$  einer Summe von solchen Rechteckflächen entspricht. Für geeignete  $\xi_j^n$  gemäß Satz 2.2.38 gilt

$$\int_a^b f(t) dt = \sum \int_{x_{i-1}}^{x_i} f(t) dt = \sum f(\xi_j^n)(x_i^n - x_{i-1}^n) = R_{f;a,b}(Z_n, \vec{\xi}_n).$$

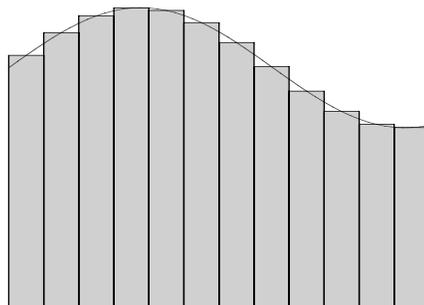
Das liefert dann die folgende geometrische Interpretation des Integrals, siehe auch Abbildung 2.2.1.

**Bezeichnung 2.2.40.** Für zwei Funktionen  $f_1, f_2 : [a, b] \rightarrow \mathbb{R}$  bezeichne  $A(f_1, f_2; a, b)$  den Inhalt der Fläche, die von den Graphen der beiden Funktionen eingeschlossen wird. Als Spezialfall bezeichnet  $A(f; a, b) := A(f, 0; a, b)$  den Inhalt der Fläche zwischen der  $x$ -Achse und dem Graphen von  $f : [a, b] \rightarrow \mathbb{R}$ .

**Satz 2.2.41.** Für eine nicht-negative, stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}^{\geq 0}$  gilt

$$A(f; a, b) = \int_a^b f(t) dt.$$

Abbildung 2.2.1: Fläche unter einem Graphen



**Folgerung 2.2.42.** 1. Es sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig mit den Nullstellen  $x_1 < \dots < x_{\ell-1}$  in  $]a, b[$ . Schreiben wir  $a = x_0, b = x_\ell$ , so gilt

$$A(f; a, b) = \sum_{k=1}^{\ell} \left| \int_{x_{k-1}}^{x_k} f(t) dt \right|.$$

2. Sind  $f_1, f_2 : [a, b] \rightarrow \mathbb{R}$  stetig und ist stets  $f_1 \geq f_2$ , so gilt

$$A(f_1, f_2; a, b) = \int_a^b (f_1(t) - f_2(t)) dt.$$

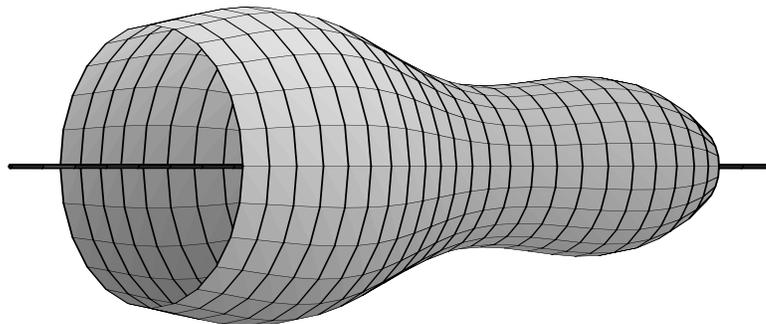
3. Es seien  $f_1, f_2 : [a, b] \rightarrow \mathbb{R}$  stetig mit den Schnittstellen  $x_1 < \dots < x_{\ell-1}$  in  $]a, b[$ , so gilt mit  $a = x_0, b = x_\ell$

$$A(f_1, f_2; a, b) = \sum_{k=1}^{\ell} \left| \int_{x_{k-1}}^{x_k} (f_1(t) - f_2(t)) dt \right|.$$

**Definition 2.2.43.** Unter einem Rotationskörper im weiteren Sinne verstehen wir einen Körper, der entsteht, wenn wir eine "stetige ebene Kurve" um eine in der Kurvenebene liegende Gerade rotieren lassen. Im engeren Sinne verstehen wir unter einem Rotationskörper hier die folgenden Spezialfälle:

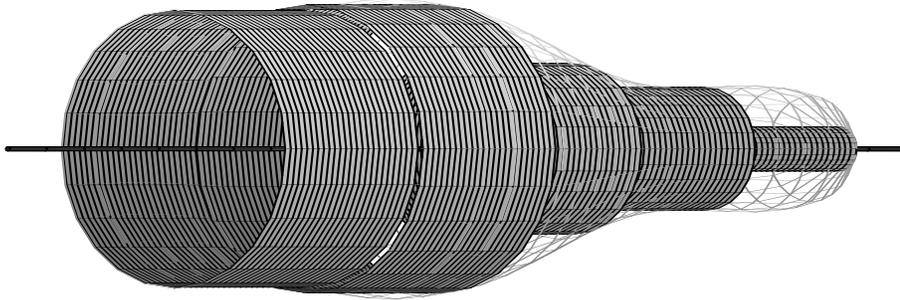
1. Ist  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion, so bezeichnet  $V_x(f; a, b)$  das Volumen des Körpers, der entsteht, wenn man den Graphen der Funktion  $f$  um die  $x$ -Achse (Abzisse) rotieren lässt, siehe Abbildung 2.2.2.
2. Ist  $f : [a, b] \rightarrow \mathbb{R}$  eine injektive, stetige Funktion, so bezeichnet  $V_y(f; a, b)$  das Volumen des Körpers, der entsteht, wenn man den Graphen der Funktion  $f$  um die  $y$ -Achse (Ordinate) rotieren lässt.

Abbildung 2.2.2: Rotationskörper



Die Argumentation zur Flächenberechnung überträgt sich nun direkt auf die Berechnung der Volumina, siehe Abbildung 2.2.3.

Abbildung 2.2.3: Riemannsummen am Rotationskörper



**Satz 2.2.44.** Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig, so gilt

$$V_x(f; a, b) = \pi A(f^2(x); a, b) = \pi \int_a^b f^2(t) dt.$$

**Folgerung 2.2.45.** Ist  $f : [a, b] \rightarrow \mathbb{R}$  injektiv und stetig differenzierbar, so gilt

$$V_y(f; a, b) = \pi \int_a^b t^2 |f'(t)| dt.$$

**Bemerkung 2.2.46.** Satz 2.2.41, Folgerung 2.2.42 und Satz 2.2.44 bleiben richtig, wenn man "stetig" durch "beschränkt mit endlich vielen Unstetigkeitsstellen" ersetzt, siehe Folgerung 2.2.29.

Eine zweidimensionale Verallgemeinerung der Riemannsummen liefert die folgende Verallgemeinerung von Satz 2.2.44:

**Satz 2.2.47.** Es sei  $K \subset \mathbb{R}^3$  ein Körper mit endlichem Volumen  $V(K)$  und  $E_t := \{(x, y, z) \mid z = t\} \subset \mathbb{R}^3$  bezeichne die Ebene durch  $(0, 0, t)$  parallel zur  $xy$ -Ebene.  $K$  werde beschränkt durch die Ebenen  $E_a$  und  $E_b$ . Es sei weiter  $A(t)$  der Flächeninhalt von  $K \cap E_t$  und die Funktion  $A : [a, b] \rightarrow \mathbb{R}$  sei stetig. Dann gilt

$$V(K) = \int_a^b A(t) dt.$$

Damit haben wir auch sofort die folgende wichtige Erkenntnis gewonnen.

**Satz 2.2.48.** *Gegeben seien zwei Körper  $K_1, K_2$ , die bezogen auf eine feste Basisebene  $E$  die gleiche Gesamthöhe haben. Weiter seien zu jeder festen Höhe  $h$  die Flächeninhalte gleich groß, die man als Schnitte von  $K_1 \cap E_h$  und  $K_2 \cap E_h$  mit der Ebene  $E_h$  durch diese Höhe und parallel zu  $E$  erhält. Dann haben die Körper die gleichen Volumina.*

**Folgerung 2.2.49** (Prinzip von Cavallieri, Scherungsprinzip). Ein Spezialfall ergibt sich, wenn die Schnittflächen  $K_1 \cap E_h$  und  $K_2 \cap E_h$  bis auf Translation gleich sind. Dann ergibt sich  $K_1$  aus  $K_2$  durch eine Scherung.

**Bemerkung 2.2.50.** In der Situation von Satz 2.2.47 können wir uns für jedes  $z$  die Projektion der Fläche  $K \cap E_z$  in die  $xy$ -Ebene anschauen. Dort lässt sich diese Fläche im Idealfall so zerlegen, dass die Begrenzung jeder Teilfläche durch Graphen von Funktionen  $y = f_1^z(x)$  und  $y = f_2^z(x)$  beschrieben wird, die ihrerseits von  $z$  abhängen. Wir können dann zur Berechnung der zugehörigen Flächeninhalte jeweils Satz 2.2.41 oder Folgerung 2.2.42 anwenden, wobei die Integrationsgrenzen in der Regel von  $z$  abhängen werden.

Wir betrachten den Spezialfall, dass für kein  $z$  eine Zerlegung notwendig ist und dass sich die Projektion der Schnittflächen  $K \cap E_z$  jeweils als Fläche zwischen der  $x$ -Achse und dem Graphen einer Funktion  $y = f^z(x) =: f(x, z)$  auf dem Intervall  $[c(z), d(z)]$  ergibt. Dann gilt

$$V(K) = \int_a^b \left( \int_{c(z)}^{d(z)} f(x, z) dx \right) dz =: \int_a^b \int_{c(z)}^{d(z)} f(x, z) dx dz .$$

**Beispiel 2.2.51** (Volumen eines Kegelsstumpfes). Wir betrachten den Rotationskörper, der sich bei Rotation des Geradenstücks, das sich als Graph von  $f(z) = mz$  in der  $zx$ -Ebene auf dem Intervall  $[a, b]$ , ergibt. Der Körper ist ein gerader Kegelstumpf der Höhe  $h = b - a$ , großem Radius  $R = mb$  und kleinem Radius  $r = ma$ . Damit ist

$$\begin{aligned} V &= \frac{\pi}{3} h (R^2 + rR + r^2) = \frac{\pi}{3} (b - a) (m^2 b^2 + m^2 ab + m^2 a^2) \\ &= \frac{\pi}{3} m^2 (b^3 - a^3) . \end{aligned}$$

Gemäß Satz 2.2.44 haben wir

$$\begin{aligned} V &= \pi \int_a^b (mt)^2 dt = \pi m^2 \int_a^b t^2 dt = \pi m^2 \cdot \frac{1}{3} t^3 \Big|_a^b \\ &= \frac{\pi}{3} m^2 (b^3 - a^3) . \end{aligned}$$

Gemäß Bemerkung 2.2.50 wählen wir als Basisebene die  $xy$ -Ebene und der Schnitt von  $K$  mit der Ebene  $E_t = \{(x, y, t) \mid x, y \in \mathbb{R}\}$  ist eine Kreisscheibe mit Radius  $r_t = mt$ . Diese Kreisscheibe wird oben und unten jeweils von den Graphen der Funktionen  $y = f_1^t(x) = \sqrt{m^2t^2 - x^2}$  und  $y = f_2^t(x) = -\sqrt{m^2t^2 - x^2}$  auf den Intervallen  $[-mt, mt]$  begrenzt. Damit ergibt sich

$$\begin{aligned} V &= \int_a^b \int_{-mt}^{mt} \left( \sqrt{m^2t^2 - x^2} - (-\sqrt{m^2t^2 - x^2}) \right) dx dt \\ &= 2 \int_a^b \int_{-mt}^{mt} \sqrt{m^2t^2 - x^2} dx dt. \end{aligned}$$

Eine Stammfunktion von  $y \mapsto \sqrt{\rho^2 - y^2}$  ist

$$F(y) = \frac{y}{2} \sqrt{\rho^2 - y^2} + \frac{\rho^2}{2} \arcsin\left(\frac{y}{\rho}\right)$$

wie man durch Ableiten nachprüft. Damit ist dann

$$\begin{aligned} V &= \int_a^b \left[ \frac{1}{2} y \sqrt{m^2t^2 - x^2} + \frac{1}{2} m^2t^2 \arcsin\left(\frac{x}{mt}\right) \Big|_{-mt}^{mt} \right] dt \\ &= \frac{1}{2} \int_a^b \left[ \left(0 + m^2t^2 \arcsin(1)\right) - \left(0 + m^2t^2 \arcsin(-1)\right) \right] dt \\ &= \frac{1}{2} \int_a^b (2m^2\pi t^2) dt = \pi m^2 \cdot \frac{1}{3} t^3 \Big|_a^b \\ &= \frac{\pi}{3} m^2 (b^3 - a^3). \end{aligned}$$

### 2.2.4.2 Wege und Weglänge

Wir verwenden in diesem Abschnitt die für Vektoren im Zahlenraum  $\mathbb{R}^m$  üblichen Schreibweisen. Das heißt für Vektoren  $\vec{x} = (x_1, \dots, x_m)$  und  $y = (y_1, \dots, y_m)$  ist

$$\begin{aligned} \langle \vec{x}, \vec{y} \rangle &= x_1 y_1 + \dots + x_m y_m = \sum_{i=1}^m x_i y_i \\ \|\vec{x}\| &= \sqrt{\langle \vec{x}, \vec{x} \rangle} = \sqrt{x_1^2 + \dots + x_m^2} = \sqrt{\sum_{i=1}^m x_i^2} \end{aligned}$$

**Definition 2.2.52.** Eine *Kurve* oder ein *Weg* im  $\mathbb{R}^m$  ist eine Abbildung  $\vec{c}: [a, b] \rightarrow \mathbb{R}^m$  mit

$$\vec{c}(t) = (c_1(t), \dots, c_m(t)).$$

Wir nennen den Weg  $\vec{c}$  stetig, differenzierbar oder stetig differenzierbar, wenn das für alle Komponentenfunktionen  $c_i: [a, b] \rightarrow \mathbb{R}$  gilt.

**Definition 2.2.53.** Ist  $\vec{c}$  ein Weg dessen Komponenten differenzierbare Funktionen sind, so definieren wir Ableitung des Weges über

$$\vec{c}'(t) := (c_1'(t), \dots, c_m'(t))$$

und die Integration über

$$\int \vec{c}(t) dt := \left( \int c_1(t) dt, \dots, \int c_m(t) dt \right).$$

**Definition/Bemerkung 2.2.54.** Es sei  $Z = \{t_0, \dots, t_k\}$  mit  $a = t_0 < t_1 < \dots < t_{k-1} < t_k = b$  eine Zerlegung des Intervalls  $[a, b]$ . Weiter sei  $\vec{c}: [a, b] \rightarrow \mathbb{R}^m$  ein Weg.

- Mit  $L(\vec{c}, Z)$  bezeichnen wir die Länge des Streckenzuges

$$\overline{\vec{c}(t_0)\vec{c}(t_1)} \cup \overline{\vec{c}(t_1)\vec{c}(t_2)} \cup \dots \cup \overline{\vec{c}(t_{k-1})\vec{c}(t_k)},$$

Ist der Weg  $\vec{c}$  differenzierbar, so gilt

$$L(\vec{c}, Z) = \sum_{j=1}^k \|\vec{c}(t_j) - \vec{c}(t_{j-1})\| = \sum_{j=1}^k \left\| \int_{t_{j-1}}^{t_j} \vec{c}'(t) dt \right\|.$$

- Als Länge des Weges bezeichnen wir die Größe

$$L(\vec{c}) := \sup\{L(\vec{c}, Z) \mid Z \text{ ist eine Zerlegung von } [a, b]\}$$

falls sie existiert.

- Es sei  $\vec{c}: [a, b] \rightarrow \mathbb{R}^m$  mit  $L(\vec{c}) < \infty$ . Dann heißt  $L_{\vec{c}}: [a, b] \rightarrow \mathbb{R}$  mit

$$L_{\vec{c}}(t) = L(\vec{c}|_{[a,t]})$$

die Längenfunktion des Weges  $\vec{c}$ .

**Bemerkung 2.2.55.** Für jeden stetigen Weg  $\vec{v}: [a, b] \rightarrow \mathbb{R}^m$  gilt

$$\left\| \int_a^b \vec{v}(t) dt \right\| \leq \int_a^b \|\vec{v}(t)\| dt.$$

Insbesondere ist für einen stetig differenzierbaren Weg  $\vec{c}: [a, b] \rightarrow \mathbb{R}^m$

$$L(\vec{c}) \leq \int_a^b \|\vec{c}'(t)\| dt.$$

**Satz 2.2.56.** Ist der Weg  $\vec{c}: [a, b] \rightarrow \mathbb{R}^m$  stetig differenzierbar, dann gilt:

1. Die Längenfunktion von  $\vec{c}$  ist differenzierbar mit  $L'_{\vec{c}}(t) = \|\vec{c}'(t)\|$ .
2. Für die Länge gilt  $L(\vec{c}) = L_{\vec{c}}(b) = \int_a^b \|\vec{c}'(t)\| dt$ .

## 2.3 Abbildungen mehrerer Variablen, Stetigkeit und topologische Grundbegriffe

Wir hatten die Stetigkeit von reellen Funktionen mit Hilfe von Intervallen definiert, siehe Definition 1.4.1. Später hatten wir dann gezeigt, dass sich die Stetigkeit auch mit Hilfe von Folgen charakterisieren ließ, siehe Satz 1.4.30. Wir wollen den ersten Ansatz nun auf Funktionen in mehreren Variablen verallgemeinern. Dabei werden wir einige Bezeichnungen im Zahlenraum  $\mathbb{R}^n$  verwenden, die wir im Folgenden wiederholen bzw. einführen.

### 2.3.1 Grundbegriffe des Zahlenraums

Der Zahlenraum  $\mathbb{R}^n$  besteht aus  $n$ -tupeln reeller Zahlen, die *Punkte*. Wir schreiben  $x = (x_1, \dots, x_n)$  für  $x \in \mathbb{R}^n$ .<sup>(i)</sup> Identifizieren wir die Punkte des

---

<sup>(i)</sup>Aus Gründen der Platzersparnis, werden wir in den kommenden zwei Kapiteln die Elemente des Zahlenraums als Zeilenvektoren schreiben.

$\mathbb{R}^n$  mit Ihren Aufpunktvektoren, so ist der  $\mathbb{R}^n$  zusammen mit der komponentenweisen Addition und skalaren Multiplikation ein reeller Vektorraum: Für  $a \in \mathbb{R}$  und  $x, y \in \mathbb{R}^n$  ist

$$\begin{aligned}x + y &= (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n), \\ax &= a(x_1, \dots, x_n) = (ax_1, \dots, ax_n).\end{aligned}$$

Im  $\mathbb{R}^n$  haben wir den natürlichen Abstandsbegriff, der durch die geometrische Länge des Abstandsvektors gegeben ist. Die Länge oder Norm von  $x \in \mathbb{R}^n$  ist

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

und damit der Abstand zweier Punkte

$$d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}.$$

Damit haben wir natürlich Begriffe des  $n$ -dimensionalen offenen und abgeschlossenen Balls und der  $(n - 1)$ -dimensionalen Sphäre mit Radius  $r$  um einen Punkt  $x^0 \in \mathbb{R}^n$ . Diese Bälle und die Sphäre bezeichnen wir mit<sup>(i)</sup>

$$\begin{aligned}B_r(x^0) &= \{x \in \mathbb{R}^n \mid \|x - x^0\| < r\}, \\ \bar{B}_r(x^0) &= \{x \in \mathbb{R}^n \mid \|x - x^0\| \leq r\}, \\ S_r(x^0) &= \{x \in \mathbb{R}^n \mid \|x - x^0\| = r\},\end{aligned}$$

also  $\bar{B}_r(x^0) = B_r(x^0) \cup S_r(x^0)$ .

### 2.3.2 Abbildungen in mehreren Variablen und Stetigkeit

**Definition 2.3.1.** • Unter einer *Funktion (in mehreren Variablen)* verstehen wir eine Abbildung die auf einer Teilmenge  $M \subset \mathbb{R}^n$  definiert ist, also

$$f : M \rightarrow \mathbb{R}.$$

Wir schreiben auch  $f(x) = f(x_1, \dots, x_n)$ .

---

<sup>(i)</sup>Wenn wir die Dimension des umgebenden Raums eines Balls oder einer Sphäre betonen wollen, werden wir das als Index ergänzen:  $B_r^n(x^0), \bar{B}_r^n(x^0), S_r^{n-1}(x^0) \subset \mathbb{R}^n$ .

- Als den *Graphen* der Funktion  $f$  bezeichnen wir die Menge

$$G_f = \{(x, z) \mid x \in M, z = f(x)\} \subset \mathbb{R}^{n+1}.$$

- Als *Höhenmengen* der Funktion  $f$  bezeichnen wir die Mengen

$$H_{f,z} = \{x \in M \mid f(x) = z\} \subset \mathbb{R}^n.$$

Im Fall  $n = 2$  spricht man auch von *Höhenlinien* und im Fall  $n = 3$  von *Höhenflächen*.

**Beispiel 2.3.2.** 1. Die Abbildung, die  $x \in \mathbb{R}^n$  die  $i$ -te Koordinate  $x_i$  zuordnet

$$\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \pi_i(x) = x_i$$

heißt  *$i$ -te Koordinatenfunktion*. Wenn wir den Definitionsbereich betonen wollen, dann schreiben wir auch  $\pi_i^n$ .

2. Eine Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  der Form

$$p(x) = \sum_{i_1=0}^{m_1} \sum_{i_2=0}^{m_2} \cdots \sum_{i_n=0}^{m_n} a_{i_1 i_2 \dots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$$

mit  $a_{i_1 i_2 \dots i_n} \in \mathbb{R}$  und  $m_j \in \mathbb{N}$  heißt *Polynom* (in  $n$  Variablen).

Die Zahl  $m := \max\{j_1 + j_2 + \dots + j_n \mid a_{j_1 j_2 \dots j_n} \neq 0\}$  heißt der *Grad* des Polynoms.

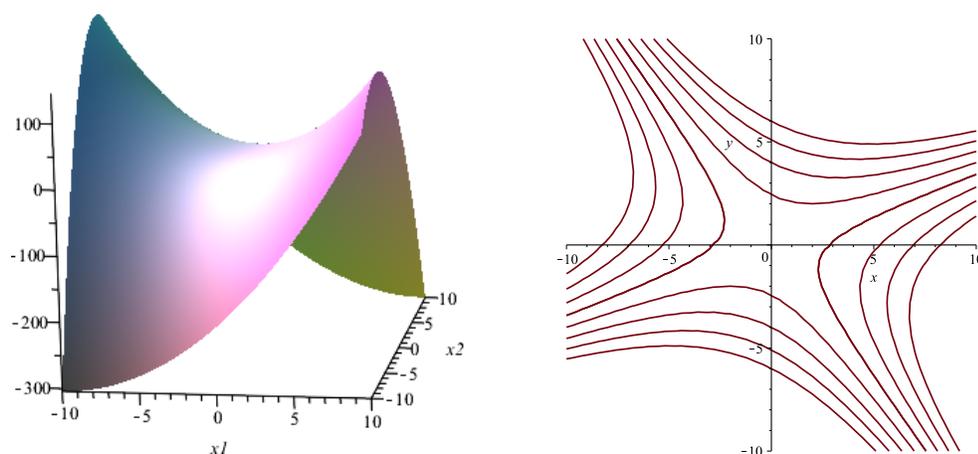
3. Beispiel eines Polynoms:

$$f(x_1, x_2) = x_1^2 - 2x_2^2 - 2x_1x_2 - 4 = (x_1 - x_2)^2 - 3x_2^2 - 4.$$

Es ist hier  $a_{20} = 1$ ,  $a_{02} = -2$ ,  $a_{11} = -2$  und  $a_{00} = -4$  und der Grad von  $f$  ist 2. Die Höhenlinien von  $f$  sind spezielle Quadriken im  $\mathbb{R}^2$  und der Graph von  $f$  hat die Form einer Sattelfläche, siehe Abbildung 2.3.1.

4. Die Menge aller Polynome in  $n$  Variablen ist eine unendlichdimensionale reelle Algebra, also insbesondere ein unendlichdimensionaler reeller Vektorraum. Wir bezeichnen ihn mit  $\mathbb{R}[x_1, x_2, \dots, x_n]$ .

Abbildung 2.3.1: Graph und Höhenlinien der Funktion  $f(x_1, x_2) = x_1^2 - 2x_2^2 - 2x_1x_2 - 4$



5. Die Menge aller Polynome in  $n$  Variablen mit maximalem Grad  $m$  ist ein Untervektorraum  $\mathbb{R}_m[x_1, \dots, x_n] \subset \mathbb{R}[x_1, \dots, x_n]$ . Er ist endlichdimensional mit Dimension  $\binom{m+n}{m}$ .
6. Rationale Funktionen in  $n$  Variablen sind Funktionen der Form  $\frac{p}{q} : M \setminus \{x \in \mathbb{R}^n \mid q(x) = 0\} \rightarrow \mathbb{R}$  mit  $p, q \in \mathbb{R}[x_1, \dots, x_n]$ .

**Definition 2.3.3.** Eine *Abbildung (in mehreren Variablen)*<sup>(i)</sup> ist eine Abbildung, die auf einer Teilmenge  $M \subset \mathbb{R}^n$  definiert ist, und ihre Werte in  $\mathbb{R}^k$  annimmt, also

$$f : M \rightarrow \mathbb{R}^k.$$

Schreiben wir mit Hilfe der Projektionen  $\pi_i : \mathbb{R}^k \rightarrow \mathbb{R}$  von  $\mathbb{R}^k$

$$f_i : M \rightarrow \mathbb{R}, \quad f_i(x) := \pi_i(f(x)),$$

so lässt sich die Abbildung  $f$  schreiben als

$$f(x) = (f_1(x), f_2(x), \dots, f_k(x)).$$

<sup>(i)</sup>Wir verwenden im folgenden diesen strengeren Begriff der "Abbildung" und sprechen von "Funktionen", wenn wir betonen wollen, dass der Wertebereich  $\mathbb{R}$  ist. Im Zusammenhang mit dem allgemeinen Begriff der Abbildung sollte es nicht zu Verwirrung kommen.

Die  $f_i$  heißen dann die *Komponentenfunktionen von  $f$* . In dieser Form werden Abbildungen in der Regel vorgegeben sein.

**Beispiel 2.3.4.** 1. Die Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad f(x, y) = (x^2 - y^2, 2xy)$$

hat das folgende Verhalten:

- Kreise vom Radius  $r$  um den Ursprung werden auf Kreise mit Radius  $r^2$  um den Ursprung abgebildet.  
Sei nämlich  $(x, y)$  ein Punkt mit  $\|(x, y)\|^2 = x^2 + y^2 = r^2$ , dann ist  $x = r \cos(\alpha)$  und  $y = r \sin(\alpha)$  für einen Winkel  $\alpha$ . Damit ist

$$\begin{aligned} \|f(x, y)\|^2 &= \|(r^2 \cos^2(\alpha) - r^2 \sin^2(\alpha), 2r^2 \cos(\alpha) \sin(\alpha))\|^2 \\ &= \|(r^2 \cos(2\alpha), r^2 \sin(2\alpha))\|^2 \\ &= r^4 \sin^2(2\alpha) + r^4 \cos^2(2\alpha) \\ &= r^4. \end{aligned}$$

- Halbgeraden vom Ursprung, die mit der  $x$ -Achse einen Winkel  $\phi$  einschließen, werden auf Halbgeraden vom Ursprung abgebildet, die mit der  $x$ -Achse einen Winkel  $2\phi$  einschließen.  
Sei nämlich  $y = \tan(\phi)x$ , dann ist  $f_1(x, y) = x^2 - x^2 \tan^2(\phi)$  und  $f_2(x, y) = 2x^2 \tan(\phi)$ , sodass

$$\frac{f_2(x, y)}{f_1(x, y)} = \frac{2 \tan(\phi)}{1 - \tan^2(\phi)} = \tan(2\phi),$$

also  $f_2(x, y) = \tan(2\phi) f_1(x, y)$ .

Identifiziert man  $\mathbb{R}^2$  mit der komplexen Ebene gemäß  $(x, y) \leftrightarrow z = x + iy$ , so entspricht die Funktion  $f$  dem komplexen Quadrieren  $z \mapsto z^2$ .

2. Auch im Fall allgemeinerer Abbildungen kann man in Ausnahmefällen den Verlauf noch graphisch darstellen. Ist etwa  $M \subset \mathbb{R}^3$  und  $f : M \rightarrow \mathbb{R}^2$ , so heißt für  $z \in \mathbb{R}^2$  die Menge  $H_{f,z} = \{x \in M \mid f(x) = z\} \subset \mathbb{R}^3$  *Höhenlinie*.

Für  $f(x_1, x_2, x_3) = (x_1^2 + x_2^2 - x_3^2, x_3)$  ist  $H_{f,(\rho^2, h)}$  ein Kreis in der Ebene  $\{(x_1, x_2, h) \mid x_1, x_2 \in \mathbb{R}\}$  mit Radius  $\sqrt{\rho^2 + h^2}$  und Mittelpunkt  $(0, 0, h)$ .

Die Definition der Stetigkeit geschieht analog zur Stetigkeit in Definition 1.4.1.

**Definition 2.3.5.** Eine Abbildung  $f : M \rightarrow \mathbb{R}^k$  mit  $M \subset \mathbb{R}^n$  heißt stetig in  $x^0 \in M$ , wenn

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in B_\delta(x^0) \cap M : f(x) \in B_\epsilon(f(x^0)).$$

Sie heißt *stetig auf  $M$* , wenn Sie stetig in jedem  $x \in M$  ist.

Die Definition der Stetigkeit hat folgende Umformulierungen:

**Bemerkung 2.3.6.** Eine Abbildung  $f : M \rightarrow \mathbb{R}^k$  ist stetig in  $x^0 \in M \subset \mathbb{R}^n$  genau dann, wenn

- $\forall \epsilon > 0 \exists \delta > 0 \forall x \in M : (|x - x^0| < \delta \implies \|f(x) - f(x^0)\| < \epsilon)$
- $\forall B_\epsilon(f(x^0)) \exists B_\delta(x^0) : f(M \cap B_\delta(x^0)) \subset B_\epsilon(f(x^0))$

**Satz 2.3.7.** Die Funktion  $f : M \rightarrow \mathbb{R}^k$  mit  $M \subset \mathbb{R}^n$  ist genau dann stetig, wenn sämtliche Komponentenfunktionen  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig sind.

Wie bei den reellen Funktionen gilt auch bei den stetigen Abbildungen die folgende Bemerkung:

**Bemerkung 2.3.8.** Es sei  $f : M \rightarrow \mathbb{R}^k$  stetig in  $x^0 \in M \subset \mathbb{R}^n$ , dann gilt:

- Es gibt einen Ball  $B_\delta(x^0)$ , sodass  $f|_{M \cap B_\delta(x^0)}$  beschränkt ist.
- Ist  $f(x^0) > 0$ , so gibt es einen Ball  $B_\delta(x^0)$ , sodass  $f|_{M \cap B_\delta(x^0)} > 0$ .

Auch der Beweis des folgenden Satzes verläuft analog zum Beweis im Fall einer Variablen, siehe Sätze 1.4.3 und 1.4.4.

**Satz 2.3.9.** 1. Es seien  $f : M \rightarrow \mathbb{R}^k$  und  $\hat{f} : M \rightarrow \mathbb{R}^k$  Abbildungen auf  $M \subset \mathbb{R}^n$ . Weiter seien  $f$  und  $\hat{f}$  beide stetig in  $x^0 \in M$ . Dann sind auch die Summe  $f + \hat{f} : M \rightarrow \mathbb{R}^k$  und die Differenz  $f - \hat{f} : M \rightarrow \mathbb{R}^k$  stetig in  $x^0 \in M$ .

2. Es sei  $g : M \rightarrow \mathbb{R}$  eine Funktion auf  $M \subset \mathbb{R}^n$ . Weiter sei  $g$  stetig in  $x^0 \in M$  und  $g(x^0) \neq 0$ . Dann ist auch  $\frac{1}{g} : M \rightarrow \mathbb{R}$  mit  $\frac{1}{g}(x) = \frac{1}{g(x)}$  stetig in  $x_0$ .

3. Es sei  $f : M \rightarrow \mathbb{R}^k$  eine Abbildung und  $g : M \rightarrow \mathbb{R}$  eine Funktion auf  $M \subset \mathbb{R}^n$ . Weiter seien  $f$  und  $g$  beide stetig in  $x^0 \in M$ . Dann ist auch das Produkt  $gf : M \rightarrow \mathbb{R}^k$  stetig in  $x^0 \in M$ .
4. Es seien  $M \subset \mathbb{R}^n$  und  $N \subset \mathbb{R}^k$  und  $f : M \rightarrow \mathbb{R}^k$ ,  $h : N \rightarrow \mathbb{R}^m$  Abbildungen mit  $f(M) \subset N$ . Ist  $f$  stetig in  $x^0 \in M$  und  $h$  stetig in  $f(x^0) \in N$ , so ist auch die Verkettung  $h \circ f : M \rightarrow \mathbb{R}^m$  mit  $(h \circ f)(x) = h(f(x))$  stetig in  $x^0$ .

**Beispiel 2.3.10.** 1. Die konstanten Abbildungen  $\underline{c} : \mathbb{R}^n \rightarrow \mathbb{R}^k$  mit  $\underline{c}(x) := c$  für ein festes  $c \in \mathbb{R}^k$  sind stetig.

2. Die Projektionen  $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$  sind stetig.
3. Die Multiplikation und die Addition auf  $\mathbb{R}$  können als Funktionen  $p : \mathbb{R}^2 \rightarrow \mathbb{R}$  und  $m : \mathbb{R}^2 \rightarrow \mathbb{R}$  interpretiert werden:

$$p(x_1, x_2) = x_1 + x_2, \quad m(x_1, x_2) = x_1 x_2$$

Diese sind stetig. Bemerkung: Man kann nun die Rechenregeln 2. und 3. aus dem vorigen Satz neu beweisen, indem man dieses Beispiel und Teil 1. aus dem Satz anwendet.

4. Polynome in mehreren Variablen sind stetig.
5. Rationale Funktionen in mehreren Variablen sind stetig.
6. Es sei  $A \in \mathbb{R}^{k \times \ell}$  eine Matrix und  $f : M \rightarrow \mathbb{R}^k$  und  $g : M \rightarrow \mathbb{R}^\ell$  seien stetig auf  $M \subset \mathbb{R}^n$ . Dann ist auch die Funktion  $F_A : M \rightarrow \mathbb{R}$  mit

$$F_A(x) = f(x)^T \cdot A \cdot g(x)$$

stetig auf  $M$ .

Ein Spezialfall erhält man für  $\ell = k$  und  $A = \mathbb{1}$ . Es ist dann

$$F_{\mathbb{1}}(x) = \langle f(x), g(x) \rangle = \sum_{i=1}^k f_i(x) g_i(x).$$

**Beispiel 2.3.11.** 1. (**Polarkoordinaten**). Die Abbildung

$$\Phi_P : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \Phi_P(r, \phi) = \begin{pmatrix} r \cos(\phi) \\ r \sin(\phi) \end{pmatrix}$$

ist stetig und surjektiv, aber nicht injektiv. Die Einschränkung auf die Teilmenge

$$\mathbb{R}^{>0} \times ]0, 2\pi[ \subset \mathbb{R}^2$$

ist injektiv, aber nicht mehr surjektiv. Schränkt man hier den Wertebereich auf das Bild

$$\text{Bild}\left(\Phi_P|_{\mathbb{R}^{>0} \times ]0, 2\pi[}\right) = \mathbb{R}^2 \setminus \{(x, 0) \mid x \geq 0\}$$

ein, so ergibt sich die Umkehrfunktion zu

$$\begin{aligned} & \left(\Phi_P|_{\mathbb{R}^{>0} \times ]0, 2\pi[}\right)^{-1} : \mathbb{R}^2 \setminus \{(x, 0) \mid x \geq 0\} \rightarrow \mathbb{R}^2 \\ & \left(\Phi_P|_{\mathbb{R}^{>0} \times ]0, 2\pi[}\right)^{-1}(x, y) = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arg(x, y) \end{pmatrix} \end{aligned}$$

mit

$$\arg(x, y) := \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{falls } x > 0, y > 0 \\ \frac{\pi}{2} & \text{falls } x = 0, y > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{falls } x < 0 \\ \frac{3\pi}{2} & \text{falls } x = 0, y < 0 \\ \arctan\left(\frac{y}{x}\right) + 2\pi & \text{falls } x > 0, y < 0 \end{cases}$$

2. (**Zylinderkoordinaten**). Die Abbildung

$$\Phi_Z : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \Phi_Z(r, \varphi, z) = \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \\ z \end{pmatrix}$$

ist stetig und surjektiv, aber nicht injektiv. Die Einschränkung auf die Teilmenge

$$\mathbb{R}^{>0} \times ]0, 2\pi[ \times \mathbb{R} \subset \mathbb{R}^3$$

ist injektiv, aber nicht mehr surjektiv. Schränkt man den Wertebereich auf das Bild

$$\text{Bild}\left(\Phi_Z|_{\mathbb{R}^{>0} \times ]0, 2\pi[ \times \mathbb{R}}\right) = \mathbb{R}^3 \setminus \{(x, 0, z) \mid x \geq 0, z \in \mathbb{R}\} .$$

ein, so ergibt sich die Umkehrfunktion

$$\begin{aligned} & \left( \Phi_Z \Big|_{\mathbb{R}^{>0} \times ]0, 2\pi[ \times \mathbb{R}} \right)^{-1} : \mathbb{R}^3 \setminus \{(x, 0, z) \mid x \geq 0, z \in \mathbb{R}\} \rightarrow \mathbb{R}^3 \\ & \left( \Phi_Z \Big|_{\mathbb{R}^{>0} \times ]0, 2\pi[ \times \mathbb{R}} \right)^{-1} (x, y, z) = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arg(x, y) \\ z \end{pmatrix} \end{aligned}$$

3. (**Kugelkoordinaten**). Die Abbildung

$$\Phi_K : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \Phi_K(r, \varphi, \vartheta) = \begin{pmatrix} r \cos(\varphi) \cos(\vartheta) \\ r \sin(\varphi) \cos(\vartheta) \\ r \sin(\vartheta) \end{pmatrix}$$

ist stetig und surjektiv, aber nicht injektiv. Die Einschränkung auf die Teilmenge

$$\mathbb{R}^{>0} \times ]0, 2\pi[ \times ]-\frac{\pi}{2}, \frac{\pi}{2}[ \subset \mathbb{R}^3$$

ist injektiv, aber nicht mehr surjektiv. Schränkt man hier den Wertebereich auf das Bild

$$\text{Bild} \left( \Phi_K \Big|_{\mathbb{R}^{>0} \times ]0, 2\pi[ \times ]-\frac{\pi}{2}, \frac{\pi}{2}[} \right) = \mathbb{R}^3 \setminus \{(x, 0, z) \mid x \geq 0, z \in \mathbb{R}\} .$$

ein, so ergibt sich die Umkehrfunktion zu

$$\begin{aligned} & \left( \Phi_K \Big|_{\mathbb{R}^{>0} \times ]-\frac{\pi}{2}, \frac{\pi}{2}[ \times ]0, 2\pi[} \right)^{-1} : \mathbb{R}^3 \setminus \{(x, 0, z) \mid x \geq 0, z \in \mathbb{R}\} \rightarrow \mathbb{R}^3 \\ & \left( \Phi_K \Big|_{\mathbb{R}^{>0} \times ]-\frac{\pi}{2}, \frac{\pi}{2}[ \times ]0, 2\pi[} \right)^{-1} (x, y, z) = \begin{pmatrix} \sqrt{x^2 + y^2 + z^2} \\ \arg(x, y) \\ \arcsin \left( \frac{x}{\sqrt{x^2 + y^2 + z^2}} \right) \end{pmatrix} \end{aligned}$$

### 2.3.3 Folgen und Grenzwerte im Zahlenraum

Mit Hilfe der Definition des Balls, bzw. des Abstandsbegriffs  $(x, y) \mapsto \|x - y\|$ , können wir auch Begriffe, die wir im Zusammenhang mit Zahlenfolgen erhalten haben, auf den Fall von Folgen im Zahlenraum  $\mathbb{R}^n$  übertragen.

**Definition 2.3.12.** • Eine Folge im Zahlenraum ist eine Abbildung von  $\mathbb{N}$  nach  $\mathbb{R}^n$  und wir schreiben dafür  $(x^{(n)})_{n \in \mathbb{N}}$ .

- Eine Folge  $(x^{(n)})_{n \in \mathbb{N}}$  heißt konvergent gegen  $x^0 \in \mathbb{R}^n$ , wenn

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n > n_0 : x^{(n)} \in B_\epsilon(x^0).$$

**Satz 2.3.13.** 1. Eine Folge  $(x^{(k)})_{k \in \mathbb{N}}$  in  $\mathbb{R}^n$  konvergiert genau dann, wenn alle ihre Komponentenfolgen  $(x_i^{(k)})_{k \in \mathbb{N}}$  für  $i = 1, \dots, n$  konvergieren.

2. (**Cauchy-Kriterium**). Eine Folge  $(x^{(k)})_{k \in \mathbb{N}}$  in  $\mathbb{R}^n$  konvergiert genau dann, wenn sie ein Cauchy-Folge ist, d.h.

$$\forall \epsilon > 0 \exists k_0 \in \mathbb{N} \forall k, \ell > k_0 : x^{(k)} - x^{(\ell)} \in B_\epsilon(0).$$

3. Eine Folge  $(x^{(k)})_{k \in \mathbb{N}}$  in  $\mathbb{R}^n$  ist genau dann eine Cauchy-Folge, wenn das für ihre Komponentenfolgen gilt.

Auch die Äquivalenz zwischen Folgenstetigkeit und Stetigkeit überträgt sich wortwörtlich aus dem Fall einer Variablen.

**Satz 2.3.14.** Eine Funktion  $f : M \rightarrow \mathbb{R}^k$  ist genau dann stetig in  $x^0 \in M \subset \mathbb{R}^n$ , wenn für alle Folgen  $(x^{(k)})_{k \in \mathbb{N}}$  in  $M$  die gegen  $x^0$  konvergieren die Folge  $(f(x^{(k)}))_{k \in \mathbb{N}}$  gegen  $f(x^0)$  konvergiert.

Wir übertragen nun noch einen weiteren Begriff aus dem Fall einer reellen Variablen.

**Definition 2.3.15.** Ein Punkt  $x^0 \in \mathbb{R}^n$  heißt *Häufungspunkt* der Menge  $M \subset \mathbb{R}^n$ , wenn in jedem Ball  $B_\epsilon(x^0)$  um  $x^0$  noch ein Punkt von  $M \setminus \{x^0\}$  liegt, d.h.

$$\forall \epsilon > 0 \exists x \in M \setminus \{x^0\} : x \in B_\epsilon(x^0).$$

Alle Punkte von  $M$ , die keine Häufungspunkte sind, heißen *isolierte Punkte*.

Auch die Definition des Grenzwertes und die Folgerungen sind analog zu den Aussagen in Kapitel 1.4.3.

**Definition 2.3.16.** Es sei  $M \subset \mathbb{R}^n$ ,  $x^0 \in \mathbb{R}^n$  ein Häufungspunkt von  $M$ . Ein Wert  $y^0 \in \mathbb{R}^k$  heißt Grenzwert der Funktion  $f : M \rightarrow \mathbb{R}^k$ , wenn

$$\forall B_\epsilon(y^0) \exists B_\delta(x^0) \forall x \in (M \setminus \{x^0\}) \cap B_\delta(x^0) : f(x) \in B_\epsilon(y^0).$$

Wir schreiben dann  $\lim_{x \rightarrow x^0} f(x) = y^0$ .

- Bemerkung 2.3.17.** 1. Es gibt in jedem Häufungspunkt  $x^0$  von  $M$  höchstens einen Grenzwert.
2. Die Abbildung  $f : M \rightarrow \mathbb{R}^k$  besitzt genau dann im Häufungspunkt  $x^0$  von  $M \subset \mathbb{R}^n$  den Grenzwert  $y^0$ , wenn die Abbildung

$$\hat{f} : M \cup \{x^0\} \rightarrow \mathbb{R}^k \quad \text{mit} \quad \hat{f}(x) = \begin{cases} f(x) & \text{für } x \in M \setminus \{x^0\} \\ y^0 & \text{für } x = x^0 \end{cases}$$

stetig in  $x^0$  ist.

3.  $f : M \rightarrow \mathbb{R}^k$  ist in dem Häufungspunkt  $x^0 \in M \subset \mathbb{R}^n$  genau dann stetig, wenn  $\lim_{x \rightarrow x^0} f(x) = f(x^0)$ .
4.  $f : M \rightarrow \mathbb{R}^k$  besitzt im Häufungspunkt  $x^0$  von  $M \subset \mathbb{R}^n$  genau dann einen Grenzwert, wenn dies für alle ihre Komponentenfunktionen gilt. In diesem Fall gilt  $\lim_{x \rightarrow x^0} f(x) = \left( \lim_{x \rightarrow x^0} f_1(x), \dots, \lim_{x \rightarrow x^0} f_k(x) \right)$ .

Auch die Rechenregeln gelten analog:

- Bemerkung 2.3.18.** 1.  $f : M \rightarrow \mathbb{R}$  habe in dem Häufungspunkt  $x^0$  von  $M \subset \mathbb{R}^n$  den Grenzwert  $y^0 = \lim_{x \rightarrow x^0} f(x) \in N \subset \mathbb{R}^k$ . Weiter sei  $g : N \rightarrow \mathbb{R}^\ell$  stetig in  $y^0$ . Dann hat  $g \circ f : M \rightarrow \mathbb{R}^\ell$  einen Grenzwert in  $x^0$  und es gilt

$$\lim_{x \rightarrow x^0} g(f(x)) = g\left(\lim_{x \rightarrow x^0} f(x)\right).$$

2. Haben die Abbildungen  $f_1, f_2 : M \rightarrow \mathbb{R}^k$ ,  $g : M \rightarrow \mathbb{R}$  im Häufungspunkt  $x^0$  von  $M \subset \mathbb{R}^n$  einen Grenzwert, so auch die Abbildungen  $f_1 \pm f_2$ ,  $gf_1$  und  $\frac{1}{g}$  und es gilt<sup>(i)</sup>

$$\begin{aligned} \lim_{x \rightarrow x^0} (f_1(x) \pm f_2(x)) &= \lim_{x \rightarrow x^0} f_1(x) \pm \lim_{x \rightarrow x^0} f_2(x), \\ \lim_{x \rightarrow x^0} g(x)f_1(x) &= \lim_{x \rightarrow x^0} g(x) \lim_{x \rightarrow x^0} f_1(x), \quad \lim_{x \rightarrow x^0} \frac{1}{g(x)} = \frac{1}{\lim_{x \rightarrow x^0} g(x)}, \end{aligned}$$

---

<sup>(i)</sup>Im letzten Fall muss natürlich  $\lim_{x \rightarrow x^0} g(x) \neq 0$  vorausgesetzt werden.

## 2.3.4 Topologische Grundbegriffe des Zahlenraums

**Definition 2.3.19.** Es sei  $M \subset \mathbb{R}^n$ . Ein Punkt  $x \in \mathbb{R}^n$  heißt

1. *innerer Punkt von  $M$* , wenn

$$\exists \epsilon > 0 : B_\epsilon(x) \subset M;$$

d.h.: Es gibt einen Ball um  $x$ , der ganz in  $M$  liegt. Insbesondere gehört jeder innerer Punkt von  $M$  zu  $M$ .

2. *Randpunkt von  $M$*

$$\forall \epsilon > 0 : B_\epsilon(x) \cap M \neq \emptyset \text{ und } B_\epsilon(x) \cap M^c \neq \emptyset;$$

das heißt, jeder Ball um  $x$  enthält sowohl Punkte aus  $M$  als auch Punkte aus  $M^c = \mathbb{R}^n \setminus M$ . Ein Randpunkt von  $M$  kann – aber muss nicht – zu  $M$  gehören.

**Definition 2.3.20.** Es sei  $M \subset \mathbb{R}^n$ .

1. Die Menge  $\overset{\circ}{M} \subset M$  der inneren Punkte von  $M$  heißt das *Innere von  $M$* .
2. Die Menge  $\partial M \subset \mathbb{R}^n$  der Randpunkte von  $M$  heißt der *Rand* von  $M$ .
3. Die Menge  $\bar{M} := M \cup \partial M$  heißt *Abschluss von  $M$* .
4. Eine Menge  $M$  heißt *offen*, wenn  $M = \overset{\circ}{M}$ , d.h. sie besteht nur aus inneren Punkte, bzw. keiner ihrer Randpunkte gehört zu ihr.
5. Eine Menge  $M$  heißt *abgeschlossen*, wenn  $M = \bar{M}$ , d.h. alle Randpunkte gehören zu  $M$ .

Statt zu sagen "sei  $U \subset \mathbb{R}^n$  eine offene Menge" schreiben wir kurz "sei  $U \overset{\circ}{\subset} \mathbb{R}^n$ ".

**Beispiel 2.3.21.** 1. Die leere Menge und  $\mathbb{R}^n$  sind offen.

2.  $B_r(x)$  ist offen,  $\bar{B}_r(x)$  ist abgeschlossen und  $\partial B_r(x) = S_r(x)$ . Deswegen sprechen wir auch von offenen und abgeschlossenen Bällen.

3. Für  $n = 1$  sind die offenen und abgeschlossenen Bälle gerade die offenen und abgeschlossenen Intervalle. Der Rand eines Intervalls besteht aus den beiden Randpunkten, sodass die vorab eingeführten Begriffe sinnvoll sind.
4. Eine Menge muss weder offen noch abgeschlossen sein.

**Bemerkung 2.3.22.** 1.  $\partial M = \partial(M^c)$ ,  $(\partial M)^c = \overset{\circ}{M} \cup (\bar{M})^c$

2.  $M$  ist genau dann offen, wenn  $M^c$  abgeschlossen ist.  $M$  ist genau dann abgeschlossen, wenn  $M^c$  offen ist.
3. Eine Menge  $M$  ist genau dann offen, wenn es für jeden Punkt  $x \in M$  einen offenen Ball<sup>(i)</sup>  $B \subset M$  gibt, mit  $x \in B$ .
4. Die Vereinigung von offenen Bällen ist offen. Der Schnitt zweier offener Bälle ist offen.
5. Die offenen Mengen sind genau die Vereinigung offener Bällen.

Im Zusammenhang mit Häufungspunkten gilt das Folgende:

- Bemerkung 2.3.23.** 1. Jeder innere Punkt von  $M$  ist ein Häufungspunkt von  $M$ .
2. Jeder Häufungspunkt von  $M$ , der nicht zu  $M$  gehört, ist ein Randpunkt von  $M$ .
  3. Eine Menge  $M$  ist genau dann abgeschlossen, wenn Sie alle ihre Randpunkte enthält.

Mit Hilfe von Bemerkung 2.3.22 beweist man leicht den folgenden Satz.

**Satz 2.3.24.** *Die Vereinigung beliebig vieler offener Mengen und der Schnitt endlich vieler offener Mengen ist offen. Das heißt, ist  $\mathcal{U} \subset \mathcal{P}(\mathbb{R}^n)$  eine nicht-leere Menge von offenen Teilmengen von  $\mathbb{R}^n$ , so gilt:*

1.  $\bigcup \mathcal{U} \subset \mathbb{R}^n$  ist offen.

---

<sup>(i)</sup>Die Bezeichnung  $B_r$  oder  $B$  für einen Ball benutzen wir, wenn der Mittelpunkt oder auch der Radius nicht spezifiziert wird.

2. Ist  $\mathcal{U}$  eine endliche Menge, so ist  $\bigcap \mathcal{U} \subset \mathbb{R}^n$  offen.

Durch den Übergang zum Komplement erhält man die folgende Aussage für abgeschlossene Mengen.

**Bemerkung 2.3.25.** 1. Endliche Vereinigungen und beliebige Schnitte abgeschlossener Mengen sind abgeschlossen.

2. Ist  $U \overset{\circ}{\subset} \mathbb{R}^n$  offen und  $A \subset \mathbb{R}^n$  abgeschlossen, so ist  $U \setminus A$  offen und  $A \setminus U$  abgeschlossen.

3. Unendliche Schnitte offener Mengen müssen nicht offen sein, und unendliche Vereinigungen abgeschlossener Mengen müssen nicht abgeschlossen sein.

- Sei  $(r_n)_{n \in \mathbb{N}}$  eine monoton fallende, reelle Nullfolge und betrachte  $\mathcal{U} = \{B_{r_n}(x^0) \mid n \in \mathbb{N}\}$ . Dann ist

$$\bigcap \mathcal{U} = \bigcap_{n \in \mathbb{N}} B_{r_n}(x^0) = \{x^0\}$$

abgeschlossen.

- Sei  $(r_n)_{n \in \mathbb{N}}$  mit  $0 < r_n < 1$  eine monoton steigende Zahlenfolge mit Grenzwert 1. Dann ist für  $\mathcal{V} = \{\bar{B}_{r_n}(x^0) \mid n \in \mathbb{N}\}$

$$\bigcup \mathcal{V} = \bigcup_{n \in \mathbb{N}} \bar{B}_{r_n}(x^0) = B_1(x^0)$$

offen.

Insbesondere sieht man an diesen Beispielen, dass nicht einmal die Abzählbarkeit von  $\mathcal{U}$  in Satz 2.3.24.2 und Bemerkung 2.3.25.1 hinreichend ist.

Mit Hilfe offener Mengen, kann man die Stetigkeit einer Abbildung in mehreren Variablen gut charakterisieren

**Satz 2.3.26.** *Es sei  $M \subset \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}^k$ .  $f$  ist genau dann stetig auf  $M$ , wenn für alle offenen Mengen  $V \subset \mathbb{R}^k$  das Urbild  $f^{-1}(V)$  ein Schnitt von  $M$  mit einer offenen Menge  $U$  des  $\mathbb{R}^n$  ist, d.h.*

$$f \text{ stetig auf } M \iff \forall V \overset{\circ}{\subset} \mathbb{R}^k \exists U \overset{\circ}{\subset} \mathbb{R}^n : f^{-1}(V) = U \cap M.$$

- Bemerkung 2.3.27.** 1. Ist  $M \subset \mathbb{R}^n$  eine Teilmenge, dann sagen wir eine Teilmenge  $U \subset M$  ist *relativ offen bezüglich  $M$*  oder *offen in  $M$* , wenn es eine offene Menge  $V \subset \mathbb{R}^n$  gibt mit  $U = M \cap V$ .
2. Die Menge aller relativ offenen Mengen in  $M$  verhalten sich bezüglich Schnitt und Vereinigung wie offene Mengen.
3. Damit ist  $f : M \rightarrow \mathbb{R}^k$  genau dann stetig, wenn die Urbilder offener Mengen relativ offen sind.<sup>(i)</sup>
4. Ist  $M \subset \mathbb{R}^n$  selbst offen, dann ist die obige Charakterisierung der Stetigkeit ohne den Zusatz "relativ" richtig.

**Bemerkung 2.3.28.** Jetzt haben wir alles zusammen, um die Konvergenzaussagen im Zusammenhang mit Zahlenfolgen und alle dortigen Argumente wortwörtlich oder mit leichter Modifikation zu übernehmen.

Ebenso können wir im Wesentlichen alles, was wir über Funktionenfolgen im Zusammenhang mit gleichmäßiger Konvergenz bewiesen haben, auf Abbildungen in mehreren Variablen übertragen. Dazu müssen wir statt des Betrags die Norm  $\|\cdot\|$  verwenden.

Typische Beispiele für das "Ins-Spiel-kommen" von Folgen sind neben der Stetigkeit in Satz 2.3.14 die abschließenden charakterisierenden Eigenschaften für Offenheit und Abgeschlossenheit.

**Satz 2.3.29.** 1.  $M \subset \mathbb{R}^n$  ist offen  $\iff$

$$\forall x \in M \forall (x^{(n)})_{n \in \mathbb{N}} \subset \mathbb{R}^n, \lim_{n \rightarrow \infty} x^{(n)} = x \exists n_0 \forall k > n_0 : x^{(k)} \in M$$

2.  $M \subset \mathbb{R}^n$  ist abgeschlossen  $\iff$

$$\forall x \in \mathbb{R}^n \forall (x^{(n)})_{n \in \mathbb{N}} \subset M : \left( \lim_{n \rightarrow \infty} x^{(n)} = x \implies x \in M \right).$$

**Beweisskizze.** • zu 1.  $\implies$  Diesen Beweis führt man direkt.

Zu zeigen: Ist  $M \subset \mathbb{R}^n$  offen,  $x \in M$  und  $(x^{(n)})$  mit  $\lim_{n \rightarrow \infty} x^{(n)} = x$ , dann  $\exists n_0 \forall k > n_0 : x^{(k)} \in M$ .

---

<sup>(i)</sup>Im Rahmen des allgemeinen Topologiebegriffs wird dies zur Definition der Stetigkeit von Abbildungen zwischen zwei topologischen Räumen

- i)  $M$  ist offen und  $x \in M \implies \exists \delta_0 \forall \delta < \delta_0 : B_\delta(x) \subset M$   
 ii)  $x \in M$  und  $\lim_{n \rightarrow \infty} x^{(n)} = x \implies \forall \delta \exists n_0 \forall k > n_0 : x^{(k)} \in B_\delta(x)$

Wegen ii) gilt:  $\forall \delta < \delta_0$  gemäß i)  $\exists n_0 \forall k > n_0 : x^{(k)} \in B_\delta(x) \subset M$ .

Insgesamt gilt also:  $\exists n_0 \forall k > n_0 : x^{(k)} \in M$ .

- zu 1.  $\Leftarrow$  Diesen Beweis führt man indirekt.

Zu zeigen:  $M$  nicht offen  $\implies \exists x \in M \exists (x^{(n)})_{n \in \mathbb{N}}, \lim_{n \rightarrow \infty} x^{(n)} = x \forall n_0 \exists k \geq n_0 : x^{(k)} \notin M$ .

- i) Sei also  $M$  nicht offen. Dann gibt es  $x \in M \cap \partial M \neq \emptyset \implies \exists x \in M \forall \delta \exists y \notin M : y \in B_\delta(x)$

Wähle nun zu  $\delta = \frac{1}{n}$  ein  $y^{(n)} \in B_{\frac{1}{n}}(x)$  mit  $y_n \notin M$  gemäß i). Das gibt eine Folge  $(y^{(n)})$  mit

- ii)  $y^{(n)} \notin M$  für alle  $n$  und  
 iii)  $y_n \in B_{\frac{1}{n}}(x)$ , d.h.  $\|y_n - x\| < \frac{1}{n}$ , also  $\lim_{n \rightarrow \infty} y^{(n)} = x$ .

Für die Folge  $y^{(n)}$  gilt: zu  $n_0 = 1$  gibt es ein  $k \geq n_0$  mit  $y^{(k)} \notin M$  – nämlich jedes  $k$  erfüllt das.

- zu 2.  $\implies$  Diesen Beweis führt man indirekt.

Zu zeigen:  $(\exists x \in \mathbb{R}^n \exists (x^{(n)})_{n \in \mathbb{N}} \subset M, \lim_{n \rightarrow \infty} x^{(n)} = x : x \notin M) \implies M$  ist nicht abgeschlossen.

- i)  $(x^{(n)})_{n \in \mathbb{N}} \subset M, \lim_{n \rightarrow \infty} x^{(n)} = x$ , d.h.  $\forall \epsilon \exists n_0 \forall k \geq n_0 : x^{(k)} \in B_\epsilon(x)$   
 ii)  $x \notin M$ , d.h.  $x \in M^c$ , d.h.  $x \in \partial M$  oder  $x \in (\bar{M})^c$ . Dann gilt in beiden Fällen  $\forall \epsilon \exists y \notin M : y \in B_\epsilon(x)$ .

Wegen i) gibt es also in jedem Ball um  $x$  Punkte aus  $M$  und wegen ii) Punkte aus  $M^c$ . Damit ist  $x \in \partial M$  und damit auch  $x \in \bar{M}$ . Da nun also  $x \in \bar{M}$  und  $x \notin M$  ist  $M \neq \bar{M}$ , also  $M$  nicht abgeschlossen.

- zu 2.  $\Leftarrow$  Diesen Beweis führt man indirekt.

Zu zeigen:  $M$  nicht abgeschlossen  $\implies \exists x \exists (x^{(n)}) \subset M, \lim_{n \rightarrow \infty} x^{(n)} = x : x \notin M$ .

- i) Da  $M$  nicht abgeschlossen ist, ist  $\partial M \not\subset M$ . Sei nun  $x \in \partial M \setminus M$ , also

insbesondere  $x \notin M$ .

ii)  $x \in \partial M \implies \forall \delta \exists y \in M : y \in B_\delta(x)$ .

Wähle nun zu  $\delta = \frac{1}{n}$  ein  $y^{(n)} \in B_{\frac{1}{n}}(x)$  mit  $y^{(n)} \in M$  gemäß ii). Die so erhaltene Folge  $(y^{(n)})$  erfüllt

iii)  $\|y^{(n)} - x\| < \frac{1}{n}$ , d.h.  $\lim_{n \rightarrow \infty} y^{(n)} = x$  und

iv)  $y^{(n)} \in M$  für alle  $n$

i), iii) und iv) ist genau das, was die Folge erfüllen muss.  $\square$

### 2.3.5 Exkurs: Topologie

Die zwei Eigenschaften aus Satz 2.3.24 zusammen mit Beispiel 2.3.21.1 benutzt man als Basiseigenschaften für die folgende allgemeine Definition:

**Definition 2.3.30.** Es sei  $M$  eine beliebige Menge. Dann heißt eine Familie  $\mathcal{T} \subset \mathcal{P}(M)$  von Teilmengen von  $M$  eine *Topologie auf  $M$* , wenn folgende drei Punkte erfüllt sind:

- $\emptyset, M \in \mathcal{T}$ ,
- die beliebige Vereinigung von Elementen aus  $\mathcal{T}$  ist in  $\mathcal{T}$ ,
- der Schnitt endlich vieler Elemente aus  $\mathcal{T}$  ist in  $\mathcal{T}$ .

Die Elemente aus  $\mathcal{T}$  nennt man *offenen Mengen* der Topologie.

**Beispiel 2.3.31.** 1. Die im vorigen Abschnitt eingeführten offenen Mengen des  $\mathbb{R}^n$  bilden eine Topologie auf dem  $\mathbb{R}^n$ . Diese wird durch die offenen Bälle erzeugt und heißt die natürliche Topologie des  $\mathbb{R}^n$ , siehe Bemerkung 2.3.22.5.

2. Es sei  $A \subset \mathbb{R}^n$  und  $\mathcal{T} := \{A \cap U \mid U \overset{\circ}{\subset} \mathbb{R}^n\}$ . Dann ist  $\mathcal{T}$  eine Topologie auf  $A$ . Sie heißt die von der natürlichen Topologie auf  $\mathbb{R}^n$  induzierte Teilraumtopologie oder Relativtopologie, siehe auch Bemerkung 2.3.27.

Eine ganze Klasse weiterer Beispiele für Topologien auf dem  $\mathbb{R}^n$  erhält man wie folgt:

**Beispiel 2.3.32.** Es sei  $A \subset \mathbb{R}^n$  und  $\mathcal{T}_A = \{U \subset \mathbb{R}^n \mid U = \emptyset \text{ oder } A \subset U\}$ . Dann ist  $\mathcal{T}$  eine Topologie auf  $\mathbb{R}^n$ . Zwei wichtige Spezialfälle sind

- $\mathcal{T}_{\mathbb{R}^n} = \{\emptyset, \mathbb{R}^n\}$  (*Klumpentopologie*)
- $\mathcal{T}_{\emptyset} = \mathcal{P}(M)$  (*diskrete Topologie*)

Eine Topologie auf einer etwas abstrakteren Menge erhält man wie folgt.

**Beispiel 2.3.33.** Sei  $\tilde{M} := (\mathbb{R} \times \{-1\}) \cup (\mathbb{R} \times \{1\}) \subset \mathbb{R}^2$  und  $\sim$  die durch

$$(x, t) \sim (y, s) \iff x = y \wedge s = t \wedge xy \neq 0$$

definierte Äquivalenzrelation. Weiter sei  $\pi : \tilde{M} \rightarrow M$  die natürliche Projektion, die jedem Element aus  $\tilde{M}$  die zugehörige Äquivalenzklasse zuordnet. Wir bezeichnen die Äquivalenzklassen als

$$\begin{aligned} [x] &:= \pi(x, 1) = \pi(x, 0) = \{(x, 0), (x, 1)\} \quad \text{falls } x \neq 0, \\ [0] &:= \pi(0, 0) = \{(0, 0)\}, \\ [*] &:= \pi(0, 1) = \{(0, 1)\} \end{aligned}$$

und es sei  $M = \tilde{M} / \sim = \{[x], [0], [*] \mid x \neq 0\}$  die zugehörige Äquivalenzklasseneinteilung.

Wir nennen eine Teilmenge  $U \subset M$  nun offen, wenn ihr Urbild  $\pi^{-1}(U)$  in  $\tilde{M}$  offen ist bezüglich der Relativtopologie im  $\mathbb{R}^2$ :

$$\begin{aligned} U \subset M \text{ offen} \\ : \iff \pi^{-1}(U) \subset \tilde{M} \subset \mathbb{R}^2 \text{ offen bezüglich der Relativtopologie} \end{aligned}$$

Die offenen Mengen sind somit von einem der folgenden Typen

$$\begin{aligned} U &= \{[x] \mid x \in V, 0 \notin V, V \overset{\circ}{\subset} \mathbb{R}\}, \\ U &= \{[x] \mid x \in V, 0 \in V, V \overset{\circ}{\subset} \mathbb{R}\} \cup \{[*]\}. \end{aligned}$$

**Definition 2.3.34.** Eine Topologie  $\mathcal{T}$  auf einer Menge  $M$  heißt auf *hausdorffsch* oder *Hausdorff-Topologie*, wenn sich zwei verschiedene Punkte stets durch zwei offene Mengen trennen lassen, d.h.

$$\forall x, y \in M, x \neq y \exists U, V \in \mathcal{T} : x \in U \wedge y \in V \wedge U \cap V = \emptyset.$$

- Beispiel 2.3.35.** 1. Die natürliche Topologie auf  $\mathbb{R}^n$  ist hausdorffsch: Wähle zu zwei Punkten  $x \neq y$  die offenen Mengen als  $U = B_\epsilon(x)$  und  $V = B_\epsilon(y)$  mit  $\epsilon = \frac{\|x-y\|}{4}$ .
2. Die Topologie aus Beispiel 2.3.32 ist nur für  $A = \emptyset$  hausdorffsch: die dann offenen Mengen  $\{x\}$  und  $\{y\}$  trennen die Punkte  $x$  und  $y$ .  
Für alle anderen Wahlen von  $A$  lassen sich keine zwei Punkte durch offene Mengen trennen.
3. Die Topologie aus Beispiel 2.3.33 ist nicht hausdorffsch: die Punkte  $[0]$  und  $[*]$  lassen sich nicht trennen, aber es lassen sich alle Punkte  $[x], [0]$  für  $[x] \neq [0]$  und alle Punkte  $[x], [*]$  mit  $[x] \neq [0]$  trennen.

## 2.4 Topologie und Eigenschaften stetiger Abbildungen

Wir werden in diesem Kapitel weitere topologische Begriffe kennenlernen und untersuchen wie diese sich unter stetigen Abbildungen verhalten. Wir werden dabei einige Aussagen der Analysis in einer Variablen wiederfinden und erweitern.

### 2.4.1 Zusammenhang

Umgangssprachlich nennt man eine Menge zusammenhängend, wenn man sie nicht in zwei Teile trennen kann, bzw. wenn man von jedem Punkt der Menge durch eine Verbindung jeden anderen erreichen kann. Wir wollen untersuchen, inwiefern diese Formulierung zu Missverständnissen führen kann. Dazu präzisieren wir die zwei umgangssprachlichen Formulierungen:

**Definition 2.4.1.** 1. Eine Menge  $M \subset \mathbb{R}^n$  heißt *zusammenhängend*, wenn sie die folgende Eigenschaft besitzt:

Für alle offenen Mengen  $U, V \overset{\circ}{\subset} \mathbb{R}^n$  mit  $M \subset U \cup V$ ,  $U \cap M \neq \emptyset$  und  $M \cap V \neq \emptyset$  gilt  $M \cap U \cap V \neq \emptyset$ .

2. Eine Menge  $M \subset \mathbb{R}^n$  heißt *wegzusammenhängend*, wenn sie die folgende Eigenschaft hat:

Für alle  $x, y \in M$  gibt es einen stetigen Weg  $w : [a, b] \rightarrow M$  mit  $w(a) = x$ ,  $w(b) = y$  und  $w([a, b]) \subset M$ .

Den zweiten Begriff kann man noch etwas verstärken:

**Definition 2.4.2.** 1. Eine Menge  $M \subset \mathbb{R}^n$  heißt *sternförmig*, wenn es einen Punkt  $x^0 \in M$  gibt, sodass für alle Punkte  $y \in M$  auch die Strecke zwischen  $x^0$  und  $y$ , also  $\{x^0 + t(y - x^0) \mid t \in [0, 1]\}$ , ganz in  $M$  liegt.

2. Eine Menge  $M \subset \mathbb{R}^n$  heißt *konvex*, wenn sie bezüglich jedes Punktes in  $M$  sternförmig ist. D.h. mit allen Punkten  $x, y \in M$  liegt auch die Strecke zwischen  $x$  und  $y$ , also  $\{x + t(y - x) \mid t \in [0, 1]\}$ , ganz in  $M$ .

**Beispiel 2.4.3.** •  $\emptyset$  und  $\mathbb{R}^n$  sind konvex, sternförmig, wegzusammenhängend und zusammenhängend.

- Die punktierte Ebene  $\mathbb{R}^n \setminus \{0\}$  ist nicht sternförmig, nicht konvex, aber wegzusammenhängend. Wegen der folgenden Bemerkung 2.4.6 ist sie auch zusammenhängend.

Ein wichtiges erstes Beispiel ist anschaulich sofort klar.

**Beispiel 2.4.4.**  $M \subset \mathbb{R}$  ist genau dann zusammenhängend, wenn  $M$  ein Intervall ist.

Der Zusammenhang von Mengen ist mit der Stetigkeit von Abbildungen verträglich:

**Satz 2.4.5.** *Ist  $M \subset \mathbb{R}^n$  zusammenhängend und  $f : M \rightarrow \mathbb{R}^k$  stetig, dann ist das Bild  $f(M) \subset \mathbb{R}^k$  ebenfalls zusammenhängend.*

Mit Beispiel 2.4.4 und Satz 2.4.5 ergeben sich einige direkte Folgerungen:

**Bemerkung 2.4.6.** 1. Eine wegzusammenhängende Menge ist auch zusammenhängend.

2. Eine konvexe Menge ist wegzusammenhängend, und daher ebenfalls zusammenhängend.

3. Da Intervalle konvex sind, sind Teilmengen von  $\mathbb{R}$  genau dann zusammenhängend, wenn sie wegzusammenhängend sind.

**Beispiel 2.4.7.** Die (offenen und abgeschlossenen) Bälle im  $\mathbb{R}^n$  sind als konvexe Mengen zusammenhängend und wegzusammenhängend.

Es gilt die folgende Verallgemeinerung des Zwischenwertsatzes 1.4.10:

**Satz 2.4.8.** *Es sei  $M \subset \mathbb{R}^n$  zusammenhängend und  $f : M \rightarrow \mathbb{R}$  stetig. Für  $x, y \in M$  sei weiter  $f(x) = a$  und  $f(y) = b$ . Dann gibt es für alle  $c \in [a, b] \cup [b, a]$  ein  $z \in M$  mit  $f(z) = c$ .*

**Bemerkung 2.4.9.** Die Aussage 3. in Bemerkung 2.4.6 gilt im  $\mathbb{R}^n$  mit  $n \geq 2$  nicht mehr: Eine Menge, die zusammenhängend ist, muss nicht wegzusammenhängend sein:

Die abgeschlossene Teilmenge des  $\mathbb{R}^2$

$$M := \overline{M_0} \quad \text{mit} \quad M_0 := \left\{ (x, y) \in \mathbb{R}^2 \mid x > 0, y = \sin\left(\frac{1}{x}\right) \right\}$$

ist zusammenhängend, aber nicht wegzusammenhängend. Zur Begründung nutzt man die folgenden Überlegungen:

- Man weiß, dass  $M_0 \subset \partial M_0$ . Weiter zeigt man, dass für  $M_1 := \{0\} \times [-1, 1]$

$$M = M_0 \cup M_1 = \partial M_0.$$

Das sieht man, indem man ausnutzt, dass für alle  $-1 < r < 1$  und für alle  $\delta > 0$  ein  $m \in \mathbb{N}$  existiert mit  $\frac{1}{(m+1)\pi} < \frac{1}{m\pi} < \delta$  und dass dazu ein  $\hat{x} \in [\frac{1}{(m+1)\pi}, \frac{2}{(2m+1)\pi}]$  existiert mit  $\sin(\frac{1}{\hat{x}}) = r$ , also  $(\hat{x}, r) \in B_\delta((0, r)) \cap M_0$ . Andererseits ist  $\sin(\frac{2}{(2m+1)\pi}) = \pm 1 \neq r$ , sodass  $(\frac{2}{(2m+1)\pi}, r) \in B_\delta((0, r)) \cap M^c$ . Für  $r = 1$  modifiziere man das Argument.

- Insbesondere folgt daraus, dass jede offene Menge, die einen nichtleeren Schnitt mit  $M_1$  hat, auch einen nichtleeren Schnitt mit  $M_0$  hat.
- $M_0$  ist als Graph einer stetigen Funktion wegzusammenhängend und  $M_1$  ist als Intervall ebenfalls wegzusammenhängend.
- Um zu zeigen, dass  $M$  nicht wegzusammenhängend ist, reicht es, zu zeigen, dass kein stetiger Weg ein beliebiges  $(x, \sin(\frac{1}{x})) \in M_0$  und  $(0, 0)$  verbindet. Gäbe es diesen Weg, so zeigt man, dass  $0 \mapsto 0$  die stetige Ergänzung von  $x \mapsto \sin(\frac{1}{x})$  wäre, im Widerspruch zu Satz 1.7.11.

- Um zu zeigen, dass  $M$  zusammenhängend ist, nutzt man aus, dass  $M_0$  als wegzusammenhängende Menge zusammenhängend ist, und dass  $M_1$  abgeschlossen ist:

Es seien  $U_1, U_2 \subset \mathbb{R}^2$  offene Teilmengen mit  $U_1 \cap M \neq \emptyset$ ,  $U_2 \cap M \neq \emptyset$  und  $M \subset U_1 \cup U_2$ . Dann sind  $V_j := U_j \setminus M_1$  ebenfalls offen und nichtleer, und wegen der Vorüberlegung ist  $V_j \cap M_0 \neq \emptyset$  für  $j = 1, 2$ .

Dann ist  $M_0 \subset (U_1 \cup U_2) \setminus M_1 = V_1 \cup V_2$ . Damit ist  $M_0 \cap V_1 \cap V_2 \neq \emptyset$ , da  $M_0$  zusammenhängend ist. Damit ist aber auch  $M \cap U_1 \cap U_2$  als Obermenge von  $M_0 \cap V_1 \cap V_2$  nichtleer.

Für Teilmengen des  $\mathbb{R}^n$  gilt jedoch die folgende Aussage

**Satz 2.4.10.** *Eine offene Teilmenge des  $\mathbb{R}^n$  ist genau dann zusammenhängend, wenn sie wegzusammenhängend ist.*

Wie wir es schon zum Beweis einiger der vorstehenden Aussagen gemacht haben, kann man mit dem Verträglichkeitssatz 2.4.5 auch den Zusammenhang von weiteren Mengen zeigen.

**Beispiel 2.4.11.**  $S_r(x^0) \subset \mathbb{R}^n$  ist für  $n > 1$  zusammenhängend als Bild der Abbildung  $f : B_1(0) \setminus \{0\} \rightarrow S_r(x^0)$ ,  $f(x) = x^0 + r \frac{x}{\|x\|}$ , denn  $B_1(0) \setminus \{0\}$  ist zusammenhängend und  $f$  ist stetig.

#### 2.4.2 Kompaktheit und der Satz von Heine-Borel

Der Halbraum  $\{(x, y) \mid x \in \mathbb{R}^{\geq 0}, y \in \mathbb{R}^{n-1}\} \subset \mathbb{R}^n$  ist eine abgeschlossene Menge genau wie der achsenparallele Quader  $\{x \in \mathbb{R}^n \mid |x_i| \leq 1\} \subset \mathbb{R}^n$ . Die erste ist jedoch unbeschränkt und die zweite ist beschränkt. Diesen Unterschied wollen wir als nächstes untersuchen.

**Definition 2.4.12.** Es sei  $\mathcal{U} \subset \mathcal{P}(\mathbb{R}^n)$  eine Familie offener Teilmengen von  $\mathbb{R}^n$ .

- $\mathcal{U}$  heißt *offene Überdeckung* von  $M \subset \mathbb{R}^n$ , wenn  $M \subset \bigcup \mathcal{U}$ .
- Ist  $\mathcal{U}$  eine offene Überdeckung von  $M$  und  $\mathcal{U}' \subset \mathcal{U}$  ebenfalls, so heißt  $\mathcal{U}'$  eine *Teilüberdeckung*.

- Ist  $\mathcal{U} = \{U_1, U_2, \dots, U_\ell\}$  eine offene Überdeckung von  $M$ , so heißt  $\mathcal{U}$  eine *endliche Überdeckung*.

**Beispiel 2.4.13.**  $\mathcal{U} := \{B_r(0) \mid 0 < r \leq 1\}$  ist eine offene Überdeckung von  $B_1(0)$ .

- Die folgenden Teilmengen von  $\mathcal{U}$  sind Teilüberdeckungen von  $B_1(0)$ :

$$\mathcal{U} \supset \{B_r(0) \mid \frac{1}{2} < r \leq 1\} \supset \{B_r(0) \mid r = \frac{3}{4}, 1\} \supset \{B_1(0)\},$$

wobei die letzten beiden endliche Teilüberdeckungen sind.

- $\mathcal{U}' := \{B_r(0) \mid 0 < r < 1\} \subset \mathcal{U}$  ist ebenfalls eine Teilüberdeckung von  $B_1(0)$ . Diese besitzt jedoch keine endliche Teilüberdeckung.

**Definition 2.4.14.** Eine Teilmenge  $M \subset \mathbb{R}^n$  heißt *kompakt*, wenn jede offene Überdeckung von  $M$  eine endliche Teilüberdeckung besitzt.

Auch dieser Begriff ist mit stetigen Abbildungen verträglich:

**Satz 2.4.15.** *Ist  $M \subset \mathbb{R}^n$  kompakt und  $f : M \rightarrow \mathbb{R}^k$  stetig, so ist auch  $f(M) \subset \mathbb{R}^k$  kompakt.*

**Bemerkung 2.4.16.** 1. Die Kompaktheit hält nun die eingangs beschriebenen Mengen auseinander, denn es gilt:

Ist  $M \subset \mathbb{R}^n$  kompakt, so ist  $M$  auch abgeschlossen und beschränkt.

2. Umgekehrt ist das auch richtig, siehe Satz 2.4.17, aber nicht klar. Es gilt jedoch zunächst:

Es sei  $M \subset \mathbb{R}^n$  kompakt und  $M' \subset M$ .  $M'$  ist genau dann kompakt, wenn  $M'$  abgeschlossen ist.

**Satz 2.4.17** (Satz von Heine-Borel). *Eine Teilmenge  $M \subset \mathbb{R}^n$  ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Der Beweis benötigt das folgende fundamentale Beispiel, das wir deshalb als Satz formulieren:

**Satz 2.4.18.** *Die abgeschlossenen Bälle im  $\mathbb{R}^n$  sind kompakt.*

Im Beweis zu Satz 2.4.18 haben wir nicht explizit die verwendete Norm erwähnt, mit Hilfe derer die Bälle definiert sind, sondern nur ihre Stetigkeit verwendet. Da jede Norm auf dem  $\mathbb{R}^n$  stetig ist, ist der Satz somit ebenfalls richtig, wenn wir die euklidische Norm durch eine beliebige andere ersetzen. Als Konsequenz erhalten wir die folgende Aussage.

**Folgerung 2.4.19** (Äquivalenz von Normen). Im Zahlenraum  $\mathbb{R}^n$  sind alle Normen im folgenden Sinne äquivalent: Sind  $\|\cdot\|_1$  und  $\|\cdot\|_2$  Normen auf dem  $\mathbb{R}^n$ , so gibt es positive Zahlen  $c, C$ , sodass für alle  $x \in \mathbb{R}^n$

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

Benutzen wir für eine Argumentation eine Norm, so ist es egal, für welche wir uns entscheiden!

Aus der Kompaktheit ergeben sich einige interessante Eigenschaften für stetige Funktionen.

**Satz 2.4.20.** 1. Es sei  $M \subset \mathbb{R}^n$  kompakt und  $f : M \rightarrow \mathbb{R}$  eine stetige Funktion. Dann ist  $f$  beschränkt.

2. Es sei  $M \subset \mathbb{R}^n$  kompakt und  $f : M \rightarrow \mathbb{R}$  eine stetige Funktion. Dann nimmt  $f$  auf  $M$  ihr Maximum und ihr Minimum an, vgl. Satz 1.4.13.

3. Es sei  $M \subset \mathbb{R}^n$  kompakt und  $f : M \rightarrow \mathbb{R}^k$  eine stetige, injektive Abbildung. Dann ist  $f^{-1} : f(M) \rightarrow \mathbb{R}^n$  ebenfalls stetig.

4. Es sei  $M \subset \mathbb{R}^n$  kompakt und  $f : M \rightarrow \mathbb{R}^k$  eine stetige Abbildung. Dann ist  $f$  bereits gleichmäßig stetig, d.h.

$$\forall \epsilon > 0 \exists \delta > 0 \forall x, y \in M : (\|x - y\| < \delta \implies \|f(x) - f(y)\| < \epsilon).$$

Die folgende Aussage lässt sich auch elegant mit Hilfe der Kompaktheit beweisen.

**Satz 2.4.21.** Es seien  $M \subset \mathbb{R}^k$  und  $I = [a, b] \subset \mathbb{R}$  ein Intervall. Die Funktion  $f : M \times I \rightarrow \mathbb{R}$  sei stetig. Dann ist auch  $\hat{f} : M \rightarrow \mathbb{R}$  mit

$$\hat{f}(x) := \int_a^b f(x, t) dt$$

stetig.

Zum Abschluss wollen wir auch die Kompaktheit mit Hilfe von Folgen charakterisieren.

**Definition 2.4.22.** Eine Menge  $M \subset \mathbb{R}^n$  heißt *folgenkompakt*, wenn jede Folge  $(x^{(n)})_{n \in \mathbb{N}}$  in  $M$  eine konvergente Teilfolge besitzt, deren Grenzwert ebenfalls in  $M$  liegt.

**Satz 2.4.23.** *Die kompakten Teilmengen von  $\mathbb{R}^n$  sind genau die folgenkompakten Mengen.*

**Beweisskizze.** Im Beweis nutzen wir den Satz von Heine Borel und zeigen, dass eine Menge  $M \subset \mathbb{R}^n$  genau dann folgenkompakt ist, wenn sie beschränkt und abgeschlossen ist.  $\square$

Einen Teil des Beweises wollen wir hier noch einmal herausstellen:

**Satz 2.4.24** (Satz von Bolzano-Weierstraß). *Jede beschränkte Folge im  $\mathbb{R}^n$  besitzt eine konvergente Teilfolge.*

**Beweisskizze.** Es sei  $(x^{(n)})$  eine beschränkte Folge. Dann sind insbesondere auch alle Komponentenfolgen  $(x_i^{(n)})$  beschränkt. Starte mit  $(x_1^{(n)})$  und wähle gemäß Satz 1.2.51.2 eine konvergente Teilfolge  $(x_1^{(f_1(n))})$ . Damit haben wir die Teilfolge  $(x^{(f_1(n))})$  von  $(x^{(n)})$ , deren erste Komponente konvergiert. Betrachte nun  $(x_2^{(f_1(n))})$  und wähle genauso dort eine konvergente Teilfolge  $(x_2^{((f_2 \circ f_1)(n))})$ . Dann ist auch  $(x_1^{((f_2 \circ f_1)(n))})$  eine Teilfolge von  $(x_1^{(f_1(n))})$  und besitzt den gleichen Limes. Wir haben damit insgesamt die Folge  $(x^{((f_2 \circ f_1)(n))})$ , von der die erste und zweite Komponente konvergiert. So erhält man induktiv eine konvergente Teilfolge  $(x^{((f_n \circ \dots \circ f_1)(n))})$  der Folge  $(x^{(n)})$ .  $\square$

**Bemerkung 2.4.25.** 1. In der Formulierung von Satz 2.4.23 spielt der Zahlenraum  $\mathbb{R}^n$  eine zentrale Rolle. Die Aussage ist in allgemeinen topologischen Räumen nicht richtig: Es gibt folgenkompakte Mengen, die nicht kompakt sind, und umgekehrt.

2. Allerdings kann man den Beweis von Satz 2.4.23 abändern, sodass er in jedem metrischen Raum gültig bleibt, siehe dazu auch Bemerkung 2.4.31.

Zum Abschluss diskutieren wir noch eine anschaulich sofort einzusehende Tatsache.

**Bemerkung 2.4.26.** Zwei Teilmengen  $K \subset \mathbb{R}^k$  und  $L \subset \mathbb{R}^\ell$  sind genau dann kompakt, wenn dies für ihr Kreuzprodukt  $K \times L \subset \mathbb{R}^{k+\ell}$  gilt.

Das folgt direkt aus Satz 2.4.17 oder aus Satz 2.4.23.

Möchte man nur Definition 2.4.14 zur Begründung zurückgreifen, so folgt "⇐" aus Satz 2.4.15 und der Tatsache, dass die Projektionen  $\pi^{(k)} : \mathbb{R}^{k+\ell} \rightarrow \mathbb{R}^k$  und  $\pi^{(\ell)} : \mathbb{R}^{k+\ell} \rightarrow \mathbb{R}^\ell$  stetig sind.

Für "⇒" startet man mit einer offenen Überdeckung von  $K \times L$  und zeigt zunächst, dass die Projektionen davon auf die beiden Faktoren dort offene Überdeckungen liefern – dazu nutzt man etwa, dass  $\pi^{(k)}(B_r^{k+\ell}(x^0)) = B_r^k(\pi^{(k)}(x^0))$  ist. Nach Auswahl zweier endlicher Teilüberdeckungen von  $K$  und  $L$  liefert das schließlich eine endliche Teilüberdeckung von  $K \times L$ .

### 2.4.3 Der Banachsche Fixpunktsatz

Viele der oben definierten Begriffe (Kugeln, offen, abgeschlossen, beschränkt, Folgen und Konvergenz, Kompaktheit = Folgenkompaktheit ⇒ Abgeschlossenheit und Beschränktheit, Abbildung, Stetigkeit = Folgenstetigkeit) in den beiden letzten Kapiteln können wir direkt auf allgemeinere Klassen von Räumen einführen. Es lassen sich allerdings nicht alle Sätze und Eigenschaften übertragen. Das betrifft insbesondere diejenigen, wo man explizit benutzt hat, dass in dem zugrundeliegenden Raum  $\mathbb{R}^n$  das  $n$  endlich ist.

**Definition 2.4.27.** Es sei  $(B, \|\cdot\|)$  ein normierter Vektorraum.  $B$  heißt *Banachraum*, wenn er *vollständig* ist, d.h., wenn in  $B$  jede Cauchy-Folge konvergiert.

**Beispiel 2.4.28.** 1.  $\mathbb{R}^n$  ist ein Banachraum.

2. Es sei  $\ell = \{(x_n)_{n \in \mathbb{N}} \mid x_k \in \mathbb{R}\}$  der Vektorraum aller reellen Zahlenfolgen. Weiter sei

$$\|\cdot\|_\infty : \ell \rightarrow \mathbb{R} \cup \{\infty\} \quad \text{mit} \quad \|(x_n)\|_\infty := \sup_{n \in \mathbb{N}} |x_n|.$$

Dann ist die Menge  $\ell^\infty := \{(x_n)_{n \in \mathbb{N}} \mid \|(x_n)\|_\infty < \infty\}$  aller beschränkten Zahlenfolgen mit der Norm  $\|\cdot\|_\infty$  ein Banachraum

**Bemerkung 2.4.29.** In einem allgemeinen Banachraum muss der Satz von Heine-Borel 2.4.17 nicht gelten. Zu dessen Beweis hatten wir eine Induktion

über die endliche Dimension der abgeschlossenen Einheitskugeln gemacht und so gezeigt, dass diese kompakt sind. Dieser Beweis lässt sich auch nicht reparieren, wie das folgende Beispiel zeigt:

In  $\ell^\infty$  ist die abgeschlossene Einheitskugel nicht (folgen)kompakt. Das sieht man am besten, indem man eine Folge in  $\ell^\infty$  (also eine Folge von Folgen) konstruiert, die keine konvergente Teilfolge hat. Hier ist das etwa die Folge

$$(\mathbf{e}_n)_{n \in \mathbb{N}} \quad \text{mit} \quad \mathbf{e}_k = (0, \dots, 0, \overbrace{1}^{\text{k-te Stelle}}, 0, \dots) \in \ell^\infty.$$

Jede Teilfolge  $(\mathbf{e}_{f(n)})$  erfüllt  $\|\mathbf{e}_{f(n)} - \mathbf{e}_{f(m)}\|_\infty = 1$  für  $m \neq n$  und kann somit keine Cauchy-Folge sein. Insbesondere ist sie dann aber auch nicht konvergent.

In einem Banachraum gilt die folgende Aussage, von der wir schon eine Variante in Satz 1.4.39 kennengelernt haben.

**Satz 2.4.30** (Banachscher Fixpunktsatz). *Es sei  $M \subset B$  eine nichtleere, abgeschlossene Menge des Banachraums  $(B, \|\cdot\|)$ . Weiter sei  $f : M \rightarrow M$  eine Abbildung, für die es eine Zahl  $L < 1$  gibt, mit*

$$\|f(x) - f(y)\| \leq L\|x - y\|.$$

Dann gilt

1. Die Abbildung hat einen eindeutigen Fixpunkt  $x^* \in M$ , das heißt  $f(x^*) = x^*$ .
2. Dieser lässt sich als Grenzwert der rekursiv definierten Folge  $(x^{(n)})_{n \in \mathbb{N}}$  berechnen mit  $x^{(k)} := f(x^{(k-1)})$  für  $k \geq 1$  und  $x^{(0)} \in M$  beliebig. Es gilt dabei

$$\begin{aligned} (a) \quad & \|x^{(k)} - x^*\| \leq L \|x^{(k-1)} - x^*\|, \\ (b) \quad & \|x^{(k)} - x^*\| \leq \frac{L^k}{1-L} \|x^{(1)} - x^{(0)}\|, \\ (c) \quad & \|x^{(k)} - x^*\| \leq \frac{L}{1-L} \|x^{(k)} - x^{(k-1)}\|. \end{aligned}$$

**Bemerkung 2.4.31.** Die Aussage des Banachschen Fixpunktsatzes bleibt auch für eine größere Klasse von Räumen richtig. Man benötigt in seiner

Formulierung und in der Konstruktion der notwendigen Begriffe die Norm gar nicht, sondern lediglich den Begriff des Abstandes zweier Punkten.

Ein Raum  $M$  mit einer Abbildung  $d : M \times M \rightarrow \mathbb{R}^{\geq 0}$  heißt *metrischer Raum*, wenn

- $d(x, y) = 0 \iff x = y$ ,
- $d(x, y) = d(y, x)$  für alle  $x, y \in M$ , und
- $d(x, y) \leq d(x, z) + d(z, y)$  für alle  $x, y, z \in M$ .

Insbesondere ist jede normierte Raum durch  $d_{\|\cdot\|}(x, y) := \|x - y\|$  ein metrischer Raum.

In metrischen Räumen kann man nun den Begriff der Folge, Konvergenz, Cauchy-Folge und damit Vollständigkeit analog zu unserem Vorgehen definieren. Ebenso kann man Stetigkeit zwischen metrischen Räumen und die topologischen Begriffe einführen – Bälle werden nun mit Hilfe der Metrik  $d$  definiert:  $B_r(x^0) = \{x \in M \mid d(x, x^0) < r\}$ .

Der obige Fixpunktsatz bleibt also auch richtig, wenn man einen vollständigen metrischen Raum als Grundlage nimmt und an alle Stellen, an denen die Norm benutzt wird, diese durch die Metrik ersetzt: z.B. lautet die Lipschitzbedingung nun

$$d(f(x), f(y)) \leq L d(x, y).$$

## 2.5 Differenzierbare Abbildungen des Zahlenraums

### 2.5.1 Richtungsableitung und partielle Differentiation

Für Abbildungen  $f : I \rightarrow \mathbb{R}^n$ , die auf einem Intervall  $I \subset \mathbb{R}$  definiert sind, haben wir schon eine naheliegende Erweiterung der Differentiation eingeführt,

siehe Definition 2.2.53. Wir werden diese nun benutzen um einen Begriff der Ableitung für Abbildungen von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$  einzuführen.<sup>(i)</sup>

Es sei  $U \overset{\circ}{\subset} \mathbb{R}^n$  offen,  $f : U \rightarrow \mathbb{R}^k$  eine Abbildung und  $x^0 \in M$ . Sei weiter  $v \in \mathbb{R}^n$  ein beliebiger Vektor. Dann gibt es ein  $\epsilon > 0$ , sodass  $B_{\epsilon\|v\|}(x^0) \subset U$ . Das definiert eine Abbildung

$$f_v : ]-\epsilon, \epsilon[ \rightarrow \mathbb{R}^k, \quad f_{x^0, v}(t) = f(x^0 + tv).$$

**Definition 2.5.1.** Es sei  $M \overset{\circ}{\subset} \mathbb{R}^n$  offen,  $f : M \rightarrow \mathbb{R}^k$  eine Abbildung,  $x^0 \in M$  ein innerer Punkt,<sup>(i)</sup>  $v \in \mathbb{R}^n$  und  $\epsilon > 0$  derart, dass  $x^0 + tv \in M$  für  $-\epsilon < t < \epsilon$ . Existiert nun  $(f_{x^0, v})'(0)$  so heißt  $f$  in Richtung  $v$  differenzierbar und wir schreiben

$$D_v f(x^0) := (f_{x^0, v})'(0)$$

und nennen dies die *Richtungsableitung von  $f$  in Richtung  $v$* .

Im kartesischen Koordinatensystem gibt es ausgezeichnete Richtungen: die Koordinatenrichtungen. Wir bezeichnen die Einheitskoordinatenvektoren mit  $e_1, e_2, \dots, e_n$ .

**Definition 2.5.2.** Es sei  $M \subset \mathbb{R}^n$ ,  $f : M \rightarrow \mathbb{R}^k$  und  $x^0 \in M$ .

- Existiert in  $x^0$  die Richtungsableitung in die Koordinatenrichtung  $e_j$  so schreiben wir

$$D_{e_j} f(x^0) =: D_j f(x^0) =: \frac{\partial f}{\partial x_j}(x^0).$$

Wir nennen dies die  *$j$ -te partielle Ableitung von  $f$  in  $x^0$* .

- Existieren in  $x^0$  die Richtungsableitung in alle Koordinatenrichtungen, so sagen wir  $f$  ist in  $x^0$  *partiell differenzierbar*.
- Wir sagen  $f$  ist *partiell differenzierbar*, wenn  $f$  dies in allen Punkten aus  $M$  ist

---

<sup>(i)</sup>Da in diesem Kapitel die lineare Algebra ein wichtige Hilfsmittel sein wird, werden wir nun wieder dazu übergehen, Elemente des Zahlenraums  $\mathbb{R}^n$  als Spaltenvektoren zu schreiben.

<sup>(i)</sup>Im Folgenden schreiben wir nur  $x^0 \in M$  und setzen im Zusammenhang mit partieller Differentiation stets voraus, dass  $x^0$  ein innerer Punkt ist.

**Bemerkung 2.5.3.** 1. Schreiben wir für  $M \subset \mathbb{R}^n$  die Funktion  $f : M \rightarrow \mathbb{R}^k$  als  $f(x_1, \dots, x_n)$ , so ist

$$\begin{aligned} \frac{\partial f}{\partial x_j}(x^0) &= \left( t \mapsto f(x_1^0, \dots, x_{j-1}^0, x_j^0 + t, x_{j+1}^0, \dots, x_n^0) \right)'(0) \\ &= \left( x_j \mapsto f(x_1^0, \dots, x_{j-1}^0, x_j, x_{j+1}^0, \dots, x_n^0) \right)'(x_j^0). \end{aligned}$$

Das heißt praktisch: Wir erhalten die  $j$ -te partielle Ableitung, indem wir in  $f$  alle Variablen außer der  $j$ -ten festhalten und die so erhaltene Abbildung in einer Variablen ableiten.

2. Ist  $f = \begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} : M \rightarrow \mathbb{R}^k$  mit  $M \subset \mathbb{R}^n$  gegeben durch seine Koordinatenfunktionen und ist  $f$  in  $x^0 \in M$  partiell differenzierbar, so ist

$$\frac{\partial f}{\partial x_j}(x^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(x^0) \\ \vdots \\ \frac{\partial f_k}{\partial x_j}(x^0) \end{pmatrix}$$

**Beispiel 2.5.4.** Für  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  mit  $f(x, y, z) = \begin{pmatrix} x^2 y + z \sin(x) \\ y \cos(x) + e^x + 1 \end{pmatrix}$  ist

$$\begin{aligned} \frac{\partial f}{\partial x}(x_0, y_0, z_0) &= \begin{pmatrix} 2x_0 y_0 + z_0 \cos(x_0) \\ -y_0 \sin(x_0) + e^{x_0} \end{pmatrix}, \\ \frac{\partial f}{\partial y}(x_0, y_0, z_0) &= \begin{pmatrix} x_0^2 \\ \cos(x_0) \end{pmatrix}, \\ \frac{\partial f}{\partial z}(x_0, y_0, z_0) &= \begin{pmatrix} \sin(x_0) \\ 0 \end{pmatrix}. \end{aligned}$$

**Definition 2.5.5.** Es sei  $f = \begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} : M \rightarrow \mathbb{R}^k$  mit  $M \subset \mathbb{R}^n$  in  $x^0 \in M$  partiell differenzierbar. Die Matrix, deren Spalten aus den partiellen Ableitungen

besteht, also

$$Df(x^0) = \left( \frac{\partial f}{\partial x_1}(x^0), \dots, \frac{\partial f}{\partial x_n}(x^0) \right) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^0) & \dots & \frac{\partial f_1}{\partial x_n}(x^0) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1}(x^0) & \dots & \frac{\partial f_k}{\partial x_n}(x^0) \end{pmatrix} \in M_{n,k}\mathbb{R},$$

heißt *Funktionalmatrix von  $f$  in  $x^0$* .

**Bezeichnung 2.5.6.** Ist  $f : M \rightarrow \mathbb{R}$  eine partiell differenzierbare Funktion auf  $M \subset \mathbb{R}^n$ , so ist die Funktionalmatrix ein Zeilenvektor. Das transponierte davon ist ein Spaltenvektor und diesen bezeichnen wir als den *Gradienten von  $f$* . Wir schreiben in diesem speziellen Fall

$$\text{grad}(f) := (Df)^T = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Speziell in der Physik sind Funktionen in mehreren Veränderlichen die natürlichen Objekte und der Gradient dort die gebräuchliche Notation. Wir werden später noch sehen, dass der Gradient auch eine geometrische Interpretation besitzt.

Für eine partiell differenzierbare Abbildung  $f : M \rightarrow \mathbb{R}^k$  lässt sich für jede Komponentenfunktion der Gradient bestimmen. Mit Hilfe dieser Gradienten gilt dann

$$Df = \begin{pmatrix} \text{grad}(f_1)^T \\ \vdots \\ \text{grad}(f_k)^T \end{pmatrix} \quad \text{und}$$

**Beispiel 2.5.7.** 1. Es sei  $f(x, y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}$ , dann ist

$$\frac{\partial f}{\partial x}(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}, \quad \frac{\partial f}{\partial y}(x, y) = \begin{pmatrix} -2y \\ 2x \end{pmatrix}$$

also

$$Df(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

2. Es sei  $V(x, y, z) = -\frac{1}{\sqrt{x^2+y^2+z^2}}$ . Dann ist

$$\text{grad}(V)(x, y, z) = \frac{1}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

und  $\|\text{grad}(V)\| = V^2$ .

**Definition 2.5.8.** Es sei  $f : M \rightarrow \mathbb{R}^k$  partiell differenzierbar und alle partiellen Ableitungen  $\frac{\partial f}{\partial x_j} : M \rightarrow \mathbb{R}^k$  seien selbst wieder partiell differenzierbar. Dann schreiben wir für die zweiten Ableitungen

$$\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial \left( \frac{\partial f}{\partial x_j} \right)}{\partial x_i}.$$

Dies sind  $kn^2$  viele Funktionen auf  $M$ . Wir ordnen die zweiten partiellen Ableitungen in einer  $n \times n$ -Matrix an, und nennen diese die Hessematrix  $Hf$  von  $M$ . Ihre Komponenten sind

$$Hf_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Die Einträge dieser Matrix sind selbst Abbildungen auf  $M$  mit Werten in  $\mathbb{R}^k$ . In dem speziellen Fall  $k = 1$ , also  $f : M \rightarrow \mathbb{R}$ , sind die Einträge der Hessematrix  $Hf$  reelle Funktionen und daher  $Hf(x) \in M_n \mathbb{R}$  für alle  $x \in M$ . Wenn die höheren partiellen Ableitungen existieren, dann schreiben wir induktiv

$$\frac{\partial^\ell f}{\partial x_{i_\ell} \dots \partial x_{i_1}} := \frac{\partial \left( \frac{\partial^{\ell-1} f}{\partial x_{i_{\ell-1}} \dots \partial x_{i_1}} \right)}{\partial x_{i_\ell}}.$$

**Satz 2.5.9** (Lemma von Schwarz). *Es sei  $M \subset \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}^k$ . Existieren die zweiten partiellen Ableitungen  $\frac{\partial^2 f}{\partial x_i \partial x_j}$  und  $\frac{\partial^2 f}{\partial x_j \partial x_i}$  in einer offenen Umgebung von  $x^0$  und sind diese stetig in  $x^0$ , dann gilt*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x^0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x^0),$$

das heißt, dass in diesem Fall die Reihenfolge der Ableitungen keine Rolle spielt.

## 2.5.2 Das totale Differential

Im ersten Abschnitt haben wir die Differentiation von Abbildungen in mehreren Variablen auf die Differentiation in einer Variablen zurückgeführt. Wir wollen als nächstes den Ansatz aus Satz 1.5.5, also die lineare Approximation, auf unsere allgemeine Situation verallgemeinern. Wir werden anschließend untersuchen, wie diese mit der partiellen Differentiation zusammenhängt.

**Definition 2.5.10.** Es sei  $M \subset \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}^k$ . Die Abbildung  $f$  heißt *total differenzierbar im Punkt*  $x^0 \in M \subset \mathbb{R}^n$ , wenn es eine Matrix  $\Psi_{x^0} \in M_{n,k}\mathbb{R}$  und eine in  $x^0$  stetige Abbildung  $r_{x^0} : M \rightarrow \mathbb{R}^k$  mit  $r_{x^0}(x^0) = 0$  gibt mit

$$f(x) = f(x^0) + \Psi_{x^0}(x - x^0) + r_{x^0}(x)\|x - x^0\| \quad \text{für alle } x \in M.$$

Die Matrix bzw. lineare Abbildung  $\Psi_{x^0}$  heißt das *totale Differential der Abbildung  $f$  im Punkt  $x^0$* .

Analog zu Satz 1.5.5 erhalten wir hier auch eine weitere Charakterisierung der totalen Differenzierbarkeit

**Bemerkung 2.5.11.** Die Funktion  $f : M \rightarrow \mathbb{R}^k$  ist genau dann in  $x^0 \in M \subset \mathbb{R}^n$  total differenzierbar, wenn es eine in  $x^0$  stetige matrixwertige Abbildung  $\hat{\Psi} : M \rightarrow M_{n,k}\mathbb{R}$  gibt mit

$$f(x) = f(x^0) + \hat{\Psi}(x)(x - x_0)$$

für alle  $x \in M$ . Es gilt dann  $\hat{\Psi}(x_0)(\vec{v}) = \Psi_{x^0}(\vec{v})$  für alle  $\vec{v} \in \mathbb{R}^n$ .

**Beweisskizze.** Für " $\implies$ " definiere  $\hat{\Psi}(x) := \Psi_{x^0} + \frac{1}{\|x - x^0\|} r_{x^0}(x) \cdot (x - x^0)^T$  und für " $\impliedby$ " definiere  $r_{x^0}(x) := \frac{1}{\|x - x^0\|} \left( \hat{\Psi}(x) - \hat{\Psi}(x^0) \right) (x - x^0)$  und  $\Psi_{x^0} := \hat{\Psi}(x^0)$ .  
□

**Bemerkung 2.5.12.** Wenn  $x^0$  ein Häufungspunkt von  $M$  ist, dann gibt es die folgende zu Definition 2.5.10 äquivalente Formulierung:  $f$  ist genau dann in  $x^0$  total differenzierbar, wenn eine lineare Abbildung  $\Psi_{x^0} : \mathbb{R}^n \rightarrow \mathbb{R}^k$  gibt mit

$$\lim_{x \rightarrow x^0} \frac{1}{\|x - x^0\|} \left( f(x) - f(x^0) - \Psi_{x^0}(x - x^0) \right),$$

vergleiche auch mit Bemerkung 1.5.6.

Hier sehen wir, dass eine Abbildung genau dann total differenzierbar ist, wenn alle Komponenten total differenzierbar sind:

- Sind  $\hat{\Psi}$ ,  $\Psi_{x^0}$  und  $r_{x^0}$  die Abbildung zu  $f$  in  $x^0$ , dann sind  $\pi_j \circ \hat{\Psi}$ ,  $\pi_j \circ \Psi_{x^0}$  und  $\pi_j \circ r_{x^0}$  diejenigen zur Komponentenfunktion  $f_j$  in  $x^0$ .
- sind  $\hat{\Psi}_j$ ,  $\Psi_{j,x^0}$  und  $r_{j,x^0}$  die Abbildungen zu  $f_j$  in  $x^0$ , dann sind  $\hat{\Psi} = \begin{pmatrix} \hat{\Psi}_1 \\ \vdots \\ \hat{\Psi}_k \end{pmatrix}$ ,  $\Psi_{x^0} = \begin{pmatrix} \Psi_{1,x^0} \\ \vdots \\ \Psi_{k,x^0} \end{pmatrix}$  und  $r_{x^0} = \begin{pmatrix} r_{1,x^0} \\ \vdots \\ r_{k,x^0} \end{pmatrix}$  diejenigen zu  $f$  in  $x^0$ .

**Folgerung 2.5.13.** Für Abbildungen  $f : \mathbb{R} \supset M \rightarrow \mathbb{R}^k$ , d. h. für Wege, sind die Konzepte "komponentenweises Ableiten" gemäß Definition 2.2.53 und "totale Differentiation" gemäß Definition 2.5.10 äquivalent.

**Bemerkung 2.5.14.** Es sei  $A \in M_{n,k}\mathbb{R}$  eine  $k \times n$ -Matrix und  $L : \mathbb{R}^n \rightarrow \mathbb{R}^k$  mit  $L(x) = Ax$  die zugehörige lineare Abbildung. Dann wissen wir, dass  $L$  stetig ist, da alle Komponentenfunktionen von  $L$  Polynome vom Grad 1 sind.

Eine lineare Abbildung  $L$  heißt *beschränkt*, wenn es eine Zahl  $\lambda > 0$  gibt mit  $\|Ax\| \leq \lambda\|x\|$ . Dabei ist es unerheblich, welche Norm wir hier zugrunde legen.

Es gilt nun

- $L$  ist in  $x^0 = 0$  stetig  $\iff L$  ist in jedem Punkt stetig.
- $L$  ist beschränkt  $\iff L$  ist Lipschitz-stetig  $\iff A$  ist stetig.

Diese Aussagen sind auch wahr, wenn man statt  $\mathbb{R}^n$  und  $\mathbb{R}^k$  beliebige normierte Vektorräume zugrunde legt.

In endlichdimensionalen Vektorräumen, also insbesondere in der vorliegenden Situation, haben wir sogar:

- Jede lineare Abbildung ist beschränkt und daher stetig.

Nutzen wir die euklidische Norm, so erhalten wir

$$\frac{\|Ax\|^2}{\|x\|^2} = \frac{\langle x, A^T Ax \rangle}{\langle x, x \rangle} = R_{A^T A}(x) \leq \lambda_{\max}.$$

wobei  $\lambda_{\max}$  der größte Eigenwert von  $A^T A$  ist.<sup>(i)</sup> Es gilt somit die Beschränktheit  $\|Ax\| \leq \sqrt{\lambda_{\max}}\|x\|$ .

Die Zuordnung  $A \mapsto \|A\| := \sqrt{\lambda_{\max}}$  ist gerade die Matrixnorm bezüglich der euklidischen Norm in Urbild- und Bildraum, da die Abschätzung für einen normierten Eigenvektor von  $A^T A$  zum Eigenwert  $\lambda_{\max}$  angenommen wird. Diese Matrixnorm heißt auch *Spektralnorm* und  $\lambda_{\max}$  der *Spektralradius* von  $A^T A$ . Ein paar mehr Details zu Normen und Matrixnormen sind in Kapitel 2.8.1 zusammengetragen.

Analog zu der Situation reeller Funktionen in Folgerung 1.5.7 erhalten wir hier:

**Folgerung 2.5.15.** Ist  $x^0 \in M \subset \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}^k$  in  $x^0$  total differenzierbar, dann ist  $f$  in  $x^0$  stetig.

**Beweisskizze.** Wir nutzen die Stetigkeit von  $x \mapsto \Psi_{x^0}(x)$  und  $x \mapsto \hat{r}_{x^0}(x) := r_{x^0}(x)\|x - x^0\|$  in  $x^0$  und die Abschätzung

$$\|f(x) - f(x^0)\| \leq \|\Psi_{x^0}(x) - \Psi_{x^0}(x^0)\| + \|\hat{r}_{x^0}(x) - \hat{r}_{x^0}(x^0)\|. \quad \square$$

**Beispiel 2.5.16.** Es sei  $A = (A_{ij})_{\substack{1 \leq i \leq k \\ 1 \leq j \leq n}} \in M_{n,k}\mathbb{R}$  eine Matrix und  $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$  mit  $f(x) = A \cdot x$ . Für  $f$  existieren sämtliche partiellen Ableitung und es gilt

$$\frac{\partial f}{\partial x_j}(x^0) = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{kj} \end{pmatrix}, \quad \text{also} \quad Df(x^0) = A.$$

$f$  ist sogar in alle Richtungen differenzierbar mit Richtungsableitung

$$D_v f(x^0) = \left( t \mapsto f(x^0 + tv) = Ax^0 + tAv \right)'(0) = Av.$$

Außerdem ist  $f$  wegen

$$f(x) = f(x_0) + A \cdot (x - x^0)$$

total differenzierbar mit  $\Psi_{x^0} = A$  und  $r_{x^0} = \underline{0}$ . Wir können auch sagen: Lineare Abbildungen haben eine konstante Ableitung.

<sup>(i)</sup>Als symmetrische positiv-semidefinite Matrix besitzt  $A^T A$  nur nicht-negative Eigenwerte, und es gilt  $\lambda_{\min} \leq R_{A^T A}(x) \leq \lambda_{\max}$ . Die Funktion  $R_{A^T A}$  heißt auch der *Rayleigh-Quotient* der Matrix  $A^T A$ .

Ein erster Zusammenhang zwischen totaler Differenzierbarkeit und partieller Differentiation bzw. Richtungsableitung ergibt sich wie folgt:

**Satz 2.5.17.** *Ist  $f : M \rightarrow \mathbb{R}^k$  in einem inneren Punkt  $x^0 \in M \subset \mathbb{R}^k$  total differenzierbar, dann existieren alle Richtungsableitungen von  $f$  in  $x^0$  und es gilt*

$$D_v f(x^0) = \Psi_{x^0}(v) = Df(x^0) \cdot v.$$

**Bemerkung 2.5.18.** Die Umkehrung von Satz 2.5.17 ist in der Regel falsch, wie das folgende Beispiel zeigt:

$$\text{Es sei } f(x, y) = \begin{cases} \frac{xy^3}{x^2 + y^4} & \text{falls } (x, y) \neq (0, 0) \\ 0 & \text{falls } (x, y) = (0, 0) \end{cases}.$$

Für diese Funktion existieren zwar alle Richtungsableitungen, sie ist jedoch nicht total differenzierbar:

- Die Richtungsableitung in  $(0, 0)$  in Richtung  $v = (v_1, v_2)^T$  ist

$$D_v f(0, 0) = \lim_{t \rightarrow 0} \frac{1}{t} (f(tv) - f(0, 0)) = \lim_{t \rightarrow 0} t \frac{v_1 v_2^3}{v_1^2 + t^2 v_2^4} = 0$$

- Da die Richtungsableitungen durch  $v \mapsto D_v f(0, 0) = 0$  gegeben sind, definieren sie eine lineare Abbildung.

Wäre  $f$  in  $(0, 0)$  total differenzierbar, so wäre – wegen Satze 2.5.17 –  $\Psi_{(0,0)} = 0$  und es gäbe noch eine Restfunktion  $r_{(0,0)}$ , die stetig in  $(0, 0)$  wäre mit dem Funktionswert  $r_{(0,0)} = 0$ .

Sei nun  $(\alpha_n)_{n \in \mathbb{N}}$  eine beliebige Nullfolge und wir betrachten die gegen  $(0, 0)$  konvergierende Folge  $(\alpha_n^2, \alpha_n)_{n \in \mathbb{N}}$  in  $\mathbb{R}^2$ . Für diese gilt jedoch – im Widerspruch zur Stetigkeit von  $r_{(0,0)}$  –

$$\begin{aligned} \lim_{n \rightarrow \infty} r_{(0,0)}(\alpha_n^2, \alpha_n) &= \lim_{n \rightarrow \infty} \frac{f(\alpha_n^2, \alpha_n) - f(0, 0) - \Psi_{(0,0)}(\alpha_n^2, \alpha_n)}{\|(\alpha_n^2, \alpha_n)\|} \\ &= \lim_{n \rightarrow \infty} \frac{\frac{\alpha_n^5}{\alpha_n^4 + \alpha_n^4}}{\sqrt{\alpha_n^4 + \alpha_n^2}} = \lim_{n \rightarrow \infty} \frac{\frac{\alpha_n}{2}}{\alpha_n \sqrt{\alpha_n^2 + 1}} \\ &= \lim_{n \rightarrow \infty} \frac{1}{2\sqrt{\alpha_n^2 + 1}} = \frac{1}{2} \neq 0. \end{aligned}$$

Das Beispiel aus der vorigen Bemerkung hat die besondere Eigenschaft, dass die Einträge der Funktionalmatrix nicht stetig sind. Die partiellen Ableitungen sind

$$\frac{\partial f}{\partial x} = y^3 \frac{y^4 - x^2}{(x^2 + y^4)^2}, \quad \frac{\partial f}{\partial y} = xy^2 \frac{3x^2 - y^4}{(x^2 + y^4)^2}.$$

Beide sind in  $(0, 0)$  nicht stetig. Betrachte dazu  $\frac{\partial f}{\partial x}(s, 0)$  und  $\frac{\partial f}{\partial x}(0, s)$  sowie  $\frac{\partial f}{\partial y}(s, 0)$  und  $\frac{\partial f}{\partial y}(s^2, s)$ .

Der folgende Satz zeigt uns nun, dass genau diese fehlende Stetigkeit die Eigenschaft ist, welche die totale Differenzierbarkeit verhindert:

**Satz 2.5.19.** *Es sei  $M \overset{\circ}{\subset} \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}^k$ . Auf  $M$  sei  $f$  partiell differenzierbar und die partiellen Ableitungen seien in  $x_0$  stetig. Dann ist  $f$  in  $x_0$  total differenzierbar.*

Die folgende Situation ist die, die bei uns in der Regel vorliegt.

**Folgerung 2.5.20.**  $f : M \rightarrow \mathbb{R}^k$  ist genau dann in jedem Punkt aus  $M \overset{\circ}{\subset} \mathbb{R}^n$  total differenzierbar mit stetiger Ableitung  $x \mapsto \Psi_x$ , wenn  $f$  partiell differenzierbar ist mit stetigen partiellen Ableitungen. Wir können auf offenen Mengen sagen:

Die stetig total differenzierbaren Abbildungen sind genau die partiell differenzierbaren Abbildungen mit stetigen partiellen Ableitungen.

In dieser Situation sagen wir kurz:  $f$  ist stetig differenzierbar.

Wir beweisen nun die üblichen Ableitungsregeln, vergleiche dazu Satz 1.5.8 und Satz 1.5.10.<sup>(i)</sup>

**Satz 2.5.21** (Kettenregel). *Es seien  $M \subset \mathbb{R}^n$ ,  $N \subset \mathbb{R}^k$ ,  $f : M \rightarrow \mathbb{R}^k$  mit  $f(M) \subset N$  und  $g : N \rightarrow \mathbb{R}^\ell$ . Weiter seien  $f$  in  $x^0 \in M$  und  $g$  sei in  $f(x^0)$  total differenzierbar. Dann ist auch  $g \circ f : M \rightarrow \mathbb{R}^\ell$  in  $x^0$  total differenzierbar und es gilt*

$$D(g \circ f)(x^0) = Dg(f(x^0)) \cdot Df(x^0).$$

---

<sup>(i)</sup>Wir werden ab jetzt die lineare Abbildung  $\Psi_{x^0}$  mit  $Df(x^0)$  bezeichnen, auch wenn keine explizite Matrixdarstellung von  $\Psi_{x^0}$  gemeint ist.

**Satz 2.5.22.** *Es sei  $M \subset \mathbb{R}^n$  und  $f_1, f_2 : M \rightarrow \mathbb{R}^k$ ,  $g : M \rightarrow \mathbb{R}$  seien in  $x^0$  total differenzierbar. Dann gilt:*

1.  $f_1 + f_2 : M \rightarrow \mathbb{R}^k$  ist total differenzierbar in  $x^0$  mit

$$D(f_1 + f_2)(x^0) = Df_1(x^0) + Df_2(x^0).$$

2.  $gf_1 : M \rightarrow \mathbb{R}^k$  ist total differenzierbar in  $x^0$  mit

$$D(gf_1)(x^0) = f_1(x^0) \cdot Dg(x^0) + g(x^0) \cdot Df_1(x^0).$$

**Beweisskizze.** Zu 1.: Wir setzen die totale Differenzierbarkeit von  $f_1$  und  $f_2$  ein und erhalten  $f_1(x) + f_2(x) = f_1(x^0) + f_2(x^0) + \Psi(x - x^0) + r(x)\|x - x^0\|$  mit

$$\Psi(x - x^0) := (\Psi_1 + \Psi_2)(x - x^0), \quad r(x) := r_1(x) + r_2(x)$$

und  $r(x) \xrightarrow{x \rightarrow x^0} 0$ .

Zu 2.: Wir setzen die totale Differenzierbarkeit von  $f_1$  und  $g$  ein und erhalten  $g(x)f_1(x) = g(x^0)f_1(x^0) + \Psi(x - x^0) + r(x)\|x - x^0\|$  mit

$$\Psi(x - x^0) := g(x^0)\Psi_1(x - x^0) + \Psi_2(x - x^0)f_1(x^0)$$

und

$$r(x) := r_2(x^0)(f_1(x^0) + \Psi_1(x - x^0)) + (g(x^0) + \Psi_2(x - x^0))r_1(x)$$

mit  $r(x) \xrightarrow{x \rightarrow x^0} 0$ . □

Eine Verallgemeinerung der Produktregel aus Satz 2.5.22.2 ist durch die folgende Aussage gegeben:

**Satz 2.5.23.** *Es seien  $f : M \rightarrow \mathbb{R}^k$  und  $g : M \rightarrow \mathbb{R}^\ell$  in  $x^0 \in M$  total differenzierbar und  $B : \mathbb{R}^k \times \mathbb{R}^\ell \rightarrow \mathbb{R}^m$  sei eine bilineare Abbildung. Dann ist auch*

$$h : M \rightarrow \mathbb{R}^m \quad \text{mit} \quad h(x) := B(f(x), g(x))$$

in  $x^0$  total differenzierbar mit

$$Dh(x^0) \cdot v = B(Df(x^0) \cdot v, g(x^0)) + B(f(x^0), Dg(x^0) \cdot v).$$

**Bemerkung 2.5.24.** • Für den Beweis von 2.5.23 benötigen wir neben der Norm für die linearen Abbildungen auch eine Norm für die bilineare Abbildung  $B$ . Dies klappt analog zu den Normen linearer Abbildungen, siehe Anhang 2.8.1.

- Wir können einerseits die Produktregel aus Satz 2.5.22.2 als Folgerung aus Satz 2.5.23 betrachten. Dazu nehmen wir  $m = \ell$  und  $k = 1$  und wählen  $B$  als gewöhnliche skalare Multiplikation.

Andererseits können wir Satz 2.5.23 auch mit Hilfe von Satz 2.5.22 beweisen. Dazu betrachten wir die Komponentenfunktionen  $B_1, \dots, B_m$  von  $B$ . Dies sind dann gewöhnliche bilineare Abbildungen von  $\mathbb{R}^k \times \mathbb{R}^\ell \rightarrow \mathbb{R}$ . Sie besitzen bezüglich der Standardbasen eine Matrixdarstellung  $(B_i^{rs})_{\substack{1 \leq r \leq k, \\ 1 \leq s \leq \ell}}$ , sodass mit Hilfe der Komponentenfunktionen  $f_1, \dots, f_k$  von  $f$  und  $g_1, \dots, g_\ell$  von  $g$  gilt:

$$B(f, g) = \begin{pmatrix} \sum_{r=1}^k \sum_{s=1}^{\ell} B_1^{rs} f_r g_s \\ \vdots \\ \sum_{r=1}^k \sum_{s=1}^{\ell} B_m^{rs} f_r g_s \end{pmatrix},$$

worauf wir nun 2.5.22 anwenden können.

**Beispiel 2.5.25.** Einen wichtigen Spezialfall liefert das euklidische Skalarprodukt. Wir wählen dazu  $k = \ell$ ,  $m = 1$  und  $B = \langle \cdot, \cdot \rangle$ . Dann gilt mit  $\langle f, g \rangle = f^T \cdot g = g^T \cdot f$

$$D(f^T \cdot g) = g \cdot Df^T + f^T \cdot Dg.$$

**Folgerung 2.5.26.** Ist  $h : M \rightarrow \mathbb{R}^*$  in  $x^0$  total differenzierbar, so ist auch  $\frac{1}{h} : M \rightarrow \mathbb{R}$  total differenzierbar in  $x^0$  mit

$$D\left(\frac{1}{h}\right)(x^0) = -\frac{1}{h^2(x^0)} Dh(x^0).$$

Zum Beweis schreiben wir  $\frac{1}{h} = q \circ h$ , wobei  $q : \mathbb{R}^* \rightarrow \mathbb{R}$  mit  $q(t) = \frac{1}{t}$ , und wenden die Kettenregel an.

**Satz 2.5.27.** Es sei  $M \subset \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}^n$  sei injektiv und in  $x^0 \in M$  total differenzierbar.  $f^{-1} : f(M) \rightarrow \mathbb{R}^n$  ist genau dann im Punkt  $f(x^0)$  total

differenzierbar, wenn  $f^{-1}$  in  $f(x^0)$  stetig ist und  $\det(Df(x^0)) \neq 0$ . In diesem Fall ist

$$D(f^{-1})(f(x^0)) = (Df(x^0))^{-1}.$$

Für stetig differenzierbare Funktionen formulieren wir eine Variante des Mittelwertsatzes als Folgerung aus seiner eindimensionalen Variante.

**Satz 2.5.28.** *Es sei  $M \subset \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}^n$  stetig differenzierbar. Weiter seien  $x^0 \in M$  und  $v \in \mathbb{R}^n$  derart, dass die Strecke  $\{x^0 + tv \mid t \in [0, 1]\}$  ganz in  $M$  verläuft. Dann gilt*

1.  $f(x^0 + tv) - f(x^0) = \left( \int_0^1 Df(x^0 + tv) dt \right) \cdot v$
2.  $\|f(x^0 + v) - f(x^0)\| \leq L_0 \|v\|$  mit  $L_0 = \max_{t \in [0,1]} \|Df(x^0 + tv)\|$ .
3. Ist  $M$  konvex und kompakt, so folgt daraus  $\|f(x) - f(y)\| \leq L \|x - y\|$  mit  $L = \|Df\| = \max_{x \in M} \|Df(x)\|$ . Insbesondere ist  $f$  in diesem Fall Lipschitz-stetig.<sup>(i)</sup>

Abschließend schauen wir uns noch eine Beziehung zwischen (partieller) Differentiation und Integration an.

**Satz 2.5.29.** *Es seien  $I \subset \mathbb{R}$  und  $J = [a, b] \subset \mathbb{R}$  Intervalle. Die Funktion  $f : I \times J \rightarrow \mathbb{R}$  sei stetig und in der ersten Variable partiell differenzierbar, sodass  $\frac{\partial f}{\partial x} : I \times J \rightarrow \mathbb{R}$  stetig ist. Dann ist auch  $\hat{f} : I \rightarrow \mathbb{R}$  mit*

$$\hat{f}(x) := \int_a^b f(x, t) dt$$

differenzierbar mit

$$\hat{f}'(x) = \int_a^b \frac{\partial f}{\partial x}(x, t) dt.$$

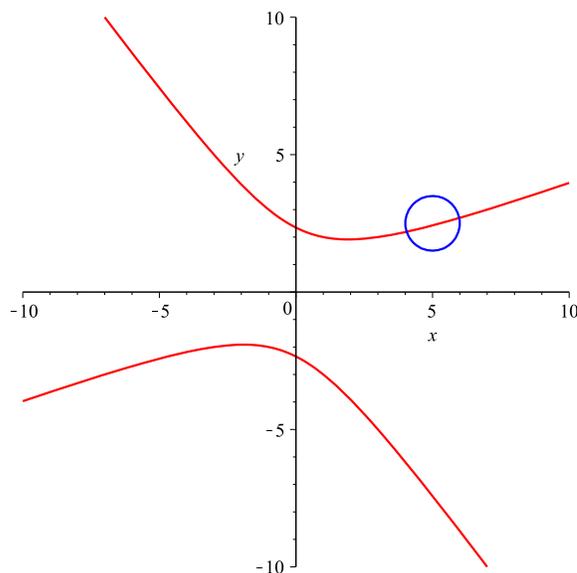
---

<sup>(i)</sup>Wir schreiben für die Konstanten  $L_0$  und  $L$  auch kurz  $\|Df\|$  und meinen damit eine Kombination aus Supremumsnorm und Matrixnorm. Dabei sollte aus dem Kontext klar sein, über welche Menge das Supremum gebildet wird.

## 2.5.3 Der Satz über implizite Funktionen und der Umkehrsatz

Eine Möglichkeit Funktionen graphisch darzustellen, ist die Angabe von Höhenmengen, siehe Abbildung 2.5.1. Schauen wir uns für  $f(x, y) = x^2 - 2y^2 - 2xy - 4$

Abbildung 2.5.1: Höhenlinie  $H_{f,-15}$  von  $f(x, y) = x^2 - 2y^2 - 2xy - 4$



die Höhenlinie  $H_{f,-15} = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = -15\}$  an und darin einen Punkt, etwa  $(x_0, y_0) = (5, \frac{1}{2}(\sqrt{97} - 5))$ . Dann ist – zumindest in einer kleinen Umgebung des Punktes<sup>(ii)</sup> – die Höhenlinie der Graph einer Funktion  $g$  über (einem Teil) der  $x$ -Achse. Es ist dann  $f(x, g(x)) = -15$ . Wir sagen dann auch: Die Gleichung  $f(x, y) = -15$  lässt sich um den Punkt  $(x_0, y_0)$  nach  $y$  auflösen. In unserem Beispiel ist das explizit machbar, indem wir die quadratische Gleichung klassisch nach  $y$  auflösen:

- Für jeden Punkt des "oberen" Teils der Höhenlinie wählen wir als offene Umgebung  $\mathbb{R} \times \mathbb{R}^{>0}$  und  $g(x) = \frac{1}{2}(\sqrt{3x^2 + 22} - x)$ .

<sup>(ii)</sup>Eine zusammenhängende Menge  $M \subset \mathbb{R}^n$  nennen wir *Umgebung von*  $x^0 \in \mathbb{R}^n$ , wenn es einen offenen Ball  $B_\epsilon(x^0)$  gibt mit  $B_\epsilon(x^0) \subset M$ . Ist  $M$  offen, so sprechen wir von einer *offenen Umgebung*.

Insbesondere ist eine offene Menge, eine Umgebung all ihrer Punkte.

- Für jeden Punkt des "unteren" Teils der Höhenlinie wählen wir als offene Umgebung  $\mathbb{R} \times \mathbb{R}^{<0}$  und  $g(x) = -\frac{1}{2}(\sqrt{3x^2 + 22} + x)$ .

Wir wollen im Folgenden für eine Abbildung  $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$  überlegen, unter welchen Bedingungen sich die Höhenmengen als Graphen darstellen lassen, sich also das Gleichungssystem  $f(x) = c$  oder

$$\begin{aligned} f_1(x_1, \dots, x_n) &= c_1 \\ &\vdots \\ f_k(x_1, \dots, x_n) &= c_k \end{aligned}$$

nach einigen der Variablen auflösen lässt.

Wir dürfen in  $H_{f,c}$  stets  $c = 0$  annehmen, indem wir zur Funktion  $\tilde{f} := f - c$  übergehen.

Wir benutzen in diesem Abschnitt die folgende Notation. Wir betrachten für  $M \subset \mathbb{R}^n$  eine Funktion

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} : M \rightarrow \mathbb{R}^k$$

mit  $k < n$ . Wir setzen nun  $\ell := n - k$ , nutzen, dass  $\mathbb{R}^n = \mathbb{R}^\ell \times \mathbb{R}^k$ , und schreiben die Elemente des  $\mathbb{R}^n$  als

$$\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^n \quad \text{mit} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_\ell \end{pmatrix} \in \mathbb{R}^\ell, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} \in \mathbb{R}^k.$$

Es sind  $\pi_1 : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$  und  $\pi_2 : \mathbb{R}^n \rightarrow \mathbb{R}^k$  die zwei Projektionen mit  $\pi_1(x, y) = x$  und  $\pi_2(x, y) = y$ .

Wir definieren für alle  $(x^0, y^0) \in M$  die folgenden Untermatrizen der Funktio-

nalmatrix  $Df(x^0, y^0)$ , von denen die zweite quadratisch ist:

$$\frac{\partial f}{\partial x}(x^0, y^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^0, y^0) & \cdots & \frac{\partial f_1}{\partial x_\ell}(x^0, y^0) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1}(x^0, y^0) & \cdots & \frac{\partial f_k}{\partial x_\ell}(x^0, y^0) \end{pmatrix} \in M_{\ell, k} \mathbb{R},$$

$$\frac{\partial f}{\partial y}(x^0, y^0) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x^0, y^0) & \cdots & \frac{\partial f_1}{\partial y_k}(x^0, y^0) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial y_1}(x^0, y^0) & \cdots & \frac{\partial f_k}{\partial y_k}(x^0, y^0) \end{pmatrix} \in M_{k, k} \mathbb{R}.$$

Es ist also

$$Df(x^0, y^0) = \left( \frac{\partial f}{\partial x}(x^0, y^0), \frac{\partial f}{\partial y}(x^0, y^0) \right) \in M_{n, k} \mathbb{R}.$$

**Bemerkung 2.5.30.** Die Untermatrix  $\frac{\partial f}{\partial x}(x^0, y^0)$  ist die Funktionalmatrix der Abbildung  $x \mapsto f(x, y^0)$  im Punkt  $x^0 \in \pi_1(M) \subset \mathbb{R}^\ell$ .

Analog ist  $\frac{\partial f}{\partial y}(x^0, y^0)$  die Funktionalmatrix der Abbildung  $y \mapsto f(x^0, y)$  im Punkt  $y^0 \in \pi_2(M) \subset \mathbb{R}^k$

Die Frage, ob sich  $f(x, y) = 0$  nach  $y$  auflösen lässt, bzw. ob sich das Gleichungssystem

$$\begin{aligned} f_1(x_1, \dots, x_\ell, y_1, \dots, y_k) &= 0 \\ &\vdots \\ f_k(x_1, \dots, x_\ell, y_1, \dots, y_k) &= 0 \end{aligned}$$

nach  $y_1, \dots, y_k$  auflösen lässt, läuft auf die Frage hinaus, ob es eine Menge  $U \subset \mathbb{R}^\ell$  gibt und eine Abbildung

$$g = \begin{pmatrix} g_1 \\ \vdots \\ g_k \end{pmatrix} : U \rightarrow \mathbb{R}^k$$

mit  $f(x, g(x)) = 0$  für alle  $x \in U$  bzw.

$$\begin{aligned} f_1(x_1, \dots, x_\ell, g_1(x_1, \dots, x_\ell), \dots, g_k(x_1, \dots, x_\ell)) &= 0 \\ &\vdots \\ f_k(x_1, \dots, x_\ell, g_1(x_1, \dots, x_\ell), \dots, g_k(x_1, \dots, x_\ell)) &= 0 \end{aligned}$$

Die Antwort liefert der folgende Satz.

**Satz 2.5.31** (Satz über implizite Funktionen). *Es seien  $U \overset{\circ}{\subset} \mathbb{R}^\ell, V \overset{\circ}{\subset} \mathbb{R}^k$  offene Mengen. Die Abbildung  $f : M \rightarrow \mathbb{R}^k$  sei auf der offenen Menge  $M = U \times V \overset{\circ}{\subset} \mathbb{R}^n$  stetig differenzierbar. Weiter seien  $(x^0, y^0) \in M$  mit*

- $f(x^0, y^0) = 0$  und
- $\det \left( \frac{\partial f}{\partial y}(x^0, y^0) \right) \neq 0$ .

*Dann gibt es  $B_\delta(x^0) \subset U$  und  $B_\epsilon(y^0) \subset V$  und eine eindeutige stetige Abbildung  $g : B_\delta(x^0) \rightarrow B_\epsilon(y^0) \subset \mathbb{R}^k$  mit*

- i)  $g(x^0) = y^0$ ,
- ii)  $f(x, g(x)) = 0$  für alle  $x \in B_\delta(x^0)$ , und
- iii)  $H_{f,0} \cap (B_\delta(x^0) \times B_\epsilon(y^0)) = \{(x, g(x)) \mid x \in B_\delta(x^0)\}$ .

*Teil iii) besagt, dass für jedes  $x \in B_\delta(x^0)$  der Wert  $g(x)$  die einzige Lösung von  $f(x, \cdot) = 0$  in  $B_\epsilon(y^0)$  ist.*

**Beweisskizze.** Zum Beweis nutzen wir den Banachschen Fixpunktsatz [2.4.30](#). Dazu schreiben wir die potentielle implizite Funktion in der etwas aufgeblähten Form

$$g(x) = g(x) + \left( \frac{\partial f}{\partial y}(x^0, y^0) \right)^{-1} f(x, g(x)).$$

Für eine "geeignete Teilmenge"  $\mathcal{X} \subset \{g : B_\delta(x^0) \rightarrow \mathbb{R}^k\}$  motiviert das die Abbildung

$$F : \mathcal{X} \rightarrow \mathcal{X} \quad \text{mit} \quad F(g) := g + \left( \frac{\partial f}{\partial y}(x^0, y^0) \right)^{-1} (f \circ (\text{id} \times g)).$$

diese "geeignete Teilmenge" ist

$$\mathcal{X} = \left\{ g : B_\delta(x^0) \rightarrow \mathbb{R}^k \mid g \text{ stetig, } g(x^0) = y^0, \|g - \underline{y^0}\| \leq \frac{1}{2}\epsilon \right\}.$$

Für geeignete  $\delta, \epsilon$  sind dann die Voraussetzung an den Banachschen Fixpunktsatz erfüllt, und die Abbildung  $F$  besitzt einen Fixpunkt  $g \in \mathcal{X}$ .  $\square$

Wenn wir nun annehmen, dass die implizite Funktion  $g$  differenzierbar ist, so erhalten wir ihre Funktionalmatrix  $Dg$  wie folgt: Wir leiten die Gleichung  $f \circ (\text{id} \times g) = 0$  ab und lösen die so erhaltene neue Gleichung  $D(f \circ (\text{id} \times g)) = 0$  nach  $Dg$  auf. Dass das tatsächlich klappt, zeigt uns der folgende Satz.

**Satz 2.5.32.** *Die implizite Funktion  $g$  aus Satz 2.5.31 ist auf einer – eventuell kleineren – Umgebung von  $x^0$  stetig differenzierbar mit Funktionalmatrix*

$$Dg(x) = -\left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1} \frac{\partial f}{\partial x}(x, g(x)).$$

**Folgerung 2.5.33.** Es sei  $M \overset{\circ}{\subset} \mathbb{R}^n$ ,  $n > k$  und  $f : M \rightarrow \mathbb{R}^k$ . Für den Punkt  $x^0 \in H_{f,c} = \{x \in M \mid f(x) = c\}$  habe die Funktionalmatrix  $Df(x^0) \in M_{n,k}\mathbb{R}$  den vollen Rang  $k$ . Dann gibt zu  $x^0 \in H_{f,c}$  eine Auswahl  $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ , sodass  $f(x) = 0$  lokal um  $x^0$  nach den Variablen  $x_{i_1}, \dots, x_{i_k}$  auflösbar ist.

Die Aussage folgt aus Satz 2.5.31: Wenn  $Df(x^0) \in M_{n,k}\mathbb{R}$  den vollen Rang  $k$  hat, gibt es eine Auswahl  $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ , sodass die entsprechende Untermatrix von  $Df(x^0)$  invertierbar ist, d.h.

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial x_{i_1}}(x^0) & \cdots & \frac{\partial f_1}{\partial x_{i_k}}(x^0) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_{i_1}}(x^0) & \cdots & \frac{\partial f_k}{\partial x_{i_k}}(x^0) \end{pmatrix} \neq 0.$$

**Bemerkung 2.5.34.** 1. Auch Gleichungen, die der Regularitätsbedingung an die Jacobimatrix von Satz 2.5.31 nicht genügen, können auflösbar sein. Diese erfüllen jedoch in der Regel nicht Satz 2.5.32.

Als Beispiel betrachten wir  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) = x - y^3$  und die Menge  $H_{f,0} = \{(x, y) \mid x - y^3 = 0\}$ . Es gilt hier  $(0, 0) \in H_{f,0}$  und  $\frac{\partial f}{\partial y} = -3y^2$ , was an der Stelle  $(0, 0)$  verschwindet. Somit ist die Frage nach der Auflösbarkeit von  $f(x, y) = 0$  um diesen Punkt im Sinne des Satzes über implizite Funktionen nicht zu beantworten. Nichtsdestotrotz gilt

$$H_{f,0} = \{(x, g(x)) \mid x \in \mathbb{R}\}$$

für die stetige, in  $x = 0$  nicht differenzierbare Funktion

$$g : \mathbb{R} \rightarrow \mathbb{R}, \quad g(x) = \sqrt[3]{x}.$$

2. Der Satz über implizite Funktionen 2.5.31 mag zunächst lediglich einen theoretischen Nutzen haben, da er nur die Existenz der Abbildung  $g$  liefert und nicht ihre explizite Darstellung. Wir erhalten jedoch mit Hilfe des Banachschen Fixpunktsatzes zumindest ein Iterationsverfahren zur näherungsweise Bestimmung der Abbildung  $g$ . Aber auch dies liefert in der Regel keine explizite Darstellung. Auch wenn uns  $g$  nicht explizit bekannt ist, wir jedoch die Lösungsmenge  $H_{f,0}$  von  $f(x, y) = 0$  kennen, so kennen wir zumindest alle Werte von  $g$  und Satz 2.5.32 liefert uns zudem zusätzlich die Werte der Ableitung der impliziten Abbildung  $g$ . In späteren Anwendungen werden uns spezielle Eigenschaften der Ableitung der impliziten Abbildung  $g$  genügen, die wir aus der Formel aus Satz 2.5.32 erhalten.

**Beispiel 2.5.35.** Wir betrachten die Höhenmenge  $H_{f,0}$  der Funktion

$$f(x, y, z) = x^2 + y^3 - 3xy$$

und untersuchen die Frage nach der Auflösbarkeit der Gleichung  $f(x, y, z) = 0$  in der Nähe eines Punktes  $(x_0, y_0, z_0) \in H_{f,0}$ . Da die beschreibende Gleichung nicht von  $z$  abhängt, ist die Schnittmenge von  $H_{f,0}$  mit der zur  $xy$ -Ebene parallelen Ebene  $E_c := \{(x, y, c) \mid x, y \in \mathbb{R}\}$  stets durch

$$H_{f,0} \cap E_c = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^3 - 3xy = 0\} \times \{c\} =: \hat{H}_{f,0} \times \{c\},$$

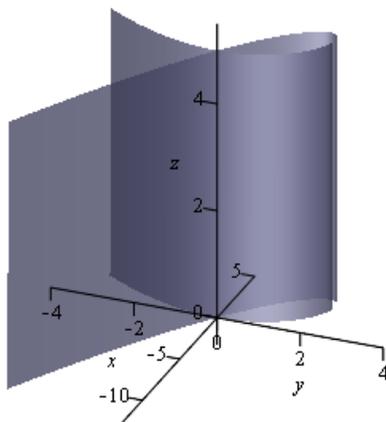
sodass  $H_{f,0} = \hat{H}_{f,0} \times \mathbb{R}$ .

gegeben. Damit ist die Menge  $H_{f,0}$  ein gerader Zylinder über der Menge  $\hat{H}_{f,0}$ , siehe Abbildung 2.5.2.

Die Funktionalmatrix von  $f$  ist

$$Df(x, y, z) = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = (2x - 3y, 3(y^2 - x), 0).$$

- Wegen  $\frac{\partial f}{\partial z} = 0$  ist  $f(x, y, z) = 0$  in keinem Punkt nach  $z$  auflösbar.

Abbildung 2.5.2: Die Menge  $H_{f,0}$  für  $f(x, y, z) = x^2 + y^3 - 3xy$ 

- Die Punkte, in denen  $f(x, y, z) = 0$  nicht nach  $x$  auflösbar (im Sinne von Satz 2.5.31) ist, sind durch die Gleichung  $\frac{\partial f}{\partial x} = 0$ , also durch  $x = \frac{3}{2}y$  gegeben. Setzen wir das in  $f(x, y, z) = 0$  ein, so ergibt sich

$$0 = y^2 \left( y - \frac{9}{4} \right) \iff y = 0 \vee y = \frac{9}{4}.$$

Setzen wir das ein in die Ausgangsgleichung  $2 = 3y$  so ergeben sich die Punkte

$$(0, 0, z_0) \quad \text{und} \quad \left( \frac{27}{8}, \frac{9}{4}, z_0 \right).$$

- Die Punkte, in denen  $f(x, y, z) = 0$  nicht nach  $y$  auflösbar (im Sinne von Satz 2.5.31) ist, sind durch die Gleichung  $\frac{\partial f}{\partial y} = 0$ , also durch  $x = y^2$  gegeben. Setzen wir das in  $f(x, y, z) = 0$  ein, so ergibt sich

$$0 = y^3(y - 2) \iff y = 0 \vee y = 2.$$

Das Zurücksetzen in  $x^2 = y$  liefert die Punkte

$$(0, 0, z_0) \quad \text{und} \quad (4, 2, z_0).$$

- Insbesondere lässt sich  $f(x, y, z) = 0$  um  $(0, 0, z_0)$  nach keiner der Variablen auflösen.

Wir untersuchen jetzt die "klassischen Auflösungen" nach  $x$ . Wir wissen, dass es um jeden Punkt  $(x_0, y_0, z_0)$  mit Ausnahme von  $(0, 0, z_0)$  und  $(\frac{27}{8}, \frac{9}{4}, z_0)$  lokale Funktionen gibt. "Klassische Auflösung" meint hier: Wir suchen explizite Funktionen  $x = g(y, z)$  mit  $f(g(y, z), y, z) = 0$ .

Dazu lösen wir die quadratische Gleichung  $x^2 + y^3 - 3xy = 0$  nach  $x$  auf und erhalten die zwei Lösungen

$$g_{\pm}(y, z) = \frac{3}{2}y \pm \sqrt{\frac{9}{4}y^2 - y^3} = \frac{3}{2} \left( y \pm |y| \sqrt{1 - \frac{4y}{9}} \right)$$

Statt dieser Funktionen können wir auch die modifizierten Funktionen

$$\hat{g}_{\pm}(y, z) = \frac{3y}{2} \left( 1 \pm \sqrt{1 - \frac{4y}{9}} \right)$$

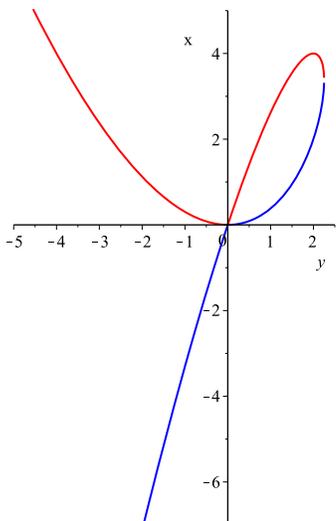
betrachten. Sie haben den Vorteil, dass sie auch in  $(0, z)$  stetig differenzierbar sind.

Die lokalen impliziten Funktionen, gemäß Satz 2.5.31 sind dann durch die jeweiligen Zweige gegeben. Ist z.B.  $(x_0, y_0, z_0)$  mit  $y_0 < 0, x_0 > 0$ , so ist  $g_+|_{\mathbb{R}^{<0} \times \mathbb{R}}$  die gesuchte Funktion.

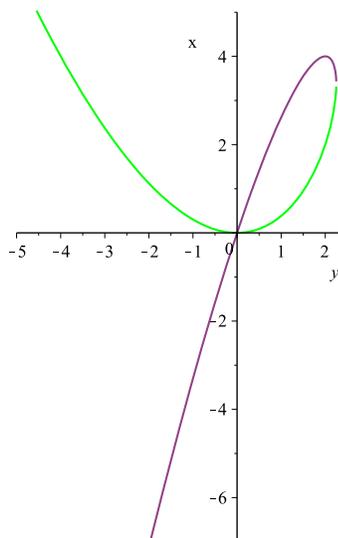
Da das Problem im wesentlichen ein Problem in der  $xy$ -Ebene ist, werden wir uns in den Skizzen auf die Projektionen in diese Ebene beschränken; das bedeutet, wir betrachten  $\hat{H}_{f,0}$ . Alle obigen Funktionen sind auf der Menge  $] -\infty, \frac{9}{4}]$  definiert und ihr Verlauf ist in Abbildung 2.5.3 skizziert.

Abbildung 2.5.3: Die Graphen zu  $\hat{H}_{f,0} = \{(x, y) \mid x^2 + y^3 - 3xy = 0\}$

Die Projektion der Graphen von  $g_+$  (rot) und  $g_-$  (blau) in die  $yx$ -Ebene



Die Projektion der Graphen von  $\hat{g}_+$  (grün) und  $\hat{g}_-$  (violett) in die  $yx$ -Ebene



Als Anwendung des Satzes über implizite Funktionen, Satz 2.5.31 lässt sich nun die Frage beantworten, unter welchen Bedingungen eine Abbildung  $f : M \rightarrow \mathbb{R}^n$  auf  $M \subset \mathbb{R}^n$  – zumindest im Kleinen – eine Umkehrabbildung  $f^{-1} : M' \rightarrow \mathbb{R}^n$  besitzt, wobei  $M' \subset f(M) \subset \mathbb{R}^n$ .

**Satz 2.5.36** (Umkehrsatz). *Es sei  $M \overset{\circ}{\subset} \mathbb{R}^n$  offen und  $f : M \rightarrow \mathbb{R}^n$  stetig differenzierbar. Weiter gebe es ein  $x^0 \in M$  mit  $\det(Df(x^0)) \neq 0$  und wir schreiben  $y^0 := f(x^0)$ . Dann gibt es offene Umgebungen  $U \subset M$  von  $x^0$  und  $V \subset \mathbb{R}^n$  von  $y^0$ , sodass gilt:*

- i)  $f : U \rightarrow V$  ist bijektiv und
- ii)  $g := (f|_U)^{-1} : V \rightarrow U$  ist stetig differenzierbar mit
- iii)  $Dg(y) = (Df(x))^{-1}$  für alle  $y \in V$  und  $y = f(x)$  (bzw.  $x = g(y)$ ).

**Beweisskizze.** Betrachte  $F(x, y) = y - f(x)$  und wende darauf den Satz über implizite Funktionen an.  $\square$

**Bemerkung 2.5.37.** Im Gegensatz zur Situation in Dimension  $n = 1$  gemäß Folgerung 1.5.23, garantiert in höheren Dimensionen selbst die Regularität von  $Df$  in jedem Punkt lediglich die lokale Umkehrbarkeit der Abbildung  $f$ . Dazu betrachte  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $f(x, y) = \begin{pmatrix} \rho(x) \cos(y) \\ \rho(x) \sin(y) \end{pmatrix}$ , wobei  $\rho : \mathbb{R} \rightarrow \mathbb{R}$  mit  $\rho(x)\rho'(x) \neq 0$  für alle  $x \in \mathbb{R}$ . Für diese Abbildung gilt

$$Df(x, y) = \begin{pmatrix} \rho'(x) \cos(y) & -\rho(x) \sin(y) \\ \rho'(x) \sin(y) & \rho(x) \cos(y) \end{pmatrix}$$

also  $\det(Df(x, y)) = \rho(x)\rho'(x) \neq 0$ . Jedoch ist  $f$  sicher nicht injektiv, denn es ist  $f(x, y) = f(x, y + 2\pi)$ .

**Definition/Bemerkung 2.5.38.** • Es sei  $M \overset{\circ}{\subset} \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}^n$ . Die Abbildung  $f$  heißt Diffeomorphismus (auf  $f(M)$ ), wenn  $f(M) \subset \mathbb{R}^n$  offen ist,  $f : M \rightarrow f(M)$  bijektiv ist und  $f$  und  $f^{-1}$  stetig differenzierbar sind.

- Insbesondere ist eine injektive, stetig differenzierbare Abbildung  $f : M \rightarrow \mathbb{R}^n$  genau dann ein Diffeomorphismus, wenn  $\det(Df(x)) \neq 0$  für alle  $x \in M$ .

## 2.6 Höhere Ableitungen und Extremwerte

### 2.6.1 Höhere Ableitungen und der Satz von Taylor

Wir werden uns hier auf Abbildungen beschränken, die auf offenen Teilmengen definiert sind.

**Definition 2.6.1.** Es sei  $M \overset{\circ}{\subset} \mathbb{R}^n$ . Wir definieren und interpretieren – falls die Ausdrücke existieren –

$$Df : M \rightarrow M_{n,k} \mathbb{R} \simeq \mathbb{R}^{nk},$$

$$D^2 f := D(Df) : M \rightarrow M_{n,nk} \mathbb{R} \simeq \mathbb{R}^{n^2 k},$$

und induktiv weiter

$$D^r f := D(D^{r-1} f) : M \rightarrow M_{n,n^{r-1}k} \mathbb{R} \simeq \mathbb{R}^{n^r k}.$$

- Eine Abbildung  $f : M \rightarrow \mathbb{R}^k$  heißt  $\ell$ -mal total differenzierbar, wenn die Funktionalmatrizen  $D^r f$  für  $r = 1, \dots, \ell - 1$  existieren und stetig sind, sowie  $D^{\ell-1} f$  total differenzierbar ist.
- $f$  heißt  $\ell$ -mal stetig differenzierbar, wenn  $D^\ell f$  existiert und stetig ist.

**Bemerkung 2.6.2.** • Eine Matrixdarstellung von  $D^2 f$  ist die Hessematrix  $Hf$  von  $f$ .

- Die Einträge der "Matrix"-wertigen Abbildung  $D^\ell f : M \rightarrow \mathbb{R}^{n^\ell k}$  sind die  $\ell$ -ten partiellen Ableitungen von  $f$ :

$$(D^\ell f)_{i_1 \dots i_\ell j} = \frac{\partial^\ell f_j}{\partial x_{i_1} \dots \partial x_{i_\ell}}$$

für  $1 \leq i_1, \dots, i_\ell \leq n, 1 \leq j \leq k$ .

**Beispiel 2.6.3.**  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3, f(x, y) = \begin{pmatrix} \sin(y) + e^x \\ x^2 y \\ y^2 x + x^2 \end{pmatrix}$ . Dann ist

$$Df = \begin{pmatrix} e^x & \cos(y) \\ 2xy & x^2 \\ y^2 + 2x & 2yx \end{pmatrix} \simeq \begin{pmatrix} e^x \\ 2xy \\ y^2 + 2x \\ \cos(y) \\ x^2 \\ 2yx \end{pmatrix}$$

$$D^2 f = \begin{pmatrix} e^x & 0 \\ 2y & 2x \\ 2 & 2y \\ 0 & -\sin(y) \\ 2x & 0 \\ 2y & 2x \end{pmatrix} \simeq \begin{pmatrix} e^x \\ 2y \\ 2 \\ 0 \\ 2x \\ 2y \\ 0 \\ 2x \\ 2y \\ -\sin(y) \\ 0 \\ 2x \end{pmatrix}$$

$$D^3 f = \begin{pmatrix} e^x & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \\ 2 & 0 \\ 0 & 2 \\ 0 & 0 \\ 2 & 0 \\ 0 & 2 \\ 0 & -\cos(y) \\ 0 & 0 \\ 2 & 0 \end{pmatrix}, \quad D^4 f(x) = \begin{pmatrix} e^x & 0 \\ \mathbf{0}_{18 \times 2} \\ 0 & \sin(y) \\ \mathbf{0}_{4 \times 2} \end{pmatrix}, \text{ etc.}$$

**Bemerkung 2.6.4.** Wie im Fall  $\ell = 1$  gilt auch für höhere Ableitungen:

- Jede  $\ell$ -mal total differenzierbare Abbildung ist  $\ell$ -mal partiell differenzierbar.
- Die  $\ell$ -mal stetig differenzierbaren Abbildungen sind genau die  $\ell$ -mal partiell differenzierbaren Abbildungen mit stetigen partiellen Ableitungen.

Eine Verallgemeinerung des Lemmas von Schwarz, Satz 2.5.9, lautet wie folgt

**Satz 2.6.5.** Ist  $f : M \rightarrow \mathbb{R}^k$  auf  $M \overset{\circ}{\subset} \mathbb{R}^n$   $\ell$ -mal stetig differenzierbar, so ist für jede Permutation  $\tau \in \mathcal{S}_\ell$

$$\frac{\partial^\ell f}{\partial x_{i_1} \dots \partial x_{i_\ell}} = \frac{\partial^\ell f}{\partial x_{\tau(i_1)} \dots \partial x_{\tau(i_\ell)}}.$$

**Bezeichnung 2.6.6.** Wegen des Lemmas von Schwarz, können wir die Ableitungen "sortieren". Ist  $\ell = \ell_1 + \dots + \ell_n$  und kommen in der partiellen Ableitung  $\frac{\partial^\ell f}{\partial x_{i_1} \dots \partial x_{i_\ell}}$  die Ableitungen nach  $x_1$  genau  $\ell_1$ -mal, die nach  $x_2$  genau  $\ell_2$ -mal,  $\dots$ , und die nach  $x_n$  genau  $\ell_n$ -mal, dann gilt

$$\frac{\partial^\ell f}{\partial x_{i_1} \dots \partial x_{i_\ell}} = \frac{\partial^\ell f}{\partial x_1^{\ell_1} \dots \partial x_n^{\ell_n}}$$

Für jedes  $n$ -Tupel  $\vec{\ell} = (\ell_1, \dots, \ell_n)$  schreiben wir mit  $|\vec{\ell}| = \sum_{i=1}^n \ell_i$

$$\frac{\partial^{|\vec{\ell}|} f}{\partial x^{\vec{\ell}}} := \frac{\partial^{|\vec{\ell}|} f}{\partial x_1^{\ell_1} \dots \partial x_n^{\ell_n}}.$$

Weiter schreiben wir  $\vec{\ell}! := \ell_1! \ell_2! \cdot \dots \cdot \ell_n!$  und  $x^{\vec{\ell}} = x_1^{\ell_1} x_2^{\ell_2} \cdot \dots \cdot x_n^{\ell_n}$ .

Mit Hilfe dieser Notation lässt sich die Verallgemeinerung der Taylorformel aus Satz 1.5.31 elegant formulieren.

**Satz 2.6.7** (Satz von Taylor). *Sei  $M \overset{\circ}{\subset} \mathbb{R}^n$  offen und  $f : M \rightarrow \mathbb{R}$  eine  $(m+1)$ -mal total differenzierbare Funktion. Dann gibt es zu je zwei Punkten  $x^0, x \in M$ , für die  $\{x^0 + t(x - x^0) \mid t \in [0, 1]\} \subset M$ , einen Punkt  $y$  auf dieser Strecke, sodass*

$$f(x) = f(x^0) + \sum_{\vec{\ell}, 1 \leq |\vec{\ell}| \leq m} \frac{1}{\vec{\ell}!} \frac{\partial^{|\vec{\ell}|} f}{\partial x^{\vec{\ell}}}(x^0) (x - x^0)^{\vec{\ell}} + \sum_{\vec{\ell}, |\vec{\ell}| = m+1} \frac{1}{\vec{\ell}!} \frac{\partial^{|\vec{\ell}|} f}{\partial x^{\vec{\ell}}}(y) (x - x^0)^{\vec{\ell}}.$$

Schreiben wir  $x = x^0 + v$  so ergibt sich

$$f(x^0 + v) = f(x^0) + \sum_{\vec{\ell}, 1 \leq |\vec{\ell}| \leq m} \frac{1}{\vec{\ell}!} \frac{\partial^{|\vec{\ell}|} f}{\partial x^{\vec{\ell}}}(x^0) v^{\vec{\ell}} + \sum_{\vec{\ell}, |\vec{\ell}| = m+1} \frac{1}{\vec{\ell}!} \frac{\partial^{|\vec{\ell}|} f}{\partial x^{\vec{\ell}}}(y) v^{\vec{\ell}}.$$

**Definition 2.6.8.** Das Polynom

$$T_{f,m,x^0}(x) := f(x^0) + \sum_{\vec{\ell}, 1 \leq |\vec{\ell}| \leq m} \frac{1}{\vec{\ell}!} \frac{\partial^{|\vec{\ell}|} f}{\partial x^{\vec{\ell}}}(x^0) (x - x^0)^{\vec{\ell}}$$

heißt das  $m$ -te Taylorpolynom der Funktion  $f$  um  $x^0$ . Die Differenz

$$f(x) - T_{f,m,x^0}(x) = \sum_{\vec{\ell}, |\vec{\ell}| = m+1} \frac{1}{\vec{\ell}!} \frac{\partial^{|\vec{\ell}|} f}{\partial x^{\vec{\ell}}}(y) (x - x^0)^{\vec{\ell}}$$

heißt das  $m$ -te Restglied.

**Beispiel 2.6.9.**  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  so lässt sich das zweite Taylorpolynom wie folgt schreiben

$$\begin{aligned} T_{f,2,x^0}(x_0 + v) &= f(x^0) + \sum_{i=1}^n \frac{\partial f}{\partial x^i}(x^0) v_i + \sum_{i,j=1}^n \frac{1}{2} \frac{\partial^2 f}{\partial x^i \partial x^j}(x^0) v_i v_j \\ &= f(x^0) + \langle \text{grad}(f)(x^0), v \rangle + \frac{1}{2} v^T H f(x^0) v \end{aligned}$$

**Bemerkung 2.6.10.** Die  $m$ -te Ableitung einer Funktion  $f : M \rightarrow \mathbb{R}$  mit  $M \subset \mathbb{R}^n$  definiert in  $x^0 \in M$  eine symmetrische,  $m$ -fache Linearform  $Q_{f,x^0,m}$  durch

$$Q_{f,x^0,m}(v^1, \dots, v^m) := \sum_{i_1, \dots, i_m=1}^n \frac{\partial^m f}{\partial x_{i_1} \dots \partial x_{i_m}} v_{i_1}^1 \cdot \dots \cdot v_{i_m}^m$$

d.h. insbesondere<sup>(i)</sup>

$$Q_{f,x^0,m}(v, \dots, v) = m! \sum_{\vec{\ell}, |\vec{\ell}|=m} \frac{1}{\ell!} \frac{\partial^{|\vec{\ell}|} f}{\partial x^{\vec{\ell}}} v^{|\vec{\ell}|}.$$

Es ist speziell

$$Q_{f,x^0,1}(v) = Df(x^0) \cdot v \quad \text{und} \quad Q_{f,x^0,2}(v, w) = v^T \cdot Hf(x^0) \cdot w.$$

und das Taylorpolynom  $m$ -ter Ordnung schreibt sich als

$$T_{f,m,x^0}(x) = f(x^0) + \sum_{\ell=1}^m \frac{1}{\ell!} Q_{f,x^0,\ell}(\underbrace{x - x^0, \dots, x - x^0}_{\ell\text{-mal}})$$

### 2.6.2 Extremwerte ohne und mit Nebenbedingungen

Wir beginnen mit der aus der Anschauung stammenden Definition für ein Extremum.

**Definition 2.6.11.** Es sei  $f : M \rightarrow \mathbb{R}$  eine Funktion auf der Menge  $M \subset \mathbb{R}^n$ . Ein Punkt  $x^0 \in M$  heißt

- *lokales Maximum*, wenn es eine Umgebung  $B_\epsilon(x^0)$  von  $x^0$  gibt, sodass  $f(x^0) \geq f(x)$  für alle  $x \in B_\epsilon(x^0) \cap M$ .
- *globales Maximum*, wenn  $f(x^0) \geq f(x)$  für alle  $x \in M$ .

<sup>(i)</sup>Es sei  $\vec{\ell} \in \mathbb{N}^n$  und  $|\vec{\ell}| = \ell_1 + \ell_2 + \dots + \ell_n = m$ . Dann entspricht  $\binom{m}{\ell_1, \dots, \ell_n} = \frac{m!}{\ell_1! \dots \ell_n!} = \frac{m!}{\ell!}$  der Anzahl der verschiedenen Möglichkeiten, die Zahlen  $\{1, 2, \dots, n\}$  so anzuordnen, dass genau  $\ell_1$ -mal die 1,  $\ell_2$ -mal die 2,  $\dots$  und  $\ell_n$ -mal die  $n$  vorkommt. Es ist z.B.  $(a_1 + \dots + a_r)^n = \sum_{\substack{i_1, \dots, i_r \\ i_1 + \dots + i_r = n}} \binom{n}{i_1, \dots, i_r} a_1^{i_1} \cdot \dots \cdot a_r^{i_r}$ .

- *lokales Minimum*, wenn es eine Umgebung  $B_\epsilon(x^0)$  von  $x^0$  gibt, sodass  $f(x^0) \leq f(x)$  für alle  $x \in B_\epsilon(x^0) \cap M$ .
- *globales Minimum*, wenn  $f(x^0) \leq f(x)$  für alle  $x \in M$ .

**Bemerkung 2.6.12.** Ist  $f$  stetig und  $M$  kompakt, dann wissen wir bereits, dass ein globales Maximum und ein globales Minimum existiert, siehe Satz 2.4.20.2.

### 2.6.2.1 Extremwerte ohne Nebenbedingungen

Der folgende Satz ergibt sich sofort aus seiner eindimensionalen Variante Satz 1.5.18:

**Satz 2.6.13.** *Es sei  $f : M \rightarrow \mathbb{R}$  und  $x^0$  ein innerer Punkt und ein lokales Minimum oder lokales Maximum. Existieren alle partiellen Ableitungen von  $f$  in  $x_0$ , so ist  $\text{grad}f(x^0) = 0$ .*

**Bemerkung 2.6.14.** Zur Berechnung der möglichen lokalen Extrema müssen wir das Gleichungssystem

$$\frac{\partial f}{\partial x_1} = 0, \quad \frac{\partial f}{\partial x_2} = 0, \quad \dots, \quad \frac{\partial f}{\partial x_n} = 0$$

von  $n$  Gleichungen für die  $n$  Variablen  $x_1, \dots, x_n$  untersuchen.

Ähnlich wie im eindimensionalen Fall erhalten wir hier mit Hilfe der Taylorformel ein hinreichendes Kriterium für die Existenz von Extremwerten.

**Satz 2.6.15.** *Es sei  $x^0 \in M \subset \mathbb{R}^n$  ein innerer Punkt und ein lokales Extremum der Funktion  $f : M \rightarrow \mathbb{R}$ . Weiter sei  $m > 1$  die kleinste Zahl mit  $D^m f(x^0) \neq 0$  und  $D^m f$  sein in  $x^0$  stetig.*

1. Ist  $m$  ungerade, dann ist  $x^0$  kein Extremum.
2. Ist  $m$  gerade, dann ist
  - $x^0$  ein lokales Maximum, wenn  $Q_{f,x^0,m}$  negativ definit ist.
  - $x^0$  ein lokales Minimum, wenn  $Q_{f,x^0,m}$  positiv definit ist.

- $x^0$  einen Sattelpunkt, wenn  $Q_{f,x^0,m}$  indefinit ist.

**Bemerkung 2.6.16.** Der obige Satz liefert insofern keine völlige Charakterisierung, als dass er keine Aussage über die Eigenschaften von  $x^0$  zulässt, wenn  $Q_{f,x^0,m}$  positiv oder negativ semidefinit ist.

Im Fall  $m = 2$  entspricht die Untersuchung des Punktes  $x^0$  mit  $Df(x^0) = 0$  auf Extrema der Untersuchung der Hessematrix  $Hf(x^0)$  auf Definitheit. Ist darüber hinaus die Dimension  $n = 2$ , so können wir den speziellen Fall  $m = 2$  wie folgt formulieren:

**Folgerung 2.6.17.** Es sei  $(a, b) \in M \subset \mathbb{R}^2$  ein innerer Punkt mit  $\frac{\partial f}{\partial x}(a, b) = \frac{\partial f}{\partial y}(a, b) = 0$ . Weiter sei die Hessematrix  $Hf(a, b)$  nicht die Nullmatrix. Dann ist

- $(a, b)$  ein lokales Maximum, wenn  $\left(\frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2\right)(a, b) > 0$  und  $\frac{\partial^2 f}{\partial x^2}(a, b) < 0$ .
- $(a, b)$  ein lokales Minimum, wenn  $\left(\frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2\right)(a, b) > 0$  und  $\frac{\partial^2 f}{\partial x^2}(a, b) > 0$ .
- $(a, b)$  kein lokales Extremum, wenn  $\left(\frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2\right)(a, b) < 0$ .

### 2.6.2.2 Extremwerte mit Nebenbedingungen

In praktischen Anwendungen interessiert man sich in der Regel nicht für die allgemeinen Extremwerte einer Funktion, sondern für speziellere Werte.

Wir betrachten eine Leuchtquelle in einem Zimmer und die Leuchtstärke in den Punkten des Raums werde durch eine Funktion  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  beschrieben. Bei einer punktförmigen Lichtquelle wird die Lichtstärke im Idealfall mit dem Abstand von der Lichtquelle sinken, sodass  $f(\vec{x}) < f(\vec{y})$  für  $\|\vec{x} - \vec{a}^0\| > \|\vec{y} - \vec{a}^0\|$ , wobei  $\vec{a}^0$  der Ort der Lichtquelle ist. In dem Zimmer befindet sich ein Schreibtisch, und gesucht ist der Punkt auf dem Tisch, der am hellsten ausgeleuchtet wird. Das bedeutet: Wir suchen nicht das Maximum der Leuchtstärke, sondern das Maximum unter der Nebenbedingung, dass

der gesuchte Punkt auf dem Schreibtisch liegen soll. Die mathematische Formulierung könnte etwa wie folgt lauten:

Wir suchen das Maximum der Funktion  $f(\vec{x}) = \frac{b}{\|\vec{x} - \vec{a}^0\|^2}$  unter der Nebenbedingung  $g(\vec{x}) = \langle \vec{x}, \vec{n} \rangle - c = 0$ , wobei  $\langle \vec{x}, \vec{n} \rangle = c$  eine die Tischoberfläche beschreibende Ebenengleichung ist –  $\vec{n}$  ist also eine Normale der Ebene.

Wir wollen dieses Problem formalisieren.

**Definition 2.6.18.** Gegeben sei eine Menge  $M \subset \mathbb{R}^n$ , eine Funktion  $f : M \rightarrow \mathbb{R}$  und für  $k < n$  eine Abbildung  $g : M \rightarrow \mathbb{R}^k$ . Wir sagen, dass  $f$  im Punkt  $x^0$  ein lokales Maximum (lokales Minimum) unter der Nebenbedingung  $g(x) = 0$  besitzt, wenn:

1.  $g(x^0) = 0$  d.h.  $x^0 \in H_{g,0}$ .
2. Es gibt ein  $B_\epsilon(x^0)$ , sodass  $f(x^0) > f(y)$  für alle  $y \in H_{g,0} \cap B_\epsilon(x^0)$ .

**Bemerkung 2.6.19.** Betrachten wir die Komponentenfunktionen der Nebenbedingung  $g : M \rightarrow \mathbb{R}^k$ , also

$$g_1(x_1, \dots, x_n) = 0, \dots, g_k(x_1, \dots, x_n) = 0,$$

so können wir diese in speziellen Situationen gegebenenfalls explizit auflösen und erhalten<sup>(i)</sup>

$$x_1 = \tilde{g}_1(x_{k+1}, \dots, x_n), \dots, x_k = \tilde{g}_k(x_{k+1}, \dots, x_n).$$

Setzen wir diese nun in  $f$  ein, so erhalten wir eine neue Funktion

$$\tilde{f}(x_{k+1}, \dots, x_n) := f(\tilde{g}_1(x_{k+1}, \dots, x_n), \dots, \tilde{g}_k(x_{k+1}, \dots, x_n), x_{k+1}, \dots, x_n).$$

Das Problem der Suche nach Extremwerten unter Nebenbedingungen reduziert sich so zur Suche nach "gewöhnlichen" Extremwerten der Funktion  $\tilde{f}$ . Dies ist die Situation, wie sie üblicherweise in der Schule behandelt wird.

---

<sup>(i)</sup>Nach Umbenennung der Variablen geschieht diese Auflösung nach den ersten  $k$  Variablen.

**Beispiel 2.6.20.** In unserem obigen Beispiel mit der Lichtquelle sei etwa  $\vec{a}^0 = \vec{0}$  und  $c \neq 0$  und  $\|\vec{n}\|^2 = n_x^2 + n_y^2 + n_z^2 = 1$ , sodass die Lichtquelle nicht in der Ebene liegt. Dann ist

$$f(x, y, z) = \frac{1}{x^2 + y^2 + z^2} \quad \text{und} \quad g(x, y, z) = xn_x + yn_y + zn_z - c = 0$$

Die Nebenbedingung lässt sich (falls  $n_z \neq 0$  ist) nach  $z$  auflösen mit

$$\tilde{g}(x, y) = \frac{1}{n_z}(c - n_x x - n_y y).$$

Das setzen wir in  $f$  ein und erhalten

$$\tilde{f}(x, y) = \frac{n_z^2}{n_z^2 x^2 + n_z^2 y^2 + (c - n_x x - n_y y)^2}$$

mit

$$\begin{aligned} \frac{\partial \tilde{f}}{\partial x} &= -2n_z^2 \frac{x(n_x^2 + n_z^2) + n_x n_y y - cn_x}{(n_z^2 x^2 + n_z^2 y^2 + (c - n_x x - n_y y)^2)^2}, \\ \frac{\partial \tilde{f}}{\partial y} &= -2n_z^2 \frac{y(n_y^2 + n_z^2) + n_x n_y x - cn_y}{(n_z^2 x^2 + n_z^2 y^2 + (c - n_x x - n_y y)^2)^2}. \end{aligned}$$

Das Verschwinden dieser Ausdrücke liefert ein lineares Gleichungssystem für  $(x, y)$  mit der Lösung

$$\begin{pmatrix} x \\ y \end{pmatrix} = c \begin{pmatrix} n_x \\ n_y \end{pmatrix},$$

was nach Einsetzen in die Nebenbedingung  $z = \tilde{g}(x, y) = cn_z$  liefert. Das mögliche Extremum wird also an der Stelle

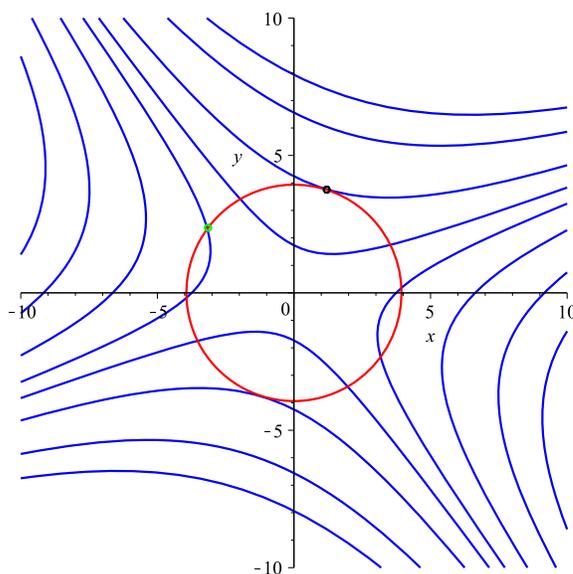
$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = c \cdot \vec{n}$$

angenommen. Das war auch zu erwarten: Das Extremum liegt in dem Punkt senkrecht unter der Lichtquelle.<sup>(i)</sup>

<sup>(i)</sup>Der Wert  $\frac{|c|}{\|\vec{n}\|}$  entspricht dem Abstand der Ebene zum Ursprung.

Wir betrachten die Situation für eine Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  und eine Nebenbedingung  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ .<sup>(ii)</sup> In Abb. 2.6.1 sind blau die Höhenmengen der Funktion  $f$  eingezeichnet und rot die Menge  $H_{g,0}$  auf der das Extremum gesucht wird. Wenn sich eine Höhenlinie von  $f$  und die Nebenbedingungsmenge

Abbildung 2.6.1: Die zweidimensionale Situation



$H_{g,0}$  schneiden, wie etwa im grün markierten Punkt, bedeutet das, dass die Funktion  $f$  entlang  $H_{f,0}$  dort steigt oder fällt. Der schwarz markierte Punkt, etwa  $(x_0, y_0)$ , hat die Eigenschaft, dass sich die Höhenmenge zu  $f$ , etwa  $H_{f,c}$ , und  $H_{g,0}$  nicht schneiden, sondern berühren. Bezeichnen wir die Auflösungen von  $f(x, y) = c$  und  $g(x, y) = 0$  mit  $s(x)$  und  $r(x)$ , dann gilt  $r'(x_0) = s'(x_0)$ . Wegen Satz 2.5.31 ist

$$r'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)} \quad \text{und} \quad s'(x_0) = -\frac{\frac{\partial g}{\partial x}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)}.$$

Mit  $\lambda := -\frac{\frac{\partial f}{\partial y}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)}$  ist somit  $\frac{\partial f}{\partial x}(x_0, y_0) + \lambda \frac{\partial g}{\partial x}(x_0, y_0) = 0$ . Es gilt aber

<sup>(ii)</sup>Ein typisches Beispiel ist etwa eine Temperaturverteilung  $f$  in der Ebene und ein Wanderweg  $H_{g,0}$ , auf dem der kühlfte oder wärmfste Ort gesucht wird. Die Höhenmengen der Funktion  $f$  entsprechen dann den Isothermen.

sowieso  $\frac{\partial f}{\partial y}(x_0, y_0) + \lambda \frac{\partial g}{\partial y}(x_0, y_0) = 0$ , sodass in dem Punkt, wo sich die Mengen berühren

$$\text{grad}f(x_0, y_0) + \lambda \text{grad}g(x_0, y_0) = 0.$$

Die Verallgemeinerung davon liefert nun den folgenden Satz:

**Satz 2.6.21.** *Es sei  $M \subset \mathbb{R}^n$  eine offene Menge,  $f : M \rightarrow \mathbb{R}$  eine Funktion und  $g : M \rightarrow \mathbb{R}^k$  eine Abbildung,  $k < n$ .  $f$  besitze in  $x^0$  ein lokales Extremum unter der Nebenbedingung  $g(x) = 0$  und  $Dg(x^0)$  habe den vollen Rang  $k$ . Dann gibt es einen Vektor  $\vec{\lambda}^0 \in \mathbb{R}^k$ , sodass*

$$\text{grad}f(x^0) + (\vec{\lambda}^0)^T Dg(x^0) = 0.$$

Sind  $g_1, \dots, g_k$  die Komponentenfunktionen von  $g$  und  $\lambda_1^0, \dots, \lambda_k^0$  die Komponenten von  $\vec{\lambda}^0$ , so schreibt sich das als

$$\text{grad}f(x^0) + \sum_{i=1}^k \lambda_i^0 \text{grad}g_i(x^0) = 0.$$

Die Werte  $\lambda_i^0$  heißen Lagrange-Multiplikatoren.

**Bemerkung 2.6.22.** 1. Um mögliche lokale Extremwerte zu bestimmen, suchen wir nach Lösungen des Gleichungssystems

$$\text{grad}f(x) + \sum_{i=1}^k \lambda_i \text{grad}g_i(x) = 0, \quad g_1(x) = \dots = g_k(x) = 0.$$

Dies sind  $n + k$  Gleichungen für die  $n + k$  Variablen  $x_1, \dots, x_n$ , und  $\lambda_1, \dots, \lambda_k$ .

Wir können sagen, dass wir die möglichen Extrema der Funktion

$$L : \mathbb{R}^k \times M \rightarrow \mathbb{R} \quad \text{mit} \quad L(\vec{\lambda}, x) := f(x) + \sum_{i=1}^k \lambda_i g_i(x)$$

suchen, also die Lösungen von  $\text{grad}L(\vec{\lambda}, x) = 0$ . Die Funktion  $L$  heißt *Lagrangefunktion* des Extremwertproblems.

2. Ist die Nebenbedingungs Menge kompakt so wissen wir, dass die Funktion  $f$  dort ihr Maximum und Minimum annimmt.

- Besitzt darüber hinaus die Nebenbedingungsmenge keinen Rand, so brauchen wir bei der Suche nach globalen Extrema lediglich unter den möglichen lokalen Extremwerten durch Einsetzen in  $f$  den größten und den kleinsten Wert aussortieren. Das ist insbesondere der Fall, wenn die kompakte Nebenbedingungsmenge als Nullstellenmenge einer stetigen Funktion gegeben ist, z.B.  $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1 \wedge x + y + z = 1\}$ .
- Hat die Nebenbedingung einen Rand, dann müssen wir bei der Suche nach globalen Extrema die Randpunkte gegebenenfalls gesondert untersuchen. Das ist insbesondere der Fall, wenn die kompakte Nebenbedingungsmenge mit Hilfe stetigen Funktion und Ungleichungen gegeben ist, z.B.  $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1 \wedge x + y + z \leq 1\}$ .

**Beispiel 2.6.23.** Wir gehen zurück zu unserem Ausleuchtungsproblem. Dort sind mit  $\vec{x} = (x, y, z)^T$  und  $\vec{n} = (n_x, n_y, n_z)^T$

$$f(\vec{x}) = \frac{1}{\|\vec{x}\|^2} \quad \text{und} \quad g(\vec{x}) = \langle \vec{n}, \vec{x} \rangle - c$$

mit

$$\text{grad}f(\vec{x}) = -\frac{2}{\|\vec{x}\|^4}\vec{x} \quad \text{und} \quad \text{grad}g(\vec{x}) = \vec{n}.$$

Die zu untersuchenden Gleichungen sind somit

$$-\frac{2}{\|\vec{x}\|^4}\vec{x} + \lambda\vec{n} = 0 \quad \text{und} \quad \langle \vec{x}, \vec{n} \rangle = c.$$

Die erste Gleichung liefert, dass die Lösung  $\vec{x}$  parallel zu  $\vec{n}$  sein muss, also  $\vec{x} = \alpha \cdot \vec{n}$ . Das in die zweite Gleichung eingesetzt gibt  $\alpha\|\vec{n}\|^2 = c$  oder  $\alpha = \frac{c}{\|\vec{n}\|^2}$ . Zurück einsetzen in die erste Gleichung liefert zusätzlich

$$\lambda = \frac{2\alpha}{\alpha^4\|\vec{n}\|^4} = \frac{2}{c^3}.$$

Somit ist die Lösung von  $\text{grad}L(\lambda, \vec{x}) = 0$  durch  $(\frac{2}{c^3}, c\vec{n})$  gegeben.

**Bemerkung 2.6.24.** 1. Sind die Extrema der Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  auf dem Abschluss  $\bar{U}$  einer offenen beschränkten Menge  $U \subset \mathbb{R}^n$  zu suchen, deren Rand durch eine implizite Funktion gemäß  $\partial U = \{x \in \mathbb{R}^n \mid g(x) = 0\}$  beschrieben wird, so kann man wie folgt vorgehen:

- Suche zunächst nach lokalen Extrema der Funktion  $f$  auf der offenen Menge  $U$ , gemäß Satz 2.6.13.
  - Suche nach weiteren Extrema auf dem Rand  $\partial U$  gemäß Satz 2.6.21.
2. Eine kleine Modifikation davon klappt auch, wenn die Nebenbedingungsmenge ein Graph mit Rand ist.

Gesucht seien die Extrema einer Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  wobei die Nebenbedingungsmenge in Graphenform als  $\{x = (y, z) \in \mathbb{R}^n \mid z = h(y), y \in \bar{U}, U \subset \mathbb{R}^\ell\}$  gegeben ist. Weiter sei der Rand  $\partial U$  von  $\bar{U} \subset \mathbb{R}^\ell$  durch eine Funktion gemäß  $\partial U = \{y \in \mathbb{R}^\ell \mid g(y) = 0\}$  gegeben. Dann können wir wie folgt vorgehen:

- Suche die Extrema  $x = (y, z)$  der Funktion  $f$  im (relativen) Inneren der Nebenbedingungsmenge. Das sind gerade die Punkte  $(y, h(y))$ , wobei  $y$  den Extrema von  $\tilde{f}(y) = f(y, h(y))$  auf der offenen Menge  $U \subset \mathbb{R}^\ell$  gemäß Satz 2.6.13 entspricht.
- Suche anschließend die Extrema von  $f$  auf dem Rand  $\{(y, h(y)) \mid g(y) = 0\}$  der Nebenbedingungsmenge. Die erhalten wir als Punkte  $(y, h(y))$ , wobei  $y$  ein Extremum von  $\tilde{f}$  unter der Nebenbedingung  $g(y) = 0$  ist.

Es folgt noch ein kleiner praktischer Tipp

**Bemerkung 2.6.25.** Sind die Extrema einer Funktion  $f$  gesucht und hat die Funktion eine komplizierte Form, so ist es gegebenenfalls angebracht die folgende

Eigenschaft zu benutzen: Ist  $g : \mathbb{R} \rightarrow \mathbb{R}$  monoton  $\begin{cases} \text{steigend} \\ \text{fallend} \end{cases}$  so gilt

$$f \text{ hat in } x_0 \text{ ein Maximum/Minimum} \\ \iff (g \circ f) \text{ hat in } x_0 \text{ ein } \begin{cases} \text{Maximum/Minimum} \\ \text{Minimum/Maximum} \end{cases}$$

Z.B. können wir statt der Funktion  $f(x) = \sqrt[n]{x_1^2 + \dots + x_n^2}$  die Funktion  $\hat{f} = f^n$ , also  $\hat{f}(x) = x_1^2 + \dots + x_n^2$ , betrachten, oder statt der Funktion  $f(x) = \ln(|g(x)|)$  die Funktion  $\hat{f} = e^f$ , also  $\hat{f}(x) = |g(x)|$ .

Ohne Beweis geben wir hier zum Abschluss ein hinreichendes Kriterium für Extrema unter Nebenbedingungen an. Dieses nutzt die Hessematrix der Lagrangefunktion  $L$ . Diese ist gegeben durch

$$HL(\vec{\lambda}, x) = \begin{pmatrix} 0 & Dg(x) \\ Dg(x)^T & Hf(x) + \sum_{i=1}^k \lambda_i Hg_i(x) \end{pmatrix} \in M_{k+n}\mathbb{R}.$$

Wir bezeichnen wieder mit  $HL^{[i]}$  die Untermatrizen von  $HL$  gemäß Definition 2.8.22, also insbesondere

$$HL^{[n+k]} = HL, \quad HL^{[i]} = 0 \text{ für } 1 \leq i \leq k$$

**Satz 2.6.26.** *Es sei  $M \subset \mathbb{R}^n$ ,  $f : M \rightarrow \mathbb{R}$  und  $g : M \rightarrow \mathbb{R}^k$ . Weiter sei  $x^0$  ein möglicher Extremwert von  $f$  unter der Nebenbedingung  $g(x) = 0$ , gemäß Satz 2.6.21. Dann gibt es also ein  $\vec{\lambda}^0$ , sodass  $(\vec{\lambda}^0, x^0)$  eine Lösung von  $\text{grad}L(\vec{\lambda}, x) = 0$  ist. Weiter sei  $Dg(x^0)$  derart, dass die ersten  $k$  Spalten von  $Dg(x^0)$  linear unabhängig sind.<sup>(i)</sup> Dann ist*

1.  $x^0$  ein lokales Minimum, wenn  $(-1)^k \det HL^{[k+i]} > 0$  für  $k+1 \leq i \leq n$ .
2.  $x^0$  ein lokales Maximum, wenn  $(-1)^i \det HL^{[k+i]} > 0$  für  $k+1 \leq i < n$ .

**Beispiel 2.6.27.** Hier schauen wir uns wieder das Beispiel der Ausleuchtung an, also  $f(\vec{x}) = \frac{1}{\|\vec{x}\|^2}$  und  $g(\vec{x}) = \langle \vec{n}, \vec{x} \rangle - c$ .

Die Lösung von  $\text{grad}L(\lambda, \vec{x}) = 0$  ist  $(\lambda_0, \vec{x}^0) = (\frac{2}{c^3}, c \cdot \vec{n})$ . Da  $g$  eine lineare Bedingung ist, ist die Hessematrix von  $L$  unabhängig von  $\lambda$ . Sie ist gegeben durch

$$HL(\lambda, \vec{x}) = \begin{pmatrix} 0 & \vec{n}^T \\ \vec{n} & Hf(\vec{x}) \end{pmatrix}$$

mit  $Hf(\vec{x}) = \frac{2}{\|\vec{x}\|^6} (4\vec{x} \cdot \vec{x}^T - \|\vec{x}\|^2 \mathbb{1})$ . An der Stelle  $(\lambda_0, \vec{x}^0)$  ist das

$$HL(\lambda_0, \vec{x}^0) = \begin{pmatrix} 0 & n_x & n_y & n_z \\ n_x & \frac{2(4n_x^2-1)}{c^6} & \frac{8n_x n_y}{c^6} & \frac{8n_x n_z}{c^6} \\ n_y & \frac{8n_x n_y}{c^6} & \frac{2(4n_y^2-1)}{c^6} & \frac{8n_y n_z}{c^6} \\ n_z & \frac{8n_x n_z}{c^6} & \frac{8n_y n_z}{c^6} & \frac{2(4n_z^2-1)}{c^6} \end{pmatrix}$$

<sup>(i)</sup>Unter der Voraussetzung, dass  $Dg(x^0)$  vollen Rang  $k$  hat, ist das keine Einschränkung, da wir das nach Umbenennung der Variablen stets erreichen können.

Die Bedingung an  $Dg(\vec{x}^0)$  ist hier bei  $n_x \neq 0$  erfüllt, und zur Untersuchung der Art des Extremwertes benötigen wir  $\det HL^{[k+i]}$  mit  $k = 1$  und  $i = 2, 3$ , also  $\det HL^{[3]}$ ,  $\det HL^{[4]}$ . Es ist

$$(-1)^3 \det HL^{[4]} = -\det HL = \frac{4}{c^{12}} > 0$$

$$(-1)^2 \det HL^{[3]} = \det \begin{pmatrix} 0 & n_x & n_y \\ n_x & \frac{2(4n_x^2-1)}{c^6} & \frac{8n_x n_y}{c^6} \\ n_y & \frac{8n_x n_y}{c^6} & \frac{2(4n_y^2-1)}{c^6} \end{pmatrix} = \frac{2(n_x^2 + n_y^2)}{c^6} > 0$$

Somit handelt es sich bei unserem Extremum um ein Maximum.

**Beispiel 2.6.28.** Es soll die Funktion  $f(x, y, z) = x^2 + y^2 + z^2$  unter den Nebenbedingungen  $g_1(x, y, z) = x^2 + 2y^2 + 2z^2 - 1 = 0$  und  $g_2(x, y, z) = x + 2y + 3z = 0$  auf Extremstellen untersucht werden. Das zu untersuchende Gleichungssystem ergibt sich aus dem Gradienten der Lagrangefunktion  $L = f + \lambda_1 g_1 + \lambda_2 g_2$  zu

$$\begin{aligned} 2(1 + \lambda_1)x + \lambda_2 &= 0 \\ 2(1 + 2\lambda_1)y + 2\lambda_2 &= 0 \\ 2(1 + 2\lambda_1)z + 3\lambda_2 &= 0 \\ x + 2y + 3z &= 0 \\ x^2 + 2y^2 + 2z^2 - 1 &= 0 \end{aligned}$$

Wir nummerieren die Gleichungen mit 1-5 durch und untersuchen mehrere Fälle

- $\lambda_1 \neq -1, -\frac{1}{2}$ , sodass 1-3  $x = \frac{-\lambda_2}{2(1+\lambda_1)}, y = \frac{-2\lambda_2}{2(1+2\lambda_1)}, z = \frac{-3\lambda_2}{2(1+2\lambda_1)}$  liefert. Setzt man das in 4 ein so gibt das  $\lambda_2(14 + 15\lambda_1) = 0$ , also  $\lambda_2 = 0$  oder  $\lambda_1 = -\frac{14}{15}$ .  
 $\lambda_2 = 0$  liefert dann  $x = y = z = 0$  im Widerspruch zu Gleichung 5.  
 $\lambda_1 = -\frac{14}{15}$  liefert aus 1-3 die Lösung  $x = -\frac{15}{2}\lambda_2, y = \frac{15}{13}\lambda_2, z = \frac{45}{26}\lambda_2$  von 1-4. In 5 eingesetzt liefert das  $\lambda_2 = \pm \frac{2}{15}\sqrt{\frac{13}{15}}$  und somit  $x = \mp \sqrt{\frac{13}{15}}, y = \pm \frac{2}{13}\sqrt{\frac{13}{15}}, z = \pm \frac{3}{13}\sqrt{\frac{13}{15}}$ .

- $\lambda = -1$  liefert in 1  $\lambda_2 = 0$  und damit in 2-3  $x = y = 0$ . Mit  $g_1 = 0$  liefert das auch  $x = 0$ , sodass dieser Fall entfällt.
- $\lambda = -\frac{1}{2}$  liefert in 2-3  $\lambda_2 = 0$  und dann in 1  $x = 0$ . In 4 eingesetzt haben wir also  $y = -\frac{3}{2}z$  und damit aus 5  $z = \pm\sqrt{\frac{2}{13}}$  und zurück in 4  $y = \mp\frac{3}{2}\sqrt{\frac{2}{13}}$ . Der Wert zu

Die potentiellen Extremwerte sind damit

$$(x, y, z) = \left( \mp\sqrt{\frac{13}{15}}, \pm\frac{2}{13}\sqrt{\frac{13}{15}}, \pm\frac{3}{13}\sqrt{\frac{13}{15}} \right) \quad \text{zu} \quad (\lambda_1, \lambda_2) = \left( -\frac{14}{15}, \pm\frac{2}{15}\sqrt{\frac{13}{15}} \right)$$

$$(x, y, z) = \left( 0, \mp\frac{3}{2}\sqrt{\frac{2}{13}}, \pm\sqrt{\frac{2}{13}} \right) \quad \text{zu} \quad (\lambda_1, \lambda_2) = \left( -\frac{1}{2}, 0 \right)$$

Zur weiteren Untersuchung der Extremstellen gehen wir auf zwei Arten vor:

1. (Mit Hilfe der Hessematrix der Lagrangefunktion)

Die Hessematrix der Lagrangefunktion ist

$$HL = \begin{pmatrix} 0 & 0 & 2x & 4y & 4z \\ 0 & 0 & 1 & 2 & 3 \\ 2x & 1 & 2(1 + \lambda_1) & 0 & 0 \\ 4y & 2 & 0 & 2(1 + 2\lambda_1) & 0 \\ 4z & 3 & 0 & 0 & 2(1 + 2\lambda_1) \end{pmatrix}$$

Wegen  $n = 3$  und  $k = 2$  haben wir hier nur  $\det HL^{[5]} = \det HL$  zu untersuchen. Für die ersten beiden Lösungen ist  $(-)^3 \det HL > 0$  und für die zweiten beiden ist  $(-1)^2 \det HL > 0$ , sodass die ersten beiden Maxima und die zweiten Minima liefern.

2. (Mit Hilfe eines Kompaktheitsarguments)

Als Schnitt eines Ellipsoiden mit einer Ebene ist die Nebenbedingungs-  
menge kompakt und besitzt, da in Gleichungsform gegeben, keinen Rand.  
Daher brauchen wir lediglich die Funktionswerte der möglichen Extrem-  
werte berechnen und erhalten

$$f \left( \mp\sqrt{\frac{13}{15}}, \pm\frac{2}{13}\sqrt{\frac{13}{15}}, \pm\frac{3}{13}\sqrt{\frac{13}{15}} \right) = \frac{13}{15} + \frac{4}{13} \frac{1}{15} + \frac{9}{13} \frac{13}{15} = \frac{14}{15}$$

$$f\left(0, \mp \frac{3}{2}\sqrt{\frac{2}{13}}, \pm\sqrt{\frac{2}{13}}\right) = \frac{9}{2} \frac{1}{13} + \frac{2}{13} = \frac{1}{2}.$$

Damit liefern auch bei dieser Untersuchung die ersten beiden Punkte die Maxima und die zweiten beiden die Minima.

## 2.7 Grundlagen gewöhnlicher Differentialgleichungen

### 2.7.1 Problemstellung und elementare Beispiele

#### 2.7.1.1 Problemstellung

Wir haben im vorigen Kapitel Gleichungen der Form  $F(x, y) = 0$  untersucht und diskutiert, ob und unter welchen Bedingungen es eine Funktion  $g : x \mapsto g(x)$  gibt, sodass mit  $y = g(x)$  die Gleichung  $F(x, g(x)) = 0$  erfüllt ist.

Allgemein sprechen wir von einer Differentialgleichung, wenn eine Relation oder ein Problem gegeben ist, deren Lösung eine Funktion ist, und die Formulierung des Problems neben der Funktion selbst auch deren Ableitungen enthält.

Das einfachste Beispiel ist die Suche nach einer Stammfunktion: Hier ist eine Funktion  $f : x \mapsto f(x)$  gegeben und gesucht ist eine Funktion  $g : x \mapsto g(x)$  mit  $g'(x) = f(x)$ . Für so eine gesuchte Funktion schreiben wir in diesem Abschnitt in der Regel  $y : x \mapsto y(x)$ , sodass sich das obige Problem wie folgt formulieren lässt:

Gegeben eine Funktion  $f$ ; gesucht ist eine Funktion  $y$  mit  $y' - f(x) = 0$ .

Genauso können wir natürlich allgemeinere Probleme mit höheren Ableitungen formulieren:

**Beispiel 2.7.1.** 1. Die Gleichung  $y'' - 2c = 0$  hat die Lösung  $y(x) = cx^2 + bx + a$ .

2. Die Gleichung  $y' - y = 0$  hat die Lösung  $y(x) = ce^x$ .

3. Die Gleichung  $y'' + y = 0$  hat die Lösung  $y(x) = \cos(x)$ , aber auch  $y(x) = \sin(x)$ .

4. die Gleichung  $(y^{(4)})^2 + y^2 = 2yy^{(4)}$  hat die Lösung  $y = e^x$ , aber auch  $y(x) = e^{-x}$  oder  $y(x) = \sin(x)$  oder  $y(x) = \cos(x)$ .
5. Die Gleichung  $y' - y^2 - 1 = 0$  hat die Lösung  $y(x) = \tan(x)$ .
6. Die Gleichung  $(y')^2 - 4y = 0$  hat die Lösung  $y(x) = (x + c)^2$ , aber auch  $y(x) = 0$ .
7. Die Gleichung  $x^2(y')^2 + y^2 - (y')^2 = 0$  hat die Lösung  $y(x) = e^{\arcsin(x)}$ , aber auch  $y(x) = 0$ .

**Beispiel 2.7.2.** Die Kreise mit Radius  $r$  und Mittelpunkt auf der  $x$ -Achse sind von der Form  $(x - a)^2 + y^2 - r^2 = 0$ . Nehmen wir  $y$  als Funktion von  $x$  an und differenzieren die Kreisgleichung so erhalten wir  $(x - a) + yy' = 0$  oder  $(x - a)^2 = y^2 y'^2$ . Einsetzen in die Ausgangsgleichung eliminiert den Parameter  $a$  und liefert die DGL

$$y^2(y')^2 + y^2 - r^2 = 0.$$

Lösungen dieser DGL sind die den Kreis beschreibenden Funktionen  $y(x) = \pm\sqrt{r^2 - (x - a)^2}$  und zusätzlich noch  $y(x) = \pm R$ .

Diese ersten Beispiele motivieren die folgende Definition

**Definition 2.7.3.** Es sei  $D \subset \mathbb{R}^{k+2}$  und  $F : D \rightarrow \mathbb{R}$  stetig.<sup>(i)</sup>

1. Die Gleichung

$$F(x, y, y', y'', \dots, y^{(k)}) = 0$$

heißt *gewöhnliche Differentialgleichung (DGL)*. Als *Ordnung* der DGL bezeichnen wir die größte Zahl  $k$  derart, dass  $y^{(k)}$  nicht-trivial in der impliziten Funktion  $F$  auftritt. Ist die Gleichung in der Form

$$y^{(k)} = f(x, y, y', \dots, y^{(k-1)})$$

gegeben, so sprechen wir von einer *expliziten DGL*.

---

<sup>(i)</sup>Diese Bedingung können wir abschwächen. Hier und im Folgenden sind jedoch alle betrachteten Funktionen mindestens stetig, es sei denn, es wird explizit etwas anderes vorausgesetzt.

2. Unter einer Lösung der DGL  $F(x, y, y', \dots, y^{(k)}) = 0$  verstehen wir eine Funktion  $y : x \mapsto y(x)$ , die auf einem Intervall  $I \subset \mathbb{R}$  definiert ist, und für die Folgendes gilt:

- $y$  ist  $k$ -mal differenzierbar,
- $(x, y(x), y'(x), \dots, y^{(k)}(x)) \in D$  für alle  $x \in I$  und
- $F(x, y(x), y'(x), \dots, y^{(k)}(x)) = 0$  für alle  $x \in I$ .

**Beispiel 2.7.4.** Die DGLn aus den vorigen Beispielen 1., 2., 3. und 5. sind explizit und haben die Ordnungen 2, 1, 2 und 1, die DGLn aus den Beispielen 4., 6. und 7. sind von den Ordnungen 4, 1 und 1 und sind nicht explizit.

**Definition/Bemerkung 2.7.5.** Die Lösungen, die wir in den Anfangsbeispielen erhalten haben hängen in der Regel von freien Parametern ab. Diese ergeben sich typischerweise bei der Integration. Suchen wir spezielle Lösungen, so müssen wir diese Parameter fixieren. Das kann man tun, indem man angibt, welche Werte die gesuchte Funktion an vorgegebenen Stellen annehmen soll. Diese Vorgaben  $(x_i, y(x_i))$  nennt man *Anfangswerte*. Da bei einer DGL  $k$ -ter Ordnung im Wesentlichen  $k$  Integrationen nötig sind, werden zur Festlegung der  $k$  freien Parameter  $k$  solcher Anfangswerte vorgegeben.

Die Suche nach einer Lösung  $y : x \mapsto y(x)$  der Gleichungen

$$\begin{aligned} F(x, y(x), y'(x), \dots, y^{(k)}(x)) &= 0, \\ y(x_0) = \eta_0, \quad y'(x_0) = \eta_1, \quad \dots, \quad y^{(k-1)}(x_0) &= \eta_{k-1} \end{aligned}$$

nennt man *Anfangswertproblem (AWP)*.

Wir verallgemeinern den Begriff aus der Definition 2.7.3 etwas:

**Definition 2.7.6.** Ein *DGL-System vom Rang  $n$  und Ordnung  $k$*  ist gegeben durch  $n$  Funktionen  $F_i : D \rightarrow \mathbb{R}$ ,  $1 \leq i \leq n$ , mit  $D \subset \mathbb{R}^{n(k+1)+1}$  und ist von der Form

$$\begin{aligned} F_1(x, y_1, \dots, y_n, y_1', \dots, y_n', \dots, y_1^{(k)}, \dots, y_n^{(k)}) &= 0 \\ &\vdots \\ F_n(x, y_1, \dots, y_n, y_1', \dots, y_n', \dots, y_1^{(k)}, \dots, y_n^{(k)}) &= 0. \end{aligned}$$

Nutzen wir die Vektorschreibweise  $F = (F_1, \dots, F_n)^T$ ,  $y = (y_1, \dots, y_n)^T, \dots$ ,  $y^{(k)} = (y_1^{(k)}, \dots, y_n^{(k)})^T$ , so schreibt sich das System kurz als

$$F(x, y, \dots, y^{(k)}) = 0$$

und hat formal die gleiche Form wie im Fall  $n = 1$ . Auch hier heißt das DGL-System *explizit*, wenn es sich in der aufgelösten Form

$$\begin{aligned} y_1^{(k)} &= f_1(x, y_1, \dots, y_n, \dots, y_1^{(k-1)}, \dots, y_n^{(k-1)}) \\ &\vdots \\ y_n^{(k)} &= f_n(x, y_1, \dots, y_n, \dots, y_1^{(k-1)}, \dots, y_n^{(k-1)}) \end{aligned}$$

darstellen lässt. Als Lösung des DGL-Systems bezeichnen wir hier eine Menge von Funktionen  $y_i : I \rightarrow \mathbb{R}$ ,  $1 \leq i \leq n$ , die auf einem gemeinsamen Intervall  $I$  definiert sind, und für die gilt

- $y_i$  ist  $k$ -mal differenzierbar für alle  $1 \leq i \leq n$ ,
- $(x, y_1(x), \dots, y_n(x), y_1'(x), \dots, y_n'(x), \dots, y_1^{(k)}(x), \dots, y_n^{(k)}(x)) \in D$  für alle  $x \in I$  und
- $F_i(x, y_1(x), \dots, y_n(x), y_1'(x), \dots, y_n'(x), \dots, y_1^{(k)}(x), \dots, y_n^{(k)}(x)) = 0$  für alle  $x \in I$  und  $1 \leq i \leq n$ .

**Beispiel 2.7.7.** 1. Ist  $A = (a_{ij}) \in M_n \mathbb{R}$  eine quadratische Matrix, so ist für vorgegebene Funktionen  $f_1, \dots, f_n$

$$\begin{aligned} a_{11} y_1' + a_{12} y_2' + \dots + a_{1n} y_n' + f_1(x, y_1, \dots, y_n) &= 0 \\ a_{21} y_1' + a_{22} y_2' + \dots + a_{2n} y_n' + f_2(x, y_1, \dots, y_n) &= 0 \\ &\vdots \\ a_{n1} y_1' + a_{n2} y_2' + \dots + a_{nn} y_n' + f_n(x, y_1, \dots, y_n) &= 0 \end{aligned}$$

ein DGL-System vom Rang  $n$  und erster Ordnung. Ist  $A$  invertierbar mit inverser Matrix  $B = (b_{ij}) \in M_n \mathbb{R}$ , so lässt sich das System äquivalent als explizites DGL-System schreiben:

$$\begin{aligned} y_1' &= b_{11} f_1(x, y_1, \dots, y_n) + \dots + b_{1n} f_n(x, y_1, \dots, y_n) \\ y_2' &= b_{21} f_1(x, y_1, \dots, y_n) + \dots + b_{2n} f_n(x, y_1, \dots, y_n) \end{aligned}$$

$$\begin{aligned} & \vdots \\ y'_n &= b_{n1} f_1(x, y_1, \dots, y_n) + \dots + b_{nn} f_n(x, y_1, \dots, y_n) \end{aligned}$$

Das wird insbesondere klar, wenn wir die Vektornotation benutzen: das erste System ist dann gerade  $Ay' + f = 0$  oder eben  $y' = Bf$ .

2. Wir betrachten den Fall einer sich schnell ausbreitenden Zombieepidemie, die sich durch Bisse ausbreitet. Es bezeichne

- $y_1$  die Zahl der Menschen, die infiziert werden können,
- $y_2$  die Zahl der Zombies,
- $y_3$  die Zahl derer, die immun sind gegen das Zombievirus oder bereits gestorben sind,

und

- $a$  den Infektionskoeffizienten,
- $b$  den Gesundungs- und Sterbekoeffizienten.

Wir nehmen an, dass während der gesamten Zeit der Ausbreitung, die Gesamtbevölkerung konstant bleibt, dass gesundene Personen immun sind, und, dass tote Zombies nicht mehr infektiös sind. Außerdem gebe es einen Impfstoff gegen das Zombievirus und  $\rho(x)$  sei die Zahl der zur Zeit  $x$  geimpften Personen. Dann erfüllen  $(y_1, y_2, y_3)$  das folgende explizite DGL-System

$$\begin{aligned} y'_1 &= -ay_1y_2 - \rho \\ y'_2 &= ay_1y_2 - by_2 \\ y'_3 &= by_2 + \rho \end{aligned}$$

Wegen  $y_1 + y_2 + y_3 = \text{const}$  ist insbesondere  $y'_1 + y'_2 + y'_3 = 0$ .

3. Es sei  $y^{(k)} = f(x, y, y', \dots, y^{(k-1)})$  eine explizite DGL  $k$ -ter Ordnung. Setzen wir nun

$$y_1 := y, \quad y_2 := y', \quad \dots, \quad y_k := y^{(k-1)}.$$

so lässt sich die obige DGL als DGL-System vom Rang  $k$  und Ordnung 1 schreiben:

$$\begin{aligned}y_1' &= y_2 \\y_2' &= y_3 \\&\vdots \\y_{k-1}' &= y_k \\y_k' &= f(x, y_1, \dots, y_k)\end{aligned}$$

4. Analog können wir ein explizites DGL-System vom Rang  $n$  und Ordnung  $k$  in ein explizites DGL-System vom Rang  $nk$  und Ordnung 1 umformen.

Fragen, die sich bisher natürlich ergeben, sind zum Beispiel

- Wie finden wir eine Lösung?
- Wie viele Lösungen gibt es?
- Gibt es Bedingungen, sodass eindeutige Lösungen existieren?
- Können wir aus vorgegebenen Lösungen neue konstruieren?

Wir wollen uns zunächst der ersten Frage widmen und einige Beispielklassen betrachten.

### 2.7.1.2 Die separierbare DGL

**Definition/Satz 2.7.8.** *Ein DGL erster Ordnung heißt separierbar, wenn sie sich in der Form*

$$h(y)y' = g(x)$$

*schreiben lässt. Diese hat die implizite Lösung*

$$\int h(y)dy - \int g(x)dx = 0.$$

**Beispiel 2.7.9.** Die DGL  $xyy' + (1 + y^2) = 0$  lässt sich umschreiben zu  $\frac{y}{1 + y^2}y' = -\frac{1}{x}$ . Sie hat damit die implizite Lösung  $\int \frac{2y dy}{1 + y^2} = -\int \frac{2}{x} dx$  oder  $\ln(1 + y^2) = -\ln(x^2) + c$  oder  $y = \pm\sqrt{\frac{a^2}{x^2} - 1}$  mit  $a^2 = e^c > 0$ . Sie ist auf den Intervallen  $]0, a[$  und  $] -a, 0[$  definiert.

**Beispiel 2.7.10** (Homogene DGL). Eine DGL vom Typ

$$y' = f\left(\frac{y}{x}\right)$$

heißt *homogene DGL*. Durch die Einführung der neuen Funktion  $u$  mit  $y = xu$  also  $y' = u + xu'$  wird die Ausgangs-DGL zu  $u + xu' = f(u)$  oder  $xu' = f(u) - u$ , also zu einer separierbaren DGL für die Funktion  $u$ .

$$y' = f\left(\frac{y}{x}\right) \quad \begin{array}{c} y = xu \\ \longleftarrow \longrightarrow \end{array} \quad xu' = f(u) - u.$$

Als Beispiel betrachten wir die DGL  $(x + y) - (x - y)y' = 0$  diese lässt sich für  $y \neq x$  umschreiben zu  $y' = \frac{x+y}{x-y}$  oder, für  $x \neq 0$ , zu  $y' = \frac{1+\frac{y}{x}}{1-\frac{y}{x}}$ . Damit erhalten wir für die Funktion  $u$  mit  $y = xu$  die DGL  $xu' = \frac{1+u}{1-u} - u = -\frac{1+u^2}{u-1}$  oder  $\frac{u-1}{1+u^2}u' = -\frac{1}{x}$ . Damit erhalten, z.B. durch Partialbruchzerlegung wir die implizite Lösung  $\frac{1}{2}\ln(1 + u^2) - \arctan(u) = -\ln|x| + c$ . Für die gesuchte Funktion  $y$  liefert das  $\frac{1}{2}\ln(x^2 + y^2) - \arctan\left(\frac{y}{x}\right) = c$ .

**Beispiel 2.7.11.** Es seien  $a, b, c, \bar{a}, \bar{b}, \bar{c} \in \mathbb{R}$  und die DGL

$$y' = f\left(\frac{ax + by + c}{\bar{a}x + \bar{b}y + \bar{c}}\right)$$

gegeben.

- Ist  $\bar{a}\bar{b} - \bar{a}b = 0$ , dann führt Substitution  $u = ax + by + c$  auf eine separierbare DGL.  
In diesem Fall ist  $y' = \frac{1}{\bar{b}}(u' - a)$  und  $\frac{ax+by+c}{\bar{a}x+\bar{b}y+\bar{c}} = \frac{bu}{bu+\bar{c}b-\bar{c}b}$ , sodass die Ausgangs-DGL zu  $u' = bf\left(\frac{bu}{bu+\bar{c}b-\bar{c}b}\right) - a$  wird.
- ist  $\bar{a}\bar{b} - \bar{a}b \neq 0$ , dann haben die Geraden  $\{(x, y) \mid ax + by + c = 0\}$  und  $\{(x, y) \mid \bar{a}x + \bar{b}y + \bar{c} = 0\}$  einen Schnittpunkt  $(p, q)$ . In diesem Fall führt

die Substitution  $t = x - p$  und  $u(t) = y(t + p) - q = y(x) - q$  auf eine homogene DGL.

Es ist  $u'(t) = y'(t + p) = y'(x)$  und  $\frac{ax+by(x)+c}{\bar{a}x+\bar{b}y(x)+\bar{c}} = \frac{at+ap+bu(t)+bq+c}{\bar{a}t+\bar{a}p+\bar{b}u(t)+\bar{b}q+\bar{c}} = \frac{at+bu}{\bar{a}t+\bar{b}u} = \frac{a+b\frac{u}{t}}{\bar{a}+\bar{b}\frac{u}{t}}$ . Damit liefert die Ausgangs-DGL für die Funktion  $u(t)$  die

homogene Gleichung  $u' = f\left(\frac{a + b\frac{u}{t}}{\bar{a} + \bar{b}\frac{u}{t}}\right)$ .

**Beispiel 2.7.12** (Radioaktiver Zerfall). Die Zerfallsrate eines radioaktiven Materials zu einem Zeitpunkt  $t$  ist proportional zu dem zu dieser Zeit vorhandenen Menge des Materials. Ist also  $y(t)$  die Menge des Materials, so erfüllt die Funktion  $y : t \rightarrow y(t)$  die separierbare DGL

$$y' = -\gamma y,$$

wobei  $\gamma > 0$  eine vom Material abhängige Konstante ist. Dies hat die Lösung  $y(t) = Ce^{-\gamma t}$ . Ist etwa zur Zeit  $t_0$  die Materialmenge  $y(t_0) = y_0$ , so ist  $C = y_0 e^{\gamma t_0}$ , sodass die Lösung die Form  $y(x) = y_0 e^{-\gamma(t-t_0)}$  bekommt.

Als Halbwertszeit  $T_H$  bezeichnet man die Zeit, die benötigt wird, um durch Zerfall die Menge des Stoffs zu halbieren:  $y(t + T_H) = \frac{1}{2}y(t)$  oder  $e^{-\gamma T_H}y(t) = e^{-\ln(2)}y(t)$  oder  $\gamma T_H = \ln(2)$ . Wählen wir nun diese Materialgröße als Parameter, so lautet die Lösung

$$y(t) = y_0 e^{-\ln(2)\frac{t-t_0}{T_H}}.$$

### 2.7.1.3 Verlauf von Lösungen am Beispiel der logistischen DGL

Die Funktion  $y : t \mapsto y(t)$  beschreibe die Anzahl einer Population zur Zeit  $t$ . Weiter seien die Geburten- und Sterberaten durch  $B : t \mapsto B(t)$  und  $D : t \mapsto D(t)$  gegeben. Dann entwickelt sich eine ideale Population gemäß der DGL

$$y' = (B - D)y.$$

Ist die Differenz aus Geburten- und Sterberate konstant, so ergibt sich analog zu Beispiel 2.7.12 ein exponentielles Populationswachstum oder ein exponentielles Populationssterben, je nachdem, ob  $B - D > 0$  oder  $B - D < 0$ .

Ziehen wir Konflikte innerhalb der Population mit in die Diskussion ein und nehmen an, dass  $R = B - D$  konstant ist, so erhalten wir die *logistische DGL*

$$y' = Ry \left(1 - \frac{y}{K}\right).$$

Hierbei nehmen wir an, dass die Rate der Konflikte zu einem "nicht-gutem" Ende für die Population führt und proportional zum Quadrat  $y^2$  der Populationszahl ist.

**Bemerkung 2.7.13** (Logistische DGL). Die logistische DGL  $y' = Ry(1 - \frac{y}{K})$  mit  $R, K > 0$  ist separierbar mit impliziter Lösung

$$\left| \frac{y - K}{y} \right| = C e^{-Rt}.$$

0. Die konstante Funktion  $y = \underline{K}$  ist eine Lösung der logistischen DGL.
1. Ist  $y > K$ , so eine Lösung durch  $y(t) = \frac{K}{1 - C e^{-Rt}}$  gegeben.
2. Ist  $y < K$ , so ist eine Lösung durch  $y = \frac{K}{1 + C e^{Rt}}$  gegeben.

Ist  $y(t_0) = y_0 > K$  als Anfangswert vorgegeben, so sind wir im ersten Fall und es ist  $C = \frac{1}{y_0}(y_0 - K)e^{Rt_0}$ . Ist  $y(t_0) = y_0 < K$  als Anfangswert vorgegeben, so sind wir im zweiten Fall und es ist  $C = -\frac{1}{y_0}(y_0 - K)e^{Rt_0}$ . Nehmen wir noch den nullten Fall dazu so ergibt sich als eine Lösung des Anfangswertproblems

$$\begin{aligned} y(t) &= K && \text{falls } y_0 = K \\ y(t) &= \frac{y_0 K}{y_0 - (y_0 - K)e^{-R(t-t_0)}} && \text{falls } y_0 \neq K \end{aligned}$$

Für die nicht-konstante Lösung gilt  $\lim_{t \rightarrow \infty} y(t) = K$ . Später in Satz 2.7.23 werden wir zeigen, dass die Lösungen der Anfangswertprobleme eindeutig sind. Die Lösungen der sind in Abb. 2.7.1 skizziert.

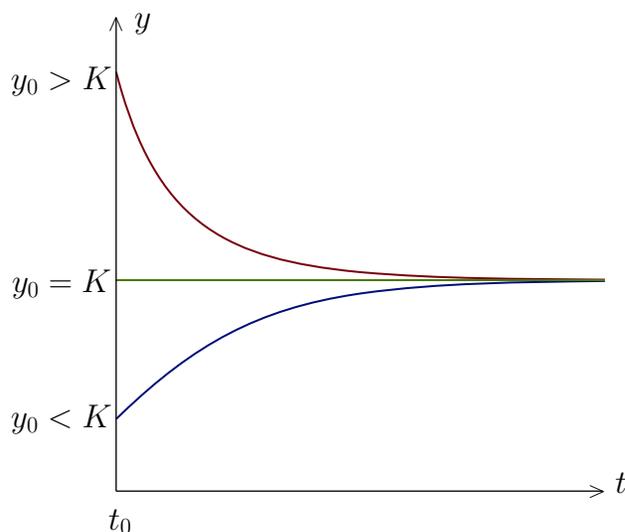
#### 2.7.1.4 Die exakte DGL

**Definition 2.7.14.** Eine DGL der Form

$$f(x, y) + g(x, y)y' = 0$$

heißt

Abbildung 2.7.1: Der Verlauf der Lösungen der logistischen DGL



1. *exakt*, wenn es eine Funktion  $h(x, y)$  gibt,  $\frac{\partial h}{\partial x} = f$  und  $\frac{\partial h}{\partial y} = g$ .
2. *exakt mit integrierendem Faktor*  $\mu(x, y)$ , wenn die DGL  $\mu(x, y)f(x, y) + \mu(x, y)g(x, y)y' = 0$  exakt ist.

**Satz 2.7.15.** Sind  $f$  und  $g$  stetig differenzierbar, so ist die DGL  $f(x, y) + g(x, y)y' = 0$  genau dann exakt, wenn  $\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x}$ .

**Beweisskizze.** Die Richtung " $\implies$ " folgt aus dem Lemma von Schwarz. Die Rückrichtung " $\impliedby$ " werden wir im dritten Teil der Vorlesung beweisen.  $\square$

**Bemerkung 2.7.16.** 1. Lösungen der exakten DGL  $f(x, y) + g(x, y)y' = 0$  sind implizit durch die Gleichung  $h(x, y) = c$  gegeben.

2. Wird eine Gleichung durch einen integrierenden Faktor exakt, so kann es vorkommen, dass die original DGL mehr oder auch weniger Lösungen hat als die dann exakte DGL. Zum Beispiel kommen die Lösungen  $y : x \mapsto y(x)$  mit  $m(x, y) = 0$  hinzu.
3. Einen integrierenden Faktor zu finden ist in der Regel schwer. In speziellen Fällen lässt sich mit von Satz 2.7.15 eine gewöhnliche DGL für  $\mu$  angeben, mit deren Hilfe man  $\mu$  bestimmen kann.

## 2.7.1.5 Die lineare DGL erster Ordnung

**Definition 2.7.17.** Als lineare DGL erster Ordnung bezeichnet man eine DGL der Form

$$y' + p(x)y = q(x)$$

mit zwei Funktionen  $p, q$ , die nur von  $x$  abhängen.

**Bemerkung 2.7.18.** 1. Die lineare DGL erster Ordnung ist exakt mit integrierendem Faktor

$$\mu(x) = e^{\int p(x)dx}.$$

2. Es sei  $\mu$  wie oben. Dann sind die Funktionen

$$y(x) = \frac{1}{\mu(x)} \int \mu(x)q(x) dt$$

Lösungen der linearen DGL  $y' + py = q$ . Die Integrale sind hier als unbestimmte Integrale zu verstehen.

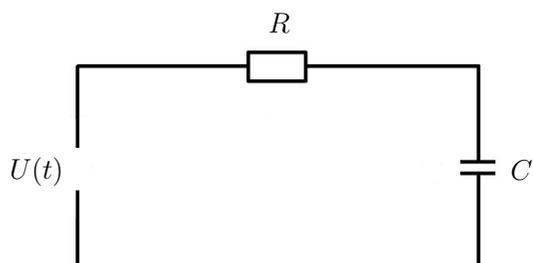
3. Die lineare DGL erster Ordnung zusammen mit dem Anfangswert  $y(x_0) = y_0$  hat die Lösung

$$y(x) = e^{-\int_{x_0}^x p(t)dt} \left( y_0 + \int_{x_0}^x \mu(t)q(t) dt \right).$$

**Beispiel 2.7.19** (Elektrische Schaltkreise). Wir betrachten eine Reihenschaltung aus Widerstand und Kondensator, die von einer externen Spannungsquelle versorgt wird. Die Spannung habe den zeitlichen Verlauf  $U(t)$  und der Widerstand und die Kapazität des Kondensators seien konstante Größen  $R$  und  $C$ , siehe Abb. 2.7.2 Weiter bezeichne  $I(t)$  den Strom im Stromkreis und  $Q(t)$  die Ladungsmenge des Kondensators zur Zeit  $t$ . Dann gilt für die Spannungen  $U_R$  und  $U_C$  am Widerstand und am Kondensator  $U_C = \frac{Q}{C}$  und  $U_R = RI$ . Die Stromstärke im Stromkreis entspricht genau der Änderung der Ladung des Kondensators, also  $Q'(t) = I(t)$ . Weiter wissen wir, dass sich die Gesamtspannung der Komponenten in der Reihenschaltung zur angelegten Spannung addieren,  $U = U_R + U_C$ . Somit erhalten wir eine lineare DGL für die Ladungsmenge  $Q$

$$RQ' + \frac{1}{C}Q = U(t)$$

Abbildung 2.7.2: Die RC-Schaltung



mit der Lösung

$$Q(t) = \left( Q(0) + \frac{1}{R} \int_0^t U(v) e^{\frac{v}{RC}} dv \right) e^{-\frac{t}{RC}}.$$

Im Fall einer konstanten Spannungsquelle und einem nicht-geladenen Kondensator  $Q(0) = 0$  liefert das

$$Q(t) = U_0 C (1 - e^{-\frac{t}{RC}})$$

und damit

$$I(t) = Q'(t) = \frac{U_0}{R} e^{-\frac{t}{RC}}.$$

Für  $t \rightarrow \infty$  nimmt der Ladestrom immer weiter ab und die Ladungsmenge des Kondensators nähert sich dem Maximalwert  $Q_0 = U_0 C$ , siehe Abb. 2.7.3.

**Beispiel 2.7.20** (Bernoulli-DGL und Riccati-DGL).

1. Die *Bernoulli-DGL*

$$y' + f(x)y = g(x)y^n$$

wird mit der Substitution  $z(x) = y^{1-n}(x)$  zu der linearen DGL

$$z' + (1-n)f(x)z = (1-n)g(x).$$

Diese Substitution ist nur für Lösungen  $y$  mit  $y(x) > 0$  durchführbar.

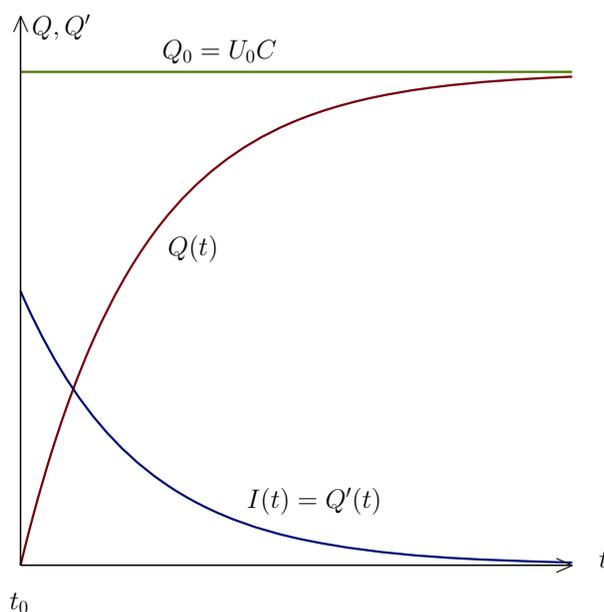
2. Die DGLn der Form

$$y' = f(x)y^2 + g(x)y + h(x)$$

heißen *Riccati-DGLn*. Kennen wir eine spezielle Lösung  $\tilde{y}$  der Riccati-DGL, so führt die Substitution  $z(x) = y(x) - \tilde{y}(x)$  auf die Bernoulli-DGL

$$z' - (2f(x)\tilde{y}(x) + g(x))z = f^2(x)y^2.$$

Abbildung 2.7.3: Spannungsverlauf der RC-Schaltung



### 2.7.2 Existenz und Eindeutigkeit von Lösungen

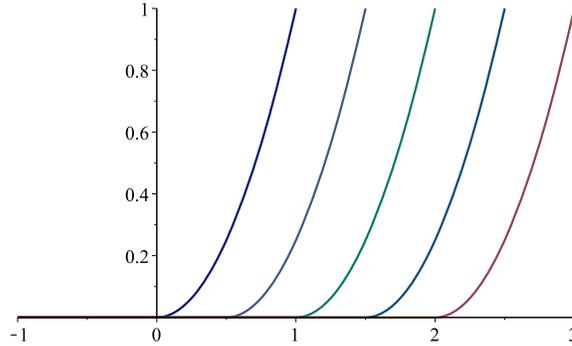
Im vorigen Abschnitt haben wir uns einige Beispiele für DGLn angeschaut sowie Lösungswege und -ansätze diskutiert. Dabei blieb stets die Frage offen, ob die angegebene Lösung eindeutig war, oder ob es mehrere Lösungen gibt. Die Antwort dafür ist der zentrale Teil dieses Kapitels. Die Eindeutigkeit einer Lösung kann schon allein wegen der auftretenden Integrationskonstanten nicht die richtige allgemeine Fragestellung sein; vielmehr müssen wir uns zur Festlegung der Konstanten ein Anfangswertproblem im Sinne von Definition 2.7.5 ansehen.

Aber auch hier ist die Eindeutigkeit nicht ohne weitere Voraussetzung an die DGL klar, wie das folgende Beispiel zeigt.

**Beispiel 2.7.21.** Wir betrachten die DGL  $y' = 2\sqrt{y}$  diese ist separierbar und hat wegen  $y, y' \geq 0$  die Lösungen

$$y_c(x) = \begin{cases} 0 & \text{falls } x \leq c \\ (x - c)^2 & \text{falls } x > c \end{cases} \quad \text{sowie} \quad y(x) = 0.$$

Abbildung 2.7.4: Die Lösungen  $y_c$  des AWP  $y' = 2\sqrt{y}$ ,  $y(0) = 0$ , für  $c = 0, \frac{1}{2}, 1, \frac{3}{2}, 2$



Das AWP  $y' = 2\sqrt{y}$ ,  $y(0) = 0$  hat damit unendlich viele Lösungen, die auf ganz  $\mathbb{R}$  definiert sind: Für jedes  $c \in \mathbb{R}^{\geq 0}$  ist  $y_c(x) = \begin{cases} 0 & \text{falls } x \leq c \\ (x - c)^2 & \text{falls } x > c \end{cases}$  eine Lösung, siehe Abb. 2.7.4.

Dagegen hat das AWP  $y' = 2\sqrt{y}$ ,  $y(0) = 1$  unter den obigen allgemeinen Lösungen die einzige Lösung  $y_{-1}(x) = \begin{cases} 0 & \text{falls } x \leq -1 \\ (x + 1)^2 & \text{falls } x > -1 \end{cases}$ .

Bevor, wir den zentralen Satz formulieren, benötigen wir noch etwas Handwerkzeug.

**Definition/Bemerkung 2.7.22.** 1. Es sei  $U \overset{\circ}{\subset} \mathbb{R}^k$  und  $f : \mathbb{R}^k \rightarrow \mathbb{R}^n$ . Dann heißt  $f$  *lokal Lipschitz-stetig*, wenn es für alle konvexen, beschränkten, offenen Mengen  $N$  mit  $\bar{N} \subset U$  eine Zahl  $L_N > 0$  gibt, sodass für alle  $x, y \in N$

$$\|f(x) - f(y)\| < L_N \|x - y\|.$$

2. Jede stetig differenzierbare Abbildung  $f : U \rightarrow \mathbb{R}^n$ ,  $U \overset{\circ}{\subset} \mathbb{R}^k$ , ist lokal Lipschitz-stetig. Als Lipschitz-Konstante können wir auf der konvexen, beschränkten Menge  $N \overset{\circ}{\subset} U$  die Zahl

$$L_N = \max_{x \in \bar{N}} \|Df(x)\|$$

wählen, siehe Satz 2.5.28.

Mit Hilfe des Banachschen Fixpunktsatzes beweisen wir das folgende Resultat.

**Satz 2.7.23** (Existenz- und Eindeutigkeitsatz von Picard-Lindelöf). *Gegeben sei das explizite DGL-System erster Ordnung*

$$y' = f(x, y)$$

mit  $f : I \times U \rightarrow \mathbb{R}^n$  definiert auf dem Produkt des offenen Intervalls  $I \subset \mathbb{R}$  und der offenen Menge  $U \subset \mathbb{R}^n$ . Weiter sei  $f$  lokal Lipschitz-stetig. Dann gilt:

1. Zu jedem Punkt  $(x_0, y^0) \in I \times U$  gibt es ein Intervall  $I_\epsilon(x_0) = [x_0 - \epsilon, x_0 + \epsilon] \subset I$  und eine Abbildung  $g : I_\epsilon(x_0) \rightarrow \mathbb{R}^n$  mit
  - $g'(x) = f(x, g(x))$  für alle  $x \in I_\epsilon(x_0)$  und
  - $g(x_0) = y^0$ .
2. Ist  $\tilde{g} : I_\epsilon(x_0) \rightarrow \mathbb{R}^n$  eine weitere Lösung des Anfangswertproblems  $y' = f(x, y)$ ,  $y(x_0) = y^0$ , so gilt bereits  $g = \tilde{g}$ .

**Folgerung 2.7.24.** 1. Die Umformulierung des AWP zur Gleichung

$$g(x) = y^0 + \int_{x_0}^x f(t, g(t)) dt$$

liefert mit dem Beweis des Banachschen Fixpunktsatzes (BFS) ein Iterationsverfahren zur Bestimmung der eindeutigen Lösung:

$$g_0(x) := y^0 \quad g_n(x) := y^0 + \int_{x_0}^x f(t, g_{n-1}(t)) dt.$$

Gemäß Konstruktion der Menge, auf der wir die obige Fixpunktgleichung definieren, konvergiert die so definierte Funktionenfolge  $(g_n)$  bezüglich der Supremumsnorm gegen die eindeutige Lösung  $g$  des AWP. Mit anderen Worten: die Funktionenfolge  $(g_n)$  konvergiert gleichmäßig gegen  $g$ .

2. Der Beweis des Existenz- und Eindeutigkeitsatzes zeigt, dass es ausreicht, die lokale Lipschitz-Stetigkeit von  $f$  im zweiten Argument zu fordern, also  $\|f(x, y) - f(x, \bar{y})\| < L\|y - \bar{y}\|$  für ein  $L > 0$ .

3. Haben wir die Lösung des AWP auf dem Intervall  $I_\epsilon(x_0)$  erhalten, so können wir den Iterationsprozess mit dem Anfangswert  $(x_0 + \epsilon, y(x_0 + \epsilon))$  erneut starten. Dadurch erhalten wir eine Lösung auf einem größeren Intervall  $I' \supset I_\epsilon(x_0)$ . Wir können dabei in der Regel nicht erwarten, dass wir bei diesem Prozess eine Lösung auf ganz  $I$  gewinnen. Nichtsdestotrotz ist das in speziellen Fällen möglich, siehe Kapitel 3.1.

Zum Abschluss folgen noch einige Bemerkungen, die hier ohne Beweis bleiben.

**Bemerkung 2.7.25.** 1. (Satz von Peano). Schwächt man in Satz 2.7.23 die Bedingung an die Regularität ab und fordert lediglich die Stetigkeit von  $f : I \times U \rightarrow \mathbb{R}$ , so folgt immer noch die Existenz der Lösung  $g : I_\epsilon(x_0) \rightarrow \mathbb{R}^n$  der Lösung des AWP. Die Eindeutigkeit der Lösung ist in der Regel nicht mehr gegeben, siehe Beispiel 2.7.21.

2. Eine wünschenswerte Eigenschaft der Lösungen des AWP ist: Ändern wir den Anfangswert  $(x_0, y^0)$  nur ein wenig und ändern wir die Abbildung  $f$  nur ein wenig, so ändert sich die Lösung auch nur ein wenig. Mathematischer formulieren wir das wie folgt.

Die Abbildung  $f : I \times U \rightarrow \mathbb{R}^n$  erfülle die Bedingungen aus Satz 2.7.23 und  $\tilde{f} : I \times U \rightarrow \mathbb{R}^n$  sei stetig. Weiter seien  $g$  die eindeutige Lösung des AWP  $y' = f(x, y)$ ,  $y(x_0) = y^0$  und  $\tilde{g}$  eine Lösung des AWP  $y' = \tilde{f}(x, y)$ ,  $y(x_0) = \tilde{y}^0$ . Es gelten die Abschätzungen

$$\|y^0 - \tilde{y}^0\| \leq \gamma \quad \text{und} \quad \|f(x, y) - \tilde{f}(x, y)\| \leq \delta.$$

Dann folgt

$$\|g(x) - \tilde{g}(x)\| \leq \gamma e^{L|x-x_0|} + \frac{\delta}{L} (e^{L|x-x_0|} - 1).$$

mit der Lipschitz-Konstante  $L$  von  $f$ .

Das bedeutet: Wir finden stets eine kleine Umgebung um unseren Startwert  $x_0$ , sodass sich die beiden Lösungen dort nur wenig unterscheiden. Man kann das jedoch nicht auf dem gesamten Definitionsbereich der Lösungen erwarten: Sind beide Lösungen für  $x$  mit großem  $|x - x_0|$  definiert, so können schon kleine Änderungen in den Daten des AWP "auf lange Sicht" große Fehler in der Lösung liefern.

## 2.8 Anhang Analysis II: Werkzeuge aus der linearen Algebra

Wir betrachten in diesem Anhang nur reelle Vektorräume, obwohl sich viele der Aussagen mit kleinen Modifikationen auch auf komplexe Vektorräume übertragen lassen. Ebenso betrachten wir nur endlichdimensionale Vektorräume, obwohl auch hier viele der betrachteten Begriffe auch auf allgemeinen Vektorräumen erklärt sind.

### 2.8.1 Normen und Matrixnormen

In diesem Anhang sammeln wir einige Begriffe aus der linearen Algebra zum Thema Normen und Matrixnormen.

**Definition 2.8.1.** Eine Norm auf einem Vektorraum  $V$  ist eine Abbildung  $\|\cdot\| : V \rightarrow \mathbb{R}^{\geq 0}$  mit

1.  $\|v\| = 0$  genau dann, wenn  $v = 0$ .
2.  $\|\alpha v\| = |\alpha| \|v\|$  für alle  $v \in V$  und  $\alpha \in \mathbb{R}$ .
3.  $\|v + w\| \leq \|v\| + \|w\|$  für alle  $v, w \in V$ .

Die letzte Eigenschaft heißt *Dreiecksungleichung*. Das Paar  $(V, \|\cdot\|)$  heißt *normierten Raum*.

**Beispiel 2.8.2.** 1. Auf dem Vektorraum  $\mathbb{R}^n$  definieren die folgenden Ausdrücke Normen

$$\bullet \|x\|_1 := \sum_{i=1}^n |x_i|, \quad \bullet \|x\|_2 := \sqrt{\sum_{i=1}^n x_i^2}, \quad \bullet \|x\|_\infty := \max_{1 \leq i \leq n} |x_i|,$$

oder allgemeiner für  $p \in \mathbb{N}^*$

$$\bullet \|x\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p}.$$

Diese Normen heißen *p-Normen*,  $p \in \mathbb{N}^* \cup \{\infty\}$ .

2. Ist  $\|\cdot\|$  eine Norm auf  $\mathbb{R}^n$ , dann ist für jedes  $c > 0$  die Abbildung  $x \mapsto c\|x\|$  ebenfalls eine Norm.

3. Sind  $\|\cdot\|_V$  bzw.  $\|\cdot\|_W$  Normen auf  $\mathbb{R}^n$  bzw.  $\mathbb{R}^m$ , so wird durch

$$\|(v, w)\| := \|v\|_V + \|w\|_W$$

für alle  $v \in \mathbb{R}^n, w \in \mathbb{R}^m$  die direkte Summe  $\mathbb{R}^{n+m} \simeq \mathbb{R}^n \oplus \mathbb{R}^m$  zu einem normierten Raum.

4. Identifizieren wir den Vektorraum  $M_{n,m}\mathbb{R}$  der  $m \times n$ -Matrizen auf die übliche Weise mit dem  $\mathbb{R}^{mn}$ , so können wir auf ihm die Normen aus Beispiel 1. analog definieren. Als ausgezeichnete Norm bezeichnen wir das analoge zur 2-Norm als *Frobeniusnorm* und schreiben

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2}.$$

Ebenso nennen wir das (skalierte) analoge zur  $\infty$ -Norm die *Totalnorm* auf  $M_{n,m}\mathbb{R}$

$$\|A\|_T := \sqrt{mn} \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} |A_{ij}|.$$

**Satz 2.8.3.** *Es seien  $\|\cdot\|_V$  bzw.  $\|\cdot\|_W$  Normen auf  $\mathbb{R}^n$  bzw.  $\mathbb{R}^m$ . Dann wird durch*

$$\|A\|_{V,W} := \max_{\|x\|_V=1} \|Ax\|_W$$

eine Norm auf  $M_{n,m}\mathbb{R}$  definiert. Sie heißt die zu  $\|\cdot\|_V$  und  $\|\cdot\|_W$  assoziierte *Matrixnorm*.

**Bemerkung 2.8.4.** Die assoziierte Matrixnorm hat die folgenden Eigenschaften.

1.  $\|A\|_{V,W} = \sup_{x \neq 0} \frac{\|Ax\|_W}{\|x\|_V}$ .
2. Für alle  $x \in V$  ist  $\|Ax\|_W \leq \|A\|_{V,W} \|x\|_V$ .

**Bemerkung 2.8.5.** Wir können die expliziten Formen der assoziierten Matrixnormen wie folgt erhalten:

- Schätze zunächst  $\|Ax\|_W$  ab, d.h. finde eine Zahl  $\lambda$  mit  $\|Ax\|_W \leq \lambda\|x\|_V$ . Dieses  $\lambda$  erfüllt dann  $\|A\|_{V,W} \leq \lambda$ .
- Dann gib ein  $x \in \mathbb{R}^n$  an mit  $\|x\|_V = 1$  und  $\|Ax\|_W = \lambda$ . Damit folgt  $\lambda \leq \|A\|_{V,W}$ .

Zusammen gilt somit  $\|A\|_{V,W} = \lambda$ .

**Beispiel 2.8.6.** Auf dem  $\mathbb{R}^n$  und  $\mathbb{R}^m$  betrachten wir jeweils die  $p$ -Normen mit dem gleichen  $p$ . Dann sind die assoziierten Normen auf  $M_{n,m}$  gegeben durch

$$1. \|A\|_{1,1} = \max_{1 \leq j \leq n} \sum_{i=1}^m |A_{ij}| \quad (\text{Spaltensummennorm})$$

$$2. \|A\|_{\infty,\infty} = \max_{1 \leq i \leq m} \sum_{j=1}^n |A_{ij}| \quad (\text{Zeilensummennorm})$$

$$3. \|A\|_{2,2} = \sqrt{\lambda_{\max}(A^T A)} \quad (\text{Spektralnorm})$$

Hier bezeichnet  $\lambda_{\max}(A^T A)$  den größten Eigenwert der Matrix  $A^T A$ .

Auf der Menge der quadratischen Matrizen haben wir eine zusätzliche Struktur, nämlich die Matrixmultiplikation. Um diese geeignet mit in die Diskussion einzubinden definieren wir wie folgt.

**Definition 2.8.7.** Eine Norm  $\|\cdot\|$  auf dem Vektorraum  $M_n\mathbb{R}$  der quadratischen  $n \times n$ -Matrizen heißt *Matrixnorm*, wenn sie zusätzlich *submultiplikativ* ist, d.h. es ist

$$\|AB\| \leq \|A\| \|B\|$$

für alle  $A, B \in M_n\mathbb{R}$ .

**Satz 2.8.8.** Es seien  $\|\cdot\|_V$  und  $\|\cdot\|_W$  Normen auf dem  $\mathbb{R}^n$ .

1. Auf der Menge der quadratischen Matrizen  $M_n\mathbb{R}$  ist die assoziierten Norm  $\|\cdot\|_{V,W}$  submultiplikativ, also eine Matrixnorm. Diese nennen wir die assoziierte Matrixnorm

2. Ist speziell  $\|\cdot\|_V = \|\cdot\|_W$ , so erfüllt die assoziierte Matrixnorm stets  $\|\mathbb{1}\|_{V,V} = 1$ .

**Beispiel 2.8.9.** 1. Die Frobeniusnorm ist submultiplikativ und erfüllt  $\|\mathbb{1}\|_F = \sqrt{n}$ . Sie ist also eine Matrixnorm, aber keine assoziierte Matrixnorm.

2. Die (unskalierte)  $\infty$ -Norm  $\|A\| := \max\{|A_{ij}| \mid i, j = 1, \dots, n\}$  auf  $M_n\mathbb{R}$  erfüllt  $\|\mathbb{1}\| = 1$ . Sie ist jedoch nicht submultiplikativ, denn für die Matrizen  $A$  mit  $A_{ij} = 1$  für alle  $i, j$  ist  $\|A\| = 1$ . Außerdem gilt  $A^2 = nA$ , also  $\|A^2\| = n$ . Zusammen haben wir also  $\|A^2\| > \|A\|^2$ . Damit ist  $\|\cdot\|$  sie keine Matrixnorm.
3. Die Totalnorm als skalierte Variante der  $\infty$ -Norm ist zwar submultiplikativ, sie erfüllt jedoch  $\|\mathbb{1}\|_T = n$  und ist somit ebenfalls keine assoziierte Matrixnorm.

**Definition 2.8.10.** 1. Es seien  $\|\cdot\|_V$  bzw.  $\|\cdot\|_W$  Normen auf  $\mathbb{R}^n$  bzw.  $\mathbb{R}^m$ . Eine Norm  $\|\cdot\|$  auf  $M_{n,m}\mathbb{R}$  heißt mit  $\|\cdot\|_V$  und  $\|\cdot\|_W$  verträglich, wenn

$$\|Ax\|_W \leq \|A\| \|x\|_V$$

für alle  $A \in M_{n,m}$  und  $x \in \mathbb{R}^n$ .

2. Eine Norm  $\|\cdot\|$  auf  $M_m\mathbb{R}$  heißt mit der Norm  $\|\cdot\|_0$  auf  $M_{n,m}\mathbb{R}$  verträglich, wenn

$$\|AB\|_0 \leq \|A\| \|B\|_0$$

für alle  $A \in M_m\mathbb{R}$  und  $B \in M_{n,m}\mathbb{R}$ . Dies ist eine Verallgemeinerung der Submultiplikativität.

**Beispiel 2.8.11.** 1. Per Definition ist die assoziierte Norm  $\|\cdot\|_{V,W}$  mit  $\|\cdot\|_V$  und  $\|\cdot\|_W$  verträglich.

2. Die Frobeniusnorm  $\|\cdot\|_F$  auf  $M_{n,m}\mathbb{R}$  ist mit  $\|\cdot\|_2$  auf  $\mathbb{R}^n$  und  $\mathbb{R}^m$  verträglich, d.h.

$$\|Ax\|_2 \leq \|A\|_F \|x\|_2$$

für alle  $A \in M_{n,m}\mathbb{R}$  und  $x \in \mathbb{R}^n$ .

3. Die Frobeniusnorm auf  $M_m\mathbb{R}$  ist mit der Spektralnorm  $\|\cdot\|_{2,2}$  auf  $M_{n,m}\mathbb{R}$  verträglich, d.h.

$$\|AB\|_{2,2} \leq \|A\|_F \|B\|_{2,2}$$

für alle  $A \in M_m\mathbb{R}$  und  $B \in M_{n,m}\mathbb{R}$ .

4. Die Totalnorm auf  $M_m\mathbb{R}$  ist für alle  $p \in \mathbb{N}^* \cup \{\infty\}$  mit den assoziierten Matrixnormen  $\|\cdot\|_{p,p}$  auf  $M_{n,m}\mathbb{R}$  verträglich, d.h.

$$\|AB\|_{p,p} \leq \|A\|_T \|B\|_{p,p}$$

für alle  $A \in M_m\mathbb{R}$  und  $B \in M_{n,m}\mathbb{R}$ .

**Bemerkung 2.8.12.** 1. Da bekanntlich im Zahlenraum alle Normen äquivalent sind, können wir sämtliche Normen auf  $M_{n,m}\mathbb{R}$  gegeneinander abschätzen, siehe 2.4.19. Das liefert interessante Ungleichungen im Zusammenhang mit Matrizen.

2. Eine Ungleichung die nicht aus der Äquivalenz herrührt ist die folgende, die für alle  $A \in M_n\mathbb{R}$  gilt:

$$(\|A\|_{2,2})^2 \leq \|A\|_{1,1} \|A\|_{\infty,\infty}.$$

**Bemerkung 2.8.13.** 1. Statt in Satz 2.8.3 die Argumentation mit Matrizen zu machen, können wir allgemeiner die assoziierte Abbildungsnorm  $\|\cdot\|_{X,Y}$  Norm genauso auf der Menge  $L(X,Y)$  der linearen Abbildungen zwischen zwei normierten Vektorräumen  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$  definieren. Für  $\Phi \in L(X,Y)$  ist

$$\|\Phi\|_{X,Y} = \max_{\|x\|_X=1} \|\Phi(x)\|_Y.$$

2. Auf die gleiche Weise bekommen wir auch eine Norm  $\|\cdot\|_{X,Y,Z}$  auf der Menge  $Bil(X,Y;Z)$  der bilinearen Abbildungen von den normierten Vektorräumen  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$  in den normierten Vektorraum  $(Z, \|\cdot\|_Z)$ . Für  $\beta \in Bil(X,Y;Z)$  ist

$$\|\beta\|_{X,Y,Z} := \max_{\substack{\|x\|_X=1 \\ \|y\|_Y=1}} \|\beta(x,y)\|_Z.$$

Diese Norm hat ebenfalls die Verträglichkeitseigenschaft

$$\|\beta(x,y)\|_Z \leq \|\beta\|_{X,Y,Z} \|x\|_X \|y\|_Y.$$

3. Betrachten wir in 2. den Spezialfall  $X = Y = \mathbb{R}^n$  und  $Z = \mathbb{R}$ . In diesem Fall ist  $\beta$  eine gewöhnliche Bilinearform und kann durch eine Matrix, etwa  $B \in M_n \mathbb{R}$ , beschrieben werden und es ist  $\beta(x, y) = x^T B y = \langle x, B y \rangle$ . Wählt man in  $\mathbb{R}^n$  die gewöhnliche 2-Norm so erhält man mit der Cauchy-Schwarz-Ungleichung

$$|\beta(x, y)| = |\langle x, B y \rangle| \leq \|x\|_2 \|B y\|_2 \leq \|B\| \|x\|_2 \|y\|_2$$

für jede mit  $\|\cdot\|_2$  verträgliche Matrixnorm  $\|\cdot\|$ . Hier kann man selbstverständlich die assoziierte Matrixnorm  $\|\cdot\|_{2,2}$  wählen, aber ebenso die Frobeniusnorm  $\|\cdot\|_F$  oder die Totalnorm  $\|\cdot\|_T$ .

### 2.8.2 Symmetrische Multilinearformen

**Definition 2.8.14.** Eine  $\ell$ -fache Multilinearform auf dem Vektorraum  $V$  ist eine Abbildung  $\mu : \overbrace{V \times \dots \times V}^{\ell\text{-mal}} \rightarrow \mathbb{R}$  mit

$$\begin{aligned} \mu(v^1, \dots, v^{i-1}, v^i + w, v^{i+1}, \dots, v^\ell) &= \mu(v^1, \dots, v^{i-1}, v^i, v^{i+1}, \dots, v^\ell) \\ &\quad + \mu(v^1, \dots, v^{i-1}, w, v^{i+1}, \dots, v^\ell) \end{aligned}$$

und

$$\mu(v^1, \dots, v^{i-1}, \alpha v^i, v^{i+1}, \dots, v^\ell) = \alpha \mu(v^1, \dots, v^{i-1}, v^i, v^{i+1}, \dots, v^\ell)$$

für alle  $1 \leq i \leq \ell$ ,  $v^1, \dots, v^\ell, w \in V$ ,  $\alpha \in \mathbb{R}$ .

**Bemerkung 2.8.15.** Ist  $\{e_i, \dots, e_n\}$  eine Basis des  $n$ -dimensionalen Vektorraums  $V$ , dann ist die  $\ell$ -fache Multilinearform  $\mu$  durch Angabe der  $n^\ell$  Zahlen

$$\mu_{i_1 \dots i_\ell} := \mu(e_{i_1}, \dots, e_{i_\ell}) \quad \text{für } 1 \leq i_1, \dots, i_\ell \leq n$$

eindeutig festgelegt. Entwickeln wir für  $1 \leq i \leq \ell$  die Vektoren  $v^i := \sum_{j=1}^n v_j^i e_j$  bezüglich der Basis  $\{e_j\}_{1 \leq j \leq n}$ , so gilt

$$\mu(v^1, \dots, v^\ell) = \sum_{i_1=1}^n \cdots \sum_{i_\ell=1}^n \mu_{i_1 \dots i_\ell} v_{i_1}^1 \cdots v_{i_\ell}^\ell.$$

Wir nennen  $(\mu_{i_1 \dots i_\ell})_{1 \leq i_1, \dots, i_\ell \leq n} \in \mathbb{R}^{n^\ell}$  auch die verallgemeinerte Matrixdarstellung von  $\mu$  bezüglich der Basis  $\{e_1, \dots, e_n\}$ .

**Beispiel 2.8.16.** Ist  $\ell = 2$  so liefert Definition 2.8.14 die Bilinearformen auf  $V$ . In diesem Fall erhalten wir mit der Bemerkung die Matrixdarstellung von  $\mu$  bezüglich der angegebenen Basis.

**Definition 2.8.17.** Eine  $\ell$ -fache Multilinearform heißt *symmetrisch*, wenn

$$\begin{aligned} & \mu(v^1, \dots, v^{i-1}, v^i, v^{i+1}, \dots, v^{j-1}, v^j, v^{j+1}, \dots, v^\ell) \\ &= \mu(v^1, \dots, v^{i-1}, v^j, v^{i+1}, \dots, v^{j-1}, v^i, v^{j+1}, \dots, v^\ell) \end{aligned}$$

für alle  $1 \leq i, j \leq \ell$ .

**Bemerkung 2.8.18.** Im Fall  $\ell = 1$  ist die Bilinearform genau dann symmetrisch, wenn das für eine Matrixdarstellung gilt. In diesem Fall hat die Matrixdarstellung nur reelle Eigenwerte.

**Definition 2.8.19.** Es sei  $\ell$  gerade. Eine  $\ell$ -fache Multilinearform heißt

- *positiv definit*, wenn  $\mu(v, \dots, v) > 0$  für alle  $v \in V \setminus \{0\}$ .
- *positiv semidefinit*, wenn  $\mu(v, \dots, v) \geq 0$  für alle  $v \in V \setminus \{0\}$ .
- *negativ definit*, wenn  $\mu(v, \dots, v) < 0$  für alle  $v \in V \setminus \{0\}$ .
- *negativ semidefinit*, wenn  $\mu(v, \dots, v) \leq 0$  für alle  $v \in V \setminus \{0\}$ .
- *indefinit*, wenn keiner der obigen Fälle zutrifft.

**Bemerkung 2.8.20.** • Eine symmetrische Multilinearform ist genau dann positiv (semi)definit, wenn  $-\mu$  negativ(semi)definit ist.

- Ist die symmetrische Multilinearform  $\mu$  indefinit, dann gibt es stets ein  $v \in V \setminus \{0\}$ , so dass  $\mu(v, \dots, v) = 0$ .

Im Fall  $\ell = 2$  gibt es dazu eine Reihe von Kriterien, die wir hier kurz wiederholen wollen

**Satz 2.8.21.** *Es sei  $\mu$  eine symmetrische Bilinearform mit Matrixdarstellung  $H$ . Dann gilt*

- $\mu$  ist positiv definit  $\iff H$  hat nur positiv Eigenwerte.

- $\mu$  ist positiv semidefinit  $\iff H$  hat nur nichtnegative Eigenwerte.
- $\mu$  ist negativ definit  $\iff H$  hat nur negative Eigenwerte.
- $\mu$  ist negativ semidefinit  $\iff H$  hat nur nichtpositive Eigenwerte.
- $\mu$  ist indefinit  $\iff H$  hat sowohl positive als auch negative Eigenwerte.

**Definition 2.8.22.** Es sei  $H = (H_{ij})_{1 \leq i, j \leq n}$  eine symmetrische  $n \times n$ -Matrix. Dann heißen für  $r = 1, \dots, n$  die symmetrischen  $r \times r$ -Matrizen  $H^{[r]} \in M_r \mathbb{R}$  mit

$$H^{[1]} = (H_{11}), \quad H^{[2]} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}, \dots,$$

$$H^{[r]} = \begin{pmatrix} H_{11} & \dots & H_{12} \\ \vdots & & \vdots \\ H_{r1} & \dots & H_{rr} \end{pmatrix}, \dots, \quad H^{[n]} = H$$

die Hauptminoren von  $H$ .

**Bemerkung 2.8.23.** Eine symmetrische Matrix  $H \in M_n \mathbb{R}$  ist regulär  $\iff \det H \neq 0 \iff$  kein Eigenwert verschwindet.

**Satz 2.8.24.** Es sei  $\mu$  eine symmetrische Bilinearform mit regulärer Matrixdarstellung  $H$ . Dann gilt

- $\mu$  ist positiv definit  $\iff \det H^{[r]} > 0$  für alle  $1 \leq r \leq n$ .
- $\mu$  ist negativ definit  $\iff (-1)^r \det H^{[r]} > 0$  für alle  $1 \leq r \leq n$ .
- $H$  ist indefinit  $\iff$  keiner der beiden vorigen Fälle trifft zu.

**Bemerkung 2.8.25.** • Im letzten Fall des Satzes können einige Untermatrizen tatsächlich verschwindende Determinante haben, z.B. bei der

regulären indefiniten Matrix  $H = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . Es ist jedoch stets

$$\det H^{[n]} = \det H \neq 0.$$

- Im nichtregulären Fall können wir die indefiniten Matrizen von den semidefiniten nicht mit Hilfe der Untersuchung der Determinanten unterscheiden.

- Die semidefiniten Fälle können wir im Allgemeinen weder im regulären noch im nicht-regulären Fall mit Hilfe der Determinanten auseinanderhalten.

In Dimension 2 ist die Situation sehr speziell, da es nur zwei Eigenwerte gibt.

**Bemerkung 2.8.26.** Es sei  $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{pmatrix}$  symmetrisch. Dann gilt

- $H$  ist positiv definit  $\iff \det H > 0$  und  $H_{11} > 0$ .
- $H$  ist negativ definit  $\iff \det H > 0$  und  $H_{11} < 0$ .
- $H$  ist indefinit  $\iff \det H < 0$ .
- $H$  ist positiv semidefinit  $\iff \det H = 0$  und  $H_{11} > 0$
- $H$  ist negativ semidefinit  $\iff \det H = 0$  und  $H_{11} < 0$

Außerdem ist  $H = 0 \iff \det H = 0$  und  $H_{11} = 0$ .

### 2.8.3 Matrixexponential

Es sei  $\|\cdot\|$  eine Matrixnorm auf der Menge  $M_n\mathbb{R}$  der  $n \times n$ -Matrizen, siehe Definition 2.8.7. Wegen der Ergebnisse von Abschnitt 2.3.3 wissen wir:

Eine Folge  $(A_n)$  von Matrizen konvergiert genau dann, wenn sie eine Cauchy-Folge ist, d.h.

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n > n_0 \forall \ell \in \mathbb{N} : \|A_{n+\ell} - A_n\| < \epsilon.$$

**Beispiel 2.8.27.** Es sei  $A \in M_n\mathbb{R}$  eine feste Matrix mit  $\|A\| = a$ . Dann ist die Folge  $(A_n)$ , die durch

$$A_n := \sum_{k=0}^n \frac{1}{k!} A^k$$

definiert ist, eine Cauchy-Folge. Es ist nämlich

$$\|A_{n+\ell} - A_n\| = \left\| \sum_{k=n+1}^{n+\ell} \frac{1}{k!} A^k \right\| \leq \sum_{k=n+1}^{n+\ell} \frac{1}{k!} \|A^k\| \leq \sum_{k=n+1}^{n+\ell} \frac{1}{k!} \|A\|^k = \sum_{k=n+1}^{n+\ell} \frac{a^k}{k!}$$

Die rechte Seite konvergiert für  $n \rightarrow \infty$  gegen Null, da es sich hier um einen Teil der reellen Exponentialreihe handelt.

**Definition 2.8.28.** Die Abbildung, die jeder Matrix  $A \in M_n \mathbb{R}$  den Grenzwert der "Matrixreihe"  $\sum \frac{1}{k!} A^k$  zuordnet, heißt *Matrixexponential*. Wir schreiben

$$\exp(A) := \sum_{n=0}^{\infty} \frac{1}{k!} A^k.$$

**Satz 2.8.29.** Gilt für zwei Matrizen  $A, B \in M_n \mathbb{R}$  die Relation

$$AB = BA$$

so gilt

$$\exp(A + B) = \exp(A) \exp(B) = \exp(B) \exp(A).$$

Deshalb ist für eine Blockdiagonalmatrix  $A = \begin{pmatrix} B & \\ & C \end{pmatrix}$

$$\exp(A) = \begin{pmatrix} \exp(B) & \\ & \exp(C) \end{pmatrix}.$$

Als nächstes schauen wir uns für  $k = 0, \dots, n-1$  die Matrizen  $N_k \in M_n \mathbb{R}$  an mit

$$(N_k)_{ij} = \begin{cases} 1 & \text{falls } j - i = k \\ 0 & \text{sonst} \end{cases},$$

also insbesondere  $N_0 = \mathbb{1}$ . Diese Matrizen erfüllen

$$N_k N_\ell = \begin{cases} N_{k+\ell} & \text{falls } k + \ell \leq n - 1 \\ 0 & \text{sonst} \end{cases}.$$

Insbesondere folgt daraus

- $N_k N_\ell = N_\ell N_k$
- $(N_1)^m = \begin{cases} \mathbb{1} & \text{falls } m = 0 \\ N_m & \text{falls } 0 < m < n \\ 0 & \text{falls } m \geq n \end{cases}$

Es sei nun weiter

$$\Lambda := \lambda \mathbb{1} + N_1 = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

eine Jordan-Elementarmatrix. Dann ist

$$\begin{aligned} \exp(t\Lambda) &= e^{t\lambda} \exp(tN_1) = e^{t\lambda} \sum_{k=0}^{n-1} \frac{t^k}{k!} N_k \\ &= e^{t\lambda} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \frac{t^3}{3!} & \cdots & \frac{t^{n-2}}{(n-2)!} & \frac{t^{n-1}}{(n-1)!} \\ & 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{n-3}}{(n-3)!} & \frac{t^{n-2}}{(n-2)!} \\ & & 1 & t & \cdots & \frac{t^{n-4}}{(n-4)!} & \frac{t^{n-3}}{(n-3)!} \\ & & & \ddots & \ddots & \vdots & \vdots \\ & & & & \ddots & t & \frac{t^2}{2} \\ & & & & & 1 & t \\ & & & & & & 1 \end{pmatrix}. \end{aligned}$$

Schreiben wir  $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$  für die Standardvektoren des  $\mathbb{R}^n$ , so gilt

$$\exp(t\Lambda)\vec{e}_j = e^{t\lambda} \sum_{k=0}^{n-1} \frac{t^k}{k!} N_k \vec{e}_j = e^{t\lambda} \sum_{k=0}^{j-1} \frac{t^k}{k!} \vec{e}_{j-k}.$$

#### 2.8.4 Einige Eigenschaften der Determinante

**Satz 2.8.30** (Cramersche Regel). *Es sei  $A \in M_n \mathbb{R}$  mit  $\det(A) \neq 0$  und  $\vec{b} \in \mathbb{R}^n$ . Die eindeutige Lösung  $\vec{x} = (x_1, \dots, x_n)$  des linearen Gleichungssystems  $A\vec{x} = \vec{b}$  ist durch*

$$x_i = \frac{1}{\det(A)} \det(\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{b}, \vec{a}_{i+1}, \dots, \vec{a}_n)$$

gegeben, wobei die  $\vec{a}_i$  die Spaltenvektoren der Matrix  $A$  sind.

**Bemerkung 2.8.31.** • Betrachten wir in der Cramerschen Regel als rechte Seite den  $i$ -ten Einheitsvektor, also  $\vec{b} = \vec{e}_i$ , so liefert die zugehörige Lösung  $\vec{b}_i$  den  $i$ -ten Spaltenvektor von  $A^{-1}$ . Das heißt

$$A_{ij}^{-1} = \frac{(-1)^{i+j}}{\det(A)} \det A^{[ji]}$$

wobei  $A^{[ji]} \in M_{n-1}\mathbb{R}$  die Matrix bezeichnet, die man aus  $A$  nach streichen der  $j$ -ten Zeile und  $i$ -ten Spalte erhält.

- Die Matrix

$$A^{\text{ad}} := \left( (-1)^{i+j} \det(A^{[ij]}) \right)_{i,j=1,\dots,n} \in M_n\mathbb{R}$$

heißt die *Adjunkte* zu  $A$ ; damit gilt  $A^{-1} = \frac{1}{\det(A)} (A^{\text{ad}})^T$ .

**Satz 2.8.32.** *Es seien  $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$  beliebig und  $A \in M_n\mathbb{R}$ , dann ist*

$$\sum_{k=1}^n \det(\vec{v}_1, \dots, \vec{v}_{k-1}, A\vec{v}_k, \vec{v}_{k+1}, \dots, \vec{v}_n) = \text{Spur}(A) \det(\vec{v}_1, \dots, \vec{v}_n).$$

**Satz 2.8.33.** *Die Determinante als Abbildung  $\det : M_n\mathbb{R} \rightarrow \mathbb{R}$  ist differenzierbar. Die Ableitung im Punkt  $A \in M_n\mathbb{R}$  in Richtung  $B \in M_n\mathbb{R}$  ist durch*

$$D \det(A) \cdot B = \sum_{k=1}^n \det(\vec{a}_1, \dots, \vec{a}_{k-1}, \vec{b}_k, \vec{a}_{k+1}, \dots, \vec{a}_n)$$

gegeben, wobei  $\vec{a}_i$  und  $\vec{b}_i$  die Spaltenvektoren der Matrizen  $A$  und  $B$  sind. Ist  $A$  invertierbar, dann ist

$$D \det(A) \cdot B = \det(A) \text{Spur}(BA^{-1}).$$

**Satz 2.8.34.** *Ist  $\{\vec{v}_1, \dots, \vec{v}_{n-1}\} \subset \mathbb{R}^n$  eine linear unabhängige Menge, dann gibt es einen eindeutig definierten Vektor  $\vec{n}$  mit*

$$\vec{n} \perp \vec{v}_i, \quad \|\vec{n}\| = 1, \quad \det(\vec{v}_1, \dots, \vec{v}_{n-1}, \vec{n}) > 0.$$

Schreiben wir  $V = (\vec{v}_1 \dots \vec{v}_{n-1}) \in M_{n-1,n}\mathbb{R}$ , dann gilt

$$\det(\vec{v}_1, \dots, \vec{v}_{n-1}, \vec{n}) = \sqrt{\det(V^T V)}.$$

### 3.1 Lineare DGLn und lineare DGL-Systeme

Wir hatten im vorigen Kapitel bereits die lineare Differentialgleichung erster Ordnung diskutiert, ohne den gewählten Namen zu begründen oder zu motivieren, siehe Kapitel 2.7.1.5. Wir hatten diese DGL als Spezialfall einer exakten DGL behandelt. In diesem Kapitel wollen wir dieses Beispiel verallgemeinern. Wir werden Lösungen konstruieren, grundlegende Eigenschaften besprechen und uns Gedanken über die Eindeutigkeit der Lösungen machen.

#### 3.1.1 Die lineare DGL $n$ -ter Ordnung

**Definition 3.1.1.** • Es seien  $f_0, f_1, \dots, f_n, g : I \rightarrow \mathbb{R}$  reelle Funktionen auf dem Intervall  $I \subset \mathbb{R}$  mit  $f_n \neq 0$ . Dann heißt

$$f_n(x)y^{(n)} + f_{n-1}(x)y^{(n-1)} + \dots + f_1(x)y' + f_0(x)y = g(x)$$

eine *lineare Differentialgleichung  $n$ -ter Ordnung*.

- Die  $f_k$  heißen *Koeffizientenfunktionen* und  $g$  heißt die *rechte Seite*.

- Ist  $g = \underline{0}$  die Nullfunktion, so nennen wir die lineare DGL *homogen*, andernfalls *inhomogen*.
- Ist die inhomogene DGL  $\sum_{k=0}^n f_k(x)y^{(k)} = g(x)$  gegeben, so nennen wir  $\sum_{k=0}^n f_k(x)y^{(k)} = 0$  die *zugehörige* homogene DGL.

- Sind die Koeffizientenfunktionen sämtlich konstant, so sprechen wir von einer *linearen DGL n-ter Ordnung mit konstanten Koeffizienten*. Wir schreiben dann

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = g(x)$$

und dürfen gegebenenfalls annehmen, dass  $a_n = 1$ .

**Bemerkung 3.1.2.** 1. Sind  $y_H$  und  $\tilde{y}_H$  Lösungen der gleichen homogenen linearen DGL und ist  $\alpha \in \mathbb{R}$ , so sind die Funktionen

$$y_H + \tilde{y}_H \quad \text{und} \quad \alpha y_H$$

ebenfalls Lösungen der DGL. Somit ist die Menge  $L_H$  der Lösungen einer homogenen DGL ein linearer Vektorraum.

2. Sind  $y_S$  und  $\tilde{y}_S$  Lösungen einer inhomogenen linearen DGL so ist die Funktion

$$y := y_S - \tilde{y}_S$$

eine Lösung der zugehörigen homogenen DGL. Somit ist die Menge  $L_I$  der Lösungen einer inhomogenen DGL ein affiner Raum. Ist  $y_S$  eine spezielle Lösung der inhomogenen DGL und  $L_H$  der Lösungsraum der homogenen DGL, so gilt

$$L_I = y_S + L_H.$$

**Definition 3.1.3.** Ist  $a_n y^{(n)} + \dots + a_1 y' + a_0 y = g(x)$  eine lineare DGL mit konstanten Koeffizienten, so heißt das Polynom

$$\chi(t) := a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0$$

das zugehörige *charakteristische Polynom*.

**Satz 3.1.4.** Gegeben sei die homogene lineare Differentialgleichung  $a_n y^{(n)} + \dots + a_1 y' + a_0 y = 0$  mit konstanten Koeffizienten und charakteristischem Polynom  $\chi(t)$ .

1. Ist  $\lambda$  eine  $r$ -fache Nullstelle des charakteristischen Polynoms, so sind die  $r$  Funktionen

$$e^{\lambda x}, xe^{\lambda x}, \dots, x^{r-1}e^{\lambda x}$$

linear unabhängige Lösungen der homogenen DGL.

2. Ist  $t^2 - 2\mu t + (\mu^2 + \omega^2)$  ein quadratischer Faktor des charakteristischen Polynoms der Ordnung  $r$  ohne Nullstellen, so sind die  $2r$  Funktionen

$$e^{\mu x} \sin(\omega x), xe^{\mu x} \sin(\omega x), \dots, x^{r-1}e^{\mu x} \sin(\omega x), \\ e^{\mu x} \cos(\omega x), xe^{\mu x} \cos(\omega x), \dots, x^{r-1}e^{\mu x} \cos(\omega x)$$

linear unabhängige Lösungen der homogenen DGL.

**Beweisskizze.** Wir nutzen die Darstellung mit Hilfe komplexer Zahlen, siehe Abschnitt 3.7. Wir zerlegen  $\chi(t)$  in seine komplexen Linearfaktoren: dabei ist dann  $t^2 - 2\mu t + (\mu^2 + \omega^2) = (t - \mu - i\omega)(t - \mu + i\omega)$ . Damit ist  $x^m e^{(\mu+i\omega)x} = x^m e^{\mu x} \cos(\omega x) + i x^m e^{\mu x} \sin(\omega x)$  und wir können Fall 1. als Spezialfall von 2. auffassen mit  $\omega = 0$ . Oder umgekehrt formuliert: es reicht zu zeigen, dass Fall 1. für komplexe Nullstellen richtig ist. Da das Polynom  $\chi(t)$  reell ist, treten die komplexen Nullstellen stets in komplex konjugierten Paaren auf. Wegen der Linearität der DGL, sind jeweils Real- und Imaginärteil der komplexen Lösung reelle Lösungen. Da diese sich bei den Paaren komplexer Nullstellen höchstens im Vorzeichen unterscheiden, reicht es aus, nur eine von beiden zu betrachten.

Wir haben also durch Einsetzen nachzuprüfen: Ist  $\lambda$  eine – vielleicht komplexe – Nullstelle von  $\chi(t)$  der Vielfachheit  $r$ , so lösen die – vielleicht komplexen – Funktionen  $e^{\lambda x}, xe^{\lambda x}, \dots, x^{r-1}e^{\lambda x}$  die homogene DGL.

Die lineare Unabhängigkeit der Funktionen zeigen wir, indem wir in Fall 1.  $r$  verschiedene Werte  $x_1, \dots, x_r$  einsetzen und das so erhaltene lineare Gleichungssystem für die unbekanntenen Koeffizienten lösen. Die Determinante der Koeffizientenmatrix ist dann ein Vielfaches einer Vandermonde-Determinante.

In Fall 2. wählen wir  $2r$  verschiedene Werte  $x_i = \frac{\pi}{2\omega} + \frac{2\pi}{\omega}k_i$ ,  $x_{i+r} = \frac{\pi}{\omega} + \frac{2\pi}{\omega}k_i$  wobei  $k_1, \dots, k_r \in \mathbb{Z}$ . Die Determinante der Koeffizientenmatrix ist dann ein Vielfaches des Quadrates einer Vandermonde-Determinante.  $\square$

**Folgerung 3.1.5.** Es sei  $\sum_{k=0}^n a_k y^{(k)} = 0$  eine homogene lineare DGL  $n$ -ter Ordnung mit charakteristischem Polynom  $\chi(t)$ .

1. Der Lösungsraum  $L_H$  enthält einen  $n$ -dimensionalen Unterraum  $L'_H$ , der von Lösungen der Form

$$x^k e^{\lambda x}, x^\ell e^{\mu x} \sin(\omega x), x^m e^{\mu x} \cos(\omega x)$$

aufgespannt wird.

2. Hat das charakteristische Polynom  $n$  verschiedene Nullstellen, so ist

$$L'_H = \text{span}_{\mathbb{R}} \{e^{\lambda_1 x}, \dots, e^{\lambda_n x}\}$$

**Bemerkung 3.1.6.** Es bleiben bis jetzt z.B. folgende Fragen offen

- a) Ist im Fall konstanter Koeffizienten vielleicht  $L'_H = L_H$ ?
- b) Geben wir "Anfangswerte"  $y(x_0) = \alpha_0, y'(x_0) = \alpha_1, \dots, y^{(n-1)}(x_0) = \alpha_{n-1}$  vor, gibt es dann eine eindeutige Lösung der DGL?
- c) Wie berechnen wir eine spezielle Lösung einer inhomogenen linearen DGL mit konstanten Koeffizienten für eine feste rechte Seite  $g(x)$ ? Diese Fragen werden wir in Abschnitt 3.1.3 angehen.

Zumindest auf die zweite Frage haben wir eine, wenn auch nur partielle Antwort: Im Fall ausgewählter rechter Seiten führen spezielle Ansätze auf eine Lösung der inhomogenen DGL mit linearen Koeffizienten.

**Bemerkung 3.1.7** (Lösungen spezieller inhomogener linearer DGLn). Es sei  $\sum_{k=0}^n a_k y^{(k)} = g(x)$  eine inhomogene lineare DGL mit konstanten Koeffizienten und es sei  $\chi(t)$  ihr charakteristisches Polynom. Weiter sei die rechte Seite von der Form

- a)  $g(x) = q(x)e^{\lambda x}$

$$\text{b) } g(x) = q_s(x)e^{\mu x} \sin(\omega x) + q_c(x)e^{\mu x} \cos(\omega x)$$

wobei  $\lambda, \mu, \nu \in \mathbb{R}$  und  $q, q_s, q_c$  Polynome vom Grad  $m, m_s, m_c$  sind. Es sei nun  $r \geq 0$  die Vielfachheit von  $t - \lambda$  bzw. von  $t^2 - 2\mu t + (\mu^2 + \omega^2)$  in der Zerlegung von  $\chi(t)$ .<sup>(i)</sup> Dann gibt es eine Lösung der inhomogenen DGL der Form

$$\text{a) } y_S(x) = x^r p_0(x) e^{\lambda x}$$

$$\text{b) } y_S(x) = x^r p_{0,s}(x) e^{\mu x} \sin(\omega x) + x^r p_{0,c}(x) e^{\mu x} \cos(\omega x).$$

Hierbei sind  $p_0$  ein Polynom vom Grad  $\leq m$  und  $p_{s,0}, p_{c,0}$  sind Polynome vom Grad  $\leq \max\{m_s, m_c\}$ .

**Beweisskizze.** Wieder nutzen wir die komplexe Darstellung. Wir schreiben  $p = x^r p_0, p_c = x^s p_{c,0}$  und  $p_s = x^s p_{s,0}$ . Durch Einsetzen des Ansatzes erhalten wir die Polynomgleichungen

$$\text{a) } q(x) = \sum_{k=0}^n \frac{\chi^{(k)}(\lambda)}{k!} p^{(k)}(x)$$

$$\text{b) } q_s(x) = \sum_{k=0}^n \frac{\operatorname{Re}(\chi^{(k)}(\mu + i\omega))}{k!} p_s^{(k)}(x) - \sum_{k=0}^n \frac{\operatorname{Im}(\chi^{(k)}(\mu + i\omega))}{k!} p_c^{(k)}(x)$$

$$q_c(x) = \sum_{k=0}^n \frac{\operatorname{Re}(\chi^{(k)}(\mu + i\omega))}{k!} p_c^{(k)}(x) + \sum_{k=0}^n \frac{\operatorname{Im}(\chi^{(k)}(\mu + i\omega))}{k!} p_s^{(k)}(x)$$

Durch Koeffizientenvergleich erhalten wir ein lineares Gleichungssystem für die Koeffizienten von  $p$  bzw. von  $p_s, p_c$ .  $\square$

### 3.1.2 Die Schwingungsgleichung

Wir betrachten die folgende inhomogene lineare DGL zweiter Ordnung:

$$y'' + \overbrace{2\gamma y'}^{\text{Dämpfung}} + \underbrace{\omega^2 y}_{\text{harmonische Schwingung}} = \overbrace{g(x)}^{\text{äußere Kraft}},$$

wobei  $\gamma, \omega \in \mathbb{R}^{\geq 0}$ . Das charakteristische Polynom dieser DGL ist  $\chi(t) = t^2 + 2\gamma t + \omega^2$  mit den Nullstellen

$$t_{\pm} = -\gamma \pm \sqrt{\gamma^2 - \omega^2}.$$

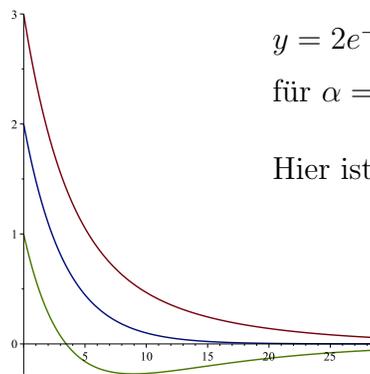
<sup>(i)</sup>Dabei bedeuten  $r = 0$ , dass der jeweilige Faktor nicht vorkommt.

Die Lösungen der homogenen DGL mit  $g(x) = 0$  ergeben sich zu

$$\begin{aligned} y_1(x) &= e^{-(\gamma - \sqrt{\gamma^2 - \omega^2})x}, & y_2(x) &= e^{-(\gamma + \sqrt{\gamma^2 - \omega^2})x} & \text{falls } \gamma^2 > \omega^2 \\ y_1(x) &= e^{-\gamma x}, & y_2(x) &= x e^{-\gamma x} & \text{falls } \gamma^2 = \omega^2 \\ y_1(x) &= e^{-\gamma x} \cos(\omega_0 x), & y_2(x) &= e^{-\gamma x} \sin(\omega_0 x) & \text{falls } \gamma^2 < \omega^2 \end{aligned}$$

wobei wir im letzten Fall  $\omega_0 := \sqrt{\omega^2 - \gamma^2}$  setzen.<sup>(i)</sup>

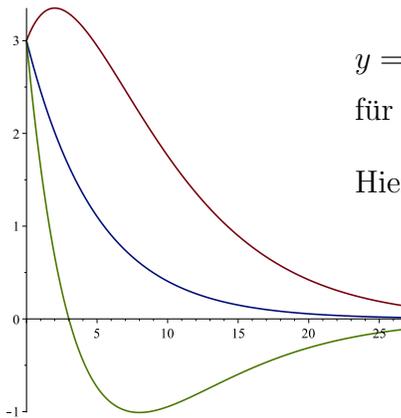
Abbildung 3.1.1: Lösungen der homogenen Schwingungsgleichung



$$y = 2e^{-\frac{3}{10}x} + \alpha e^{-\frac{1}{10}x}$$

für  $\alpha = 1$  (rot)  $\alpha = 0$  (blau)  $\alpha = -1$  (grün)

Hier ist  $\gamma = \frac{2}{10}$ ,  $\omega = \frac{\sqrt{3}}{10}$ , d.h.  $\omega_0 = \frac{1}{10}$

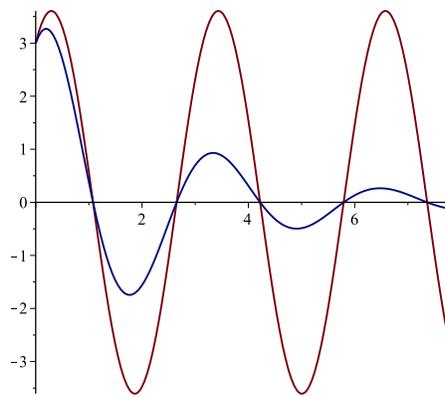


$$y = 3e^{-\frac{2}{10}x} + \alpha x e^{-\frac{2}{10}x}$$

für  $\alpha = 1$  (rot)  $\alpha = 0$  (blau)  $\alpha = -1$  (grün)

Hier ist  $\gamma = \omega = \frac{1}{5}$ , d.h.  $\omega_0 = 0$

<sup>(i)</sup> $\omega_0$  heißt in diesem Fall auch die *Eigenfrequenz* des schwingenden Systems.



$$y = 2e^{-\gamma x} \sin(2x) + 3e^{-\gamma x} \cos(2x)$$

Hier ist  $\gamma = 0, \omega = 2$  (rot)

und  $\gamma = -\frac{2}{5}, \omega = 2$  (blau),

d.h. in beiden Fällen  $\omega_0^2 < 0$

Wenden wir uns dem Fall einer periodischen äußeren Kraft zu,  $g(x) = \cos(\varpi x)$ , so müssen wir bei der Suche nach Lösungen zwei Fälle unterscheiden: 1.  $\gamma = 0$  und  $\varpi^2 = \omega^2$  und 2. jeden anderen Fall.

In Fall 1. ist ohne Einschränkung  $\varpi = \omega$  eine einfache Nullstelle des charakteristischen Polynoms. Der Ansatz  $y_S(x) = ax \cos(\omega x) + bx \sin(\omega x)$  eingesetzt in die Ausgangs-DGL liefert  $b = \frac{1}{2\omega}$  und  $a = 0$ , also

$$y_S(x) = \frac{x}{2\omega} \sin(\omega x).$$

In Fall 2. ist  $\varpi$  keine Nullstelle des charakteristischen Polynoms. Der Ansatz  $y_S(x) = a \cos(\varpi x) + b \sin(\varpi x)$  liefert das lineare Gleichungssystem

$$\begin{pmatrix} \omega^2 - \varpi^2 & 2\gamma\varpi \\ -2\gamma\varpi & \omega^2 - \varpi^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

mit der Lösung  $a = \frac{\omega^2 - \varpi^2}{(\omega^2 - \varpi^2)^2 + 4\gamma^2\varpi^2}$ ,  $b = \frac{2\gamma\varpi}{(\omega^2 - \varpi^2)^2 + 4\gamma^2\varpi^2}$ .

Wir normieren diese Lösung gemäß  $a = \alpha a_0$  und  $b = \alpha b_0$  so, dass  $a_0^2 + b_0^2 = 1$  ist, also

$$a_0 = \frac{\omega^2 - \varpi^2}{\sqrt{(\omega^2 - \varpi^2)^2 + 4\gamma^2\varpi^2}}, \quad b_0 = \frac{2\gamma\varpi}{\sqrt{(\omega^2 - \varpi^2)^2 + 4\gamma^2\varpi^2}}.$$

Setzen wir weiter  $A(\varphi) := \alpha$ , also

$$A(\varpi) = \frac{1}{\sqrt{(\omega^2 - \varpi^2)^2 + 4\gamma^2\varpi^2}}$$

$$= \frac{1}{\sqrt{(\gamma^2 + \omega_0^2)^2 + \varpi^2(\varpi^2 + 2\gamma^2 - 2\omega_0^2)}},$$

so ergibt sich als Lösung der inhomogenen DGL

$$y_S(x) = A(\varpi)(a_0 \cos(\varpi x) + b_0 \sin(\varpi x)) = A(\varpi) \cos(\varpi x - \phi(\varpi))$$

mit  $\cot(\phi(\varpi)) = \frac{a_0}{b_0}$  oder

$$\phi(\varpi) = \operatorname{arccot}\left(\frac{\omega^2 - \varpi^2}{2\gamma\varpi}\right).$$

Als Resonanzfrequenz  $\omega_R$  dieses Systems bezeichnet man die Frequenz  $\varpi$ , bei der die Amplitude  $A(\varpi)$  ihr Maximum annimmt, also

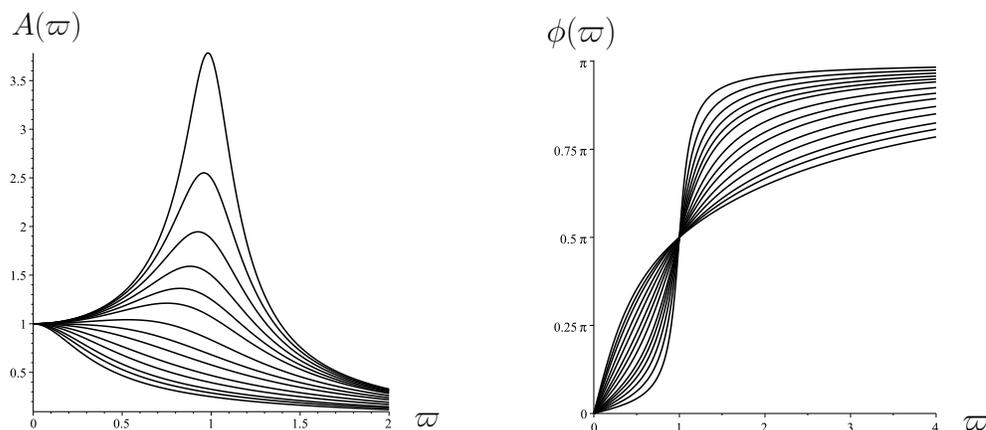
$$\omega_R^2 = \omega_0^2 - \gamma^2 = \omega^2 - 2\gamma^2.$$

Im Fall, dass die Resonanzfrequenz existiert, also im Fall  $\omega^2 > 2\gamma^2$ , ist die maximale Amplitude gegeben durch

$$A(\omega_R) = \frac{1}{2\omega_0\gamma} = \frac{1}{2\gamma\sqrt{\omega^2 - \gamma^2}},$$

siehe Abbildung 3.1.2.

Abbildung 3.1.2: Der Amplitudenverlauf  $A(\varpi)$  und die Phaseverschiebung  $\phi(\varpi)$  für  $\omega = 1$  und ausgewählte  $\frac{2}{15} \leq \gamma \leq \frac{29}{15}$



Im Fall eines Systems mit kleiner Dämpfung  $\gamma \ll 1$  und Frequenz nahe der Resonanzfrequenz  $\varpi \sim \omega_R$  spricht man auch von Resonanzkatastrophe. Die

Auswirkungen einer solchen Resonanzkatastrophe werden in dem Video zu Abbildung 3.1.3 sehr anschaulich deutlich.

Abbildung 3.1.3: Die Tacoma-Narrows-Brücke



### 3.1.3 Lineare DGL-Systeme

Wir können wegen Beispiel 2.7.7.3. jede lineare DGL  $n$ -ter Ordnung, z.B.

$$y^{(n)} = b - a_{n-1}y^{(n-1)} - \dots - a_1y' - a_0y$$

in ein DGL-System erster Ordnung übersetzen. Setzen wir  $u_1 := y$ ,  $u_2 := y'$ ,  $\dots$ ,  $u_n := y^{(n-1)}$  dann ist das DGL-System gegeben durch

$$\begin{pmatrix} u_1 \\ \vdots \\ u_{n-2} \\ u_{n-1} \\ u_n \end{pmatrix}' = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ -a_0 & \dots & \dots & \dots & -a_{n-1} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_{n-2} \\ u_{n-1} \\ u_n \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ b \end{pmatrix}$$

Das begründet auch die Bezeichnung charakteristisches Polynom für das zur DGL zugehörige Polynom  $\chi(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1t + a_0$ : Das Polynom

ist genau das charakteristische Polynom der obigen Matrix; oder umgekehrt: Die Matrix ist gerade die Begleitmatrix zum Polynom  $\chi$ .

**Definition 3.1.8.** Es sei  $I \subset \mathbb{R}$  ein Intervall,  $A : I \rightarrow M_n \mathbb{R}$  eine stetige Abbildung mit Werten in den  $n \times n$ -Matrizen und  $b : I \rightarrow \mathbb{R}^n$  ebenfalls stetig.

- Ein *lineares DGL-System erster Ordnung* ist eine DGL der Form

$$u'(x) = A(x)u(x) + b(x)$$

für die  $\mathbb{R}^n$ -wertige Abbildung  $u$ .

- Das lineare DGL-System heißt *homogen*, wenn  $b(x) = 0$  für alle  $x \in I$ , andernfalls *inhomogen* und  $b$  heißt die *rechte Seite*
- Ist  $A$  eine konstante Matrix, so heißt  $Au = b(x)$  ein *lineares DGL-System mit konstanten Koeffizienten*.

**Satz 3.1.9.** Es seien  $A : I \rightarrow M_n \mathbb{R}$  und  $b : I \rightarrow \mathbb{R}^n$  stetig.

1. Das AWP  $u' = A(x)u + b(x)$ ,  $u(x_0) = u^0$  hat genau eine Lösung in  $I$ .
2. Die Lösungsmenge  $L_H$  des homogenen linearen DGL-Systems  $u' = A(x)u$  ist ein  $n$ -dimensionaler Vektorraum.
3. Ist  $u_S$  eine Lösung des inhomogenen linearen DGL-Systems  $u' = A(x)u + b(x)$ , so ist seine Lösungsmenge  $L_I$  der affine Raum  $L_I = u_S + L_H$ .

**Folgerung 3.1.10.** Damit sind die ersten zwei Fragen aus Bemerkung 3.1.6 beantwortet. Ist

$$f_n(x)y^{(n)} + f_{n-1}(x)y^{(n-1)} + \dots + f_1(x)y' + f_0(x)y = g(x)$$

eine lineare DGL  $n$ -ter Ordnung auf dem Intervall  $I \subset \mathbb{R}$ , so gilt:

- a) Für die Mengen  $L_H$  und  $L'_H$  aus Bemerkung 3.1.2 und Folgerung 3.1.5 gilt tatsächlich  $L_H = L'_H$ .
- b) Es gibt eine eindeutige Lösung der DGL, wenn wir die Bedingungen  $y(x_0) = \alpha_0, y'(x_0) = \alpha_1, \dots, y^{(n-1)}(x_0) = \alpha_{n-1}$  an die Lösung vorgeben.

Es seien nun  $u^1, \dots, u^n : I \rightarrow \mathbb{R}^n$  linear unabhängige Lösungen des homogenen linearen DGL-Systems

$$u' = A(x)u.$$

Diese schreiben wir als Spalten in eine Matrix und erhalten so eine matrixwertige Abbildung  $U : I \rightarrow M_n \mathbb{R}$

$$U(x) = \left( u^1(x), u^2(x), \dots, u^n(x) \right).$$

Diese Abbildung erfüllt die DGL

$$U' = A(x)U.$$

Es sei nun  $x_0 \in I$  und  $U_0 := U(x_0)$  invertierbar<sup>(i)</sup> und weiter sei  $u^0 \in \mathbb{R}^n$  ein beliebiger Vektor. Dann ist

$$u(x) := U(x)U_0^{-1}u^0$$

die eindeutige Lösung des AWP  $u' = A(x)u$ ,  $u(x_0) = u^0$ .

**Definition 3.1.11.** Eine Basis  $\{u^1, \dots, u^n\}$  des Lösungsraums eines homogenen DGL-Systems heißt auch *Fundamentalsystem* und die daraus gebildete Matrix  $U$  die *Fundamentalmatrix*.

**Satz 3.1.12.** *Es sei  $u^1$  eine Lösung des homogenen linearen DGL-Systems  $u' = A(x)u$ , deren  $i$ -te Koordinate nicht identisch verschwindet. Weiter sei  $z : I \rightarrow \mathbb{R}^n$  eine Abbildung, deren  $i$ -te Koordinate verschwindet, und  $\alpha : I \rightarrow \mathbb{R}$  eine Funktion mit*

$$\alpha'(x)u^1(x) + z'(x) = A(x)z(x).$$

*Eliminiert man nun  $\alpha'(x)$  mit Hilfe der  $i$ -ten Zeile aus diesem DGL-System, so erhält man für  $(z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n)$  ein homogenes lineares DGL-System der Größe  $n-1$  und eine gewöhnliche DGL für  $\alpha$ . Sind  $z(x)$  und  $\alpha(x)$  Lösungen dieser beiden, so erhält man durch*

$$u^2(x) := \alpha(x)u^1(x) + \tilde{z}(x)$$

*eine zu  $u^1$  linear unabhängige Lösung des homogenen Ausgangssystems. Hierbei erhält man  $\tilde{z}$  aus  $z$  indem man an der  $i$ -ten Stelle eine Null einfügt.*

---

<sup>(i)</sup>Wir werden später sehen, dass die Fundamentalmatrix an jeder Stelle invertierbar ist, siehe Bemerkung 3.1.14.

Dieses Reduktionsverfahren lässt sich induktiv fortsetzen:

**Bemerkung 3.1.13.** Es seien  $u^1, \dots, u^k$  linear unabhängige Lösungen des homogenen linearen DGL-Systems  $u' = A(x)u$ . Dann gibt es eine  $k \times k$ -Unterdeterminante der  $(k \times n)$ -Matrix  $\hat{U}(x) := ((u^1(x), \dots, u^k(x)))$ , die nicht identisch verschwindet – dies sei ohne Einschränkung der obere  $k \times k$ -Block. Weiter sei  $z : I \rightarrow \mathbb{R}^n$  eine Abbildung, deren ersten  $k$  Koordinaten verschwinden, und  $\alpha_1, \dots, \alpha_k : I \rightarrow \mathbb{R}$  Funktionen mit

$$\hat{U}(x) \begin{pmatrix} \alpha_1'(x) \\ \vdots \\ \alpha_k'(x) \end{pmatrix} + z'(x) = A(x)z(x).$$

Eliminiert man nun  $\alpha_1'(x), \dots, \alpha_k'(x)$  mit Hilfe der ersten  $k$  Zeilen aus diesem DGL-System, so erhält man für  $(z_{k+1}, \dots, z_n)$  ein homogenes lineares DGL-System der Größe  $n - k$  sowie  $k$  gewöhnliche DGLn für die Funktionen  $\alpha_1, \dots, \alpha_k$ . Ihre Lösungen liefern dann durch

$$u^{k+1}(x) := \hat{U}(x) \begin{pmatrix} \alpha_1(x) \\ \vdots \\ \alpha_k(x) \end{pmatrix} + \tilde{z}(x)$$

eine zu  $u^1, \dots, u^k$  linear unabhängige Lösung des homogenen Ausgangssystems, wobei sich  $\tilde{z}$  aus  $z$  ergibt, indem  $k$  Nullen vorangestellt werden.

**Bemerkung 3.1.14.** Ist  $U(x)$  eine Fundamentalmatrix des homogenen linearen DGL-Systems  $u' = A(x)u$ , so nennen wir die Funktion

$$w : I \rightarrow \mathbb{R} \quad \text{mit} \quad w(x) := \det(U(x))$$

die *Wronski-Determinante* des linearen DGL-Systems. Die Funktion  $w$  hängt nur vom DGL-System ab und nicht von der Wahl des Fundamentalsystems. Die Wronski-Determinante erfüllt die DGL

$$w' = \text{Spur}(A(x))w.$$

Wegen Bemerkung 2.7.18.3 ist die Lösung dieser gewöhnlichen DGL

$$w(x) = w(x_0)e^{\int_{x_0}^x \text{Spur}(A(t))dt}.$$

Insbesondere ist die Fundamentalmatrix somit in jedem Punkt invertierbar, wenn sie es nur in einem ist.

Das Obige lässt sich nun auf lineare Differentialgleichungen übertragen: Wenn wir bereits  $n - 1$  linear unabhängige Lösungen einer linearen DGL  $n$ -ter Ordnung haben, dann können wir die Wronski-Determinante nutzen, um die noch fehlende Lösung mit Hilfe einer linearen DGL  $(n - 1)$ -ter Ordnung zu finden.

**Beispiel 3.1.15.** Es sei  $y'' + f_1y' + f_0y = 0$  eine lineare DGL 2-ter Ordnung und es sei  $y_1$  mit  $y_1(x_0) = \alpha_1, y_1'(x_0) = \alpha_2$  eine Lösung. Gesucht ist eine zu  $y_1$  linear unabhängige Lösung  $y_2$  mit  $y_2(x_0) = \beta_1, y_2' = \beta_2$ . Es sei weiter  $\gamma := \alpha_1\beta_2 - \beta_1\alpha_2$ .

Das zur linearen DGL zweiter Ordnung gehörige lineare DGL-System mit  $u_1 = y, u_2 = y'$  hat die Form

$$\begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -f_0 & -f_1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

Weiter sei  $w(x)$  die Wronski-Determinante dieses DGL-Systems, also die Lösung der DGL erster Ordnung  $w' = -f_1w$  zum Anfangswert  $w(x_0) = \gamma$ , siehe dazu auch Bemerkung 2.7.18.3. Dann ergibt sich die gesuchte Funktion  $y_2$  als Lösung der linearen DGL erster Ordnung

$$y_1(x)y_2' - y_1'(x)y_2 = -w(x)$$

zum Anfangswert  $y_2(x_0) = \beta_1$ .

**Satz 3.1.16.** 1. Es sei  $A \in M_n\mathbb{R}$  eine Matrix und  $u^0 \in \mathbb{R}^n$ . Dann ist die Lösung des linearen DGL-System mit konstanten Koeffizienten zum Anfangswert  $u(x_0) = u^0$  gegeben durch

$$u(x) = e^{(x-x_0)A} u^0,$$

denn  $(e^{(x-x_0)A}u^0)' = Ae^{(x-x_0)A}u^0$ . Zur Definition des Matrixexponentials siehe Abschnitt 2.8.3.

2. Es sei  $A \in M_n\mathbb{R}$  und  $u' = Au$  ein homogenes lineares DGL-System mit konstanten Koeffizienten mit Lösungsraum  $L_H^{(A)}$ . Ist weiter  $S \in M_n\mathbb{R}$  invertierbar und  $B = SAS^{-1}$  eine zu  $A$  ähnliche Matrix, so ist der Lösungsraum des DGL-Systems  $u' = Bu$  gegeben durch

$$L_H^{(B)} = S L_H^{(A)}.$$

Zur Bestimmung des Lösungsraums reicht es also, wenn wir uns auf Matrizen beschränken, die eine besonders einfache Gestalt haben.

**Bemerkung 3.1.17.** Es sei im Folgenden mit den Bezeichnungen aus Abschnitt 2.8.3

$$\Lambda := \lambda \mathbb{1} + N_1 = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in M_r \mathbb{R}$$

eine Jordan-Elementarmatrix mit  $\lambda \in \mathbb{R}$ . Dann wissen wir, dass der Eigenraum sowie alle Haupträume bis zur Stufe  $r - 1$  jeweils die Dimension 1 haben. Eine Basis aus Eigen- und Hauptvektoren ist durch die Standardbasis  $\{\vec{e}_1, \dots, \vec{e}_r\}$  gegeben; es ist also

$$\Lambda \vec{e}_1 = \lambda \vec{e}_1, \quad \Lambda \vec{e}_j = \lambda \vec{e}_j + \vec{e}_{j-1} \quad \text{für } j = 2, \dots, r$$

Damit ist dann  $\{u^1, \dots, u^r\}$  mit

$$u^j(x) = \exp(x\Lambda) \vec{e}_j = e^{\lambda x} \left( \vec{e}_j + x \vec{e}_{j-1} + \dots + \frac{x^{j-1}}{(j-1)!} \vec{e}_1 \right)$$

ein Fundamentalsystem des homogenen DGL-Systems

$$u' = \Lambda u.$$

Mit diesen Vorbemerkungen können wir nun ein Fundamentalsystem eines linearen DGL-Systems mit konstanten Koeffizienten angeben.

**Folgerung 3.1.18.** 1. Sei  $J = S^{-1}AS$  die Jordandarstellung der Matrix  $A$  mit nur reellen Diagonalelementen. Für die Jordanmatrix lässt sich mit Hilfe der Standardbasis blockweise nach dem obigen Schema ein Fundamentalsystem von  $u' = Ju$  angeben. Wenden wir auf diese nun die Matrix  $S$  an, so erhalten wir das gesuchte Fundamentalsystem von  $u' = Au$ .

2. Statt die Transformationsmatrix  $S$  zu bestimmen, kann man auch nach dem bekannten Schema direkt die Jordanbasis von  $A$  bestimmen und den obigen Ansatz für die  $u^j$  nutzen; dabei ersetze man die  $\vec{e}_j$  durch die entsprechenden Elemente der Jordanbasis.

3. Besitzt die Jordanmatrix auch komplexe Diagonalelemente, dann gibt es auch komplexe Eigen- und Hauptvektoren. Diese kommen immer in konjugierten Paaren vor, da  $A \in M_n \mathbb{R}$  als reell vorausgesetzt ist. Um nun ein Fundamentalsystem zu erhalten, geht man genau wie oben vor, rechnet jedoch mit komplexen Zahlen. Als Zwischenergebnis erhält man so eine Menge von komplexen Abbildungen der Form  $\{u^1, \dots, u^m, v^1, \overline{v^1}, \dots, v^p, \overline{v^p}\}$  mit  $m + 2p = n$ , hierbei sind die  $u^j$  reelle Abbildungen. Der Übergang zum reellen Fundamentalsystem geschieht nun, indem wir von den komplexen Lösungen Real- und Imaginärteil getrennt betrachten:

$$\left\{ u^1, \dots, u^m, u^{m+1} = \operatorname{Re}(v^1), \dots, u^{m+p} = \operatorname{Re}(v^p), \right. \\ \left. u^{m+p+1} = \operatorname{Im}(v^1), \dots, u^{m+2p} = \operatorname{Im}(v^p) \right\}$$

4. Da es in der Regel ein unschönes Problem ist, die Jordanbasis zu berechnen, kann man sich auch wie folgt behelfen:
- Berechne zuerst die Eigenwerte und Eigenvektoren. Dabei sei  $k_\lambda$  die Vielfachheit des Eigenwertes  $\lambda$  und  $r_\lambda$  die Dimension des zu  $\lambda$  gehörigen Eigenraums.<sup>(i)</sup> Insbesondere ist  $k_\lambda \geq r_\lambda$ .
  - Es sei  $\{v^1, \dots, v^{r_\lambda}\}$  eine Eigenraumbasis von  $\lambda$  und es sei  $k_\lambda > r_\lambda$ . Dann erhalten wir einen Teil des zu  $\lambda$  gehörigen Fundamentalsystems durch  $u^j(x) = e^{\lambda x} v^j$  für  $j = 1, \dots, r_\lambda$ .
  - Ist nun  $s_\lambda$  die Größe des größten Jordan-Elementarblockes zum Eigenwert  $\lambda$ , so kann man den folgenden Ansatz zur Berechnung der fehlenden Lösungsabbildungen wählen:<sup>(ii)</sup>

$$y = \begin{pmatrix} p_1(x) \\ \vdots \\ p_n(x) \end{pmatrix} e^{\lambda x}$$

---

<sup>(i)</sup>Die Vielfachheit eines Eigenwertes entspricht der Vielfachheit von  $(t - \lambda)$  im charakteristischen Polynom  $\chi(t)$ . Sie stimmt überein mit der Anzahl der Jordan-Elementarblöcke zu  $\lambda$ .

<sup>(ii)</sup>Die Größe des größten Jordan-Elementarblockes zum Eigenwert  $\lambda$  erhält man als Vielfachheit von  $(t - \lambda)$  im Minimalpolynom  $\mu(t)$  der Ausgangsmatrix.

wobei die  $p_j$  Polynome vom Grad  $\leq s_\lambda - 1$  sind. Kennt man das Minimalpolynom nicht, so kann man  $s_\lambda - 1$  durch  $k_\lambda - r_\lambda$  nach oben abschätzen.

5. Eine besonders einfache Situation tritt ein, wenn die Matrix  $A$  diagonalisierbar ist. In diesem Fall existiert eine Basis  $\{v^i\}_{i=1,\dots,n}$  aus Eigenvektoren zu den (nicht unbedingt unterschiedlichen) Eigenwerten  $\{\lambda_i\}_{i=1,\dots,n}$ . Ein Fundamentalsystem lässt sich dann in der Form  $\{u^i(x) = e^{\lambda_i x} v^i\}_{i=1,\dots,n}$  angeben.

**Beispiel 3.1.19.**

1. Wir betrachten das DGL-System  $u' = \begin{pmatrix} 7 & -3 & 21 \\ 2 & -1 & 6 \\ -1 & 0 & -3 \end{pmatrix} u$  mit dem charakteristischen Polynom  $\chi(t) = -t(t-1)(t-2)$ . Die Eigenräume zu den Eigenwerten  $\lambda_1 = 0$ ,  $\lambda_2 = 1$  und  $\lambda_3 = 2$  sind eindimensional und Eigenvektoren sind jeweils gegeben durch  $\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$ ,  $\begin{pmatrix} -4 \\ -1 \\ 1 \end{pmatrix}$ , und  $\begin{pmatrix} -15 \\ -4 \\ 3 \end{pmatrix}$ .

Damit ergibt sich eine Lösungsbasis zu

$$\left\{ \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}, e^x \begin{pmatrix} -4 \\ -1 \\ 1 \end{pmatrix}, e^{2x} \begin{pmatrix} -15 \\ -4 \\ 3 \end{pmatrix} \right\}.$$

2. Wir betrachten das DGL-System  $u' = \begin{pmatrix} 0 & 4 & -3 \\ 1 & -3 & 3 \\ 1 & -4 & 4 \end{pmatrix} u$  mit dem charakteristischen Polynom  $\chi(t) = -(t+1)(t-1)^2$ . Die Eigenräume zu den Eigenwerten  $\lambda_1 = -1$  und  $\lambda_2 = 1$  sind eindimensional und zweidimensional und gegeben durch  $E(-1) = \text{span} \left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}$  und

$$E(1) = \text{span} \left\{ \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\}. \text{ Damit ergibt sich eine Lösungsbasis zu}$$

$$\left\{ e^{-x} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, e^x \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}, e^x \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

3. Wir betrachten das DGL-System  $u' = \begin{pmatrix} 2 & 2 & 9 \\ 1 & 0 & 3 \\ 0 & -1 & -1 \end{pmatrix} u$  mit dem charakteristischen Polynom  $\chi(t) = -(t+1)(t-1)^2$ . Die Eigenräume zu den Eigenwerten  $\lambda_1 = -1$  und  $\lambda_2 = 1$  sind beide eindimensional und Eigenvektoren sind gegeben durch  $\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$  zu  $-1$  und  $\begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}$  zu  $1$ . Zwei Abbildungen der Lösungsbasis sind damit  $e^{-x} \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$  und  $e^x \begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}$ . Eine weitere Lösung findet man durch den Polynomansatz

$$u(x) = e^x \begin{pmatrix} ax + d \\ bx + e \\ cx + f \end{pmatrix}.$$

Das Einsetzen in die DGL liefert nach Koeffizientenvergleich ein LGS mit der Lösung

$$(a, b, c, d, e, f) = \rho(-5, -2, 1, -3, -1, 0) + \sigma(0, 0, 0, -5, -2, 1)$$

und damit

$$u(x) = \rho e^x \begin{pmatrix} -5x - 3 \\ -2x - 1 \\ x \end{pmatrix} + \sigma e^x \begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}.$$

Wenn wir etwa  $\rho = 1, \sigma = 0$  wählen, liefert das eine Lösungsbasis:

$$\left\{ e^{-x} \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}, e^x \begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}, e^x \begin{pmatrix} -3 \\ -1 \\ 0 \end{pmatrix} + x e^x \begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

4. Wir betrachten das DGL-System  $u' = \begin{pmatrix} 1 & -1 & -3 \\ -5 & 6 & -15 \\ -1 & 2 & -1 \end{pmatrix} u$  mit dem charakteristischen Polynom  $\chi(t) = -(t-2)(t^2 - 4t + 13)$ . Der Eigenraum zum reellen Eigenwert  $\lambda_1 = 2$  ist eindimensional. Das gleiche gilt für die beiden komplex konjugierten Eigenwerte  $\lambda_2 = 2 + 3i$  und  $\lambda_3 = 2 - 3i$ .

Eigenvektoren sind gegeben durch  $\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$  zu 2 und  $\begin{pmatrix} -1 \pm i \\ 1 \pm 2i \\ 1 \end{pmatrix}$  zu  $2 \pm 3i$ .

Wir teilen die komplexe Abbildung

$$e^{(2+3i)x} \begin{pmatrix} -1+i \\ 1+2i \\ 1 \end{pmatrix} = e^{2x} (\cos(3x) + i \sin(3x)) \begin{pmatrix} -1+i \\ 1+2i \\ 1 \end{pmatrix}$$

in Real- und Imaginärteil auf und erhalten insgesamt eine Lösungsbasis zu

$$\left\{ e^{2x} \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}, e^{2x} \cos(3x) \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} - e^{2x} \sin(3x) \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \right. \\ \left. e^{2x} \cos(3x) \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + e^{2x} \sin(3x) \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

5. Wir betrachten das DGL-System  $u' = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{pmatrix} u$  mit dem charakteristischen Polynom  $\chi(t) = -(t-1)^3$ . Der Eigenraum zum Eigenwert  $\lambda_1 = 1$  ist eindimensional und ein Eigenvektor ist gegeben durch  $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ .

Eine Abbildung der Lösungsbasis ist damit  $e^x \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ . Weitere Lösungen findet man durch den Polynomansatz

$$u(x) = e^x \begin{pmatrix} ax^2 + dx + g \\ bx^2 + ex + h \\ cx^2 + fx + j \end{pmatrix}.$$

Das Einsetzen in die DGL liefert nach Koeffizientenvergleich ein LGS mit der Lösung

$$(a, b, c, d, e, f, g, h, j) = \rho(1, -2, 1, -2, 2, 0, 2, 0, 0)$$

$$\begin{aligned}
& + \sigma(0, 0, 0, 1, -2, 1, -1, 1, 0) \\
& + \tau(0, 0, 0, 0, 0, 0, 1, -2, 1)
\end{aligned}$$

und damit

$$u(x) = \rho e^x \begin{pmatrix} x^2 - 2x - 2 \\ -2x^2 + 2x \\ x^2 \end{pmatrix} + \sigma e^x \begin{pmatrix} x - 1 \\ -2x + 1 \\ x \end{pmatrix} + \tau e^x \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

Wenn wir etwa  $\rho = \frac{1}{2}, \sigma = \tau = 0$  und  $\rho = \tau = 0, \sigma = 1$  wählen, liefert das eine Lösungsbasis:<sup>(i)</sup>

$$\left\{ e^x \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, e^x \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x e^x \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, e^x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x e^x \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{2} x^2 e^x \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

**Bemerkung 3.1.20.** Im Fall spezieller rechter Seiten kann man auch für die Lösung eines inhomogenes linearen DGL-System mit konstanten Koeffizienten einen speziellen Ansatz wählen. Wir betrachten dazu für  $A \in M_n \mathbb{R}$  das DGL-System  $u'(x) = Au(x) + b(x)$  mit

$$b(x) = \begin{pmatrix} q_{c,1}(x) \\ \vdots \\ q_{c,n}(x) \end{pmatrix} e^{\lambda x} \cos(\omega x) + \begin{pmatrix} q_{s,1}(x) \\ \vdots \\ q_{s,n}(x) \end{pmatrix} e^{\lambda x} \sin(\omega x)$$

wobei  $m = \max\{\text{grad}(q_{s,i}), \text{grad}(q_{c,i})\}$ . Weiter sei  $\lambda + i\omega$  (und damit auch  $\lambda - i\omega$ ) ein  $k$ -facher Eigenwert von  $A$ ,  $r$  sei die Dimension des Eigenraums von  $\lambda + i\omega$  und  $s$  sei die Größe des größten Jordan-Elementarblocks zu  $\lambda + i\omega$ . Dann liefert der Ansatz

$$u(x) = \begin{pmatrix} p_{c,1}(x) \\ \vdots \\ p_{c,n}(x) \end{pmatrix} e^{\lambda x} \cos(\omega x) + \begin{pmatrix} p_{s,1}(x) \\ \vdots \\ p_{s,n}(x) \end{pmatrix} e^{\lambda x} \sin(\omega x)$$

mit Polynomen  $p_{c,i}, p_{s,i}$  vom Grad  $\leq m + s$  spezielle Lösungen des inhomogenen linearen DGL-Systems. Auch hier lässt sich  $s$  gemäß  $s \leq k - r + 1$  abschätzen.

<sup>(i)</sup>Hier und in Beispiel 3. zeigt ein Test, dass die so erhaltenen Vektoren tatsächlich eine Jordanbasis bilden, vergleiche Bemerkung 3.1.17.

**Beispiel 3.1.21.** Wir überprüfen diesen Ansatz an zwei Beispielen:

- 1)  $A = \lambda \mathbb{1}$  und die rechte Seite sei  $\vec{q}(x)e^{\lambda x}$ .
- 2)  $A \in M_n \mathbb{R}$  sei die Begleitmatrix eines Polynoms mit  $k$  facher Nullstelle  $\lambda$ ,  $k \leq n$ . Weiter sei die rechte Seite von der Form  $\vec{q}(x)e^{\lambda x}$  und  $\vec{q}(x) = (0, \dots, 0, g(x))^T$ .

zu 1) In diesem Fall schreiben wir  $u(x) = \vec{p}(x)e^{\lambda x}$ , also  $u'(x) = \vec{p}'(x)e^{\lambda x} + \lambda \vec{p}(x)e^{\lambda x}$ . Eingesetzt in das System liefert das

$$\begin{aligned} \vec{p}'(x)e^{\lambda x} + \lambda \vec{p}(x)e^{\lambda x} &= \lambda \vec{p}(x)e^{\lambda x} + \vec{q}(x)e^{\lambda x} \\ \iff \vec{p}'(x) &= \vec{q}(x). \end{aligned}$$

Es reicht also  $\text{grad}(\vec{p}) = \text{grad}(\vec{q}) + 1$  zu wählen.<sup>(i)</sup> Da in diesem Fall für den Eigenwert  $\lambda$  gerade  $k = r = n$  und  $s = 1$  gilt, entspricht das dem Ansatz aus der vorigen Bemerkung.

zu 2) Das System aus 2) ist die Übersetzung einer inhomogenen linearen DGL mit rechter Seite  $g(x)e^{\lambda x}$  deren charakteristisches Polynom gerade das der Matrix  $A$  ist. Der Ansatz für die Lösung der inhomogenen DGL lautet  $y = x^k p(x)e^{\lambda x}$  mit  $\text{grad}(p) = \text{grad}(g)$ . Wir nutzen

$$\begin{aligned} y' &= (kx^{k-1}p(x) + x^k p'(x) + \lambda x^k p(x))e^{\lambda x} \\ &=: x^{k-1} \hat{p}_1(x)e^{\lambda x} \quad \text{mit } \text{grad}(\hat{p}_1) = m + 1, \\ y'' &= ((k-1)x^{k-2} \hat{p}_1(x) + x^{k-1} \hat{p}_1'(x) + \lambda x^{k-1} \hat{p}_1(x))e^{\lambda x} \\ &=: x^{k-2} \hat{p}_2(x)e^{\lambda x} \quad \text{mit } \text{grad}(\hat{p}_2) = m + 2, \\ &\vdots \\ y^{(k)} &= \hat{p}_k(x)e^{\lambda x} \quad \text{mit } \text{grad}(\hat{p}_k) = m + r, \\ y^{(k+1)} &= \hat{p}_{k+1}(x)e^{\lambda x} \quad \text{mit } \text{grad}(\hat{p}_{k+1}) = m + r, \\ &\vdots \\ y^{(n-1)} &= \hat{p}_{n-1}(x)e^{\lambda x} \quad \text{mit } \text{grad}(\hat{p}_{n-1}) = m + r. \end{aligned}$$

Schreiben wir nun noch  $p_1 = x^k p$ ,  $p_i = x^{k-i+1} \hat{p}_{i-1}$  für  $i = 2, \dots, k$  und  $p_i = \hat{p}_{i-1}$  für  $i = k+1, \dots, n$  so ist  $\text{grad}(p_i) = m + r$  für alle

<sup>(i)</sup>  $\text{grad}(\vec{q})$  meint hier das Maximum der Grade der Einträge von  $\vec{q}$ .

$i = 1, \dots, n$ . Der Ansatz für die spezielle Lösung des Systems hat damit die Form

$$u(x) = \vec{p}(x)e^{\lambda x}$$

mit  $\vec{p} = (p_1, \dots, p_n)^T$  und  $\text{grad}(\vec{p}) = n + r$ , was in diesem Fall wegen  $s = r$  wieder dem Ansatz aus der vorigen Bemerkung entspricht.<sup>(i)</sup>

Ein Verfahren zur Bestimmung einer Lösung eines inhomogenen linearen DGL-Systems erster Ordnung liefert die *Variation der Konstanten*:

**Satz 3.1.22.** *Es sei  $u'(x) = A(x)u(x) + b(x)$  ein inhomogenes lineares System und  $U(x)$  eine Fundamentalmatrix des zugehörigen linearen DGL-System. Dann liefert der Ansatz*

$$u(x) := U(x)c(x)$$

mit  $c : I \rightarrow \mathbb{R}^n$  eine Lösung des inhomogenen DGL-Systems. Die Abbildung  $c$  erhält man durch Integration der Gleichung

$$c'(x) = U(x)^{-1}b(x).$$

**Folgerung 3.1.23.** Im Fall linearer DGL-Systeme mit konstanten Koeffizienten ist die Lösung tatsächlich auf dem ganzen Definitionsintervall  $I$  der Inhomogenität  $b(x)$  definiert, siehe auch Folgerung 2.7.24.3.

### 3.2 Das Jordan-Maß und messbare Mengen

In diesem Kapitel werden wir uns mit dem Volumen von Teilmengen des  $\mathbb{R}^n$  beschäftigen. Dabei machen wir von einem intuitiven Begriff von Volumina - nämlich dem von Quadern - Gebrauch.

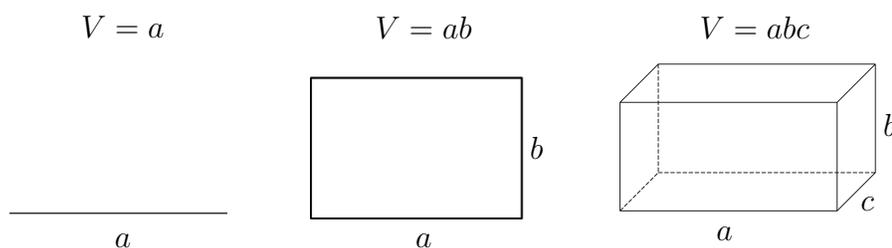
---

<sup>(i)</sup>Für eine Begleitmatrix gilt für jeden Eigenwert stets  $r = s$ .

## 3.2.1 Quaderapproximation

Wir haben ein intuitives Verständnis für einen Begriff von Volumen von Quadern. Im Folgenden nehmen wir diesen als Definition.

Abbildung 3.2.1: Das Volumen von Quadern in Dimension 1, 2 und 3



**Definition 3.2.1.** 1. Ein Standardquader  $Q_0$  im  $\mathbb{R}^n$  ist ein achsenparalleles Parallelepipiped. Bezeichnet  $\{e_j\}_{j=1,\dots,n}$  die Standardbasis des  $\mathbb{R}^n$ , so wird  $Q_0$  von den paarweise orthogonalen Vektoren  $\{\alpha_i e_i\}_{i=1,\dots,n}$  aufgespannt, wobei  $\alpha_1, \dots, \alpha_n \in \mathbb{R}^{>0}$ . Das Volumen eines solchen Quaders ist

$$V(Q_0) := \alpha_1 \alpha_2 \cdot \dots \cdot \alpha_n.$$

2. Ein *Quader* ist ein allgemeiner achsenparalleles Parallelepipiped. Er entsteht aus einem Standardquader durch eine Translation. Ist  $Q = v + Q_0$  eine solcher Quader, dann ist  $V(Q) := V(Q_0)$ . Ein *Würfel* ist ein Quader, dessen Seiten alle gleiche Länge haben. Ein Standardwürfel ist also durch die Angabe einer positiven reellen Zahl, seiner *Seitenlänge*, eindeutig festgelegt.
3. Es sei  $Q = v + Q_0$  der Quader, der durch  $v = (v_1, \dots, v_n) \in \mathbb{R}^n$  und den Standardquader  $Q_0$ , der von  $\{\alpha_i e_i\}_{i=1,\dots,n}$  aufgespannt wird, definiert wird.  $Q$  entspricht der Punktmenge

$$\begin{aligned} Q &= \{x \in \mathbb{R}^n \mid x_i = v_i + t\alpha_i, t \in [0, 1], i = 1, \dots, n\} \\ &= \{x \in \mathbb{R}^n \mid v_i \leq x_i \leq v_i + \alpha_i, i = 1, \dots, n\}. \end{aligned}$$

Die  $2^n$  Punkte

$$\{p \mid p_i = v_i + t\alpha_i, t = 0, 1, i = 1, \dots, n\}$$

heißen *Eckpunkte* des Quaders.

Wir wollen uns nicht alle Quader oder Würfel anschauen, sondern etwas speziellere.

**Definition 3.2.2.** Ein *Elementarwürfel der Ordnung  $r$*  ist ein Würfel der Seitenlänge  $\frac{1}{2^r}$ , der durch Verschieben des Standardwürfels der gleichen Seitenlänge um den Vektor  $\frac{1}{2^r}(a_1, a_2, \dots, a_n)$  entsteht, wobei  $a_i \in \mathbb{Z}$  für alle  $i = 1, \dots, n$ . Für  $\vec{a} = (a_1, \dots, a_n) \in \mathbb{Z}^n$  bezeichnen wir diesen Elementarwürfel mit

$$W_{r,\vec{a}} = \left\{ x \in \mathbb{R}^n \mid \frac{a_i}{2^r} \leq x_i \leq \frac{a_i + 1}{2^r}, i = 1, \dots, n \right\}.$$

Es ist  $V(W_{r,\vec{a}}) = 2^{-rn}$ .

**Definition/Bemerkung 3.2.3.** 1. Die Elementarwürfel haben die folgenden Eigenschaften

(a)  $W_{r,\vec{a}} \cap W_{r,\vec{b}} = \emptyset \iff \max_{1 \leq i \leq n} \{|a_i - b_i|\} \geq 2$

(b) Es sei  $\max_{1 \leq i \leq n} \{|a_i - b_i|\} = 1$  also  $\vec{b} = \vec{a} + e_{i_1} + \dots + e_{i_\ell}$  für  $\ell > 0$  und  $\{i_1, \dots, i_\ell\} \subset \{1, \dots, n\}$ . Dann ist  $W_{r,\vec{a}} \cap W_{r,\vec{b}}$  (bei geeigneter Interpretation) ein Elementarwürfel der Ordnung  $r$  in Dimension  $n - \ell$ . Im Spezialfall  $\ell = n$  treffen sich  $W_{r,\vec{a}}$  und  $W_{r,\vec{b}}$  in genau einem Punkt.

2. Ist  $M \subset \mathbb{R}^n$  eine Teilmenge, dann ist eine Würfelüberdeckung von  $M$  der Ordnung  $r$  eine Teilmenge  $\mathcal{W} \subset \{W_{r,\vec{a}} \mid \vec{a} \in \mathbb{Z}^n\}$  mit  $M \subset \bigcup \mathcal{W}$ .

**Bemerkung 3.2.4.** 1. Der Rand  $\partial W_{r,\vec{a}}$  eines Elementarwürfels besteht (bei geeigneter Interpretation) aus der Vereinigung von Elementarwürfeln  $r$ -ter Ordnung in Dimension  $n - 1$ .

2. Alle Elementarwürfel  $r$ -ter Ordnung bilden eine Würfelüberdeckung des  $\mathbb{R}^n$ . <sup>(i)</sup>

3. Die  $2^n$ -elementige Menge  $\{W_{r+1,2\vec{a}+\vec{v}} \mid v_i \in \{0, 1\}, i = 1, \dots, n\}$  ist eine Würfelüberdeckung der Ordnung  $r + 1$  von  $W_{r,\vec{a}}$  mit

$$W_{r,\vec{a}} = \bigcup \{W_{r+1,2\vec{a}+\vec{v}} \mid v_i \in \{0, 1\}, i = 1, \dots, n\}.$$

Diese Überdeckung erhält man durch "Halbieren der Kanten" von  $W_{r,\vec{a}}$ .

<sup>(i)</sup>Eine Würfelüberdeckung ist insbesondere keine offene Überdeckung!

**Bezeichnung 3.2.5.** Es sei  $M \subset \mathbb{R}^n$  eine Menge und  $r \in \mathbb{N}$  eine feste Zahl.

- Mit  $\mathcal{W}^r(M)$  bezeichnen wir die Menge aller Elementarwürfel der Ordnung  $r$ , die einen nichtleeren Schnitt mit  $\bar{M}$  haben.
- Mit  $\mathcal{W}_r(M)$  bezeichnen wir die Menge aller Elementarwürfel der Ordnung  $r$ , die ganz in  $\overset{\circ}{M}$  enthalten sind.

Es ist  $\mathcal{W}_r(M) \subset \mathcal{W}^r(M)$ .

**Beispiel 3.2.6.** Wir betrachten  $B_R(0, 0) \subset \mathbb{R}^2$ . Für diese Mengen gilt

$$\mathcal{W}^0(B_1(0, 0)) = \left\{ W_{0, \binom{0}{0}}, W_{0, \binom{1}{0}}, W_{0, \binom{-1}{0}}, W_{0, \binom{-2}{0}}, W_{0, \binom{0}{-1}}, W_{0, \binom{1}{-1}}, \right. \\ \left. W_{0, \binom{-1}{-1}}, W_{0, \binom{-2}{-1}}, W_{0, \binom{-1}{1}}, W_{0, \binom{0}{1}}, W_{0, \binom{-1}{2}}, W_{0, \binom{0}{2}} \right\}$$

und allgemeiner für  $N \in \mathbb{N}$  und  $N - 1 < R \leq N$

$$\mathcal{W}^0(B_R(0, 0)) = \{W_{0, \vec{a}} \mid \|\vec{a} + \vec{1}\| \leq R, \text{ wobei} \\ \nu_i = 1 \text{ falls } a_i < 0 \text{ und } \nu_i = 0 \text{ falls } a_i \geq 0\} .$$

**Bemerkung 3.2.7.** 1.  $\mathcal{W}^r(M)$  ist eine Würfelüberdeckung der Ordnung  $r$  von  $M$ . Wie das obige Beispiel zeigt, ist  $\mathcal{W}^r(M)$  in der Regel nicht die "kleinste" Würfelüberdeckung von  $M$  der Ordnung  $r$ .<sup>(i)</sup>

2. Für eine beschränkte Menge  $M \subset \mathbb{R}^n$  sind  $\mathcal{W}_r(M)$  und  $\mathcal{W}^r(M)$  endliche Mengen.

**Definition 3.2.8.** Für eine Menge  $M \subset \mathbb{R}^n$  seien die Mengen  $\mathcal{W}^r(M)$  und  $\mathcal{W}_r(M)$  endlich. Dann definieren wir

$$\mu_r(M) := \frac{1}{2^{nr}} \#\mathcal{W}_r(M) \quad \text{und} \quad \mu^r(M) := \frac{1}{2^{nr}} \#\mathcal{W}^r(M) .$$

$\mu_r(M)$  und  $\mu^r(M)$  heißen *inneres* und *äußeres  $r$ -Maß* von  $M$ .

**Bemerkung 3.2.9.** Es seien  $M, N \subset \mathbb{R}^n$  Mengen für die  $\mu_r(M)$  und  $\mu^r(M)$  existieren. Dann existieren auch weitere  $r$ -Maße und für diese gilt:

<sup>(i)</sup>Eine Würfelüberdeckung  $\mathcal{W}'$  heißt "die kleinste" wenn folgende Eigenschaft erfüllt ist: Für jede weitere Würfelüberdeckung  $\mathcal{W}$  von  $M$  der Ordnung  $r$  gilt  $\mathcal{W}' \subset \mathcal{W}$ .

1. Es ist  $\mu^r(M \cup N) \leq \mu^r(M) + \mu^r(N)$
2. Gilt  $M \cap N = \emptyset$ , so ist  $\mu_r(M \cup N) \geq \mu_r(M) + \mu_r(N)$ .
3. Gilt  $M \subset N$ , so ist  $\mu_r(M) \leq \mu_r(N)$  und  $\mu^r(M) \leq \mu^r(N)$ .
4. Für den Rand gilt  $\mu^r(\partial M) = \mu^r(M) - \mu_r(M)$ .

**Satz 3.2.10.** *Es sei  $M \subset \mathbb{R}^n$  eine beschränkte Menge. Dann gilt:*

1.  $\mu_{r+1}(M) \geq \mu_r(M)$  für alle  $r \in \mathbb{N}$
2.  $\mu^{r+1}(M) \leq \mu^r(M)$  für alle  $r \in \mathbb{N}$
3. Es gibt eine von  $r$  unabhängige Zahl  $\nu(M)$ , sodass  $0 \leq \mu_r(M) \leq \mu^r(M) \leq \nu(M)$  für alle  $r$ .

### 3.2.2 Messbare Teilmengen des $\mathbb{R}^n$ und das Jordan-Maß

Wegen Satz 3.2.10 sind für eine beschränkte Menge  $M$  die Zahlenfolgen  $(\mu_r(M))_{r \in \mathbb{N}}$  und  $(\mu^r(M))_{r \in \mathbb{N}}$  monoton und beschränkt. Das motiviert nun die folgende Definition.

**Definition/Satz 3.2.11.** *Es sei  $M \subset \mathbb{R}^n$  eine beschränkte Menge. Dann heißen die existierenden Grenzwerte*

$$\mu^\infty(M) := \lim_{r \rightarrow \infty} \mu^r(M) \quad \text{und} \quad \mu_\infty(M) := \lim_{r \rightarrow \infty} \mu_r(M)$$

das innere und äußere Jordan-Maß von  $M$ .

**Bemerkung 3.2.12.** Die Eigenschaften aus Bemerkung 3.2.9.1-3 gelten für  $r = \infty$  analog.

**Definition 3.2.13.** Es sei  $M \subset \mathbb{R}^n$  eine beschränkte Menge.

1.  $M$  heißt (*Jordan-*)messbar, wenn  $\mu^\infty(M) = \mu_\infty(M)$  und die so definierte Zahl  $\mu(M)$  heißt das Jordan-Maß von  $M$ .
2.  $M$  heißt Nullmenge, wenn  $\mu(M) = 0$ .

**Bemerkung 3.2.14.** 1. Sind  $M \subset N$  messbar, so gilt  $\mu(M) \leq \mu(N)$ .

2. Teilmengen, endliche Vereinigungen und beliebige Schnitte von Nullmengen sind Nullmengen.
3. Endliche Mengen sind Nullmengen.
4. Ist  $(x_n)$  eine beschränkte Folge in  $\mathbb{R}^n$  mit endlich vielen Häufungspunkten, dann ist  $M = \{x_n \mid n \in \mathbb{N}\}$  eine Nullmenge. Das ist insbesondere der Fall, wenn  $(x_n)$  eine konvergente Folge ist.
5.
  - Ist  $M$  messbar, so ist  $\partial M$  eine Nullmenge.
  - Ist  $M$  zusätzlich beschränkt so gilt auch die Umkehrung: Eine beschränkte Menge  $M$  ist genau dann messbar, wenn  $\partial M$  eine Nullmenge ist.
6. Sind  $M, N$  messbar, so auch  $M \cup N$ ,  $M \cap N$  und  $M \setminus N$ . Insbesondere gilt  $\mu(M \cup N) \leq \mu(M) + \mu(N)$ .
7. Für die Vereinigung  $M \cup N$  zweier messbarer Mengen mit  $M \cap N = \emptyset$  gilt  $\mu(M \cup N) = \mu(M) + \mu(N)$ .
8. Ist  $M$  messbar und  $N$  eine Nullmenge so gilt  $\mu(M) = \mu(M \cup N)$ .
9. Ist  $M$  messbar,  $v \in \mathbb{R}^n$  und  $N = v + M$  eine Translation der Menge  $M$ , dann gilt  $\mu(N) = \mu(M)$ .

Für den letzten Punkt der vorigen Bemerkung benötigen wir das folgende elementare, aber sehr wichtige Ergebnis. Es zeigt insbesondere, dass das Jordan-Maß einen sinnvollen Volumenbegriff definiert.

**Bemerkung 3.2.15.** • Ist  $Q \subset \mathbb{R}^n$  ein Quader, so gilt  $\mu(Q) = V(Q)$ .

- Insbesondere sind die Seiten eines Quaders Nullmengen. Damit ist eine Menge  $M$ , die eine Zerlegung in Quader der Form  $M = Q_1 \cup \dots \cup Q_k$  besitzt mit  $\overset{\circ}{Q}_i \cap Q_j = \emptyset$ , messbar und es ist

$$\mu(M) = \sum_{i=1}^k \mu(Q_i).$$

- Zum Beispiel ist die Menge  $W := \bigcup \mathcal{W}^0(B_1(0,0)) + (\sqrt{2}, \sqrt{3})$  vom obigen Typ und es ist  $\mu(W) = 12$ .

**Beispiel 3.2.16.** Die Menge  $M = [0, 1]^n \cap \mathbb{Q}^n$  hat das untere Jordan-Maß  $\mu_\infty(M) = 0$  und das obere Jordan-Maß  $\mu^\infty(M) = 1$ , sie ist also nicht messbar.

Wie das obige Beispiel zeigt, ist die Messbarkeit bezüglich des Jordan-Maßes in der Regel nicht auf unendliche Vereinigungen erweiterbar:  $M$  ist eine abzählbare Vereinigung von Punktmengen, nämlich rationaler Zahlen. Diese haben jeweils das Maß 0, sodass die Summe der Maße ebenfalls verschwindet; andererseits ist die Vereinigung  $M$  dieser Punkte selbst nicht messbar. Es gilt aber

**Satz 3.2.17.** *Ist  $M = \bigcup_{i \in \mathbb{N}} M_i$  eine abzählbare Vereinigung messbarer Mengen derart, dass  $M_i \cap M_j$  für verschiedenen  $i, j \in \mathbb{N}$  eine Nullmenge ist, und ist  $M$  selbst messbar, dann gilt*

$$\mu(M) = \sum_{i=1}^{\infty} \mu(M_i).$$

### 3.2.3 Das Verhalten des Jordan-Maßes unter Abbildungen

Dass das Maß einer messbaren Menge nach Anwenden einer Abbildung beliebige Werte annehmen kann, zeigt bereits das Beispiel  $f_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}^n$  mit  $f_\alpha(x) = \alpha x$  für  $\alpha \in \mathbb{R}$ . Dann ist

$$\mu\left(f_\alpha([0, 1]^n)\right) = |\alpha|^n.$$

Als weiteres Beispiel hatten wir schon gesehen, dass das Maß jedoch invariant unter Translationen ist.

Eine wichtige Eigenschaft ist die folgende, die Aussagen über Bilder "niedrigdimensionaler" Mengen macht.

**Satz 3.2.18.** *Es seien  $m < n$ , die Menge  $M \subset \mathbb{R}^m$  messbar im  $\mathbb{R}^m$  und  $f : M \rightarrow \mathbb{R}^n$  Lipschitz-stetig. Dann ist  $f(M)$  eine Nullmenge im  $\mathbb{R}^n$ . Das ist insbesondere der Fall, wenn  $M$  kompakt und messbar und  $f$  stetig differenzierbar ist.*

Hiermit und mit Hilfe von Folgerung 3.2.14.6 kann man nun ganze Klassen von Teilmengen des  $\mathbb{R}^n$  als messbar identifizieren.

**Bemerkung 3.2.19.** 1. Beschränkte Teilmengen echter affiner Unterräume des  $\mathbb{R}^n$  sind Nullmengen.

2. Eine Menge  $M \subset \mathbb{R}^n$ , deren Rand sich zerlegen lässt als  $\partial M = \bigcup_{i=1}^k K_i$  mit  $K_i = g_i(N_i)$  wobei

- $N_i \subset \mathbb{R}^{n_i}$  kompakt und messbar,
- $n_i < n$  und
- $g_i$  stetig differenzierbar,

ist messbar.

3. Alle Bälle, Vollellipsoiden und alle Polyeder des  $\mathbb{R}^3$  sind messbar, genauso wie ihre Verallgemeinerungen in andere Dimensionen.

Ein wichtige Klasse von Abbildungen sind die linearen Abbildungen.

**Satz 3.2.20.** *Es sei  $A \in M_n\mathbb{R}$ ,  $b \in \mathbb{R}^n$  und  $f_{A,b}(x) := Ax + b$  die durch  $A$  und  $b$  vermittelte affine Abbildung. Weiter sei  $M \subset \mathbb{R}^n$  messbar. Dann gilt*

$$\mu(f_{A,b}(M)) = |\det(A)| \mu(M).$$

**Beispiel 3.2.21.** Das von den linear unabhängigen Vektoren  $v^1, \dots, v^n \in \mathbb{R}^n$  aufgespannte Parallelepipid  $P$  hat das Maß

$$\mu(P) = |\det(v^1, \dots, v^n)|.$$

Es stimmt somit mit dem in der linearen Algebra eingeführten Volumen überein.

### 3.3 Integration im $\mathbb{R}^n$

Bevor wir zur Integration in mehreren Dimensionen kommen, werden wir eine weitere Darstellung der Riemann-Integration in einer Dimension diskutieren. Für die im Folgenden verwendeten Bezeichnungen sei auf Kapitel 2.2.3 verwiesen.

### 3.3.1 Das Darboux-Integral in $\mathbb{R}$

Für eine reelle, beschränkte Funktion  $f$  auf einem Intervall  $[a, b]$  betrachten wir zu einer Zerlegung  $Z = \{x_0, x_1, \dots, x_k\} \subset \mathbb{R}$  mit  $a = x_0 < x_1 < \dots < x_{k-1} < x_k = b$  nicht beliebige Zwischenvektoren  $\vec{\xi} = \{\xi_1, \dots, \xi_k\}$  und deren Funktionswerte  $f(\xi_i)$ , um eine Summe zu definieren, sondern wir schauen uns lediglich die "größten" und "kleinsten" Werte auf den Teilintervallen  $[x_{j-1}, x_j]$  an. Da wir die Stetigkeit von  $f$  nicht voraussetzen wollen, müssen diese nicht existieren und wir können lediglich von Supremum und Infimum sprechen.

Für  $j = 1, \dots, k$  setzen wir  $h_j = x_j - x_{j-1}$ , sodass  $\sum_{j=1}^k h_j = b - a$ .

**Definition 3.3.1.** Es sei  $f : [a, b] \rightarrow \mathbb{R}$  beschränkt und  $Z = \{x_0, x_1, \dots, x_k\}$  eine Zerlegung des Intervalls  $[a, b]$ . Weiter sei zu dieser Zerlegung  $h_j$  wie oben und

$$m^j(f) := \sup\{f(x) \mid x_{j-1} \leq x \leq x_j\},$$

$$m_j(f) := \inf\{f(x) \mid x_{j-1} \leq x \leq x_j\}.$$

Dann heißen

$$O_f(Z) := \sum_{j=1}^k m^j(f)h_j \quad \text{und} \quad U_f(Z) := \sum_{j=1}^k m_j(f)h_j$$

*Obersumme* und *Untersumme* von  $f$  bezüglich  $Z$ .

**Bemerkung 3.3.2.** Es sei  $f : [a, b] \rightarrow \mathbb{R}$  eine beschränkte Funktion und  $Z, \hat{Z}$  seien Zerlegungen von  $[a, b]$ . Dann gilt

1.  $O_f(Z) \geq U_f(Z)$
2.  $O_f(Z') \leq O_f(Z)$ ,  $U_f(Z') \geq U_f(Z)$  und  $O_f(Z') - U_f(Z') \leq O_f(Z) - U_f(Z)$  für jede Verfeinerung  $Z'$  von  $Z$
3.  $U_f(Z) \leq O_f(\hat{Z})$
4.  $\sup_Z U_f(Z) \leq \inf_Z O_f(Z)$

**Definition/Bemerkung 3.3.3.** Es sei  $f : [a, b] \rightarrow \mathbb{R}$  beschränkt.

1. Die Werte

$$\int_a^{b*} f(x) dx := \inf_Z O_f(Z) \quad \text{und} \quad \int_{a*}^b f(x) dx := \sup_Z U_f(Z)$$

heien *oberes* und *unteres Darboux-Integral*.

2.  $f$  heit *Darboux-integrierbar*, wenn  $\int_a^{b*} f(x) dx = \int_{a*}^b f(x) dx$ . Dieser Wert heit das *Darboux-Integral* und wir schreiben

$$\int_a^b f(x) dx.$$

**Satz 3.3.4.** *Eine beschrnkte Funktion ist genau dann Darboux-integrierbar, wenn gilt*

$$\forall \epsilon > 0 \exists Z : O_f(Z) - U_f(Z) < \epsilon.$$

Wir knnen nun die Darboux-Integration mit der Riemann-Integration vergleichen. Es zeigt sich, dass wir keinen neuen Integrationsbegriff bekommen, sondern "nur" die Riemann-Integration wiederfinden.

**Satz 3.3.5.** *Die Riemann-integrierbaren Funktionen auf einem Intervall  $[a, b]$  sind genau die Darboux-integrierbaren und wir schreiben*

$$\int_a^b f(x) dx = \int_a^b f(x) dx = \int_a^b f(x) dx.$$

Der Vorteil ist, dass wir ein handliches Verfahren zur Bestimmung des Integrals mit Hilfe der Ober- und Untersummen erhalten.

**Satz 3.3.6.** *Wir knnen die Integrierbarkeit mit Hilfe von Zerlegungsnullfolgen charakterisieren. Diese knnen wegen der Vertrglichkeit unter Verfeinerung stets als Zerlegungsnullfolgen angenommen werden.*

1. Ist  $f : [a, b] \rightarrow \mathbb{R}$  integrierbar, so gibt es eine Zerlegungsfolge  $(Z_n)$  von  $[a, b]$ , sodass

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} U_f(Z_n) = \lim_{n \rightarrow \infty} O_f(Z_n).$$

2. Es sei  $f : [a, b] \rightarrow \mathbb{R}$  eine beschränkte Funktion. Weiter sei  $(Z_n)$  eine Folge von Zerlegungen des Intervalls  $[a, b]$ , sodass die beiden Grenzwerte  $\lim_{n \rightarrow \infty} U_f(Z_n)$  und  $\lim_{n \rightarrow \infty} O_f(Z_n)$  existieren. Stimmen diese überein, so ist  $f$  integrierbar und es gilt

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} U_f(Z_n) = \lim_{n \rightarrow \infty} O_f(Z_n).$$

### 3.3.2 Integration im $\mathbb{R}^n$

Eine mehrdimensionale Variante der Riemann-Integration ist ohne Weiteres formulierbar: In den eindimensionalen Definitionen lassen sich etwa Intervalle durch Quader (also kartesische Produkte von Intervallen) und Zerlegungen entsprechend durch kartesische Produkte von Zerlegungen ersetzen. Die Zwischenvektoren werden zu Vektoren deren Einträge selbst Vektoren sind. Dann erhält man für Quadergebiete analoge Aussagen, wie im Eindimensionalen und vor allem eine mehrdimensionale Variante von Folgerung 3.3.6 bleibt für Quadergebiete gültig. Um von Quadern auf allgemeinere Definitionsmengen zu schließen (z.B. auf Vereinigungen von Quadern als Verallgemeinerung der Vereinigung von Intervallen) nutzt man die folgende Konstruktion: Man betrachte eine beschränkte Teilmenge  $M \subset \mathbb{R}^n$  und eine Funktion  $f : M \rightarrow \mathbb{R}$ . Die Funktion  $f$  besitzt eine Erweiterung  $f_M : \mathbb{R}^n \rightarrow \mathbb{R}$ , die insbesondere auf jedem Quader  $Q$  mit  $M \subset Q$  erklärt ist. Sie ist definiert durch

$$f_M(x) := \begin{cases} f(x) & \text{falls } x \in M \\ 0 & \text{falls } x \notin M \end{cases}.$$

Hat man nun so eine Definition des Integrals über Quader, so definiert man das "Integral von  $f$  über  $M$ " als das "Integral von  $f_M$  über  $Q$ ".

**Definition/Bemerkung 3.3.7.** 1. Für  $M \subset \mathbb{R}^n$  definieren wir die *charakteristische Funktion*  $\chi_M : \mathbb{R}^n \rightarrow \mathbb{R}$  von  $M$  durch

$$\chi_M(x) := \begin{cases} 1 & \text{falls } x \in M \\ 0 & \text{falls } x \notin M \end{cases}$$

Die charakteristische Funktion ist in allen Punkten  $x \notin \partial M$  stetig.

2. Ist  $f : M \rightarrow \mathbb{R}$  eine Funktion mit Definitionsbereich  $M \subset \mathbb{R}^n$ , dann schreiben wir die Erweiterung  $f_M : \mathbb{R}^n \rightarrow \mathbb{R}$  von  $f$  als Produkt  $f_M = \chi_M f$ . Diese Schreibweise ist sinnvoll, wenn wir vereinbaren, dass das Produkt einer Funktion an einer Stelle, wo diese nicht definiert ist, mit 0 verschwindet.

Wir werden die mehrdimensionale Integration hier ähnlich zur eindimensionalen einführen. Wir werden jedoch, motiviert durch Folgerung 3.3.6, zur Definition direkt spezielle Zerlegungsfolgen. Wir beschränken uns in der Begründung dafür, dass die Konstruktion das obige liefert darauf, zu bemerken, dass unsere Definition im Fall einer Dimension das bereits Bekannte liefert.

Wir gehen im Folgenden wie im vorigen Abschnitt vor. Wir ersetzen dabei im Wesentlichen die Intervalle durch Würfel/Quader und die Intervalllänge durch das Maß der Würfel/Quader.

**Definition/Bemerkung 3.3.8.** Es sei  $f : M \rightarrow \mathbb{R}$  eine reelle, beschränkte Funktion auf der beschränkten Menge  $M \subset \mathbb{R}^n$ . Weiter sei  $W_{r,\vec{a}}$  ein Elementarwürfel. Damit sei

$$m^{r,\vec{a}}(f) := \sup \{f_M(x) \mid x \in W_{r,\vec{a}}\},$$

$$m_{r,\vec{a}}(f) := \inf \{f_M(x) \mid x \in W_{r,\vec{a}}\},$$

und die endlichen Summen

$$O_f(r) := \sum_{\vec{a} \in \mathbb{Z}^n} m^{r,\vec{a}}(f) \cdot \frac{1}{2^{rn}},$$

$$U_f(r) := \sum_{\vec{a} \in \mathbb{Z}^n} m_{r,\vec{a}}(f) \cdot \frac{1}{2^{rn}}.$$

Dann gilt Bemerkung 3.3.2 sinngemäß und wir definieren das *obere* und *untere* (Riemann-)Integral als

$$\int_M^* f(x) d^n x := \lim_{r \rightarrow \infty} O_f(r) \quad \text{und} \quad \int_{M^*} f(x) d^n x := \lim_{r \rightarrow \infty} U_f(r).$$

Wir nennen  $f$  integrierbar, wenn oberes und unteres Integral übereinstimmen und schreiben dann

$$\int_M f(x) d^n x := \int_M^* f(x) d^n x = \int_{M^*} f(x) d^n x.$$

**Beispiel 3.3.9.** Ist  $M = [a, b] \subset \mathbb{R}$  und  $f : [a, b] \rightarrow \mathbb{R}$  integrierbar, dann gilt

$$\int_{[a,b]} f(x) d^1x = \int_a^b f(x) dx$$

Ebenso wie in Satz 3.3.4 haben wir das folgende Integrationskriterium

**Satz 3.3.10.** Die beschränkte Funktion  $f : M \rightarrow \mathbb{R}^n$  ist auf der beschränkten Menge  $M \subset \mathbb{R}^n$  genau dann integrierbar, wenn

$$\forall \epsilon > 0 \exists r_0 \in \mathbb{N} \forall r > r_0 : O_f(r, \vec{a}) - U_f(r, \vec{a}) < \epsilon.$$

Dass eine große Menge von Funktionen integrierbar ist zeigt der folgende Satz

**Satz 3.3.11.** Es sei  $f : M \rightarrow \mathbb{R}$  eine beschränkte Funktion auf der beschränkten Menge  $M \subset \mathbb{R}^n$ .

1. Verschwindet  $f$  auf  $M$  mit Ausnahme einer Nullmenge, so gilt

$$\int_M f(x) d^n x = 0$$

2. Ist  $f$  in allen Punkten von  $M$  mit Ausnahme einer Nullmenge stetig, so ist  $f$  integrierbar.

Das Integral hat die üblichen Eigenschaften.

**Bemerkung 3.3.12.** In den folgenden Punkten genügen die Mengen  $M, N$  und die Funktionen  $f, g$  den notwendigen Voraussetzungen.

1. Für alle  $\alpha \in \mathbb{R}$  ist

$$\int_M (f(x) + \alpha g(x)) d^n x = \int_M f(x) d^n x + \alpha \int_M g(x) d^n x$$

2. Ist  $f(x) \geq 0$  für alle  $x \in M$ , so gilt  $\int_M f(x) d^n x \geq 0$

3.  $\left| \int_M f(x) d^n x \right| \leq \int_M |f(x)| d^n x$

4. Ist  $M \cap N$  eine Nullmenge, so gilt

$$\int_{M \cup N} f(x) d^n x = \int_M f(x) d^n x + \int_N f(x) d^n x.$$

Ebenfalls gilt die folgende zu erwartende Aussage

**Satz 3.3.13.** *Ist  $M \subset \mathbb{R}^n$  messbar, so ist*

$$\mu(M) = \int_M \chi_M(x) d^n x.$$

*Wir schreiben auch  $\mu(M) = \int_M d^n x$ .*

**Folgerung 3.3.14** (Mittelwertsatz der Integralrechnung). *Es sei  $M \subset \mathbb{R}^n$  messbar und  $f : M \rightarrow \mathbb{R}$  integrierbar. Weiter seien  $c := \inf\{f(x) \mid x \in M\}$  und  $C := \sup\{f(x) \mid x \in M\}$ .*

1. Unter diesen Voraussetzungen gilt  $c\mu(M) \leq \int_M f(x) d^n x \leq C\mu(M)$ .
2. Nimmt  $f$  jeden Wert zwischen  $c$  und  $C$  an, so gibt es ein  $\xi \in M$  mit

$$\int_M f(x) d^n x = f(\xi)\mu(M).$$

Zur praktischen Berechnung von Integralen ist der folgende Satz fundamental. Er erlaubt es uns das  $n$ -dimensionale Integral unter leichten Voraussetzungen auf das eindimensionale zurückzuführen.

In den nächsten beiden Sätzen nutzen wir die folgende Bezeichnung: Mit  $M_1 \subset \mathbb{R}^p$ ,  $M_2 \subset \mathbb{R}^q$  sei  $M = M_1 \times M_2 \subset \mathbb{R}^n$ . Die Punkte aus  $M_1$ ,  $M_2$  und  $M$  bezeichnen wir mit  $x$ ,  $y$  und  $z = (x, y)$ . Ist  $f : M \rightarrow \mathbb{R}$  eine Funktion, so schreiben wir  $f(z) = f(x, y)$  für  $z = (x, y) \in M = M_1 \times M_2$ .

**Satz 3.3.15** (Satz von Fubini). *Es seien  $M_1 \subset \mathbb{R}^p$  und  $M_2 \subset \mathbb{R}^q$  beschränkte Mengen. Auf der Menge  $M = M_1 \times M_2 \subset \mathbb{R}^n$  betrachten wir die Funktion  $f : M \rightarrow \mathbb{R}$ . Ist  $f$  integrierbar so gilt*

$$\begin{aligned} \int_M f(z) d^n z &= \int_{M_1} \left( \int_{M_2}^* f(x, y) d^q y \right) d^p x = \int_{M_1} \left( \int_{M_2^*} f(x, y) d^q y \right) d^p x \\ &= \int_{M_2} \left( \int_{M_1}^* f(x, y) d^p x \right) d^q y = \int_{M_2} \left( \int_{M_1^*} f(x, y) d^p x \right) d^q y. \end{aligned}$$

Wie der Beweis von Satz 3.3.15 zeigt, brauchen wir die Gleichheit der (existierenden!) inneren Integrale nicht vorauszusetzen. Anwenden werden wir den Satz in der Regel jedoch in der folgenden schärferen Variante, die den gleichen Namen tragen soll.

**Folgerung 3.3.16** (Satz von Fubini – Variante). Es gelten die Bezeichnungen aus Satz 3.3.15. Ist  $f : M \rightarrow \mathbb{R}$  integrierbar und existiert das Teilintegral

$$\int_{M_2} f(x, y) d^q y \text{ so gilt}^{(i)}$$

$$\int_M f(z) d^n z = \int_{M_1} \left( \int_{M_2} f(x, y) d^q y \right) d^p x.$$

Dass  $M$  die Form eines Produktes hat, ist keine wirkliche Einschränkung, wie die folgende Bemerkung zeigt.

**Bemerkung 3.3.17** (Praktische Anwendung des Satzes von Fubini). Es seien  $M \subset \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$  und  $M_1 := \pi_1(M) \subset \mathbb{R}^p$ ,  $M_2 = \pi_2(M) \subset \mathbb{R}^q$ . Weiter sei für  $x \in M_1$  die Menge  $M^x \subset \mathbb{R}^q$  definiert als

$$M^x := \{y \in \mathbb{R}^q \mid (x, y) \in M\}.$$

Dann ist  $M_2 = \bigcup_{x \in M_1} M^x$ . Ist nun  $x \in M_1$  fest gewählt, so gilt

$$\chi_{M_2} f(x, \cdot) = \chi_{M^x} f(x, \cdot).$$

Damit ist mit  $z = (x, y)$

$$\int_M f(z) d^n z = \int_{M_1} \left( \int_{M^x} f(x, y) d^q y \right) d^p x.$$

Siehe dazu Abbildung 3.3.1.

---

<sup>(i)</sup>Liegt eine Zerlegung des Punktes  $z \in \mathbb{R}^n$  in der Form  $z = (x, y)$  mit  $x \in \mathbb{R}^p, y \in \mathbb{R}^q$  vor, so schreiben wir auch  $d^n z = d^n(x, y)$ .



$$\begin{aligned}
&= \int_{[-1,1]^2} \sqrt{2-x^2-y^2} d^2(x,y) \\
&= \int_{-1}^1 \left( \int_{-1}^1 \sqrt{2-x^2-y^2} dy \right) dx \\
&= \int_{-1}^1 \left( \sqrt{1-x^2} + (2-x^2) \arcsin \left( \frac{1}{\sqrt{2-x^2}} \right) \right) dx \\
&= 2 \cdot \left( \frac{1}{4}\pi + \frac{7}{12}\pi - \frac{2\sqrt{2}}{3}\pi \right) = \frac{5}{3}\pi + \frac{4\sqrt{2}}{3}\pi
\end{aligned}$$

Dabei haben wir von der vierten zur fünften Zeile mit  $a^2 = 2 - x^2$  genutzt, dass

$$\int \sqrt{a^2 - t^2} dt = \frac{1}{2}t\sqrt{a^2 - t^2} + \frac{a^2}{2} \arctan \left( \frac{t}{\sqrt{a^2 - t^2}} \right)$$

und von der fünften zur sechsten das gleiche nochmal jedoch mit  $a^2 = 1$ , sowie

$$\begin{aligned}
&\int (a^2 - t^2) \arcsin \left( \frac{1}{\sqrt{a^2 - t^2}} \right) dt \\
&= \frac{t}{6} \sqrt{a^2 - t^2} - \frac{1}{6} + \left( a^2 t - \frac{t^3}{3} \right) \arcsin \left( \frac{1}{\sqrt{a^2 - t^2}} \right) \\
&\quad - \frac{a^3}{3} \left( \arctan \left( \frac{a(t-a)+1}{\sqrt{a^2 - t^2} - 1} \right) + \arctan \left( \frac{a(t+a)-1}{\sqrt{a^2 - t^2} - 1} \right) \right) \\
&\quad + \frac{3a^2 + 1}{6} \arctan \left( \frac{t}{\sqrt{a^2 - t^2} - 1} \right)
\end{aligned}$$

mit  $a^2 = 2$ .

**Beispiel 3.3.19.** Das Volumen des  $n$ -dimensionalen Balls  $B_R^n(0)$  erfüllt die Rekursionsformel

$$\mu(B_R^n(0)) = \frac{2\pi R^2}{n} \mu(B_R^{n-2}(0))$$

also

$$\mu(B_R^{2k}(0)) = \frac{\pi^k R^{2k}}{k!} \quad \mu(B_R^{2k+1}(0)) = \frac{2^{2k+1} k! \pi^k R^{2k+1}}{(2k+1)!}.$$

Für den Beweis erweist sich die folgende Tatasche als nützlich:

Es sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  stetig und  $f$  hänge nicht echt von  $(x, y) \in \mathbb{R}^2$  ab, sondern nur von dem Ausdruck  $r := \sqrt{x^2 + y^2}$ . Setzen wir nun  $\tilde{f}(r) := f(x, y)$  für  $(x, y)$  mit  $x^2 + y^2 = r^2$ , so gilt

$$\int_{B_R^2(0)} f(x, y) d^2(x, y) = 2\pi \int_0^R \tilde{f}(r) r dr.$$

### 3.3.3 Die Transformationsformel

Wir betrachten im  $\mathbb{R}^3$  die Kugelkoordinaten aus<sup>(i)</sup> Beispiel 2.3.11:

$$\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \Phi(r, \theta, \varphi) = \begin{pmatrix} r \cos(\varphi) \sin(\theta) \\ r \sin(\varphi) \sin(\theta) \\ r \cos(\theta) \end{pmatrix}$$

Dabei bezeichnen wir den Urbildraum und den Bildraum auch als  $(r\varphi\theta)$ - und  $(xyz)$ -Raum und bezeichnen ihn mit  $\mathbb{R}_{r\varphi\theta}^3$  und  $\mathbb{R}_{xyz}^3$ .

Die Funktionalmatrix

$$D\Phi(r, \theta, \varphi) = \begin{pmatrix} \cos(\varphi) \sin(\theta) & r \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \sin(\theta) \\ \sin(\varphi) \sin(\theta) & r \sin(\varphi) \cos(\theta) & r \cos(\varphi) \sin(\theta) \\ \cos(\theta) & -r \sin(\theta) & 0 \end{pmatrix}$$

hat die Determinante

$$\det D\Phi = r^2 \sin(\theta).$$

Damit ist  $\Phi$  um jeden Punkt  $(r, \varphi, \theta) \in \mathbb{R}^3$  mit Ausnahme von  $(0 \times \mathbb{R}^2) \cup (\mathbb{R}^2 \times \mathbb{Z}\pi)$  lokal umkehrbar, siehe Satz 2.5.36. Die Abbildung  $\Phi$  ist sogar ein Diffeomorphismus auf das Bild, wenn wir sie auf die Menge

$$\mathbb{R}^{>0} \times ]0, \pi[ \times ]0, 2\pi[ \subset \mathbb{R}_{r\varphi\theta}^3$$

einschränken. Das Bild ist in diesem Fall

$$\mathbb{R}^3 \setminus \{(x, 0, z) \in \mathbb{R}^3 \mid x \geq 0\} \subset \mathbb{R}_{xyz}^3.$$

$\Phi$  hat die folgenden Eigenschaften:

<sup>(i)</sup>Die Variable  $\vartheta$  ist hier durch die Variable  $\theta = \frac{\pi}{2} - \vartheta$  ersetzt worden.

- $\Phi$  bildet den Quader  $Q_R = [0, R] \times [0, \pi] \times [0, 2\pi]$  auf den Ball  $\bar{B}_R(0)$  ab.
- Ein "kleiner" Quader  $Q := [r_1, r_2] \times [\varphi_1, \varphi_2] \times [\theta_1, \theta_2]$  wird durch  $\Phi$  auf ein Kugelschalensegment abgebildet. Vom Ursprung aus betrachtet hat es die folgende Gestalt:
  - seine "Unterseite" hat die "Seitenlängen"  $r_1 \sin \theta_1 (\varphi_2 - \varphi_1)$ ,  $r_1 \sin \theta_2 (\varphi_2 - \varphi_1)$  und zwei mal  $r_1(\theta_2 - \theta_1)$ ,
  - seine "Oberseite" hat die "Seitenlängen"  $r_2 \sin \theta_1 (\varphi_2 - \varphi_1)$ ,  $r_2 \sin \theta_2 (\varphi_2 - \varphi_1)$ , und zwei mal  $r_2(\theta_2 - \theta_1)$ ,
  - es hat die Höhe  $r_2 - r_1$ ,

siehe Abbildung 3.3.2.

Wollen wir nun eine Funktion  $f(x, y, z)$  über die Kugel  $\bar{B}_R(0)$  integrieren, so wissen wir, dass das Ergebnis wie folgt aussieht

$$\int_{\bar{B}_R(0)} f(x, y, z) d^3(x, y, z) = \int_{-R}^R \left( \int_{-\sqrt{R^2-z^2}}^{\sqrt{R^2-z^2}} \left( \int_{-\sqrt{R^2-y^2-z^2}}^{\sqrt{R^2-y^2-z^2}} f(x, y, z) dx \right) dy \right) dz.$$

Bringen wir nun die Abbildung  $\Phi$  ins Spiel und beachten statt  $f$  die Funktion

$$\tilde{f}(r, \theta, \varphi) = (f \circ \Phi)(r, \theta, \varphi) = f(r \cos(\varphi) \cos(\theta), r \sin(\varphi) \cos(\theta), r \sin(\theta))$$

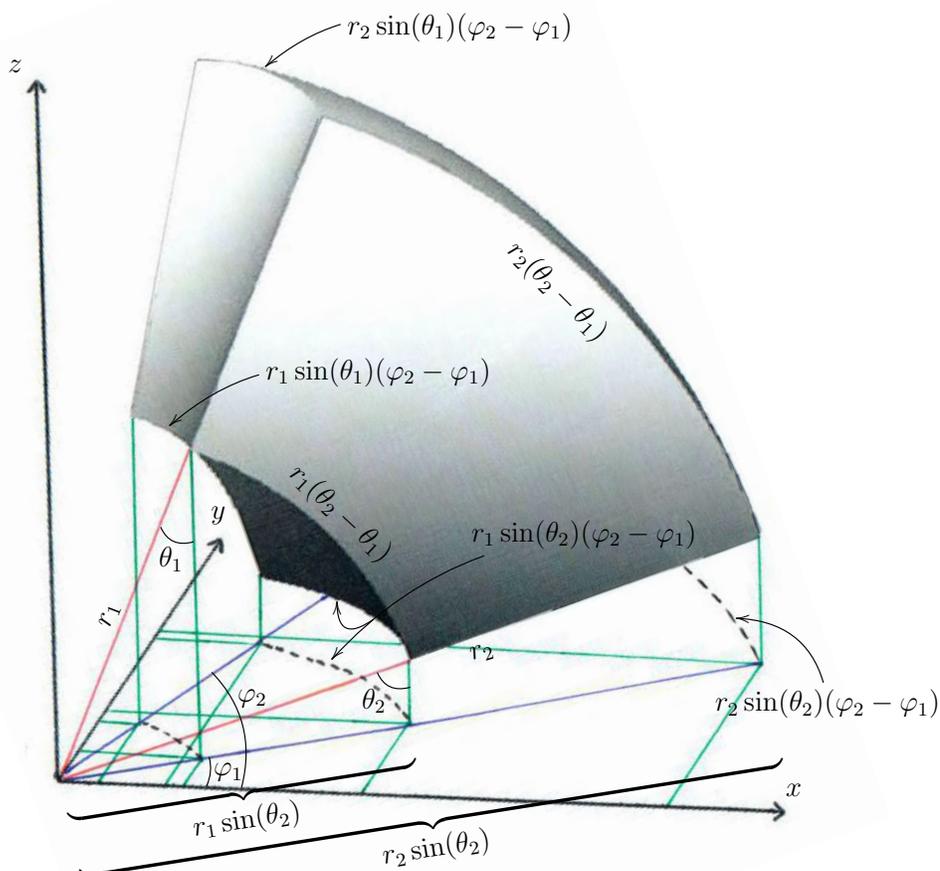
und beachten, dass  $\Phi^{-1}(\bar{B}_R(0)) = Q_R$  ist, so können wir statt des obigen Integrals das einfacher zu berechnende Integral

$$\int_{Q_R} \tilde{f}(r, \theta, \varphi) d^3(r, \theta, \varphi) = \int_0^R \left( \int_0^\pi \left( \int_0^{2\pi} \tilde{f}(r, \theta, \varphi) d\varphi \right) d\theta \right) dr$$

betrachten. Die beiden Integrale liefern jedoch unterschiedliche Werte, wie die Funktion  $f(x, y, z) = 1$ , also auch  $\tilde{f}(r, \theta, \varphi) = 1$ , zeigt

$$\int_{\bar{B}_R(0)} d^3(x, y, z) = \frac{4}{3}\pi R^3 \quad \text{aber} \quad \int_{Q_R} d^3(r, \theta, \varphi) = 2\pi^2 R.$$

Wir können diese Idee trotzdem nutzen, und wir erhalten die folgende gültige Umrechnungsformel für mehrdimensionale Integrale:

Abbildung 3.3.2: Das Kugelschalensegment als Bild des Würfels unter  $\Phi$ .

**Satz 3.3.20** (Transformationsformel). *Es seien  $M \subset \mathbb{R}_u^n$  und  $N \subset \mathbb{R}_x^n$  kompakt und messbar. Weiter sei  $\Phi : M \rightarrow N$  surjektiv und stetig differenzierbar.  $\Phi$  sei lediglich auf einer Nullmenge  $M \setminus M' \subset M$  nicht injektiv, sodass  $\Phi|_{M'}$  injektiv. Für alle  $u \in M'$  sei  $D\Phi(u) \in M_n \mathbb{R}$  invertierbar und es sei  $f : N \rightarrow \mathbb{R}$  stetig. Dann gilt*

$$\int_N f(x) d^n x = \int_M (f \circ \Phi)(u) |\det(D\Phi(u))| d^n u.$$

**Beweisskizze.** Der Hauptpunkt im Beweis ist die Begründung der folgenden Tatsache: Es sei  $M \subset \mathbb{R}^n$  eine kompakte messbare Menge und  $\Phi : M \rightarrow \mathbb{R}^n$  stetig differenzierbar derart, dass  $D\Phi$  auf ganz  $M$  regulär ist. Dann gilt:

$$\forall u_0 \in M \forall \epsilon > 0 \exists \delta > 0 \forall a < \delta \forall W = u_0 + [-a, a]^n : \\ |\mu(\Phi(W)) - |\det(D\Phi(u_0))| \mu(W)| < \epsilon \mu(W)$$

□

**Bemerkung 3.3.21.** Der Fall  $n = 1$  in Satz 3.3.20 entspricht der Substitutionsregel für eindimensionale Integrale, siehe Satz 1.5.8.1. Der Unterschied der dortigen Formel zur Transformationsformel ist das Fehlen des Betrags. Das wird an der Stelle durch die Anpassung der Integralgrenzen kompensiert.

**Beispiel 3.3.22.** Für die Menge  $M \subset \mathbb{R}^3$ , die durch die Mengen

$$O = \{(x, y, z) \mid z = 1\} , \\ U = \{(x, y, z) \mid z = x^2 + y^2\}$$

begrenzt wird, wollen wir das Integral

$$I = \int_M f(x, y, z) d^3(x, y, z)$$

berechnen. Mit dem Satz von Fubini gilt

$$I = \int_0^1 \left( \int_{B_{\sqrt{z}}(0)} f(x, y, z) d^2(x, y) \right) dz = \int_0^1 \int_{-\sqrt{z}}^{\sqrt{z}} \int_{-\sqrt{z-y^2}}^{\sqrt{z-y^2}} f(x, y, z) dx dy dz$$

oder

$$I = \int_{B_1(0)} \left( \int_{x^2+y^2}^1 f(x, y, z) dz \right) d^2(x, y) = \int_0^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{x^2+y^2}^1 f(x, y, z) dz dy dx$$

Nutzt man Zylinderkoordinaten  $\Phi(r, \varphi, z) = (r \cos(\varphi), r \sin(\varphi), z)$ , so ist

$$\Phi^{-1}(O) = \{(r, \varphi, z) \mid z = 1\} = \mathbb{R}^{\geq 0} \times [0, 2\pi] \times \{1\} , \\ \Phi^{-1}(U) = \{(r, \varphi, z) \mid z = r^2\} .$$

Nutzt man noch  $\det(D\Phi) = r$  und schreibt  $\tilde{f}(r, \varphi, z) = f(r \cos \varphi, r \sin(\varphi), z)$ , so gilt wegen der Transformationsformel

$$\begin{aligned} I &= \int_{\Phi^{-1}(M)} r \tilde{f}(r, \varphi, z) d^3(r, \varphi, z) = \int_0^{2\pi} \left( \int_{M^\varphi} r \tilde{f}(r, \varphi, z) d^2(r, z) \right) d\varphi \\ &= \int_0^{2\pi} \int_0^1 \int_0^{r^2} r \tilde{f}(r, \varphi, z) dz dr d\varphi \end{aligned}$$

wobei im letzten Schritt  $M^\varphi = \{(r, z) | r \in [0, 1], z \in [r^2, 1]\}$  nicht von  $\varphi$  abhängt. Ebenso können wir wie folgt rechnen:

$$I = \int_0^1 \left( \int_{[0, \sqrt{z}] \times [0, 2\pi]} r \tilde{f}(r, \varphi, z) d^2(r, \varphi) \right) dz = \int_0^1 \int_0^{\sqrt{z}} \int_0^{2\pi} r \tilde{f}(r, \varphi, z) d\varphi dr dz$$

oder

$$I = \int_{[0, 1] \times [0, 2\pi]} \left( \int_{r^2}^1 r \tilde{f}(r, \varphi, z) dz \right) d^2(r, \varphi) = \int_0^1 \int_0^{2\pi} \int_0^{r^2} r \tilde{f}(r, \varphi, z) dz d\varphi dr$$

Wählt man speziell  $f(x, y, z) = \sqrt{x^2 + y^2}$ , so ist  $I = \frac{4}{15}\pi$ . Während die Integration hier in kartesischen Koordinaten einen nicht unerheblichen Aufwand erfordert, verläuft diese in Zylinderkoordinaten nahezu problemlos.

Der Satz von Fubini lässt sich oft in der folgenden Form anwenden. Im vorigen Beispiel haben wir das an mehreren Stellen genutzt.

**Beispiel 3.3.23.** Es sei  $U \subset \mathbb{R}^{n-1}$  messbar und  $f_u, f_o : U \rightarrow \mathbb{R}$  Funktionen mit  $f_u(x) \leq f_o(x)$  für alle  $x \in U$ . Ist  $M \subset \mathbb{R}^n$  in der Form<sup>(i)</sup>

$$M = \{(x, t) \in \mathbb{R}^n \mid f_u(x) \leq t \leq f_o(x)\} .$$

gegeben, so ist

$$\int_M f(x, t) d^n(x, t) = \int_U \left( \int_{f_u(x)}^{f_o(x)} f(x, t) dt \right) d^{n-1}x .$$

<sup>(i)</sup>In der Literatur heissen Bereiche dieser Forma uch *Normalbereiche*.

Solch eine Beschreibung kann man oft durch geschickte Zerlegung des Integrationsbereichs erhalten.

### 3.4 Reguläre Teilmengen des $\mathbb{R}^n$

#### 3.4.1 Vorbereitungen

Unter regulären Punktmenge verstehen wir in diesem Kapitel Teilmengen des  $\mathbb{R}^n$ , die bestimmte natürliche, durch die Anschauung motivierte Eigenschaften haben. Bevor wir uns diesen "schönen" Teilmengen des  $\mathbb{R}^n$  widmen, wollen wir noch zwei Eigenschaften von Abbildungen nachholen, die eng verwandt sind mit dem Umkehrsatz 2.5.36 und sich aus diesem ableiten lassen. Darüber hinaus kann die erste auch als Variante des Satzes über implizite Funktionen 2.5.31 angesehen werden.

**Satz 3.4.1** (Submersionssatz/Immersionssatz). *1. Es sei  $m > n$ ,  $M \subset \mathbb{R}^m$  und  $f : M \rightarrow \mathbb{R}^n$  stetig differenzierbar. Weiter sei  $x^0 \in M$  derart, dass die Funktionalmatrix  $Df(x^0)$  den vollen Rang  $n$  hat. Dann gibt es einen Ball  $B = B_r(x^0) \subset M$  und einen Diffeomorphismus  $h : B \rightarrow h(B) \subset \mathbb{R}^m$ , sodass*

$$f(x) = (h_1(x), \dots, h_n(x)) \quad \text{für alle } x \in B,$$

wobei  $(h_1, \dots, h_n, h_{n+1}, \dots, h_m)$  die Komponentenfunktionen von  $h$  bezeichnen. Mit Hilfe der Projektion  $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n, \pi(x_1, \dots, x_m) = (x_1, \dots, x_n)$  schreibt sich das als

$$f|_B = \pi \circ h.$$

*2. Es sei  $k < n$ ,  $M \subset \mathbb{R}^k$  und  $f : M \rightarrow \mathbb{R}^n$  stetig differenzierbar. Weiter sei  $x^0 \in M$  derart, dass die Funktionalmatrix  $Df(x^0)$  den vollen Rang  $k$  hat. Dann gibt es einen Ball  $B = B_r(x^0) \subset M$  und eine offene Menge  $U \subset \mathbb{R}^n$  mit  $B \times \underbrace{\{(0, \dots, 0)\}}_{\in \mathbb{R}^{n-k}} \subset U$  und einen Diffeomorphismus*

$h : U \rightarrow h(U) \subset \mathbb{R}^n$ , sodass

$$f(x) = h(x, 0, \dots, 0) \quad \text{für alle } x \in B.$$

Mit Hilfe der Inklusion  $\iota : \mathbb{R}^k \rightarrow \mathbb{R}^n, \iota(x) = (x, \underbrace{0, \dots, 0}_{(n-k)\text{-mal}}$  schreibt sich das als

$$f|_B = h \circ \iota|_B.$$

Abbildung 3.4.1: Zum Submersionsatz 3.4.1.1

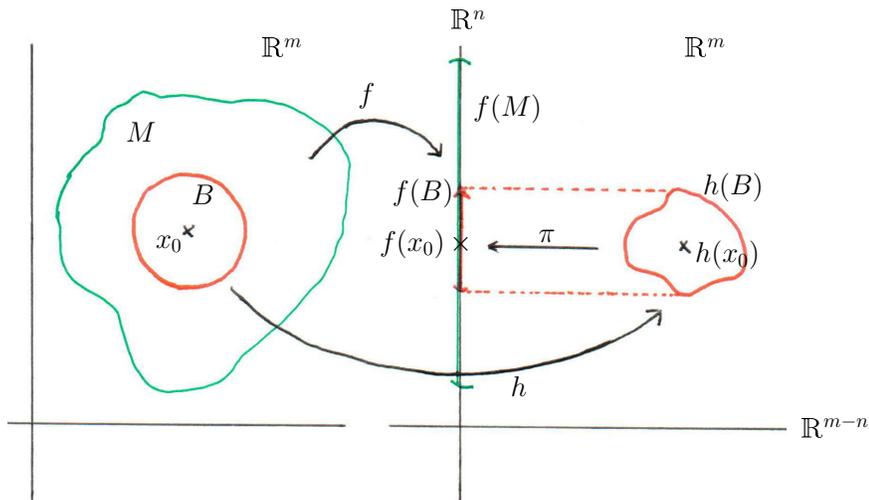
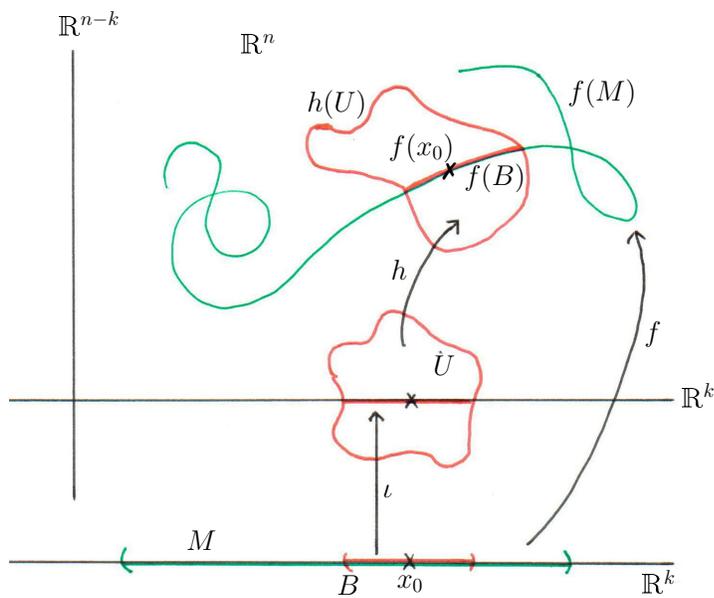


Abbildung 3.4.2: Zum Immersionsatz 3.4.1.2



**Beweisskizze.**

1. Es seien  $(f_1, \dots, f_n)$  die Komponentenfunktionen von  $f$  und ohne Beschränkung gelte  $\det \left( \left( \frac{\partial f_i}{\partial x_j}(x^0) \right)_{i,j=1, \dots, n} \right) \neq 0$ . Dann hat eine geeignete Einschränkung der Abbildung

$$h(x_1, \dots, x_m) = (f_1(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m), x_{n+1}, \dots, x_m)$$

die gewünschten Eigenschaften.

2. Es seien  $v^1, \dots, v^{n-k} \in \mathbb{R}^n$  eine Ergänzung von  $\frac{\partial f}{\partial x_1}(x^0), \dots, \frac{\partial f}{\partial x_k}(x^0)$  zu einer Basis des  $\mathbb{R}^n$ . Dann hat die Abbildung

$$h(x, t) = f(x) + \sum_{j=1}^{n-k} t_j v^j$$

die gewünschten Eigenschaften. □

**Bemerkung 3.4.2.** • Schreiben wir in Satz 3.4.1.2.  $h^{-1} = (g_1, \dots, g_n)$ , so ist  $(f|_B)^{-1} = (g_1, \dots, g_k)$ .

- Aus Satz 3.4.1.1. folgt der Satz über implizite Funktionen. Das liefert insbesondere die Äquivalenz des Umkehrsatzes und des Satzes über implizite Funktionen:

Das sieht man, indem man die Abbildung aus dem Beweis heranzieht. Wir schreiben die Element aus dem Definitionsbereich von  $f$  und  $h$  als  $x$  und die aus dem Bildbereich in der Form  $(y, z)$  mit  $y \in \mathbb{R}^n$  und  $z \in \mathbb{R}^{m-n}$ . Damit schreibt sich die im Beweis konstruierte Funktion als

$$h(x) = (y, z) \text{ mit } \begin{cases} y = f(x) \\ z = (x_{n+1}, \dots, x_m) \end{cases} .$$

Setzen wir weiter  $y_0 = f(x_0) \in \mathbb{R}^n$  so gilt

$$h(f^{-1}(\{y_0\}) \cap B) = \{y \in h(B) \mid \pi(y, z) = y^0\} = \{y^0\} \times V$$

mit  $V = \{z \in \mathbb{R}^{m-n} \mid (y^0, z) \in h(U)\}$ . Dann ist für alle  $z \in V$

$$h^{-1}(y^0, z) = (h_1^{-1}(y^0, z), \dots, h_n^{-1}(y^0, z), z) ,$$

denn es ist

$$\begin{aligned} (y^0, z) &= h(h^{-1}(y^0, z)) \\ &= (f(\underbrace{h^{-1}(y^0, z)}_{\in f^{-1}(\{y^0\})}, h_{n+1}^{-1}(y^0, z), \dots, h_m^{-1}(y^0, z)) \\ &= (y^0, h_{n+1}^{-1}(y^0, z), \dots, h_m^{-1}(y^0, z)). \end{aligned}$$

Schreiben wir noch  $g : V \rightarrow \mathbb{R}^n$  mit  $g(z) := (h_1^{-1}(y^0, z), \dots, h_n^{-1}(y^0, z))$ , so ist

$$\begin{aligned} f^{-1}(\{y^0\}) \cap B &= h^{-1}(h(f^{-1}(\{y^0\}) \cap B)) \\ &= h^{-1}(\{y^0\} \times V) \\ &= \{(g(z), z) \mid z \in V\}. \end{aligned}$$

Dies ist die Aussage aus dem Satz über implizite Funktionen.

### 3.4.2 Untermannigfaltigkeiten des $\mathbb{R}^n$

**Definition 3.4.3.** 1. Ein Punkt  $x^0 \in M \subset \mathbb{R}^n$  heißt *regulärer Punkt von  $M$  der Dimension  $k$* , wenn es eine Menge  $U \subset M$  mit  $x^0 \in U$  und eine offene Menge  $V \subset \mathbb{R}^k$  sowie eine Abbildung  $\Phi : V \rightarrow U$  gibt mit

- $\Phi$  ist bijektiv und stetig.
  - $\Phi^{-1} : U \rightarrow V$  ist stetig (bezüglich der relativen Offenheit bezüglich  $M$  in  $\mathbb{R}^n$ )
2. Ist  $x^0$  regulär der Dimension  $k$ , so heißt  $x^0$  ein  *$r$ -regulärer Punkt*, wenn  $\Phi$   $r$ -mal stetig differenzierbar ist und  $\text{Rang}(D\Phi(x^0)) = k$ . Ist  $r = \infty$ , so heißt  $x^0$  ein *glatter Punkt*
3. Die Abbildung  $\Phi$  heißt *Parametrisierung (von  $M$  um  $x^0$ )* und die Abbildung  $\Phi^{-1}$  heißt *Karte (von  $M$  um  $x^0$ )*.<sup>(i)</sup>

**Bemerkung 3.4.4.** • Ist  $x^0 \in M$  regulär und  $\Phi : \Phi^{-1}(U) \rightarrow U \subset M$  eine Parametrisierung um  $x^0$ , dann sind alle Punkte aus  $U$  regulär derselben Dimension  $k$ .

<sup>(i)</sup>Genauer  $C^r$ -Parametrisierung oder  $C^r$ -Karte, wenn  $\Phi$   $r$ -mal stetig differenzierbar ist.

- Die regulären Punkte der Dimension  $k$  bilden eine (relativ) offene Teilmenge von  $M$ .
- Die nicht-regulären Punkte bilden eine (relativ) abgeschlossene Teilmenge von  $M$ .
- Ist  $k$  die Dimension des regulären Punktes  $x^0 \in M$ , dann müssen nicht alle regulären Punkte von  $M$  diese Dimension haben.

Das motiviert nun die folgende Definition.

**Definition 3.4.5.** 1. Eine Teilmenge  $M \subset \mathbb{R}^n$  heißt *k-dimensionale Untermannigfaltigkeit des  $\mathbb{R}^n$* , wenn jeder Punkt aus  $M$  regulär der Dimension  $k$  ist.

2. Als *Atlas* bezeichnen wir eine Menge von Karten von  $M$ , deren Definitionsbereiche ganz  $M$  überdecken.<sup>(i)</sup>

**Bemerkung 3.4.6.** Es seien  $V_1, V_2 \overset{\circ}{\subset} \mathbb{R}^k$  und  $\Phi_1 : V_1 \rightarrow M$  und  $\Phi_2 : V_2 \rightarrow M$  zwei Parametrisierungen eines  $C^r$ -Atlanten der Untermannigfaltigkeit  $M \subset \mathbb{R}^n$ . Dann ist  $\Phi_1^{-1} \circ \Phi_2 : \Phi_2^{-1}(\Phi_1(V_1) \cap \Phi_2(V_2)) \rightarrow \Phi_1^{-1}(\Phi_1(V_1) \cap \Phi_2(V_2))$  eine  $r$ -mal stetig differenzierbare Abbildung zwischen offenen Mengen des  $\mathbb{R}^k$ , siehe Abbildung 3.4.3.

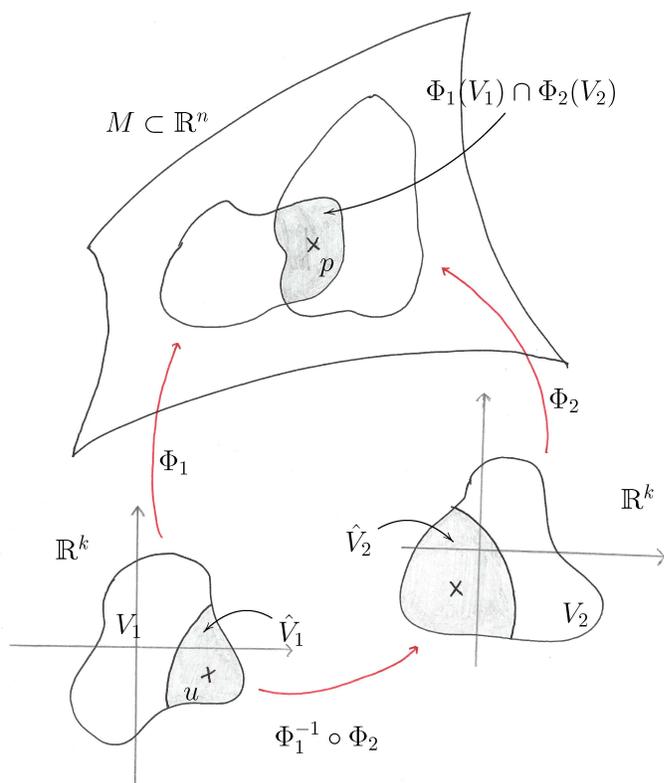
**Bezeichnung 3.4.7.** • Statt Untermannigfaltigkeit sagen wir ab jetzt kürzer nur Mannigfaltigkeit.

- Wenn wir im Folgenden von Mannigfaltigkeiten sprechen sind diese – wenn nicht anders angegeben – in der Regel glatt, jedoch stets mindestens stetig differenzierbar.

**Verabredung:** Um uns technische Probleme zu ersparen, setzen wir in diesem und den folgenden Kapiteln voraus, dass die betrachteten Mannigfaltigkeiten endliche Atlanten besitzen.

<sup>(i)</sup>Genauer sprechen wir von einer  $C^r$ -Untermannigfaltigkeit, wenn alle Punkte  $r$ -regulär sind, und von einem  $C^r$ -Atlas wenn alle Karten  $r$ -mal stetig differenzierbar sind. Im Fall  $r = \infty$  sprechen wir auch von einer glatten Untermannigfaltigkeit und einem glatten Atlas.

Abbildung 3.4.3: Der Kartenwechsel



$$\hat{V}_1 = \Phi_1^{-1}(\Phi_1(V_1) \cap \Phi_2(V_2)), \quad \hat{V}_2 = \Phi_2^{-1}(\Phi_1(V_1) \cap \Phi_2(V_2))$$

- Beispiel 3.4.8.**
1. Die Sphären  $S_{\mathbb{R}}^{n-1}(x^0) \subset \mathbb{R}^n$  sind Mannigfaltigkeiten der Dimension  $n - 1$ .
  2. Die affinen  $k$ -dimensionalen Unterräume des  $\mathbb{R}^n$  sind Mannigfaltigkeiten der Dimension  $k$ .
  3. Die offenen Teilmengen des  $\mathbb{R}^n$  sind genau die  $n$ -dimensionalen Untermannigfaltigkeiten des  $\mathbb{R}^n$ .
  4. Der Schnitt von  $S_{\mathbb{R}}^{n-1}(0) \subset \mathbb{R}^n$  mit dem affinen Unterraum  $E_\ell := \alpha(e_n) + \{x \in \mathbb{R}^n \mid x_1 = \dots = x_\ell = 0\}$  für  $\ell \in \{1, \dots, n - 2\}$  ist entweder leer  $|\alpha| > 1$ , besteht aus einem Punkt  $|\alpha| = 1$ , oder ist eine Mannigfaltigkeit der Dimension  $n - \ell - 1$ . Der Schnitt ist gerade

$$\alpha e_n + S_{1-\alpha^2}^{n-\ell-1}(0).$$

Einige der Eigenschaften aus dem vorigen Beispiel sind so fundamental, dass wir sie hier herausstellen wollen.

**Satz 3.4.9.** 1. Es sei  $U \subset \mathbb{R}^k$  offen und  $f : U \rightarrow \mathbb{R}^\ell$  eine  $r$ -mal stetig differenzierbare Abbildung. Dann ist ihr Graph

$$G_f = \{y \in \mathbb{R}^{k+\ell} \mid y = (x, f(x)) \text{ für ein } x \in U\}$$

eine  $k$ -dimensionale  $C^r$ -Mannigfaltigkeit in  $\mathbb{R}^{k+\ell}$ . Die Abbildung  $\Phi : U \rightarrow \mathbb{R}^{k+\ell}$  mit  $\Phi(x) := (x, f(x))$  ist eine globale Karte.

2. Es sei  $F : \mathbb{R}^{k+\ell} \rightarrow \mathbb{R}^\ell$   $r$ -mal stetig differenzierbar und für alle  $x^0 \in F^{-1}(y^0)$  sei der Rang von  $DF(x^0)$  maximal, also  $k$ . Dann ist  $F^{-1}(y^0)$  eine  $C^r$ -Mannigfaltigkeit der Dimension  $k$ .

Mit Satz 3.4.1 besitzt das zuvor gesagte auch eine – zumindest lokale – Umkehrung, die wir mit der vorigen Aussage wie folgt zusammenfassen.

**Satz 3.4.10** (Darstellungssatz für Mannigfaltigkeiten). Die folgenden drei Charakterisierungen einer Teilmenge  $M \subset \mathbb{R}^n$  sind äquivalent:

1.  $M$  ist eine  $C^r$ -Mannigfaltigkeit der Dimension  $k$ .
2. Zu jedem Punkt  $x^0 \in M$  gibt es eine offene zusammenhängende Umgebung  $U \subset M$  und eine offene Teilmenge  $V \subset \mathbb{R}^k$ , sodass  $U$  der Graph einer  $r$ -mal stetig differenzierbaren Abbildung  $f : V \rightarrow \mathbb{R}^{n-k}$  ist (gegebenenfalls nach Umsortierung der Koordinaten).
3. Zu jedem Punkt  $x^0 \in M$  gibt es eine offene Umgebung  $U \subset M$ , eine offene Menge  $V \subset \mathbb{R}^n$  mit  $U \subset V$  und eine  $r$ -mal stetig differenzierbare Abbildung  $F : V \rightarrow \mathbb{R}^{n-k}$ , sodass  $U = F^{-1}(0)$ .

Etwas plakativ kann man sagen: Eine  $k$ -dimensionale Mannigfaltigkeit  $M \subset \mathbb{R}^n$  lässt sich lokal

- als Bild einer Parametrisierung  $\Phi : \mathbb{R}^k \rightarrow \mathbb{R}^n$  oder
- als Graph einer Abbildung  $f : \mathbb{R}^k \rightarrow \mathbb{R}^{n-k}$  oder

- als Nullstellenmenge einer Abbildung  $F : \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$

darstellen. Sind im letzten Fall  $(F_1, \dots, F_{n-k})$  die Komponentenfunktionen von  $F$ , so kann man sagen, dass  $M$  sich lokal

- als Lösungsmenge des Gleichungssystems 
$$\begin{cases} F_1(x_1, \dots, x_n) = 0 \\ \vdots \\ F_{n-k}(x_1, \dots, x_n) = 0 \end{cases}$$

darstellen lässt. Wie wir im Laufe der Vorlesung sehen werden, haben alle drei Darstellungen ihre Vorteile.

**Bezeichnung 3.4.11.** Die  $n - 1$  dimensionalen Mannigfaltigkeiten im  $\mathbb{R}^n$  nennen wir auch *Hyperflächen*. Sie sind lokal Nullstellenmengen von reellen Funktionen  $F$  auf  $\mathbb{R}^n$  oder lokal Graphen reeller Funktionen  $f$  auf  $\mathbb{R}^{n-1}$ .

Es folgen einige wichtige Beispiele für Hyperflächen im  $\mathbb{R}^n$

**Beispiel 3.4.12.** 1. Die Sphären  $S_R^{n-1}(x^0) \subset \mathbb{R}^n$  sind Hyperflächen

2. Die affinen Hyperebenen  $E_{v,\delta} = \{x \in \mathbb{R}^n \mid \langle v, x \rangle = \delta\}$  mit Normalenvektor  $v$  und Abstand  $|\delta|$  zum Ursprung sind Hyperflächen.
3. Ist  $A \in M_n \mathbb{R}$  eine symmetrische Matrix,  $b \in \mathbb{R}^n$  und  $\gamma \in \mathbb{R}$ , dann sind alle Punkte der Quadrik

$$Q(A, b, \gamma) := \{x \in \mathbb{R}^n \mid x^t A x - 2b^t x - \gamma = 0\}$$

regulär mit Ausnahme der Punkte die das lineare Gleichungssystem  $Ax = b$  lösen. Durch die Transformation  $x \mapsto y = x - b$  erhalten wir

$$Q(A, b, \gamma) = b + Q(A, 0, \hat{\gamma})$$

mit  $\hat{\gamma} = \gamma - b^t A b$ . Mit dem Sylvesterschen Trägheitssatz und mit Hilfe einer Skalierung erreichen wir, dass es reicht Diagonalmatrizen mit Einträgen aus  $\{0, 1, -1\}$  und Skalare  $\gamma$  aus der gleichen Menge zu betrachten.

Ist speziell  $A$  symmetrisch so schreiben  $E_{n,\ell} := \begin{pmatrix} \mathbb{1}_\ell & 0 \\ 0 & -\mathbb{1}_{n-\ell} \end{pmatrix}$  und  $Q_\gamma^{n,\ell} := Q(E_{n,\ell}, 0, \gamma)$ . Für  $\gamma = \pm 1$  erhalten wir eine Hyperfläche. Für  $\gamma = 0$  jedoch erst nach Entfernung des Ursprungs

4. Im  $\mathbb{R}^3$  heißen die Hyperflächen  $Q_{-1}^{3,2}$  und  $Q_1^{3,2}$  zweischaliges und einschaliges Hyperboloid und  $Q_0^{3,2} \setminus \{0\}$  heißt Doppelkegel.

Ist der Atlas einer Mannigfaltigkeit differenzierbar, dann macht es Sinn zu fragen, ob eine Funktion auf der Mannigfaltigkeit differenzierbar ist.

**Definition/Bemerkung 3.4.13.** Es sei  $M \subset \mathbb{R}^n$  eine  $k$ -dimensionale Mannigfaltigkeit. Eine Abbildung  $f : M \rightarrow \mathbb{R}^m$  heißt differenzierbar in  $p \in M$ , wenn es eine Parametrisierung  $\Phi : U \rightarrow M$ ,  $u \in U \stackrel{\circ}{\subset} \mathbb{R}^k$ ,  $\Phi(u) = p$  gibt, sodass die Abbildung  $f \circ \Phi : U \rightarrow \mathbb{R}^m$  in  $u \in U$  differenzierbar im gewöhnlichen Sinn ist. Diese Definition ist unabhängig von der Wahl der Parametrisierung.

**Bemerkung 3.4.14.** Ist  $M \subset \mathbb{R}^n$  eine Mannigfaltigkeit und ist die Funktion  $f : M \rightarrow \mathbb{R}$  als Einschränkung einer differenzierbaren Funktion auf einer offenen Menge  $V \subset \mathbb{R}^n$  mit  $M \subset V$  gegeben, dann ist  $f$  differenzierbar.

**Beispiel 3.4.15.** Es sei  $x^0 \in \mathbb{R}^n$  mit  $\|x^0\| \neq R$ . Dann ist die Funktion

$$f : S_R^{n-1}(0) \rightarrow \mathbb{R}, \quad f(x) = \frac{1}{\|x - x^0\|}$$

differenzierbar.

Wieder mit Hilfe der Konstruktion aus Satz 3.4.1 folgt auch hier die – zumindest lokale – Umkehrung von Bemerkung 3.4.14.

**Satz 3.4.16.** *Es sei  $M \subset \mathbb{R}^n$  eine Mannigfaltigkeit und  $f : M \rightarrow \mathbb{R}$  eine in  $p \in M$  differenzierbare Funktion. Dann gibt es eine offene Umgebung  $V \subset \mathbb{R}^n$  um  $p$  und eine differenzierbare Abbildung  $\hat{f} : V \rightarrow \mathbb{R}$  mit  $f|_{M \cap V} = \hat{f}|_{M \cap V}$ .*

### 3.4.3 Kurven und Flächen

Wir werden uns in diesem Abschnitt einigen speziellen Teilmengen des  $\mathbb{R}^3$  und  $\mathbb{R}^2$  zuwenden. Dazu führen wir im Folgenden Sprechweisen ein, die Spezialfälle und Verallgemeinerungen von niedrigdimensionalen Mannigfaltigkeiten beschreiben.

Die erste Definition ist bereits bekannt, siehe Definition 2.2.52, wir wollen sie in diesem Kontext trotzdem wiederholen.

- Definition 3.4.17** (Kurven). 1. Eine *Kurve* ist eine Abbildung  $c : I \rightarrow \mathbb{R}^n$ , wobei der Definitionsbereich  $I \subset \mathbb{R}$  ein Intervall ist. Wenn wir betonen möchten, dass  $c$   $r$ -mal stetig differenzierbar ist, so sprechen wir von einer  $C^r$ -Kurve.
2. Eine  $C^r$ -Kurve  $c : [a, b] \rightarrow \mathbb{R}^n$  heißt *geschlossen*, wenn  $c(a) = c(b)$  und  $c^{(j)}(a) = c^{(j)}(b)$  für alle  $j = 1, \dots, r$ .  $c : [a, b] \rightarrow \mathbb{R}^n$  heißt *einfach geschlossen*, wenn  $c$  geschlossen und  $c|_{]a,b[}$  injektiv ist.

Die Bilder von Kurven als Teilmengen des  $\mathbb{R}^n$  sind in der Regel keine eindimensionalen Mannigfaltigkeit.

- Bemerkung 3.4.18.** 1. Eine Kurve ist nicht notwendigerweise injektiv.
2. Ist  $c' \neq 0$ , dann ist  $c$  eingeschränkt auf ein geeignetes offenes Teilintervall wegen Satz 3.4.1 injektiv.
3. Das Bild einer injektiven Kurve ist nicht notwendigerweise eine Mannigfaltigkeit.
4. Ist das Bild einer Kurve eine Mannigfaltigkeit, so ist  $c$  ihre Parametrisierung.

**Beispiel 3.4.19.** Wir betrachten die Kurve  $c : \mathbb{R} \rightarrow \mathbb{R}^2$  mit

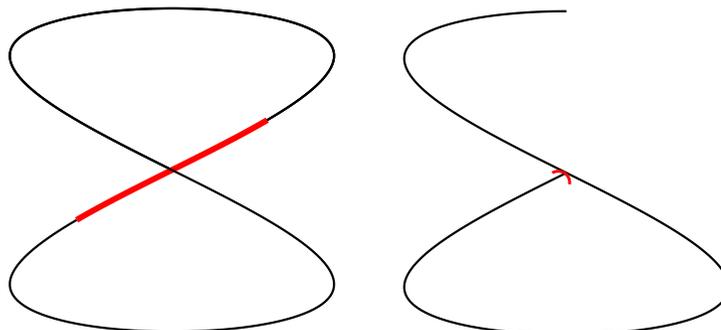
$$c(t) = (\sin(2t), \sin(t)).$$

- i. Die Einschränkung auf  $[0, 2\pi]$  ist geschlossen.
- ii. Die Einschränkung auf  $J = ] - \frac{\pi}{10}, \frac{\pi}{10}[$  ist injektiv und ihr Bild eine Mannigfaltigkeit.
- iii. Die Einschränkung auf  $] - \frac{3}{2}\pi, 0[$  ist injektiv, aber keine Mannigfaltigkeit.

Abbildung 3.4.4: Kurven, die keine Mannigfaltigkeiten sind

zu i. (schwarz) und ii. (rot)

zu iii.



**Definition 3.4.20.** Es sei  $M \subset \mathbb{R}^n$  eine Mannigfaltigkeit und  $c : I \rightarrow \mathbb{R}^n$  eine Kurve.  $c$  heißt *Kurve auf  $M$* , wenn  $c(I) \subset M$ .

**Definition 3.4.21 (Flächen).** 1. Unter einem Flächenstück verstehen wir eine zweidimensionale Mannigfaltigkeit  $M \subset \mathbb{R}^3$ , die durch genau eine Parametrisierung beschrieben wird.<sup>(i)</sup>

2. Eine (echte) Fläche ist eine zweidimensionale Mannigfaltigkeit im  $\mathbb{R}^3$ .
3. Unter einer (allgemeinen) Fläche verstehen wir eine Menge  $M \subset \mathbb{R}^3$  der folgenden Form:

$M$  ist die Vereinigung von endlich vielen Teilmengen  $\hat{M}_\alpha = M_\alpha \cup C_\alpha$  mit den Eigenschaften

- (a)  $M_\alpha$  ist ein Flächenstück, das über einer Menge  $U_\alpha \subset \mathbb{R}^2$  parametrisiert ist, deren Rand  $\partial U_\alpha$  eine endliche Vereinigung von einfach geschlossenen stückweise stetig differenzierbaren Kurven ist. Ist  $\Phi_\alpha : U_\alpha \rightarrow M_\alpha$  diese Parametrisierung, so gilt weiter  $C_\alpha = \Phi_\alpha(\partial U_\alpha)$ .
- (b)  $\hat{M}_\alpha \cap \hat{M}_\beta$  ist leer, endlich oder Teilmenge von  $C_\alpha \cap C_\beta$ . Die Kurven, die sich im letzten Fall ergeben nennen wir *Kanten von  $M$*
- (c)  $\hat{M}_\alpha \cap \hat{M}_\beta \cap \hat{M}_\gamma$  ist eine endliche Menge.

<sup>(i)</sup>Wenn wir betonen wollen, dass die Parametrisierung  $r$ -mal stetig differenzierbar ist, so sprechen wir genauer von einem  $C^r$ -Flächenstück.

Der Punkt 3. lässt sich ohne Umwege zur Definition einer *allgemeinen zweidimensionalen Untermannigfaltigkeit* des  $\mathbb{R}^n$  erweitern, indem wir in der ersten Zeile lediglich  $\mathbb{R}^3$  durch  $\mathbb{R}^n$  ersetzen.

**Bemerkung 3.4.22.** 1. Durch diese Definition wird insbesondere die Oberfläche eines Quaders zu einer allgemeinen Fläche. Ebenso gilt dies für die Oberfläche eines Vollzylinders.

2. Aber auch die Halbsphäre  $\hat{M}_1$  mit angesetztem Zylindermantel  $\hat{M}_2$  ist eine verallgemeinerte Fläche. Die Kante dieser Fläche ist dann die Verklebungslinie. Der nicht im Schnitt der Teilflächen vorkommende Rand von  $\hat{M}_2$  liefert schließlich den Rand von  $M$ .

3. Im Allgemeinen besteht der Rand von  $M$  aus den Teilen der Bilder der  $C_\alpha$  die nicht in den Schnitten der Teilflächen vorkommen.

**Bemerkung 3.4.23.** 1. Ein  $C^r$ -Flächenstück ist das Bild einer  $r$ -mal stetig differenzierbaren Abbildung  $\Psi : \mathbb{R}^2 \supset U \rightarrow \Psi(U) = M \subset \mathbb{R}^3$  derart, dass  $\Psi^{-1} : \Psi(U) \rightarrow U$  stetig ist.

2. Eine (echte) Fläche ist eine Vereinigung von Flächenstücken, derart, dass der Schnitt zweier Flächenstücke entweder leer oder selbst ein Flächenstück ist.

#### 3.4.4 Der Tangentialraum

Den Graphen  $G_f$  einer Funktion  $f : U \rightarrow \mathbb{R}$  für  $U \subset \mathbb{R}^{n-1}$  können wir, wie wir in Satz 3.4.10 gesehen haben, als  $(n-1)$ -dimensionale Mannigfaltigkeit auffassen. Bereits früher haben wir für solche Funktionen die Richtungsableitung von  $f$  im Punkt  $x^0 \in U$  in Richtung  $v \in \mathbb{R}^{n-1}$  eingeführt, siehe Definition 2.5.1. Die Richtungsableitung lässt sich als letzte Komponente der Ableitung einer Kurve auf dieser Mannigfaltigkeit interpretieren. Dabei ist die betrachtete Kurve gerade  $t \mapsto (x^0 + tv, f(x^0 + tv))$  und ihr Bild liegt im Schnitt von  $M$  und der Ebene  $\text{span} \left\{ \begin{pmatrix} v \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^n$ . Die Gerade in dieser Ebene, die tangential an die Kurve ist, bezeichnen wir als eine Tangente an  $G_f$  im Punkt  $(x_0, f(x^0))$ , analog zu Funktionen in einer Variablen. Die Tangente hat den Aufpunkt  $(x^0, f(x^0))$  und den Richtungsvektor  $(D_v f(x^0))$ . Schreiben wir  $\Phi(x) = (x, f(x))$

für die Parametrisierung von  $G_f$ , dann ist

$$\begin{aligned} \begin{pmatrix} v \\ D_v f(x^0) \end{pmatrix} &= \frac{d}{dt}(x^0 + tv, f(x^0 + tv))\big|_{t=0} = \frac{d}{dt}\Phi(x^0 + tv)\big|_{t=0} \\ &= D\Phi(x^0, f(x^0)) \cdot v. \end{aligned}$$

Alle so erhaltenen Richtungsvektoren von möglichen Tangenten an Kurven auf der Mannigfaltigkeit bilden somit einen Untervektorraum des  $\mathbb{R}^n$  der Dimension  $n - 1$ .

**Definition 3.4.24.** Es sei  $M \subset \mathbb{R}^n$  eine  $k$ -dimensionale Mannigfaltigkeit und  $p \in M$ . Mit  $T_p M$  bezeichnen wir den *Tangentialraum an  $M$  im Punkt  $p$*  und definieren ihn wie folgt: Ist  $\Phi : U \rightarrow M$  eine Parametrisierung um  $p$  und schreiben wir  $u = \Phi^{-1}(p)$ , so ist

$$T_p M := \text{Bild}(D\Phi(u)) = D\Phi(u)(\mathbb{R}^k) \subset \mathbb{R}^n.$$

Dass Definition 3.4.24 sinnvoll ist, zeigt die folgende Bemerkung:

**Bemerkung 3.4.25.** 1. Der Tangentialraum ist ein Untervektorraum des  $\mathbb{R}^n$  der Dimension  $k$ . Eine Basis ist durch die Spalten von  $D\Phi(u)$  gegeben.

2. Die Definition des Tangentialraums hängt nicht von der Wahl der Karte ab.

**Bemerkung 3.4.26.** Ist  $M \subset \mathbb{R}^n$  eine Mannigfaltigkeit und  $T_p M$  der Tangentialraum in  $p \in M$ , dann nennen wir den affinen Raum  $p + T_p M \subset \mathbb{R}^n$  den geometrischen Tangentialraum. In der Regel werden wir diese beiden Räume miteinander identifizieren.

**Beispiel 3.4.27.** Es sei  $c : I \rightarrow \mathbb{R}^n$  derart, dass ihr Bild  $M = c(I)$  eine Mannigfaltigkeit ist. Dann ist der Tangentialraum im Punkt  $c(t) \in M$  gegeben durch  $T_{c(t)} M = \mathbb{R}c'(t)$  und der geometrische Tangentialraum stimmt mit der Tangente überein:  $c(t) + \mathbb{R}c'(t)$ .

Es folgt noch eine nützliche Beziehung zwischen Tangentialvektoren und Kurven auf Mannigfaltigkeiten:

**Satz 3.4.28.** 1. Es sei  $c : I \rightarrow M \subset \mathbb{R}^n$  eine Kurve auf der Mannigfaltigkeit  $M$  die auf dem offenen Intervall  $I \subset \mathbb{R}$  definiert ist. Dann ist  $c'(t) \in T_{c(t)}M$  für alle  $t \in M$ .

2. ist  $v \in T_pM \subset \mathbb{R}^n$ , dann gibt es eine Kurve  $c : ]-\epsilon, \epsilon[ \rightarrow \mathbb{R}^n$  mit  $c(0) = p$  und  $c'(0) = v$ .

**Satz 3.4.29.** Ist eine Mannigfaltigkeit  $M \subset \mathbb{R}^n$  als Nullstellenmenge einer Abbildung  $F : \mathbb{R}^n \supset U \rightarrow \mathbb{R}^{n-k}$  gegeben, dann ist

$$T_pM = \ker(DF(p))$$

für alle  $p \in M$ .

**Satz 3.4.30.** Es sei  $M \subset \mathbb{R}^n$  eine Mannigfaltigkeit der Dimension  $k$ . Weiter sei  $p \in M$ ,  $\Phi : U \rightarrow M$  eine Parametrisierung um  $p$  und  $u = \Phi^{-1}(p)$ .

Die Abbildung  $D\Phi(u) : \mathbb{R}^k \rightarrow T_pM \subset \mathbb{R}^n$  bildet den Würfel  $[0, 1]^k$  auf ein Parallelepiped  $P_{\Phi(u)} \subset T_pM$  ab, dessen Seiten durch die Spalten von  $D\Phi(u)$ , also durch  $\left\{ \frac{\partial \Phi}{\partial u_1}(u), \dots, \frac{\partial \Phi}{\partial u_k}(u) \right\}$ , gegeben sind.

Nun sei  $\{w_1, \dots, w_{n-k}\}$  eine Orthonormalbasis des orthogonalen Komplements von  $T_pM$  in  $\mathbb{R}^n$ . Dann lässt sich das Volumen des obigen Parallelepipeds berechnen zu

$$\begin{aligned} \text{vol}(P_{\Phi(u)}) &= \left| \det \left( \frac{\partial \Phi}{\partial u_1}(u), \dots, \frac{\partial \Phi}{\partial u_k}(u), w_1, \dots, w_{n-k} \right) \right| \\ &= \sqrt{\det \begin{pmatrix} \left\langle \frac{\partial \Phi}{\partial u_1}(u), \frac{\partial \Phi}{\partial u_1}(u) \right\rangle & \cdots & \left\langle \frac{\partial \Phi}{\partial u_1}(u), \frac{\partial \Phi}{\partial u_k}(u) \right\rangle \\ \vdots & & \vdots \\ \left\langle \frac{\partial \Phi}{\partial u_k}(u), \frac{\partial \Phi}{\partial u_1}(u) \right\rangle & \cdots & \left\langle \frac{\partial \Phi}{\partial u_k}(u), \frac{\partial \Phi}{\partial u_k}(u) \right\rangle \end{pmatrix}}. \end{aligned}$$

**Definition 3.4.31.** Ist  $\Phi : U \rightarrow M$  eine Parametrisierung der Mannigfaltigkeit  $M$ , dann bezeichnen wir im Folgenden die Funktion die  $u \in U$  den Wert  $\text{vol}(P_{\Phi(u)})$  zuordnet mit

$$g_\Phi : U \rightarrow \mathbb{R}, \quad g_\Phi(u) = \sqrt{\det(D\Phi(u)^T \cdot D\Phi(u))}.$$

**Beispiel 3.4.32.** 1. Es sei  $M \subset \mathbb{R}^3$  eine Fläche und  $\Phi : \mathbb{R}^2 \supset U \rightarrow M$  eine Parametrisierung um  $p \in M$  und  $p = \Phi(u)$ . Weiter sei  $P_{\Phi(u)}$  das von  $\frac{\partial \Phi}{\partial u_1}(u)$  und  $\frac{\partial \Phi}{\partial u_2}(u)$  aufgespannte Parallelepiped. Dann ist

$$\begin{aligned} g_{\Phi}(u) &= \sqrt{\left\| \frac{\partial \Phi}{\partial u_1}(u) \right\| \left\| \frac{\partial \Phi}{\partial u_2}(u) \right\| - \left\langle \frac{\partial \Phi}{\partial u_1}(u), \frac{\partial \Phi}{\partial u_2}(u) \right\rangle^2} \\ &= \left\| \frac{\partial \Phi}{\partial u_1}(u) \times \frac{\partial \Phi}{\partial u_2}(u) \right\|. \end{aligned}$$

2. Es sei  $M \subset \mathbb{R}^n$  eine eindimensionale Mannigfaltigkeit und  $c : I \rightarrow M$  eine Parametrisierung – also insbesondere eine Kurve auf  $M$ . Dann ist

$$g_c(t) = \|c'(t)\|$$

und damit die Länge des Tangentialvektors an diese Kurve im Punkt  $c(t)$ .

**Bemerkung 3.4.33.** Es seien  $\Phi : U \rightarrow M$  und  $\Psi : V \rightarrow M$  Parametrisierungen der Mannigfaltigkeit  $M$  um den gleichen Punkt  $p \in M$  mit Urbildern  $u = \Phi^{-1}(p) \in U$  und  $v = \Psi^{-1}(p) \in V$ , siehe Abbildung 3.4.3. Dann ist

$$g_{\Psi}(v) = \left| \det \left( D(\Phi^{-1} \circ \Psi)(v) \right) \right| g_{\Phi}(u).$$

Das sieht man, wenn man  $\Psi = \Phi \circ (\Phi^{-1} \circ \Psi)$  ableitet.

## 3.5 Integralsätze und die Existenz von Potentialen

### 3.5.1 Integration über Mannigfaltigkeiten im $\mathbb{R}^n$

Ein Spezialfall der Integration über Mannigfaltigkeiten wird auch hier wieder die Messung ihres Volumens sein – für eine Fläche im  $\mathbb{R}^3$  somit die Bestimmung des Flächeninhalts. Wir verfolgen hier die Idee, die wir bei der Integration stets verfolgt haben: Wir zerlegen die Mannigfaltigkeit in kleine Stücke und ersetzen diese durch Objekte, von denen wir das Volumen bestimmen können. Mit den Ergebnissen des vergangenen Abschnitts bieten sich Parallelepipede

an. Von diesen können wir das Volumen einfach bestimmen, da sie als Bilder von Würfeln im  $\mathbb{R}^k$  konstruiert sind.

Wir betrachten eine Parametrisierung  $\Phi : \mathbb{R}^k \supset U \rightarrow \mathbb{R}^n$  der Mannigfaltigkeit und im Parameterbereich einen Würfel  $W \subset U$ . Wir ersetzen nun das Bild  $\Phi(W)$  durch das Parallelepipèd, das wir erhalten, wenn wir  $\Phi$  durch seine Taylorentwicklung erster Ordnung um den Punkt  $u \in W$  ersetzen. Ist  $W$  "klein", so ist zu erwarten, dass der Fehler dieser Ersetzung ebenfalls "klein" ist, und wir erhalten für Volumen des Bildes des Würfels  $W \subset U$  um  $u$

$$\text{vol}(\Phi(W)) \approx g_{\Phi}(u)\text{vol}(W).$$

Dieser geometrische Ansatz motiviert die folgende Definition:

**Definition 3.5.1.** Es sei  $M \subset \mathbb{R}^n$  eine  $k$ -dimensionale Mannigfaltigkeit und  $\Phi : \mathbb{R}^k \supset U \rightarrow U' := \Phi(U) \subset M$  eine Parametrisierung. Eine Funktion  $f : M \rightarrow \mathbb{R}$  heißt integrierbar über  $U'$ , wenn der Wert

$$\int_{U'} f(x) dS_x := \int_U (f \circ \Phi)(u) g_{\Phi}(u) d^k u$$

existiert. Dieser heißt das Integral von  $f$  über  $U'$ .

**Bemerkung 3.5.2.** Wenn das Integral gemäß Definition 3.5.1 existiert, dann ist der Wert unabhängig von der Wahl der Karte, die  $U' \subset M$  vollständig beschreibt. Dies ist eine Folgerung aus der Transformationsformel aus Satz 3.3.20.

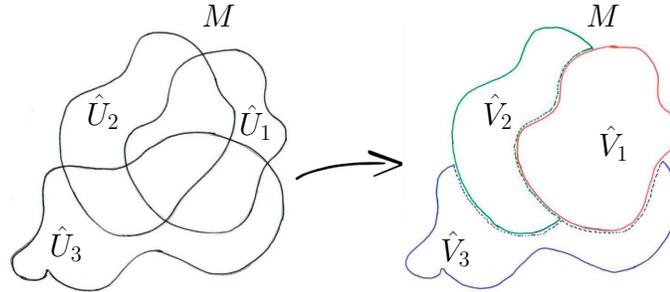
Wenn die Mannigfaltigkeit nicht von einer Karte überdeckt wird verwenden wir die folgende Idee: Ersetze die Kartenüberdeckung von  $M$  durch eine disjunkte Überdeckung mit eventuell nicht-offenen Mengen, auf denen die Einschränkungen der vorgegebenen Karten definiert sind.

Es sei  $\left\{ \Phi_i : U_i \rightarrow \hat{U}_i = \Phi_i(U_i) \subset M \right\}_{i=1, \dots, \nu}$  ein endlicher Atlas der  $k$ -dimensionalen Mannigfaltigkeit  $M \subset \mathbb{R}^n$ . Wir konstruieren die Mengen

$$\hat{V}_1 := \hat{U}_1 \quad \hat{V}_i := \hat{U}_i \setminus (\hat{U}_1 \cup \dots \cup \hat{U}_{i-1}) \quad i = 2, \dots, \nu,$$

siehe Abbildung 3.5.1 – die gestrichelten Randteile gehören zur jeweiligen Menge dazu. Dann ist  $\{\hat{V}_i\}_{i=1, \dots, \nu}$  ebenfalls eine Überdeckung von  $M$  und wir nennen sie die dem Atlas zugeordnete disjunkte Überdeckung. Schreiben wir nun  $V_i := \Phi_i^{-1}(\hat{V}_i) \subset \mathbb{R}^k$ , dann bleibt  $\Phi|_{V_i} : V_i \rightarrow M$  eine Parametrisierung im umgangssprachlichen Sinn.

Abbildung 3.5.1: Konstruktion der disjunkten Überdeckung



Das führt zu der folgenden naheliegenden Definition.

**Definition 3.5.3.** Es sei  $M \subset \mathbb{R}^n$  eine  $k$ -dimensionale Mannigfaltigkeit mit einem endlichen Atlas  $\{\Phi_i : U_i \rightarrow M\}_{i=1, \dots, \nu}$ . Weiter sei  $\{\hat{V}_i\}_{i=1, \dots, \nu}$  die zugeordnete disjunkte Überdeckung von  $M$  und  $V_i := \Phi_i^{-1}(\hat{V}_i)$ .

1. Eine Funktion  $f$  heißt *integrierbar über  $M$* , wenn der Wert

$$\int_M f(x) dS_x := \sum_{i=1}^{\nu} \int_{\hat{V}_i} f(x) dS_x = \sum_{i=1}^{\nu} \int_{V_i} (f \circ \Phi_i)(u) g_{\Phi_i}(u) d^k u$$

existiert. Er heißt das *Integral von  $f$  über  $M$* .

2. Ist  $f : M \rightarrow \mathbb{R}$  eine Funktion und ist  $A \subset M$  eine Teilmenge, dann heißt

$$\int_A f(x) dS_x := \int_M \chi_A(x) f(x) dS_x$$

das *Integral von  $f$  über  $A$* , falls dieser Wert existiert.

3. Ist die konstante Funktion  $f = \underline{1}$  integrierbar über  $M$ , dann heißt

$$\text{vol}(M) := \int_M dS_x := \int_M 1 dS_x$$

das *Volumen von  $M$* . Analog heißt für eine Teilmenge  $A \subset M$

$$\text{vol}(A) := \int_A dS_x$$

das *Volumen von  $A$* , falls das Integral existiert.

**Bemerkung 3.5.4.** Das in Definition 3.5.3 eingeführte Integral ist unabhängig von der Wahl des endlichen Atlanten.

Auch wenn Kurven und verallgemeinerte Flächen im Sinne des vorigen Abschnitts keine Mannigfaltigkeiten sein müssen, macht die Erweiterung des obigen Begriffs auf diese Objekte jedoch ebenfalls Sinn, siehe auch Satz 2.2.56.

**Definition 3.5.5.** 1. Ist  $c : [a, b] \rightarrow \mathbb{R}^n$  eine Kurve mit  $c' \neq 0$  – jedoch nicht notwendig die Parametrisierung einer eindimensionalen Mannigfaltigkeit – und ist  $f : U \rightarrow \mathbb{R}$  eine Funktion auf der Menge  $U \subset \mathbb{R}^n$ , wobei  $\text{Bild}(c) \subset U$ , so heißt

$$\int_c f ds := \int_a^b (f \circ c)(t) \|c'(t)\| dt$$

das *Kurvenintegral von  $f$  längs  $c$* .

2. Die Länge einer Kurve  $c : I \rightarrow \mathbb{R}^n$  ist gegeben durch

$$L(c) = \int_a^b \|c'(t)\| dt$$

und stimmt im Fall, dass die Kurve eine Mannigfaltigkeit liefert, mit ihrem Volumen überein.

3. Ist  $M \subset \mathbb{R}^3$  eine verallgemeinerte Fläche mit der Zerlegung  $M = \bigcup_{\alpha=1}^m M_\alpha$  gemäß Definition 3.4.21 und ist  $f : M \rightarrow \mathbb{R}$ , dann definieren wir

$$\int_M f(x) dS_x := \sum_{\alpha=1}^m \int_{\mathring{M}_\alpha} f(x) dS_x.$$

Das Volumen der verallgemeinerten Fläche ist dann

$$\text{vol}(M) = \sum_{\alpha=1}^m \text{vol}(\mathring{M}_\alpha).$$

Den Spezialfall eines Flächenstücks wollen wir wegen seiner Bedeutung in den kommenden Abschnitten hier noch hervorheben.

**Bemerkung 3.5.6.** Es sei  $M \subset \mathbb{R}^3$  ein Flächenstück, das durch seine Parametrisierung  $\Phi : U \rightarrow \mathbb{R}^3$  mit  $U \subset \mathbb{R}^2$  definiert ist.

1. In diesem Fall gilt

$$\int_M f(x) dS_x = \int_U (f \circ \Phi)(u, v) \left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\| d^2(u, v)$$

2. Der Vektor  $\frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v)$  steht als Kreuzprodukt senkrecht auf den von den beiden Ableitungen aufgespannten Tangentialraum  $T_p M$  im Punkt  $p = \Phi(u, v) \in M$ . Daher nennt man den Vektor

$$\vec{N}_\Phi(u, v) = \frac{1}{\left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\|} \left( \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right)$$

auch einen (*lokalen*) *Normalenvektor* der Fläche im Punkt  $p = \Phi(u, v)$ .

3. Den (*lokalen*) Normalenvektor kann man ganz analog auch für Hyperflächen im  $M \subset \mathbb{R}^n$  definieren: Ist  $\Phi : U \subset \mathbb{R}^{n-1} \rightarrow M$  eine Parametrisierung von  $M$ , so bilden  $\left\{ \frac{\partial \Phi}{\partial u_i}(u) \right\}_{i=1, \dots, n-1}$  eine Basis des Tangentialraums  $T_p M \subset \mathbb{R}^n$  in  $p = \Phi(u) \in M$ . Dann ist für alle  $u \in U$  ein Normalenvektor  $\vec{N}_\Phi(u)$  an  $M$  in  $p = \Phi(u)$  definiert durch die folgenden Eigenschaften

$$\begin{aligned} \|\vec{N}_\Phi(u)\| &= 1, & \vec{N}_\Phi(u) &\perp \frac{\partial \Phi}{\partial u_i}(u), \\ \det \left( \frac{\partial \Phi}{\partial u_1}(u), \dots, \frac{\partial \Phi}{\partial u_{n-1}}(u), \vec{N}_\Phi(u) \right) &> 0. \end{aligned}$$

**Definition 3.5.7.** 1. Eine Hyperfläche  $M \subset \mathbb{R}^n$  heißt *orientierbar*, wenn es ein stetiges Vektorfeld  $\nu : M \rightarrow \mathbb{R}^n$  gibt mit

$$\|\nu(p)\| = 1, \quad \nu(p) \perp T_p M$$

für alle  $p \in M$ .  $\nu$  heißt *Normalenfeld auf  $M$* , und es gilt stets  $\nu(p) = \pm \vec{N}_\Phi(\Phi^{-1}(p))$ .

2. Ist eine Hyperfläche orientierbar, so nennt man die Wahl eines der beiden existierenden Normalenfelder eine *Orientierung*. Fixiert man eine Wahl, so heißt die Hyperfläche *orientiert*.

**Satz 3.5.8.** *Eine Hyperfläche im  $\mathbb{R}^n$  ist genau dann orientierbar, wenn es einen Atlas gibt, für den die Funktionalmatrizen aller Übergangsabbildungen positive Determinante haben.*

**Bemerkung 3.5.9.** Eine beschränkte, allgemeine Fläche  $M \subset \mathbb{R}^3$  mit  $\partial M = \emptyset$  ist orientierbar. Insbesondere ist jede ihrer Teilflächen orientierbar und das Normalenfeld kann so gewählt werden, dass er entweder in jedem Punkt nach außen zeigt oder in jedem Punkt nach innen zeigt.<sup>(i)</sup>

### 3.5.2 Vektorfelder und Potentiale

**Definition 3.5.10.** Es sei  $M \subset \mathbb{R}^n$  eine Teilmenge. Ein *Vektorfeld auf  $M$*  ist eine Abbildung  $X : M \rightarrow \mathbb{R}^n$ . Ist  $M$  eine Untermannigfaltigkeit des  $\mathbb{R}^n$ , so heißt das Vektorfeld  $X$  *tangential*, wenn  $X(p) \in T_p M \subset \mathbb{R}^n$  ist.

**Bemerkung 3.5.11.** Die folgenden drei Fälle für Vektorfelder werden uns im Besonderen interessieren:

- $M$  ist eine offene Teilmenge des  $\mathbb{R}^n$
- $M$  ist das Bild einer Kurve.
- $M$  ist eine  $(n - 1)$ -dimensionale Mannigfaltigkeit, also eine Hyperfläche.

Ist  $A$  keine offene Teilmenge des  $\mathbb{R}^n$ , so werden die behandelten Vektorfelder jedoch oft (lokal) als Einschränkung eines Vektorfeldes auf einer offenen Menge gegeben sein.

- Beispiel 3.5.12.**
1. Die Windgeschwindigkeit ist ein Vektorfeld: Sie ordnet jedem Punkt des Raumes die Geschwindigkeit des Windes in diesem Punkt zu.
  2. Die Windgeschwindigkeit an der Erdoberfläche lässt sich als Vektorfeld auf  $S_{\mathbb{R}}^2(0)$  auffassen.

---

<sup>(i)</sup>Eine beschränkte, allgemeine Fläche  $M \subset \mathbb{R}^3$  mit  $\partial M = \emptyset$  mit nur einer Zusammenhangskomponente teilt den  $\mathbb{R}^3$  in zwei Teilmengen, von denen die eine unbeschränkt ist – das Äußere von  $M$  – und eine beschränkt – das Innere von  $M$ . Bezeichnen wir das Innere von  $M$  mit  $N$ , so gilt  $M = \partial N$ .

Mit diesen Bezeichnungen zeigt der Normalenvektor  $\nu(x)$  von  $M$  nach aussen, wenn es ein  $\epsilon > 0$  gibt, sodass  $x + t\nu(x) \notin N$  für alle  $0 < t < \epsilon$ .

3. Ist  $c : I \rightarrow M \subset \mathbb{R}^n$  eine Kurve auf der Mannigfaltigkeit  $M$ , dann ist die Abbildung  $X : I \rightarrow \mathbb{R}^n$  mit  $X(t) = c'(t)$  tangential in dem Sinne, dass  $X(t) \in T_{c(t)}M$ .
4. Ist allgemeiner  $c : I \rightarrow \mathbb{R}^n$  eine Kurve, dann nennt man eine Abbildung  $X : I \rightarrow \mathbb{R}^n$  auch *Vektorfeld längs  $c$* .

Ein Spezialfall ergibt sich wie folgt: Ist  $X : M \rightarrow \mathbb{R}^n$  ein Vektorfeld auf der Mannigfaltigkeit  $M \subset \mathbb{R}^n$  und ist  $c : I \rightarrow M$  eine Kurve auf  $M$ , so ist  $X \circ c : I \rightarrow \mathbb{R}^n$  ein Vektorfeld längs  $c$ . Ist weiter  $X$  als Einschränkung eines auf einer offenen Menge  $U \subset \mathbb{R}^n$  definierten Vektorfelds  $Y : U \rightarrow \mathbb{R}^n$  gegeben, also  $X = Y|_M$ , so ist

$$(X \circ c)'(t) = DY(c(t)) \cdot c'(t).$$

5. Ist  $f : U \rightarrow \mathbb{R}$  eine differenzierbare Funktion auf der offenen Teilmenge  $U \subset \mathbb{R}^n$ , dann ist der Gradient von  $f$

$$\nabla f := \text{grad}(f) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} : U \rightarrow \mathbb{R}^n$$

ein Vektorfeld.

Das letzte Beispiel ist so zentral, dass wir diesem eine eigene Bezeichnung geben

**Definition 3.5.13.** Es sei  $X : U \rightarrow \mathbb{R}^n$  ein Vektorfeld auf der offenen Teilmenge  $U \subset \mathbb{R}^n$  für das es eine Funktion  $f : U \rightarrow \mathbb{R}$  gibt mit  $X(x) = \nabla f(x)$  für alle  $x \in U$ . Dann nennen wir  $X$  *konservatives Vektorfeld* oder *Gradientenfeld* und die Funktion  $f$  heißt *ein Potential von  $X$* .

**Beispiel 3.5.14.** Es sei  $g : \mathbb{R} \rightarrow \mathbb{R}$  eine integrierbare Funktion und  $X : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n$  gegeben durch

$$X(x) = \frac{g(\|x\|)}{\|x\|} x.$$

Dann ist  $X$  konservativ. Ist nämlich  $v : \mathbb{R} \rightarrow \mathbb{R}$  eine Stammfunktion von  $g$ , so ist die Funktion

$$f(x) = v(\|x\|)$$

ein Potential von  $X$ .

Bisher haben wir uns darauf beschränkt, Funktionen zu integrieren. Wir wollen die Integration nun in speziellen Situationen auch für Vektorfelder erklären.

**Definition 3.5.15.** 1. Es sei  $X : I \rightarrow \mathbb{R}^n$  ein Vektorfeld längs der Kurve  $c : I \rightarrow \mathbb{R}^n$ . Dann heißt

$$\int_c X d\vec{s} := \int_I \langle X(t), c'(t) \rangle dt$$

das *Linienintegral* von  $X$  längs  $c$ .

2. Es sei  $M \subset \mathbb{R}^n$  eine orientierbare Hyperfläche mit fest gewähltem Normalenfeld  $\nu$  und  $X$  ein Vektorfeld auf  $M$ . Dann heißt

$$\int_M X d\vec{S} := \int_M \langle X(p), \nu(p) \rangle dS_p$$

das *Flussintegral* von  $X$  über  $M$ .

Die naheliegende Verallgemeinerungen dieser Definitionen für stückweise differenzierbare Kurven und verallgemeinerte Flächen geschieht über Summierung der beteiligten Teilstücke.

**Beispiel 3.5.16.** 1. Ist  $M \subset \mathbb{R}^3$  eine Fläche, die durch eine einzige Parametrisierung  $\Phi : U \rightarrow \mathbb{R}^3$  beschrieben wird, für die  $\vec{N} = \nu$  ist, so gilt

$$\begin{aligned} \int_M X d\vec{S} &= \int_U \left\langle X(\Phi(u, v)), \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\rangle d^2(u, v) \\ &= \int_U \det \left( X(\Phi(u, v)), \frac{\partial \Phi}{\partial u}(u, v), \frac{\partial \Phi}{\partial v}(u, v) \right) d^2(u, v). \end{aligned}$$

Der Spezialfall  $X = \nu$  liefert

$$\int_M \nu(x) d\vec{S} = \text{vol}(M).$$

2. Wir betrachten den Fall  $n = 2$ . In diesem Fall wird eine Hyperfläche  $C$  durch eine Kurve  $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} : I \rightarrow C$  parametrisiert. Ein Normalenfeld für diese "Hyperfläche" ist durch den Vektor  $\nu = -\frac{1}{\|c'_1\|} c'_1 = \frac{1}{\|c'\|} \begin{pmatrix} c'_2 \\ -c'_1 \end{pmatrix}$

gegeben<sup>(i)</sup>. Damit ist

$$\begin{aligned}\int_C \begin{pmatrix} P \\ Q \end{pmatrix} d\vec{S} &= \int_I \langle \begin{pmatrix} P \\ Q \end{pmatrix}, \nu \rangle dS_t = \int_I \langle \begin{pmatrix} P \\ Q \end{pmatrix}, \nu \rangle \|c'\| dt \\ &= \int_I \langle \begin{pmatrix} P \\ Q \end{pmatrix}, \begin{pmatrix} c'_2 \\ -c'_1 \end{pmatrix} \rangle dt = \int_I \langle \begin{pmatrix} -Q \\ P \end{pmatrix}, c' \rangle dt = \int_c \begin{pmatrix} -Q \\ P \end{pmatrix} d\vec{s}\end{aligned}$$

**Bemerkung 3.5.17.** • Die Definition des Linienintegrals hängt nicht von der Wahl der Parametrisierung der Kurve ab. Dies gilt allerdings nur, wenn man die Orientierung der Kurve beibehält; andernfalls ändert sich das Vorzeichen.

- Genauso ändert sich das Vorzeichen im Flussintegral, wenn man die Wahl des Normalenvektors – also die *Orientierung* der Fläche – ändert.

Aus dem Hauptsatz der Differential- und Integralrechnung ergibt sich direkt das folgende Resultat.

**Satz 3.5.18.** *Ist  $f$  ein Potential des Vektorfeldes  $X : U \rightarrow \mathbb{R}^n$  auf der offenen Menge  $U \subset \mathbb{R}^n$  und ist  $c : [a, b] \rightarrow U$  ein Weg, dann hängt  $\int_c X d\vec{s}$  nur vom Anfangs- und Endpunkt von  $c$  ab. Es gilt nämlich*

$$\int_c \nabla f d\vec{s} = f(c(b)) - f(c(a)).$$

Wie man sich leicht überlegt, ist diese Aussage für beliebige Vektorfelder in der Regel falsch. Dazu wählen wir  $X(x, y) = \begin{pmatrix} x \\ xy \end{pmatrix}$  und  $c, \tilde{c} : [0, 1] \rightarrow \mathbb{R}^2$  mit  $c(t) = (t, t)$  und  $\tilde{c} = (t^2, t)$ . Beide Kurven verbinden die Punkte  $(0, 0)$  und  $(1, 1)$  und es gilt  $\int_c X d\vec{s} = \int_0^1 (t + t^2) dt = \frac{5}{6}$  aber  $\int_{\tilde{c}} X d\vec{s} = \int_0^1 (2t^3 + t^3) dt = \frac{3}{4}$ .

Wir werden später sehen, dass die Eigenschaft aus Satz 3.5.18 sogar eine Charakterisierung der konservativen Vektorfelder liefert.

Der vorige Satz hat ein wenig Ähnlichkeit mit der Beziehung zwischen Ableitung und Integration von Funktionen, siehe Definition/Bemerkung 2.2.4. Wir wollen untersuchen, ob es solche Zusammenhänge auch für die Ableitung von Vektorfeldern gibt.

<sup>(i)</sup>Interpretiert man  $c : I \rightarrow C \subset \mathbb{R}^2$  als Parametrisierung der Kurve  $C$ , dann ist  $\frac{1}{\|c'\|} c'_\perp$  der (lokale) Normalenvektor der "Hyperfläche"  $C$  gemäß Bemerkung 3.5.6.3

Wir wissen, dass die Ableitung eines Vektorfeldes  $X : U \rightarrow \mathbb{R}^n$  auf eine offenen Menge  $U \subset \mathbb{R}^n$  in jedem Punkt seines Definitionsbereiches eine lineare Abbildung  $DX(x) \in M_n \mathbb{R}$  definiert. Diese definiert zwei spezielle Ableitungsoperationen.

**Definition 3.5.19.** Es sei  $X : U \rightarrow \mathbb{R}^n$  ein differenzierbares Vektorfeld auf eine offenen Menge  $U \subset \mathbb{R}^n$ . Dann sind *Divergenz*  $\operatorname{div} X$  und *Rotation*  $\operatorname{Rot} X$  von  $X$  wie folgt definiert

1.  $\operatorname{div} X : U \rightarrow \mathbb{R}$

$$\operatorname{div} X = \operatorname{Spur}(DX) = \sum_{i=1}^n \frac{\partial X^i}{\partial x_i}.$$

2.  $\operatorname{Rot} X : U \rightarrow \operatorname{Bil}_{\text{skew}}(\mathbb{R}^3)$

$$\operatorname{Rot} X(\vec{v}, \vec{w}) = \langle DX\vec{v}, \vec{w} \rangle - \langle \vec{v}, DX\vec{w} \rangle.$$

Die Rotation ist in niedrigen Dimensionen  $n = 2, 3$  durch einfachere Vorschriften beschreibbar:

**Bemerkung 3.5.20.** 1. Ist  $X$  ein Vektorfeld und bezeichnet  $DX$  die Funktionalmatrix mit  $(DX)_{ij} = \frac{\partial X^i}{\partial x_j}$  für  $1 \leq i, j \leq n$ , dann hat  $\operatorname{Rot} X$  die Matrixdarstellung  $DX^T - DX$  also  $(\operatorname{Rot} X)_{ij} = \frac{\partial X^j}{\partial x_i} - \frac{\partial X^i}{\partial x_j}$ .

2. In Dimension 2 ist der Raum der schiefsymmetrischen bilinearen Abbildungen eindimensional. Da die Determinante eine solche Abbildung ist, gibt es eine Funktion  $\operatorname{rot}^{(2)} X : U \rightarrow \mathbb{R}$ , so dass

$$\operatorname{Rot} X = \operatorname{rot}^{(2)} X \cdot \det$$

mit  $\operatorname{rot}^{(2)} X = \operatorname{Rot} X(\vec{e}_1, \vec{e}_2)$ . Schreiben wir  $X(x, y) = \begin{pmatrix} P(x, y) \\ Q(x, y) \end{pmatrix}$ , so gilt

$$\operatorname{rot}^{(2)} X = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

3. In Dimension 3 ist der Raum der schiefsymmetrischen bilinearen Abbildungen dreidimensional. Es gibt somit einen Isomorphismus  $\Omega : \operatorname{Bil}_{\text{skew}}(\mathbb{R}^3) \rightarrow \mathbb{R}^3$ . Dieser Isomorphismus ist explizit durch

$$B(\vec{v}, \vec{w}) = \langle \Omega(B), \vec{v} \times \vec{w} \rangle$$

gegeben.<sup>(i)</sup> Wir setzen  $\text{rot}^{(3)}X = \Omega(\text{Rot}X)$  und schreiben wir  $X(x, y, z)$

$$= \begin{pmatrix} P(x, y, z) \\ Q(x, y, z) \\ R(x, y, z) \end{pmatrix}, \text{ so gilt}$$

$$\text{rot}^{(3)}X = \begin{pmatrix} \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \\ \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \\ \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \end{pmatrix}.$$

**Bemerkung 3.5.21.** Die Operationen Rotation, Divergenz und Gradient erfüllen die folgenden Rechenregeln, wobei  $f, g$  stets Funktionen und  $X, Y$  Vektorfelder sind:

1.  $\text{div } \nabla f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} =: \Delta f$
2.  $\text{Rot } \nabla f = 0$  in allen Dimensionen
3.  $\text{div } \text{rot}^{(3)}X = 0$  in Dimension 3
4.  $\nabla \text{div}X - \text{rot}^{(3)} \text{rot}^{(3)}X = \Delta X$  in Dimension 3, wobei die rechte Seite komponentenweise zu verstehen ist.

Außerdem gelten die folgenden Produktregeln

5.  $\nabla(fg) = g\nabla f + f\nabla g$  in allen Dimensionen
6.  $\text{div}(fX) = f \text{div}X + \langle \nabla f, X \rangle$  in allen Dimensionen
7.  $\text{Rot}(fX) = f \text{Rot}X + X \cdot (\nabla f)^T - \nabla f \cdot X^T$  in allen Dimensionen
  - $\text{rot}(fX) = f \text{rot}^{(3)}X + \nabla f \times X$  in Dim. 3
  - $\text{rot}(fX) = f \text{rot}^{(2)}X + \det(\nabla f, X)$  in Dim. 2

---

<sup>(i)</sup>Hat  $B$  die Matrixdarstellung  $\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$ , so ist  $\Omega(B) = \begin{pmatrix} c \\ -b \\ a \end{pmatrix}$ .

8.  $\nabla\langle X, Y \rangle = DY^T \cdot X + DX^T \cdot Y$  in allen Dimensionen
- $\nabla\langle X, Y \rangle = X \times \operatorname{rot}^{(3)}Y + Y \times \operatorname{rot}^{(3)}X + DY \cdot X + DX \cdot Y$  in Dim. 3
9.  $\operatorname{div}(X \times Y) = \langle Y, \operatorname{rot}X \rangle - \langle X, \operatorname{rot}Y \rangle$  in Dim. 3
10.  $\operatorname{rot}^{(3)}(X \times Y) = \operatorname{div}Y X - \operatorname{div}X Y - DY \cdot X + DX \cdot Y$  in Dim. 3

Geben wir eine Abbildung von  $\mathbb{R}^2$  nach  $\mathbb{R}^3$  vor, so lassen sich die Operationen  $\operatorname{rot}^{(2)}$  und  $\operatorname{rot}^{(3)}$  verbinden:

- 11.a Es sei  $X : U \rightarrow \mathbb{R}^3$  ein Vektorfeld auf der offenen Menge  $U \subset \mathbb{R}^3$  und  $\Phi : V \rightarrow U$  zweimal stetig differenzierbar auf der offenen Menge  $V \subset \mathbb{R}^2$ . Wir definieren  $\hat{X} : V \rightarrow \mathbb{R}^3$  durch

$$\hat{X}(v) = D\Phi(v)^T X(\Phi(v)).$$

Dann gilt

$$\operatorname{rot}^{(2)}\hat{X}(v) = \left\langle \operatorname{rot}^{(3)}X(\Phi(v)), \frac{\partial\Phi}{\partial v_1}(v) \times \frac{\partial\Phi}{\partial v_2}(v) \right\rangle.$$

- 11.b Indem wir  $\mathbb{R}^3$  in diesem Beispiel durch  $\mathbb{R}^n$  ersetzen, erhalten wir analog

$$\operatorname{rot}^{(2)}\hat{X}(v) = \operatorname{Rot}X(\Phi(v)) \left( \frac{\partial\Phi}{\partial v_1}(v), \frac{\partial\Phi}{\partial v_2}(v) \right).$$

**Folgerung 3.5.22.** Ist  $U \subset \mathbb{R}^n$  und  $X : U \rightarrow \mathbb{R}^n$  konservativ, so ist  $\operatorname{Rot}X = 0$ .

Wir werden zum Abschluss dieses Abschnitts die Umkehrung von Satz 3.5.18 beweisen:

**Satz 3.5.23.** Für ein Vektorfeld  $X : U \rightarrow \mathbb{R}^n$  auf der offenen, zusammenhängenden Menge  $U \subset \mathbb{R}^n$  sind die folgenden Aussagen äquivalent

1.  $X$  ist ein Gradientenfeld.
2.  $\int_c X d\vec{s}$  hängt nur vom Anfangs- und Endpunkt der Kurve  $c$  ab.

$$3. \int_c X d\vec{s} = 0 \text{ für alle geschlossenen Kurven } c.$$

Insbesondere ist das Potential bis auf eine additive Konstante eindeutig bestimmt.

Für die Umkehrung von Folgerung 3.5.22 benötigen wir noch ein wenig Vorbereitung und verschieben sie deshalb auf die folgenden Abschnitte.

### 3.5.3 Die geometrische Interpretation der Divergenz und der Satz von Gauß

Wir wollen zu Anfang dieses Abschnitts direkt eine recht einfache Situation diskutieren

**Satz 3.5.24.** *Es sei  $U \overset{\circ}{\subset} \mathbb{R}^n$  und  $X : U \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld. Weiter sei  $Q \subset \mathbb{R}^n$  ein achsenparalleler Quader. Dann gilt*

$$\int_{\partial Q} X d\vec{S} = \int_Q \operatorname{div} X d^n x.$$

Hierbei ist im linken Integral das Normalenfeld der orientierbaren, allgemeinen Hyperfläche  $\partial Q$  als äußeres Normalenfeld zu wählen.

Eine Verallgemeinerung ergibt sich unmittelbar.

**Folgerung 3.5.25.** *Besitzt  $R \subset \mathbb{R}^n$  eine Quaderüberdeckung  $R = \bigcup_{\alpha=1}^k Q_\alpha$ , dann gilt*

$$\int_{\partial R} X d\vec{S} = \int_R \operatorname{div} X d^n x.$$

für alle stetig differenzierbaren Vektorfelder  $X$  auf  $R$ .

Mit Hilfe des Mittelwertsatzes der Integralrechnung erhalten wir als Folgerung eine Beziehung zwischen der Divergenz eines Vektorfeldes und dem Flussintegral.

**Folgerung 3.5.26.** *Es sei  $U \overset{\circ}{\subset} \mathbb{R}^n$  und  $X : U \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld. Für alle  $x \in U$  gilt*

$$\operatorname{div} X(x) = \lim_{|Q| \rightarrow 0} \frac{1}{\operatorname{vol}(Q)} \int_{\partial Q} X d\vec{S}.$$

Hierbei sind  $Q \subset \mathbb{R}^n$  achsenparallele Würfel mit  $x \in \overset{\circ}{Q}$  und  $|Q|$  bezeichne ihren Durchmesser. Die Integration erfolgt mit Hilfe des äußeren Normalenfeldes von  $\partial Q$ .

- Definition 3.5.27** (Erlaubte Bereiche im  $\mathbb{R}^n$ ). 1. Ein *erlaubter Bereich*  $B \subset \mathbb{R}^2$  ist eine kompakte Menge, deren Rand eine endliche Vereinigung von stückweise differenzierbaren, einfach geschlossenen Kurven ist.
2. Ein *erlaubter Bereich*  $B \subset \mathbb{R}^3$  ist eine beschränkte Teilmenge, sodass  $\partial B$  eine orientierte allgemeine Fläche ohne Rand ist, für die jede Teilfläche über einem erlaubten Bereich im  $\mathbb{R}^2$  parametrisiert ist.<sup>(i)</sup>
3. Wir fahren induktiv fort und definieren so *erlaubte Bereiche* im  $\mathbb{R}^n$ .

**Beispiel 3.5.28.** • Alle Bälle, Ellipsoide, Zylinder und Polyeder sind erlaubte Bereiche.

- Ist  $M \subset \mathbb{R}^n$  eine orientierte, beschränkte Hyperfläche ohne Rand, so ist ihr Inneres ein erlaubter Bereich, der von  $M$  berandet wird.

**Bezeichnung 3.5.29.** Ein erlaubter Bereich  $B \subset \mathbb{R}^n$  mit nach außen gerichtetem Normalenfeld heißt *Gauß-Menge*, wenn

$$\int_{\partial B} X d\vec{S} = \int_B \operatorname{div} X d^n x$$

für alle stetig differenzierbaren Vektorfelder  $X$  auf  $B$ .

Mit dieser Bezeichnung sind Quader und Mengen, die eine Quaderzerlegung zulassen, Gauß-Mengen. "Verzerren" wir eine Gauß-Menge, so erhalten wir wieder eine solche. Dazu brauchen wir noch die folgende Bemerkung.

**Bemerkung 3.5.30** (Transformationsformel für Flussintegrale).

1. Es sei  $X : U \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld auf der offenen Menge  $U \subset \mathbb{R}^n$  und  $\Psi : \hat{U} \rightarrow U$  ein zweimal stetig differenzierbarer Diffeomorphismus. Für das Vektorfeld  $\hat{X} : \hat{U} \rightarrow \mathbb{R}^n$  mit

$$\hat{X}(\hat{x}) := \det(D\Psi(\hat{x})) (D\Psi(\hat{x}))^{-1} X(\Psi(\hat{x}))$$

<sup>(i)</sup>Für die Beweise benötigen wir, dass die Parametrisierungen der Seitenflächen mindestens zweimal stetig differenzierbar sind, in der Regel sind sie sogar glatt.

gilt

$$\operatorname{div} \hat{X}(\hat{x}) = \det(D\Psi(\hat{x})) \operatorname{div} X(\Psi(\hat{x})).$$

Hierbei ist die Divergenz links bezüglich der Koordinaten  $\hat{x}_1, \dots, \hat{x}_n$  und rechts bezüglich der Koordinaten  $x_1, \dots, x_n$  zu bilden.

2. Ist  $\partial U$  ein Hyperflächenstück, dann auch  $\partial \hat{U}$  und es gilt

$$\int_{\partial U} X d\vec{S}_x = \int_{\partial \hat{U}} \hat{X} d\vec{S}_{\hat{x}}$$

für alle stetig differenzierbaren Vektorfelder  $X$  auf  $\partial U$ .

**Satz 3.5.31.** 1. Ist  $B \subset \mathbb{R}^n$  eine Gauß-Menge und  $\Psi : \hat{B} \rightarrow B$  ein zweimal stetig differenzierbarer Diffeomorphismus mit  $\det(D\Psi) > 0$ , so ist  $\hat{B}$  eine Gauß-Menge.

2. Sind  $B_1, B_2$  Gauß-Mengen, und ist  $B_1 \cap B_2$  ein Flächenstück, dann ist auch  $B_1 \cup B_2$  eine Gauß-Menge.

Schließlich kommen wir zur Hauptaussage dieses Abschnitts, der besagt, dass alle erlaubten Bereiche Gauß-Mengen sind:

**Satz 3.5.32** (Satz von Gauß). Es sei  $B \subset \mathbb{R}^n$  ein erlaubter Bereich mit nach außen zeigendem Normalenfeld. Dann ist  $B$  ein Gauß-Bereich, das heißt

$$\int_{\partial B} X d\vec{S} = \int_B \operatorname{div} X d^n x$$

für alle stetig differenzierbaren Vektorfelder  $X$  auf  $B$ .

Zum Abschluss wollen wir noch einige Folgerungen aus dem Satz von Gauß angeben, die jede für sich in speziellen Anwendungen von Interesse ist.

- Betrachten wir zwei Funktionen  $f, g : B \rightarrow \mathbb{R}$  auf dem erlaubtem Bereich  $B \subset \mathbb{R}^n$  und setzen wir  $X = f\nabla g$  in Satz 3.5.32, so liefert das die Greenschen Formeln.

**Folgerung 3.5.33** (Greensche Formeln). Es sei  $B \subset \mathbb{R}^n$  ein erlaubter Bereich und  $f, g : B \rightarrow \mathbb{R}$  stetig differenzierbare Funktionen. Dann gelten die folgenden Formeln:

(1. Greensche Formel)

$$\int_{\partial B} f \nabla g \, d\vec{S} = \int_B (\langle \nabla f, \nabla g \rangle + f \Delta g) \, d^n x$$

(2. Greensche Formel)

$$\int_{\partial B} (f \nabla g - g \nabla f) \, d\vec{S} = \int_B (f \Delta g - g \Delta f) \, d^n x$$

mit dem Spezialfall

$$\int_{\partial B} \nabla f \, d\vec{S} = \int_B \Delta f \, d^n x$$

Mit Hilfe der Rechenregeln aus Bemerkung 3.5.21 bekommen wir noch weitere spezielle Formen des Satzes von Gauß.

- Betrachten wir ein Vektorfeld der Form  $f \vec{a}$  mit einem konstanten Vektor  $\vec{a} \in \mathbb{R}^n$  und einer Funktion  $f : B \rightarrow \mathbb{R}$ , so erhalten wir die folgende vektorielle Variante des Gaußschen Satzes für Funktionen.

**Folgerung 3.5.34.** Es sei  $B \subset \mathbb{R}^n$  ein erlaubter Bereich und  $f : B \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion. Dann gilt

$$\int_{\partial B} f \, d\vec{S} = \int_B \nabla f \, d^n x.$$

- In Dimension 3 konstruieren wir mit dem konstanten Vektor  $\vec{a}$  und dem Vektorfeld  $X : B \rightarrow \mathbb{R}^3$  ein neues Vektorfeld  $\vec{a} \times X$ . Das liefert die folgende vektorielle Variante des Gaußschen Satzes für Vektorfelder

**Folgerung 3.5.35.** Es sei  $B \subset \mathbb{R}^3$  ein erlaubter Bereich und  $X : B \rightarrow \mathbb{R}^3$  ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_{\partial B} d\vec{S} \times X = \int_{\partial B} (\nu \times X) \, dS = \int_B \operatorname{rot} X \, d^3 x$$

- Diese Variante hat in beliebiger Dimension eine Entsprechung mit Werten in den schiefsymmetrischen Matrizen. Zur Herleitung dieser Variante betrachten wir für das Vektorfeld  $X : B \rightarrow \mathbb{R}^n$  und für  $\vec{a}, \vec{b} \in \mathbb{R}^n$  das Vektorfeld  $\langle X, \vec{a} \rangle \vec{b} - \langle X, \vec{b} \rangle \vec{a}$ .

**Folgerung 3.5.36.** Es sei  $B \subset \mathbb{R}^n$  ein erlaubter Bereich und  $X : B \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_{\partial B} (X \cdot d\vec{S}^T - d\vec{S} \cdot X^T) = \int_{\partial B} (X\nu^T - \nu X^T) dS = \int_B \text{Rot}X \, d^n x.$$

### 3.5.4 Der Satz von Stokes und die Existenz von Potentialen

**Definition 3.5.37.** Ein erlaubter Bereich  $B \subset \mathbb{R}^2$  heißt *positiv berandet*, wenn alle Teile des Randes  $\partial B_\alpha$  so parametrisiert sind, dass stets das Innere  $\overset{\circ}{B}_\alpha$  beim Durchlaufen links liegt. Andernfalls heißt  $B$  *negativ berandet*.

**Bemerkung 3.5.38.** Ist  $c(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix}$  die Parametrisierung eines Teils des positiven Randes des erlaubten Bereichs  $B \subset \mathbb{R}^2$ , dann ist dort  $\nu(t) = \frac{1}{\|c'\|} \begin{pmatrix} c_2'(t) \\ -c_1'(t) \end{pmatrix}$  der äußere Normalenvektor des Kurvenstücks interpretiert als Hyperflächenstück im  $\mathbb{R}^2$ , siehe auch Beispiel 3.5.16.2.

Die folgende zweidimensionale Variante des Satzes von Gauß heißt wahlweise *Satz von Stokes in der Ebene* bzw. *Satz von Green*. Er ergibt sich mit Hilfe von Beispiel 3.5.16.2 aus Satz 3.5.32 oder direkt aus Folgerung 3.5.36.

**Folgerung 3.5.39.** Es sei  $B \subset \mathbb{R}^2$  ein positiv berandeter erlaubter Bereich und  $X : B \rightarrow \mathbb{R}^2$  ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_{\partial B} X \, d\vec{s} = \int_B \text{rot}X \, d^2(x, y).$$

Ist der Bereich negativ berandet, so stimmt die Aussage nur bis auf das Vorzeichen.

In Ergänzung zu Definition 3.5.7 und Bemerkung 3.5.9 geben wir hier einen Orientierungsbegriff für allgemeine Flächen.

**Definition 3.5.40.** 1. Ist  $M$  ein Flächenstück (mit Rand), das über einem erlaubten Bereich parametrisiert ist, dann heißt  $M$  *positiv berandet*, wenn für jeden Randpunkt  $p \in \partial M$  der Tangentialvektor der Randbeschreibung, ein beliebiger nach innen zeigender Tangentialvektor und der Normalenvektor eine positive Basis bilden.

2. Ein allgemeine Fläche (mit Rand)  $M = \bigcup M_\alpha \subset \mathbb{R}^3$  heißt *positiv berandet*, wenn
- alle Teilflächenstücke  $M_\alpha$  über erlaubten Bereichen  $B_\alpha \subset \mathbb{R}^2$  parametrisiert sind,
  - die Randorientierung der Parameterbereiche  $B_\alpha$  so gewählt ist, dass auf gemeinsamen Kanten im Bild die Orientierungen gegenläufig sind und
  - es ein Normalenfeld gibt, sodass alle  $M_\alpha$  positiv berandet sind.

**Bemerkung 3.5.41.** Schreiben wir  $X = \begin{pmatrix} P \\ Q \end{pmatrix}$ , das heißt  $\operatorname{rot} X = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$ , und ist der Rand  $\partial B$  von  $B$  parametrisiert durch eine Kurve  $c : [a, b] \rightarrow \mathbb{R}^2$ ,  $t \mapsto c(t) = (x(t), y(t))$ , dann schreibt sich die Formel aus dem vorigen Satz als

$$\int_a^b (P x' + Q y') dt = \int_B (Q_x - P_y) d^2(x, y).$$

Für das linke Integral findet man wegen  $\frac{dx}{dt} = x'$ ,  $\frac{dy}{dt} = y'$  auch die Schreibweise  $\int_c P dx + Q dy$ .

Mit Hilfe von Folgerung 3.5.39 und Bemerkung 3.5.21.11.a formulieren wir nun den Satz von Stokes für (allgemeine) Flächen im Raum.

**Satz 3.5.42** (Satz von Stokes im Raum). *Es sei  $M \subset \mathbb{R}^3$  eine positiv berandete allgemeine Fläche. Dann gilt*

$$\int_{\partial M} X d\vec{s} = \int_M \operatorname{rot} X d\vec{S}.$$

Als Folgerung erhält man hier analog zu Folgerung 3.5.26 die geometrische Interpretation der Rotation in Dimension 3.

**Folgerung 3.5.43.** Es sei  $U \subset \mathbb{R}^3$  und  $X : U \rightarrow \mathbb{R}^3$  ein stetig differenzierbares Vektorfeld. Für alle  $x \in U$  und konstante Einheitsvektoren  $\vec{n} \in \mathbb{R}^3$  gilt

$$\langle \vec{n}, \operatorname{rot} X(x) \rangle = \lim_{|Q| \rightarrow 0} \frac{1}{\operatorname{vol}(Q)} \int_{\partial Q} X d\vec{s}.$$

Hierbei sind  $Q \subset \mathbb{R}^3$  Quadrate in der Ebene  $E_{\vec{n}} := \{y \in \mathbb{R}^3 \mid \langle y, \vec{n} \rangle = \text{const}\}$  mit  $x \in Q \setminus \partial Q$  und  $|Q|$  bezeichne den Durchmesser des Quadrats. Für die Integration wählt man  $Q$  positiv berandet.

Wir haben nun das Rüstzeug um eine Umkehrung von Folgerung 3.5.22 zu formulieren. Dies klappt allerdings im Allgemeinen nur, wenn wir für den Definitionsbereich des Vektorfeldes die folgende Einschränkung machen.

**Definition/Bemerkung 3.5.44.** Es seien  $c_0 : [a, b] \rightarrow \mathbb{R}^n$  und  $c_1 : [a, b] \rightarrow \mathbb{R}^n$  zwei stetige Wege.

1. Eine stetige Abbildung  $H : [a, b] \times [0, 1] \rightarrow \mathbb{R}^n$  heißt *Homotopie zwischen*  $c_0$  und  $c_1$ , wenn

- $H(t, 0) = c_0(t)$  für alle  $t \in [a, b]$
- $H(t, 1) = c_1(t)$  für alle  $t \in [a, b]$

Für  $s_0 \in [0, 1]$  sind die *Zwischenkurven*  $c_{s_0} : [a, b] \rightarrow \mathbb{R}^n$  mit  $c_{s_0}(t) := H(t, s_0)$  selbst stetige Kurven. Existiert so eine Homotopie  $H$ , dann sagen wir  $c_0$  und  $c_1$  sind *homotop*.

2. Haben  $c_0$  und  $c_1$  den selben Anfangs- und Endpunkt und sind sie homotop, so heißt  $H$  eine *Homotopie mit festen Anfangs- und Endpunkt*, wenn

- $H(a, s) = c_0(a) = c_1(a)$  und  $H(b, s) = c_0(b) = c_1(b)$  für alle  $s \in [0, 1]$ .

3. Eine einfach geschlossene Kurve heißt *nullhomotop*, wenn sie homotop zu ihrem als konstanten Weg interpretiertem Anfangspunkt ist.

**Definition 3.5.45.** Eine zusammenhängende Menge  $A \subset \mathbb{R}^n$  heißt *einfach zusammenhängend*, wenn jede einfach geschlossene, stetige Kurve  $c : [a, b] \rightarrow A$  nullhomotop ist.

**Bemerkung 3.5.46.** 1. Eine Menge ist genau dann einfach zusammenhängend, wenn jeder einfach geschlossene Weg zu einem beliebigem Punkt  $x^0$  homotop ist.

2. Eine Teilmenge des  $\mathbb{R}^2$  ist genau dann einfach zusammenhängend, wenn für jede einfach geschlossene Kurve das Innere der Kurve ein erlaubter Bereich ist.

**Beispiel 3.5.47.** •  $B^n \subset \mathbb{R}^n$  ist einfach zusammenhängend.

- $B^n \setminus \{0\} \subset \mathbb{R}^n$  ist einfach zusammenhängend  $\iff n > 2$ . Im Fall  $n = 2$  lassen sich die verschiedenen "Arten" einfach geschlossener Kurven mit  $\mathbb{Z}$  "durchnummerieren".
- $B^3 \setminus ([-1, 1] \times \{0, 0\}) \subset \mathbb{R}^3$  ist nicht einfach zusammenhängend
- $S^2 \subset \mathbb{R}^3$  ist einfach zusammenhängend
- Der Torus im  $\mathbb{R}^3$  ist nicht einfach zusammenhängend. Die verschiedenen "Arten" einfach geschlossener Kurven lassen mit  $\mathbb{Z}^2$  "durchnummerieren".<sup>(i)</sup>

**Bemerkung 3.5.48.** Ist  $U \subset \mathbb{R}^n$  eine offene Menge und  $c : [a, b] \rightarrow \mathbb{R}^n$  eine einfach geschlossene stückweise stetig differenzierbare Kurve mit Teilkurven  $c_\alpha : I_\alpha = [a_\alpha, a_{\alpha+1}] \rightarrow U$  für  $\alpha = 0, 1, \dots, k$ ,  $a_0 := a$ ,  $a_{k+1} := b$ . Dann können wir die Homotopie  $H : [a, b] \times [0, 1] \rightarrow \mathbb{R}^n$  so wählen, dass das Bild von  $H$  eine allgemeine zweidimensionale Mannigfaltigkeit im  $\mathbb{R}^n$  im Sinne von Definition 3.4.21 ist, für die  $H|_{I_\alpha \times ]0, 1[}$  die Parametrisierungen der differenzierbaren Teilstücke liefert.

Aus Folgerung 3.5.39 zusammen mit Bemerkung 3.5.21.11.b folgt nun das Gewünschte:

**Satz 3.5.49** (Integrabilitätsbedingung für Gradientenfelder). *Es sei  $U \subset \mathbb{R}^n$  eine einfach zusammenhängende offene Menge und  $X : U \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld. Dann sind folgende Aussagen äquivalent:*

1.  $X$  ist ein Gradientenfeld
2.  $\text{Rot}X = 0$  (bzw.  $DX$  ist symmetrisch)

**Bemerkung 3.5.50.** Der Beweis von Satz 3.5.23 zeigt, dass wir das Potential eines Gradientenfeldes auf der Menge  $U$  erhalten, wenn wir für einen festen Punkt  $p^0 \in U$  zu  $p \in U$  einen beliebigen Weg  $c : [a, b] \rightarrow U$  wählen, der  $p^0$  und  $p$  verbindet. Wir haben dann

$$f(p) = \int_c f d\vec{s}.$$

---

<sup>(i)</sup>Mit Hilfe des Begriffs der Homotopie kann man eine Äquivalenzrelation zwischen Wegen definieren, sodass die hier anschaulich begründeten "Arten" von Kurven gerade die Äquivalenzklassen sind. Es stellt sich dann heraus, dass die Klasseneinteilung eine Gruppenstruktur trägt; in diesem Sinne ist das "Durchnummerieren" zu verstehen. Die erhaltene Gruppe heißt *Fundamentalgruppe*.

1. Ist  $U$  aus Satz 3.5.49 eine bezüglich  $p^0$  sternförmige Menge, so können wir als Weg die Verbindungsstrecke von  $p^0$  zu  $p$  wählen. Das gibt

$$f(p) = \sum_{k=1}^n \int_0^1 X^k(p^0 + t(p - p^0))(p_k^0 - p_k) dt.$$

2. Ist  $U$  aus Satz 3.5.49 ein achsenparalleler Quader, so können wir als Weg die Quaderkanten wählen. Das gibt

$$f(p) = \sum_{k=1}^n \int_{p_k^0}^{p_k} X^k(p_1, \dots, p_{k-1}, t, p_{k+1}^0, \dots, p_n^0) dt.$$

Als Folgerung aus Satz 3.5.49 erhalten wir auch einen Beweis für die Aussage von Satz 2.7.15 zu exakten Differentialgleichungen

**Folgerung 3.5.51.** Sind  $P$  und  $Q$  stetig differenzierbare Funktionen in zwei Variablen, so ist die DGL  $P(x, y) + Q(x, y)y' = 0$  genau dann exakt, wenn  $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ .

Oft lässt sich ein Potential einfacher mit Hilfe eindimensionaler Integrale bestimmen.

**Beispiel 3.5.52.** Das Vektorfeld  $X(x, y, z) = \begin{pmatrix} 2xy + yz + e^x \\ x^2 + 2yz + \sin(z) + xz \\ y^2 + y \cos(z) + xy \end{pmatrix}$  ist auf der einfach zusammenhängenden Menge  $\mathbb{R}^3$  definiert. Wegen

$$\operatorname{rot}(X) = \begin{pmatrix} (2y + \cos(z) + x) - (2y + \cos(z) + x) \\ (0) - (0) \\ (2x + x) - (2x + z) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

existiert somit ein Potential  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  mit  $\nabla f = X$ , d.h.

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xy + yz + e^x \\ \frac{\partial f}{\partial y} &= x^2 + 2yz + \sin(z) + xz \\ \frac{\partial f}{\partial z} &= y^2 + y \cos(z) + xy. \end{aligned}$$

Wir integrieren nacheinander diese drei Gleichungen nach  $x$ ,  $y$  und  $z$  und erhalten

$$\begin{aligned} f(x, y, z) &= \int \frac{\partial f}{\partial x} dx = \int (2xy + yz + e^x) dx \\ &= x^2y + xyz + e^x + c_1(y, z), \\ f(x, y, z) &= \int \frac{\partial f}{\partial y} dy = \int (x^2 + 2yz + \sin(z) + xz) dy \\ &= x^2y + zy^2 + y \sin(z) + xyz + c_2(x, z), \\ f(x, y, z) &= \int \frac{\partial f}{\partial z} dz = \int (y^2 + y \cos(z) + xy) dz \\ &= zy^2 + y \sin(z) + xyz + c_3(x, y). \end{aligned}$$

Hierbei ist zu beachten, dass die Integrationskonstanten von den jeweils anderen Variablen abhängen darf. Wählen wir in dem obigen Beispiel die Integrationskonstanten als

$$c_1(y, z) = zy^2 + y \sin(z) \quad c_2(x, z) = e^x, \quad c_3(x, y) = x^2y + e^x,$$

so erhalten wir als Potential

$$f(x, y, z) = x^2y + xyz + e^x + zy^2 + y \sin(z).$$

## 3.6 Grundzüge der Funktionentheorie

### 3.6.1 Komplex versus reell

Wir werden uns in diesem Abschnitt mit Funktionen beschäftigen, die ihren Definition- und ihren Wertebereich in den komplexen Zahlen haben. Wir bezeichnen dabei die komplexen Variablen in der Form  $z = x + iy$  und schreiben für eine Teilmenge  $U \subset \mathbb{C}$  verwenden dazu die Schreibweise

$$f : U \rightarrow \mathbb{C}, f : z \mapsto f(z) = f(x + iy).$$

Mit Hilfe der kanonischen Identifikation der komplexen Zahlen mit der reellen Zahlenebene über

$$\mathbb{C} \ni z = x + iy \longleftrightarrow (x, y) \in \mathbb{R}^2$$

und wegen der Eigenschaft

$$|z| = \sqrt{x^2 + iy^2} = \|(x, y)\|$$

des Betrags einer komplexen Zahl, stimmen die Topologien, die über  $|\cdot|$  auf  $\mathbb{C}$  und  $\|\cdot\|$  auf  $\mathbb{R}^2$  definiert werden, überein. Das liefert uns die Möglichkeit die topologischen Begriffe aus dem reellen Fall zu übertragen, Folgen und Reihen zu definieren und über Konvergenz zu sprechen. Die Ergebnisse aus den entsprechenden Kapiteln übertragen sich dabei sinngemäß, insbesondere behalten alle Aussagen bezüglich absoluter Konvergenz ihre Richtigkeit. Dabei müssen wir lediglich aufpassen, wenn in der Formulierung einiger Aussagen über  $\mathbb{R}$  explizit die Ordnungsrelationen Verwendung finden, da diese keine Entsprechung in  $\mathbb{C}$  hat. Ebenso erhalten wir eine Notation von Stetigkeit für komplexe Funktionen. Dabei können wir uns die folgende nützliche Beschreibung komplexer Funktionen zunutze machen.

Wegen der kanonischen Identifikation  $\mathbb{C} \supset U \longleftrightarrow \hat{U} \subset \mathbb{R}^2$  identifizieren wir eine Abbildung  $f : U \rightarrow \mathbb{C}$  mit einer Abbildung  $\hat{f} : \hat{U} \rightarrow \mathbb{C}$  durch

$$\hat{f}(x, y) := f(x + iy).$$

Für die Funktion  $f : U \rightarrow \mathbb{C}$  mit  $U \subset \mathbb{C}$  schreiben wir

$$f(z) = u(z) + iv(z)$$

mit

$$\begin{aligned} u(z) &:= \operatorname{Re}(f(z)) = \frac{1}{2}(f(z) + \overline{f(z)}), \\ v(z) &:= \operatorname{Im}(f(z)) = \frac{1}{2i}(f(z) - \overline{f(z)}), \end{aligned}$$

oder kurz

$$f = u + iv$$

mit

$$u = \operatorname{Re}(f) = \frac{1}{2}(f + \bar{f}), \quad v = \operatorname{Im}(f) = \frac{1}{2i}(f - \bar{f})$$

und

$$\bar{\bar{f}}(z) := \overline{f(z)}.$$

Wir identifizieren die Funktion  $f : U \rightarrow \mathbb{C}$  mit Hilfe des Real- und Imaginärteils mit einer Abbildung

$$f_R : \hat{U} \rightarrow \mathbb{R}^2$$

durch

$$f_R(x, y) = \begin{pmatrix} \hat{u}(x, y) \\ \hat{v}(x, y) \end{pmatrix}$$

In diesem Sinne können wir auch von reeller Differenzierbarkeit einer komplexen Funktion  $f$  sprechen, wenn wir uns auf die Darstellung  $f_R$  zurückziehen.

**Beispiel 3.6.1.** 1. Für  $f(z) = z^2$  bzw.  $f(x + iy) = (x^2 - y^2) + 2ixy$  gilt

$$\hat{u}(x, y) = u(x + iy) = x^2 - y^2, \quad \hat{v}(x, y) = v(x + iy) = 2xy$$

und

$$f_R(x, y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}.$$

2. Für  $f(z) = \bar{z}$  bzw.  $f(x + iy) = x - iy$  gilt

$$u(x + iy) = x, \quad v(x + iy) = -y$$

und

$$f_R(x, y) = \begin{pmatrix} x \\ -y \end{pmatrix}.$$

**Beispiel 3.6.2.** Einige mit Hilfe der Funktion  $f = u + iv : U \rightarrow \mathbb{C}$  definierten wichtige Abbildungen sind

$$\begin{aligned} if = -v + iu &\longleftrightarrow (if)_R = (-v, u)^T, \\ \bar{f} = u - iv &\longleftrightarrow (\bar{f})_R = (u, -v)^T, \\ i\bar{f} = v + iu &\longleftrightarrow (i\bar{f})_R = (v, u)^T, \end{aligned}$$

## 3.6.2 Komplexe Differentiation

**Bezeichnung 3.6.3.** Ist  $f : U \rightarrow \mathbb{C}$  derart, dass  $f_R : \hat{U} \rightarrow \mathbb{R}^2$  differenzierbar ist, und schreiben wir  $f(z) = u(z) + iv(z)$ , also  $f_R(x, y) = (\hat{u}(x, y), \hat{v}(x, y))^T$  so definieren wir in naheliegender Weise

$$\begin{aligned} \frac{\partial u}{\partial x}(x + iy) &:= \frac{\partial \hat{u}}{\partial x}(x, y), & \frac{\partial v}{\partial x}(x + iy) &:= \frac{\partial \hat{v}}{\partial x}(x, y), \\ \frac{\partial u}{\partial y}(x + iy) &:= \frac{\partial \hat{u}}{\partial y}(x, y), & \frac{\partial v}{\partial y}(x + iy) &:= \frac{\partial \hat{v}}{\partial y}(x, y) \end{aligned}$$

und damit

$$\begin{aligned} \frac{\partial f}{\partial x}(x + iy) &:= \frac{\partial u}{\partial x}(x, y) + i \frac{\partial v}{\partial x}(x, y), \\ \frac{\partial f}{\partial y}(x + iy) &:= \frac{\partial u}{\partial y}(x, y) + i \frac{\partial v}{\partial y}(x, y). \end{aligned}$$

Ab jetzt werden wir  $\hat{f}$ ,  $\hat{u}$  und  $\hat{v}$  mit  $f$ ,  $u$  und  $v$  identifizieren. Zwischen  $f$  und  $f_R$  werden wir jedoch unterscheiden, wenn wir den unterschiedlichen Charakter, komplex oder reell, betonen möchten.

Die reelle Differenzierbarkeit komplexer Funktionen macht, wie wir gesehen haben, keine Probleme. Im Gegensatz zu Abbildungen in zwei reellen Variablen, haben wir jedoch im Komplexen die Möglichkeit eine Definition aus einer reellen Dimension direkt zu übernehmen.

**Definition 3.6.4.** Es sei  $U \subset \mathbb{C}$  und  $f : U \rightarrow \mathbb{C}$ . Die Funktion  $f$  heißt in  $z_0 \in U$  *komplex differenzierbar*, wenn der Grenzwert

$$f'(z_0) := \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existiert. Die Zahl  $f'(z_0) \in \mathbb{C}$  heißt dann die *komplexe Ableitung von  $f$  in  $z_0$* .

**Beispiel 3.6.5.** Analog zu den Rechnungen in einer reellen Variablen erhalten wir komplexe Differenzierbarkeit für viele elementare komplexe Funktionen

1. Für  $c \in \mathbb{C}$  ist die konstante Funktion  $f(z) = \underline{c}$  an jeder Stelle  $z \in \mathbb{C}$  differenzierbar mit  $f'(z) = 0$ .

2.  $f(z) := z^k$  ist für  $k \in \mathbb{N}$  an jeder Stelle  $z \in \mathbb{C}$  differenzierbar und es ist  $f'(z) = kz^{k-1}$ .
3. Die Funktion  $f(z) = \frac{1}{z}$  ist an jeder Stelle  $z \in \mathbb{C} \setminus \{0\}$  differenzierbar mit  $f'(z) = -\frac{1}{z^2}$ .

Wie das nächste Beispiel zeigt, sind jedoch schon recht einfache, reell differenzierbare komplexe Funktionen nicht unbedingt komplex differenzierbar.

4. Die Funktion  $f(z) = \bar{z}$  ist an keiner Stelle  $z_0 \in \mathbb{C}$  differenzierbar, denn

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{(x - x_0) + i(y_0 - y)}{(x - x_0) + i(y - y_0)} = \begin{cases} -1 & \text{falls } x = x_0 \\ +1 & \text{falls } y = y_0 \end{cases}$$

Die Beweise der folgenden Aussagen übertragen sich wortwörtlich aus der reellen Situation.

**Satz 3.6.6.** *Die folgenden Aussagen sind für  $U \subset \mathbb{C}$ ,  $f : U \rightarrow \mathbb{C}$  und  $z_0 \in \mathbb{C}$  äquivalent:*

1.  $f$  ist in  $z_0$  komplex differenzierbar.
2. Es gibt eine in  $z_0$  stetige Funktion  $f^* : U \rightarrow \mathbb{C}$ , sodass

$$f(z) = f(z_0) + f^*(z_0)(z - z_0)$$

für alle  $z \in U$ .

3. Es gibt eine in  $z_0$  stetige Funktion  $r : U \rightarrow \mathbb{C}$  mit  $r(z_0) = 0$  und eine komplexe Zahl  $c$ , sodass

$$f(z) = f(z_0) + c(z - z_0) + r(z)(z - z_0)$$

für alle  $z \in U$ .

Dann gilt insbesondere  $f'(z_0) = f^*(z_0) = c$ .

**Satz 3.6.7.** 1. *Komplex differenzierbare Funktionen erfüllen die Summen-, Produkt-, und Quotientenregel:*

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

2. *Komplex differenzierbare Funktionen erfüllen die Kettenregel:*

$$(f \circ g)'(z) = f'(g(z)) g'(z).$$

3. *Ebenso gilt die Rechenregel für die Ableitung der Umkehrfunktion: Ist  $f : U \rightarrow f(U)$  bijektiv und ist  $f'(z) \neq 0$  für  $z \in U$  so gilt*

$$(f^{-1})'(f(z)) = \frac{1}{f'(z)}.$$

**Definition 3.6.8.** Eine Abbildung  $f : U \rightarrow \mathbb{C}$  auf der offenen Menge  $U \subset \mathbb{C}$  heißt *holomorph*, wenn  $f$  in jedem Punkt  $z \in U$  komplex differenzierbar ist.

Die Verbindung zwischen reeller und komplexer Differenzierbarkeit liefert nun der folgende Satz.

**Satz 3.6.9.** *Es sei  $U \subset \mathbb{C}$ . Die Abbildung  $f = u + iv : U \rightarrow \mathbb{C}$  ist genau dann komplex differenzierbar, wenn  $f_{\mathbb{R}} = (u, v)^T$  reell differenzierbar ist, und die Funktionalmatrix die Bedingung*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

*erfüllt. Diese Gleichungen heißen Cauchy-Riemann-Differentialgleichungen und lassen sich in Termen von  $f$  als*

$$\frac{\partial f}{\partial x} = -i \frac{\partial f}{\partial y}$$

*schreiben. Ist  $f$  komplex differenzierbar so gilt*

$$f'(z) = \frac{\partial f}{\partial x}(z).$$

**Beispiel 3.6.10.** 1. Wir wissen bereits, dass die Funktion  $f : \mathbb{C} \rightarrow \mathbb{C}$  mit  $f(z) = \bar{z}$  nicht holomorph ist. Es ist hier  $f(x + iy) = x - iy$ , also  $u(x, y) = x$  und  $v(x, y) = -y$ . Die Cauchy-Riemann-DGLen sind nicht erfüllt, denn es ist

$$\frac{\partial u}{\partial x} = 1 \neq -1 = \frac{\partial v}{\partial y}$$

2. Die Funktion  $f : \mathbb{C} \rightarrow \mathbb{C}$  mit  $f(z) = e^z$  ist holomorph. Es ist

$$f(x + iy) = e^{x+iy} = e^x e^{iy} = e^x (\cos(y) + i \sin(y))$$

also

$$u(x, y) = e^x \cos(y), \quad v(x, y) = e^x \sin(y).$$

Damit sind die Cauchy-Riemann-DGLen erfüllt, denn es gilt

$$\begin{aligned} \frac{\partial u}{\partial x}(x, y) &= e^x \cos(y) = \frac{\partial v}{\partial y}(x, y) \\ \frac{\partial u}{\partial y}(x, y) &= -e^x \sin(y) = -\frac{\partial v}{\partial x}(x, y) \end{aligned}$$

bzw. äquivalent dazu

$$\frac{\partial f}{\partial x} = e^x (\cos(y) + i \sin(y)) = -ie^x (-\sin(y) + i \cos(x)) = -i \frac{\partial f}{\partial y}.$$

Insbesondere ist also

$$f'(z) = \frac{\partial f}{\partial x} = e^z$$

wie zu erwarten war.

### 3.6.3 Komplexe Stammfunktion und komplexes Wegintegral

**Definition 3.6.11.** Es sei  $U \subset \mathbb{C}$  offen und  $f : U \rightarrow \mathbb{C}$ . Eine holomorphe Funktion  $g : U \rightarrow \mathbb{C}$  heißt *komplexe Stammfunktion* von  $f$ , wenn  $g'(z) = f(z)$  für alle  $z \in U$ . Besitzt  $f$  eine komplexe Stammfunktion, so heißt  $f$  *komplex integrierbar*.

Komplexe Integrierbarkeit hängt eng mit der Eigenschaft Gradientenfeld zu sein zusammen. Genauer formuliert ist dieser Zusammenhang in der folgenden Aussage.

**Satz 3.6.12.** *Für eine offene Menge  $U \subset \mathbb{C}$  und eine Funktion  $f : U \rightarrow \mathbb{C}$  sind die folgenden zwei Aussagen äquivalent:*

1.  $f$  ist komplex integrierbar mit Stammfunktion  $g$ .
2.  $(\bar{f})_R$  und  $(i\bar{f})_R$  sind Gradientenfelder mit Potentialen  $\operatorname{Re}(g)$  und  $\operatorname{Im}(g)$ .

Zusammen mit Satz 3.5.49 erhalten wir als Folgerung die folgende Formulierung des Cauchy-Integralsatzes.

**Satz 3.6.13** (Cauchy-Integralsatz). *Jede holomorphe Funktion auf einer einfach zusammenhängenden, offenen Menge ist komplex integrierbar.*

Ist die Definitionsmenge nicht einfach zusammenhängend, dann hat man zumindest noch eine lokale Variante.

**Folgerung 3.6.14.** Jede holomorphe Funktion ist lokal komplex integrierbar.

**Definition 3.6.15.** Es sei  $f : U \rightarrow \mathbb{C}$  und  $c : [a, b] \rightarrow U$  für  $U \subset \mathbb{C}$ . Das *komplexe Wegintegral von  $f$  längs  $c$*  ist definiert durch

$$\int_c f(z) dz = \int_a^b (f \circ c)(t) c'(t) dt.$$

**Bemerkung 3.6.16.** 1. Ist  $f : U \rightarrow \mathbb{C}$  komplex differenzierbar mit holomorpher Stammfunktion  $g : U \rightarrow \mathbb{C}$ , dann gilt für eine Kurve  $c : [a, b] \rightarrow U$

$$(g \circ c)'(t) = (g' \circ c)(t) c'(t) = (f \circ c)(t) c'(t),$$

so dass mit  $c(a) = z_a$  und  $c(b) = z_e$

$$\int_c f(z) dz = \int_a^b (g \circ c)'(t) dt = g(z_e) - g(z_a)$$

nur vom Anfangs- und Endpunkt der Kurve  $c$  abhängt.

2. Ist  $f : U \rightarrow \mathbb{C}$  komplex integrierbar, so verschwindet das Wegintegral über jeden geschlossenen Weg in  $U$ .

**Beispiel 3.6.17.** Es sei  $c : [0, 2\pi] \rightarrow \mathbb{C}$  mit  $c(t) = z_0 + re^{it}$  eine einfache Beschreibung der Kreislinie  $S_r(z_0) \subset \mathbb{C}$ . Wir berechnen jeweils das komplexe Wegintegral über diese einfach geschlossene Kurve:

1.  $f : \mathbb{C} \rightarrow \mathbb{C}$  mit  $f(z) = \bar{z}$

$$\begin{aligned} \int_c \bar{z} dz &= \int_0^{2\pi} (\bar{z}_0 + re^{-it}) ire^{it} dt \\ &= ir\bar{z}_0 \int_0^{2\pi} e^{it} dt + ir^2 \int_0^{2\pi} dt \\ &= 2\pi ir^2 + ir\bar{z}_0 \int_0^{2\pi} \cos(t) dt - r\bar{z}_0 \int_0^{2\pi} \sin(t) dt \\ &= 2\pi ir^2 \end{aligned}$$

Insbesondere ist also  $f(z) = \bar{z}$  auf  $\mathbb{C}$  nicht komplex integrierbar.

2.  $f : \mathbb{C} \setminus \{z_0\} \rightarrow \mathbb{C}$  mit  $f(z) = \frac{1}{(z-z_0)^2}$

$$\int_c \frac{dz}{(z-z_0)^2} = \int_0^{2\pi} \frac{ire^{it}}{r^2 e^{2it}} dt = \frac{i}{r} \int_0^{2\pi} e^{-it} dt = 0$$

Das Ergebnis war vorauszusehen, da  $f$  komplex integrierbar ist mit Stammfunktion  $g(z) = \frac{1}{z_0-z}$ .

3.  $f : \mathbb{C} \setminus \{z_0\} \rightarrow \mathbb{C}$  mit  $f(z) = \frac{1}{z-z_0}$

$$\int_c \frac{dz}{z-z_0} = \int_0^{2\pi} \frac{ire^{it}}{re^{it}} dt = i \int_0^{2\pi} dt = 2\pi i$$

Damit ist die Funktion  $f(z) = \frac{1}{z-z_0}$  auf  $\mathbb{C} \setminus \{z_0\}$  nicht komplex integrierbar.

**Bemerkung 3.6.18.** Das komplexe Wegintegral hängt mit dem reellen Wegintegral wie folgt zusammen:

$$\int_c f(z) dz = \int_c (\bar{f})_R d\vec{s} + i \int_c (if)_R d\vec{s}.$$

Nutzt man dies, so folgt die Abschätzung

$$\left| \int_c f(z) dz \right| \leq L(c) \cdot \max_{a \leq t \leq b} |f \circ c|.$$

## 3.6.4 Der Cauchy-Integralsatz und die Cauchy-Integralformel

Wir fassen das bisherige noch einmal zusammen.

**Satz 3.6.19** (Cauchy-Integralsatz – 1. Fassung). *Ist  $U \overset{\circ}{\subset} \mathbb{C}$  einfach zusammenhängend und  $f : U \rightarrow \mathbb{C}$  holomorph, so gilt für jeden geschlossenen Weg  $c$  in  $U$*

$$\oint_c f(z) dz = 0.$$

Aus diesem Satz folgt über einen Grenzwertprozess eine etwas weniger strenge Variante.

**Folgerung 3.6.20.** *Es sei  $U \overset{\circ}{\subset} \mathbb{C}$  einfach zusammenhängend und  $f : U \rightarrow \mathbb{C}$  holomorph auf  $U \setminus \{z^*\}$  aber stetig in  $z^*$ , dann gilt für jeden einfach geschlossenen Weg  $c$  in  $U$*

$$\oint_c f(z) dz = 0.$$

Integration und Differentiation lassen sich unter schwachen Voraussetzungen vertauschen.

**Satz 3.6.21.** *Es seien  $U \overset{\circ}{\subset} \mathbb{C}$ ,  $W \subset \mathbb{C}$  und  $c : [a, b] \rightarrow W$ . Weiter sei  $f : U \times W \rightarrow \mathbb{C}$  stetig und die Funktionen  $f_w : U \rightarrow \mathbb{C}$  mit  $f_w(z) := f(z, w)$  seien für alle  $w \in W$  holomorph. Weiter sei  $\frac{\partial f}{\partial z} : U \times W \rightarrow \mathbb{C}$  mit  $\frac{\partial f}{\partial z}(z, w) := f'_w(z)$  stetig. Dann ist auch  $\hat{f} : U \rightarrow \mathbb{C}$  mit*

$$\hat{f}(z) := \int_c f(z, w) dw$$

*holomorph mit*

$$\hat{f}'(z) = \int_c \frac{\partial f}{\partial z}(z, w) dw.$$

**Beispiel 3.6.22.** *Es sei  $c(t) := z_0 + re^{it}$  mit  $t \in [0, 2\pi]$  die Beschreibung der Kreislinie  $S_r(z_0)$ . Wir wollen zeigen, dass die Abbildung  $\hat{f} : \mathbb{C} \setminus S_r(z_0) \rightarrow \mathbb{C}$  mit*

$$\hat{f}(z) = \int_c \frac{1}{w - z} dw$$

lokal konstant ist. Dazu nutzen wir die obige Verträglichkeit zwischen Ableitung und Integral und zeigen  $\hat{f}' = 0$ . Es ist

$$\hat{f}'(z) = \int_c \frac{\partial}{\partial z} \left( \frac{1}{w-z} \right) dw = \int_c \frac{1}{(w-z)^2} dw = 0,$$

da  $\mathbb{C} \setminus \{z\} \ni w \mapsto \frac{1}{(w-z)^2}$  auf  $\mathbb{C} \setminus \{z\}$  eine komplexe Stammfunktion besitzt.

- Wir wissen aber bereits, dass für  $z_0 \in K_r(z_0)$  die Funktion den Wert  $\hat{f}(z_0) = 2\pi i$  annimmt. Damit ist insgesamt

$$\hat{f}|_{K_r(z_0)} = 2\pi i,$$

da  $K_r(z_0)$  zusammenhängend ist.

- Ist  $z \in \mathbb{C} \setminus \overline{K}_r(z_0)$ , so beschreibe  $\sigma(t) := z_0 + t(z - z_0)$  die Verbindungsstrecke zwischen  $z_0$  und  $z$  und es sei  $t_0 := \frac{R+|z-z_0|}{2|z-z_0|}$ . Dann liegt  $\sigma(t_0)$  außerhalb von  $\overline{K}_r(z_0)$  und wir betrachten die Gerade durch  $s(t_0)$ , die senkrecht auf dem Bild von  $\sigma$  steht. Diese zerteilt  $\mathbb{C}$  in zwei Halbebenen, von denen die eine,  $H_-$ , den Punkt  $z$  enthält und die andere,  $H_+$ , die Kreisscheibe  $\overline{K}_r(z_0)$ . Insbesondere ist  $w \mapsto \frac{1}{z-w}$  holomorph auf der einfach zusammenhängenden Menge  $H_+$ . Damit ist  $\int_c \frac{dw}{w-z} = 0$  und, da  $z$  beliebig gewählt war,

$$\hat{f}|_{\mathbb{C} \setminus \overline{K}_r(z_0)} = 0.$$

Es sei  $f : U \rightarrow \mathbb{C}$  eine holomorphe Funktion auf  $U \stackrel{\circ}{\subset} M$ . Dann ist für ein festes  $z \in U$  die Funktion  $f^* : U \rightarrow \mathbb{C}$  mit

$$f^*(w) := \begin{cases} \frac{f(w) - f(z)}{w - z} & w \in U \setminus \{z\} \\ f'(z) & w = z \end{cases}$$

auf  $U \setminus \{z\}$  holomorph und in  $z$  stetig. Für eine Kreisscheibe  $K_r(z_0) \subset U$  verschwindet das Integral  $\int_c f^*(w) dw$  für  $c : [0, 2\pi] \rightarrow \mathbb{C}$  mit  $c(t) = z_0 + re^{it}$  wegen Folgerung 3.6.20. Damit ist

$$f(z) \int_c \frac{dw}{w-z} = \int_c \frac{f(w)}{w-z} dw$$

für alle  $z \notin S_r(z_0)$ .

Diese Überlegung liefert uns die Cauchy-Integralformel.

**Satz 3.6.23** (Cauchy-Integralformel). *Es sei  $U \overset{\circ}{\subset} \mathbb{C}$  und  $f : U \rightarrow \mathbb{C}$  holomorph. Weiter sei  $\overline{K}_r(z_0) \subset U$ . Dann gilt für alle  $z \in K_r(z_0)$*

$$f(z) = \frac{1}{2\pi i} \int_c \frac{f(w)}{w - z} dw$$

wobei  $c(t) = z_0 + re^{it}$  mit  $t \in [0, 2\pi]$ .

Damit ergeben sich jetzt eine Reihe von Folgerungen

**Satz 3.6.24.** *Ist  $f : U \rightarrow \mathbb{C}$  holomorph, so ist  $f$  beliebig oft komplex differenzierbar mit*

$$f^k(z) = \frac{k!}{2\pi i} \int_c \frac{f(w)}{(w - z)^{k+1}} dw.$$

**Satz 3.6.25.** *Es sei  $f : U \rightarrow \mathbb{C}$  holomorph. Dann ist  $f$  auf jeder Zusammenhangskomponente von  $U$  konstant, wenn nur eine der drei folgenden Aussagen erfüllt ist:*

1.  $\bar{f} : U \rightarrow \mathbb{C}$  ist holomorph.
2.  $|f| : U \rightarrow \mathbb{R}$  ist konstant.
3.  $|f| : U \rightarrow \mathbb{R}$  nimmt sein Maximum an.

**Satz 3.6.26** (Satz von Cauchy-Morera). *Für eine Funktion  $f : U \rightarrow \mathbb{C}$  auf der offenen Menge  $U \subset \mathbb{C}$  sind folgende Aussagen äquivalent:*

1.  $f$  ist holomorph.
2.  $f$  ist lokal komplex integrierbar.

**Satz 3.6.27.** *Ist  $f : U \rightarrow \mathbb{C}$  holomorph auf  $U \setminus \{z_0\}$  und ist  $f$  in  $z_0$  beschränkt, dann gibt es eine holomorphe Funktion  $\tilde{f} : U \rightarrow \mathbb{C}$  mit  $\tilde{f}|_{U \setminus \{z_0\}} = f$ .*

**Satz 3.6.28** (Satz von Liouville). *Es gelten die beiden äquivalenten Aussagen*

- Jede holomorphe, nicht-konstante Funktion  $f : \mathbb{C} \rightarrow \mathbb{C}$  ist unbeschränkt.
- Ist die holomorphe Funktion  $f : \mathbb{C} \rightarrow \mathbb{C}$  beschränkt, so ist sie konstant.

### 3.7 Anhang Analysis III: Komplexe Zahlen

#### 3.7.1 Beschreibung der komplexen Zahlen

Jedes Polynom lässt sich über  $\mathbb{R}$  in Linearfaktoren und quadratische Faktoren zerlegen, siehe Satz 2.2.13. Das Standardbeispiel eines quadratischen Faktors ohne Nullstelle ist  $p(x) = x^2 + 1$ . Wir wollen nun den Körper beschreiben, der die Zerlegung eines Polynoms in ausschließlich Linearfaktoren erlaubt.

**Definition/Bemerkung 3.7.1** (Rechenoperationen auf  $\mathbb{R}^2$ ). Auf  $\mathbb{R}^2$  führen wir eine Addition und eine Multiplikation auf die folgende Art ein:

1.  $(a, b) + (c, d) := (a + c, b + d)$
2.  $(a, b)(c, d) := (ac - bd, ad + bc)$

Die Addition ist die, die wir schon in der Vektorrechnung kennengelernt haben. Nur die Multiplikation ist wirklich neu.

**Definition 3.7.2** (Komplexe Zahlen). Die Elemente der Ebene zusammen mit der in Definition 3.7.1 definierten Addition und Multiplikation nennt man die Menge der *komplexen Zahlen* und wir bezeichnen diese mit  $\mathbb{C}$ . Ein Element  $z = (a, b)$  heißt *komplexe Zahl*.  $\mathbb{C}$  nennt man auch die *Gaußsche Zahlenebene*.

Das Rechnen mit komplexen Zahlen erfüllt alle gängigen Rechenregeln:

**Satz 3.7.3** (Rechenregeln für komplexe Zahlen). Sind  $z_1 = (a_1, b_1)$ ,  $z_2 = (a_2, b_2)$  und  $z_3 = (a_3, b_3)$  komplexe Zahlen, so gilt:

1.  $z_1 + z_2 = z_2 + z_1$
2.  $z_1 z_2 = z_2 z_1$
3.  $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$
4.  $(z_1 z_2) z_3 = z_1 (z_2 z_3)$
5.  $z_1 (z_2 + z_3) = z_1 z_2 + z_1 z_3$
6. Die komplexe Zahl  $(0, 0)$  erfüllt  $(0, 0) + z = z$  für jedes  $z \in \mathbb{C}$

7. Ist  $z = (a, b) \in \mathbb{C}$  eine komplexe Zahl, so erfüllt die komplexe Zahl  $-z := (-a, -b)$  die Gleichung  $z + (-z) = (0, 0)$
8. Die komplexe Zahl  $(1, 0)$  erfüllt  $(1, 0)z = z$  für jedes  $z \in \mathbb{C}$
9. Ist  $z = (a, b) \in \mathbb{C}$  eine komplexe Zahl mit  $(a, b) \neq (0, 0)$ , so erfüllt die komplexe Zahl  $z^{-1} := \left( \frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right)$  die Gleichung  $zz^{-1} = (1, 0)$  (statt  $z^{-1}$  schreiben wir auch  $\frac{1}{z}$ )
10. Es gilt  $(0, 0)z = (0, 0)$  für jedes  $z \in \mathbb{C}$

**Folgerung 3.7.4.** Wenn wir uns auf die komplexen Zahlen beschränken, deren zweite Komponente verschwindet, also komplexen Zahlen der Form  $(a, 0)$ , so sehen wir:

1.  $(a, 0) + (b, 0) = (a + b, 0)$
2.  $(a, 0)(b, 0) = (ab, 0)$
3.  $(a, 0)^{-1} = \left(\frac{1}{a}, 0\right)$

Das heißt, die Rechenvorschriften nehmen in diesem Fall keine Notiz von dem zweiten Eintrag.

4. Für alle  $(c, d) \in \mathbb{C}$  ist  $(a, 0)(c, d) = (ac, ad)$

Folgerung 3.7.4 liefert:

- Bemerkung 3.7.5.**
- [zu 1.-3.] Wir können mit den komplexen Zahlen der Form  $(a, 0)$  wie mit den reellen Zahlen rechnen. Die komplexe Zahl  $(a, 0)$  entspricht der reellen Zahl  $a$ .
  - [zu 4.] Dieser Punkt begründet diese Interpretation weiter, denn die Multiplikation beliebiger komplexer Zahlen mit solchen der Form  $(a, 0)$  entspricht der skalaren Multiplikation der Vektorrechnung.

Wegen Satz 3.7.3 und der anschließenden Bemerkungen liegt die folgende Definition und Vereinbarung nahe:

**Definition 3.7.6** (Real- und Imaginärteil). • Wir identifizieren die reelle Zahl  $a$  und die komplexe Zahl  $(a, 0) \in \mathbb{C}$ . So wird  $\mathbb{R}$  eine Teilmenge von  $\mathbb{C}$ .

- Für  $z = (a, b) \in \mathbb{C}$  heißt  $\operatorname{Re}(z) := a$  der *Realteil* und  $\operatorname{Im}(z) := b$  der *Imaginärteil* von  $z$ .
- Die Gerade  $\{(x, 0) \mid x \in \mathbb{R}\} \subset \mathbb{C}$  heißt die *reelle Achse* und wir schreiben  $\mathbb{R}$ . Die Gerade  $\{(0, y) \mid y \in \mathbb{R}\} \subset \mathbb{C}$  heißt die *imaginäre Achse* und wir schreiben  $i\mathbb{R}$ .

Die Identifizierung aus Definition 3.7.6 liefert

**Folgerung 3.7.7.** (Zerlegung komplexer Zahlen) Jede komplexe Zahl  $z = (a, b)$  besitzt die Zerlegung

$$(a, b) = (a, 0)(1, 0) + (b, 0)(0, 1) = a + b(0, 1).$$

Diese Folgerung motiviert nun die nächste Definition

**Definition 3.7.8.** Die komplexe Zahl  $i := (0, 1)$  heißt *imaginäre Einheit*. Sie erlaubt es, jede komplexe Zahl  $(a, b) \in \mathbb{C}$  in der Form  $z = a + ib$  darzustellen.

**Satz 3.7.9.** Die imaginäre Einheit erfüllt

$$i^2 = -1.$$

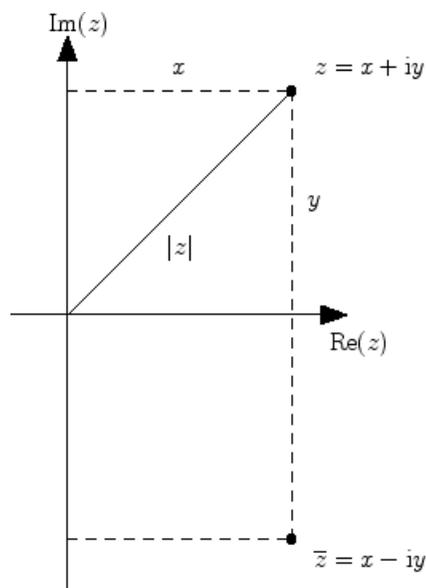
In der alten Schreibweise ist das  $(0, 1)(0, 1) = (-1, 0)$ . Das heißt:  $i$  und damit natürlich auch  $-i$  sind Lösungen der (über  $\mathbb{R}$  nicht lösbaren) Gleichung  $z^2 = -1$ .

**Definition/Bemerkung 3.7.10.** 1. Für  $z = a + bi \in \mathbb{C}$  bezeichnet  $\bar{z} := a - ib$  die zu  $z$  komplex konjugierte Zahl.

2. Für  $z = a + bi \in \mathbb{C}$  bezeichnet  $|z| := \sqrt{a^2 + b^2} \in \mathbb{R}$  den Betrag der komplexen Zahl  $z$ .

zu 1. Geometrisch entspricht die komplexe Konjugation der Spiegelung an der reellen Achse.

Zu 2. Der Betrag der komplexen Zahl entspricht der Norm des entsprechenden Vektors im  $\mathbb{R}^2$ , bzw. dem Abstand des Punktes im  $\mathbb{R}^2$  vom Ursprung.



**Satz 3.7.11.** 1.  $\overline{\bar{z}} = z$ .

2.  $\overline{z + w} = \bar{z} + \bar{w}$  und  $\overline{z - w} = \bar{z} - \bar{w}$ .

3.  $\overline{zw} = \bar{z} \bar{w}$  und  $\overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}$  (falls  $w \neq 0$ ).

4.  $\bar{z} = z$  genau dann, wenn  $z \in \mathbb{R}$  und  $\bar{z} = -z$  genau dann, wenn  $z \in i\mathbb{R}$ .

5.  $\operatorname{Re}(z) = \frac{z + \bar{z}}{2}$  und  $\operatorname{Im}(z) = \frac{z - \bar{z}}{2i}$ .

6. Insbesondere ist  $\bar{i} = -i = \frac{1}{i}$ .

Wegen der Verwandtschaft zwischen Betrag und Norm in der Bemerkung zu 2. aus Satz 3.7.11 haben wir

**Satz 3.7.12.** 1. Es ist  $|z| \geq 0$  und  $|z| = 0$  genau dann, wenn  $z = 0$ .

2.  $|zw| = |z||w|$  und  $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$  (falls  $w \neq 0$ ).

3. Für  $a \in \mathbb{R} \subset \mathbb{C}$  ist  $|a|$  der gewöhnliche Betrag und es gilt  $|\bar{z}| = |z|$ .

4.  $|\operatorname{Re}(z)| \leq |z|$  und  $|\operatorname{Im}(z)| \leq |z|$ .

5. Es gelten die Dreiecksungleichungen, das heißt

$$|z + w| \leq |z| + |w| \quad \text{und} \quad ||z| - |w|| \leq |z - w|.$$

Zwei wichtige Eigenschaften sind noch

6.  $z \bar{z} = |z|^2$ .

7.  $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$  (falls  $z \neq 0$ ).

### 3.7.2 Polarkoordinaten, Wurzeln und komplexe Exponentialfunktion

Sei  $z \in \mathbb{C} \setminus \{0\}$  eine komplexe Zahl, die wir als Punkt in der Zahlenebene betrachten. Dann lässt sich die Lage der komplexen Zahl durch die Polarkoordinaten beschreiben, also durch den Abstand zum Ursprung und den Winkel zur positiven reellen Achse.

**Definition/Bemerkung 3.7.13.** Es gibt eine reelle Zahl  $r \geq 0$  und einen Winkel  $\phi \in [0, 2\pi[$ , sodass  $z \in \mathbb{C}$  die folgende Darstellung hat

$$z = r(\cos \phi + i \sin \phi)$$

Diese Darstellung ist für  $z = 0$  mehrdeutig, und wir beschreiben  $z = 0$  durch  $r = 0$  und einen beliebigen Winkel.

Die Darstellung der komplexen Zahl heißt *Polarkoordinatendarstellung* der Zahl  $z$ . Der Winkel  $\phi \in [0, 2\pi[$  heißt ihr *Argument* und wird mit  $\arg(z)$  bezeichnet.

**Bemerkung 3.7.14.** Lässt man die Bedingung fallen, dass der Winkel, der sich aus der Polarkoordinatendarstellung ergibt, aus  $[0, 2\pi[$  stammt, so spricht man lediglich von dem (besser: einem) *Polarwinkel* von  $z$ .

Das Argument einer komplexen Zahl entspricht somit einer "Normierung" dieses Polarwinkels.

**Satz 3.7.15** (Umrechnung Polarkoordinaten  $\leftrightarrow$  kartesische Koordinaten).

1. Ist  $z = r(\cos \phi + i \sin \phi)$ , so ist  $z = a + ib$  mit

$$a = r \cos \phi \quad \text{und} \quad b = r \sin \phi.$$

2. Ist  $z = a + ib$ , so ist  $r = |z| = \sqrt{a^2 + b^2}$  und für  $z \neq 0$  ist

$$\phi = \arg(z) = \begin{cases} \arctan\left(\frac{b}{a}\right) & \text{falls } a > 0, b \geq 0, \\ \frac{\pi}{2} & \text{falls } a = 0, b > 0, \\ \arctan\left(\frac{b}{a}\right) + \pi & \text{falls } a < 0, \\ \frac{3\pi}{2} & \text{falls } a = 0, b < 0, \\ \arctan\left(\frac{b}{a}\right) + 2\pi & \text{falls } a > 0, b < 0, \end{cases}$$

**Bemerkung 3.7.16.** Wie oben erwähnt, entspricht die Beschränkung von  $\arg(z)$  auf  $[0, 2\pi[$  einer Normierung des Polarwinkels. In der Literatur finden sich auch andere Normierungen, etwa  $]-\pi, \pi]$ .

**Satz 3.7.17.**

1. Für alle reellen Zahlen  $a$  gilt:  $a > 0$  bzw.  $a < 0$  genau dann, wenn  $\arg(a) = 0$  bzw.  $\arg(a) = \pi$ .
2. Für  $a \in \mathbb{R}^{>0}$  und  $z \in \mathbb{C}$  gilt  $\arg(az) = \arg(z)$ .
3. Für die komplex konjugierte Zahl zu  $z \neq 0$  gilt<sup>(i)</sup>  $|\bar{z}| = |z|$  und  $\arg(\bar{z}) = 2\pi - \arg(z)$ .
4. Wegen  $\frac{1}{z} = \frac{1}{|z|^2} \bar{z}$  gilt für das Inverse  $\left|\frac{1}{z}\right| = \frac{1}{|z|}$  und  $\arg\left(\frac{1}{z}\right) = 2\pi - \arg(z)$ .
5. Für alle  $z \neq 0$  gilt  $\arg(-z) = \arg(z) + \pi$ .
6. Für das Produkt zweier komplexer Zahlen gilt  $|wz| = |w||z|$  und

$$\arg(wz) = \begin{cases} \arg(w) + \arg(z) & \text{falls Summe} < 2\pi \\ \arg(w) + \arg(z) - 2\pi & \text{falls Summe} \geq 2\pi \end{cases}$$

7. Für die Potenzen einer komplexen Zahl gilt

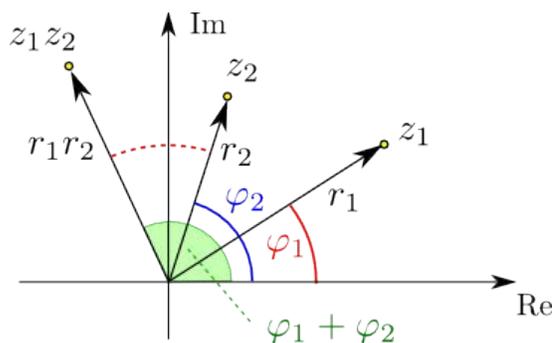
$$|z^n| = |z|^n \quad \text{und} \quad \arg(z^n) = n \arg(z).$$

<sup>(i)</sup>Hier und in den folgenden Formeln muss man um im Ergebnis tatsächlich das Argument zu erhalten ein Vielfaches von  $2\pi$  addieren.

**Bemerkung 3.7.18.** • *Geometrische Interpretation der Summe.* Die Summe zweier komplexer Zahlen entspricht der vektoriellen Addition der einzelnen Summanden, da die Summe – wie die Vektoraddition – komponentenweise definiert ist.

- *Geometrische Interpretation des Produkts.* Wegen Punkt 5 des letzten Satzes gilt: Der Polarwinkel des Produktes zweier komplexer Zahlen ist die Summe der Polarwinkel der einzelnen Faktoren und die Länge des Produktes ist das Produkt der einzelnen Längen. Man spricht auch von einer Drehstreckung.

Abbildung 3.7.1: Geometrische Interpretation der komplexen Multiplikation



**Satz 3.7.19.** Jede komplexe Zahl  $w \neq 0$  hat  $n$   $n$ -te Wurzeln. Mit anderen Worten: Die Gleichung  $z^n = w$  hat genau  $n$  verschiedene Lösungen  $w_1, w_2, \dots, w_n$ , die Wurzeln von  $w$ .

Ist  $\phi = \arg(w)$ , so sind diese Wurzeln

$$w_k = \sqrt[n]{|w|} \left( \cos \left( \frac{\phi + 2\pi k}{n} \right) + i \sin \left( \frac{\phi + 2\pi k}{n} \right) \right),$$

für  $k = 1, \dots, n$ .

**Definition 3.7.20.** Für  $z = a + ib \in \mathbb{C}$  definieren wir

$$e^z := e^a (\cos b + i \sin b).$$

**Bemerkung 3.7.21.** • Für  $a \in \mathbb{R} \subset \mathbb{C}$  erhalten wir die reelle Exponentialabbildung zurück.

- Für  $ib \in i\mathbb{R} \subset \mathbb{C}$  gilt  $\exp(ib) = \cos b + i \sin b$ .

**Satz 3.7.22** (Formel von de Moivre). Für alle  $x \in \mathbb{R}$  und  $n \in \mathbb{N}$  gilt

$$(\cos x + i \sin x)^n = \cos(nx) + i \sin(nx).$$

**Bemerkung 3.7.23.** Für alle  $z \in \mathbb{C}$  und  $n \in \mathbb{N}$  ist

$$e^{nz} = (e^z)^n.$$

Neben der letzten Eigenschaft erfüllt die komplexe Exponentialfunktion auch die anderen Eigenschaften ihrer reellen Schwester. Damit bekommt die Polarkoordinatendarstellung einer komplexen Zahl die kompakte (und übliche) Form

$$z = re^{i\phi}.$$

**Satz 3.7.24.** Sind  $p, q \in \mathbb{C}$ , so ist die Gleichung  $z^2 + pz + q = 0$  in  $\mathbb{C}$  lösbar.

1. Im Allgemeinen berechnet man die Lösungen etwa mit Hilfe der pq-Formel, nur, dass man hier die komplexe Wurzel betrachten muss.
2. Ist die quadratische Gleichung reell, also  $p, q \in \mathbb{R}$ , und schreiben wir  $\Delta := \frac{p^2}{4} - q$  für die Diskriminante, so besitzt sie
  - die eindeutige reelle Lösung  $z = -\frac{p}{2}$ , falls  $D = 0$
  - die zwei reellen Lösungen  $z_1 = -\frac{p}{2} + \sqrt{\Delta}$  und  $z_2 = -\frac{p}{2} - \sqrt{\Delta}$ , falls  $\Delta > 0$ .
  - die zwei komplex konjugierten Lösungen  $z_1 = -\frac{p}{2} + i\sqrt{-\Delta}$  und  $\bar{z}_1 = -\frac{p}{2} - i\sqrt{-\Delta}$ , falls  $\Delta < 0$ .

Damit haben wir

**Satz 3.7.25** (Fundamentalsatz der Algebra). • Jedes nicht-konstante Polynom  $p$  hat über  $\mathbb{C}$  mindestens eine Nullstelle.

- Jedes nicht-konstante Polynom lässt sich über  $\mathbb{C}$  vollständig in Linearfaktoren zerlegen.

(Hierbei sind als Koeffizienten ausdrücklich auch komplexe Zahlen zugelassen)

Das liefert die folgende Erweiterung der Faktorisierung reeller Polynome aus Satz 2.2.13

**Satz 3.7.26.** *Es sei  $p$  ein Polynom  $n$ -ten Grades mit Leitkoeffizient  $a_n = 1$ .  $p$  habe die reellen Nullstellen  $\lambda_1, \dots, \lambda_m$  mit Vielfachheiten  $r_1, \dots, r_m$  und die Paare konjugiert-komplexer Nullstellen  $\mu_1 \pm i\omega_1, \dots, \mu_p \pm i\omega_p$  mit Vielfachheiten  $s_1, \dots, s_\ell$ . Dann gilt*

$$p(x) = (x - \lambda_1)^{r_1} \cdot \dots \cdot (x - \lambda_m)^{r_m} (x - \mu_1 - i\omega_1)^{s_1} (x - \mu_1 + i\omega_1)^{s_1} \cdot \dots \cdot (x - \mu_p - i\omega_p)^{s_p} (x - \mu_p + i\omega_p)^{s_p}.$$

*Die quadratischen Faktoren von  $p$  ohne reelle Nullstellen sind*

$$(x - \mu_j - i\omega_j)(x - \mu_j + i\omega_j) = x^2 - 2\mu_j x + (\mu_j^2 + \omega_j^2).$$

---

## Abbildungsverzeichnis

---

|       |   |    |
|-------|---|----|
| 1.1.1 | Venn-Diagramme . . . . .  | 5  |
| 1.2.1 | Die Zahlenfolge $(\frac{1}{n})$ . . . . .   | 19 |
| 1.2.2 | Konvergenz von Zahlenfolgen . . . . .   | 21 |
| 1.3.1 | Einige Funktionen und ihre Graphen . . . . .  | 41 |
| 1.3.2 | Wurzelfunktionen . . . . .  | 44 |
| 1.3.3 | Die allgemeine Potenzfunktion $x \mapsto x^r$ . . . . .   | 45 |
| 1.3.4 | Polynome . . . . .  | 47 |
| 1.3.5 | Rationale Funktionen . . . . .  | 49 |
| 1.3.6 | Die Graphen der Exponentialfunktion $x \mapsto a^x$ (rot) und der<br>Logarithmusfunktion $x \mapsto \log_a(x)$ (blau) . . . . . | 53 |
| 1.3.7 | Allgemeine Potenzfunktionen . . . . .   | 54 |
| 1.4.1 | Einige Funktionsgraphen . . . . .   | 55 |
| 1.4.2 | Zwischenwertsatz . . . . .  | 59 |
| 1.5.1 | Differenzierbarkeit . . . . .   | 68 |

---

|       |  |     |
|-------|--|-----|
| 1.5.2 | Tangente . . . . .   | 69  |
| 1.5.3 | Typen von "Glattheit" . . . . .  | 85  |
| 1.7.1 | Der Einheitskreis . . . . .  | 97  |
| 1.7.2 | Sinus und Cosinus . . . . .  | 97  |
| 1.7.3 | Additionstheoreme . . . . .  | 98  |
| 1.7.4 | Tangens, Cotangens . . . . .   | 102 |
| 1.7.5 | Sekans und Kosekans . . . . .  | 103 |
| 1.7.6 | Winkelfunktionen am Kreis . . . . .  | 103 |
| 1.7.7 | Die Arkusfunktionen . . . . .  | 106 |
| 1.9.1 | Das Cantorsche Diagonalverfahren . . . . .   | 115 |
| 2.1.1 | $f_n(x)$ aus Bsp. 2.1.4.4. für $\alpha_n = \frac{1}{n}$ , $\alpha_n = 1$ und $\alpha_n = \sqrt{n}$ . . . . .     | 121 |
| 2.2.1 | Fläche unter einem Graphen . . . . .   | 143 |
| 2.2.2 | Rotationskörper . . . . .  | 144 |
| 2.2.3 | Riemannsummen am Rotationskörper . . . . .   | 145 |
| 2.3.1 | Graph und Höhenlinien der Funktion $f(x_1, x_2) = x_1^2 - 2x_2^2 - 2x_1x_2 - 4$ . . . . .                        | 152 |
| 2.5.1 | Höhenlinie $H_{f,-15}$ von $f(x, y) = x^2 - 2y^2 - 2xy - 4$ . . . . .  | 189 |
| 2.5.2 | Die Menge $H_{f,0}$ für $f(x, y, z) = x^2 + y^3 - 3xy$ . . . . .   | 195 |
| 2.5.3 | Die Graphen zu $\hat{H}_{f,0} = \{(x, y) \mid x^2 + y^3 - 3xy = 0\}$ . . . . .                                   | 197 |
| 2.6.1 | Die zweidimensionale Situation . . . . .   | 207 |
| 2.7.1 | Der Verlauf der Lösungen der logistischen DGL . . . . .  | 223 |
| 2.7.2 | Die RC-Schaltung . . . . .   | 225 |
| 2.7.3 | Spannungsverlauf der RC-Schaltung . . . . .  | 226 |
| 2.7.4 | Die Lösungen $y_c$ des AWP $y' = 2\sqrt{y}$ , $y(0) = 0$ , für $c = 0, \frac{1}{2}, 1, \frac{3}{2}, 2$ . . . . . | 227 |
| 3.1.1 | Lösungen der homogenen Schwingungsgleichung . . . . .  | 247 |

---

|       |  |     |
|-------|--|-----|
| 3.1.2 | Der Amplitudenverlauf $A(\varpi)$ und die Phaseverschiebung $\phi(\varpi)$<br>für $\omega = 1$ und ausgewählte $\frac{2}{15} \leq \gamma \leq \frac{29}{15}$ . . . . . | 249 |
| 3.1.3 | Die Tacoma-Narrows-Brücke . . . . .  | 250 |
| 3.2.1 | Das Volumen von Quadern in Dimension 1, 2 und 3 . . . . .  | 263 |
| 3.3.1 | Der Satz von Fubini . . . . .  | 277 |
| 3.3.2 | Das Kugelschalensegment als Bild des Würfels unter $\Phi$ . . . . .  | 281 |
| 3.4.1 | Zum Submersionssatz 3.4.1.1 . . . . .  | 285 |
| 3.4.2 | Zum Immersionssatz 3.4.1.2 . . . . .   | 285 |
| 3.4.3 | Der Kartenwechsel . . . . .  | 289 |
| 3.4.4 | Kurven, die keine Mannigfaltigkeiten sind . . . . .  | 294 |
| 3.5.1 | Konstruktion der disjunkten Überdeckung . . . . .  | 300 |
| 3.7.1 | Geometrische Interpretation der komplexen Multiplikation . . . . .   | 337 |

---

## Literaturverzeichnis

---

- [Bl1] Christian Blatter: *Analysis I*. (Heidelberger Taschenbücher 151), Springer Verlag, I: 3. Aufl. 1980
- [Bl2] Christian Blatter: *Analysis II*. (Heidelberger Taschenbücher 152), Springer Verlag, 2. Aufl. 1979
- [Bl3] Christian Blatter: *Analysis III*. (Heidelberger Taschenbücher 153), Springer Verlag, 3. Aufl. 1981
- [Br] I. N. Bronstein und K. A. Semendjajew: *Taschenbuch der Mathematik*. B.G. Teubner Verlagsgesellschaft, 25. Aufl. 1991
- [Fo1] Otto Forster: *Analysis 1: Differential- und Integralrechnung einer Veränderlichen*. Springer Verlag, 11. Aufl. 2013
- [Fo2] Otto Forster: *Analysis 2: Differentialrechnung im  $\mathbb{R}^n$ , gewöhnliche Differentialgleichungen*. Springer Spektrum, 10. Aufl. 2013
- [Fo3] Otto Forster: *Analysis 3: Maß- und Integrationstheorie, Integralsätze im  $\mathbb{R}^n$  und Anwendungen*. Vieweg+Teubner Verlag, 7. Aufl. 2012
- [He1] Harro Heuser: *Lehrbuch der Analysis Teil 1*. Vieweg+Teubner Verlag, Teil 1: 17. Aufl. 2009
- [He2] Harro Heuser: *Lehrbuch der Analysis Teil 2*. Vieweg+Teubner Verlag, Teil 2: 14. Aufl. 2008
- [Ko] Hans-Joachim Kowalsky: *Lineare Algebra*. De Gruyter Verlag, 9. Aufl. 1979

- [La] Edmund Landau: *Grundlagen der Analysis*. Chelsea Publishing Company, N.Y., 4. Aufl. 1965
- [Lo1] Falko Lorenz: *Lineare Algebra I*. Spektrum Akademischer Verlag, I: 4. Aufl. 2003
- [Lo2] Falko Lorenz: *Lineare Algebra II*. Spektrum Akademischer Verlag, II: 3. Aufl. 1992, Nachdruck 2008
- [QSS] Alfio Quarteroni, Riccardo Sacco und Fausto Saleri: *Numerical Mathematics* Springer Verlag, 2nd ed. 2007