



UNIVERSITÄT  
DES  
SAARLANDES

Michael Bildhauer

Mathematik für Naturwissenschaftler I & II

Wintersemester 2010/2011, Sommersemester 2011

Universität des Saarlandes  
F.R. 6.1 Mathematik  
Postfach 151150  
D-66041 Saarbrücken



# Inhaltsverzeichnis

<b>1</b>	<b>Mengen und Funktionen</b>	<b>7</b>
1.1	Mengen . . . . .	7
1.2	Funktionen . . . . .	10
<b>2</b>	<b>Natürliche, ganze und rationale Zahlen</b>	<b>15</b>
2.1	Natürliche Zahlen . . . . .	15
2.2	Ganze und rationale Zahlen . . . . .	19
<b>3</b>	<b>Reelle Zahlen und Funktionen</b>	<b>21</b>
3.1	Reelle Zahlen . . . . .	21
3.2	Rechnen mit Ungleichungen und Beträgen . . . . .	24
3.3	Eigenschaften reeller Funktionen . . . . .	25
3.4	Spezielle reelle Funktionen . . . . .	28
3.4.1	Potenzfunktionen und enge Verwandte . . . . .	29
3.4.2	Exponentialfunktion und Logarithmus . . . . .	33
3.4.3	Trigonometrische Funktionen . . . . .	37
3.4.4	Hyperbolische Funktionen . . . . .	41
<b>4</b>	<b>Komplexe Zahlen</b>	<b>45</b>
<b>5</b>	<b>Folgen und Reihen</b>	<b>53</b>
5.1	Folgen . . . . .	54
5.2	Reihen . . . . .	63
<b>6</b>	<b>Potenzreihen</b>	<b>71</b>
<b>7</b>	<b>Stetige Funktionen</b>	<b>77</b>
<b>8</b>	<b>Differentialrechnung in einer Veränderlichen</b>	<b>85</b>
8.1	Grundlagen . . . . .	85

8.2	Lokale Extrema, Mittelwertsatz . . . . .	94
8.3	Der Satz von Taylor . . . . .	103
<b>9</b>	<b>Integralrechnung in einer Veränderlichen</b>	<b>109</b>
9.1	Das bestimmte Riemannsche Integral . . . . .	109
9.2	Der Hauptsatz der Differential- und Integralrechnung . .	114
9.3	Integrationstechniken . . . . .	117
9.3.1	Einfache Integrationstechniken . . . . .	117
9.3.2	Partielle Integration . . . . .	119
9.3.3	Substitutionsregel . . . . .	120
9.3.4	Partialbruchzerlegung . . . . .	122
9.4	Uneigentliche Integrale . . . . .	127
9.5	Numerische Integration . . . . .	131
<b>10</b>	<b>Lineare Gleichungssysteme: Das Gaußsche Eliminationsverfahren</b>	<b>135</b>
<b>11</b>	<b>Der Vektorraum <math>\mathbb{R}^n</math></b>	<b>143</b>
11.1	Vektorräume . . . . .	143
11.2	Norm und Skalarprodukt . . . . .	150
11.3	Analytische Geometrie . . . . .	156
<b>12</b>	<b>Matrizen und lineare Gleichungssysteme</b>	<b>167</b>
12.1	Matrizenkalkül . . . . .	167
12.2	Zur Lösbarkeit linearer Gleichungssysteme . . . . .	174
12.3	Quadratische Matrizen . . . . .	181
12.3.1	Invertierbare Matrizen . . . . .	181
12.3.2	Die Determinante . . . . .	183
12.3.3	Spezialfälle quadratischer Matrizen . . . . .	189
12.3.4	Hauptachsentransformation: Eigenwerte und Eigenvektoren . . . . .	190
<b>13</b>	<b>Grenzwerte und Stetigkeit im <math>\mathbb{R}^n</math></b>	<b>199</b>
<b>14</b>	<b>Differentialrechnung von Funktionen mehrerer Veränderlicher</b>	<b>207</b>
14.1	Kurven im $\mathbb{R}^n$ . . . . .	208
14.2	Ableitungen einer Funktion mehrerer Veränderlicher . . .	215

14.3	Extremwertaufgaben in mehreren Veränderlichen . . . .	228
<b>15</b>	<b>Kurvenintegrale</b>	<b>233</b>
<b>16</b>	<b>Integrale im <math>\mathbb{R}^n</math></b>	<b>241</b>
16.1	Das Riemannsches Integral . . . . .	241
16.2	Der Transformationssatz . . . . .	256
<b>17</b>	<b>Der Gaußsche Integralsatz in der Ebene</b>	<b>263</b>
<b>18</b>	<b>Elementares zur Wahrscheinlichkeitsrechnung und Sta-</b>	
	<b>tistik</b>	<b>269</b>
18.1	Grundbausteine der Kombinatorik . . . . .	269
18.2	Wahrscheinlichkeitsrechnung . . . . .	274



# Kapitel 1

## Mengen und Funktionen

### 1.1 Mengen

Die **Mengenlehre** ist die Sprache der Mathematik und geht als “naive” (im Gegensatz zur strengen Axiomatik) Mengenlehre auf Georg Cantor (1845–1918) zurück.

#### Intuitive Beispiele für Mengen.

- i)* Zahlen:  $\mathbb{N}$  (natürliche),  $\mathbb{Z}$  (ganze),  $\mathbb{Q}$  (rationale),  $\mathbb{R}$  (reelle),  $\mathbb{C}$  (komplexe).
- ii)* Endliche Mengen als **Aufzählung**, z.B.  $\{1, 2, 3\}$ , wobei die Reihenfolge keine Rolle spielt:  $\{1, 2, 3\} = \{3, 1, 2\}$ .
- iii)* **Angedeutete Aufzählungen** von Mengen wie etwa  $\{1, 2, 3, \dots, 50\}$  oder  $\{1, 2, 3, \dots, \} = \mathbb{N}$ .
- iv)* Charakterisierung einer Menge über eine **gemeinsame Eigenschaft**, z.B.  $\{n \in \mathbb{N} : n \text{ ist Primzahl}\}$ .
- v)* Die **leere Menge**  $\emptyset$ , die kein Element enthält.

**Definition 1.1.** *Eine Menge  $A$  ist eine Zusammenfassung von wohl bestimmten und wohl unterschiedenen Objekten unserer Anschauung oder unseres Denkens, den Elementen der Menge, zu einem Ganzen.*

**Kann man Mengen vergleichen bzw. mit Mengen operieren (Mengenalgebra)?**

Zur Veranschaulichung bedient man sich meist sogenannter Venn-Diagramme (Abbildung 1.1).

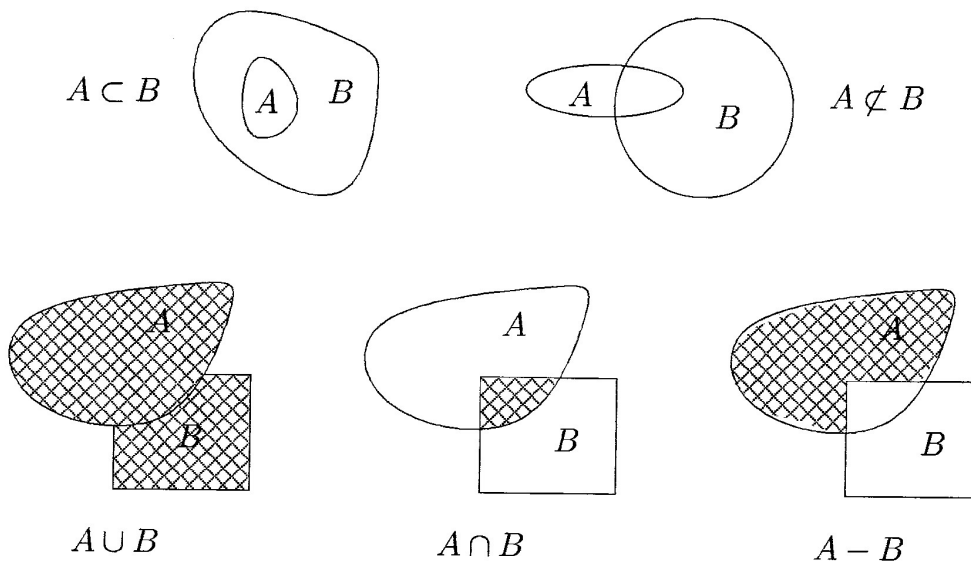


Abbildung 1.1: Venn-Diagramme.

**Interpretation von Abbildung 1.1.** Es seien  $A$  und  $B$  zwei Mengen.

Zum **Vergleich der Mengen** dienen die Eigenschaften:

- i)  $A$  ist gleich  $B$ , d.h.  
 $A = B :\Leftrightarrow A$  und  $B$  haben dieselben Elemente.
- ii)  $A$  ist Teilmenge von  $B$ , d.h.  
 $A \subset B :\Leftrightarrow$  Jedes Element von  $A$  gehört zu  $B$ .

Es ist  $A = B \Leftrightarrow$  (genau dann, wenn)  $A \subset B$  und  $B \subset A$ . In der Tat werden nicht-offensichtliche Gleichheiten meist mithilfe **beider** Teilmengenbeziehungen überprüft.

**Beispiel.** Es sei  $A$  die Menge aller ungeraden Zahlen und  $B$  die Menge aller Primzahlen  $\geq 3$ .

Ist  $x \in B$ , ist also  $x$  Primzahl  $\geq 3$ , so muss  $x$  ungerade sein (sonst wäre  $x$  durch 2 teilbar) und folglich ist  $x \in A$ . Dies gilt für **jedes beliebige**  $x \in B$ , also  $B \subset A$ .

Umgekehrt gibt es aber ein  $x \in A$ , welches keine Primzahl ist



(z.B.  $x = 15$ ), und deshalb gilt  $A \not\subset B$ .

**Übung.** Warum gilt für jede Menge  $M$ :  $M \subset M$  und  $\emptyset \subset M$ ?

Die in Abbildung 1.1 skizzierten **Mengenoperationen** sind

i) die **Vereinigung**

$$A \cup B := \{x : x \in A \vee x \in B\},$$

ii) der **Durchschnitt**

$$A \cap B := \{x : x \in A \wedge x \in B\},$$

iii) und die **Differenz** (das **Komplement** von  $B$  in  $A$ )

$$A - B = A \setminus B := \{x : x \in A \wedge x \notin B\}.$$

Beispiele werden in den Übungen diskutiert.

**Sprechweise.** Zwei Mengen  $A$  und  $B$  mit  $A \cap B = \emptyset$  heißen **disjunkt**.

**Wie kann der dreidimensionale Euklidische Raum geeignet als Menge geschrieben werden?**

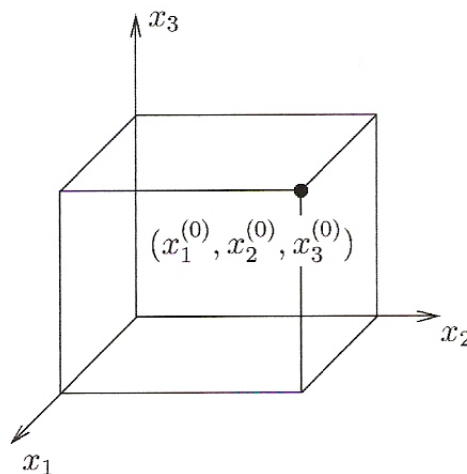


Abbildung 1.2: Der dreidimensionale Euklidische Raum  $\mathbb{R}^3$ .

Dazu sollte jedem Punkt ein **geordnetes Tupel** mit drei Eintragungen (den **Koordinaten**) zugeordnet werden (vgl. Abbildung 1.2). Geordnet bedeutet, dass die **Reihenfolge wesentlich ist**.

Ein solches Tupel wird über das **kartesische Produkt von Mengen** definiert: Es seien  $A$  und  $B$  zwei Mengen. Dann ist

$$A \times B := \{(a, b) : a \in A \wedge b \in B\}$$

die Menge der geordneten Paare (Zur Veranschaulichung siehe Abbildung 1.3).

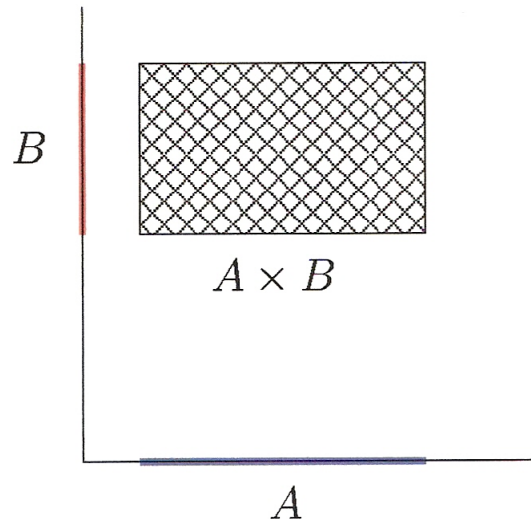


Abbildung 1.3: Veranschaulichung des kartesischen Produktes.

Somit ist  $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$  bzw. für  $n \in \mathbb{N}$

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}} .$$

## 1.2 Funktionen

Die Beschreibung von Vorgängen in der Natur erfolgt mittels Funktionen (auch Abbildungen genannt). Betrachtet sei beispielsweise die Temperaturverteilung in einem Raum, d.h. jedem Punkt des Raumes wird eine Temperatur zugeordnet.

**Beobachtung.** Einem Raumpunkt können nicht zwei Temperaturen zugeordnet werden, verschiedenen Raumpunkten kann aber derselbe Temperaturwert zugeordnet werden.

**Definition 1.2.** Eine *Abbildung* oder *Funktion*  $f$  von einer Menge  $A$  in eine Menge  $B$  ordnet *jedem* Element  $x$  aus  $A$  *genau ein* Element  $y$  aus  $B$  zu. Notation:

$$f : A \rightarrow B, \quad y = f(x),$$

oder auch

$$f : A \ni x \mapsto f(x) \in B.$$

$A$  heißt der *Definitionsbereich* oder die *Urbildmenge* von  $f$ ,  $B$  der *Bildbereich* und  $y = f(x)$  der *Bildpunkt* des *Urbildes*  $x$  unter  $f$ .

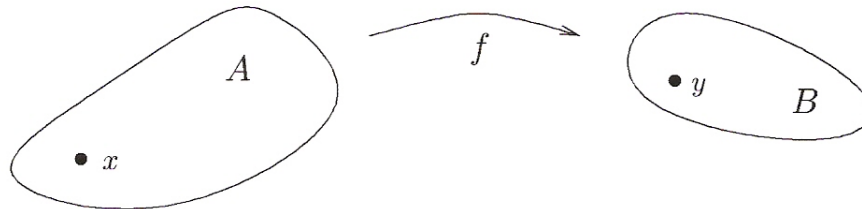


Abbildung 1.4: Eine Abbildung zwischen zwei Mengen.

### Bemerkungen.

- $x \in A$  wie oben bezeichnet man auch als *Argument* oder *unabhängige Variable* von  $f$ , die Menge der geordneten Paare

$$\{(x, f(x)) : x \in A\}$$

heißt der *Graph* von  $f$  (vgl. Abbildung 1.5).

- Statt  $x$  kann ein beliebiges anderes Symbol für die unabhängige Variable stehen, gleiches gilt für  $y$  und  $f$ .

**Einfache Beispiele.** Die einfachsten Beispiele sind (vgl. Abbildung 1.6)

- i) die *Identität*  $f: A \rightarrow A, f(x) = x \forall x \in A$ ;
- ii) die *konstante Abbildung*  $f: A \rightarrow B, f(x) = b \forall x \in A$  mit fixiertem  $b \in B$ .

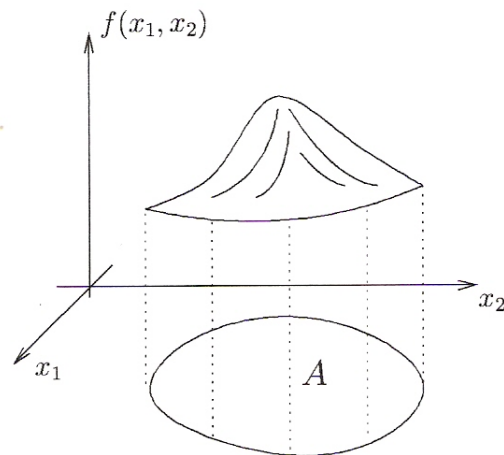


Abbildung 1.5: Der Graph einer Funktion.

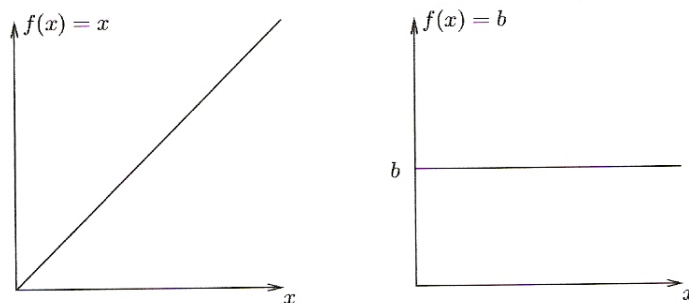


Abbildung 1.6: Identität und konstante Abbildung.

## Erste charakteristische Eigenschaften einer Abbildung.

### Beispiele.

- An obigem Beispiel der Temperaturverteilung erkennt man: Zwei Punkten kann der gleiche Temperaturwert zugeordnet werden. In einem gegebenen Raum wird aber sicher nicht die komplette Temperaturskala beobachtet (z.B. kommen in einem beheizten Raum keine Temperaturen unter  $18^\circ \text{C}$  vor).
- Gegeben sei ein Skatblatt mit 32 Karten. 3 Karten werden nacheinander gezogen und jeweils herausgelegt. Diese Ziehung kann als Funktion beschrieben werden, indem der Menge  $\{1, 2, 3\}$  (erstes, zweites, drittes Ziehen) die jeweils gezogene Karte zugeordnet

wird.

Da keine Karte zweimal gezogen werden kann, wird jeder Zahl aus der Menge  $\{1, 2, 3\}$  ein anderes Bild zugeordnet.

Klarerweise werden nicht alle Karten gezogen (nur 3 von 32).

- Betrachtet sei die Funktion  $f: \mathbb{N} \rightarrow \{0, 1\}$  mit

$$f(n) = \begin{cases} 0, & \text{falls } n \text{ ungerade;} \\ 1, & \text{falls } n \text{ gerade.} \end{cases}$$

Hier wird z.B. der 1 und der 3 das gleiche Bild (die 0) zugeordnet. Alle möglichen Bildwerte (0 und 1) werden angenommen.

- Kann etwa ein symmetrisches zweidimensionales Problem leichter in Polarkoordinaten  $(r, \varphi)$  (Radius, Winkel) als in kartesischen Koordinaten  $(x, y)$  gelöst werden, so muss die Umformung des Problems auf Polarkoordinaten eindeutig sein, und die Lösung muss wiederum eindeutig in den eigentlich gesuchten kartesischen Koordinaten auszudrücken sein, d.h. benötigt wird eine **1-1 Abbildung**  $f: (x, y) \mapsto (r, \varphi)$ .

**Definition 1.3.** Es seien  $A, B$  zwei Mengen und  $f$  eine Abbildung von  $A$  nach  $B$ . Dann heißt  $f$

- i) *injektiv*, falls aus  $x_1, x_2 \in A$  mit  $x_1 \neq x_2$   $f(x_1) \neq f(x_2)$  folgt;
- ii) *surjektiv*, falls jeder Punkt von  $B$  ein Bildpunkt ist, d.h. falls  $\text{Bild } f = B$ ;
- iii) *bijektiv*, falls  $f$  sowohl injektiv als auch surjektiv ist. In diesem Fall existiert eine *Umkehrabbildung*  $f^{-1}: B \rightarrow A$ , die über die Vorschrift

$$f^{-1}(y) := x \text{ für } y = f(x)$$

definiert ist. Insbesondere gilt also

$$f^{-1}(f(x)) = x \quad \forall x \in A \quad \text{und} \quad f(f^{-1}(y)) = y \quad \forall y \in B.$$

## Übungen.

- i) Man mache sich die Begriffe anhand obiger Beispiele klar.

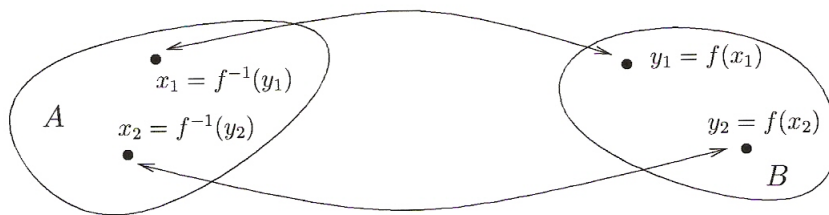


Abbildung 1.7: Zur Umkehrabbildung.

ii) **Definitions- und Bildbereich sind wesentliche Voraussetzungen in Definition 1.3.** Es sei z.B.  $f: \mathbb{R} \supset A \rightarrow B \subset \mathbb{R}, x \mapsto x^2$ . Finden Sie jeweils  $A, B$ , sodass

- $f$  surjektiv aber nicht injektiv ist;
- $f$  bijektiv ist;
- $f$  weder injektiv noch surjektiv ist.

### Beschreibung von aufeinanderfolgenden Vorgängen.

**Definition 1.4.** Sind drei Mengen  $A, B, C$  und zwei Funktionen  $f: A \rightarrow B$  und  $g: B \rightarrow C$  gegeben, so heißt die Abbildung

$$g \circ f: A \rightarrow C, \quad (g \circ f)(x) := g(f(x))$$

die *Hintereinanderschaltung* oder *Verkettung* von  $f$  und  $g$ .

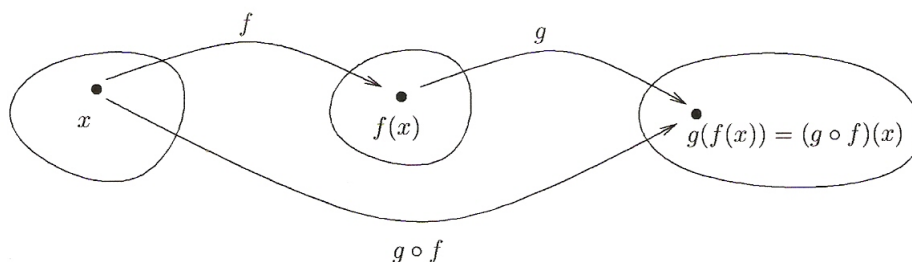


Abbildung 1.8: Zur Verkettung von Abbildungen.

# Kapitel 2

## Natürliche, ganze und rationale Zahlen

### 2.1 Natürliche Zahlen

Jeder hat eine intuitive Vorstellung von der Menge der **natürlichen Zahlen**

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

Sie sind im Wesentlichen charakterisiert durch die Existenz einer **Nachfolgeabbildung** (vgl. Abbildung 2.1), d.h.:

- i) 1 ist eine natürliche Zahl;
- ii) zu jeder natürlichen Zahl  $n$  gibt es eine nachfolgende natürliche Zahl.

**Bemerkung.** Null ist im Sinne dieser Charakterisierung **keine** natürliche Zahl, auch wenn dies in der Literatur nicht einheitlich vereinbart ist. Die Menge der natürlichen Zahlen vereinigt mit der Null wird hier mit  $\mathbb{N}_0$  bezeichnet.

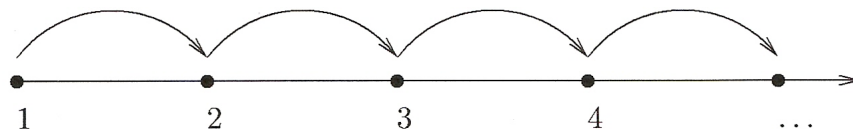


Abbildung 2.1: Zur Nachfolgeabbildung.

**Konsequenz.** Die Existenz einer Nachfolgeabbildung führt unmittelbar auf das Beweisprinzip der **vollständigen Induktion**. Es erinnert

an eine Reihe von Dominosteinen, die so aufgestellt sind, dass wenn irgendein Stein fällt (etwa der mit der Nummer  $n$ ) auch der nächste Stein fällt (der mit der Nummer  $n + 1$ ). Mathematisch spricht man vom **Induktionsschluss**. Wird also der erste Stein umgeworfen, d.h. gibt es einen **Induktionsanfang**, so fallen alle Steine.

**Beispiel.** Die Behauptung

$$1 + 2 + 3 + 4 + \cdots + n =: \sum_{k=1}^n k = \frac{n(n+1)}{2} \quad (1)$$

soll bewiesen werden.

**Exkurs Summenzeichen.** Es seien  $a_1, a_2, \dots, a_n$  Summanden mit Werten in den natürlichen, ganzen, rationalen, reellen oder komplexen Zahlen (Für  $k = 1, \dots, n$  ist im Beispiel  $a_k = k$ ). Dann wird die Summe

$$a_1 + a_2 + a_3 + \cdots + a_n$$

**kürzer geschrieben** (und mehr passiert hier nicht) als

$$\sum_{k=1}^n a_k .$$

Man nennt  $k$  den **Summationsindex** und die Menge aller  $k$ , über die summiert wird, die **Indexmenge**.

Der Summationsindex kann **umbenannt oder durch einen anderen Ausdruck ersetzt werden**, wobei sich die Indexmenge evtl. verändert. Die Indexmenge muss nicht bei 1 anfangen, es ist jede ganze Zahl zulässig:

$$\begin{aligned} \sum_{k=1}^n a_k &= \sum_{l=1}^n a_l , \\ \sum_{k=1}^n a_k &= \sum_{j=2}^{n+2} a_{j-1} = \sum_{i=-3}^{n-4} a_{i+4} , \\ \sum_{m=0}^2 b_{2m} &= b_0 + b_2 + b_4 \quad (\text{gerade}) , \\ \sum_{m=0}^2 b_{2m+1} &= b_1 + b_3 + b_5 \quad (\text{ungerade}) . \end{aligned}$$



Im Falle **konstanter Summanden** ( $a_1 = a_2 = \dots = a_n = a$ ) gilt offensichtlich

$$\sum_{k=1}^n a_k = \underbrace{a + a + \dots + a}_{n\text{-mal}} = na .$$

Ebenso offensichtlich sind die **Rechenregeln**

$$\sum_{k=n}^m a_k + \sum_{k=n}^m b_k = \sum_{k=n}^m (a_k + b_k) , \quad c \sum_{k=n}^m a_k = \sum_{k=n}^m (ca_k) .$$

**Übung.** Man schreibe das Produkt zwei Summen als **Doppelsumme**.

Zurück zur Behauptung (1) aus obigem Beispiel: Zunächst ist der erste Stein umzuwerfen, d.h. die Behauptung ist für  $n = 1$  zu verifizieren.

**Induktionsanfang** ( $n = 1$ ): In diesen Fall lautet die Behauptung

$$\sum_{k=1}^1 k = \frac{1(1+1)}{2}$$

und ist offensichtlich richtig. Damit ist der Induktionsanfang gemacht.

Anschließend ist zu zeigen, dass alle Steine im richtigen Abstand voneinander aufgestellt sind. Nimmt man also an, dass der Stein mit der Nummer  $n$  fällt (“ $n$  beliebig aber fest”), so muss zwingend daraus folgen, dass dann auch der nächste Stein – der mit der Nummer  $n + 1$  – fällt.

**Induktionsschluss**(“ $n \rightsquigarrow n + 1$ ”): Man nehme an, dass die Behauptung für ein fixiertes  $n$  gültig ist:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2} \quad \text{für dieses fixierte } n . \tag{2}$$

Mit dieser Annahme ist die Behauptung für  $n + 1$  zu zeigen, d.h. zu zeigen ist **unter der Annahme (2)**

$$\sum_{k=1}^{n+1} k = \frac{(n+1)((n+1)+1)}{2} = \frac{(n+1)(n+2)}{2} . \tag{3}$$

In der Tat gilt

$$\begin{aligned} \sum_{k=1}^{n+1} k &= \underbrace{\sum_{k=1}^n k}_{=n(n+1)/2 \text{ nach der Annahme (2)}} + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) = (n+1) \left( \frac{n}{2} + 1 \right) = \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Das ist genau Gleichung (3), die zu zeigen war.  $\square$

**Bemerkung.** Wichtig ist, dass der Induktionsschluss für jedes fixierte  $n$  richtig sein muss, alle Steine müssen den richtigen Abstand voneinander haben, damit die Kette nicht abreißt.

**Beispiel.** Nun soll die Behauptung

$$\prod_{i=1}^n \frac{i}{i+1} = \frac{1}{n+1} \quad (4)$$

gezeigt werden. Das kann entweder direkt durch Kürzen geschehen oder mit analogen Argumenten wie oben mittels vollständiger Induktion gezeigt werden.

**Exkurs Produktzeichen.** Es seien  $a_1, a_2, \dots, a_n$  Faktoren analog zu den Summanden im obigen Exkurs zu Summenzeichen. Die abkürzende Schreibweise für das Produkt

$$a_1 \cdot a_2 \cdot \dots \cdot a_n$$

ist

$$\prod_{k=1}^n a_k.$$

Der **Produktindex** und die Indexmenge werden genauso behandelt wie der Summationsindex und die Indexmenge.

Zurück zum Beispiel:

**Induktionsanfang** ( $n = 1$ ):

$$\prod_{i=1}^1 \frac{i}{i+1} = \frac{1}{1+1}$$

ist wieder offensichtlich richtig und der Induktionsanfang ist gemacht.

**Induktionsschluss** (“ $n \rightsquigarrow n + 1$ ”): Es sei  $n \in \mathbb{N}$  fixiert (beliebig aber fest) und als Annahme gelte

$$\prod_{i=1}^n \frac{i}{i+1} = \frac{1}{n+1}.$$

Dann ist

$$\prod_{i=1}^{n+1} \frac{i}{i+1} = \prod_{i=1}^n \frac{i}{i+1} \cdot \frac{n+1}{n+2} = \frac{1}{n+1} \frac{n+1}{n+2} = \frac{1}{n+2}.$$

Das ist genau die “Behauptung für  $n + 1$ ”, der Induktionsschluss ist ebenfalls richtig und (4) ist bewiesen.  $\square$

## 2.2 Ganze und rationale Zahlen

**Was ist mit der Gleichung  $x + 4 = 3$  in  $\mathbb{N}$ ?**

Diese Gleichung ist in der Menge der natürlichen Zahlen nicht lösbar. Man benötigt die Menge der **ganzen Zahlen**

$$\mathbb{Z} := \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}.$$

**Bemerkung.** Die ganzen Zahlen versehen mit der Addition haben eine besondere Struktur:  $(\mathbb{Z}, +)$  ist eine **kommutative Gruppe** (siehe Übung).

**Was ist mit der Gleichung  $4x = 3$  in  $\mathbb{Z}$ ?**

Diese Gleichung ist in der Menge der ganzen Zahlen nicht lösbar. Man benötigt die Menge der **rationalen Zahlen** (die Menge der Brüche)

$$\mathbb{Q} := \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N} \right\}.$$

**Bemerkung.** Die rationalen Zahlen versehen mit der Addition **und** der Multiplikation haben eine besondere Struktur:  $(\mathbb{Q}, +, \cdot)$  ist ein **Körper** (Übung).

### Zusammenhang mit Dezimalzahlen?

Es gilt z.B.

$$\frac{1}{2} = 0.5, \quad \frac{1}{3} = 0.333\cdots = 0.\bar{3}, \quad \frac{1}{6} = 0.1\bar{6}, \quad \frac{8}{7} = 1.\overline{142857}.$$

Analoges ist auch im Allgemeinen richtig: Jede rationale Zahl lässt sich als **abbrechende** oder **periodische Dezimalzahl** schreiben und umgekehrt stellt jede abbrechende oder periodische Dezimalzahl eine rationale Zahl dar.

### Sind rationale Zahlen eindeutig als Dezimalzahl darstellbar?

Die Antwort ist **nein**, denn es gilt etwa

$$\frac{1}{2} = 0.5 = 0.4\bar{9}, \quad \frac{9}{20} = 0.45 = 0.44\bar{9}.$$

Verlangt man von der Dezimaldarstellung **zusätzlich**, dass sie nicht abbricht, so erhält man die Charakterisierung

$$\mathbb{Q} = \{\text{nicht-abbrechend periodische Dezimalzahlen}\}.$$

# Kapitel 3

## Reelle Zahlen und Funktionen

### 3.1 Reelle Zahlen

Was ist mit der Gleichung  $x^2 = 2$  in  $\mathbb{Q}$ ?

Diese Gleichung ist in der Menge der rationalen Zahlen nicht lösbar. Man benötigt die Menge der **reellen Zahlen**

$$\mathbb{R} := \{\text{nicht-abbrechende Dezimalzahlen}\}.$$

Zu den rationalen Zahlen kommen nicht nur die Wurzeln wie

$$\sqrt{2} = 1.141421356 \dots \in \mathbb{R} - \mathbb{Q}$$

hinzu, auch sogenannte **transzendente Zahlen** wie

$$\pi = 3.141592653 \dots \in \mathbb{R} - \mathbb{Q}$$

gehören zu den reellen Zahlen. Sie sind **keine Lösungen algebraischer Gleichungen**.

Die Menge  $\mathbb{R} - \mathbb{Q}$  heißt die Menge der **irrationalen Zahlen**.

**Bemerkung.** Mathematisch exakt werden die reellen Zahlen durch ein System von **Axiomen** eingeführt. Axiome sind nicht herzuleiten. Es handelt sich um grundlegende Aussagen, die als **wahr angenommen werden**. Alle weiteren Aussagen sind daraus abzuleiten.

## Addition, Subtraktion, Multiplikation und Division von reellen Zahlen.

Diese grundlegenden Operationen wurden für  $\mathbb{N}$ ,  $\mathbb{Z}$  und  $\mathbb{Q}$  als intuitiv bekannt vorausgesetzt und nicht diskutiert (beispielsweise können rationale Zahlen als Brüche addiert werden).

Im Reellen sind nach obiger aber Dezimalzahlen mit unendlich vielen Stellen etwa zu addieren – wie ist das erklärt?

Beim üblichen Additionsschema kann kein Übertrag erklärt werden, da man nicht ab einer konkreten Stelle anfangen kann zu addieren.

**Ausweg.** Durch Approximation mit rationalen Zahlen (s.u.) können die Rechenoperationen auf die in  $\mathbb{Q}$  zurückgeführt werden. Sie entsprechen dann nach wie vor der Vorstellung.

## Vergleich von reellen Zahlen.

Wie in  $\mathbb{Q}$  gibt es eine **Anordnungsrelation** “ $<$ ” (man kann reelle Zahlen der Größe nach ordnen) mit den Eigenschaften (vgl. Rechnen mit Ungleichungen)

- i) **Transitivität**, d.h.  
aus  $x < y$  und  $y < z$  folgt  $x < z$ ;
- ii) **Verträglichkeit mit der Addition**, d.h.  
aus  $x < y$  folgt  $x + z < y + z$  für alle  $z \in \mathbb{R}$ ;
- iii) **Verträglichkeit mit der Multiplikation**, d.h.  
aus  $x < y$  und  $z > 0$  folgt  $x \cdot z < y \cdot z$ .

Diese Anordnungsrelation wird durch den Pfeil auf der Zahlengeraden symbolisiert.

## Die Verteilung der Zahlenmengen auf dem Zahlenstrahl.

Jeder reellen Zahl entspricht **genau ein** Punkt auf der Zahlengeraden.

**Beispiel.** Wie in Abbildung 3.1 angedeutet betrachte man die Menge

$$M = \{x \in \mathbb{R} : x < \sqrt{2}\}.$$

Die ‘Grenze’ zwischen den Elementen aus  $M$  und den Zahlen, die nicht in  $M$  liegen, ist  $\sqrt{2}$ .

Das sogenannte **Vollständigkeitsaxiom** besagt grob gesprochen, dass alle solchen ‘Grenzen’ reelle Zahlen sein müssen, d.h. im Gegensatz zu der Menge der rationalen Zahlen gibt es in der Menge der reellen Zahlen keine Lücken auf dem Zahlenstrahl.

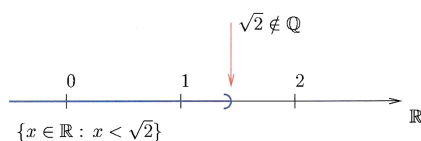


Abbildung 3.1: Zum Vollständigkeitsaxiom.

Als Konsequenz folgt das sogenannte **Intervallschachtelungsprinzip**, das – ebenfalls grob gesprochen – besagt, dass im Durchschnitt immer kleiner werdender Intervalle wie in Abbildung 3.2 eine reelle Zahl liegen muss. Auch dies ist in der Menge der rationalen Zahlen nicht richtig (Beispiel?).

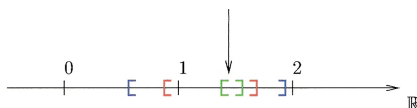


Abbildung 3.2: Zum Intervallschachtelungsprinzip.

Es gibt ‘viel mehr’ irrationale als rationale Zahlen, andererseits können **irrationale Zahlen aber beliebig gut mit rationalen Zahlen approximiert werden**. Man sagt,  $\mathbb{Q}$  **liegt dicht** in  $\mathbb{R}$ .

Anschaulich wird das, indem man eine reelle Zahl durch eine Dezimalzahl mit **nur endlich vielen Stellen** approximiert, was in Rechnern tatsächlich der Fall sein muss. Je mehr Stellen zur Verfügung stehen, desto besser nähert man sich der gegebenen Zahl.

### 3.2 Rechnen mit Ungleichungen und Beträgen

Für das Rechnen mit Ungleichungen gelten die folgenden elementaren Regeln, die man sich leicht mithilfe von einfachen Beispielen klar machen kann.

$$i) \quad x \leq y \Rightarrow -x \geq -y;$$

$$ii) \quad x \leq y \Rightarrow x^{-1} = \frac{1}{x} \geq \frac{1}{y} = y^{-1};$$

$$iii) \quad (x \leq y) \wedge (z \leq 0) \Rightarrow x \cdot z \geq y \cdot z;$$

$$iv) \quad x^2 \geq 0;$$

$$v) \quad (x \leq y) \wedge (u \leq v) \Rightarrow x + u \leq y + v;$$

$$vi) \quad (0 \leq x \leq y) \wedge (0 \leq u \leq v) \Rightarrow x \cdot u \leq y \cdot v.$$

#### Wie werden Längen und Abstände gemessen?

Das geeignete Werkzeug hierzu ist die **Betragsfunktion**, die in Abbildung 3.3 skizziert ist: Für  $x \in \mathbb{R}$  ist der Betrag von  $x$  definiert durch

$$|x| = \begin{cases} x & \text{für } x \geq 0; \\ -x & \text{für } x < 0. \end{cases}$$

Es ist beispielsweise  $|-2| = -(-2) = 2$ . Der Betrag  $|x - y|$  misst den Abstand zwischen zwei Punkten  $x, y$  auf der Zahlengeraden.

Elementare Regeln für das Rechnen mit Beträgen sind:

$$i) \quad |x| \geq 0;$$

$$ii) \quad |x| = 0 \Leftrightarrow x = 0;$$

$$iii) \quad |x \cdot y| = |x| |y|;$$

$$iv) \quad |x + y| \leq |x| + |y| \text{ (Dreiecksungleichung);}$$

$$v) \quad ||x| - |y|| \leq |x - y| \text{ (umgekehrte Dreiecksungleichung).}$$

**Übung.** Woher kommt der Name “Dreiecksungleichung”?



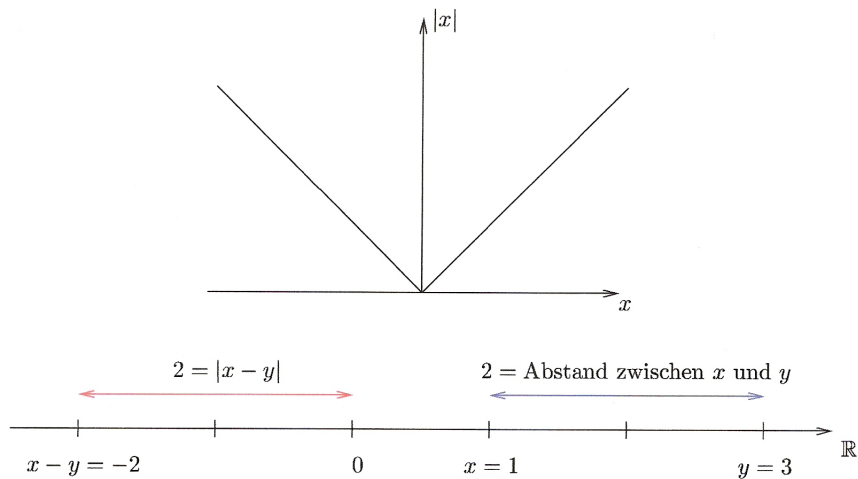


Abbildung 3.3: Zur Abstandsfunktion.

### 3.3 Eigenschaften reeller Funktionen

Neben den in Kapitel 1 erwähnten charakteristischen Eigenschaften von Funktionen gibt es im Fall reeller Funktionen weitere besondere Merkmale, die hier kurz aufgelistet werden sollen.

Es sei dazu  $f: \mathbb{R} \supset A \rightarrow B \subset \mathbb{R}$ , wobei man sich unter  $A$  und  $B$  reelle Intervalle vorstellen kann (nicht muss).

- i) (a) Die Funktion  $f$  heißt **monoton wachsend**, falls für alle  $x, y \in A$  mit  $x < y$  gilt

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

Gilt in der Ungleichung sogar das strikte “<”, so heißt  $f$  **streng monoton wachsend**.

- (b) Die Funktion  $f$  heißt **monoton fallend**, falls für alle  $x, y \in A$  mit  $x < y$  gilt

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

Gilt in der Ungleichung sogar das strikte “>”, so heißt  $f$  **streng monoton fallend**.

- ii) (a) Die Funktion  $f$  heißt **gerade**, falls mit  $x \in A$  auch  $-x \in A$  und falls für alle  $x \in A$  gilt

$$f(-x) = f(x) .$$

- (b) Die Funktion  $f$  heißt **ungerade**, falls mit  $x \in A$  auch  $-x \in A$  und falls für alle  $x \in A$  gilt

$$f(-x) = -f(x) .$$

- iii) Die Funktion  $f$  heißt **periodisch mit der Periode  $T > 0$** ,  $T \in \mathbb{R}$  fixiert, falls mit  $x \in A$  auch  $x + T \in A$  und falls für alle  $x \in A$  gilt

$$f(x + T) = f(x) .$$

- iv) (a) Die Funktion  $f$  heißt **nach oben beschränkt**, falls eine Konstante  $k \in \mathbb{R}$  existiert mit

$$f(x) \leq k \quad \text{für alle } x \in A .$$

Dann heißt  $k$  eine **obere Schranke**.

Die **kleinste obere Schranke** von  $f$  heißt das **Supremum** von  $f$ .

Notation:  $\sup_{x \in A} f(x)$ .

Existiert ein  $x_0 \in A$  mit

$$f(x_0) = \sup_{x \in A} f(x) ,$$

so heißt  $x_0$  ein **Maximierer** (oder eine **Maximalstelle**) und  $f(x_0)$  das Maximum von  $f$ . Notation:  $\max_{x \in A} f(x)$ .

- (b) Die Funktion  $f$  heißt **nach unten beschränkt**, falls eine Konstante  $k \in \mathbb{R}$  existiert mit

$$f(x) \geq k \quad \text{für alle } x \in A .$$

Dann heißt  $k$  eine **untere Schranke**.

Die **größte untere Schranke** von  $f$  heißt das **Infimum** von  $f$ .

Notation:  $\inf_{x \in A} f(x)$ .

Existiert ein  $x_0 \in A$  mit

$$f(x_0) = \inf_{x \in A} f(x) ,$$

so heißt  $x_0$  ein **Minimierer** (oder eine **Minimalstelle**) und  $f(x_0)$  das Minimum von  $f$ . Notation:  $\min_{x \in A} f(x)$ .

- (c) Die Funktion  $f$  heißt **beschränkt**, falls  $f$  nach oben und nach unten beschränkt ist.

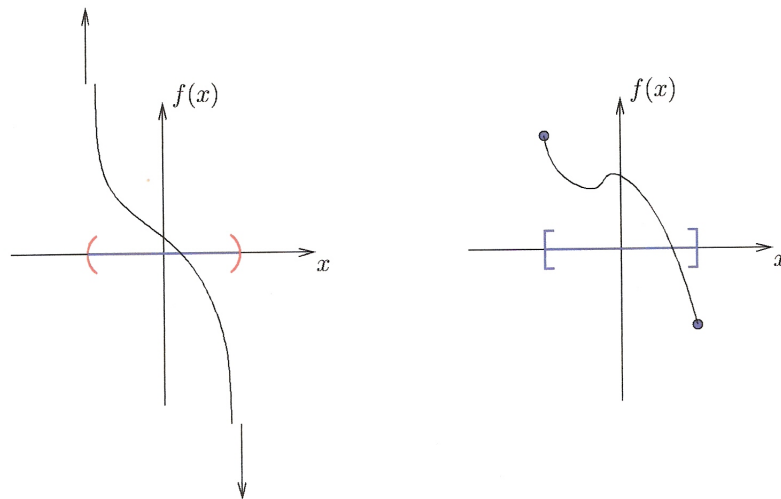


Abbildung 3.4: Die linke Funktion ist weder nach oben noch nach unten beschränkt, die rechte Funktion ist sowohl nach oben als auch nach unten beschränkt.

### Übung.

- i) Gesucht sind Beispiele von Funktionen, die eine oder mehrere der obigen Eigenschaften haben. Zur Illustration sind dabei Skizzen sehr hilfreich.
- ii) Betrachtet sei insbesondere die Funktion auf der rechten Seite von Abbildung 3.4, die auf einem abgeschlossenen Intervall definiert ist und sowohl ihr Supremum als auch ihr Infimum annimmt. Wie ändert sich die Situation, wenn das abgeschlossene Intervall durch ein offenes oder ein halboffenes ersetzt wird? **Muss eine Funktion ihr Supremum bzw. Infimum annehmen?**

### Umkehrfunktion: Explizite Berechnung und geometrische Interpretation.

Im Falle reeller Funktionen kann man die Umkehrfunktion zumindest prinzipiell berechnen. Betrachtet sei eine **bijektive** Funktion  $f: \mathbb{R} \supset A \rightarrow B \subset \mathbb{R}$  und ihre Umkehrfunktion  $f^{-1}: B \rightarrow A$  (vgl. Kapitel 1). Die Gleichung

$$y = f(x)$$

versucht man dann nach  $x$  aufzulösen, d.h. auf die Form

$$x = \dots = f^{-1}(y)$$

zu bringen. Die formale Vertauschung von  $x$  und  $y$  liefert schließlich die gesuchte Abbildungsvorschrift in der üblichen Notation.

**Einfaches Beispiel.** Es sei

$$y = 2x + 4 .$$

Dann ist

$$x = \frac{y}{2} - 2$$

und die Abbildungsvorschrift der Umkehrabbildung lautet in der üblichen Notation

$$y = f^{-1}(x) = \frac{x}{2} - 2 .$$

Probe: In der Tat ist im Beispiel

$$f^{-1}(f(x)) = f^{-1}(2x + 4) = \frac{2x + 4}{2} - 2 = x$$

und analog  $f(f^{-1}(x)) = x$ .

Wie in diesem Beispiel (man skizziere  $y = 2x + 4$  und  $y = (x/2) - 2$  in einem Schaubild) gilt auch allgemein:

Der Graph von  $y = f^{-1}(x)$  entsteht durch Spiegelung von  $y = f(x)$  an der Winkelhalbierenden (des Graphen der Funktion  $y = x$ ) (vgl. Abbildung 3.5).

### 3.4 Spezielle reelle Funktionen

Nach der konstanten Abbildung und der Identität (vgl. Abbildung 1.6) sowie den sogenannten **affin linearen Abbildungen** – der Graph einer affin linearen Abbildung ist eine Gerade -  $f(x) = ax + b$ ,  $a, b \in \mathbb{R}$  fixiert, werden nun weitere “elementare” reelle Funktionen kurz vorgestellt. Man beachte: Die Einschränkung “reelle Funktionen” ist wichtig, da beispielsweise Vektoren nicht potenziert werden können.

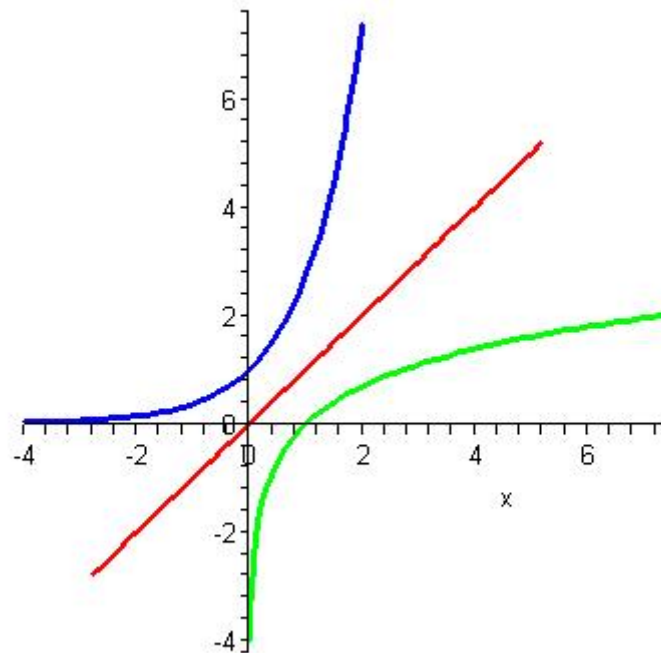


Abbildung 3.5: Zur geometrischen Interpretation der Umkehrfunktion.

**Übung.** Man diskutiere die charakteristischen Eigenschaften reeller Funktionen für die folgenden Beispiele.

### 3.4.1 Potenzfunktionen und enge Verwandte

- i) **Potenzfunktionen** mit **positiven ganzzahligen Exponenten**.
- $f: \mathbb{R} \rightarrow \mathbb{R}_0^+$ ,  $f(x) = x^2$  (vgl. grün dargestellte Funktion in Abbildung 3.6). Die Funktion wächst **quadratisch** und der Graph ist eine **Parabel**. Es existiert genau eine **Nullstelle**, d.h. es existiert genau ein  $x_0$  mit  $f(x_0) = 0$ , nämlich der Punkt  $x_0 = 0$ .
  - Von **kubischem Wachstum** ist die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^3$ , mit ebenfalls genau der einen Nullstelle  $x_0 = 0$  (vgl. rot dargestellte Funktion in Abbildung 3.6).
  - Die Funktion  $f(x) = x^n$ ,  $n \in \mathbb{N}$  fixiert, ist für gerades  $n$  eng verwandt mit Fall (a), für ungerades  $n$  mit Fall (b).

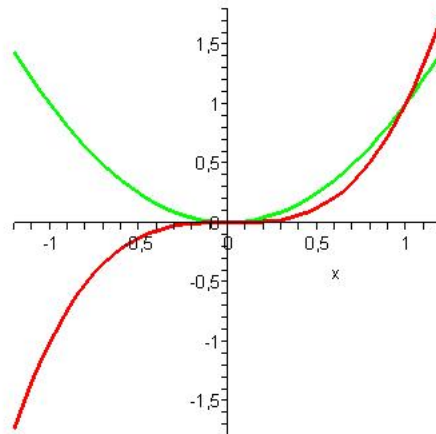


Abbildung 3.6: Die Funktionen  $f(x) = x^2$  und  $f(x) = x^3$ .

ii) **Polynome.**

Eine Funktion  $p: \mathbb{R} \rightarrow \mathbb{R}$  der Form

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad n \in \mathbb{N},$$

für fixierte  $a_i \in \mathbb{R}$ ,  $i = 1, \dots, n$ , heißt Polynom oder **ganzrationale Funktion**. Das Polynom ist **vom Grad  $n$** , falls  $a_n \neq 0$ .

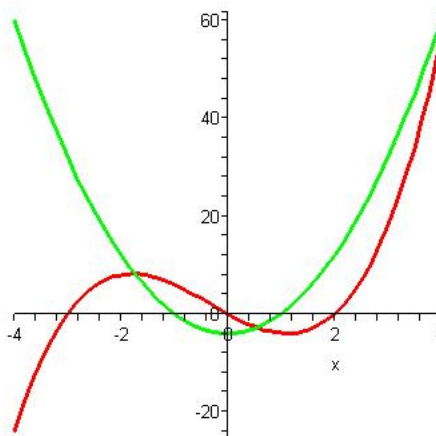


Abbildung 3.7: Ein Polynom **zweiten Grades** und ein Polynom **dritten Grades**.

Potenzfunktionen mit positiven ganzzahligen Exponenten sind als Spezialfall inbegriffen.

Ein Polynom vom Grad  $n$  hat zwischen 0 und **maximal  $n$**  Nullstellen in  $\mathbb{R}$ .

iii) Potenzfunktionen mit **negativen ganzzahligen Exponenten**.

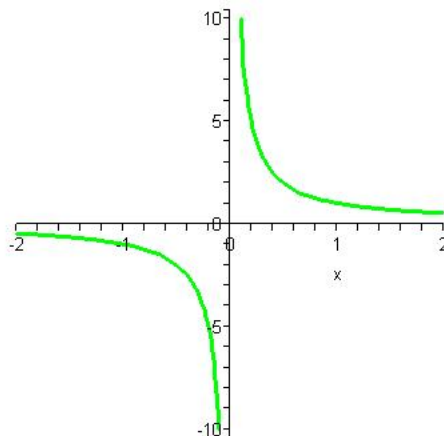


Abbildung 3.8: Die Funktion  $f(x) = 1/x$ .

Eine Potenzfunktion mit negativem ganzzahligen Exponenten  $f: \mathbb{R} - \{0\} \rightarrow \mathbb{R}$  kann als

$$f(x) = x^{-n} := \frac{1}{x^n}, \quad x \neq 0, \quad n \in \mathbb{N},$$

geschrieben werden.

Von besonderer Relevanz sind nun die Nullstellen des Nenners, in denen die Funktion nicht definiert ist und in deren Nähe der Betrag der Funktion beliebig groß wird (siehe Abbildung 3.8).

Diese Punkte heißen **Polstellen** der Funktion  $f$ .

iv) **Gebrochen-rationale Funktionen**.

Eine gebrochen-rationale Funktion ist von der Form

$$f(x) = \frac{p(x)}{q(x)},$$

wobei  $p(x)$  und  $q(x)$  Polynome sind.

Als Definitionsbereich sind die reellen Zahlen **ohne die Nullstellen von  $q$**  zu wählen.

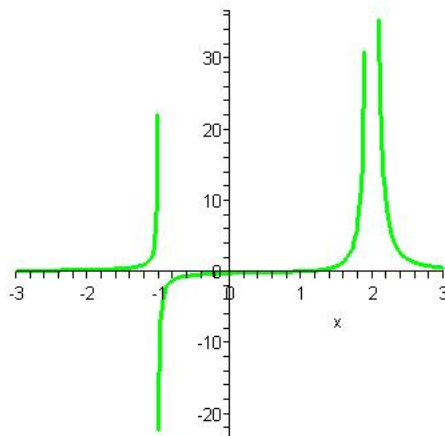


Abbildung 3.9: Eine gebrochen-rationale Funktion.

Nach einer **Polynomdivision** (vgl. Übungen) bleibt als interessantester Fall

$$\text{grad } p < \text{grad } q .$$

Wieder ist den Polstellen besondere Aufmerksamkeit zu widmen.

v) **Wurzelfunktionen.**

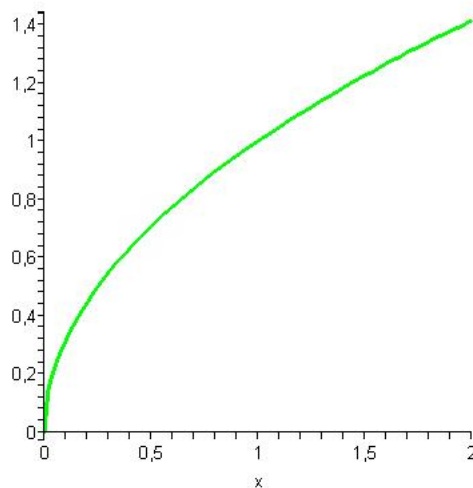


Abbildung 3.10: Die Quadratwurzel.



(a) Die **Quadratwurzel**.

Als Funktion von  $\mathbb{R}_0^+$  nach  $\mathbb{R}_0^+$  ist die Funktion  $f(x) = x^2$  bijektiv (vgl. Kapitel 1). Dementsprechend existiert eine Umkehrfunktion  $f^{-1}: \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ ,

$$f^{-1}(x) =: \sqrt{x} =: x^{\frac{1}{2}}.$$

Die Rechtfertigung für die Schreibweise  $x^{1/2}$  wird in Kürze deutlich.

(b) Die Funktion  $f(x) = x^3$  ist als Funktion von  $\mathbb{R}$  nach  $\mathbb{R}$  bijektiv. Die Umkehrfunktion (**dritte Wurzel**) wird mit

$$f^{-1}(x) =: \sqrt[3]{x} =: x^{\frac{1}{3}}$$

bezeichnet.

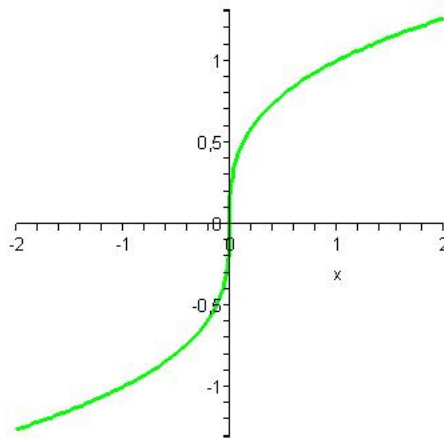


Abbildung 3.11: Die Funktion  $f(x) = \sqrt[3]{x}$ .

(c) Ist  $n \in \mathbb{N}$  fixiert, so wird für gerades  $n$  die Funktion  $\sqrt[n]{x} =: x^{1/n}$  ( $n^{\text{te}}$  Wurzel) analog zu Fall (a) definiert, für ungerades  $n$  dient Fall (b) als Vorlage.

### 3.4.2 Exponentialfunktion und Logarithmus

i) Die **Exponentialfunktion**.

Ist beispielsweise die Zahl  $2^n$  kanonisch definiert,

$$2^n = \underbrace{2 \cdot 2 \cdots 2}_{n \text{ mal}},$$

so ist dennoch völlig unklar, welcher Wert etwa dem Ausdruck  $2^{\sqrt{\pi}}$  zugeordnet werden soll.

Einen Hinweis erhält man aus der folgenden Beobachtung: Sind  $n, m \in \mathbb{N}$  gegeben, so gilt

$$2^{n+m} = 2^n \cdot 2^m .$$

Das Verhalten der Exponentialfunktion  $e^x: \mathbb{R} \rightarrow \mathbb{R}^+$  (vgl. Abbildung 3.12, die Eulersche Zahl  $e = 2.718\dots$  wird später exakt eingeführt) leitet sich im Wesentlichen aus einer analogen sogenannten **Funktionalgleichung** ab: Für alle  $x, y \in \mathbb{R}$  gilt

$$e^{x+y} = e^x \cdot e^y .$$

Zudem gilt  $e^0 = 1$  und wie die Notation suggeriert

$$e^x = \underbrace{e \cdot e \cdots e}_{x \text{ mal}}, \quad \text{falls } x \in \mathbb{N} .$$

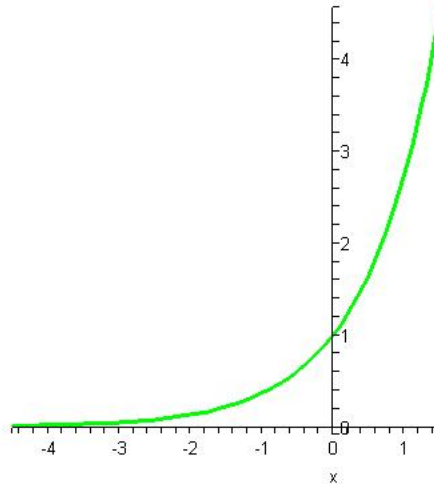


Abbildung 3.12: Die Exponentialfunktion.

Exakt wird die Exponentialfunktion später als Potenzreihe definiert.

ii) Die **allgemeine Exponentialfunktion**.

Für fixiertes  $a \in \mathbb{R}$ ,  $a > 0$ , bezeichnet die Funktion  $a^x$  die allgemeine Exponentialfunktion. Für  $a = 1$  ist die Funktion konstant 1, für  $a \neq 1$  ist eine Fallunterscheidung notwendig.

(a) Der Fall  $a > 1$ .

In diesem Fall verhält sich die Funktion qualitativ wie die Exponentialfunktion, wobei nun

$$a^x = \underbrace{a \cdot a \cdots a}_{x \text{ mal}}, \quad \text{falls } x \in \mathbb{N}.$$

(b) Der Fall  $a < 1$ .

Formal ist die Definition die gleiche wie im Fall (a). Man überlege sich aber, warum das qualitative Verhalten jetzt in Abbildung 3.13 wiedergegeben wird.

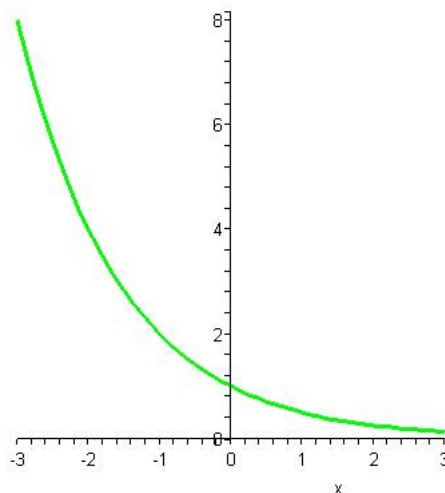


Abbildung 3.13: Beispiel einer allgemeinen Exponentialfunktion für  $0 < a < 1$ .

iii) Der [natürliche Logarithmus](#).

Die Umkehrfunktion der bijektiven Exponentialfunktion  $e^x: \mathbb{R} \rightarrow \mathbb{R}^+$  heißt natürlicher Logarithmus (vgl. Abbildung 3.14 und Abbildung 3.5):

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

Die Rechenregeln für den Logarithmus werden in den Übungen besprochen.

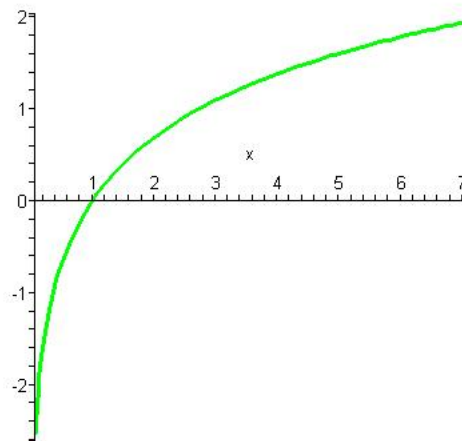
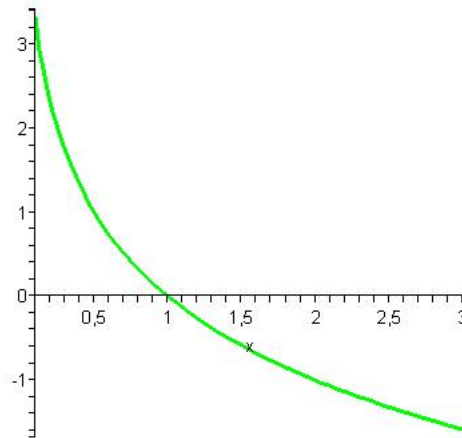


Abbildung 3.14: Der natürliche Logarithmus.

Abbildung 3.15: Beispiel eines Logarithmus zur Basis  $0 < a < 1$ .

*iv)* Der **Logarithmus zur Basis  $a$** .

Die Umkehrfunktion der bijektiven allgemeinen Exponentialfunktion  $a^x: \mathbb{R} \rightarrow \mathbb{R}^+$  heißt Logarithmus zur Basis  $a$  ( $0 < a$ ,  $a \neq 1$ ,  $a \in \mathbb{R}$  fixiert - warum  $a \neq 1$ ?)

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

Wieder ist im Fall  $0 < a < 1$  das qualitative Verhalten anders als im Fall  $a > 1$  (vgl. Abbildung 3.15).

Die Umrechnung zwischen Logarithmen mit unterschiedlichen Ba-

sen  $a > 0$ ,  $b > 0$ ,  $a, b \neq 1$  genügt der Regel

$$\log_b x = \frac{\log_a x}{\log_a b}.$$

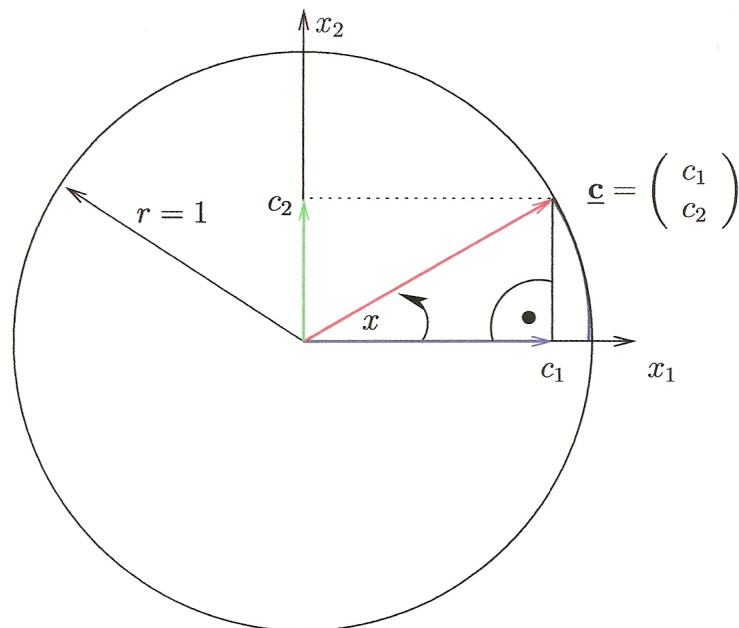


Abbildung 3.16: Zur Definition von Sinus und Cosinus am Einheitskreis.

### 3.4.3 Trigonometrische Funktionen

i) Der **Sinus** und der **Cosinus**.

Elementargeometrisch werden Sinus und Cosinus über den so genannten Einheitskreis definiert (vgl. Abbildung 3.16).

Ein **Einheitsvektor** (Länge 1)  $\underline{c}$  in der Ebene (im  $\mathbb{R}^2$ ), der mit der  $x_1$ -Achse den **orientierten Winkel**  $x$  einschließt, wird wie in der Abbildung angedeutet in die Summe von zwei achsenparallelen orthogonalen Vektoren  $\underline{c}_1$  und  $\underline{c}_2$  zerlegt (vgl. die Einführung in die Vektorrechnung).

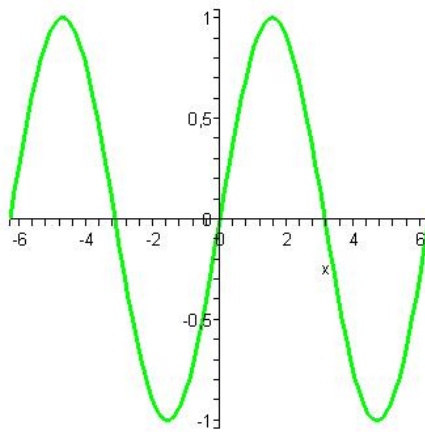


Abbildung 3.17: Der Sinus.

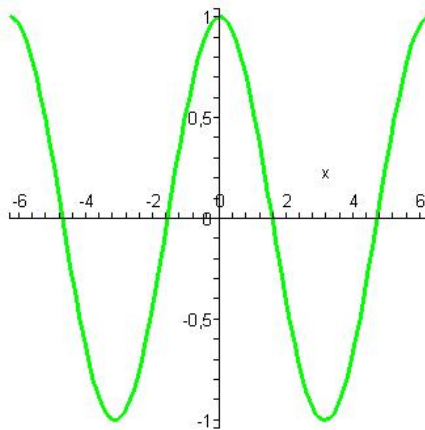


Abbildung 3.18: Der Cosinus.

Werden die Längen von  $\underline{c}_1$ ,  $\underline{c}_2$  mit  $c_1$ ,  $c_2$  bezeichnet, so definiert man

$$\sin(\alpha) := c_2 \quad (\text{“Gegenkathete durch Hypothenuse”})$$

und

$$\cos(\alpha) := c_1 \quad (\text{“Ankathete durch Hypothenuse”}).$$

Dabei wird der Winkel  $x$  in der Regel im **Bogenmaß** gemessen. Das ist die Länge des in Abbildung 3.16 dunkelblau markierten Kreisbogens.

Die Umrechnung vom Bogenmaß  $x$  in die zugehörige Gradzahl  $\alpha$  erfolgt mittels

$$\frac{x}{2\pi} = \frac{\alpha}{360^\circ} .$$

Die Definition impliziert unmittelbar (jeweils für alle  $x \in \mathbb{R}$ )

$$\sin(-x) = -\sin(x) \quad \text{und} \quad \cos(-x) = \cos(x)$$

sowie die  $2\pi$ -Periodizität

$$\sin(x + 2\pi) = \sin(x) , \quad \cos(x + 2\pi) = \cos(x) .$$

Die Eigenschaft

$$\sin^2(x) + \cos^2(x) = 1 \quad \text{für alle } x \in \mathbb{R}$$

folgt ebenfalls unmittelbar aus der Definition.

Für konkrete Rechnungen sind die folgenden Rechenregeln, die so genannten [Additionstheoreme](#), oft sehr hilfreich:

$$\begin{aligned} \sin(x \pm y) &= \sin(x) \cos(y) \pm \cos(x) \sin(y) , \\ \cos(x \pm y) &= \cos(x) \cos(y) \mp \sin(x) \sin(y) , \\ 2 \sin(x) \cos(x) &= \sin(2x) . \end{aligned}$$

Z.B. folgt direkt (wie?)

$$\sin\left(x + \frac{\pi}{2}\right) = \cos(x) .$$

ii) Der [Tangens](#).

Der Tangens ist als Quotient

$$\tan(x) := \frac{\sin(x)}{\cos(x)}$$

außerhalb der Nullstellen des Cosinus definiert, d.h. als Abbildung

$$\tan : \mathbb{R} - \left\{ \frac{\pi}{2} + k\pi : k \in \mathbb{Z} \right\} \rightarrow \mathbb{R} .$$

Es handelt sich um eine ungerade Funktion der Periode  $\pi$ , wie obige Additionstheoreme belegen.

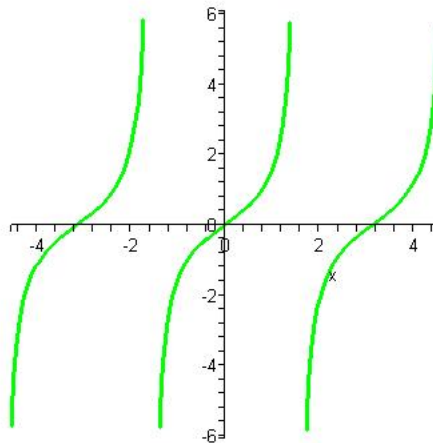


Abbildung 3.19: Der Tangens.

iii) Der **Arcussinus** und der **Arcuscosinus**.

Die Einschränkungen des Sinus als Funktion

$$\sin : \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right] \rightarrow [-1, 1]$$

und des Cosinus als Funktion

$$\cos : [0, \pi] \rightarrow [-1, 1]$$

sind bijektive Funktionen.

Diese Einschränkungen heißen **Hauptzweig des Sinus** bzw. **Hauptzweig des Cosinus**.

Die Umkehrfunktionen

$$\arcsin : [-1, 1] \rightarrow \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right]$$

und

$$\arccos : [-1, 1] \rightarrow [0, \pi]$$

heißen **Arcussinus** bzw. **Arcuscosinus**.

iv) Der **Arcustangens**

Analog ist der Arcustangens als Umkehrfunktion des Hauptzweiges des Tangens ( $\tan : ] -\pi/2, \pi/2[ \rightarrow \mathbb{R}$ ) definiert:

$$\arctan : \mathbb{R} \rightarrow \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[ .$$



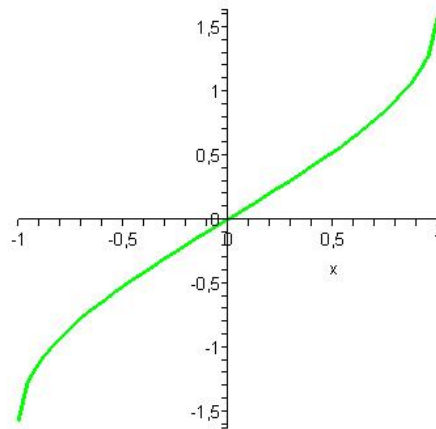


Abbildung 3.20: Der Arcussinus.

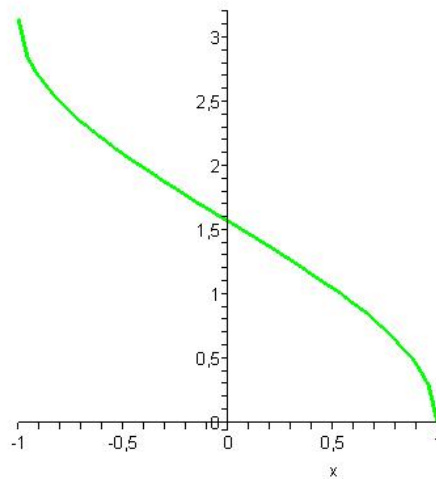


Abbildung 3.21: Der Arcuscosinus.

### 3.4.4 Hyperbolische Funktionen

Die in den Abbildungen 3.23, 3.24 und 3.25 dargestellten hyperbolischen Funktionen

$$\sinh(x) := \frac{e^x - e^{-x}}{2},$$

$$\cosh(x) := \frac{e^x + e^{-x}}{2},$$

$$\tanh(x) := \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

werden in den Übungen diskutiert.

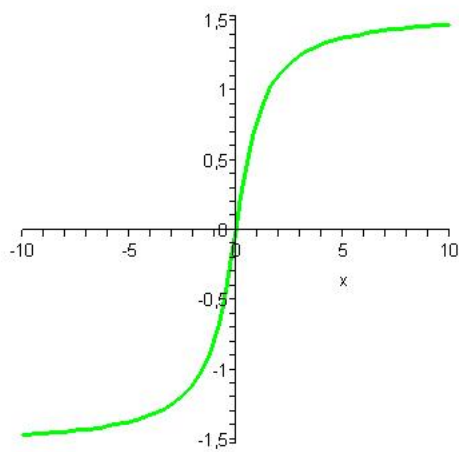


Abbildung 3.22: Der Arcustangens.

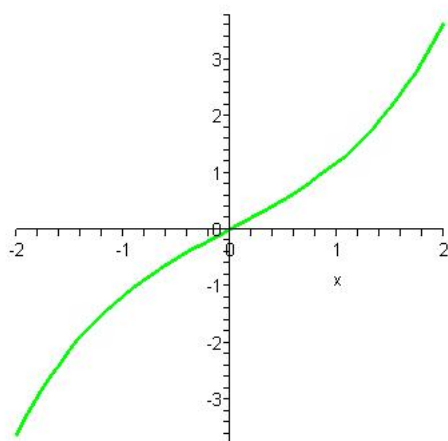


Abbildung 3.23: Der Sinus hyperbolicus.

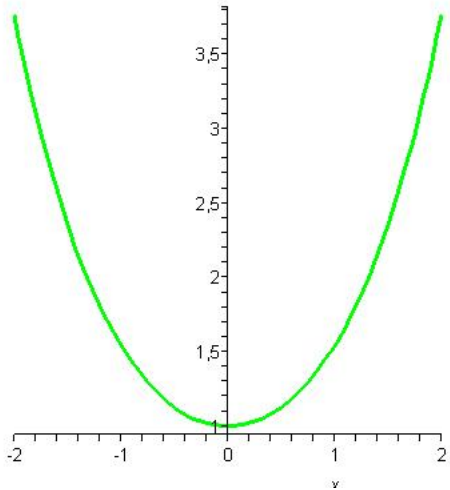


Abbildung 3.24: Der Cosinus hyperbolicus.

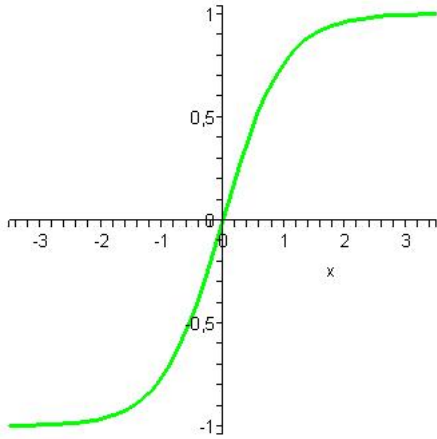


Abbildung 3.25: Der Tangens hyperbolicus.



# Kapitel 4

## Komplexe Zahlen

**Was ist mit der Gleichung  $x^2 = -1$  in  $\mathbb{R}$ ?**

Diese Gleichung ist in der Menge der reellen Zahlen nicht lösbar. Man benötigt die Menge der **komplexen Zahlen**.

Dabei ist zu beachten, dass der eindimensionale Zahlenstrahl keine “Lücken” mehr aufweist und ein neues Konzept zu suchen ist.

**Eröffnet der Übergang zu einer Zahlenebene neue Möglichkeiten?**

Es ist insbesondere **die Multiplikation neu zu interpretieren**, wobei alle **bisherigen Rechenregeln ihre Gültigkeit behalten müssen**.

**Formale Interpretation.**

Eine Paar  $(x, y) \in \mathbb{R}^2$  schreibe man als

$$x + iy, \quad x, y \in \mathbb{R},$$

wobei  $i$  Lösung der Gleichung

$$i^2 = -1,$$

sei. Die geometrische Deutung in der so genannten **Gaußschen Zahlenebene** folgt in Kürze (vgl. Abbildung 4.1).

Dabei heißt  $x$  der **Realteil** und  $y$  der **Imaginärteil** der komplexen Zahl  $z = x + iy$ :

$$x =: \operatorname{Re} z, \quad y =: \operatorname{Im} z .$$

Der Realteil und der Imaginärteil einer komplexen Zahl sind wie bei einem geordneten Paar unabhängig voneinander zu betrachten, d.h.

$$x + iy = u + iv \quad \Leftrightarrow \quad x = u \quad \text{und} \quad y = v .$$

Als Rechenregeln ergeben sich aus der formalen Definition:

$$\begin{aligned} (x + iy) + (u + iv) &= (x + u) + i(y + v) , \\ (x + iy) \cdot (u + iv) &= (xu - yv) + i(xv + yu) . \end{aligned}$$

**Beispiel.** Es ist etwa

$$\frac{1 + i}{2 - i} = \frac{(1 + i)(2 + i)}{(2 - i)(2 + i)} = \frac{1}{5}(2 + 3i + i^2) = \frac{1}{5} + \frac{3}{5}i .$$

In diesem Sinne ist die Menge der komplexen Zahlen definiert als

$$\mathbb{C} := \{x + iy : x, y \in \mathbb{R}\} .$$

**Bemerkungen.**

i) Auch die komplexen Zahlen haben eine Körperstruktur. Insbesondere existiert für  $z \neq 0$  ein Inverses bzgl. der Multiplikation (vgl. Übungen). Dies hat weitreichende Konsequenzen und ist Grundlage der sogenannten **Funktionentheorie**.

ii) In  $\mathbb{C}$  gibt es **keine Anordnungsrelation**, d.h. man kann nicht sagen, ob eine komplexe Zahl kleiner als eine andere ist. Lediglich **Beträge von komplexen Zahlen können verglichen werden**.

**Reelle Zahlen versus komplexe Zahlen.**

Die reellen Zahlen sind eine Teilmenge der komplexen Zahlen,

$$\mathbb{R} = \{z \in \mathbb{C} : \operatorname{Im} z = 0\} .$$

Alle Operationen mit reellen Zahlen übertragen sich mit dieser Interpretation.

## Geometrische Veranschaulichung.

Wie bereits angedeutet können die komplexen Zahlen in der Gaußschen Zahlenebene veranschaulicht werden (vgl. Abbildung 4.1), wobei die Menge der reellen Zahlen der “horizontalen Achse” entspricht.

In dieser Veranschaulichung beachte man die Entsprechungen

$$1 \leftrightarrow (1, 0), \quad i \leftrightarrow (0, 1), \quad -1 \leftrightarrow (-1, 0), \quad -i \leftrightarrow (0, -1).$$

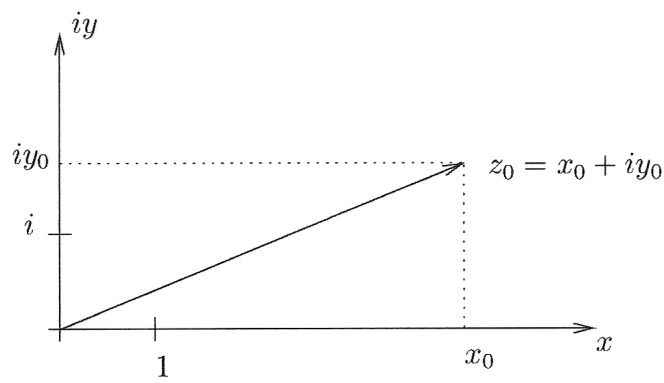


Abbildung 4.1: Die Gaußsche Zahlenebene.

Die Addition zweier komplexer Zahlen wird geometrisch in einem Parallelogramm realisiert (siehe Abbildung 4.2).

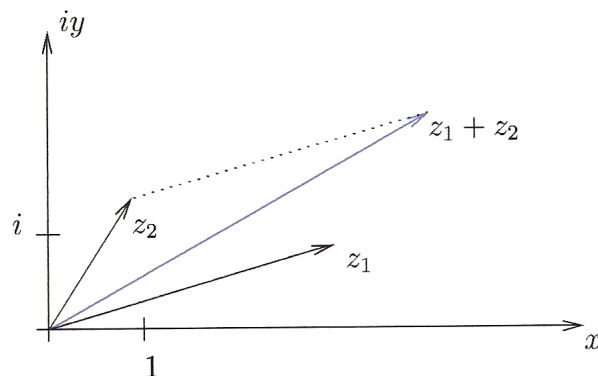


Abbildung 4.2: Zur Addition komplexer Zahlen.

## Konjugation und Betrag.

Eine wichtige Operation ist die **Konjugation** komplexer Zahlen, die geometrisch durch eine **Spiegelung an der reellen Achse** realisiert wird (vgl. Abbildung 4.3).

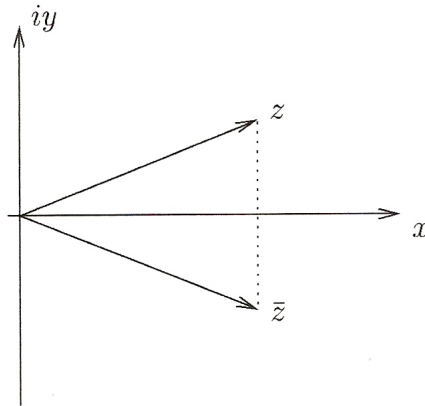


Abbildung 4.3: Zur Konjugation komplexer Zahlen.

Ist  $z = x + iy \in \mathbb{C}$ , so heißt

$$\bar{z} := x - iy$$

die zu  $z$  **konjugiert komplexe Zahl**.

Der **Betrag** einer komplexen Zahl  $z = x + iy$  ist definiert als

$$|z| := \sqrt{x^2 + y^2}.$$

Die Rechenregeln

$$\begin{aligned} \overline{\bar{z}} &= z, & \overline{z+w} &= \bar{z} + \bar{w}, & \overline{z \cdot w} &= \bar{z} \cdot \bar{w}, \\ \operatorname{Re} z &= \frac{1}{2}(z + \bar{z}), & \operatorname{Im} z &= \frac{1}{2i}(z - \bar{z}), \\ z \in \mathbb{R} &\Leftrightarrow z = \bar{z}, \\ |z| &= \sqrt{z\bar{z}}, & \text{d.h. } z\bar{z} &= x^2 + y^2 \quad (\text{insbesondere gilt } z\bar{z} \in \mathbb{R}_0^+), \\ |z| &\geq 0, & |z| = 0 &\Leftrightarrow z = 0, \\ |z_1 z_2| &= |z_1| |z_2|, \\ |z_1 + z_2| &\leq |z_1| + |z_2| \quad (\text{Dreiecksungleichung}) \end{aligned}$$



werden in den Übungen diskutiert.

### Polarkoordinaten.

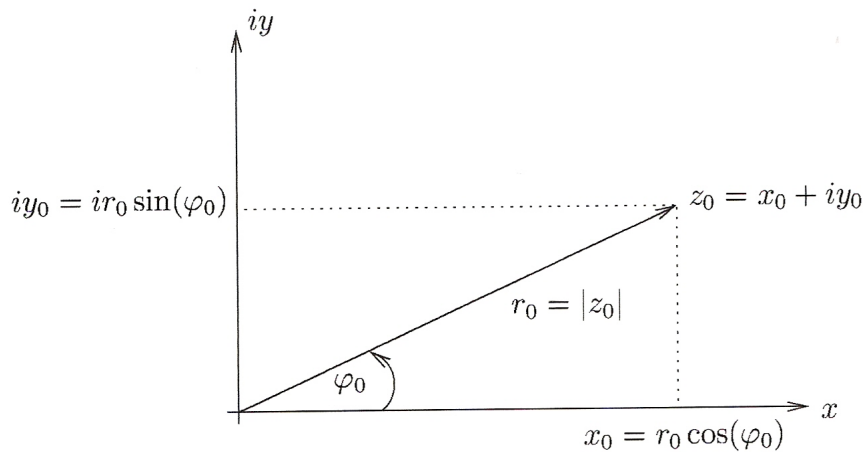


Abbildung 4.4: Komplexe Zahlen in Polarkoordinaten.

Genau wie bei der Definition der trigonometrischen Funktionen mithilfe des Einheitskreises können komplexe Zahlen in Polarkoordinaten dargestellt werden (vgl. Abbildung 4.4):

Schließt eine komplexe Zahl  $z = x + iy$  in der Gaußschen Zahlenebene den Winkel  $\varphi$  mit der reellen Achse ein, so heißt  $\varphi$  ein **Argument** von  $z$ . Notation:

$$\varphi = \arg z .$$

**Bemerkung.** Wegen der  $2\pi$ -Periodizität des Sinus und des Cosinus ist ein **Argument nur bis auf ganzzahlige Vielfache von  $2\pi$  bestimmt**.

Beispielsweise ist für jedes  $k \in \mathbb{Z}$

$$\frac{\pi}{2} + k2\pi = \arg i .$$

Bezeichnet  $r$  den Betrag von  $z$ ,

$$r = |z| = \sqrt{x^2 + y^2} ,$$

so folgt

$$x = r \cos(\varphi) , \quad y = r \sin(\varphi) , \quad \text{d.h. } z = r(\cos(\varphi) + i \sin(\varphi)) .$$

Anhand der [Eulerschen Formel](#)

$$e^{i\varphi} = \cos(\varphi) + i \sin(\varphi) ,$$

die an dieser Stelle nur formal benutzt wird und die später näher erläutert wird, erkennt man die [Exponentialdarstellung](#) einer komplexen Zahl  $z$ :

$$z = r e^{i\varphi} .$$

### Geometrische Deutung der Multiplikation.

Es seien

$$z_1 = x_1 + iy_1 = r_1 e^{i\varphi_1} \quad \text{und} \quad z_2 = x_2 + iy_2 = r_2 e^{i\varphi_2} .$$

Dann ist

$$z_1 \cdot z_2 = r_1 r_2 e^{i\varphi_1} e^{i\varphi_2} = r_1 r_2 e^{i(\varphi_1 + \varphi_2)} .$$

Bei der Multiplikation werden **Beträge multipliziert** und **Argumente addiert**, wie es in [Abbildung 4.5](#) angedeutet ist.

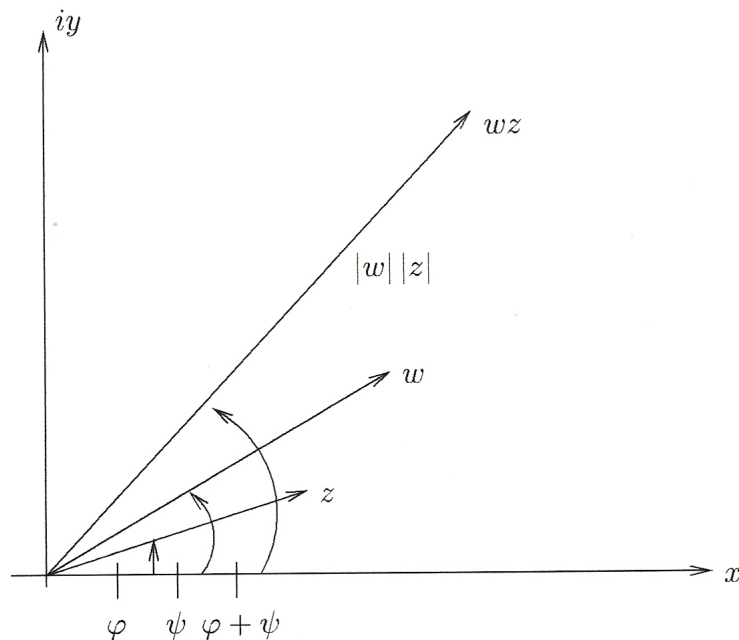


Abbildung 4.5: Zur Multiplikation in der Gaußschen Zahlenebene.

**Fundamentalsatz der Algebra.**

**Definition 4.1.** Es sei  $n \in \mathbb{N}$  und

$$p_n(z) := a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

ein Polynom  $n^{\text{ten}}$  Grades ( $a_n \neq 0$ ) mit komplexen Koeffizienten  $a_k \in \mathbb{C}$ ,  $k = 1, \dots, n$ . Dann gibt es **mindestens eine Nullstelle**  $\lambda \in \mathbb{C}$  von  $p_n$ , d.h. ein  $\lambda \in \mathbb{C}$  mit  $p_n(\lambda) = 0$ .

**Beispiel.**

$$z^4 - 1 = (z - 1) \cdot (z + 1) \cdot (z - i) \cdot (z + i).$$

**Korollar 4.1.** Jedes Polynom  $n^{\text{ten}}$  Grades wie oben lässt sich als Produkt von  $n$  **Linearfaktoren** schreiben. Es gibt also Nullstellen  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$  (nicht notwendig verschieden), sodass für alle  $z \in \mathbb{C}$

$$p_n(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n),$$

wobei  $a_n = 1$  angenommen werden kann (sonst teile man das Polynom zunächst durch  $a_n$ ).



# Kapitel 5

## Folgen und Reihen

**Beispiel.** (vgl. “Achill und die Schildkröte”)

$A$  bewege sich mit einer Geschwindigkeit von  $2m/s$ .

In einer Entfernung von  $10m$  vor sich sieht er  $B$ , der sich mit  $1m/s$  nach vorne bewegt.

Bezeichnet  $t_0$  die Zeit des Überholvorgangs und ist  $s_0$  die bis dahin von  $A$  zurückgelegte Strecke, so gilt nach der Gesetzmäßigkeit “Weg = Geschwindigkeit  $\cdot$  Zeit”

$$t_0 = \frac{s_0}{2m/s} = \frac{s_0 - 10m}{1m/s}.$$

Dementsprechend hat  $A$  nach  $20m$  den Vorsprung von  $B$  aufgeholt.

**Aber.** Wenn  $A$  die  $10m$  zurückgelegt hat, die  $B$  Vorsprung hatte, dann ist  $B$  um  $5m$  weitergekommen und hat wiederum Vorsprung.

Nachdem  $A$  den Vorsprung von  $5m$  aufgeholt hat, ist  $B$  um  $2.5m$  weitergekommen und hat immer noch Vorsprung.

Nachdem  $A$  den Vorsprung von  $2.5m$  aufgeholt hat, ...

Also kann  $A$  niemals  $B$  überholen?

**Idee zur Auflösung.**  $B$  hat zunächst  $10m$  Vorsprung, dann  $5m$ , dann  $2.5m$ , dann ..., insgesamt also

$$10m + 5m + 2.5m + 1.25m + \dots = 10m \left( 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots \right).$$

In den Übungen wird mit vollständiger Induktion gezeigt

$$\begin{aligned} s_n &:= \sum_{k=0}^n \left(\frac{1}{2}\right)^k = 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n} \\ &= \frac{1 - \left(\frac{1}{2}\right)^{n+1}}{1 - \frac{1}{2}} \end{aligned}$$

und als **Grenzwert der Folge (Reihe)**  $\{s_n\}$  findet man 2 (vgl. die nachfolgenden Ausführungen).

Für das Beispiel bedeutet das

$$10m + 5m + 2.5m + 1.25m + \cdots = 20m .$$

Tatsächlich gibt es also keinen Widerspruch. Vielmehr wird der Überholvorgang mithilfe der so genannten **geometrischen Reihe** berechnet.

## 5.1 Folgen

### Was genau ist eine Folge?

**Definition 5.1.** *Es sei  $A \neq \emptyset$  eine Menge. Eine **Folge** in  $A$  ist eine Abbildung  $\mathbb{N} \rightarrow A$ , die jedem  $n \in \mathbb{N}$  ein Element  $a_n \in A$ , das  $n^{\text{te}}$  **Glied der Folge**, zuordnet.*

*Notation:  $\{a_n\}$  oder aufzählend  $a_1, a_2, a_3, \dots$  oder  $a_n = \dots$  oder  $\{a_n\}_{n \in \mathbb{N}}$  oder  $(a_n)_{n \in \mathbb{N}}$  ...*

*Ist  $A = \mathbb{R}$  bzw.  $A = \mathbb{C}$ , so spricht man von einer **reellen bzw. komplexen Zahlenfolge**.*

**Bemerkung.** Statt  $\mathbb{N}$  kann eine andere **abzählbar unendliche Indexmenge** gewählt werden.

**Beispiele.**

i)  $f: \mathbb{N} \rightarrow \mathbb{R}, f(n) = 1/n$ . Dann ist  $\{a_n\} = \{1/n\}$ ;

ii)  $f: \mathbb{N} \rightarrow \mathbb{R}, f(n) = n$ . Dann ist  $\{a_n\} = \{n\}$ ;

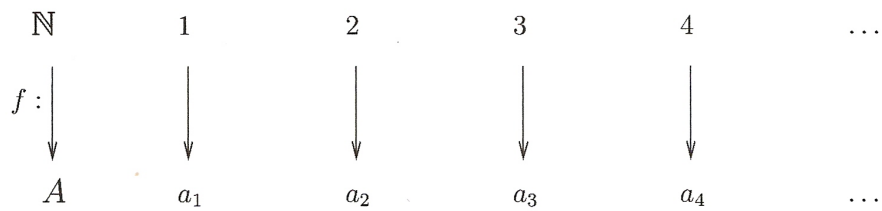


Abbildung 5.1: Eine Folge ist eine Abbildung.

iii)  $f: \mathbb{N} \rightarrow \mathbb{R}$ ,  $f(n) = a$ ,  $a \in \mathbb{R}$  fixiert. Dann ist  $\{a_n\} = \{a\}$ , **konstante Folge**.

iv)  $f: \mathbb{N} \rightarrow \mathbb{R}$ ,  $f(n) = (-1)^{n+1}$ . Hier handelt es sich um eine so genannte **alternierende Folge**, d.h. das Vorzeichen wechselt:  $1, -1, 1, -1, \dots$ .

### Folgen in anderen Mengen.

Von besonderer Bedeutung sind so genannte **Funktionenfolgen**, d.h.  $A$  ist eine **Menge von Funktionen**.

**Beispiel.** Es sei

$$A = \{p : p \text{ ist ein reelles Polynom}\}.$$

Dann ist die Abbildung

$$f: \mathbb{N} \rightarrow A, \quad f(n) = x^n$$

eine Funktionenfolge,  $\{a_n\} = \{x^n\}$ .

### Rekursiv definierte Folgen.

Folgen können auch rekursiv definiert sein, d.h. ein Folgenglied berechnet sich aus dem Vorgänger oder aus mehreren vorherigen Folgengliedern.

**Beispiel.** (**Fibonaccizahlen**, eine Folgenglied ist die Summe der beiden vorherigen))

**Rekursionsanfang:**  $a_0 := 0$ ,  $a_1 := 1$ .

**Rekursionsvorschrift:**  $a_n := a_{n-1} + a_{n-2}$  für  $n \geq 2$ .

Die Folge lautet: 0, 1, 1, 2, 3, 5, 8, ...

(Ein Kaninchenpaar befinde sich in einem geschlossenen Gehege. Wie viele Kaninchenpaare gibt es nach einem Jahr, wenn jedes Paar jeden Monat zwei Junge zeugt, welche sich ihrerseits vom zweiten Monat an vermehren?)

Im weiteren Verlauf dieses Kapitels werden stets reelle Zahlenfolgen betrachtet. Komplexe Zahlenfolgen werden völlig analog behandelt, solange die Aussagen kein “<” enthalten, da die Menge der komplexen Zahlen nicht angeordnet ist.

### Anschauliche Vorstellung einer konvergenten Folge.

Zur Veranschaulichung betrachte man die Folge  $\{1/n\}$ , die nach der intuitiven Vorstellung gegen 0 konvergiert. Das bedeutet, dass die Folgenglieder  $a_n$  der 0 beliebig nahe kommen.

Befindet man sich auf dem Zahlenstrahl also in **irgendeinem festen Abstand**  $\varepsilon > 0$  jenseits der 0, dann werden beim Durchlaufen der Folge die Folgenglieder irgendwann (d.h. ab einem gewissen Index abhängig von  $\varepsilon$ ) näher bei der 0 sein als man selbst, wie es in Abbildung 5.2 skizziert ist.

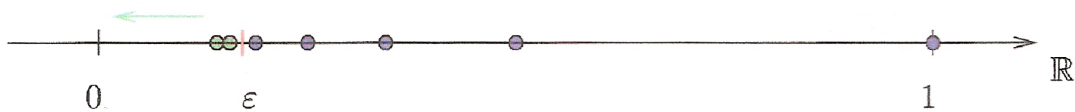


Abbildung 5.2: Bei diesem  $\varepsilon$  sind ab  $n = 6$  alle Folgenglieder näher bei 0 als  $\varepsilon$ .

### Konvergenz reeller Zahlenfolgen.

**Definition 5.2.** Eine Folge  $\{a_n\}$  heißt konvergent, wenn es ein  $a \in \mathbb{R}$  gibt mit der Eigenschaft:



*Zu jedem (noch so kleinen)  $\varepsilon > 0$  existiert eine (meist von  $\varepsilon$  abhängende) Zahl  $N = N(\varepsilon) \in \mathbb{N}$  mit*

$$|a_n - a| < \varepsilon \quad \text{für alle } n \geq N(\varepsilon).$$

Dann heißt  $a$  der **Grenzwert** oder **Limes** der Folge. Notation:

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{oder} \quad a_n \rightarrow a \quad \text{für } n \rightarrow \infty.$$

Eine Folge, die nicht konvergiert, heißt **divergent**.

**Bemerkung..** Aus der Definition der Konvergenz folgt mit einer kurzen Rechnung unmittelbar, dass (falls existent) **der Grenzwert eindeutig bestimmt ist**, d.h. eine Folge kann nicht zwei verschiedene Grenzwerte haben.

### Standardbeispiel einer konvergenten Folge.

Betrachtet sei die Folge  $\{1/n\}$  mit dem **vermuteten Grenzwert  $a = 0$** . Zu gegebenem  $\varepsilon > 0$  (man stelle sich  $\varepsilon$  sehr klein vor) wähle man eine natürliche Zahl  $N(\varepsilon)$ , die der Bedingung

$$N(\varepsilon) > \frac{1}{\varepsilon}$$

genügt, d.h.  $N(\varepsilon)$  ist hinreichend groß zu wählen. Für alle  $n \geq N(\varepsilon)$  folgt:

$$|1/n - 0| = \frac{1}{n} \leq \frac{1}{N(\varepsilon)} < \frac{1}{\varepsilon^{-1}} = \varepsilon$$

und nach Definition 5.2 konvergiert die Folge tatsächlich gegen 0.

**Bemerkung.** Eine Folge, die gegen 0 konvergiert, heißt **Nullfolge**.

### Standardbeispiel einer divergenten Folge.

Betrachtet sei die Folge  $\{n\}$ , von der keine Konvergenz erwartet wird. Divergenz heißt (man ersetze unter Beibehaltung der Reihenfolge den Quantor “ $\forall$ ” durch “ $\exists$ ” und vice versa):

Für alle  $a \in \mathbb{R}$  existiert ein  $\varepsilon = \varepsilon(a) > 0$ , sodass für alle  $N \in \mathbb{N}$  ein  $n = n(N) \geq N$  existiert mit

$$|a_n - a| \geq \varepsilon .$$

Für die Folge  $\{n\}$  kann für beliebiges  $a \in \mathbb{R}$  z.B.  $\varepsilon = 1$  gewählt werden. Dann existiert für alle  $N \in \mathbb{N}$  ein  $n \geq N$ , z.B.

$$n > \max\{N, a + 1\} ,$$

mit

$$|a_n - a| = |n - a| \geq 1 = \varepsilon .$$

Tatsächlich divergiert die Folge.

### Bemerkungen.

- i) Die obigen Beispiele zeigen: Eine erste Idee, ob die Folge konvergiert (und gegen welchen Grenzwert) oder ob sie divergiert, ist sehr hilfreich für die konkrete Rechnung. Notwendig ist sie aber nicht.
- ii) Die gleichen Argumente, mit denen die Divergenz von  $\{n\}$  gezeigt wird, belegen:  
 Eine **konvergente Folge muss beschränkt sein**. (Was bedeutet Beschränktheit für Folgen?)  
 Oder mit anderen Worten: **Ist eine Folge unbeschränkt, so muss sie divergieren**.
- iii) **Vorsicht**: Eine beschränkte Folge muss nicht konvergieren (Beispiel?).

### Rechenregeln für Folgen.

Der folgende Satz über Rechenregeln für Folgen ist leicht zu beweisen. Dennoch ist er von großer praktischer Bedeutung, da aus bekannten Grenzwerten auf weitere Grenzwerte geschlossen werden kann, ohne die Konvergenz explizit über die Definition nachweisen zu müssen, was oft nicht so einfach ist.

**Satz 5.1.** Es seien  $\{a_n\}$  und  $\{b_n\}$  zwei *konvergente* Folgen mit den Grenzwerten  $a$  bzw.  $b \in \mathbb{R}$ . Dann gilt:

i) Ist  $a_n \leq b_n$  für alle  $n \in \mathbb{N}$ , so ist  $a \leq b$ .

ii) Die Folgen  $\{|a_n|\}$ ,  $\{k \cdot a_n\}$  für eine Konstante  $k \in \mathbb{R}$ ,  $\{a_n \pm b_n\}$  und  $\{a_n \cdot b_n\}$  konvergieren mit

$$\begin{aligned}\lim_{n \rightarrow \infty} |a_n| &= |a|, \\ \lim_{n \rightarrow \infty} k \cdot a_n &= k \cdot a, \\ \lim_{n \rightarrow \infty} (a_n \pm b_n) &= a \pm b, \\ \lim_{n \rightarrow \infty} a_n \cdot b_n &= a \cdot b.\end{aligned}$$

Ist  $b \neq 0$  und gilt  $b_n \neq 0$  für alle  $n \in \mathbb{N}$ , so konvergiert auch die Folge  $\{a_n/b_n\}$  mit

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{a}{b}.$$

iii) (*Einschließungskriterium*) Ist  $a = b$  und gilt  $a_n \leq c_n \leq b_n$  für alle  $n \in \mathbb{N}$ , wobei  $\{c_n\}$  eine weitere Folge bezeichne, so konvergiert auch  $\{c_n\}$  mit Grenzwert  $c = a = b$ .

**Bemerkung.** Die Aussagen i) und iii) bleiben richtig, wenn die Voraussetzungen nur für alle  $n \geq n_0$ ,  $n_0 \in \mathbb{N}$  fixiert, erfüllt sind.

**Beispiel.** Man betrachte die Folge  $\{c_n\}$  mit

$$c_n = \frac{2n^2 - 3n + 1}{3n^2 + 15n + 4} \quad \text{für alle } n \in \mathbb{N},$$

bzw. umgeformt

$$c_n = \frac{2 - \frac{3}{n} + \frac{1}{n^2}}{3 + \frac{15}{n} + \frac{4}{n^2}} =: \frac{a_n}{b_n}.$$

$\{a_n\}$  und  $\{b_n\}$  sind nach obigen Rechenregeln jeweils die Summe aus drei konvergenten Folgen,

$$a_n \rightarrow 2 \quad \text{und} \quad b_n \rightarrow 3 \quad \text{für } n \rightarrow \infty.$$

Wieder greift man auf die Rechenregeln zurück und erkennt

$$c_n \xrightarrow{n \rightarrow \infty} \frac{2}{3}.$$

Eine Reihe analoger Beispiele wird in den Übungen besprochen.

### Beispiele wichtiger Folgen.

- i) Die Folge  $\{1/n^\alpha\}$ ,  $\alpha > 0$  fixiert, ist eine Nullfolge.
- ii) Die Folge  $\{n^\alpha\}$ ,  $\alpha > 0$  fixiert, divergiert.
- iii) Die alternierende Folge  $1, -1, 1, \dots$  kann nicht konvergieren (warum?)
- iv) Ein Kapital  $K_0$  sei zu einem Jahreszinsfuß  $p$  ein Jahr lang verzinst. Nach diesem Jahr beläuft sich das Kapital auf

$$K_1 = K_0(1 + p) \quad \text{bei jährlicher Verzinsung;}$$

$$K_2 = K_0 \left(1 + \frac{p}{2}\right)^2 \quad \text{bei halbjährlicher Verzinsung;}$$

$$K_4 = K_0 \left(1 + \frac{p}{4}\right)^4 \quad \text{bei vierteljährlicher Verzinsung.}$$

Das Beispiel führt auf die Folge

$$\{a_n\}, \quad a_n := \left(1 + \frac{1}{n}\right)^n \quad \text{für alle } n \in \mathbb{N},$$

mit

$$a_n \xrightarrow{n \rightarrow \infty} e = 2.71828182845 \dots,$$

wobei die **Eulersche Zahl**  $e$  als Grenzwert der Folge  $\{a_n\}$  definiert ist und die Basis der bereits skizzierten Exponentialfunktion ist.

- v) Ist  $x \in \mathbb{R}$  fixiert, so gilt für die Folge  $\{x^n\}$ :
  - (a) Für  $|x| < 1$  ist die Folge **konvergent** mit  $\lim_{n \rightarrow \infty} x^n = 0$ .
  - (b) Für  $x = 1$  ist die Folge **konvergent** mit  $\lim_{n \rightarrow \infty} x^n = 1$ .
  - (c) Für  $|x| > 1$  und für  $x = -1$  ist die Folge **divergent**.
- vi) Es sei  $x \in \mathbb{R}$  fixiert. Die für alle  $n \in \mathbb{N}$  durch

$$a_n := \frac{x^n}{n!}$$

definierte Folge ist eine **Nullfolge**.

- vii) (a) Für fixiertes  $a > 0$  ist  $\lim_{n \rightarrow \infty} \sqrt[n]{a} = 1$ .  
 (b)  $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$ .

### Weitere Konvergenzkriterien.

Ein Konvergenzkriterium ist das bereits angesprochene Einschließungskriterium.

Für das zweite Kriterium benötigt man:

- i) Eine reelle Zahlenfolge  $\{a_n\}$  heißt **monoton wachsend**, falls  $a_n \leq a_{n+1}$  für alle  $n \in \mathbb{N}$ .

Gilt in der Ungleichung sogar das strikte “<”, so heißt die Folge **streng monoton wachsend**.

- ii) Eine reelle Zahlenfolge  $\{a_n\}$  heißt **monoton fallend**, falls  $a_n \geq a_{n+1}$  für alle  $n \in \mathbb{N}$ .

Gilt in der Ungleichung sogar das strikte “>”, so heißt die Folge **streng monoton fallend**.

#### Satz 5.2. (Satz der monotonen Folge)

Eine **monoton wachsende** reelle Zahlenfolge ist **genau dann** konvergent, wenn sie **nach oben beschränkt** ist.

Eine **monoton fallende** reelle Zahlenfolge ist **genau dann** konvergent, wenn sie **nach unten beschränkt** ist.

**Bemerkung.** Mit dem Satz von der monotonen Folge zeigt man beispielsweise die Konvergenz von  $\{(1 + 1/n)^n\}$ .

#### Satz 5.3. (Kriterium von Cauchy)

Eine reelle Zahlenfolge  $\{a_n\}$  ist **genau dann konvergent**, wenn sie eine **Cauchy-Folge** ist, d.h.:

Zu jedem  $\varepsilon > 0$  existiert ein  $N(\varepsilon) \in \mathbb{N}$  mit

$$|a_n - a_m| < \varepsilon \quad \text{für alle } n, m \geq N(\varepsilon) .$$

## Teilfolgen und Häufungspunkte.

**Beispiel.** Wie bereits gesehen konvergiert die Folge  $1, -1, 1, \dots$  nicht. Trotzdem sind  $1$  und  $-1$  sicherlich charakteristisch für die Folge: Es handelt sich hier um so genannte **Häufungspunkte** der Folge.

**Definition 5.3.** Eine Folge  $\{\tilde{a}_j\}$  heißt **Teilfolge** der Folge  $\{a_n\}$ , wenn es eine streng monoton wachsende Folge  $\{n_j\}$  natürlicher Zahlen gibt, sodass

$$\tilde{a}_j = a_{n_j} \quad \text{für alle } j \in \mathbb{N}.$$

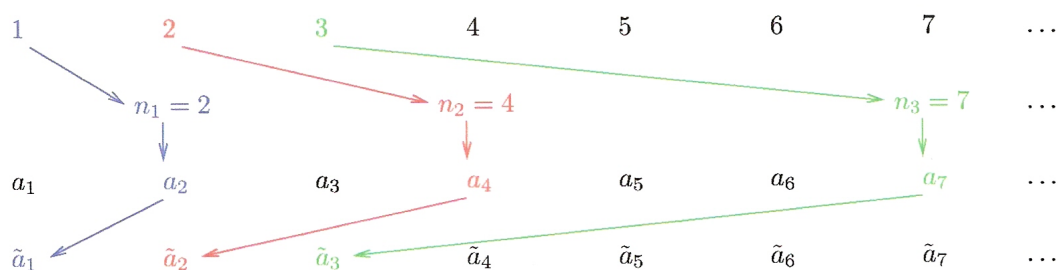


Abbildung 5.3: Zum Begriff der Teilfolge.

**Beispiel.** Man betrachte die obige Folge  $1, -1, 1, \dots$  und die Folge  $\{n_j\}$  mit  $n_1 = 1, n_2 = 3, n_3 = 5, \dots$ . Als Teilfolge erhält man die konstante Folge  $\{1\}$  mit Grenzwert  $1$ .

Analoges gilt für alle **beschränkten** Folgen im Sinne von

**Satz 5.4.** (*Satz von Bolzano Weierstraß*)

Jede **beschränkte** reelle Zahlenfolge enthält eine **konvergente Teilfolge**.

Wie oben angedeutet sind die Grenzwerte von Teilfolgen charakteristisch für eine Folge:

**Definition 5.4.** Es sei  $\{a_n\}$  eine reelle Zahlenfolge. Dann heißen die Grenzwerte konvergenter Teilfolgen von  $\{a_n\}$  **Häufungspunkte** der Folge.

Der *größte Häufungspunkt* heißt *limes superior*, Notation:

$$\limsup_{n \rightarrow \infty} a_n .$$

Der *kleinste Häufungspunkt* heißt *limes inferior*, Notation:

$$\liminf_{n \rightarrow \infty} a_n .$$

## 5.2 Reihen

**Reihen** sind eine *spezielle Art von Folgen*. Das erste und wichtigste Beispiel (*geometrische Reihe*) ist als “Achill und die Schildkröte” zu Beginn dieses Kapitels skizziert.

Wieder werden nur *reelle Zahlenreihen* diskutiert. Alle Aussagen gelten auch im Komplexen, sofern sie kein “<” enthalten.

### Definition 5.5.

i) Ist  $\{a_n\}$  eine reelle Zahlenfolge, so heißt die Folge  $\{s_k\}$ ,

$$s_k := \sum_{n=1}^k a_n ,$$

eine (unendliche) Reihe,  $s_k$  heißt die  *$k^{\text{te}}$  Partialsumme*.

ii) Eine Reihe heißt *konvergent*, wenn die Folge der Partialsummen gegen einen Grenzwert  $s \in \mathbb{R}$  konvergiert.

Konvergiert die Reihe nicht, so heißt sie *divergent*.

**Notation.** Die Symbole

$$a_1 + a_2 + a_3 + \dots \quad \text{und} \quad \sum_{n=1}^{\infty} a_n$$

werden in *doppelter Bedeutung* gebraucht.

Einerseits bezeichnen sie die Reihe, d.h. die *Folge der Partialsummen*.

Konvergiert die Reihe gegen  $s$ , so stehen sie andererseits für den **Grenzwert**

$$\sum_{n=1}^{\infty} a_n = s = \lim_{k \rightarrow \infty} s_k .$$

### Bemerkungen.

- i) Eine Reihe ist **keine unendliche Summe** (Was soll das auch sein?), sondern eine **reelle Zahlenfolge von Partialsummen**.
- ii) Der Startindex muss nicht 1 sein. Völlig analog ist für  $m \in \mathbb{Z}$

$$\sum_{n=m}^{\infty} a_n$$

definiert.

- iii) Wie beim Summenzeichen ist die Bezeichnung des Summationsindex egal.

### Eine Rechenregel.

Die folgende Rechenregel für Zahlenreihen ist elementar zu beweisen:

**Satz 5.5.** *Es seien  $\sum_{n=1}^{\infty} a_n$  und  $\sum_{n=1}^{\infty} b_n$  **konvergente** reelle Zahlenreihen. Dann ist auch für alle  $\alpha, \beta \in \mathbb{R}$  die Reihe*

$$\sum_{n=1}^{\infty} (\alpha \cdot a_n + \beta \cdot b_n)$$

*konvergent mit Grenzwert  $s$ , wobei*

$$s = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n .$$

### Die geometrische Reihe.

Als Übung ist bereits gezeigt:



**Satz 5.6.** Die geometrische Reihe *konvergiert für  $|x| < 1$*  und es ist

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Für  $|x| \geq 1$  *divergiert die Reihe.*

### Dezimaldarstellung.

Die Darstellung einer reellen Zahl  $x$  als Dezimalzahl (Zur Erinnerung: Die reellen Zahlen sind in Kapitel 3 sogar über ihre Dezimaldarstellung definiert.) lautet

$$x = n + x_0, \quad n \in \mathbb{N}_0, \quad 0 \leq x_0 < 1,$$

wobei negative Zahlen zusätzlich mit einem Vorzeichen versehen sind. Der Anteil  $x_0$  ("hinter dem Komma",  $x_0 = 0.s_1s_2s_3\dots$  mit Ziffern  $s_j$ ,  $0 \leq s_j \leq 9$  für alle  $j \in \mathbb{N}$ ) ist dabei als Reihe zu verstehen:

$$x_0 = \sum_{j=1}^{\infty} s_j \cdot 10^{-j}.$$

**Beispiel.** Es ist

$$\begin{aligned} 0.\bar{3} &:= \sum_{j=1}^{\infty} 3 \cdot 10^{-j} = 3 \cdot \sum_{j=1}^{\infty} (10^{-1})^j \\ &= 3 \cdot \left( \left( \sum_{j=0}^{\infty} (10^{-1})^j \right) - 1 \right) = 3 \cdot \left( \frac{1}{1-10^{-1}} - 1 \right) \\ &= 3 \cdot \frac{1}{9} = \frac{1}{3}. \end{aligned}$$

### Konvergenzkriterien.

i) Satz von der monotonen Folge für Reihen.

**Satz 5.7.** Eine reelle Zahlenreihe  $\sum_{j=1}^{\infty} a_j$  mit *nicht-negativen Gliedern  $a_j$*  der zugrunde liegenden Folge  $\{a_j\}$  ist *genau dann konvergent*, wenn es eine Zahl  $k > 0$  gibt mit

$$\sum_{j=1}^n a_j \leq k \quad \text{für alle } n \in \mathbb{N}.$$

ii) Cauchys Konvergenzkriterium für Reihen.

**Satz 5.8.** Eine reelle Zahlenreihe  $\sum_{n=1}^{\infty} a_n$  ist *genau dann* konvergent, wenn es zu jedem  $\varepsilon > 0$  ein  $N(\varepsilon) \in \mathbb{N}$  gibt, sodass für alle  $p \geq 1$

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+p}| < \varepsilon \quad \text{für alle } n \geq N(\varepsilon).$$

**Konsequenz.** Wenn eine Reihe  $\sum_{n=1}^{\infty} a_n$  konvergiert, so bilden ihre Glieder  $a_n$  eine **Nullfolge**.

**Vorsicht.** Die Umkehrung gilt nicht.

**Beispiel.** Harmonische Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n}.$$

Für die Partialsummen gilt

$$\begin{aligned} s_{2k} - s_k &= \sum_{n=1}^{2k} \frac{1}{n} - \sum_{n=1}^k \frac{1}{n} = \frac{1}{k+1} + \frac{1}{k+2} + \cdots + \frac{1}{2k} \\ &> k \cdot \frac{1}{2k} = \frac{1}{2}. \end{aligned}$$

Nach dem Cauchyschen Konvergenzkriterium **divergiert die harmonische Reihe**.

iii) **Konvergenzkriterium von Leibniz.**

Die Situation ändert sich, wenn die Folgenglieder der harmonischen Reihe mit einem alternierenden Vorzeichen versehen sind: Die Reihe

$$\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$$

konvergiert nach dem Konvergenzkriterium von Leibniz:

**Satz 5.9.** Ist  $\{a_n\}$  eine *monoton fallende Nullfolge*, so konvergiert die *alternierende Reihe*  $a_1 - a_2 + a_3 - a_4 + \dots$ .

Anhand dieses Beispiels wird die Bedeutung des Begriffes “*absolute Konvergenz*” deutlich:

**Definition 5.6.** Eine Reihe  $\sum_{n=1}^{\infty} a_n$  heißt *absolut konvergent*, wenn die Reihe  $\sum_{n=1}^{\infty} |a_n|$  konvergiert.

### Bemerkungen.

- (a) Die Reihe  $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$  ist also konvergent aber nicht absolut konvergent.
- (b) Absolut konvergente Reihen konvergieren, *die Umkehrung ist falsch.*
- (c) Nur bei absolut konvergenten Reihen dürfen die Glieder *umsortiert* werden, *ansonsten darf die Reihenfolge der Glieder nicht vertauscht werden.*
- (d) Für absolut konvergente Reihen gilt

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n| .$$

iv) *Majorantenkriterium.*

In den meisten Fällen wird die Konvergenz einer Reihe über einen Vergleich mit einer bekannten Reihe gezeigt.

**Definition 5.7.** Es sei  $\sum_{n=1}^{\infty} b_n$ ,  $b_n \geq 0$  für alle  $n \in \mathbb{N}$ , eine reelle Zahlenreihe.

(a) Die Reihe heißt **Majorante** einer reellen Zahlenreihe  $\sum_{n=1}^{\infty} a_n$ , wenn es einen Index  $n_0 \in \mathbb{N}$  gibt, sodass

$$|a_n| \leq b_n \quad \text{für alle } n \geq n_0 .$$

(b) Die Reihe heißt **Minorante** einer reellen Zahlenreihe  $\sum_{n=1}^{\infty} a_n$ , wenn es einen Index  $n_0 \in \mathbb{N}$  gibt, sodass

$$b_n \leq a_n \quad \text{für alle } n \geq n_0 .$$

**Satz 5.10.** Es sei  $\sum_{n=1}^{\infty} a_n$  eine reelle Zahlenreihe.

(a) Hat die Reihe eine **konvergente Majorante**, so **konvergiert** sie.

(b) Hat die Reihe eine **divergente Minorante**, so **divergiert** sie.

### Beispiele.

(a) Für eine Folge  $\{a_n\}$  gelte  $|a_n| \leq 3/2^n$  für alle  $n \in \mathbb{N}$ . Dann konvergiert die Reihe  $\sum_{n=1}^{\infty} a_n$ .

(b) Für eine Folge  $\{a_n\}$  gelte  $a_n \geq 1/n$  für alle  $n \in \mathbb{N}$ . Dann divergiert die Reihe  $\sum_{n=1}^{\infty} a_n$ .

### v) Quotienten- und Wurzelkriterium.

**Satz 5.11.** Es sei  $\sum_{n=1}^{\infty} a_n$  eine reelle Zahlenfolge mit  $a_n \neq 0$  für alle  $n \in \mathbb{N}$ . Die Reihe ist **absolut konvergent** (und damit konvergent), wenn eine der beiden folgenden Bedingungen erfüllt ist.

(a) Es existieren ein  $0 < q < 1$  und ein  $N \in \mathbb{N}$  derart, dass

$$\left| \frac{a_{n+1}}{a_n} \right| \leq q < 1 \quad \text{für alle } n \geq N .$$

(b) Es existieren ein  $0 < q < 1$  und ein  $N \in \mathbb{N}$  derart, dass

$$\sqrt[n]{|a_n|} \leq q < 1 \quad \text{für alle } n \geq N .$$

Die Reihe ist *divergent*, falls ein  $N \in \mathbb{N}$  existiert, sodass für alle  $n \geq N$  gilt

$$\left| \frac{a_{n+1}}{a_n} \right| \geq 1$$

oder im Falle des Wurzelkriteriums

$$\sqrt[n]{|a_n|} \geq 1.$$

### Bemerkungen..

- (a) Im Fall  $q = 1$  liefern weder das Quotienten noch das Wurzelkriterium eine Aussage.
- (b) Existiert der Grenzwert

$$q = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \quad \text{oder} \quad q = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|},$$

so hat für  $q < 1$  absolute Konvergenz, für  $q > 1$  Divergenz und für  $q = 1$  keine Aussage.

### Beispiele.

- (a) Die Reihe

$$\sum_{n=1}^{\infty} \frac{n!}{n^n}$$

konvergiert nach dem Quotientenkriterium (absolut), da die Definition der Eulerschen Zahl die Abschätzung

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{(n+1)!n^n}{n!(n+1)^{n+1}} = \frac{n^n}{(n+1)^n} = \frac{1}{\left(1 + \frac{1}{n}\right)^n} \leq \frac{1}{2}$$

liefert.

- (b) Betrachtet sei die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Dann gilt

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{n^2}{(n+1)^2} = \left( 1 - \frac{1}{n+1} \right)^2.$$

Hier gibt es **kein festes  $q < 1$**  mit  $|a_{n+1}/a_n| \leq q$ : **Das Quotientenkriterium liefert keine Aussage!**

**Wichtige Bemerkung.** Mithilfe des so genannten **Cauchy-schen Vedichtungskriterium** kann aber gezeigt werden:

Für alle  **$\alpha > 1$  konvergiert** die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}},$$

für alle  **$\alpha \leq 1$  divergiert** die Reihe.

Dieses Beispiel hilft oft bei der Suche nach konvergenten Majoranten bzw. divergenten Minoranten.

# Kapitel 6

## Potenzreihen

Potenzreihen sind **Funktionenreihen** mit einer **speziellen Struktur**. Mit ihrer Hilfe können viele “elementare Funktionen” eingeführt werden. Ein Beispiel ist bereits ausführlich diskutiert:

**Beispiel.** (Geometrische Reihe) Man betrachte für fixiertes  $x_0 \in \mathbb{R}$  die Funktionenfolge  $\{(x - x_0)^n\}$ . Die Reihe

$$\sum_{n=0}^{\infty} (x - x_0)^n$$

kann dort, wo sie konvergiert, als Funktion von  $x$  angesehen werden:

$$f(x) := \sum_{n=0}^{\infty} (x - x_0)^n .$$

Über die Reihe ist eine Funktion  $f: (x_0 - 1, x_0 + 1) \rightarrow \mathbb{R}$  definiert, denn: Nach Satz 5.6 konvergiert die Reihe für alle  $x \in \mathbb{R}$  mit  $|x - x_0| < 1$ , sie divergiert für alle  $x \in \mathbb{R}$  mit  $|x - x_0| > 1$ .

Ist  $|x - x_0| = 1$ , so divergiert die Reihe ebenfalls. Dies folgt aber aus einer **gesonderten Betrachtung der “Randpunkte”**  $x_0 - 1$  und  $x_0 + 1$  und nicht aus der “allgemeinen Theorie” (vgl. Satz 5.6).

**Definition 6.1.** Eine (Funktionen-) Reihe der Form

$$\sum_{n=0}^{\infty} a_n \cdot (x - x_0)^n$$

heißt **Potenzreihe** um den (fixierten) **Entwicklungspunkt**  $x_0 \in \mathbb{R}$  mit den **Koeffizienten**  $a_n \in \mathbb{R}$ .

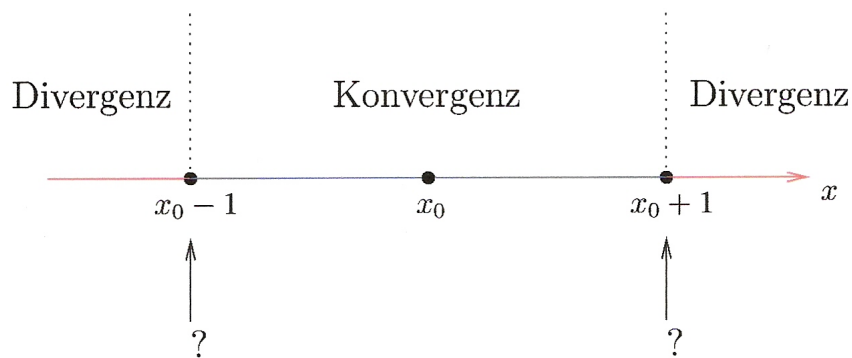


Abbildung 6.1: Zum Konvergenzverhalten der geometrischen Reihe.

### Verhalten sich allgemeine Potenzreihen qualitativ ähnlich wie die geometrische Reihe?

**Satz 6.1.** Zu jeder Potenzreihe  $\sum_{n=0}^{\infty} a_n(x - x_0)^n$  wie oben gibt es eine reelle Zahl  $\rho \geq 0$  und ein Intervall  $(x_0 - \rho, x_0 + \rho)$  mit

- i) Die Potenzreihe *konvergiert für alle  $x \in (x_0 - \rho, x_0 + \rho)$*  (punktweise).
- ii) *Außerhalb von  $[x_0 - \rho, x_0 + \rho]$  divergiert die Potenzreihe.*
- iii) Die Zahl  $\rho$  heißt der *Konvergenzradius* der Potenzreihe, das Intervall  $(x_0 - \rho, x_0 + \rho)$  (für  $\rho > 0$ ) das *Konvergenzintervall*.

Es gilt die Formel von *Cauchy-Hadamard*

$$\frac{1}{\rho} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}.$$

- iv) Im Fall  $\rho = 0$  konvergiert die Reihe nur im Punkt  $x = x_0$ , im (formalen) Fall  $\rho = \infty$  konvergiert sie für alle  $x \in \mathbb{R}$ .

### Bemerkungen.

- i) Die Randpunkte des Konvergenzintervalls sind wie im Beispiel der geometrischen Reihe *gesondert zu untersuchen*. Hier liefert der Satz *keine Aussagen* über Konvergenz oder Divergenz der Reihe.



- ii) Zur Überprüfung der Konvergenz in einem festen Punkt  $x \in \mathbb{R}$  können auch die Kriterien aus Kapitel 5 angewandt werden. Man beachte dabei, dass die Glieder der zugrunde liegenden Folge dann  $a_n \cdot (x - x_0)^n$  sind.

**Beispiel.** Für die geometrische Reihe ergibt sich wie bereits bekannt  $\rho = 1$ .

### Definition elementarer Funktionen.

**Definition 6.2.** Als reelle Funktionen sind die *Exponentialfunktion*, die *trigonometrischen Funktionen Sinus und Cosinus* sowie die *Hyperbelfunktionen Sinus Hyperbolicus und Cosinus Hyperbolicus* definiert als

$$\begin{aligned} \exp(x) &:= \sum_{n=0}^{\infty} \frac{1}{n!} x^n \quad (= e^x), \\ \sin(x) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \\ \cos(x) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \\ \sinh(x) &:= \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} x^{2n+1}, \\ \cosh(x) &:= \sum_{n=0}^{\infty} \frac{1}{(2n)!} x^{2n}. \end{aligned}$$

### Bemerkungen.

- i) Das Quotientenkriterium (für fixiertes  $x$ ) liefert die **Konvergenz der Reihen für alle  $x$** .
- ii) Anhand der Definition der Funktion  $\exp(x)$  als Reihe ist zunächst der Zusammenhang mit der Funktion  $e^x$  nicht ersichtlich.

Mithilfe des so genannten **Cauchy-Produktes** für Reihen und des binomischen Lehrsatzes verifiziert man aber, dass die Funktion  $\exp(x)$  die **charakteristische Funktionalgleichung** (vgl. Kapitel 3)

erfüllt. Dies ist das wichtigste Argument, um  $\exp(x) = e^x$  zu zeigen.

iii) Die Exponentialfunktion wächst “für große  $x$ ” schneller als jede Potenz von  $x$ .

### Definition im Komplexen.

Potenzreihen sind im Komplexen völlig analog von der Form

$$\sum_{n=0}^{\infty} c_n \cdot (z - z_0)^n$$

mit komplexen Koeffizienten  $c_n$  und mit einem Entwicklungspunkt  $z_0 \in \mathbb{C}$  definiert.

Satz 6.1 gilt in diesem Fall, wenn das Konvergenzintervall durch die **offene komplexe Kreisscheibe**

$$B_\rho(z_0) := \{z \in \mathbb{C} : |z - z_0| < \rho\}$$

ersetzt wird.

Der **Rand der Kreisscheibe**, auf dem keine Information geliefert wird ist

$$\partial B_\rho(z_0) := \{z \in \mathbb{C} : |z - z_0| = \rho\}.$$

**Definition 6.3.** *Als komplexe Funktionen sind die Exponentialfunktion, die trigonometrischen Funktionen Sinus und Cosinus sowie die Hyperbelfunktionen Sinus Hyperbolicus und Cosinus Hyperbolicus wie in Definition 6.2 definiert, wenn nur die reelle Variable  $x \in \mathbb{R}$  durch die komplexe Variable  $z \in \mathbb{C}$  ersetzt wird.*

**Übung.** Man zeige, dass für alle  $z \in \mathbb{C}$

$$e^{iz} = \cos(z) + i \sin(z)$$

gilt und interpretiere das Resultat geometrisch (vgl. Abbildung 3.16).

**Weierführende Bemerkung.** Die obige Definition der komplexen Exponentialfunktion ist tatsächlich die **einzig mögliche Definition** einer

komplex differenzierbaren Funktion, die im Reellen mit der reellen Exponentialfunktion übereinstimmt.

Analoges gilt für Sinus etc.



# Kapitel 7

## Stetige Funktionen

Grenzwerte von Folgen und Reihen bilden die Grundlage für den Einstieg in die [Infinitesimalrechnung](#), d.h. in die [Differential- und Integralrechnung](#).

Der intuitive Begriff der [Stetigkeit](#) einer Funktion ist: Der Graph einer stetigen Funktion kann “ohne Absetzen des Stiftes” gezeichnet werden (vgl. Abbildungen 7.1, 7.2, 7.3).

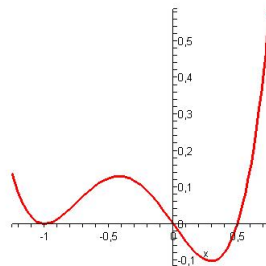


Abbildung 7.1: Der Graph einer stetigen Funktion.

Die Grenzen dieses intuitiven Begriffes werden aber deutlich, wenn etwa die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,

$$f(x) = \begin{cases} x \cos(1/x), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

betrachtet wird (vgl. Abbildung 7.4).

Hier kann nicht intuitiv anhand des Graphen entschieden werden, ob die Funktion stetig ist.

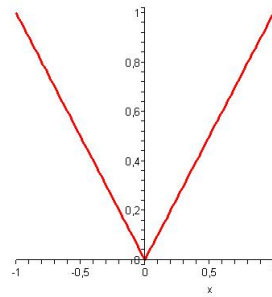


Abbildung 7.2: Der Graph der Betragsfunktion.

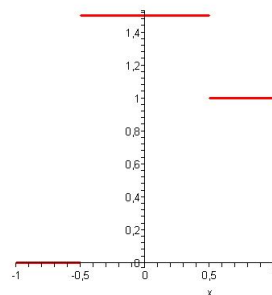


Abbildung 7.3: Der Graph einer unstetigen Funktion mit Sprungstellen.

## Grenzwert und Stetigkeit einer Funktion.

Im Folgenden sei  $I \subset \mathbb{R}$  stets ein (evtl. unendliches) Intervall und  $f: I \rightarrow \mathbb{R}$  eine Funktion von  $I$  nach  $\mathbb{R}$ .

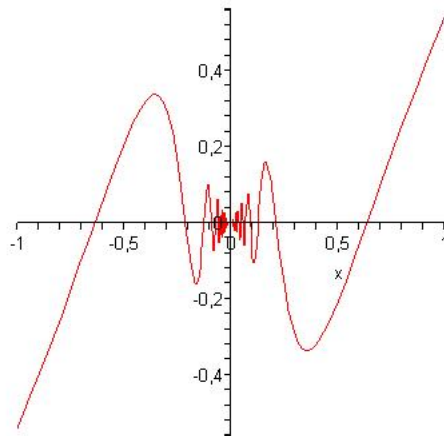
**Beobachtung.** Ist  $\{x_n\}$  eine Folge von Elementen aus  $I$ , so ist  $\{f(x_n)\}$  eine reelle Zahlenfolge, die auf Konvergenz untersucht werden kann.

**Definition 7.1.** *i) Es seien  $x_0 \in I$  und  $a \in \mathbb{R}$  fixiert. Dann hat  $f$  an der Stelle  $x_0$  den Grenzwert  $a$ , falls für jede Folge  $\{x_n\}$  aus  $I$  mit  $x_n \neq x_0$  für alle  $n \in \mathbb{N}$  und  $x_n \rightarrow x_0$  für  $n \rightarrow \infty$  gilt*

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

*Notation:*

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

Abbildung 7.4: Die Funktion  $f(x) = x \cos(1/x)$ .

ii) Die Funktion  $f$  heißt *stetig in einem Punkt*  $x_0 \in I$ , falls

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

iii) Die Funktion heißt *stetig auf  $I$* , wenn sie *in jedem Punkt*  $x_0 \in I$  stetig ist.

**Bemerkung.** “Wie immer” gilt: Falls ein Grenzwert existiert, so ist er *eindeutig bestimmt*.

**Reicht es nicht, nur eine Folge  $\{x_n\}$  zu betrachten?**

**Beispiel.** Es sei  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,

$$f(x) = \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x \geq 0. \end{cases}$$

Nach dem intuitiven Stetigkeitsbegriff ist die Funktion im Nullpunkt nicht stetig.

Dazu betrachte man zunächst etwa die gegen Null konvergente Folge  $\{1/n\}$ . Es ist

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}\right) = \lim_{n \rightarrow \infty} 1 = 1 ,$$

die Folge  $\{f(1/n)\}$  konvergiert also gegen 1.

Nun betrachte man beispielsweise die ebenfalls gegen Null konvergente Folge  $\{-1/n\}$  mit

$$\lim_{n \rightarrow \infty} f\left(-\frac{1}{n}\right) = \lim_{n \rightarrow \infty} 0 = 0.$$

Auch die Folge  $\{f(-1/n)\}$  konvergiert, allerdings mit Grenzwert 0.

Damit ist  $f$  im Nullpunkt nicht stetig.

An diesem Beispiel wird deutlich, dass es **nicht ausreicht, eine spezielle Folge zu betrachten**.

### Beispiele stetiger Funktionen.

i) Konstante Funktionen und die Identität  $f(x) = x$  sind stetig (Übung).

ii) Es sei  $f(x) = x^2$  und  $x_0 \in \mathbb{R}$  sei fixiert. Weiter sei  $\{x_n\}$  eine beliebige Folge mit  $x_n \rightarrow x_0$  für  $n \rightarrow \infty$ .

Die Stetigkeit von  $f$  ist gezeigt, falls dann

$$f(x_n) \xrightarrow{n \rightarrow \infty} f(x_0)$$

gilt.

Tatsächlich ist

$$\begin{aligned} |f(x_n) - f(x_0)| &= |x_n^2 - x_0^2| = |(x_n - x_0)(x_n + x_0)| \\ &\leq |x_n + x_0| |x_n - x_0|. \end{aligned}$$

Wegen (Dreiecksungleichung)

$$|x_n + x_0| \leq |x_n| + |x_0|$$

und da die Folge  $\{x_n\}$  als konvergente Folge beschränkt ist (vgl. Kapitel 5), existiert eine Konstante  $K \in \mathbb{R}$  mit

$$|x_n + x_0| \leq K.$$

Es folgt

$$|f(x_n) - f(x_0)| \leq K|x_n - x_0| \xrightarrow{n \rightarrow \infty} 0,$$

was zu zeigen war. □



- iii) Die Betragsfunktion  $f(x) = |x|$  ist stetig (Übung).
- iv) Aus einer tiefer gehenden Diskussion der Konvergenz von Potenzreihen folgt, dass  $\exp$ ,  $\sin$ ,  $\cos$ ,  $\sinh$ ,  $\cosh$  stetige Funktionen sind.

### Rechenregeln für stetige Funktionen.

Mithilfe des folgenden Satzes (offensichtliche Notation  $(f + g)(x) := f(x) + g(x)$  etc.) erkennt unmittelbar die Stetigkeit einer Vielzahl von Funktionen (z.B. aller Polynome).

**Satz 7.1.** *i) Es seien  $f, g$  zwei auf  $I$  stetige reellwertige Funktionen und  $c \in \mathbb{R}$  sei fixiert.*

*(a) Dann sind auch  $cf$ ,  $f + g$  und  $f \cdot g$  stetige Funktionen.*

*(b) In allen Punkten, in denen  $g$  nicht den Wert 0 annimmt, ist  $\frac{f}{g}$  stetig.*

*ii) Die Verkettung  $g \circ f$  stetiger Funktionen  $f, g$  ist stetig, sofern sie definiert ist.*

### Beispiele.

- i) Wie bereits erwähnt sind alle Polynome stetig.
- ii) Funktionen wie  $\sin^2(x)$  oder  $\sin(x) \cosh(x)$  sind stetig.
- iii) Satz 7.1 impliziert auch die Stetigkeit von  $e^{\sin(x)} = (\exp \circ \sin)(x)$ .

### Stetigkeit im Komplexen.

Alle bisherigen Betrachtungen übertragen sich wörtlich auf den Fall von Funktionen  $f: \mathbb{C} \rightarrow \mathbb{C}$ .

Dies gilt **nicht mehr** (warum?) für die folgende Diskussion.

## Weitere Eigenschaften stetiger reellwertiger Funktionen.

Wie bereits angedeutet, hängt die Existenz von Extrema einer Funktion u.A. vom Definitionsbereich ab (vgl. Abbildung 3.4). Dies kann nun mithilfe des Stetigkeitsbegriffes präzisiert werden.

**Satz 7.2.** *Ist  $I = [a, b]$  ein kompaktes Intervall (d.h. abgeschlossen und beschränkt) und  $f: I \rightarrow \mathbb{R}$  stetig auf  $I$ , so nimmt die Funktion  $f$  in  $I$  ihr absolutes Maximum und ihr absolutes Minimum an, d.h. es existieren Punkte  $x_1, x_2 \in I$  mit*

$$f(x_1) \leq f(x) \leq f(x_2) \quad \text{für alle } x \in I .$$

Die Stetigkeit etwa der Logarithmusfunktion (und vieler bereits angesprochener Umkehrfunktionen) folgt aus

**Satz 7.3.** *Eine stetige, streng monotone Funktion  $f: I \rightarrow \mathbb{R}$  besitzt eine stetige, streng monotone Umkehrfunktion  $f^{-1}: f(I) \rightarrow I$ .*

Satz 7.3 folgt wiederum aus

**Satz 7.4. (Zwischenwertsatz)** *Ist  $f: \mathbb{R} \subset [a, b] \rightarrow \mathbb{R}$  stetig, so nimmt  $f$  jeden Wert zwischen  $f(a)$  und  $f(b)$  an, d.h.:*

*Aus  $f(a) < \xi < f(b)$  folgt die Existenz von (mindestens) einem Punkt  $x_0 \in (a, b)$  mit  $f(x_0) = \xi$ .*

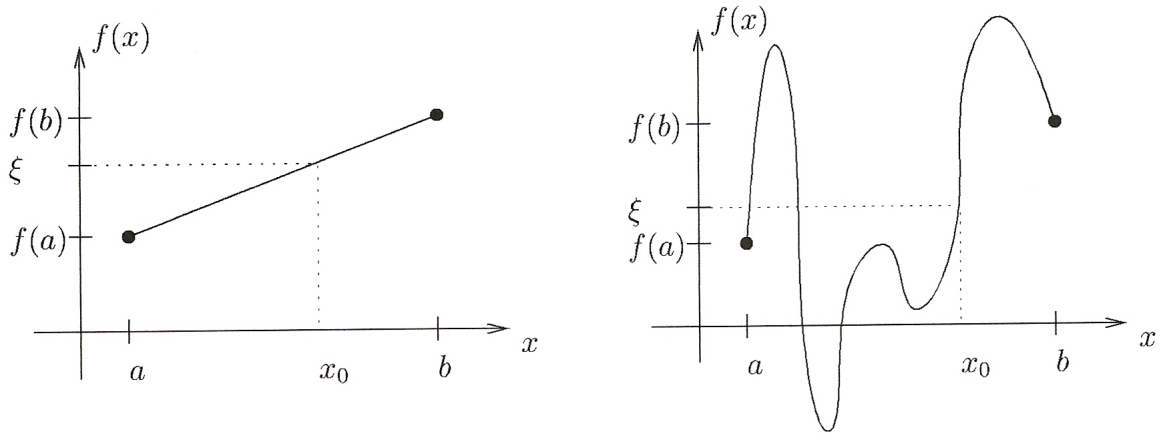


Abbildung 7.5: Zum Zwischenwertsatz.

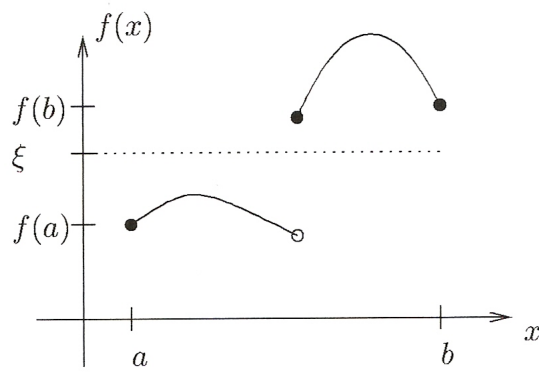


Abbildung 7.6: Die Stetigkeit muss im Zwischenwertsatz vorausgesetzt werden.



# Kapitel 8

## Differentialrechnung in einer Veränderlichen

### 8.1 Grundlagen

Der intuitive Begriff der **Differenzierbarkeit** einer Funktion ist: Der Graph einer differenzierbaren Funktion hat keinen “Knick”. Dementsprechend sollte die Betragsfunktion (vgl. Abbildung 7.2) zwar stetig aber nicht differenzierbar im Nullpunkt sein, was auch tatsächlich der Fall ist.

Für die Tauglichkeit des intuitiven Begriffs der Differenzierbarkeit gelten aber ähnliche Einschränkungen wie im letzten Kapitel zur Stetigkeit.

#### **Geometrische Interpretation der Ableitung.**

Zur geometrischen Interpretation der Ableitung wird, wie in Abbildung 8.1 angedeutet, in einem festen Punkt  $x_0$  die **Tangente** an den Graphen der Funktion als (affin) **lineare Approximation der Funktion** betrachtet. M.a.W. soll der Graph in der Nähe des Punktes  $(x_0, f(x_0))$  “möglichst gut” mit einer Geraden angenähert werden.

#### **Kinematische Interpretation der Ableitung.**

Bei der kinematischen Interpretation der Ableitung betrachtet man etwa die Bewegung eines Massenpunktes im dreidimensionalen Raum in Abhängigkeit von der Zeit.

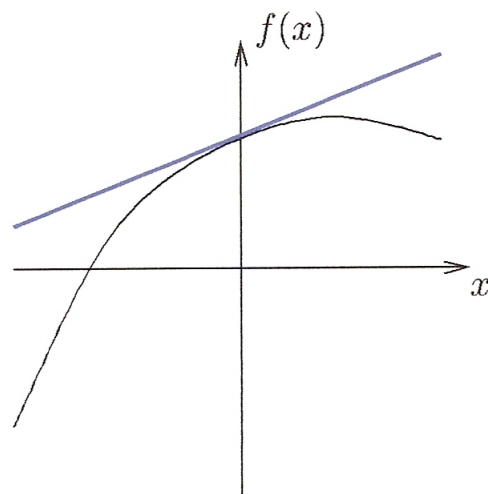


Abbildung 8.1: Zur geometrischen Interpretation der Ableitung.

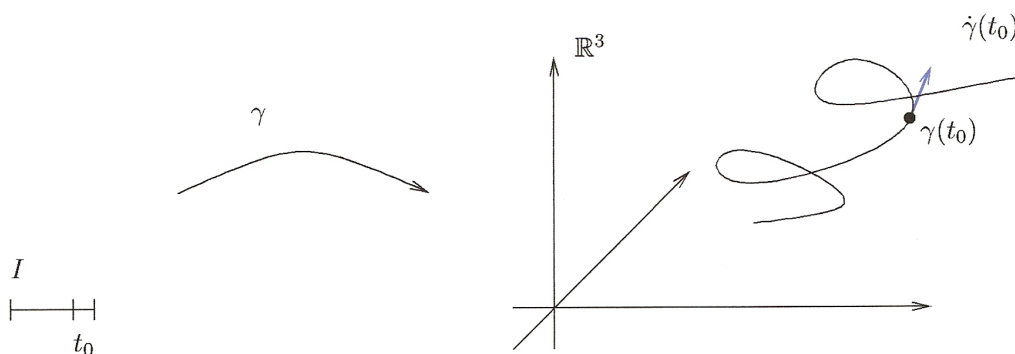


Abbildung 8.2: Zur kinematischen Interpretation der Ableitung.

Diese Bewegung wird durch eine **Kurve** im  $\mathbb{R}^3$  beschrieben.

Das ist eine Abbildung  $\gamma: \mathbb{R} \supset I \rightarrow \mathbb{R}^3$ , wobei  $I$  ein Zeitintervall bezeichnet.

Die Ableitung entspricht in diesem Fall dem **Geschwindigkeitsvektor** des Massenpunktes zu einem gegebenen Zeitpunkt  $t_0$ .

### **Differenzenquotient und Ableitung.**

Im Folgenden sei  $I \subset \mathbb{R}$  stets ein offenes (evtl. unendliches) Intervall und  $f: I \rightarrow \mathbb{R}$  eine Funktion von  $I$  nach  $\mathbb{R}$ .

**Definition 8.1.** Es sei  $f$  wie oben und  $x_0 \in I$  fixiert.

i) Ist  $x \in I$ ,  $x \neq x_0$ , so heißt der Quotient

$$\frac{f(x) - f(x_0)}{x - x_0}$$

Differenzenquotient von  $f$  bzgl. der Punkte  $x$  und  $x_0$ .

ii) Die Funktion  $f$  heißt *differenzierbar im Punkt  $x_0 \in I$* , falls der Grenzwert

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

existiert.

Der Grenzwert heißt die *Ableitung* oder der *Differentialquotient* von  $f$  bei  $x_0$ .

Notation:

$$f'(x_0) \quad \text{oder} \quad \frac{df}{dx}(x_0) \quad \text{oder} \quad \frac{df}{dx}|_{x=x_0}.$$

iii) Die Funktion heißt *differenzierbar auf  $I$* , wenn sie *in jedem Punkt  $x_0 \in I$*  differenzierbar ist.

Dann kann eine weitere Funktion (genannt die *Ableitung von  $f$* ) betrachtet werden, die jedem Punkt  $x \in I$  die Ableitung an der Stelle  $x$  zuordnet:

$$f' : I \rightarrow \mathbb{R}, \quad I \ni x \mapsto f'(x).$$

### Interpretation.

i) Geometrisch entspricht der Differenzenquotient der *Steigung der Sekante*  $s$  durch die Punkte  $(x, f(x))$  und  $(x_0, f(x_0))$  (vgl. Abbildung 8.3):

$$s(t) = f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(t - x_0).$$

Im Grenzwert erhält man die *Steigung der Tangente* an  $(x_0, f(x_0))$ .

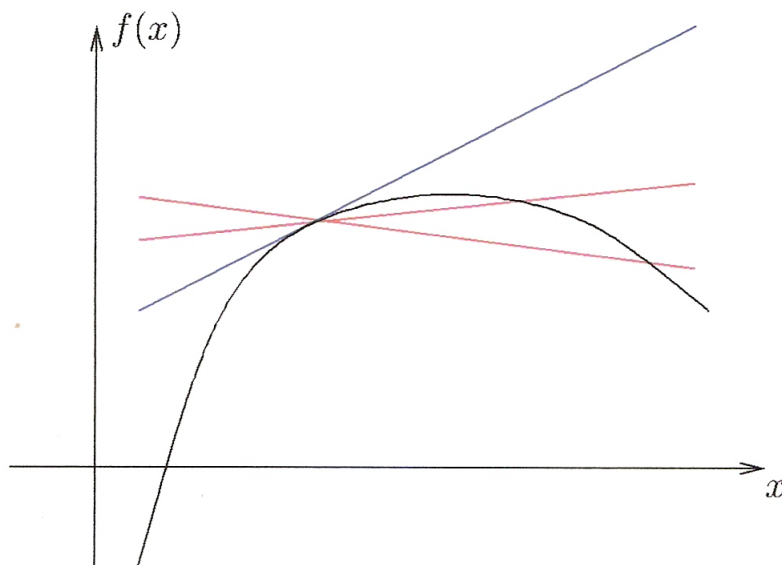


Abbildung 8.3: Die Sekantensteigungen gehen in die Steigung der Tangente über.

- ii) Kinematisch entspricht der Differenzenquotient der **Durchschnittsgeschwindigkeit** im Zeitraum zwischen  $t_0$  und  $t$ .

Der Grenzwert gibt die **Momentangeschwindigkeit** zum Zeitpunkt  $t_0$  an.

### Beispiele differenzierbarer Funktionen.

- i) Im Falle einer konstanten Funktion verschwindet der Differenzenquotient identisch. Also ist die Funktion überall differenzierbar mit der Ableitung 0.

- ii) Es sei  $f(x) = x$  für alle  $x \in \mathbb{R}$ . Der Differenzenquotient für ein fixiertes  $x_0$  lautet

$$\frac{x - x_0}{x - x_0} \equiv 1 .$$

Die Funktion ist überall differenzierbar mit der Ableitung 1.

- iii) Es sei  $f(x) = x^2$  für alle  $x \in \mathbb{R}$  und  $x_0 \in \mathbb{R}$  fixiert. Dann gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^2 - x_0^2}{x - x_0} = \frac{(x - x_0)(x + x_0)}{x - x_0} = x + x_0 .$$



Um zum Grenzwert und damit zur Ableitung überzugehen, ist in diesen Differenzenquotienten eine beliebige Folge  $\{x_n\}$  mit Grenzwert  $x_0$  einzusetzen (vgl. Definition Grenzwert einer Funktion, Definition 7.1).

Da die Folge  $\{x_n\}$  aber gegen  $x_0$  konvergiert, konvergiert die rechte Seite der obigen Gleichheit, d.h. auch der Differenzenquotient, gegen  $2x_0$ . Es ist gezeigt:

Die Funktion  $f(x) = x^2$  ist überall differenzierbar mit der Ableitung  $f'(x) = 2x$ .

*iv)* In Tabelle 8.1 sind ohne Beweis die Ableitungen einiger bekannter Funktionen festgehalten.

$f(x)$	$f'(x)$	Def. Bereich	Bem.
$x^k$	$kx^{k-1}$	$\mathbb{R}, x \neq 0$ für $k < 0$	$k \in \mathbb{Z}$
$\sqrt{x}$	$\frac{1}{2} \frac{1}{\sqrt{x}}$	$\mathbb{R}^+$	vgl. $x^\alpha$
$x^\alpha$	$\alpha x^{\alpha-1}$	$\mathbb{R}^+$	$\alpha \in \mathbb{R}$
$e^x$	$e^x$	$\mathbb{R}$	
$\sin(x)$	$\cos(x)$	$\mathbb{R}$	
$\cos(x)$	$-\sin(x)$	$\mathbb{R}$	
$\sinh(x)$	$\cosh(x)$	$\mathbb{R}$	
$\cosh(x)$	$\sinh(x)$	$\mathbb{R}$	

Tabelle 8.1: Die Ableitungen einiger bekannter Funktionen.

## Stetigkeit und Differenzierbarkeit.

**Typisches Beispiel.** Es sei  $f(x) = |x|$  und  $x_0 = 0$ .

Für die Folge  $\{x_n\} = \{1/n\}$  gilt

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} = \frac{(1/n) - 0}{(1/n) - 0} = 1 .$$

Für die Folge  $\{\tilde{x}_n\} = \{-1/n\}$  gilt hingegen

$$\frac{f(\tilde{x}_n) - f(x_0)}{\tilde{x}_n - x_0} = \frac{(1/n) - 0}{(-1/n) - 0} = -1.$$

Wie erwartet ist die **Betragsfunktion im Nullpunkt nicht differenzierbar**.

Das Beispiel der stetigen Betragsfunktion belegt:

**Aus der Stetigkeit einer Funktion folgt nicht die Differenzierbarkeit!**

Es gilt aber die Umkehrung im Sinne von

**Satz 8.1.** *Es sei  $f: I \rightarrow \mathbb{R}$  im Punkt  $x_0 \in I$  differenzierbar. Dann ist die Funktion  $f$  in  $x_0$  stetig.*

*M.a.W.: Differenzierbarkeit impliziert Stetigkeit.*

### Rechenregeln für differenzierbare Funktionen.

Wie bei der Betrachtung stetiger Funktionen sind im nächsten Schritt Summe und Produkt differenzierbarer Funktionen zu untersuchen:

**Satz 8.2.** *Es seien  $f, g$  zwei auf  $I$  differenzierbare reellwertige Funktionen und  $c \in \mathbb{R}$  sei fixiert.*

*i) Dann sind auch  $cf, f + g$  und  $fg$  auf  $I$  differenzierbar mit*

$$(a) (cf)' = cf';$$

$$(b) (f + g)' = f' + g';$$

$$(c) (fg)' = f'g + fg' \text{ (Produktregel)}.$$

*ii) Ist  $g(x_0) \neq 0$  für  $x_0 \in I$ , so ist auch  $f/g$  differenzierbar mit (Quotientenregel)*

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

**Beweisidee Produktregel.** Exemplarisch wird hier nur gezeigt, wie die zwei Summanden bei der Ableitung eines Produktes entstehen:

Es seien  $x, x_0 \in I$ . Dann gilt

$$\begin{aligned} & \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x) - f(x_0)}{x - x_0}g(x) + f(x_0)\frac{g(x) - g(x_0)}{x - x_0} \\ &\xrightarrow{x \rightarrow x_0} f'(x_0)g(x_0) + f(x_0)g'(x_0) . \end{aligned}$$

Das ist aber genau die Behauptung. □

**Beispiel.** Es sei  $f(x) = \tan(x)$ ,  $x \in (-\pi/2, \pi/2)$ . Dann ist nach der Quotientenregel

$$f'(x) = \frac{\cos(x)\cos(x) - \sin(x)(-\sin(x))}{\cos^2(x)} = \frac{1}{\cos^2(x)} .$$

### Verkettung von Funktionen.

Auch bei der Verkettung von Funktionen kann man auf Differenzierbarkeit schließen, ohne dass explizit auf die Definition zurückgegriffen werden muss.

**Satz 8.3.** (*Kettenregel*) Die Verkettung  $g \circ f$  zweier reellwertiger differenzierbarer Funktionen  $f, g$  sei wohl definiert (d.h. das Bild von  $f$  liege im Definitionsbereich von  $g$ ). Dann ist auch  $g \circ f$  differenzierbar und es gilt

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

**Beispiel.** Für alle  $x \in \mathbb{R}$  sei  $h(x) = \sqrt{1+x^2}$ , d.h.  $h = g \circ f$  mit  $g(y) = \sqrt{y}$  und  $f(x) = 1+x^2$ .

Aus der Kettenregel folgt

$$\begin{aligned} h'(x) &= g'(f(x))f'(x) \\ &= \frac{1}{2\sqrt{1+x^2}} \cdot 2x \\ &= \frac{x}{\sqrt{1+x^2}}. \end{aligned}$$

### Ableitung der Umkehrfunktion.

Auch die Ableitung einer Umkehrfunktion kann direkt aus der Ableitung der Funktion selbst berechnet werden.

Der folgende Satz folgt unmittelbar aus der Kettenregel durch Differentiation der Identität  $(f^{-1} \circ f)(x) = x$ .

**Satz 8.4.** Die Funktion  $f: I \rightarrow f(I)$  ( $y = f(x)$ ) sei bijektiv und differenzierbar mit  $f'(x) \neq 0$  für alle  $x \in I$ . Dann ist für alle  $y \in f(I)$  die Ableitung der Umkehrfunktion  $f^{-1}$  gegeben durch

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

### Beispiele.

- i) Auf  $\mathbb{R}^+$  betrachte man die Funktion  $\ln(y) = f^{-1}(y)$  als Umkehrfunktion der Exponentialfunktion  $e^x = f(x)$ .

Nach Satz 8.4 gilt ( $y = e^x$ ,  $x = \ln(y)$ )

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

- ii) Die Ableitungen der trigonometrischen Umkehrfunktionen sind ohne Beweis in Tabelle 8.2 festgehalten.

### Höhere Ableitungen.

**Beobachtung.** In Definition 8.1, *iii*), wurde bereits festgestellt, dass die Ableitung einer auf  $I$  differenzierbaren Funktion  $f$  selbst wieder eine Funktion  $f': I \rightarrow \mathbb{R}$  ist.

$f(x)$	$f'(x)$	Def. Bereich
$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$	$-1 < x < 1$
$\arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$	$-1 < x < 1$
$\arctan(x)$	$\frac{1}{1+x^2}$	$\mathbb{R}$

Tabelle 8.2: Die Ableitungen der trigonometrischen Umkehrfunktionen.

Also kann auch  $f'$  auf Differenzierbarkeit untersucht werden.

**Definition 8.2.** *i) Es sei  $f: I \rightarrow \mathbb{R}$  differenzierbar auf  $I$  mit der Ableitung  $f': I \rightarrow \mathbb{R}$ .*

*Ist auch  $f'$  differenzierbar auf  $I$ , so heißt die **Ableitung von  $f'$**  die **zweite Ableitung von  $f$** .*

*Notation:*

$$f''(x) \quad \text{oder} \quad f^{(2)}(x) \quad \text{oder} \quad \frac{d^2}{dx^2}f(x).$$

*ii) Analog wird die dritte, vierte etc. Ableitung von  $f$  definiert.*

**Interpretation.** Sagt die erste Ableitung etwas über die Steigung des Graphen einer Funktion aus, so gibt die zweite Ableitung Informationen über die **Krümmung**, wie im nächsten Abschnitt noch deutlich wird.

**Beispiel.** Es sei wieder  $f(x) = \sqrt{1+x^2}$  für alle  $x \in \mathbb{R}$ . Dann gilt (siehe oben)

$$\begin{aligned} f''(x) &= \frac{d}{dx} \frac{x}{\sqrt{1+x^2}} = \frac{d}{dx} (1+x^2)^{-\frac{1}{2}} x \\ &= \left(-\frac{1}{2}\right) (1+x^2)^{-\frac{3}{2}} 2xx + (1+x^2)^{-\frac{1}{2}} \\ &= \frac{1}{(1+x^2)^{\frac{3}{2}}}. \end{aligned}$$

## 8.2 Lokale Extrema, Mittelwertsatz

Im Folgenden sei stets  $I = [a, b]$  bzw.  $I = (a, b)$ . Die Fälle  $I = (-\infty, \infty)$  etc. werden analog behandelt.

Bei der Beschreibung von Naturvorgängen gehört es zu den wesentlichen Aufgaben, die **Maxima oder Minima bestimmter Größen zu finden** (z.B. ein Energieminimum etc.).

Hier liefert die Differentialrechnung die wesentlichen Hilfsmittel.

### Notwendige Bedingung für lokale Extrema.

**Definition 8.3.** Eine Funktion  $f: I \rightarrow \mathbb{R}$  hat an der Stelle  $x_0 \in I$  ein **lokales Maximum** (**lokales Minimum**), falls ein  $r > 0$  existiert mit

$$f(x_0) \leq f(x) \quad (\text{lok. Max.}) \quad \text{bzw.} \quad f(x_0) \geq f(x) \quad (\text{lok. Min.})$$

für alle  $x \in I \cap (x_0 - r, x_0 + r)$ .

Gilt jeweils die **strikte Ungleichung**, so spricht man von einem **strengen lokalen Maximum** (**Minimum**).

### Bemerkungen.

- i) Lokale Maxima und lokale Minima heißen **lokale Extrema**.
- ii) Ein lokales Extremum ist von einem globalen Extremum wie es etwa Satz 7.2 liefert zu unterscheiden.

Die in Abbildung 8.4 angedeutete Funktion etwa hat ihre globalen Extrema am Rande, in  $x_0$  ein lokales Maximum und in  $x_1$  ein lokales Minimum.

**Idee.** Abbildung 8.5 legt es nahe, bei der Suche nach lokalen Extrema zunächst nach Punkten mit **horizontaler Tangente an den Graphen** zu suchen.

Tatsächlich gilt die **notwendige Bedingung**:

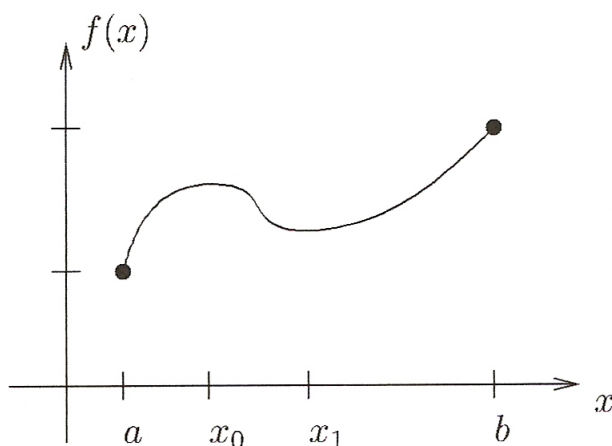


Abbildung 8.4: Lokale und globale Maxima (Minima) sind zu unterscheiden.

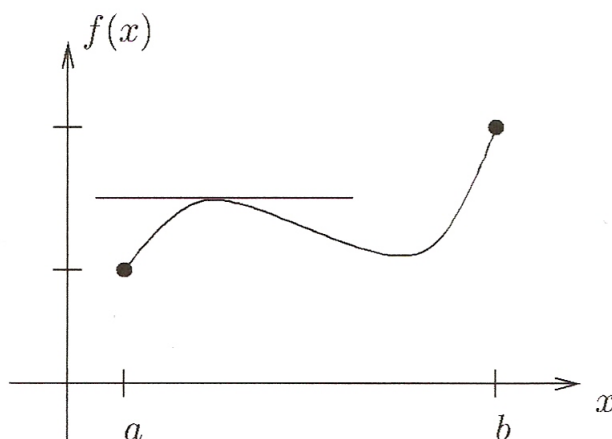


Abbildung 8.5: Die Tangente in einem lokalen Extremum.

**Satz 8.5.** Die Funktion  $f: I \rightarrow \mathbb{R}$  habe in einem Punkt  $x_0 \in (a, b)$  ein *lokales Extremum* und sei dort differenzierbar.

Dann gilt

$$f'(x_0) = 0 .$$

### Bemerkungen.

- i) Die Umkehrung von Satz 8.5 ist falsch, d.h. aus  $f'(x_0) = 0$  folgt nicht, dass ein lokales Extremum vorliegt (vgl. Abbildung 8.6).

Als Beispiel dient z.B. die Funktion  $f(x) = x^3$  im Nullpunkt.

Punkte mit horizontaler Tangente an den Graphen, d.h. mit

$f'(x_0) = 0$  heißen **stationäre** oder **kritische Punkte**.

- ii) Bei der Suche nach einem **globalen Extremum** einer differenzierbaren Funktion werden zunächst alle **kritischen Punkte** ermittelt. Deren Funktionswerte sowie die Funktionswerte in den **Randpunkten** werden anschließend verglichen.

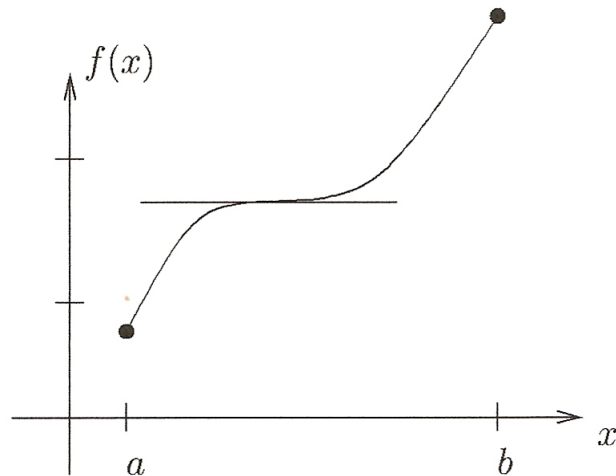


Abbildung 8.6: Kein lokales Extremum trotz horizontaler Tangente.

**Beispiel.** Betrachtet sei die Funktion  $f: [-2, 2] \rightarrow \mathbb{R}$ ,  $f(x) = (1 - x^2)^2$ .

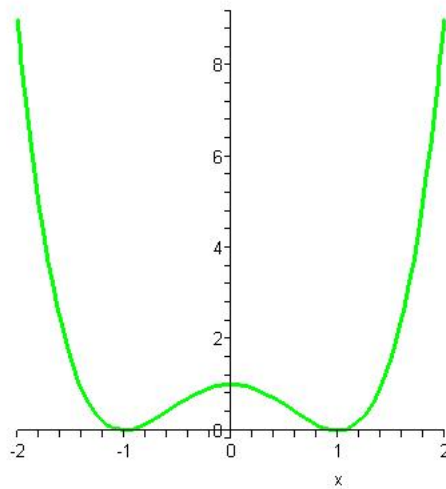


Abbildung 8.7:  $f(x) = (1 - x^2)^2$  in  $[-2, 2]$ .



Es gilt

$$f'(x) = -4x(1 - x^2),$$

woraus sich die stationären Punkte  $x_1 = -1$ ,  $x_2 = 0$  und  $x_3 = 1$  ergeben.

Wegen  $f \geq 0$  und  $f(x_1) = f(x_3) = 0$  hat  $f$  in  $x_1$  und  $x_3$  ein globales Minimum, also insbesondere auch ein lokales Minimum.

Da für alle  $x \in [-1, 1]$  die Ungleichung  $f(x) \leq 1 = f(x_2)$  gilt, liegt in  $x_2$  ein lokales Maximum vor.

Bei der Suche nach globalen Maxima sind die Funktionswerte in kritischen Punkten mit denen in den Randpunkten zu vergleichen:

$f(-2) = 9 = f(2)$ , d.h. das globale Maximum wird in den Randpunkten  $2$  und  $-2$  realisiert und nicht im Punkt  $x_2$ .

### Der Mittelwertsatz.

Eine mögliche Anwendung des folgenden Mittelwertsatzes 8.6 ist die Herleitung einer **hinreichenden Bedingung** für die Existenz lokaler Extrema.

Anschaulich sagt der Satz, dass für eine gegebene differenzierbare Funktion  $f$  im Intervall  $(a, b)$  (mindestens) ein Punkt  $x_0$  existiert, in dem die **Tangente an den Graphen von  $f$  die gleiche Steigung hat wie die Sekante durch  $(a, f(a))$  und  $(b, f(b))$**  (vgl. Abbildung 8.8).

**Satz 8.6.** Die Funktionen  $f, g: [a, b] \rightarrow \mathbb{R}$  seien stetig und auf  $(a, b)$  differenzierbar.

Dann existiert ein  $x_0 \in (a, b)$  mit

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

**Erste Konsequenzen aus dem Mittelwertsatz.** Es sei  $f$  wie in Satz 8.6.

i) Ist  $f'(x) = 0$  für alle  $x \in (a, b)$ , so ist  $f$  **konstant**.

ii) Es ist  $f'(x) \geq 0$  für alle  $x \in (a, b)$  **genau dann, wenn  $f$  auf  $[a, b]$  monoton wachsend** ist (analog “ $\leq$ ” und monoton fallend).

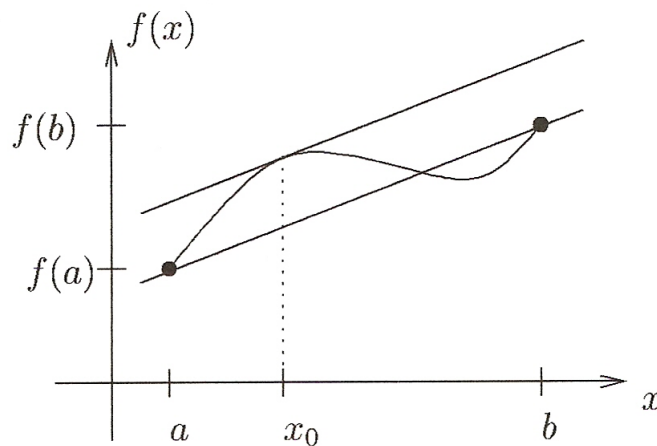


Abbildung 8.8: Zum Mittelwertsatz.

- iii) Ist  $f'(x) > 0$  für alle  $x \in (a, b)$ , so ist  $f$  **streng monoton wachsend** auf  $[a, b]$  (analog: “<” und streng monoton fallend). (Hier gilt die Umkehrung nicht, wie die Funktion  $f(x) = x^3$  belegt.)

### Konvexe und konkave Funktionen.

Um einzusehen, ob in einem kritischen Punkt tatsächlich ein Maximum oder ein Minimum vorliegt, überlegt man sich, ob die Funktion in der Nähe des Punktes konvex oder konkav ist.

Die geometrische Vorstellung ist dabei, dass der **Graph konvexer (konkaver) Funktionen stets oberhalb (unterhalb) seiner Tangenten liegt**. Dies ist in Abbildung 8.9 angedeutet.

Dementsprechend liegt die **Sekante durch die Punkte  $(a, f(a))$  und  $(b, f(b))$  bei einer konvexen (konkaven) Funktion im Intervall  $(a, b)$  oberhalb (unterhalb) des Graphen** (vgl. Abbildung 8.10):

- i) Eine Funktion  $f: I \rightarrow \mathbb{R}$  heißt konvex, falls für alle  $x, y \in I$  und für alle  $\lambda \in (0, 1)$  gilt

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) .$$

- ii) Eine Funktion  $f: I \rightarrow \mathbb{R}$  heißt konkav, falls für alle  $x, y \in I$  und

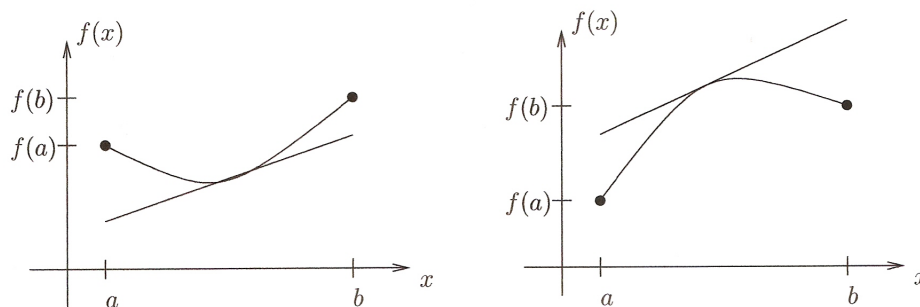


Abbildung 8.9: Der Graph einer konvexen (konkaven) Funktion liegt oberhalb (unterhalb) seiner Tangenten.

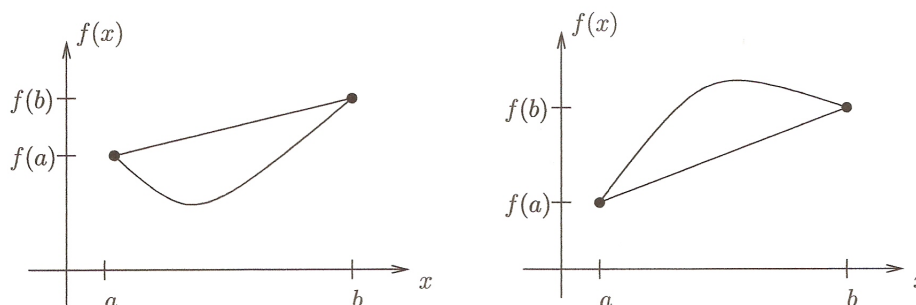


Abbildung 8.10: Die Sekante einer konvexen (konkaven) Funktion liegt in  $(a, b)$  oberhalb (unterhalb) des Graphen.

für alle  $\lambda \in (0, 1)$  gilt

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y) .$$

*iii)* Bei **streng konvexen** bzw. **streng konkaven** Funktionen gelten die strikten Ungleichungen.

### Beispiele.

*i)* Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^2$ , ist streng konvex.

*ii)* Die Funktion  $f: \mathbb{R}^+ \rightarrow \mathbb{R}$ ,  $f(x) = \ln(x)$ , ist streng konkav.

**Geometrische Idee.** Es sei  $f$  zweimal differenzierbar und **konvex**.

Wie Abbildung 8.9 nahelegt, ist die **erste Ableitung  $f'$  dann monoton wachsend**.

Für die Ableitung  $f''$  der monoton wachsenden Funktion  $f'$  gilt folglich  **$f'' \geq 0$** .

**Satz 8.7.** *Es sei  $f: (a, b) \rightarrow \mathbb{R}$  zweimal differenzierbar.*

*i) Ist  $f''(x) > 0$  für alle  $x \in (a, b)$ , so ist  $f$  **streng konvex**.*

*ii) Ist  $f''(x) < 0$  für alle  $x \in (a, b)$ , so ist  $f$  **streng konkav**.*

### **Hinreichende Bedingung für lokale Extrema.**

Mit diesen Vorbereitungen kann nun eine hinreichende Bedingung für die Existenz von Extrema formuliert werden. Wieder ist die geometrische Vorstellung in Abbildung 8.9 wiedergegeben:

Strenge Konvexität (Konkavität) in der Nähe eines kritischen Punktes impliziert die Existenz eines Minimums (Maximums).

**Satz 8.8.** *Es sei  $f: (a, b) \rightarrow \mathbb{R}$  zweimal differenzierbar ( $f''$  sei zudem eine stetige Funktion) und in  $x_0 \in (a, b)$  gelte  $f'(x_0) = 0$ .*

*i) Ist  $f''(x_0) > 0$ , so hat  $f$  in  $x_0$  ein **strenges lokales Minimum**.*

*ii) Ist  $f''(x_0) < 0$ , so hat  $f$  in  $x_0$  ein **strenges lokales Maximum**.*

**Bemerkung.** Der Satz macht **keine Aussage** über den Fall  $f''(x_0) = 0$ .

**Beispiel.** Im obigen Beispiel  $f(x) = (1 - x^2)^2$  ist

$$f''(x) = 12x^2 - 4.$$

Für die kritischen Punkte  $x_1$  und  $x_3$  gilt also  $f''(x_1) > 0$ ,  $f''(x_3) > 0$  und wie bereits gesehen handelt es sich um (strenge) lokale Minima.

Im Punkt  $x_2 = 0$  ist  $f''(x_2) < 0$ , d.h. es handelt sich um ein (strenges) lokales Maximum.

## Wendepunkte.

Neben Extrema sind Wendepunkte charakteristisch für eine Funktion. In einem Wendepunkt geht der Graph der Funktion von **einer Linkskurve in eine Rechtskurve über** (oder umgekehrt), d.h. von einem **konvexen in einen konkaven Bereich** (oder umgekehrt). Die Situation ist in Abbildung 8.11 angedeutet.

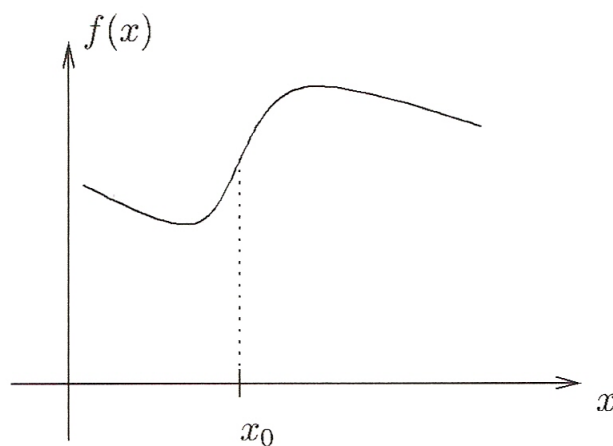


Abbildung 8.11: Ein Wendepunkt.

Es sei nun  $f: (a, b) \rightarrow \mathbb{R}$  dreimal differenzierbar ( $f'''$  sei zudem eine stetige Funktion). Ganz analog zur Diskussion von Extrema folgt:

- i) In einem Wendepunkt  $x_0 \in (a, b)$  gilt immer  $f''(x_0) = 0$ .
- ii) Ist  $f''(x_0) = 0$  und  $f'''(x_0) \neq 0$ , so ist  $x_0$  ein Wendepunkt.

## Die Regeln von l'Hospital.

Zum Abschluss dieses Kapitels 8.2 soll noch kurz auf eine wichtige Folgerung aus dem (verallgemeinerten) Mittelwertsatz eingegangen werden, mit der Grenzwerte von Quotienten berechnet werden können, die sich nicht aus den bisher bekannten Regeln erschließen.

Beispielsweise kann

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x}$$

berechnet werden mit Hilfe von

**Satz 8.9.** (*Regeln von l'Hospital*) Es seien  $f, g: (a, b) \rightarrow \mathbb{R}$  differenzierbar und  $x_0 \in (a, b)$ . Ist  $f(x_0) = g(x_0) = 0$  und ist  $g'(x_0) \neq 0$  für  $x \neq x_0$ , so gilt

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

falls der zweite Grenzwert existiert.

**Beispiele.**

i) Es ist

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = 1.$$

ii) Die Regel kann auch **mehrfach hintereinander angewandt werden**, so gilt

$$\lim_{x \rightarrow x_0} \frac{1 - \cos(x)}{x^2} = \lim_{x \rightarrow x_0} \frac{\sin(x)}{2x} = \lim_{x \rightarrow x_0} \frac{\cos(x)}{2} = \frac{1}{2}.$$

iii) Eine analoge Aussage gilt auch für **Grenzwerte " $x \rightarrow \infty$ "**, falls  $f$  und  $g$  im Unendlichen gegen Null konvergieren, z.B.

$$\begin{aligned} \lim_{x \rightarrow \infty} [\sqrt{x(x+1)} - x] &= \lim_{x \rightarrow \infty} \frac{\sqrt{1 + \frac{1}{x}} - 1}{\frac{1}{x}} \\ &= \lim_{x \rightarrow \infty} \frac{\frac{1}{2\sqrt{1 + \frac{1}{x}}} \left[ -\frac{1}{x^2} \right]}{\left[ -\frac{1}{x^2} \right]} \\ &= \lim_{x \rightarrow \infty} \frac{1}{2\sqrt{1 + \frac{1}{x}}} = \frac{1}{2}. \end{aligned}$$

iv) Ebenso gelten die Regeln für **Grenzwerte der Form**  $x \rightarrow x_0$ ,  $f(x)$ ,  $g(x) \rightarrow \pm\infty$ . Wie überall sind dabei auch einseitige Grenzwerte zugelassen:

$$\lim_{x \rightarrow 0^+} x \ln(x) = \lim_{x \rightarrow 0^+} \frac{\ln(x)}{\frac{1}{x}} = \lim_{x \rightarrow 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \rightarrow 0^+} (-x) = 0 .$$

### 8.3 Der Satz von Taylor

Die geometrische Interpretation der Ableitung ist es, eine Funktion nahe eines fixierten Punktes mit einer (affin) linearen Funktion zu approximieren.

In diesem Abschnitt soll eine Funktion in “**höherer Ordnung**”, d.h. mit Polynomen eines festen Grades, evtl. besser approximiert werden.

Die entscheidende Frage bei der Approximation bzw. beim Übergang zu einer Potenzreihe ist dabei die nach der **Größe des Fehlers**.

**Beobachtung.** Betrachtet sei ein Polynom vom Grad  $n$ ,

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n .$$

Für  $k = 0, 1, \dots, n$  ist die  $k$ -te Ableitung

$$p^{(k)}(0) = k!a_k ,$$

mit anderen Worten: Es gilt

$$p(x) = \sum_{k=0}^n \frac{1}{k!} p^{(k)}(0) x^k .$$

Zu beliebigem  $x_0 \in \mathbb{R}$  kann das Polynom ebenso geschrieben werden in der Form

$$p(x) = \sum_{k=0}^n b_k (x - x_0)^k ,$$

und völlig analog zu oben gilt

$$p(x) = \sum_{k=0}^n \frac{1}{k!} p^{(k)}(x_0) (x - x_0)^k .$$

## Können “gutartige” Funktionen in ähnlicher Weise zumindest approximiert werden?

**Definition 8.4.** Es sei  $f: [a, b] \rightarrow \mathbb{R}$   $n$ -mal differenzierbar (mit stetigen Ableitungen) und  $x_0 \in (a, b)$ . Dann heißt

$$T_n(x; x_0) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

das *Taylor-Polynom  $n$ -ten Grades zum Entwicklungspunkt  $x_0$* . Die *Taylor'sche Formel* lautet

$$f(x) = T_n(x; x_0) + R_n(x - x_0) ,$$

wobei das *Restglied*  $R_n(x - x_0)$  im Folgenden zu quantifizieren ist.

### Bemerkungen.

- i) Das Taylor-Polynom ersten Grades gibt genau die Tangente an den Graphen.
- ii) Die Ableitungen des Taylor-Polynoms stimmen an der Stelle  $x_0$  bis zur Ordnung  $n$  mit denen von  $f$  überein.

**Beispiel.** Es sei  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = \sin(x)$ ,  $x_0 = 0$ . Dann ist

$$\begin{aligned} f^{(0)}(0) &= 0 , & f^{(1)}(0) &= 1 , & f^{(2)}(0) &= 0 , \\ f^{(3)}(0) &= -1 , & f^{(4)}(0) &= 0 , & f^{(5)}(0) &= 1 . \end{aligned}$$

Für das Taylor-Polynom vom Grad 5 ergibt sich (vgl. **Reihendarstellung des Sinus** aus Kapitel 6)

$$T_n(x; 0) = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 .$$

Anhand von Abbildung 8.12 erkennt man, dass das Polynom  $T_n(x; 0)$  den Sinus schon in einem “weiten Bereich” um die 0 recht gut approximiert.

Allgemein gibt eine Antwort auf die Güte der Approximation der



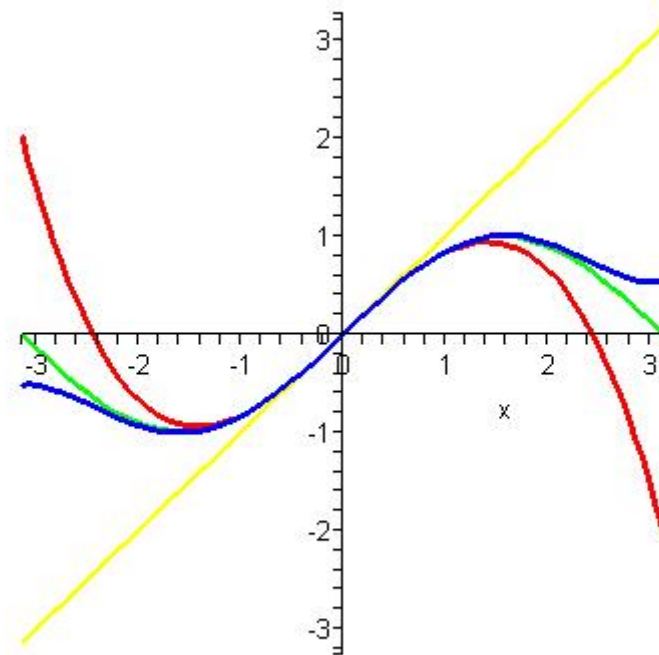


Abbildung 8.12: Der Sinus und seine Taylor-Polynome der Ordnung 1, 3, 5.

**Satz 8.10.** Die Funktion  $f: [a, b] \rightarrow \mathbb{R}$  sei  $(n + 1)$ -mal differenzierbar (mit stetigen Ableitungen) und es seien  $x_0, x \in (a, b)$ .

Dann gibt es eine von  $x$  abhängige Zahl  $\xi = \xi(x)$ , die zwischen  $x$  und  $x_0$  liegt, sodass

$$R_n(x - x_0) = \frac{f^{(n+1)}(\xi)}{(n + 1)!} (x - x_0)^{n+1}.$$

**Bemerkung.** Die Zahl  $\xi = \xi(x)$  hängt zwar auch von  $x_0$  ab, da in der Regel  $x_0$  aber fixiert ist und  $x$  variiert, wird diese Abhängigkeit in der Notation weggelassen.

### Beispiele.

i) Es sei  $f(x) = e^x$  für alle  $x \in \mathbb{R}$ . Dann ist

$$T_2(x; 0) = 1 + x + \frac{x^2}{2}$$

mit dem Restglied

$$R_2(x - 0) = \frac{e^\xi x^3}{6} .$$

Im Intervall  $[-1, 1]$  ist der Fehler beispielsweise abgeschätzt durch

$$|R_2(x)| \leq \left| \frac{e^\xi x^3}{6} \right| \leq \frac{e|x^3|}{6} .$$

ii) (**Cauchys Beispiel**, vgl. Abbildung 8.13) Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  sei gegeben durch

$$f(x) = \begin{cases} e^{-1/x^2} & \text{für } x \neq 0 , \\ 0 & \text{für } x = 0 . \end{cases}$$

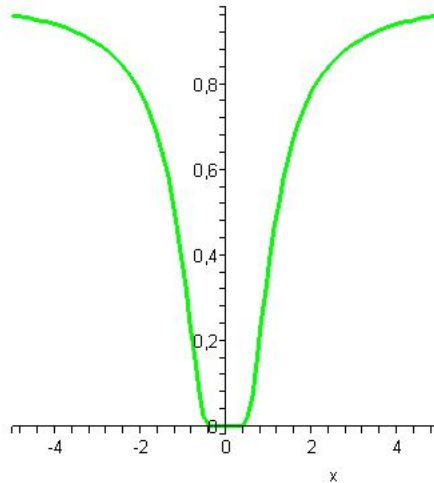


Abbildung 8.13: Die Funktion  $e^{-1/x^2}$  läuft so “flach” in den Nullpunkt, dass alle Taylor-Polynome um den Entwicklungspunkt 0 identisch verschwinden.

Man kann nachrechnen, dass die Funktion im Punkt  $x_0 = 0$  **beliebig oft differenzierbar ist** und dass alle Ableitungen im Nullpunkt verschwinden. Es folgt

$$T_n(x; 0) \equiv 0 ,$$

d.h. für alle  $n \in \mathbb{N}$  und für alle  $x \in \mathbb{R}$  gilt

$$R_n(x) = f(x) .$$

Auch im Grenzwert  $n \rightarrow \infty$  wird der Fehler für fixiertes  $x \neq 0$  also nicht klein.

## Wann kann zu einer Reihe übergegangen werden?

Cauchys Beispiel belegt, dass eine Funktion  $f$  im Allgemeinen **nicht als Grenzwert  $n \rightarrow \infty$**  der Taylor-Polynome geschrieben werden kann. Mit anderen Worten: Die Funktion  $f$  wird im Allgemeinen nicht durch ihre **Taylor-Reihe**

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

dargestellt.

Im Wesentlichen gibt es zwei offene Fragen:

- i) **Konvergiert die Reihe überhaupt?**
- ii) **Wenn ja, konvergiert sie gegen die Funktion  $f$ ?**

Das Kriterium zu einer positiven Antwort auf diese Fragen ist

$$\lim_{n \rightarrow \infty} R_n(x - x_0) = 0 .$$

Funktionen mit dieser Eigenschaft, d.h. Funktionen, die durch ihre (konvergente) Taylorsche Reihe dargestellt werden, heißen **reell analytische Funktionen**.

Zur Klasse der reell analytischen Funktionen gehören beispielsweise die Exponentialfunktion, die trigonometrischen Funktionen und die Hyperbelfunktionen, deren Taylor-Reihe um den Ursprung mit der definierenden Potenzreihe übereinstimmen.

Anhand von geeigneten **Wachstumsbedingungen für die Ableitungen einer Funktion** können hinreichende Bedingungen für reell analytische Funktionen angegeben werden, worauf hier aber nicht näher eingegangen wird.



# Kapitel 9

## Integralrechnung in einer Veränderlichen

### 9.1 Das bestimmte Riemannsches Integral

Im Folgenden ist stets  $I = [a, b]$  und  $f: I \rightarrow \mathbb{R}$  eine beschränkte Funktion.

**Idee zur Einführung des bestimmten Integrals:**

Das bestimmte Integral  $\int_a^b f(x) \, dx$  ist der Flächeninhalt der vom Graphen von  $f$  und der  $x$ -Achse eingeschlossenen Punktmenge  $F$ , so wie es in Abbildung 9.1 angedeutet ist.

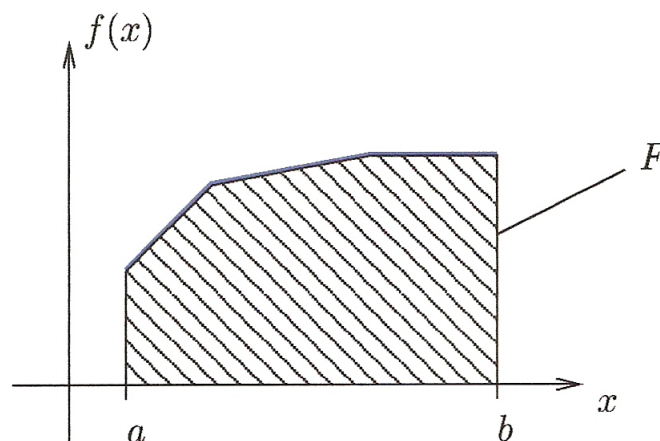


Abbildung 9.1: Zur suggestiven Vorstellung des bestimmten Integrals.

**Problem.** Das Problem bei dieser suggestiven Vorstellung ist aber:  
 Wie ist der Flächeninhalt überhaupt definiert, wenn es sich nicht um ein elementargeometrisches Objekt handelt?

**Vorgehensweise.** In der Tat sieht die Vorgehensweise wie folgt aus:

- i) Man definiere zunächst das bestimmte Integral.
- ii) Mithilfe des bestimmten Integrals können Flächeninhalte definiert werden.
- iii) Dabei soll “einfachen” geometrischen Objekten (Rechtecken, Dreiecken etc.) der Flächeninhalt zugeordnet werden, der mit der Elementargeometrie konsistent ist.
- iv) Um dies zu erreichen, approximiert man bei der Definition des Integrals die Menge  $F$  mit Rechtecken, denen elementargeometrisch ein Flächeninhalt zugeordnet werden kann.
- v) Schließlich ist zu untersuchen, ob zum Grenzwert übergegangen werden kann.

Konkret wird definiert:

**Definition 9.1.** i) Eine *Zerlegung*  $\mathcal{Z}$  des Intervalls  $I = [a, b]$  in Teilintervalle  $I_j$ ,  $j = 1, \dots, n$ , der Länge  $|I_j|$  ist eine *Menge von Punkten*

$$\mathcal{Z} = \{a = x_0 < x_1 < x_2 < \dots < x_n = b\} .$$

Hierbei ist  $I_j := [x_{j-1}, x_j]$ ,  $\Delta x_j := x_j - x_{j-1} = |I_j|$ ,  $j = 1, \dots, n$ , und

$$\Delta(\mathcal{Z}) := \max\{\Delta x_1, \dots, \Delta x_n\}$$

heißt die *Feinheit der Zerlegung*  $\mathcal{Z}$ .

ii) Ist für  $j = 1, \dots, n$

$$\begin{aligned} \underline{m}_j &:= \inf\{f(x) : x \in I_j\} , \\ \overline{m}_j &:= \sup\{f(x) : x \in I_j\} , \end{aligned}$$

so heißt

$$\underline{S}_{\mathcal{Z}}(f) := \sum_{j=1}^n \underline{m}_j \Delta x_j$$

die *Untersumme*,

$$\overline{S}_{\mathcal{Z}}(f) := \sum_{j=1}^n \overline{m}_j \Delta x_j$$

die *Obersumme* von  $f$  zur Zerlegung  $\mathcal{Z}$ .

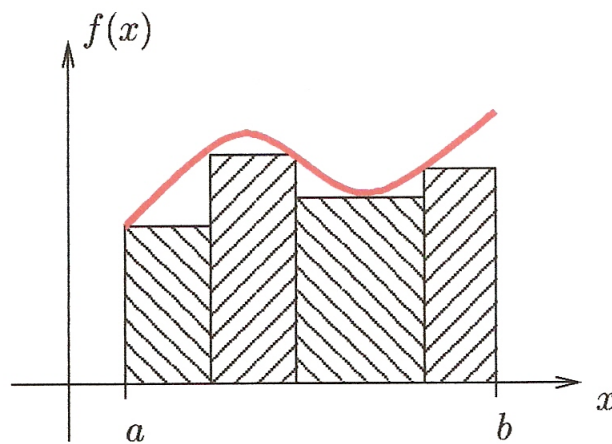


Abbildung 9.2: Eine Untersumme von  $f$ .

**Bemerkung.** Nach Definition gilt für jede Zerlegung  $\mathcal{Z}$

$$\underline{S}_{\mathcal{Z}}(f) \leq \overline{S}_{\mathcal{Z}}(f) .$$

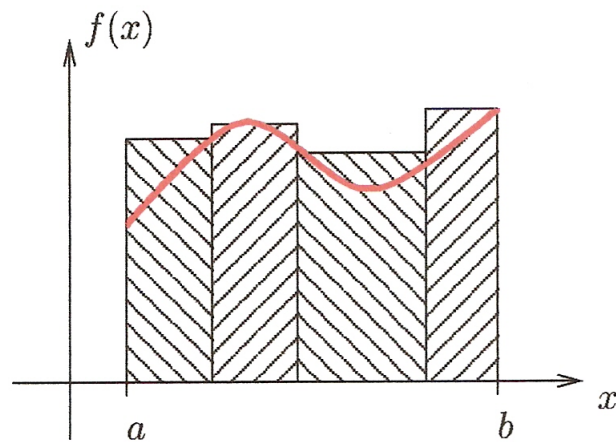
Nun definiert man:

**Definition 9.2.** Ist  $f: I = [a, b] \rightarrow \mathbb{R}$  beschränkt, so sind das *Unterin-  
tegral*  $\underline{\mathcal{I}}$  und das *Oberintegral*  $\overline{\mathcal{I}}$  definiert als

$$\begin{aligned} \underline{\mathcal{I}}(f) &:= \sup\{\underline{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } I\} , \\ \overline{\mathcal{I}}(f) &:= \inf\{\overline{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } I\} . \end{aligned}$$

Die Funktion heißt *(Riemann) integrierbar* auf  $I$ , falls gilt

$$\underline{\mathcal{I}}(f) = \overline{\mathcal{I}}(f) =: \mathcal{I}(f) =: \int_I f(x) \, dx =: \int_a^b f(x) \, dx .$$

Abbildung 9.3: Eine Obersumme von  $f$ .**Bemerkungen.**

i) Für eine beliebige Zerlegung  $\mathcal{Z}$  von  $I$  gilt

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{I}(f) \leq \overline{I}(f) \leq \overline{S}_{\mathcal{Z}}(f) .$$

ii) Die Klasse der auf  $I$  integrierbaren Funktionen wird mit  $\mathcal{R}(I)$  bezeichnet.

iii) Es heißt  $\int_a^b f(x) dx$  das bestimmte Integral von  $f$  zwischen den Grenzen  $a$  und  $b$ .

**Beispiele.**

i) Für eine Konstante  $c \in \mathbb{R}$  sei  $f(x) = c$  auf  $[a, b]$ . Es folgt für jede Zerlegung  $\mathcal{Z}$  von  $I$

$$\underline{S}_{\mathcal{Z}}(f) = \sum_{j=1}^n c \Delta x_j = c(b-a) = \overline{S}_{\mathcal{Z}}(f) .$$

Somit ist  $f$  auf  $[a, b]$  integrierbar mit

$$\int_a^b c dx = c(b-a) .$$



ii) Es sei  $I = [0, 1]$  und  $f(x) = x$  auf  $I$ . Zu  $n \in \mathbb{N}$  betrachte man die Zerlegung

$$\mathcal{Z}_n = \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, 1 \right\}.$$

Für die Ober- bzw. Untersummen gilt

$$\begin{aligned} \underline{S}_{\mathcal{Z}_n}(f) &= \sum_{j=1}^n \frac{j-1}{n} \left[ \frac{j}{n} - \frac{j-1}{n} \right] = \frac{1}{n^2} \sum_{j=1}^n (j-1) \\ &= \frac{1}{n^2} \frac{n(n-1)}{2} = \frac{1}{2} - \frac{1}{2n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}, \\ \bar{S}_{\mathcal{Z}_n}(f) &= \frac{1}{2} + \frac{1}{2n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}. \end{aligned}$$

Da hier nur eine spezielle Zerlegungsfolge betrachtet wird, ist die Integrierbarkeit von  $f$  nach wie vor unklar.

Man benötigt noch:

**Satz 9.1.** Für eine beschränkte Funktion  $f: I \rightarrow \mathbb{R}$  gilt

$$f \in \mathcal{R}(I) \Leftrightarrow \begin{cases} \text{Zu jedem } \varepsilon > 0 \text{ gibt es eine} \\ \text{Zerlegung } \mathcal{Z} \text{ von } I \text{ mit} \\ \bar{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) < \varepsilon. \end{cases}$$

**Konsequenz.** Die Funktion  $f(x) = x$  ist auf  $I = [0, 1]$  integrierbar mit

$$\int_0^1 x \, dx = \frac{1}{2}.$$

### Rechenregeln für integrierbare Funktionen.

**Satz 9.2.** Es seien  $f, g \in \mathcal{R}(I)$  und  $\alpha, \beta \in \mathbb{R}$ . Dann gilt:

i)  $\alpha f + \beta g \in \mathcal{R}(I)$  mit

$$\mathcal{I}(\alpha f + \beta g) = \alpha \mathcal{I}(f) + \beta \mathcal{I}(g).$$

ii)  $fg \in \mathcal{R}(I)$ .

iii)  $|f| \in \mathcal{R}(I)$ , wobei  $|f|(x) := |f(x)|$ .

iv) Ist  $|g| \geq c$  für eine Konstante  $c > 0$ , so ist auch  $f/g \in \mathcal{R}(I)$ .

v) Ist für alle  $x \in I$  die Ungleichung  $f(x) \leq g(x)$  richtig, so folgt

$$\mathcal{I}(f) \leq \mathcal{I}(g) .$$

vi)  $|\mathcal{I}(f)| \leq \mathcal{I}(|f|)$ .

Aus Satz 9.2 folgt insbesondere die Integrierbarkeit von Polynomen auf  $[a, b]$ .

Die Frage nach weiteren bekannten Funktionenklassen in  $\mathcal{R}(I)$  beantwortet

**Satz 9.3.** Für eine beschränkte Funktion  $f: [a, b] \rightarrow \mathbb{R}$  gilt:

i) Ist  $f$  *monoton*, so folgt  $f \in \mathcal{R}(I)$ .

ii) Ist  $f$  *stetig*, so folgt  $f \in \mathcal{R}(I)$ .

**Bemerkung.** Eine Funktion  $f \in \mathcal{R}(I)$  kann auch über Teilintervalle integriert werden im Sinne von:

i)  $f$  ist auch auf jedem Teilintervall  $I' \subset I$  integrierbar.

ii) Ist  $I$  in endlich viele Teilintervalle  $I_1, I_2, \dots, I_n$  zerlegt, die höchstens Randpunkte gemeinsam haben, so gilt

$$\int_I f(x) \, dx = \sum_{j=1}^n \int_{I_j} f(x) \, dx .$$

## 9.2 Der Hauptsatz der Differential- und Integralrechnung

Der [Hauptsatz der Differential- und Integralrechnung](#) ist das Bindeglied zwischen diesen beiden Disziplinen und gleichzeitig ein starkes

Instrument, um Integrale analytisch auszurechnen.

Die Idee des Satzes sei hier anhand eines einfachen Beispiels erläutert:

Man betrachte für  $x \geq 0$  die Funktion  $f(x) = \alpha x$ ,  $\alpha > 0$ , den Flächeninhalt ( $x_0 > 0$ )

$$A(x_0) = \int_0^{x_0} f(x) \, dx$$

und ebenso für  $h > 0$

$$A(x_0 + h) = \int_0^{x_0+h} f(x) \, dx .$$

Untersucht werden soll die relative Änderung “ $\Delta A/\Delta x$ ”, d.h. gesucht ist der Differenzenquotient von  $A$  an der Stelle  $x_0$ ,

$$\frac{1}{h} \left[ \int_0^{x_0+h} f(x) \, dx - \int_0^{x_0} f(x) \, dx \right] .$$

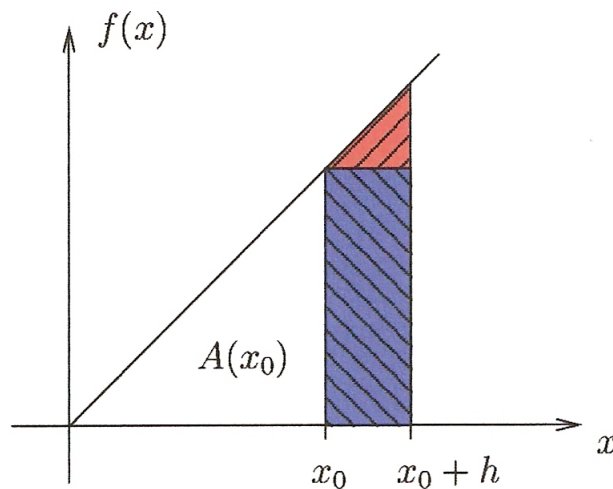


Abbildung 9.4: Zum Hauptsatz der Differential- und Integralrechnung.

Die Differenz  $A(x_0 + h) - A(x_0)$  setzt sich zusammen aus dem Flächeninhalt des in Abbildung 9.4 blau gekennzeichneten Rechtecks  $R$  und dem des roten Dreiecks  $D$ . Der Flächeninhalt von  $R$  ist  $\alpha x_0 h$ , der von  $D$  ist  $\alpha h^2/2$ .

Es ergibt sich

$$\frac{1}{h} \left[ \int_0^{x_0+h} f(x) \, dx - \int_0^{x_0} f(x) \, dx \right] = \frac{1}{h} \left[ \alpha x_0 h + \frac{\alpha h^2}{2} \right]$$

$$\xrightarrow{h \rightarrow 0} \alpha x_0 = f(x_0).$$

Die Ableitung des Flächeninhalts  $A(x_0)$  nach  $x_0$  entspricht dem Funktionswert  $f(x_0)$ .

Präzise sieht die Situation wie folgt aus:

**Definition 9.3.** Es seien  $F, f: [a, b] \rightarrow \mathbb{R}$ . Ist  $F$  differenzierbar auf  $[a, b]$  und gilt  $F'(x) = f(x)$  für alle  $x \in [a, b]$ , so heißt  $F$  eine *Stammfunktion* von  $f$ .

### Bemerkungen.

- i) Mit  $F$  sind auch alle Funktionen  $F_C = F(x) + C$ ,  $C \in \mathbb{R}$ , Stammfunktionen von  $f$ .
- ii) Die Differenz zweier Stammfunktionen einer Funktion  $f$  ist konstant.

**Satz 9.4.** (Hauptsatz der Differential- und Integralrechnung) Die Funktion  $f: I = [a, b] \rightarrow \mathbb{R}$  sei stetig.

Dann gilt:

- i) Die Funktion

$$F(x) := \int_a^x f(t) \, dt$$

ist eine *Stammfunktion von  $f(x)$* .

- ii) Ist  $F(x) \in C^1(I)$  eine Stammfunktion von  $f(x)$ , so gilt

$$\int_a^b f(t) \, dt = F(b) - F(a) =: \left[ F(x) \right]_a^b =: F(x) \Big|_a^b.$$

**Bemerkungen.**

- i) Die **Gesamtheit aller Stammfunktionen** einer stetigen Funktion  $f$  heißt das **unbestimmte Integral** der Funktion  $f$ .

Notation:

$$\{F : F \text{ ist Stammfunktion von } f\} =: \int f(x) \, dx =: F(x) + C .$$

- ii) Das Aufsuchen einer Stammfunktion wird als **Integration** von  $f$  bezeichnet. Nach dem Hauptsatz macht die Differentiation die Integration wieder rückgängig.

**Beispiele.**

- i) Als ein einfaches Beispiel sei hier die Funktion  $f(x) = x^n$ ,  $n \in \mathbb{N}$ , betrachtet. Diese hat eine Stammfunktion

$$F(x) = \frac{x^{n+1}}{n+1} ,$$

woraus

$$\int_a^b x^n \, dx = \frac{1}{n+1} [b^{n+1} - a^{n+1}]$$

folgt.

- ii) Ohne Beweis sind in Tabelle 9.1 einige unbestimmte Integrale aufgelistet.

**9.3 Integrationstechniken**

In diesem Paragraphen werden einige Integrationstechniken angesprochen, die die Integration in konkreten Fällen ermöglichen.

**9.3.1 Einfache Integrationstechniken**

Es seien  $f, g$  stetig auf  $[a, b]$ . Dann gilt (auf  $[a, b]$ )

$$i) \int cf(x) \, dx = c \int f(x) \, dx , \quad c \in \mathbb{R} ,$$

$f(x)$	$\int f(x) dx$	gültig, falls
$x^k$	$\frac{1}{k+1}x^{k+1} + C$	$k \in \mathbb{Z}, k \neq -1, x \in \mathbb{R}$
$\frac{1}{x}$	$\ln( x ) + C$	$x \neq 0$
$x^\alpha$	$\frac{1}{\alpha+1}x^{\alpha+1} + C$	$\alpha \neq -1, x > 0$
$e^x$	$e^x + C$	$x \in \mathbb{R}$
$a^x$	$\frac{a^x}{\ln(a)} + C$	$x \in \mathbb{R}, 0 < a, a \neq 1$
$\sin(x)$	$-\cos(x) + C$	$x \in \mathbb{R}$
$\cos(x)$	$\sin(x) + C$	$x \in \mathbb{R}$
$\tan(x)$	$-\ln( \cos(x) ) + C$	$x \neq (2k+1)\pi/2, k \in \mathbb{Z}$
$\cot(x)$	$\ln( \sin(x) ) + C$	$x \neq k\pi, k \in \mathbb{Z}$
$\frac{1}{\cos^2(x)}$	$\tan(x) + C$	$x \neq (2k+1)\pi/2, k \in \mathbb{Z}$
$\frac{1}{\sin^2(x)}$	$-\cot(x) + C$	$x \neq k\pi, k \in \mathbb{Z}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x) + C$	$-1 < x < 1$
$\frac{1}{\sqrt{1-x^2}}$	$\arccos(x) + C$	$-1 < x < 1$
$\frac{1}{1+x^2}$	$\arctan(x) + C$	$x \in \mathbb{R}$

Tabelle 9.1: Einige unbestimmte Integrale.

$$ii) \int (f(x) + g(x)) \, dx = \int f(x) \, dx + \int g(x) \, dx,$$

$$iii) \int \frac{f'(x)}{f(x)} \, dx = \ln(|f(x)|) + C, \text{ falls } f \text{ keine Nullstellen hat,}$$

wobei in der letzten Zeile auch  $f'$  als stetig angenommen wurde.

### Beispiele.

i) Mit  $a_j \in \mathbb{R}$ ,  $j = 1, \dots, n$ , ist

$$\int \left[ \sum_{j=0}^n a_j x^j \right] \, dx = \sum_{j=0}^n a_j \int x^j \, dx = \sum_{j=0}^n a_j \frac{x^{j+1}}{j+1} + C .$$

ii) Es sei  $f(x) = \tan(x)$ ,  $x \in [a, b] \subset (-\pi/2, \pi/2)$ . In diesem Intervall hat der Kosinus keine Nullstelle und wegen  $\cos(x) > 0$  für alle  $x \in (-\pi/2, \pi/2)$  gilt

$$\begin{aligned} \int \tan(x) \, dx &= \int \frac{\sin(x)}{\cos(x)} \, dx = - \int \frac{(-\sin(x))}{\cos(x)} \, dx \\ &= - \ln(\cos(x)) + C . \end{aligned}$$

### 9.3.2 Partielle Integration

Hierbei handelt es sich um die **Folgerung aus der Produktregel**.

**Satz 9.5.** Sind  $f, g: [a, b] \rightarrow \mathbb{R}$  differenzierbar (mit stetigen Ableitungen), so gilt

$$\begin{aligned} \int f(x)g'(x) \, dx &= f(x)g(x) - \int f'(x)g(x) \, dx , \\ \int_a^b f(x)g'(x) \, dx &= \left[ f(x)g(x) \right]_a^b - \int_a^b f'(x)g(x) \, dx . \end{aligned}$$

### Beispiele.

i) Mit  $f(x) = x$  und  $g'(x) = e^x$  folgt

$$\int x e^x \, dx = x e^x - \int e^x \, dx = x e^x - e^x + C .$$

ii) Mit  $f(x) = \ln(x)$  und  $g'(x) = 1$  ergibt sich auf  $[a, b] \subset (0, \infty)$

$$\begin{aligned} \int \ln(x) \, dx &= \int \ln(x) 1 \, dx = \ln(x)x - \int \frac{1}{x} x \, dx \\ &= x \ln(x) - x + C . \end{aligned}$$

iii) Ist  $f(x) = g'(x) = \cos(x)$ , so sieht man

$$\begin{aligned} \int \cos^2(x) \, dx &= \int \cos(x) \cos(x) \, dx \\ &= \cos(x) \sin(x) + \int \sin^2(x) \, dx \\ &= \sin(x) \cos(x) + \int 1 \, dx - \int \cos^2(x) \, dx , \end{aligned}$$

also

$$\int \cos^2(x) \, dx = \frac{1}{2} [\sin(x) \cos(x) + x] + C .$$

### 9.3.3 Substitutionsregel

Die Regel zur Substitution ist eine **Folgerung aus der Kettenregel**.

**Satz 9.6.** *Es sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig und  $g: [c, d] \rightarrow [a, b]$  sei differenzierbar (mit stetiger Ableitung). Dann gilt für alle  $t_0, t_1 \in [c, d]$*

$$\int_{g(t_0)}^{g(t_1)} f(x) \, dx = \int_{t_0}^{t_1} f(g(t))g'(t) \, dt .$$

Für die Substitutionsregel gibt es **zwei Lesarten**:

#### **Erste Lesart.**

Hier ist  $\int_{t_0}^{t_1} h(t) \, dt$  gesucht, wobei angenommen sei ( $f, g$  wie oben)

$$h(t) = f(g(t))g'(t) .$$

Dann gilt

$$\int_{t_0}^{t_1} h(t) \, dt = \int_{g(t_0)}^{g(t_1)} f(x) \, dx .$$



Kann dabei die rechte Seite explizit berechnet werden, so ist auch die linke Seite bekannt.

**Beispiel.** Gesucht sei

$$\int_0^1 (1+t^2)^n t \, dt .$$

Die Funktion  $h(t) := (1+t^2)^n t$  lässt sich in der Form

$$h(t) = \frac{1}{2}(1+t^2)^n \frac{d}{dt}(1+t^2)$$

schreiben, also

$$2h(t) = f(g(t))g'(t)$$

mit  $f(x) := x^n$  und  $g(t) := 1+t^2$ .

Es folgt

$$\begin{aligned} \int_0^1 (1+t^2)^n t \, dt &= \frac{1}{2} \int_0^1 f(g(t))g'(t) \, dt \\ &= \frac{1}{2} \int_1^2 x^n \, dx \\ &= \frac{x^{n+1}}{2(n+1)} \Big|_1^2 = \frac{2^{n+1} - 1}{2(n+1)} . \end{aligned}$$

### Zweite Lesart.

Man lese die Substitutionsregel von links nach rechts, d.h. gesucht ist  $\int_{x_0}^{x_1} f(x) \, dx$  durch geeignete [Variablentransformation](#).

Mit anderen Worten ist eine Transformation

$$x = g(t) , \quad t_0 \leq t \leq t_1 ,$$

gesucht, die die Eigenschaft

$$x_0 = g(t_0) \quad \text{und} \quad x_1 = g(t_1)$$

haben muss.

Hierbei sei  $g: [t_0, t_1] \rightarrow [x_0, x_1]$  (o.E. gelte  $t_0 < t_1$  und  $x_0 < x_1$ ) differenzierbar (mit stetiger Ableitung) und bijektiv (etwa  $g'(t) > 0$

für alle  $t \in [t_0, t_1]$ ).

Bezeichnet dann  $\psi := g^{-1}$  die Inverse von  $g$ , so gilt mit  $t_0 = \psi(x_0)$  und  $t_1 = \psi(x_1)$

$$\int_{x_0}^{x_1} f(x) \, dx = \int_{\psi(x_0)}^{\psi(x_1)} f(g(t))g'(t) \, dt .$$

**Bemerkung.** Das **formale Vorgehen** sieht wie folgt aus: Man substituiert  $x = g(t)$  in  $\int_{x_0}^{x_1} f(x) \, dx$  und schreibt in der **Leibnizschen Weise**

$$g'(t) = \frac{dx}{dt} , \quad \text{also} \quad dx = g'(t) \, dt .$$

Die untere Integrationsgrenze  $x_0 = g(t_0)$  wird zu  $t_0 = \psi(x_0)$ , die obere transformiert sich analog.

**Beispiel.** Gesucht sei zu fixiertem  $r > 0$

$$\int_0^r \sqrt{r^2 - x^2} \, dx .$$

Es bietet sich die Substitution

$$x = g(t) := r \sin(t) , \quad 0 \leq t \leq \frac{\pi}{2} ,$$

an, wobei  $g'(t) = r \cos(t)$ . Man beachtet ( $\psi = g^{-1}$ )

$$0 = g(0) , \quad r = g(\pi/2) , \quad \text{d.h.} \quad \psi(0) = 0 , \quad \psi(r) = \frac{\pi}{2} .$$

Es folgt

$$\begin{aligned} \int_0^r \sqrt{r^2 - x^2} \, dx &= \int_0^{\pi/2} \sqrt{r^2 - r^2 \sin^2(t)} \, r \cos(t) \, dt \\ &= r^2 \int_0^{\pi/2} \cos^2(t) \, dt = \frac{\pi r^2}{4} . \end{aligned}$$

### 9.3.4 Partialbruchzerlegung

Mithilfe einer sogenannten **Partialbruchzerlegung** können gebrochen-rationale Funktionen (vgl. Abschnitt 3.4.1)

$$f(x) = \frac{p(x)}{q(x)} , \quad p, q \text{ Polynome vom Grad } m \text{ bzw. } n ,$$

aufintegriert werden.

Die präzise Darstellung aller möglichen Fälle erfordert eine etwas technische Notation. Zum prinzipiellen Verständnis genügt es aber bereits, sich von den folgenden Beispielen leiten zu lassen.

### Falls nötig: Polynomdivision.

Wie in Kapitel 3.4.1 kann ohne Beschränkung der Allgemeinheit angenommen werden, dass

$$\text{grad } p < \text{grad } q ,$$

denn ansonsten wird eine Polynomdivision vorgeschaltet.

**Beispiel.** Ist

$$f(x) = \frac{x^3 + 3x^2 - 4x + 2}{x^3 - x^2 - x + 1} ,$$

so liefert die Rechnung

$$f(x) = \underbrace{1}_{r(x)} + \underbrace{\frac{4x^2 - 3x + 1}{x^3 - x^2 - x + 1}}_{p(x)/q(x)} ,$$

es ist also

$$f(x) = r(x) + \frac{p(x)}{q(x)} ,$$

wobei das Polynom  $r(x)$  elementar zu integrieren ist und wobei  $\text{grad } p < \text{grad } q$  gilt.

### Nullstellen von $q$ mit Vielfachheit 1.

**Beispiel.** Es sei

$$\frac{p(x)}{q(x)} = \frac{x}{2x^2 - 6x + 4} .$$

Zunächst sucht man die Nullstellen  $x_1, x_2$  des Nennerpolynoms  $q$ :  $x_1 = 1, x_2 = 2$ .

Beide Nullstellen sind verschieden und haben die **Vielfachheit 1**, d.h.  $q(x)$  kann als

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

in **Linearfaktoren** zerlegt werden.

In dieser Situation macht man den **Ansatz**:

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

Dieser liefert

$$\begin{aligned} \frac{p(x)}{q(x)} &= \frac{x}{2(x - 1)(x - 2)} \\ &= \frac{2A(x - 2) + 2B(x - 1)}{2(x - 1)(x - 2)} \\ &= \frac{x(2A + 2B) + (-4A - 2B)}{2(x - 1)(x - 2)} . \end{aligned}$$

Ein **Koeffizientenvergleich** zeigt

$$\begin{aligned} 2A + 2B &= 1 , \\ -4A - 2B &= 0 , \end{aligned}$$

mit der Lösung (Probe!)

$$A = -\frac{1}{2} , \quad B = 1 .$$

Insgesamt ist

$$\begin{aligned} \int \frac{x}{2x^2 - 6x + 4} dx &= -\frac{1}{2} \int \frac{1}{x - 1} dx + \int \frac{1}{x - 2} dx \\ &= -\frac{1}{2} \ln|x - 1| + \ln(|x - 2|) + C . \end{aligned}$$

**Nullstellen von  $q$  mit Vielfachheit 2 und größer.**

**Beispiel.** Es sei

$$\frac{p(x)}{q(x)} = \frac{x^2 + 1}{x^3 - 2x^2 + x} .$$

Zunächst werden wieder die Nullstellen  $x_1, x_2, x_3$  des Nennerpolynoms  $q$  gesucht:  $x_1 = 0, x_2 = 1, x_3 = 1$ .

Die Nullstelle  $x_1 = 0$  hat die Vielfachheit **1**. Die Nullstellen  $x_2 = x_3 = 1$  sind gleich, d.h. 1 ist Nullstelle der Vielfachheit **2**:  $q(x)$  wird als

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

in Linearfaktoren zerlegt.

Nun macht man den Ansatz:

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

Dieser liefert

$$\begin{aligned} \frac{p(x)}{q(x)} &= \frac{x^2 + 1}{x(x - 1)^2} \\ &= \frac{A(x - 1)^2 + Bx(x - 1) + Cx}{x(x - 1)^2} \\ &= \frac{x^2(A + B) + x(-2A - B + C) + A}{x(x - 1)^2} . \end{aligned}$$

Aus einem Koeffizientenvergleich folgt

$$\begin{aligned} A + B &= 1 , \\ -2A - B + C &= 0 , \\ A &= 1 \end{aligned}$$

mit der Lösung (Probe!)

$$A = 1 , \quad B = 0 , \quad C = 2 .$$

Insgesamt ist

$$\int \frac{x^2 + 1}{x^3 - 2x^2 + x} dx = \ln |x| - \frac{2}{x-1} + C .$$

### Irreduzible quadratische Polynome.

**Beispiel.** Es sei

$$\frac{p(x)}{q(x)} = \frac{1}{x^3 - x^2 + x - 1} .$$

Bei der Suche nach Nullstellen des Nennerpolynoms findet man nur die eine Nullstelle  $x_1 = 1$  der Vielfachheit 1.

Das Polynom **zerfällt nicht in Linearfaktoren** und kann lediglich in der Form

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

zerlegt werden, wobei  $x^2 + 1$  irreduzibel ist, d.h. keine reelle Nullstellen hat.

Der Ansatz lautet in diesem Fall

$$\frac{p(x)}{q(x)} = \frac{A}{x-1} + \frac{B+Cx}{x^2+1}$$

und führt auf

$$\begin{aligned} \frac{p(x)}{q(x)} &= \frac{1}{(x-1)(x^2+1)} \\ &= \frac{A(x^2+1) + B(x-1) + Cx(x-1)}{(x-1)(x^2+1)} \\ &= \frac{x^2(A+C) + x(B-C) + (A-B)}{(x-1)(x^2+1)} . \end{aligned}$$

Hier ergibt ein Koeffizientenvergleich

$$\begin{aligned} A + C &= 0 , \\ B - C &= 0 , \\ A - B &= 1 \end{aligned}$$

mit der Lösung (Probe!)

$$A = \frac{1}{2}, \quad B = C = -\frac{1}{2}.$$

Man berechnet damit

$$\int \frac{1}{x^3 - x^2 + x - 1} dx = \frac{1}{2} \ln|x - 1| - \frac{1}{4} \ln(|x^2 + 1|) - \frac{1}{2} \arctan(x) + C.$$

## 9.4 Uneigentliche Integrale

Bisher wurden für die Integration stets zwei Einschränkungen angenommen:

- i) Das Integrationsintervall  $I = [a, b]$  beschränkt (und abgeschlossen).
- ii) Auch die Funktion  $f: I \rightarrow \mathbb{R}$  ist beschränkt.

In diesem Abschnitt geht es darum, ob auch Integrale wie

$$\int_0^{\infty} e^{-x^2} dx \quad \text{oder} \quad \int_0^1 \frac{dx}{\sqrt{1-x^2}}$$

sinnvoll definiert werden können.

### Fall I. Unbeschränktes Integrationsintervall.

Es sei  $I = [a, \infty)$  und  $f \in \mathcal{R}([a, b])$  für alle  $b$  mit  $a < b < \infty$ . Die Funktion  $f$  heißt dann **lokal integrierbar**.

Für alle  $b \in I$  existiere also das bestimmte Integral

$$\int_a^b f(x) dx.$$

**Definition 9.4.** i) Falls der Grenzwert

$$\lim_{b \rightarrow \infty} \int_a^b f(x) dx$$

existiert, so heißt dieser das *uneigentliche Integral* von  $f$  über  $[a, \infty)$ .

Speichweise: Das uneigentliche Integral existiert oder *konvergiert*.

Notation:

$$\int_a^\infty f(x) \, dx := \lim_{b \rightarrow \infty} \int_a^b f(x) \, dx .$$

Andernfalls heißt  $\int_a^\infty f(x) \, dx$  *divergent*.

ii) Das uneigentliche Integral  $\int_a^\infty f(x) \, dx$  heißt *absolut konvergent*, falls  $\int_a^\infty |f(x)| \, dx$  konvergiert.

Absolute Konvergenz impliziert Konvergenz des uneigentlichen Integrals.

### Bemerkungen.

i) Uneigentliche Integrale der Form

$$\int_{-\infty}^b f(x) \, dx$$

sind analog zu diskutieren.

ii) Ein uneigentliches Integral der Form

$$\int_{-\infty}^\infty f(x) \, dx$$

konvergiert, wenn für ein beliebiges  $a \in \mathbb{R}$  sowohl  $\int_a^\infty f(x) \, dx$  als auch  $\int_{-\infty}^a f(x) \, dx$  konvergiert.

Es ist dann

$$\int_{-\infty}^\infty f(x) \, dx := \int_a^\infty f(x) \, dx + \int_{-\infty}^a f(x) \, dx .$$

Existiert eines der Integrale  $\int_a^\infty f(x) \, dx$ ,  $\int_{-\infty}^a f(x) \, dx$  nicht, so kann trotzdem der sogenannte *Cauchysche Hauptwert* existieren (vgl. Übungen).



Wie üblich ist die Konvergenz mit Hilfe geeigneter Kriterien zu überprüfen. Von besonderer Bedeutung ist

**Satz 9.7.** (*Majoranten- Minorantenkriterium*)

Es sei  $f: I = [a, \infty) \rightarrow \mathbb{R}$  lokal integrierbar.

- i) Ist für alle  $x \in I$  die Ungleichung  $|f(x)| \leq g(x)$  richtig und ist  $\int_a^\infty g(x) dx$  konvergent, so ist  $\int_a^\infty f(x) dx$  absolut konvergent.
- ii) Gilt umgekehrt  $0 \leq g(x) \leq f(x)$  für alle  $x \in I$  und divergiert das uneigentliche Integral  $\int_a^\infty g(x) dx$ , so divergiert auch  $\int_a^\infty f(x) dx$ .

**Beispiele.**

- i) Es ist für alle  $1 < b < \infty$

$$\begin{aligned} \int_1^b \frac{dx}{x^\alpha} &= \begin{cases} \frac{1}{1-\alpha} \frac{1}{x^{\alpha-1}} \Big|_1^b & \text{falls } \alpha \neq 1, \\ \ln(|x|) \Big|_1^b & \text{falls } \alpha = 1, \end{cases} \\ &= \begin{cases} \frac{b^{1-\alpha} - 1}{1-\alpha} & \text{falls } \alpha \neq 1, \\ \ln(b) & \text{falls } \alpha = 1. \end{cases} \end{aligned}$$

Demnach ist

$$\int_1^\infty \frac{dx}{x^\alpha} \quad \text{divergent für } \alpha \leq 1.$$

Es ist aber

$$\int_1^\infty \frac{dx}{x^\alpha} = \frac{1}{\alpha-1} \quad \text{für } \alpha > 1.$$

- ii) Zur Analyse des Integrals

$$\int_0^\infty e^{-x^2} dx$$

beachtet man zunächst, dass  $x^2 \geq 2x - 1$  aus  $(x-1)^2 \geq 0$  folgt, d.h. für alle  $x > 0$  gilt

$$e^{-x^2} \leq e^{1-2x} =: g(x).$$

Weiter ist für alle  $0 < b$

$$\int_0^b g(x) \, dx = \int_0^b e^{1-2x} \, dx = -\frac{1}{2} e^{1-2x} \Big|_0^b = \frac{e}{2} (1 - e^{-2b}),$$

die Funktion  $g$  ist somit eine konvergente Majorante von  $e^{-x^2}$  und das Majorantenkriterium impliziert die (absolute) Konvergenz von  $\int_0^\infty e^{-x^2} \, dx$ .

## Fall II. Unbeschränkte Funktionen.

Es sei nun  $f: I = [a, b) \rightarrow \mathbb{R}$  bzw.  $I = (a, b] \rightarrow \mathbb{R}$  auf allen kompakten Teilintervallen von  $[a, b)$  (bzw. von  $(a, b]$ ) integrierbar, aber nicht notwendigerweise für " $x \rightarrow b^-$ " (bzw. für " $x \rightarrow a^+$ ") beschränkt.

**Definition 9.5.** *Es sei  $f$  wie oben.*

i) *Falls der Grenzwert*

$$\lim_{\xi \rightarrow b^-} \int_a^\xi f(x) \, dx \quad \text{bzw.} \quad \lim_{\xi \rightarrow a^+} \int_\xi^b f(x) \, dx$$

*existiert, so heißt dieser das **uneigentliche Integral** von  $f$  über  $I$ .  
Speichweise: Das uneigentliche Integral existiert oder **konvergiert**.*

*Notation:*

$$\int_a^b f(x) \, dx .$$

*Andernfalls heißt  $\int_a^b f(x) \, dx$  **divergent**.*

ii) *Das uneigentliche Integral  $\int_a^b f(x) \, dx$  heißt **absolut konvergent**, falls  $\int_a^b |f(x)| \, dx$  konvergiert.*

*Absolute Konvergenz impliziert Konvergenz des uneigentlichen Integrals.*

**Bemerkung.** Im Fall II gelten analoge Bemerkungen und Kriterien wie im ersten Fall (insbesondere das Majorantenkriterium).

**Beispiele.**

i) Für  $0 < \xi < 1$  ist

$$\int_{\xi}^1 \frac{dx}{\sqrt{x}} = \left[ 2\sqrt{x} \right]_{\xi}^1 = 2 - 2\sqrt{\xi} \xrightarrow{\xi \rightarrow 0^+} 2,$$

demnach ist

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2,$$

insbesondere ist das Integral konvergent.

ii) Analog folgt die **Konvergenz** von (vgl. Übung)

$$\int_0^1 \frac{dx}{x^{\alpha}} = \frac{1}{1-\alpha} \quad \text{im Fall} \quad 0 < \alpha < 1$$

und die **Divergenz** dieses unbestimmten Integrals im Fall  $\alpha \geq 1$ .

**9.5 Numerische Integration**

Obwohl etwa jede stetige Funktion nach dem Hauptsatz der Differential- und Integralrechnung eine Stammfunktion besitzt, ist diese häufig nicht **explizit darstellbar**.

Selbst elementar aussehende Integrale wie

$$\int e^{x^2} dx$$

können nur **näherungsweise** bestimmt werden.

Die erste Idee zur **numerischen Approximation** von

$$I(f) = \int_a^b f(x) dx$$

ist es, den Integranden  $f$  durch ein Polynom  $p_n$  vom Grad  $n$  zu approximieren und

$$\int_a^b p_n(x) dx$$

als Näherung zu berechnen.

### Erster Schritt.

Im einfachsten Fall  $n = 1$ , der sogenannten **Trapezregel** (und nur diese soll hier kurz diskutiert werden), approximiert man  $\int_a^b f(x) dx$  durch die “Fläche unter der Verbindungsstrecke zwischen  $(a, f(a))$  und  $(b, f(b))$ ” (vgl. Abbildung 9.5).

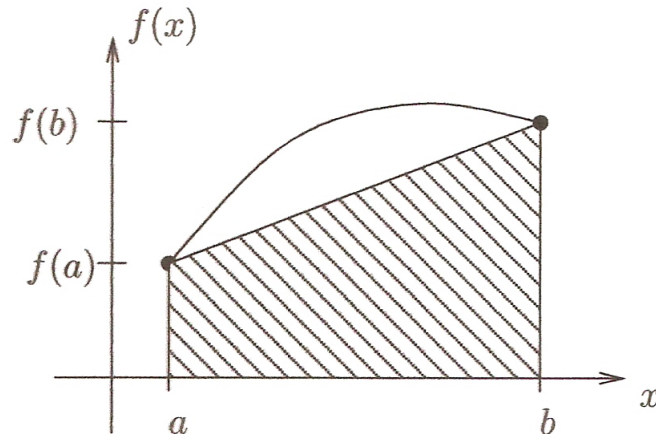


Abbildung 9.5: Zur numerischen Integration.

Man approximiert also recht grob

$$\int_a^b f(x) dx \approx \frac{b-a}{2} (f(a) + f(b)) .$$

Nur im Fall einer affinen linearen Funktion  $f(x) = \alpha x + b$  stimmt die Approximation mit dem tatsächlichen Wert überein, ansonsten kann der Fehler noch recht groß sein.

### Zweiter Schritt.

Das Intervall  $[a, b]$  wird nun in  $N$  Teilintervalle  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, N$ , gleicher Länge unterteilt:

$$x_i = a + i \frac{b-a}{N} =: a + ih, \quad i = 0, 1, \dots, N .$$

Auf jedem dieser Teilintervalle wendet man dann die Trapezregel an.

Als Approximation ergibt sich die **Trapez-Summenregel** (vgl. Abbildung 9.6)

$$\begin{aligned} \int_a^b f(x) \, dx &\approx \\ T(f, h) &:= \sum_{i=1}^N h \left[ \frac{1}{2} f(x_{i-1}) + f(x_i) \right] \\ &= h \left[ \frac{1}{2} f(x_0) + f(x_1) + f(x_2) + \dots + f(x_{N-1}) + \frac{1}{2} f(x_N) \right] \\ &= h \left[ \frac{1}{2} f(a) + f(a+h) + f(a+2h) + \dots \right. \\ &\quad \left. + f(b-h) + \frac{1}{2} f(b) \right]. \end{aligned}$$

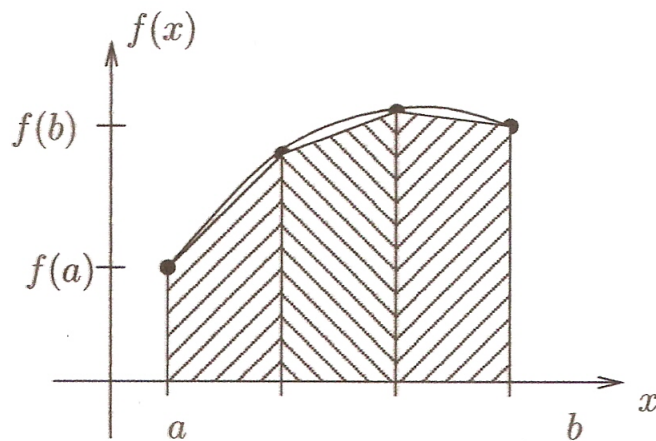


Abbildung 9.6: Zur Trapez-Summenregel.

**Beispiel.** Eine Näherung von  $\int_0^1 \frac{x}{1+x^2} \, dx$  über die Trapez-Summenregel mit  $N = 4$  ist

$$\begin{aligned} T(f, h) &= \frac{1}{4} \left[ \frac{1}{2} f(0) + f\left(\frac{1}{4}\right) + f\left(\frac{1}{2}\right) + f\left(\frac{3}{4}\right) + \frac{1}{2} f(1) \right] \\ &= \frac{1}{4} \left[ 0 + \frac{1}{4} \frac{16}{17} + \frac{1}{2} \frac{4}{5} + \frac{3}{4} \frac{16}{25} + \frac{1}{4} \right] = 0.341323529 \dots \end{aligned}$$



# Kapitel 10

## Lineare Gleichungssysteme: Das Gaußsche Eliminationsverfahren

Bei der Bearbeitung konkreter Problemstellungen in allen Anwendungsbereichen der Mathematik begegnet man immer wieder **linearen Gleichungssystemen**, deren Beherrschung in der Regel als selbstverständliches Werkzeug vorausgesetzt wird.

In diesem Kapitel wird das sogenannte **Gaußsche Eliminationsverfahren** zur expliziten Lösung linearer Gleichungssysteme vorgestellt.

Die Idee und die Durchführung des Verfahrens ist recht einfach. Die mathematisch “saubere” Formulierung sieht hingegen eher technisch aus, ohne tiefere Einsichten zu bringen.

Deshalb wird das Verfahren hier im Wesentlichen anhand von Beispielen erläutert.

Fragen nach der **Existenz und Eindeutigkeit** von Lösungen linearer Gleichungssysteme werden in diesem Kapitel noch nicht systematisch untersucht, da diese im Kontext von 12 in natürlicher Weise beantwortet werden können.

### **Ein lineares Gleichungssystem.**

Es seien  $n, m \in \mathbb{N}$  und  $m \leq n$ . Ein lineares Gleichungssystem aus  **$n$  Gleichungen** in  **$m$  Unbekannten** (d.h. gesuchten Größen)  $x_1, x_2, \dots, x_m$

sieht wie folgt aus:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m &= b_2 \\ &\vdots \quad \vdots \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m &= b_n . \end{aligned}$$

Dabei sind für  $1 \leq i \leq n$  und für  $1 \leq j \leq m$  die  $a_{ij} \in \mathbb{R}$  (die **Koeffizienten**) und die  $b_i \in \mathbb{R}$  (die **rechte Seite**) gegeben.

### Beispiele.

i) 3 Gleichungen in 3 Unbekannten: Man betrachte die Gleichungen

(a)

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ x_1 &+ x_3 &= 2 \\ &x_2 + x_3 &= 3 , \end{aligned}$$

(b) sowie für fixiertes  $\lambda \in \mathbb{R}$

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ x_1 &+ x_3 &= 1 \\ 2x_1 + 2x_2 + 2x_3 &= \lambda . \end{aligned}$$

ii) 2 Gleichungen in 4 Unbekannten:

$$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 0 \\ x_1 &+ x_3 &= 2 . \end{aligned}$$

### Schematische Darstellung.

Die Koeffizienten und die rechte Seite werden schematisch abgeordnet:

$$\left( \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{array} \right) .$$

In den Beispielen sieht das wie folgt aus:



i) (a)

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 3 \end{array} \right),$$

(b)

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 2 & 2 & \lambda \end{array} \right),$$

ii)

$$\left( \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 2 \end{array} \right).$$

### Elementare Zeilenumformungen.

Die Lösungsmenge (d.h. die Menge aller möglichen Lösungen) eines linearen Gleichungssystems ändert sich sicherlich nicht, wenn

- i) die Reihenfolge der Gleichungen geändert wird, d.h. wenn **zwei Zeilen vertauscht werden**;
- ii) das Vielfache einer Gleichung betrachtet wird, d.h. wenn **eine Zeile mit  $0 \neq \lambda \in \mathbb{R}$  multipliziert wird**;
- iii) eine Gleichung bzw. deren Vielfaches zu einer anderen Gleichung addiert wird, d.h. wenn **das Vielfache einer Zeile zu einer anderen Zeile addiert wird**.

**Bemerkung.** Elementare Zeilenumformungen beziehen sich natürlich auch auf die rechte Seite: Die **Operationen sind auch auf die entsprechenden  $b_i$  anzuwenden**.

### Ziel.

Durch elementare Zeilenumformungen soll das Schema – falls möglich – in **Zeilenstufenform** gebracht werden, anhand derer die Lösungen

direkt abgelesen werden können.

Ein Schema in Zeilenstufenform sieht wie folgt aus:

$$\left( \begin{array}{cccccccc|c} 1 & \tilde{a}_{12} & \dots & \tilde{a}_{1(r-1)} & \tilde{a}_{1r} & \tilde{a}_{1(r+1)} & \dots & \tilde{a}_{1m} & \tilde{b}_1 \\ 0 & 1 & \dots & \tilde{a}_{2(r-1)} & \tilde{a}_{2r} & \tilde{a}_{2(r+1)} & \dots & \tilde{a}_{2m} & \tilde{b}_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & \tilde{a}_{r(r+1)} & \dots & \tilde{a}_{rm} & \tilde{b}_r \\ \hline 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \tilde{b}_{r+1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \tilde{b}_{r+2} \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \tilde{b}_n \end{array} \right) .$$

### Zur Lösung der Beispiele.

i) (a) Aus

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 3 \end{array} \right)$$

wird durch Vertauschung der zweiten und dritten Zeile

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 3 \\ 1 & 0 & 1 & 2 \end{array} \right) .$$

Wird von der dritten Zeile die erste Zeile abgezogen, so folgt im nächsten Schritt

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 3 \\ 0 & -2 & 0 & 1 \end{array} \right) .$$

Nun kann das zweifache der zweiten Zeile zur dritten Zeile addiert und das Ergebnis durch zwei geteilt werden. Man erhält

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 7/2 \end{array} \right) .$$

In dieser Form kann sukzessive von unten nach oben vorgegangen werden. Man liest ab:

$$\begin{aligned}x_3 &= \frac{7}{2}, \\x_2 &= 3 - x_3 = -\frac{1}{2}, \\x_1 &= 1 - 2x_2 - x_3 = -\frac{3}{2}.\end{aligned}$$

Abschließend bestätigt eine Probe die Rechnungen.

**Bemerkung.** Die Wahl, die Anzahl und die Reihenfolge der elementaren Zeilenumformungen nicht eindeutig.

(b) Im Schema

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 2 & 2 & \lambda \end{array} \right)$$

zieht man beispielsweise von der dritten die zweite und die erste Zeile ab:

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & \lambda - 2 \end{array} \right).$$

Hier wird bereits deutlich, dass das System für  $\lambda \neq 2$  nicht lösbar ist.

Es sei also  $\lambda = 2$ . Dann ergeben die Subtraktion der zweiten Zeile von der ersten und die Division durch 2 die finale Form

$$\left( \begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

Hier ist zu erkennen, dass  $x_3$  beliebig gewählt werden kann. Man schreibt beispielsweise

$$x_3 = t \quad \text{für beliebiges } t \in \mathbb{R}.$$

Weiter folgt

$$x_2 = 0$$

sowie

$$x_1 = 1 - t \quad \text{für beliebiges } t \in \mathbb{R} .$$

Wieder kann das Ergebnis mittels einer Probe verifiziert werden.

**Bemerkung.** Wie in diesem Beispiel folgt aus obiger Zeilenstufenform:

Ist  $\tilde{b}_k = 0$  für ein  $r + 1 \leq k \leq n$ , so ist das Gleichungssystem nicht lösbar.

ii) Hier transformiert sich

$$\left( \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 2 \end{array} \right)$$

nach der Subtraktion der zweiten Zeile von der ersten und nach der anschließende Division durch 2 auf

$$\left( \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1/2 & -1 \end{array} \right) .$$

Sowohl  $x_4$  als auch  $x_3$  können beliebig gewählt werden:

$$x_4 = t \quad \text{für beliebiges } t \in \mathbb{R} ,$$

$$x_3 = s \quad \text{für beliebiges } s \in \mathbb{R} ,$$

d.h. für beliebige  $s, t \in \mathbb{R}$  (Probe!)

$$\begin{aligned} x_2 &= -1 - \frac{t}{2} , \\ x_1 &= 2 - s . \end{aligned}$$

### Elementare Spaltenumformungen.

Nicht immer ist es möglich, durch elementare Zeilenumformungen ein Schema in Zeilestufenform zu produzieren.

**Beispiel.** Das Schema zum Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_3 &= 2 \end{aligned}$$

lautet

$$\left( \begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{array} \right).$$

Um formal eine Zeilenstufenform zu erhalten, müssen die **zweite und die dritte Spalte vertauscht werden**.

**Vorsicht.** Ein Vertauschung von Spalten bedeutet **eine Umnummerierung der gesuchten  $x_i$** .

Aber auch in obiger Form können die Lösungen sukzessive abgelesen werden. Es ist

$$\begin{aligned} x_3 &= 2, \\ x_2 &= t \quad \text{für beliebiges } t \in \mathbb{R}, \\ x_1 &= -1 - t \quad \text{für beliebiges } t \in \mathbb{R}. \end{aligned}$$



# Kapitel 11

## Der Vektorraum $\mathbb{R}^n$

### 11.1 Vektorräume

#### Frage: Was ist ein Vektor?

Unter einem (freien) Vektor im  $\mathbb{R}^n$  stellt man sich eine gerichtete Größe (einen Pfeil) vor (als typisches Beispiel eine Kraft im  $\mathbb{R}^3$ ), die durch Länge und Richtung gekennzeichnet ist.

Freie Vektoren, die durch Parallelverschiebung ineinander überführt werden können, werden als gleich angesehen.

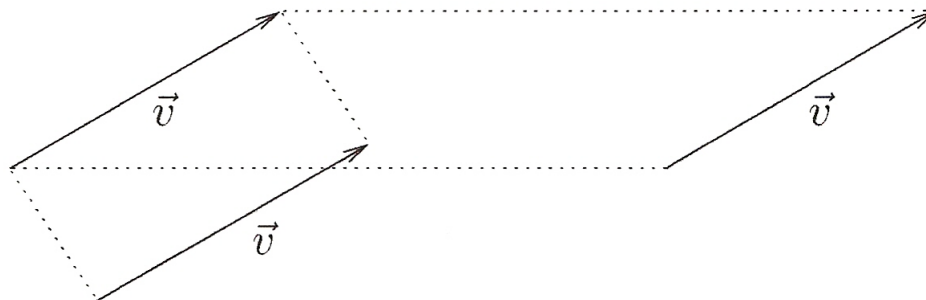


Abbildung 11.1: Parallelverschiebung freier Vektoren.

Dementsprechend kann ein Vektor so verschoben werden, dass sein Anfangspunkt im Koordinatenursprung, dem Nullpunkt, liegt.

Damit ist ein Vektor im  $\mathbb{R}^n$  durch die Lage seines Endpunktes, d.h. durch die Koordinaten (oder Komponenten) seines Endpunktes, charakterisiert.

Mit anderen Worten: Es handelt sich um ein Element  $\underline{x}$  des kartesischen

Produktes

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}},$$

wobei  $\underline{\mathbf{x}}$  in der Regel als **Spaltenvektor** (geordnetes  $n$ -Tupel) geschrieben wird:

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad x_1, x_2, \dots, x_n \in \mathbb{R}.$$

Die Komponenten werden geometrisch durch die **Projektionen auf die Koordinatenachsen** beschrieben.

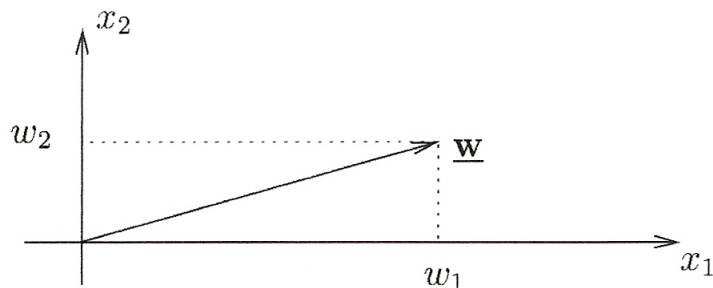


Abbildung 11.2: Der Vektor  $\underline{\mathbf{w}}$  mit den Komponenten  $w_1, w_2$  im  $\mathbb{R}^2$ .

Anschaulich werden Vektoren in einem **Kräfteparallelogramm** addiert (vgl. Abbildung 11.3).

Ebenso können Vektoren durch **Multiplikation mit einem Skalar** gestreckt oder gestaucht werden.



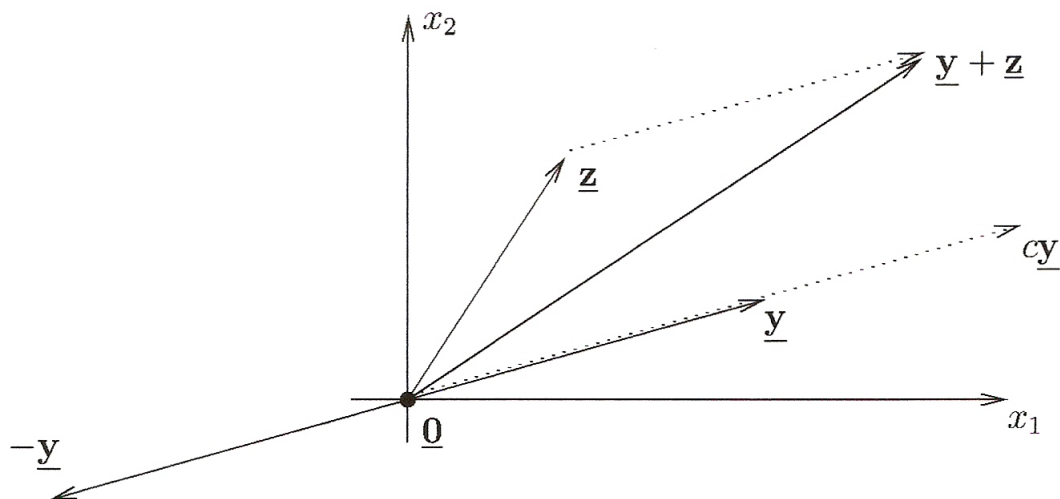


Abbildung 11.3: Ein Kräfteparallelogramm.

Es ist dabei ( $\underline{x}, \underline{y} \in \mathbb{R}^n, c \in \mathbb{R}$ )

$$c\underline{x} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}, \quad -\underline{x} = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix},$$

$$\underline{x} + \underline{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \underline{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

**Antwort: Ein Vektor ist ein Element eines Vektorraums.**

**Definition 11.1.** Es sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  und  $V$  eine Menge. Ferner existiere ein *Addition*  $+$ ,

$$+ : V \times V \rightarrow V, \\ (\underline{v}, \underline{w}) \mapsto \underline{v} + \underline{w}$$

und eine *Multiplikation mit Skalaren*  $\cdot$ ,

$$\cdot : \mathbb{K} \times V \rightarrow V, \\ (\lambda, \underline{v}) \mapsto \lambda \underline{v},$$

sodass gilt:

**V1:**  $(V, +)$  ist eine *kommutative Gruppe*, d.h.:

(a) Es gibt ein *neutrales Element*  $\underline{\mathbf{0}} \in V$  mit

$$\underline{\mathbf{x}} + \underline{\mathbf{0}} = \underline{\mathbf{x}} \quad \text{für alle } \underline{\mathbf{x}} \in V .$$

(b) Zu jedem  $\underline{\mathbf{x}} \in V$  existiert genau ein *inverses Element*, bezeichnet mit  $(-\underline{\mathbf{x}})$ , sodass

$$\underline{\mathbf{x}} + (-\underline{\mathbf{x}}) = \underline{\mathbf{0}} .$$

(c) Für alle  $\underline{\mathbf{x}}, \underline{\mathbf{y}}, \underline{\mathbf{z}} \in V$  gilt

$$(\underline{\mathbf{x}} + \underline{\mathbf{y}}) + \underline{\mathbf{z}} = \underline{\mathbf{x}} + (\underline{\mathbf{y}} + \underline{\mathbf{z}}) \quad (\text{Assoziativgesetz}) .$$

Für alle  $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in V$  gilt

$$\underline{\mathbf{x}} + \underline{\mathbf{y}} = \underline{\mathbf{y}} + \underline{\mathbf{x}} \quad (\text{Kommutativgesetz}) .$$

**V2:** Es ist

$$1 \underline{\mathbf{v}} = \underline{\mathbf{v}}$$

und es gelten die *Assoziativ-* und *Distributivgesetze* ( $\lambda, \mu \in \mathbb{K}, \underline{\mathbf{x}}, \underline{\mathbf{y}} \in V$ )

$$\begin{aligned} (\lambda \mu) \underline{\mathbf{x}} &= \lambda (\mu \underline{\mathbf{x}}) , \\ (\lambda + \mu) \underline{\mathbf{x}} &= (\lambda \underline{\mathbf{x}}) + (\mu \underline{\mathbf{x}}) , \\ \lambda (\underline{\mathbf{x}} + \underline{\mathbf{y}}) &= (\lambda \underline{\mathbf{x}}) + (\lambda \underline{\mathbf{y}}) . \end{aligned}$$

Dann heißt  $(V, +, \cdot)$  ein *Vektorraum* oder *linearer Raum* über dem Körper  $\mathbb{K}$ .

Die Elemente eines Vektorraums heißen *Vektoren*.

**Bemerkung.** Die Einschränkung  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  ist nicht notwendig zur Definition eines Vektorraums. Im Allgemeinen muss  $\mathbb{K}$  lediglich eine Körperstruktur haben (vgl. Abschnitt 2.2).

**Beispiel.** Der  $\mathbb{R}^n$  ist ein  $\mathbb{R}$ -Vektorraum (siehe oben).

Alle nachfolgenden Betrachtungen beschränken sich auf dieses Beispiel  $V = \mathbb{R}^n$ .

### Lineare Abhängigkeit von Vektoren.

Ein Vektor  $\underline{\mathbf{w}} \in \mathbb{R}^n$  heißt **Linearkombination** von  $k$  Vektoren  $\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(k)}$  im  $\mathbb{R}^n$ , falls er als Summe von Vielfachen der  $\underline{\mathbf{v}}^{(i)}$  geschrieben werden kann:

$$\underline{\mathbf{w}} = \sum_{i=1}^k \lambda_k \underline{\mathbf{v}}^{(i)}$$

für reelle Zahlen  $\lambda_1, \lambda_2, \dots, \lambda_k$ .

### Beispiele.

i) Im  $\mathbb{R}^3$  seien die drei Vektoren

$$\underline{\mathbf{e}}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{e}}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{e}}^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

gegeben.

(a) Der Vektor  $\underline{\mathbf{e}}^{(3)}$  (analog  $\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}$ ) **kann nicht als Linearkombination von  $\underline{\mathbf{e}}^{(1)}$  und  $\underline{\mathbf{e}}^{(2)}$  geschrieben werden**, da für alle  $\lambda_1, \lambda_2 \in \mathbb{R}$  gilt

$$\lambda_1 \underline{\mathbf{e}}^{(1)} + \lambda_2 \underline{\mathbf{e}}^{(2)} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ 0 \end{pmatrix} \neq \underline{\mathbf{e}}^{(3)}.$$

(b) Gleichbedeutend ist, dass der Nullvektor  $\underline{\mathbf{0}}$  **nur mit der Wahl  $\lambda_1 = \lambda_2 = \lambda_3 = 0$  als Linearkombination von  $\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(3)}$  geschrieben werden kann:**

$$\underline{\mathbf{0}} = \sum_{i=1}^3 \lambda_i \underline{\mathbf{e}}^{(i)} \Leftrightarrow \lambda_i = 0 \quad \text{für } i = 1, 2, 3.$$

ii) Nun seien die Vektoren

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(3)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

gegeben.

(a) Hier gilt beispielsweise

$$\underline{\mathbf{v}}^{(1)} = \underline{\mathbf{v}}^{(3)} - \underline{\mathbf{v}}^{(2)} .$$

(b) Der Nullvektor kann geschrieben werden als

$$\underline{\mathbf{0}} = \underline{\mathbf{v}}^{(1)} + \underline{\mathbf{v}}^{(2)} - \underline{\mathbf{v}}^{(3)} .$$

**Definition 11.2.** Die Vektoren  $\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(k)}$  im  $\mathbb{R}^n$  heißen *linear unabhängig*, falls aus

$$\underline{\mathbf{0}} = \sum_{i=1}^k \lambda_i \underline{\mathbf{v}}^{(i)}$$

*zwingend folgt*

$$\lambda_1 = \lambda_2 = \dots = \lambda_k = 0 .$$

Andernfalls heißen die Vektoren *linear abhängig*.

**Beispiele.**

i) Betrachtet sei der  $\mathbb{R}^2$  mit den Vektoren

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} , \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \underline{\mathbf{v}}^{(3)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

Aus

$$1\underline{\mathbf{v}}^{(1)} + (-1)\underline{\mathbf{v}}^{(2)} + (-2)\underline{\mathbf{v}}^{(3)} = \underline{\mathbf{0}}$$

folgt, dass die Vektoren linear abhängig sind.

ii) Sind die Vektoren

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} , \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

im  $\mathbb{R}^3$  gegeben, so gilt

$$\begin{aligned} \underline{\mathbf{0}} = \lambda_1 \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} &\Rightarrow 2\lambda_1 + \lambda_2 = 0 , \quad \lambda_2 = 0 \\ &\Rightarrow \lambda_1 = \lambda_2 = 0 , \end{aligned}$$

was die lineare Unabhängigkeit der Vektoren beweist.

### Dimension und Basen des $\mathbb{R}^n$ .

Im  $\mathbb{R}^n$  seien die Vektoren

$$\underline{\mathbf{e}}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{e}}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \dots, \quad \underline{\mathbf{e}}^{(n)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

betrachtet.

Es ist leicht nachzuprüfen, dass diese  $n$  Vektoren linear unabhängig sind.

Andererseits kann als Übung gezeigt werden, dass es im  $\mathbb{R}^n$  keine  $n + 1$  linear unabhängige Vektoren geben kann.

Das motiviert

**Definition 11.3.** i) Die maximale Zahl linear unabhängiger Vektoren eines Vektorraumes  $V$  heißt die Dimension des Vektorraumes,  $\dim V$ .

Es ist also  $\dim \mathbb{R}^n = n$ .

ii) Es sei  $n$  die Dimension eines Vektorraumes  $V$ .

Dann heißt jedes  $n$ -Tupel  $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(n)})$  von linear unabhängigen Vektoren aus  $V$  eine Basis von  $V$ .

Die obige Basis  $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(n)})$  des  $\mathbb{R}^n$  heißt die kanonische Basis oder die Standardbasis des  $\mathbb{R}^n$ .

**Beispiel.** Es seien

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(3)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Dann ist  $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \underline{\mathbf{v}}^{(3)})$  eine Basis des  $\mathbb{R}^3$  (Übung).

### Darstellung eines Vektors bzgl. einer gegebenen Basis.

Ist  $\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$  beliebig, so gilt

$$\underline{\mathbf{x}} - x_1 \underline{\mathbf{e}}^{(1)} - x_2 \underline{\mathbf{e}}^{(2)} - \dots - x_n \underline{\mathbf{e}}^{(n)} = \underline{\mathbf{0}}, \quad \text{bzw. } \underline{\mathbf{x}} = \sum_{i=1}^n x_i \underline{\mathbf{e}}^{(i)},$$

d.h. jeder Vektor  $\underline{\mathbf{x}}$  ist eine **eindeutig bestimmte Linearkombination** der Basisvektoren  $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(n)})$ .

Gleiches gilt für beliebige Basen:

Ist  $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(n)})$  eine Basis des  $\mathbb{R}^n$ , so existieren zu jedem  $\underline{\mathbf{v}} \in \mathbb{R}^n$  **eindeutig bestimmte Koeffizienten**  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$  mit

$$\underline{\mathbf{v}} = \sum_{k=1}^n \lambda_k \underline{\mathbf{v}}^{(k)}.$$

Diese Koeffizienten heißen **Koordinaten** von  $\underline{\mathbf{v}}$  bzgl. der Basis  $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(n)})$ .

**Beispiel.** Im  $\mathbb{R}^3$  seien wieder

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(3)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Dann kann jeder Vektor  $\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$  eindeutig geschrieben werden als

$$\underline{\mathbf{x}} = (x_1 - x_2) \underline{\mathbf{v}}^{(1)} + (x_2 - x_3) \underline{\mathbf{v}}^{(2)} + x_3 \underline{\mathbf{v}}^{(3)}.$$

## 11.2 Norm und Skalarprodukt

Im Vektorraum  $\mathbb{R}^n$  können (wie es der anschaulichen Vorstellung im  $\mathbb{R}^3$  entspricht) **Längen und Winkel gemessen werden**.

Dies geschieht mithilfe der **Euklidischen Norm** bzw. des **Euklidischen Skalarproduktes**.

Der Vektorraum  $\mathbb{R}^n$  versehen mit dem Euklidischen Skalarprodukt heißt der **Euklidische Raum**  $\mathbb{R}^n$ . Manchmal wird dies durch die Schreibweise  $\mathbb{E}^n$  verdeutlicht.

### Wie werden Längen gemessen?

Wenn eine Größe  $\|\underline{\mathbf{x}}\|$  die Länge eines beliebigen Vektors  $\underline{\mathbf{x}} \in \mathbb{R}^n$  messen soll, dann muss sie sicherlich den folgenden drei grundlegenden Eigenschaften einer Längenmessung genügen:

- i) Die Länge eines Vektors ist immer **größer oder gleich Null** und **nur der Nullvektor hat die Länge Null**:

$$\|\underline{\mathbf{x}}\| \geq 0 \text{ und } \|\underline{\mathbf{x}}\| = 0 \Leftrightarrow \underline{\mathbf{x}} = \underline{\mathbf{0}} \text{ für alle } \underline{\mathbf{x}} \in \mathbb{R}^n.$$

- ii) Wird ein Vektor **um einen Faktor gestreckt oder gestaucht**, so **ändert sich die Länge um den Betrag dieses Faktors**:

$$\|\lambda \underline{\mathbf{x}}\| = |\lambda| \|\underline{\mathbf{x}}\| \text{ für alle } \underline{\mathbf{x}} \in \mathbb{R}^n \text{ und für alle } \lambda \in \mathbb{R}.$$

- iii) Es gilt die **Dreiecksungleichung** (vgl. Abbildung 11.4):

$$\|\underline{\mathbf{x}} + \underline{\mathbf{y}}\| \leq \|\underline{\mathbf{x}}\| + \|\underline{\mathbf{y}}\| \text{ für alle } \underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n.$$

Sind die Eigenschaften i) – iii) erfüllt, so spricht man von einer **Norm**.

Im  $\mathbb{R}^n$  gibt es **zahlreiche Normen** zur Längenmessung. Die wichtigste Norm ist die **Euklidische Norm**, die der intuitiven Vorstellung entspricht:

$$\|\underline{\mathbf{x}}\| := \left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \text{ für alle } \underline{\mathbf{x}} \in \mathbb{R}^n.$$

Die Eigenschaften i) und ii) verifiziert man leicht für die Euklidische Norm, iii) wird in Kürze gezeigt.

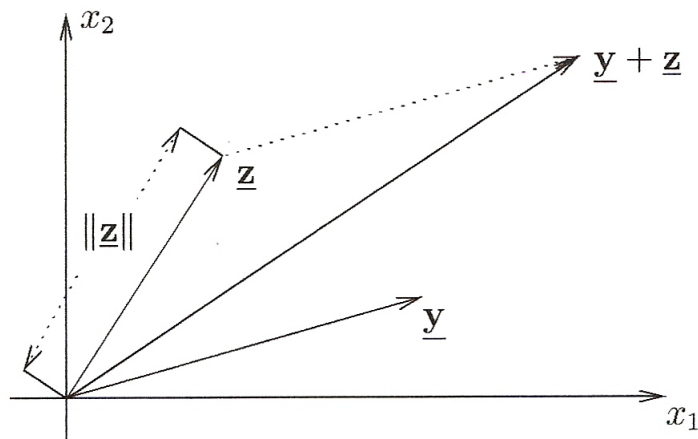


Abbildung 11.4: Es gilt die Dreiecksungleichung  $\|\underline{y} + \underline{z}\| \leq \|\underline{y}\| + \|\underline{z}\|$ .

Im Folgenden bezeichnet  $\|\underline{x}\|$  stets die Euklidische Norm, andere Normen werden nicht betrachtet.

### Wie werden Winkel gemessen?

Wesentlich für die Winkelmessung ist, dass die Euklidische Norm von einem **Skalarprodukt induziert** ist.

Dabei bezeichnet ein Skalarprodukt im  $\mathbb{R}^n$  eine Abbildung, die zwei **Vektoren**  $\underline{x}, \underline{y} \in \mathbb{R}^n$  eine **reelle Zahl** zuordnet

(Notation :  $\langle \underline{x}, \underline{y} \rangle$  oder  $\underline{x} \cdot \underline{y}$ )

und für alle  $\underline{x}, \underline{y}, \underline{z} \in \mathbb{R}^n$  sowie für alle  $\lambda \in \mathbb{R}$  die folgenden Eigenschaften hat:

- i)  $\langle \underline{x}, \underline{y} \rangle = \langle \underline{y}, \underline{x} \rangle$  (**Kommutativgesetz**);
- ii)  $\langle \underline{x} + \underline{y}, \underline{z} \rangle = \langle \underline{x}, \underline{z} \rangle + \langle \underline{y}, \underline{z} \rangle$  (**Distributivgesetz**);
- iii)  $\lambda \langle \underline{x}, \underline{y} \rangle = \langle \lambda \underline{x}, \underline{y} \rangle = \langle \underline{x}, \lambda \underline{y} \rangle$  (**Assoziativgesetz**);
- iv)  $\underline{x} \neq \underline{0} \Leftrightarrow \langle \underline{x}, \underline{x} \rangle > 0$  (**positive Definitheit**).



Alle diese Eigenschaften können für das Euklidische Skalarprodukt des  $\mathbb{R}^n$  (auch genannt das **kanonische** oder das **Standardskalarprodukt**) verifiziert werden:

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle := \sum_{j=1}^n x_j y_j, \quad \underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \underline{\mathbf{y}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n.$$

Andere Skalarprodukte werden im Folgenden nicht betrachtet.

**Wichtige Beobachtung.** Die Euklidische Norm und das Euklidische Skalarprodukt sind für alle  $\underline{\mathbf{x}} \in \mathbb{R}^n$  über die Gleichung

$$\|\underline{\mathbf{x}}\| = \langle \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle^{\frac{1}{2}}$$

miteinander verbunden.

Ähnlich wie die Dreiecksungleichung ist die sogenannte **Cauchy-Schwarzsche Ungleichung** ein wesentliches Werkzeug der Analysis:

**Satz 11.1.** (*Cauchy-Schwarzsche Ungleichung*)

Für alle  $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$  gilt

$$|\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle| \leq \|\underline{\mathbf{x}}\| \|\underline{\mathbf{y}}\|.$$

**Folgerung.** Als Konsequenz der Cauchy-Schwarzschen Ungleichung kann beispielsweise die Dreiecksungleichung gezeigt werden.

Für alle  $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$  gilt nämlich:

$$\begin{aligned} \|\underline{\mathbf{x}} + \underline{\mathbf{y}}\|^2 &= \langle \underline{\mathbf{x}} + \underline{\mathbf{y}}, \underline{\mathbf{x}} + \underline{\mathbf{y}} \rangle = \|\underline{\mathbf{x}}\|^2 + \|\underline{\mathbf{y}}\|^2 + 2\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle \\ &\leq \|\underline{\mathbf{x}}\|^2 + \|\underline{\mathbf{y}}\|^2 + 2\|\underline{\mathbf{x}}\|\|\underline{\mathbf{y}}\| = (\|\underline{\mathbf{x}}\| + \|\underline{\mathbf{y}}\|)^2. \end{aligned}$$

Nach diesen Vorbereitungen kann zur Winkelmessung der **Cosinus zwischen zwei Vektoren** definiert werden.

Für alle  $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$  definiert man

$$\cos(\underline{\mathbf{x}}, \underline{\mathbf{y}}) := \frac{\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle}{\|\underline{\mathbf{x}}\| \|\underline{\mathbf{y}}\|},$$

wobei die Cauchy-Schwarzsche Ungleichung

$$-1 \leq \cos(\underline{\mathbf{x}}, \underline{\mathbf{y}}) \leq 1 .$$

impliziert.

**Bemerkung.** Natürlich muss diese Definition **im Einklang stehen** mit

- i) der **elementargeometrischen Definition** des Cosinus (vgl. Abschnitt 3.4.3),
- ii) der Definition des Cosinus als **Potenzreihe** (siehe Kapitel 6 und dort insbesondere die Diskussion von  $e^{iz}$ ).

Dies mache man sich als Übung klar.

### Orthogonalität von Vektoren.

Mit dem Cosinus ist der Winkel zwischen zwei Vektoren definiert und als Spezialfall natürlich auch, wann zwei Vektoren **senkrecht aufeinander stehen**:

Zwei Vektoren  $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$  stehen senkrecht aufeinander (oder sind orthogonal), falls

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle = 0 .$$

**Beispiel.** Betrachtet sei die Standardbasis  $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(n)})$  des  $\mathbb{R}^n$ . Diese Basisvektoren haben die Länge 1 und stehen senkrecht aufeinander, d.h. für alle  $i, j = 1, 2, \dots, n$  ist

$$\langle \underline{\mathbf{e}}^{(i)}, \underline{\mathbf{e}}^{(j)} \rangle = \delta_{ij} := \begin{cases} 1 & \text{für } i = j , \\ 0 & \text{für } i \neq j , \end{cases}$$

wobei  $\delta_{ij}$  **Kronecker-Symbol** heißt.

Aufgrund dieser Eigenschaft sind Rechnungen oft deutlich einfacher, wenn die Koordinaten von Vektoren auf die kanonische Basis und nicht auf eine beliebige Basis bezogen sind.

Allgemein gilt:

**Satz 11.2.** Es seien  $\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)}, \dots, \underline{\mathbf{f}}^{(n)} \in \mathbb{R}^n$  und für alle  $i, j = 1, 2, \dots, n$  gelte

$$\langle \underline{\mathbf{f}}^{(i)}, \underline{\mathbf{f}}^{(j)} \rangle = \delta_{ij} .$$

i) Dann ist  $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)}, \dots, \underline{\mathbf{f}}^{(n)})$  eine **Basis** des  $\mathbb{R}^n$ .

Jede Basis mit obiger Eigenschaft heißt **Orthonormalbasis** (ONB) des  $\mathbb{R}^n$ .

ii) Jeder Vektor  $\underline{\mathbf{x}} \in \mathbb{R}^n$  ist bzgl. der Basis  $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)}, \dots, \underline{\mathbf{f}}^{(n)})$  **eindeutig als Linearkombination**

$$\underline{\mathbf{x}} = \sum_{k=1}^n \langle \underline{\mathbf{x}}, \underline{\mathbf{f}}^{(k)} \rangle \underline{\mathbf{f}}^{(k)}$$

dargestellt.

*Beweis.*

i) Man multipliziere  $\sum_{i=1}^n \lambda_i \underline{\mathbf{f}}^{(i)} = \underline{\mathbf{0}}$  skalar mit  $\underline{\mathbf{f}}^{(k)}$ .

ii) Man multipliziere  $\underline{\mathbf{x}} = \sum_{i=1}^n \alpha_i \underline{\mathbf{f}}^{(i)}$  skalar mit  $\underline{\mathbf{f}}^{(k)}$ . □

**Beispiel.** Die Vektoren

$$\tilde{\underline{\mathbf{f}}}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \tilde{\underline{\mathbf{f}}}^{(2)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$$

stehen senkrecht aufeinander.

Nach der **Normierung auf die Länge 1**

$$\underline{\mathbf{f}}^{(1)} = \frac{1}{\|\tilde{\underline{\mathbf{f}}}^{(1)}\|} \tilde{\underline{\mathbf{f}}}^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \underline{\mathbf{f}}^{(2)} = \frac{1}{\|\tilde{\underline{\mathbf{f}}}^{(2)}\|} \tilde{\underline{\mathbf{f}}}^{(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

ergibt sich die Orthonormalbasis  $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)})$  des  $\mathbb{R}^2$ .

Es ist beispielsweise (Probe!)

$$\begin{aligned} \begin{pmatrix} 1 \\ 2 \end{pmatrix} &= \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle \underline{\mathbf{f}}^{(1)} + \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle \underline{\mathbf{f}}^{(2)} \\ &= \frac{3}{\sqrt{2}} \underline{\mathbf{f}}^{(1)} + \frac{1}{\sqrt{2}} \underline{\mathbf{f}}^{(1)}. \end{aligned}$$

## 11.3 Analytische Geometrie

### Geraden im $\mathbb{R}^n$ .

Bekanntlich ist eine Gerade  $g$  bereits durch die Kenntnis von **zwei Punkten auf der Geraden vollständig bestimmt**.

Liegen zwei verschiedene Punkte  $\underline{\mathbf{x}}^{(1)}$  und  $\underline{\mathbf{x}}^{(2)} \in \mathbb{R}^n$  auf der Geraden  $g$  und ist  $\underline{\mathbf{v}} = \underline{\mathbf{x}}^{(1)} - \underline{\mathbf{x}}^{(2)}$ , so lautet die **Parameterdarstellung der Geraden**:

$$g = \{ \underline{\mathbf{x}} \in \mathbb{R}^n : \underline{\mathbf{x}} = \underline{\mathbf{x}}_0 + t\underline{\mathbf{v}}, t \in \mathbb{R} \} .$$

Dabei kann man sich  $t$  als die **Zeit** und  $\underline{\mathbf{v}}$  als die **Geschwindigkeit vorstellen**, mit der die Gerade durchlaufen wird.

Da die Geschwindigkeit ein Vektor ist, gibt  $\underline{\mathbf{v}}$  insbesondere auch die **Richtung der Geraden** vor.

**Beispiel.** Betrachtet sei die Gerade  $g$  im  $\mathbb{R}^2$ , die für  $a, b \in \mathbb{R}$  durch die Punkte

$$\begin{pmatrix} 0 \\ b \end{pmatrix}, \begin{pmatrix} 1 \\ a+b \end{pmatrix} \in \mathbb{R}^2$$

verläuft. Dann gilt für alle  $\underline{\mathbf{x}} \in g$

$$x_1 = t, \quad x_2 = b + ta \quad \text{für ein } t \in \mathbb{R},$$

d.h. als Funktionsvorschrift  $x_2 = f(x_1)$  ergibt sich

$$x_2 = ax_1 + b .$$

Zur Analyse von Ebenen im  $\mathbb{R}^3$  sind noch einige Begriffsbildungen voranzuschicken:

### Unterräume des $\mathbb{R}^n$ .

**Definition 11.4.** Ist  $U$  eine Teilmenge des  $\mathbb{R}^n$ , die **selbst ein Vektorraum ist**, so heißt  $U$  ein **Unterraum des  $\mathbb{R}^n$** .

**Typische Situation.** Gegeben seien  $k$  Vektoren  $\underline{\mathbf{x}}^{(1)}, \underline{\mathbf{x}}^{(2)}, \dots, \underline{\mathbf{x}}^{(k)} \in \mathbb{R}^n$ . Dann ist ihre **lineare Hülle**

$$L := \text{Spann}(\underline{\mathbf{x}}^{(1)}, \dots, \underline{\mathbf{x}}^{(k)}) := \left\{ \sum_{i=1}^k \alpha_i \underline{\mathbf{x}}^{(i)} : \alpha_i \in \mathbb{R}, i = 1, 2, \dots, k \right\}$$

ein Unterraum des  $\mathbb{R}^n$  (der von den  $\underline{\mathbf{x}}^{(i)}$  **aufgespannte** Raum).

**Beispiel.** Es sei  $n = 2$  und

$$L = \left\{ \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\}.$$

In diesem Beispiel ist  $L$  die “ $x_1$ ”-Achse im  $\mathbb{R}^2$ .

Für die lineare Hülle  $L$  von  $k$  Vektoren im  $\mathbb{R}^n$  gilt

$$\underline{\mathbf{0}} \in L, \quad 0 \leq \dim L \leq n,$$

und

$$\dim L = 0 \Leftrightarrow L = \{\underline{\mathbf{0}}\}, \quad \dim L = n \Leftrightarrow L = \mathbb{R}^n.$$

**Bemerkung.** Wie in obigem Beispiel sind **eindimensionale Unterräume des  $\mathbb{R}^n$  Geraden**.

Da der Nullvektor in jedem Unterraum liegen muss, Geraden aber nicht durch den Nullpunkt verlaufen müssen, **gilt die Umkehrung nicht**.

Geraden, die **nicht durch den Ursprung verlaufen**, sind **affin lineare Räume**, d.h. es handelt sich um “verschobene eindimensionale Unterräume”.

### **Orthogonales Komplement.**

**Definition 11.5.** Ist  $U \subset \mathbb{R}^n$  eine beliebige Menge, so heißt

$$U^\perp := \{\underline{\mathbf{x}} \in \mathbb{R}^n : \langle \underline{\mathbf{x}}, \underline{\mathbf{u}} \rangle = 0 \text{ für alle } \underline{\mathbf{u}} \in U\}$$

das **orthogonale Komplement** von  $U$ .

**Übung.** Man mache sich klar, dass  $U^\perp$  stets ein Unterraum ist, auch wenn  $U$  selbst kein Unterraum ist.

**Beispiel.** Es sei  $n = 3$  und

$$\underline{\mathbf{N}} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \text{und} \quad U = \{\underline{\mathbf{N}}\}.$$

Dann ist

$$U^\perp = \left\{ \underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x_1 + x_2 + 2x_3 = 0 \right\}$$

und  $U^\perp$  ist die Ebene durch den Nullpunkt, die auf dem Vektor  $\underline{\mathbf{N}}$  senkrecht steht.

In  $U^\perp$  liegen beispielsweise die Vektoren

$$\underline{\mathbf{v}} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{w}} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

Diese sind linear unabhängig und spannen die Ebene  $U^\perp$  auf (vgl. Abbildung 11.5).

### Das Vektorprodukt im $\mathbb{R}^3$ .

Ein Analogon zum Vektorprodukt in anderen Dimensionen gibt es nicht.

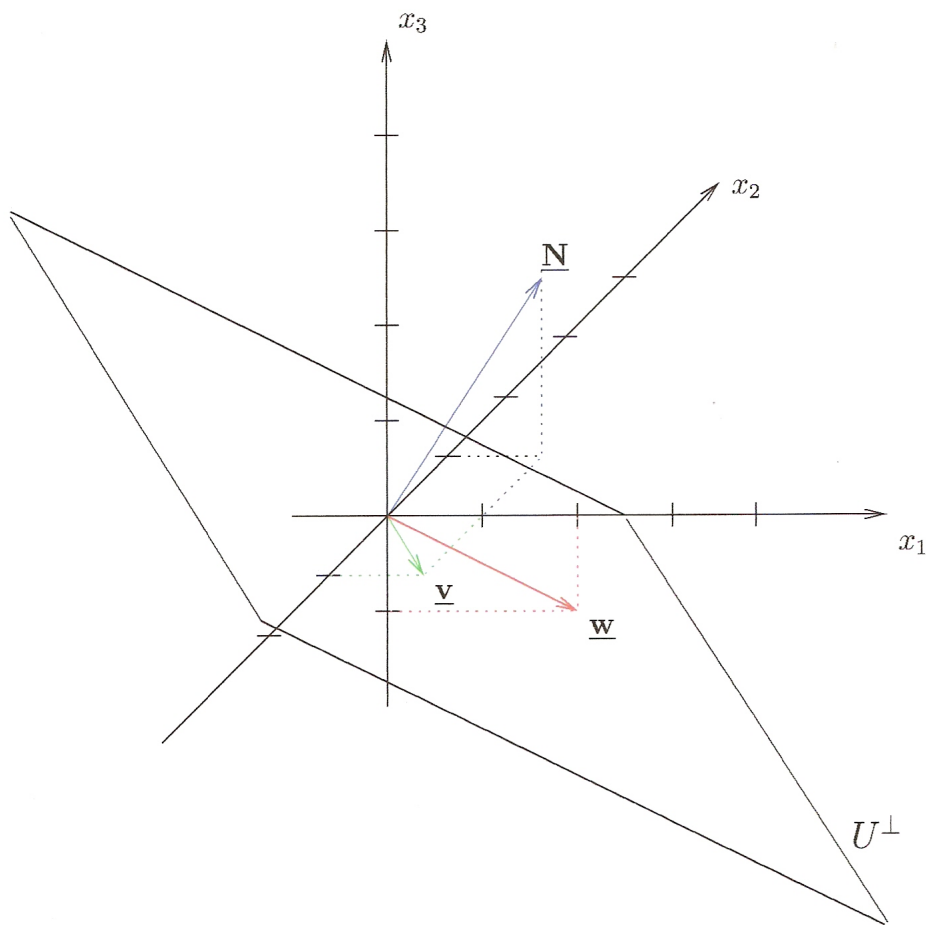
Wie üblich bezeichne  $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(3)})$  die kanonische Basis des  $\mathbb{R}^3$ .

**Definition 11.6.** Das *Vektorprodukt*  $\times$  im  $\mathbb{R}^3$  ist die Abbildung, die allen  $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^3$  einen weiteren Vektor im  $\mathbb{R}^3$  zuordnet,

$$\underline{\mathbf{x}} \times \underline{\mathbf{y}} := \underline{\mathbf{e}}^{(1)} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - \underline{\mathbf{e}}^{(2)} \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + \underline{\mathbf{e}}^{(3)} \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}.$$

Dabei bezeichnet für alle  $a, b, c, d \in \mathbb{R}$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc.$$

Abbildung 11.5: Das orthogonale Komplement von  $\{\underline{\mathbf{N}}\}$ .

**Eigenschaften.** Das Vektorprodukt ist bestimmt durch

- i)  $\|\underline{\mathbf{x}} \times \underline{\mathbf{y}}\| = \|\underline{\mathbf{x}}\| \|\underline{\mathbf{y}}\| |\sin(\varphi)|$ , wobei  $\varphi$  den von  $\underline{\mathbf{x}}$  und  $\underline{\mathbf{y}}$  eingeschlossenen Winkel bezeichne, d.h.:

Der Betrag des Vektorproduktes liefert den **Flächeninhalt des von  $\underline{\mathbf{x}}$  und  $\underline{\mathbf{y}}$  aufgespannten Parallelogramms.**

Insbesondere ist  $\underline{\mathbf{x}} \times \underline{\mathbf{y}}$  **genau dann Null, wenn  $\underline{\mathbf{x}}$  und  $\underline{\mathbf{y}}$  linear abhängig sind.**

- ii)  $\underline{\mathbf{x}} \times \underline{\mathbf{y}}$  **steht senkrecht auf  $\underline{\mathbf{x}}$  und  $\underline{\mathbf{y}}$ .**

- iii) Es gilt die **rechte-Hand-Regel:**

Zeigt der Daumen der rechten Hand in Richtung von  $\underline{\mathbf{x}}$ , der Zeigefinger in Richtung von  $\underline{\mathbf{y}}$ , so zeigt der Mittelfinger senkrecht

dazu in Richtung von  $\underline{\mathbf{x}} \times \underline{\mathbf{y}}$  (vgl. Abbildung 11.6).

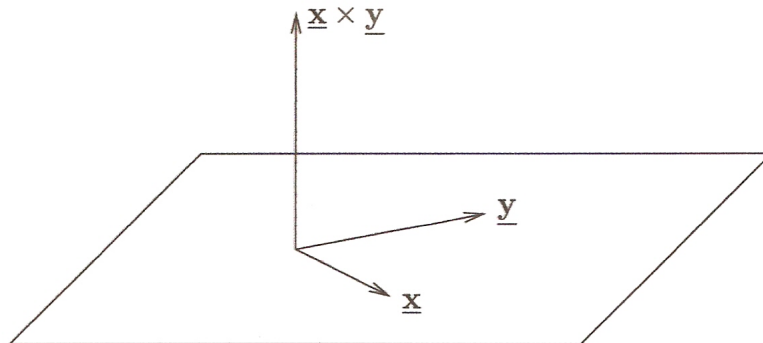


Abbildung 11.6: Zum Vektorprodukt im  $\mathbb{R}^3$ .

**Bemerkung.** Das orientierte Volumen des von  $\underline{\mathbf{a}}^{(1)}$ ,  $\underline{\mathbf{a}}^{(2)}$  und  $\underline{\mathbf{a}}^{(3)}$  im  $\mathbb{R}^3$  aufgespannten Spates (vgl. Abbildung 11.7) berechnet sich zu

$$[\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(3)}] := \langle \underline{\mathbf{a}}^{(1)} \times \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(3)} \rangle .$$

Diese Größe heißt das **Spatprodukt** der Vektoren  $\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(3)} \in \mathbb{R}^3$ .

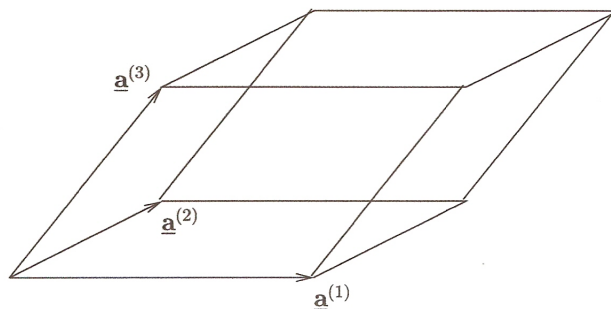


Abbildung 11.7: Ein Spat im  $\mathbb{R}^3$ .

**Rechenregeln.** Für alle  $\underline{\mathbf{x}}, \underline{\mathbf{y}}, \underline{\mathbf{z}} \in \mathbb{R}^3$  und für alle  $\lambda \in \mathbb{R}$  gilt:

- i)  $\underline{\mathbf{x}} \times \underline{\mathbf{y}} = -(\underline{\mathbf{y}} \times \underline{\mathbf{x}})$ ;
- ii)  $(\lambda \underline{\mathbf{x}}) \times \underline{\mathbf{y}} = \lambda(\underline{\mathbf{x}} \times \underline{\mathbf{y}})$ ;
- iii)  $\underline{\mathbf{x}} \times (\underline{\mathbf{y}} + \underline{\mathbf{z}}) = \underline{\mathbf{x}} \times \underline{\mathbf{y}} + \underline{\mathbf{x}} \times \underline{\mathbf{z}}$ .



**Vorsicht.** Das Vektorprodukt ist **im Allgemeinen nicht assoziativ**, d.h. im Allgemeinen gilt

$$\underline{\mathbf{x}} \times (\underline{\mathbf{y}} \times \underline{\mathbf{z}}) \neq (\underline{\mathbf{x}} \times \underline{\mathbf{y}}) \times \underline{\mathbf{z}} .$$

Dies sieht man etwa über den **Entwicklungssatz**

$$\underline{\mathbf{x}} \times (\underline{\mathbf{y}} \times \underline{\mathbf{z}}) = \langle \underline{\mathbf{x}}, \underline{\mathbf{z}} \rangle \underline{\mathbf{y}} - \langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle \underline{\mathbf{z}}$$

ein.

### **Ebenen im $\mathbb{R}^3$ : Ebenen durch den Ursprung.**

- i) Analog zu einer Geraden im  $\mathbb{R}^2$ , die durch den Ursprung verläuft, ist die **Parameterdarstellung einer Ebene**  $E$  im  $\mathbb{R}^3$ , die durch den Ursprung verläuft, durch **zwei linear unabhängige Richtungsvektoren**  $\underline{\mathbf{v}}, \underline{\mathbf{w}} \in \mathbb{R}^3$  bestimmt:

$$E = \{ \underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{x}} = \alpha \underline{\mathbf{v}} + \beta \underline{\mathbf{w}}, \alpha, \beta \in \mathbb{R} \} .$$

Es handelt sich um einen **zweidimensionalen Unterraum des  $\mathbb{R}^3$** .

**Beispiel.** Es seien

$$\underline{\mathbf{v}} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{w}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3$$

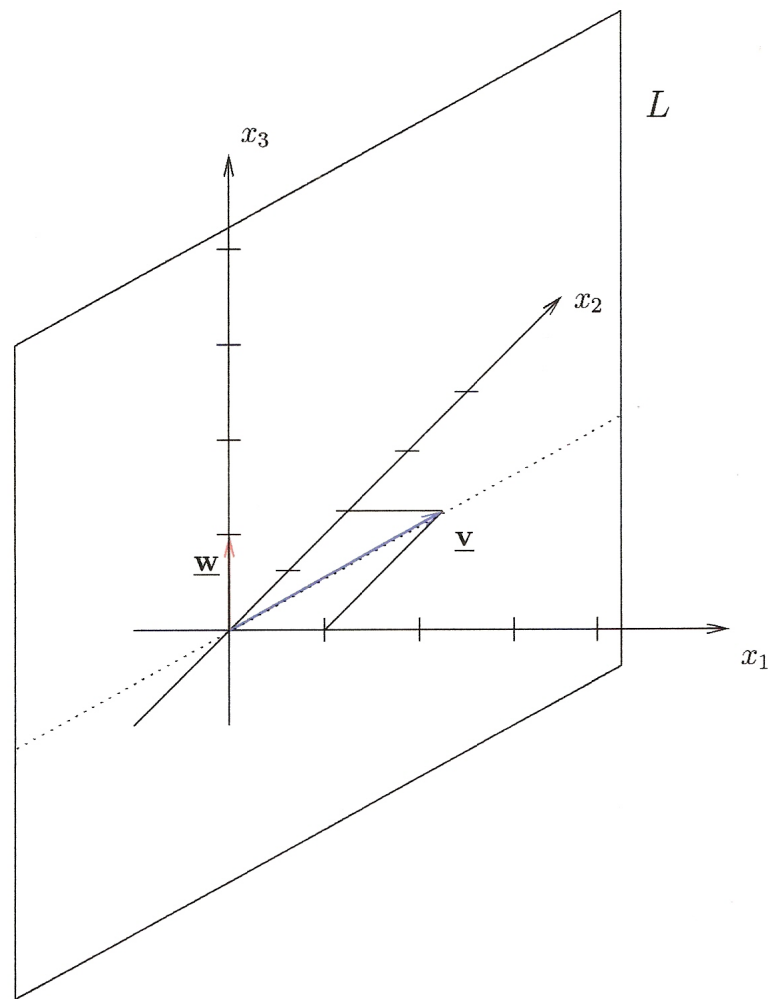
und wie oben (in abkürzender Schreibweise)

$$E = \{ \alpha \underline{\mathbf{v}} + \beta \underline{\mathbf{w}} : \alpha, \beta \in \mathbb{R} \} .$$

Der Unterraum  $E$ , d.h. die Ebene durch den Nullpunkt, die von den beiden Vektoren  $\underline{\mathbf{v}}$  und  $\underline{\mathbf{w}}$  aufgespannt wird, ist in Abbildung 11.8 dargestellt.

**Bemerkung.** Die Richtungsvektoren, die eine gegebene Ebene aufspannen, **sind nicht eindeutig bestimmt**.

Als Übung finde man im Beispiel zwei andere Richtungsvektoren, die dieselbe Ebene aufspannen.

Abbildung 11.8: Die von  $\underline{v}$  und  $\underline{w}$  aufgespannte Ebene.

- ii) Wie bereits anhand eines früheren Beispiels dargelegt (vgl. Abbildung 11.5), kann eine Ebene durch den Ursprung andererseits als eine Menge der Form (ein Unterraum)

$$\{\underline{N}\}^\perp, \quad \underline{N} \neq \underline{0},$$

interpretiert werden, d.h. als die Menge aller Punkte  $\underline{x} \in \mathbb{R}^3$ , die senkrecht auf einem **Normalenvektor**  $\underline{N}$  stehen.

**Bemerkung.** Auch  $\underline{N}$  ist **nicht eindeutig bestimmt**. Jedes Vielfache  $\lambda \underline{N}$ ,  $\lambda \neq 0$ , ist ebenso ein Normalenvektor an dieselbe Ebene.

**Beide Interpretationen sind äquivalent:**

Ergänzt man etwa  $\underline{\mathbf{N}}$  zu einer Orthogonalbasis des  $\mathbb{R}^3$ , so sind die weiteren Basisvektoren linear unabhängige Richtungsvektoren der Ebene.

Sind umgekehrt  $\underline{\mathbf{v}}$  und  $\underline{\mathbf{w}}$  linear unabhängige Richtungsvektoren einer Ebene, so ist

$$\underline{\mathbf{N}} := \underline{\mathbf{v}} \times \underline{\mathbf{w}}$$

ein Normalenvektor.

**Übung.** In den beiden oben angesprochenen Beispielen gebe man jeweils die andere Darstellung an.

**Ebenen im  $\mathbb{R}^3$ : Der allgemeine Fall.**

- i) Eine Ebene, die **nicht notwendig durch den Ursprung** verläuft, ist etwa durch drei **nicht kollineare** Punkte (liegen nicht auf einer Geraden)  $\underline{\mathbf{a}}$ ,  $\underline{\mathbf{b}}$ ,  $\underline{\mathbf{c}}$  auf der Ebene bestimmt.

Mit  $\underline{\mathbf{v}} = \underline{\mathbf{b}} - \underline{\mathbf{a}}$ ,  $\underline{\mathbf{w}} = \underline{\mathbf{c}} - \underline{\mathbf{a}}$  (diese Vektoren sind linear unabhängig, da drei nicht kollineare Punkte gewählt sind) ist diese Ebene die Menge gegeben durch die Parameterdarstellung (vgl. Abbildung 11.9)

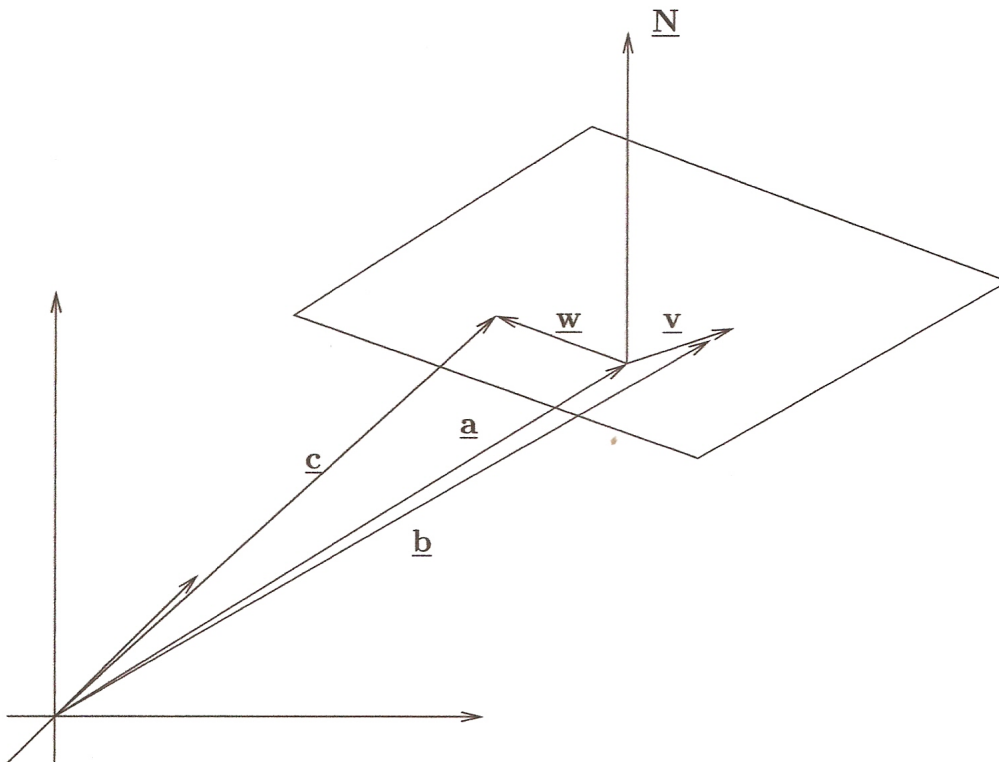
$$E = \{ \underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{x}} = \underline{\mathbf{a}} + \alpha \underline{\mathbf{v}} + \beta \underline{\mathbf{w}}, \alpha, \beta \in \mathbb{R} \} .$$

Es handelt sich nicht mehr um einen zweidimensionalen Unterraum, sondern um einen **zweidimensionalen affinen Raum** (vgl. die Diskussion von Geraden im  $\mathbb{R}^n$ ).

- ii) Wieder ist mit  $\underline{\mathbf{v}} \times \underline{\mathbf{w}}$  ein Normalenvektor an die Ebene bestimmt, dessen **Vorzeichen und Länge noch variiert werden können**.

Zunächst wird **die Länge 1 normiert**, d.h. man setzt

$$\underline{\mathbf{N}} = \pm \frac{\underline{\mathbf{v}} \times \underline{\mathbf{w}}}{\| \underline{\mathbf{v}} \times \underline{\mathbf{w}} \|} ,$$

Abbildung 11.9: Die Ebene durch die Punkte  $\underline{a}$ ,  $\underline{b}$ ,  $\underline{c}$ .

sodass  $\|\underline{N}\| = 1$ .

War die Ebene durch den Ursprung noch durch  $\langle \underline{N}, \underline{x} \rangle = 0$  gekennzeichnet, so bewirkt die Verschiebung im allgemeinen Fall

$$E = \{ \underline{x} \in \mathbb{R}^3 : \langle \underline{N}, \underline{x} \rangle - p = 0 \} .$$

Das Vorzeichen von  $N$  wird dabei so gewählt, dass  $p \geq 0$ .

Dann spricht man von der **Hesseschen Normalform** der Ebene.

**Bemerkung.** Der **Abstand**  $d$  eines beliebigen Punktes  $\underline{x}^{(0)} \in \mathbb{R}^3$  zu einer Ebene errechnet sich aus der Hesseschen Normalform zu

$$d = | \langle \underline{N}, \underline{x}^{(0)} \rangle - p | .$$

**Beispiel.** Eine Ebene sei durch die Punkte

$$\underline{a} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{b} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{c} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

festgelegt. Dann ist

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \underline{\mathbf{N}}$$

normal zu den Richtungsvektoren  $\underline{\mathbf{b}} - \underline{\mathbf{a}}$ ,  $\underline{\mathbf{c}} - \underline{\mathbf{a}}$  und bereits auf die Länge 1 normiert.

Wegen  $\langle \underline{\mathbf{N}}, \underline{\mathbf{a}} \rangle = 1 = p$  ist dabei auch schon das richtige Vorzeichen gewählt.

Die Normalform lautet (Probe!)

$$\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \underline{\mathbf{x}} \right\rangle - 1 = 0$$

**Bemerkungen.** (Vgl. Übungen)

- i)* Schneiden sich zwei ungleiche Ebenen im  $\mathbb{R}^3$ , so erhält man eine Gerade im  $\mathbb{R}^3$ , die auch interpretiert werden kann als Lösung eines Gleichungssystems der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 . \end{aligned}$$

- ii)* Geraden im  $\mathbb{R}^2$  können ebenso auf eine Hessesche Normalform gebracht werden, d.h. man findet  $\underline{\mathbf{n}} \in \mathbb{R}^2$  mit  $\|\underline{\mathbf{n}}\| = 1$  und ein  $p \geq 0$ , sodass die Gerade in der Form

$$\langle \underline{\mathbf{n}}, \underline{\mathbf{x}} \rangle - p = 0$$

geschrieben werden kann.



# Kapitel 12

## Matrizen und lineare Gleichungssysteme

### 12.1 Matrizenkalkül

**Matrizen** sind im Prinzip schon bei der schematischen Diskussion linearer Gleichungssysteme aufgetaucht. In diesem Abschnitt werden die wichtigsten Operationen mit Matrizen dargestellt.

**Im Folgenden werden stets reelle Matrizen betrachtet.** Matrizen mit beispielsweise komplexen Einträgen sind völlig analog zu behandeln.

**Definition 12.1.** *Es seien  $n, m \in \mathbb{N}$ . Ein  $n \times m$  Koeffizientenschema (d.h. eine **Tabelle aus  $n$  Zeilen und  $m$  Spalten**) der Form*

$$(a_{ij})_{i=1, \dots, n}^{j=1, \dots, m} := A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix},$$

$a_{ij} \in \mathbb{R}$  für  $1 \leq i \leq n$ ,  $1 \leq j \leq m$ , heißt eine  $n \times m$  Matrix.

*Notation:*

- i) Können bzgl. der Zeilenzahl und der Spaltenzahl keine Missverständnisse auftreten, so schreibt man oft einfach  $(a_{ij})$  für die Matrix.
- ii) Die Menge der  $n \times m$ -Matrizen wird mit  $M(n, m)$  bezeichnet; weitere Bezeichnungen:  $\mathbb{R}^{(n, m)}$ ,  $\mathbb{R}^{n \times m}$ .

**Merke.** Der Eintrag  $a_{ij}$  steht in der  $i^{\text{ten}}$  Zeile und in der  $j^{\text{ten}}$  Spalte.

### Beispiele.

- i) Die Elemente des  $\mathbb{R}^n$  (Spaltenvektoren) können als  $n \times 1$  Matrizen über  $\mathbb{R}$  aufgefasst werden:

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in M(n, 1).$$

- ii) Das Koeffizientenschema eines linearen Gleichungssystems (vgl. Kapitel 9) ohne die rechte Seite ist eine  $n \times m$  Matrix.

Eine Matrix  $(a_{ij}) \in M(n, m)$  kann mit einem Skalar  $\lambda \in \mathbb{R}$  multipliziert werden:

$$\begin{aligned} \lambda(a_{ij}) &= \lambda \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} = (\lambda a_{ij}) \\ &= \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix}, \end{aligned}$$

zwei Matrizen  $(a_{ij}), (b_{ij}) \in M(n, m)$  können addiert werden:

$$\begin{aligned} (a_{ij}) + (b_{ij}) &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2m} + b_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{nm} + b_{nm} \end{pmatrix}. \end{aligned}$$



Die **Nullmatrix**  $(0)$  ist

$$(0) = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in M(n, m)$$

und es kann elementar nachgerechnet werden:

Die Menge der  $n \times m$  Matrizen  $M(n, m)$  ist ein Vektorraum im Sinne von Definition 11.1.

### Multiplikation von Matrizen.

Matrizen mit **zueinander passenden Formaten** können multipliziert werden:

**Definition 12.2.** *Es seien  $n, m, l \in \mathbb{N}$  und  $A \in M(n, m)$ ,  $B \in M(m, l)$ , d.h. die Spaltenzahl der Matrix  $A$  ist gleich der Zeilenzahl der Matrix  $B$ .*

Dann ist das **Matrizenprodukt**  $AB$  per definitionem eine Matrix

$$C = (c_{ij})_{i=1, \dots, n}^{j=1, \dots, l} \in M(n, l),$$

deren Eintrag in der  $i^{\text{ten}}$  Zeile ( $i = 1, \dots, n$ ) und der  $j^{\text{ten}}$  Spalte ( $j = 1, \dots, l$ ) gegeben ist durch

$$c_{ij} := \sum_{k=1}^m a_{ik} b_{kj}.$$

### Merkregel.

- i) Das Produkt der  $i^{\text{ten}}$  Zeile der Matrix  $A$  mit der  $j^{\text{ten}}$  Spalte der Matrix  $B$  (gemäß  $\sum_{k=1}^m a_{ik} b_{kj}$ ) ergibt den Eintrag in der  $i^{\text{ten}}$  Zeile und der  $j^{\text{ten}}$  Spalte der resultierenden Matrix  $C$ .

ii) Schematisch sieht das wie folgt aus:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{im} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1l} \\ b_{21} & \dots & b_{2j} & \dots & b_{2l} \\ \vdots & & \vdots & & \vdots \\ b_{m1} & \dots & b_{mj} & \dots & b_{ml} \end{pmatrix} \\ = \begin{pmatrix} c_{11} & \dots & c_{1j} & \dots & c_{1l} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \dots & c_{ij} & \dots & c_{il} \\ \vdots & & \vdots & & \vdots \\ c_{n1} & \dots & c_{nj} & \dots & c_{nl} \end{pmatrix} .$$

### Beispiele.

i) Es seien

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \in M(2, 3) ,$$

$$B = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 3 & -1 & -1 & 0 \end{pmatrix} \in M(3, 4) .$$

Dann ist

$$AB = \begin{pmatrix} 5 & 0 & 1 & 2 \\ 7 & 0 & 2 & 1 \end{pmatrix} \in M(2, 4) .$$

ii) Es bezeichne  $I_m$  die quadratische  $m \times m$  Einheitsmatrix,

$$I_m := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix} ,$$

d.h. die Eintragungen  $a_{ii}$ ,  $i = 1, \dots, m$ , auf der Hauptdiagonalen sind 1, alle anderen Eintragungen sind 0.

Für  $A \in M(n, m)$  ist  $AI_m \in M(n, m)$  und es gilt

$$AI_m = A .$$

Für  $A \in M(n, m)$  ist  $I_n A \in M(n, m)$  und es gilt

$$I_n A = A .$$

iii) (a) Eine Matrix  $A = (a_{ij}) \in M(n, m)$  kann mit einer  $m \times 1$  Matrix, d.h. mit einem Vektor  $\underline{x} \in \mathbb{R}^m$  multipliziert werden.

Das Ergebnis ist ein Vektor im  $\mathbb{R}^n$ :

$$A\underline{x} := \underbrace{\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}}_{\in M(n,m)} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}}_{\in \mathbb{R}^m} := \underbrace{\begin{pmatrix} \sum_{k=1}^m a_{1k}x_k \\ \sum_{k=1}^m a_{2k}x_k \\ \vdots \\ \sum_{k=1}^m a_{nk}x_k \end{pmatrix}}_{\in \mathbb{R}^n} .$$

(b) Dementsprechend kann das lineare Gleichungssystem (vgl. Kapitel 9)

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m &= b_n \end{aligned}$$

in der Schreibweise

$$A\underline{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \underline{\mathbf{b}}$$

formuliert werden.

(c) Eine sogenannte lineare Abbildung  $L: V \rightarrow W$  von einem Vektorraum  $V$  in einen Vektorraum  $W$  ist durch die Eigenschaften

$$L(\underline{\mathbf{u}} + \underline{\mathbf{v}}) = L(\underline{\mathbf{u}}) + L(\underline{\mathbf{v}}) \quad \text{für alle } \underline{\mathbf{u}}, \underline{\mathbf{v}} \in V ,$$

$$L(\lambda \underline{\mathbf{v}}) = \lambda L(\underline{\mathbf{v}}) \quad \text{für alle } \underline{\mathbf{v}} \in V , \lambda \in \mathbb{R}$$

gekennzeichnet.

Lineare Abbildungen  $f$  von  $\mathbb{R}$  nach  $\mathbb{R}$  sind **durch eine reelle Zahl  $a \in \mathbb{R}$  charakterisiert**: Für alle  $x \in \mathbb{R}$  ist

$$f(x) = ax .$$

Lineare Abbildungen  $F$  vom  $\mathbb{R}^m$  in den  $\mathbb{R}^n$  sind analog **durch eine Matrix  $A \in M(n, m)$  festgelegt**: Für alle  $\underline{x} \in \mathbb{R}^m$  ist nun

$$F(\underline{x}) = A\underline{x} .$$

### Rechenregeln für die Matrizenmultiplikation.

Es seien  $A, B, C$  Matrizen und  $\lambda \in \mathbb{R}$ . Dann gilt:

- i)*  $(A + B)C = AC + BC$  (Distributivgesetz);
- ii)*  $A(B + C) = AB + AC$  (Distributivgesetz);
- iii)*  $A(BC) = (AB)C$  (Assoziativgesetz);
- iv)*  $A(\lambda B) = (\lambda A)B = \lambda(AB)$ .

**Übung.** Wie müssen in den einzelnen Regeln jeweils die Zeilen- und Spaltenzahlen gewählt werden, damit die Aussagen sinnvoll sind?

**Vorsicht.** Es **übertragen sich nicht alle bekannten Regeln** der Multiplikation etwa reeller Zahlen.

**Beispiel.** Es sei

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \in M(2, 2, \mathbb{R}), \quad B = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix} \in M(2, 2, \mathbb{R}).$$

Dann gilt

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

es ist aber

$$BA = \begin{pmatrix} 10 & 20 \\ -5 & -10 \end{pmatrix}.$$

Mit anderen Worten:

- i) Aus  $AB = (0)$  kann im Allgemeinen nicht gefolgert werden, dass  $A$  oder  $B$  eine Nullmatrix ist.
- ii) Auch im Falle **quadratischer Matrizen** ( $n = m$ ), bei denen sowohl das Produkt  $AB$  als auch das Produkt  $BA$  definiert ist, kann **kein Kommutativgesetz** gelten.

### Transposition von Matrizen.

Eine wichtige Operation mit Matrizen ist die **Transposition**. Dabei werden die **Zeilen und Spalten einer Matrix vertauscht**, d.h.: Ist

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \in M(n, m),$$

so heißt

$$A^T := \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix} \in M(m, n),$$

die **zu  $A$  transponierte Matrix**.

### Beispiele.

- i) Es sei  $A \in M(3, 2)$ ,

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Dann ist

$$A^T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in M(2, 3, \mathbb{R}).$$

- ii) Im Spezialfall eines **Spaltenvektors**  $\underline{x} \in M(n, 1)$ ,

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n,$$

produziert die Transposition einen **Zeilenvektor**

$$\underline{\mathbf{x}}^T = (x_1 \ x_2 \ \dots \ x_n) = \underline{\mathbf{x}} \in M(1, n) .$$

iii) Das Skalarprodukt (vgl. Kapitel 11.2) zweier Vektoren  $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$  kann somit auch als

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle = \underline{\mathbf{y}}^T \underline{\mathbf{x}} = (y_1 \ y_2 \ \dots \ y_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{k=1}^n x_k y_k$$

geschrieben werden.

### Rechenregeln.

i) Man erkennt sofort für  $A, B \in M(n, m)$  und  $\lambda \in \mathbb{R}$ :

$$(a) \ (A + B)^T = A^T + B^T;$$

$$(b) \ (\lambda A)^T = \lambda A^T;$$

$$(c) \ (A^T)^T = A.$$

ii) Für  $A \in M(n, m)$  und  $B \in M(m, l)$  ist das Matrizenprodukt  $AB \in M(n, l)$  definiert.

Wegen  $B^T \in M(l, m)$  und  $A^T \in M(m, n)$  ist aber ebenso die Bildung  $B^T A^T \in M(l, n)$  erlaubt und es gilt (Übung)

$$(AB)^T = B^T A^T .$$

## 12.2 Zur Lösbarkeit linearer Gleichungssysteme

Das Gaußsche Eliminationsverfahren zur expliziten Lösung (falls existent) linearer Gleichungssysteme ist bereits mit Kapitel 9 behandelt.

Aufbauend auf Kapitel 10 und Sektion 11.1, geht es nun um die Frage nach der **Existenz und Eindeutigkeit** von Lösungen und um die **Struktur der Lösungsmenge**.

Im Folgenden werden lineare Gleichungssysteme mit reellen (der Einfachheit halber) Koeffizienten betrachtet.

Für  $n, m \in \mathbb{N}$ ,  $m \leq n$  und  $1 \leq i \leq n$ ,  $1 \leq j \leq m$  seien  $a_{ij}$  und  $b_i \in \mathbb{R}$  fixiert. Zu diesen Daten ist

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m &= b_n \end{aligned}$$

ein lineares System aus  $n$  Gleichungen in  $m$  Unbekannten  $x_1, \dots, x_m$ .

Ist  $A = (a_{ij}) \in M(n, m)$  die Koeffizientenmatrix und  $\underline{\mathbf{b}} \in \mathbb{R}^n$  der Vektor  $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ , so lautet die äquivalente Matrixschreibweise

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}.$$

### Homogenes versus inhomogenes System.

i) Beim homogenen System verschwindet die rechte Seite:

$$A\underline{\mathbf{x}} = \underline{\mathbf{0}}.$$

#### Beobachtung.

(a) Das homogene System ist **immer lösbar** nämlich mit der **trivialen Lösung**  $\underline{\mathbf{x}} = \underline{\mathbf{0}}$ ,

$$A\underline{\mathbf{0}} = \underline{\mathbf{0}}.$$

Es stellt sich die Frage nach der Existenz bzw. der Menge **nicht-trivialer Lösungen**.

(b) **Aufgrund der Linearität** des Systems gilt das **Superpositionsprinzip**, d.h. die **Summe von zwei Lösungen** und das **Vielfache einer Lösung** sind ebenfalls Lösungen.

Gilt nämlich für  $x \in \mathbb{R}^m$  und für  $y \in \mathbb{R}^m$

$$A\underline{\mathbf{x}} = \underline{\mathbf{0}}, \quad \text{und} \quad A\underline{\mathbf{y}} = \underline{\mathbf{0}},$$

so folgt für  $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{y}}$

$$A(\underline{\mathbf{z}}) = A(\underline{\mathbf{x}} + \underline{\mathbf{y}}) = A\underline{\mathbf{x}} + A\underline{\mathbf{y}} = \underline{\mathbf{0}} + \underline{\mathbf{0}} = \underline{\mathbf{0}},$$

d.h. auch  $\underline{\mathbf{z}}$  löst das System.

ii) Beim inhomogenen System ist eine Lösung von

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}} \neq \underline{\mathbf{0}}$$

gesucht.

### Beobachtung.

- (a) Der **Nullvektor ist keine Lösung des inhomogenen Systems**. Hier stellt sich die Frage, **ob bzw. für welche  $\underline{\mathbf{b}} \in \mathbb{R}^n$  eine Lösung existiert** und wie die Lösungsmenge aussieht.
- (b) Beim inhomogenen System ist **die Summe von zwei Lösungen keine Lösung** (analog das Vielfache einer Lösung):  
Aus  $x \in \mathbb{R}^m$  und  $y \in \mathbb{R}^m$  mit

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}, \quad \text{und} \quad A\underline{\mathbf{y}} = \underline{\mathbf{b}},$$

folgt nun für  $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{y}}$

$$A(\underline{\mathbf{z}}) = A(\underline{\mathbf{x}} + \underline{\mathbf{y}}) = A\underline{\mathbf{x}} + A\underline{\mathbf{y}} = \underline{\mathbf{b}} + \underline{\mathbf{b}} = 2\underline{\mathbf{b}} \neq \underline{\mathbf{b}}.$$

(c) Es gilt jedoch:

Ist  $\underline{\mathbf{x}}$  eine **Lösung des homogenen Systems**  $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$  und ist  $\underline{\mathbf{y}}$  eine **Lösung des inhomogenen Systems**  $A\underline{\mathbf{y}} = \underline{\mathbf{b}}$ , so ist  $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{y}}$  eine **Lösung des inhomogenen Systems**:

$$A(\underline{\mathbf{z}}) = A(\underline{\mathbf{x}} + \underline{\mathbf{y}}) = A\underline{\mathbf{x}} + A\underline{\mathbf{y}} = \underline{\mathbf{0}} + \underline{\mathbf{b}} = \underline{\mathbf{b}}.$$

Die obigen Beobachtungen werden zusammengefasst in



**Satz 12.1.** *i) Die Menge der Lösungen des homogenen Systems  $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$  ist ein **Unterraum des  $\mathbb{R}^m$** , genannt der **Kern der Matrix  $A$** .*

*Notation: kern  $A$ .*

*Entweder ist kern  $A = \{\underline{\mathbf{0}}\}$  oder es ist  $\dim(\text{kern } A) = k \in \mathbb{N}$ ,  $1 \leq k \leq m$ .*

*Dann existiert eine Basis  $(\underline{\mathbf{x}}^{(1)}, \underline{\mathbf{x}}^{(2)}, \dots, \underline{\mathbf{x}}^{(k)})$  von kern  $A$  und die **Menge aller Lösungen des homogenen Systems** ist gegeben durch*

$$\underline{\mathbf{x}}_h = \sum_{j=1}^k \lambda_j \underline{\mathbf{x}}^{(j)}, \quad \lambda_i \in \mathbb{R}, \quad i = 1, \dots, k.$$

*Bezeichnung:  $\underline{\mathbf{x}}_h$  heißt die **allgemeine Lösung des homogenen Systems**.*

*ii) Ist  $\underline{\mathbf{x}}_s$  (irgend-) eine Lösung des inhomogenen Gleichungssystems  $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ ,  $\underline{\mathbf{x}}_s$  heißt dann **spezielle Lösung**, so ist*

$$\underline{\mathbf{x}} = \underline{\mathbf{x}}_s + \underline{\mathbf{x}}_h$$

*die **allgemeine Lösung des inhomogenen Systems**.*

### **Bemerkungen.**

*i) “Allgemeine Lösung” bedeutet in beiden Fällen:*

*Für beliebige  $\lambda_i$  ist eine Lösung gegeben, und jede Lösung ist von dieser Form.*

*ii) Kennt man also **alle Lösungen des homogenen Systems** und nur **eine Lösung des inhomogenen System**, so kennt man **alle Lösungen des inhomogenen Systems**.*

### **Der Rang einer Matrix.**

Satz 12.1 beantwortet zwar die Frage nach der **Struktur der Lösungsmenge** eines linearen Gleichungssystems, die Existenz nicht-trivialer Lösungen im homogenen Fall bzw. spezieller Lösungen im inhomogenen Fall wird im Satz jedoch nicht angesprochen.

**Idee.** Man schreibe die Matrix  $A \in M(n, m)$  formal als **Zeile aus Spaltenvektoren**:

$$A = (\underline{\mathbf{a}}^{(1)} \ \underline{\mathbf{a}}^{(2)} \ \dots \ \underline{\mathbf{a}}^{(m)}), \quad \underline{\mathbf{a}}^{(j)} \in \mathbb{R}^n, \quad j = 1, \dots, m.$$

Hier ist für jedes fixierte  $j = 1, \dots, m$

$$\underline{\mathbf{a}}^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} \in \mathbb{R}^n.$$

Weiterhin sei  $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(m)})$  wie üblich die kanonische Basis des  $\mathbb{R}^m$ .

Dann gilt für alle  $j = 1, 2, \dots, m$  (nachrechnen!)

$$A\underline{\mathbf{e}}^{(j)} = \underline{\mathbf{a}}^{(j)}.$$

**Beobachtung.** Identifiziert man wie oben die Matrix  $A$  mit der linearen Abbildung

$$F(\underline{\mathbf{x}}) = A\underline{\mathbf{x}},$$

so bedeutet das:

Die Bilder  $F(\underline{\mathbf{e}}^{(j)})$  der kanonischen Basisvektoren sind genau die Spaltenvektoren der Matrix  $A$ .

Nun kann jeder Vektor  $\underline{\mathbf{x}} \in \mathbb{R}^m$  geschrieben werden als

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \sum_{j=1}^m x_j \underline{\mathbf{e}}^{(j)}, \quad x_j \in \mathbb{R}, \quad j = 1, \dots, m,$$

d.h.

$$A\underline{\mathbf{x}} = \sum_{j=1}^m x_j A\underline{\mathbf{e}}^{(j)} = \sum_{j=1}^m x_j \underline{\mathbf{a}}^{(j)}.$$

Damit das System

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

lösbar sein kann, **muss es also Koeffizienten  $x_i$ ,  $i = 1, 2, \dots, m$ , geben** (die Unbekannten), sodass

$$\underline{\mathbf{b}} = \sum_{j=1}^m x_j \underline{\mathbf{a}}^{(j)} .$$

Mit anderen Worten:  **$\underline{\mathbf{b}}$  ist eine Linearkombination der Spaltenvektoren  $\underline{\mathbf{a}}^{(i)}$  der Matrix  $A$ .**

**Zusammenfassung.** Das lineare Gleichungssystem  $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$  ist genau dann lösbar, wenn

$$\underline{\mathbf{b}} \in \text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)}) ,$$

was äquivalent zu

$$\text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)}) = \text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)}, \underline{\mathbf{b}})$$

ist.

**Definition 12.3.** Es seien  $A \in M(n, m)$  und  $\underline{\mathbf{b}} \in \mathbb{R}^n$ .

i) Der *Spaltenrang* oder der *Rang* der Matrix  $A$  ist die **maximale Anzahl linear unabhängiger Spaltenvektoren** von  $A$ .

*Notation:*  $\text{rg } A$ .

ii) Die  $n \times (m + 1)$  Matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{pmatrix}$$

heißt *erweiterte Matrix*.

*Schreibweise:*

$$(A|\underline{\mathbf{b}}) := \left( \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array} \right) .$$

**Satz 12.2.** Das lineare Gleichungssystem  $A\mathbf{x} = \mathbf{b}$  ist *genau dann lösbar, wenn* gilt

$$\operatorname{rg} A = \operatorname{rg} (A|\mathbf{b}) .$$

**Beispiel.** Es sei  $n = m = 2$  und

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} .$$

Dann ist  $\operatorname{rg} A = 1$ , ebenso ist

$$\operatorname{rg} \left( \begin{array}{cc|c} 1 & 2 & 3 \\ 1 & 2 & 3 \end{array} \right) = 1 ,$$

das System

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

ist lösbar (wie lauten die Lösungen?).

Dahingegen ist

$$\operatorname{rg} \left( \begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 2 & 2 \end{array} \right) = 2 ,$$

das System

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

ist nicht lösbar (ausprobieren!).

### Bemerkungen.

- i)* Analog zum Spaltenrang kann der **Zeilenrang** einer Matrix  $A$  definiert werden. Wegen (siehe Übungen)

$$\text{Zeilenrang } A = \text{Spaltenrang } A$$

- (a) spricht man einfach vom Rang einer Matrix (siehe Definition 12.3),  
 (b) folgt insbesondere

$$\operatorname{rg} A \leq \min\{n, m\} .$$

- ii) Der Rang einer Matrix kann auch nach einer Umformung auf Zeilenstufenform (vgl. Kapitel 10) abgelesen werden.

Die eingangs gestellten Fragen beantwortet schließlich

**Satz 12.3.** *Es sei  $A \in M(n, m)$ .*

*Dann sind die folgenden Aussagen richtig:*

- i) *Es gilt die **Dimensionsformel***

$$\dim(\text{kern } A) + \text{rg } A = m .$$

- ii) *Ist das homogene Gleichungssystem  $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$  **unterbestimmt**, d.h. ist  $n < m$  (weniger Gleichungen als Unbekannte), so hat es **stets nicht-triviale Lösungen**.*

- iii) *Ist  $n = m$ , so ist das lineare Gleichungssystem  $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$  **für jede rechte Seite  $\underline{\mathbf{b}}$  genau dann eindeutig lösbar**, wenn  $\text{rg } A = n$ .*

## 12.3 Quadratische Matrizen

In diesem Abschnitt sei  $A \in M(n, n)$  **stets eine quadratische  $n \times n$  Matrix**. Für nicht-quadratische Matrizen **ergeben die folgenden Betrachtungen keinen Sinn**.

### 12.3.1 Invertierbare Matrizen

Zunächst soll die Frage untersucht werden, wann eine **inverse Matrix  $A^{-1}$**  existiert, für die per definitionem gilt

$$A^{-1}A = AA^{-1} = I_n .$$

Existiert eine solche Matrix, so ist **das lineare Gleichungssystem  $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$  für jede rechte Seite eindeutig lösbar**. Es gilt nämlich in diesem Fall

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}} \quad \Leftrightarrow \quad \underline{\mathbf{x}} = A^{-1}\underline{\mathbf{b}} .$$

Insbesondere folgt aus der Existenz einer inversen Matrix  $\operatorname{rg} A = n$ .

Die Umkehrung ist auch richtig:

**Satz 12.4.** Eine Matrix  $A \in M(n, n)$  besitzt genau dann eine inverse Matrix (*diese ist immer eindeutig bestimmt*), wenn  $A$  maximalen Rang  $\operatorname{rg} A = n$  hat.

Dann heißt die Matrix *regulär*.

Ist die Matrix nicht regulär, so heißt sie *singulär*.

**Beispiel.** Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2, 2, \mathbb{R}).$$

Man verifiziert sofort

$$A \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = I_2$$

und die inverse Matrix existiert genau dann, wenn  $ad - bc \neq 0$ .

Diese Bedingung ist wiederum äquivalent dazu, dass die Vektoren

$$\begin{pmatrix} a \\ c \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} b \\ d \end{pmatrix}$$

linear unabhängig sind (Übung), was genau der Aussage des Satzes entspricht.

**Bemerkung.** In den Übungen wird besprochen, wie inverse Matrizen analog zum Gaußschen Eliminationsverfahren (vgl. Kapitel 10) mithilfe von elementaren Zeilenumformungen berechnet werden können.

**Eigenschaften.** Es seien  $A, B \in M(n, n)$  *invertierbar*. Dann gilt

i)  $(A^{-1})^{-1} = A$ ;

ii)  $(A^T)^{-1} = (A^{-1})^T$ ;

iii)  $(AB)^{-1} = B^{-1}A^{-1}$ .

### 12.3.2 Die Determinante

Trotz der oft sehr technischen Definition der Determinante einer quadratischen Matrix, handelt es sich bei der Determinante um eine geometrisch sehr anschauliche Größe, die in vielen Bereichen der linearen Algebra und der Analysis von grundlegender Bedeutung ist.

#### Die Determinante als geometrisches Objekt.

Zur Motivation sei hier der einfachste Fall  $n = 2$ , d.h. der Fall einer  $2 \times 2$  Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

betrachtet.

Ist  $A$  regulär, so sind die Spaltenvektoren

$$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

linear unabhängig, was genau dann der Fall ist, wenn der **Flächeninhalt des von ihnen aufgespannten Parallelogramms nicht verschwindet**.

Diese Zusammenhang soll jetzt näher analysiert werden, wobei als Erstes festzuhalten ist, welche Eigenschaften von einer “natürlichen” **Volumenfunktion** erwartet werden (im Fall  $n = 2$  soll der **orientierte Flächeninhalt** geliefert werden).

Es werden vier geometrische Forderungen gestellt:

i) Die Volumenfunktion

$$\tau : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

soll zwei Spaltenvektoren ein **orientiertes Volumen** zuordnen, wobei die **Vertauschung von zwei Vektoren lediglich einen Vorzeichenwechsel** zu bewirken hat, d.h. man fordert (vgl. Abbildung 12.1)

$$\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = -\tau(\underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) .$$



Abbildung 12.1:  $\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) > 0$  entspricht der auf der linken Seite angedeuteten Orientierung,  $\tau(\underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) = -\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) < 0$  entspricht der auf der rechten Seite angedeuteten Orientierung.

- ii) Zweitens ist zu fordern, dass ein **Volumen proportional zu einer Verlängerung bzw. Verkürzung einer Seite** ist.

Mit anderen Worten soll für alle  $\lambda \in \mathbb{R}$

$$\tau(\lambda \underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = \lambda \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)})$$

erfüllt sein. Es folgt sofort

$$\tau(\underline{\mathbf{a}}^{(1)}, \lambda \underline{\mathbf{a}}^{(2)}) = -\tau(\lambda \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) = -\lambda \tau(\underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) = \lambda \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) .$$

- iii) Die Wert von

$$\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$$

soll dem **Flächeninhalt des in Abbildung 12.2 grün angedeuteten Parallelogramms** entsprechen.

Der Flächeninhalt des blauen Parallelogramms entspricht der Größe  $\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)})$ . Beachtet man nun, dass die rot bzw. schwarz angedeuteten Dreiecke jeweils gleich sind, so ist die Differenz der Flächeninhalt  $\tau(\tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$  des golden skizzierten Parallelogramms.

Demnach muss gefordert werden:

$$\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) + \tau(\tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)}) ,$$

was unmittelbar auch

$$\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)} + \tilde{\underline{\mathbf{a}}}^{(2)}) = \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) + \tau(\underline{\mathbf{a}}^{(1)}, \tilde{\underline{\mathbf{a}}}^{(2)}) .$$

ergibt.



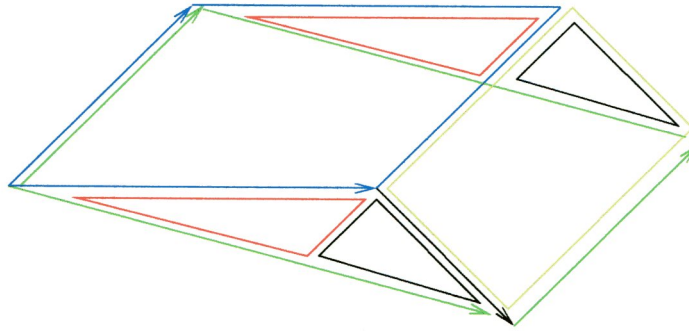


Abbildung 12.2: Zur Berechnung von  $\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$ .

*iv)* Ist schließlich  $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$  die kanonische Basis des  $\mathbb{R}^2$ , so **normiert** man

$$\tau(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}) = 1 ,$$

genau wie es dem Flächeninhalt des Quadrates mit der Seitenlänge 1 entspricht.

### Definition der Determinante.

Es gibt **genau eine** orientierte Volumenform  $\tau$ ,

$$\tau \left( \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \right) =: \det A ,$$

die alle obigen Forderungen *i)* – *iv)* erfüllt:

**Definition 12.4.** *Es sei  $A = (a_{ij}) \in M(n, n)$ .*

*Dann ist die Determinante*

$$\det A =: \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

*wie folgt definiert:*

*i)* Ist  $n = 2$ , so ist

$$\det A := a_{11}a_{22} - a_{12}a_{21} .$$

ii) Ist  $n = 3$ , so berechnet sich die Determinante aus der *Regel von Sarrus* (vgl. Abbildung 12.3)

$$\det A := a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} .$$

iii) Zur *induktiven Definition* im allgemeinen Fall  $n \geq 3$  werden zunächst die *Streichungsmatrizen*  $A_{ij}$  von  $A$  definiert:

Für  $i, j = 1, \dots, n$  geht die  $(n-1) \times (n-1)$  Matrix  $A_{ij}$  aus  $A$  durch die *Streichung der  $i^{\text{ten}}$  Zeile und der  $j^{\text{ten}}$  Spalte* hervor (vgl. Abbildung 12.4).

Dann gilt der folgende *Laplacesche Entwicklungssatz*.

(a) *Entwicklung nach der  $i^{\text{ten}}$  Zeile:*

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}$$

(b) *Entwicklung nach der  $j^{\text{ten}}$  Spalte:*

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij} .$$

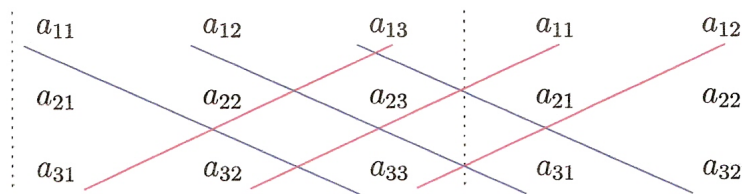


Abbildung 12.3: Zur Regel von Sarrus: **Positive** Vorzeichen hat man für die **blau unterstrichenen** Produkte, **negative** Vorzeichen für die **rot unterstrichenen**.

## Beispiele.

i) Es sei  $n = 3$  und

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in M(3, 3) .$$

$$\begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}$$

Abbildung 12.4: Die  $i^{te}$  Zeile und die  $j^{te}$  Spalte werden aus  $A$  entfernt, um die Streichungsmatrix  $A_{ij}$  zu erhalten.

Dann ist

$$\begin{aligned} \det A &= 1 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 0 + 3 \cdot 1 \cdot 1 - 3 \cdot 1 \cdot 0 - 1 \cdot 1 \cdot 1 - 2 \cdot 1 \cdot 1 \\ &= 1. \end{aligned}$$

ii) Es sei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix} \in M(3, 3).$$

Eine **Entwicklung nach der ersten Zeile** liefert

$$\begin{aligned} \det A &= 1 \begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} + 3 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} \\ &= 0 - 4 + 6 = 2. \end{aligned}$$

Eine **Entwicklung nach der zweiten Spalte** liefert

$$\begin{aligned} \det A &= -2 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} + 2 \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} - 1 \begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} \\ &= -4 + 2 + 4 = 2. \end{aligned}$$

iii) Es sei

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \in M(4, 4).$$

Bei der Entwicklung von  $\det A$  wird man **die Nullen in der ersten Spalte** ausnutzen, d.h. nach der ersten Spalte entwickeln.

Dies ergibt (vgl. Beispiel *ii*))

$$\det A = 1 \begin{vmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 0 & 1 & 1 \end{vmatrix} = 2 .$$

### Eigenschaften der Determinante.

Eine der wichtigsten Eigenschaften der Determinante ist festgehalten in

**Satz 12.5.** *Es sei  $A \in M(n, n)$ .*

*Dann sind die folgenden Aussagen äquivalent:*

- i) Es ist  $\det A \neq 0$ .*
- ii) Die **Zeilenvektoren** von  $A$  sind **linear unabhängig**.*
- iii) Die **Spaltenvektoren** von  $A$  sind **linear unabhängig**.*
- iv) Die Matrix  $A$  ist **regulär**.*

Weitere Eigenschaften sind leicht einzusehen:

- i) Im Allgemeinen gilt für  $A, B \in M(n, n)$*

$$\det (A + B) \neq \det A + \det B .$$

- ii) Ist  $A \in M(n, n)$  und  $\lambda \in \mathbb{R}$ , so gilt*

$$\det (\lambda A) = \lambda^n \det A .$$

- iii) Es gilt der **Determinantenmultiplikationssatz**: Sind  $A, B \in M(n, n)$ , so ist*

$$\det (AB) = \det A \det B .$$

- iv) Ist  $A \in M(n, n)$  regulär, so ist*

$$\det A^{-1} = \frac{1}{\det A} .$$

### 12.3.3 Spezialfälle quadratischer Matrizen

Oft begegnet man Matrizen mit einer besonderen Struktur, die beispielsweise eine geometrische Operation widerspiegelt.

Eine Matrix  $A \in M(n, n)$  heißt

- i) **symmetrisch**, wenn  $A = A^T$ ;
- ii) **orthogonal**, wenn  $AA^T = A^T A = I_n$ ;
- iii) **positiv definit**, wenn  $A$  symmetrisch ist und wenn für alle  $\underline{\mathbf{x}} \in \mathbb{R}^n$ ,  $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ , gilt

$$\underline{\mathbf{x}}^T A \underline{\mathbf{x}} > 0 .$$

Matrizen mit diesen oder ähnlichen Eigenschaften werden in den folgenden Kapiteln immer wieder auftauchen.

In diesem Abschnitt liegt ein besonderes Augenmerk auf der Klasse:

#### Orthogonale Matrizen.

**Beispiel.** Für fixiertes  $\varphi \in [0, 2\pi)$  betrachte man die Matrix

$$A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \in M(2, 2) .$$

Man beschreibe die Abbildung  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $F(\underline{\mathbf{x}}) = A\underline{\mathbf{x}}$  für alle  $\underline{\mathbf{x}} \in \mathbb{R}^2$ , geometrisch. Dieses und weitere Beispiele werden in den Übungen diskutiert.

Analog zu Kapitel 12.2 wird nun eine orthogonale Matrix  $A \in M(n, n)$  als **Zeile von Spaltenvektoren** geschrieben,

$$A = (\underline{\mathbf{a}}^{(1)} \ \underline{\mathbf{a}}^{(2)} \ \dots \ \underline{\mathbf{a}}^{(n)}) , \quad \underline{\mathbf{a}}^{(i)} \in \mathbb{R}^n , \quad i = 1, \dots, m ,$$

womit  $A^T \in M(n, n)$  als **Spalte von Zeilenvektoren** geschrieben werden kann:

$$A^T = \begin{pmatrix} (\underline{\mathbf{a}}^{(1)})^T \\ (\underline{\mathbf{a}}^{(2)})^T \\ \vdots \\ (\underline{\mathbf{a}}^{(n)})^T \end{pmatrix} .$$

Aus der Orthogonalität folgt

$$(\delta_{ij}) = I_n = A^T A = \begin{pmatrix} (\underline{\mathbf{a}}^{(1)})^T \\ (\underline{\mathbf{a}}^{(2)})^T \\ \vdots \\ (\underline{\mathbf{a}}^{(n)})^T \end{pmatrix} (\underline{\mathbf{a}}^{(1)} \ \underline{\mathbf{a}}^{(2)} \ \dots \ \underline{\mathbf{a}}^{(n)}) = (\langle \underline{\mathbf{a}}^{(i)}, \underline{\mathbf{a}}^{(j)} \rangle) :$$

Die Spaltenvektoren einer orthogonalen Matrix bilden eine Orthonormalbasis des  $\mathbb{R}^n$ .

**Bemerkung.** Aus der Definition folgt für orthogonale Matrizen

$$A^T = A^{-1},$$

und

$$\det A^T = \det A$$

(Übung: Dies gilt für alle  $A \in M(n, n)$ .) ergibt zusammen mit dem Determinantenmultiplikationssatz

$$\det A = \det A^T = \det A^{-1} = \frac{1}{\det A},$$

d.h.: Für jede orthogonale Matrix  $A \in M(n, n)$  gilt

$$\det A = \pm 1.$$

Aber: Aus  $\det A = \pm 1$  folgt nicht, dass  $A$  orthogonal ist (Beispiel?).

### 12.3.4 Hauptachsentransformation: Eigenwerte und Eigenvektoren

Alle bisherigen Betrachtungen dieses Kapitels beziehen sich im Wesentlichen auf die Standardbasis des  $\mathbb{R}^n$ .

Nun soll aufgezeigt werden, wie man sich von dieser Einschränkung lösen kann und wie bei konkreten Problemen zu geeigneteren Basen übergegangen werden kann.

Zur Anschaulichkeit der Rechnungen wird hier **o.E. der  $\mathbb{R}^2$  betrachtet** - die Rechnungen im  $\mathbb{R}^n$ ,  $n \geq 3$ , verlaufen formal identisch.

### **Neue Koordinaten.**

Es sei  $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$  die kanonische Basis des  $\mathbb{R}^2$  und  $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)})$ ,

$$\begin{aligned}\underline{\mathbf{f}}^{(1)} &= a\underline{\mathbf{e}}^{(1)} + b\underline{\mathbf{e}}^{(2)}, \\ \underline{\mathbf{f}}^{(2)} &= c\underline{\mathbf{e}}^{(1)} + d\underline{\mathbf{e}}^{(2)}.\end{aligned}$$

sei eine weitere Basis des  $\mathbb{R}^2$ . Hier sind  $a, b, c, d \in \mathbb{R}$  und es gelte (**man beachte die Anordnung der Koeffizienten!**)

$$S = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in M(2, 2) \quad \text{mit} \quad \det S \neq 0.$$

**Merke.** Die neuen Basisvektoren sind die **Spaltenvektoren** von  $S$ .

Wegen  $\det S \neq 0$  ist dabei  $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)})$  tatsächlich eine Basis des  $\mathbb{R}^2$  (warum?).

Ein beliebiger Vektor  $\underline{\mathbf{v}} \in \mathbb{R}^2$  kann **sowohl bzgl. der kanonischen Basis als auch bzgl. der neuen Basis  $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)})$  dargestellt werden** (vgl. Kapitel 11.1), d.h. es gibt eindeutig bestimmte Koeffizienten  $u_1, u_2$  mit

$$\begin{aligned}\underline{\mathbf{v}} &= \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = v_1\underline{\mathbf{e}}^{(1)} + v_2\underline{\mathbf{e}}^{(2)} \\ &= u_1\underline{\mathbf{f}}^{(1)} + u_2\underline{\mathbf{f}}^{(2)}.\end{aligned}$$

Einsetzen ergibt

$$\underline{\mathbf{v}} = (au_1 + cu_2)\underline{\mathbf{e}}^{(1)} + (bu_1 + du_2)\underline{\mathbf{e}}^{(2)}.$$

Mit der **Transformationsmatrix**  $S$  kann diese Beziehung wie folgt interpretiert werden:

**Die Koordinaten bzgl. der Standardbasis und die Koordinaten bzgl. der neuen Basis  $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)})$  erfüllen die Gleichung**

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = S \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

Mit anderen Worten: **Die neuen Koordinaten ergeben sich aus**

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = S^{-1} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} .$$

### Zwei Beispiele.

**Beispiel 1.** Im  $\mathbb{R}^2$  betrachte man die Ellipse

$$E = \left\{ \underline{\mathbf{x}} \in \mathbb{R}^2 : \frac{x_1^2}{4} + x_2^2 = 1 \right\} .$$

mit den Längen 2 und 1 der Halbachsen (vgl. Abbildung 12.5).

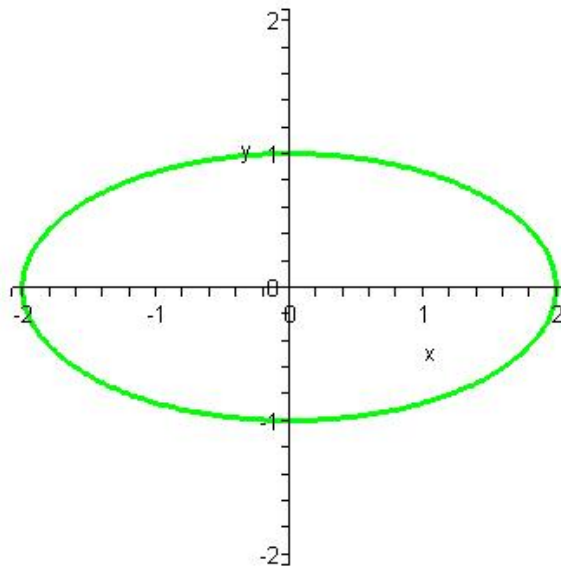


Abbildung 12.5: Die Ellipse  $E$ .

Ist

$$A = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & 1 \end{pmatrix} ,$$

so kann dies äquivalent als

$$E = \left\{ \underline{\mathbf{x}} \in \mathbb{R}^2 : \underline{\mathbf{x}}^T A \underline{\mathbf{x}} = 1 \right\}$$

geschrieben werden.



**Beobachtung.** Die Achsen der Ellipse zeigen in Richtung der kanonischen Basisvektoren und die Matrix  $A$  angewandt auf einen Basisvektoren ergibt jeweils ein Vielfaches des Basisvektors:

$$\begin{aligned} A \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \frac{1}{4} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ A \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Man sagt dazu:

Die Zahlen  $\lambda_1 = 1/4$  und  $\lambda_2 = 1$  heißen **Eigenwerte** der Matrix  $A$  und  $\underline{\mathbf{e}}^{(1)}$ ,  $\underline{\mathbf{e}}^{(2)}$  sind dazugehörige **Eigenvektoren**.

**Bemerkung.** Alle Vielfachen eines Eigenvektors sind ebenfalls Eigenvektoren zum gleichen Eigenwert.

**Beispiel 2.** Nun wird die Menge

$$\tilde{E} = \left\{ \underline{\mathbf{x}} \in \mathbb{R}^2 : \frac{1}{8} [5(x_1^2 + x_2^2) - 6x_1x_2] = 1 \right\}$$

betrachtet, deren geometrische Struktur in dieser Form nicht ersichtlich ist.

Mit der Matrix

$$\tilde{A} = \begin{pmatrix} \frac{5}{8} & -\frac{3}{8} \\ -\frac{3}{8} & \frac{5}{8} \end{pmatrix}$$

lautet die äquivalente Schreibweise in diesem Beispiel

$$\tilde{E} = \left\{ \underline{\mathbf{x}} \in \mathbb{R}^2 : \underline{\mathbf{x}}^T \tilde{A} \underline{\mathbf{x}} = 1 \right\}.$$

Wenn **Eigenwerte und Eigenvektoren eine geometrische Bedeutung haben**, dann sind diese zunächst zu berechnen:

Per definitionem ist  $\lambda \in \mathbb{R}$  ein Eigenwert der Matrix  $\tilde{A}$ , falls ein Vektor  $\underline{\mathbf{v}} \neq \underline{\mathbf{0}} \in \mathbb{R}^2$  existiert mit

$$\tilde{A} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda I_2 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

wobei  $I_2$  wie üblich die Einheitsmatrix

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

bezeichnet.

Diese Gleichung wird geschrieben als

$$(\tilde{A} - \lambda I_2) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{\mathbf{0}}.$$

Nach Satz 12.5 hat dieses homogene lineare Gleichungssystem **genau dann eine Lösung**, wenn

$$\det(\tilde{A} - \lambda I_2) = \begin{vmatrix} \frac{5}{8} - \lambda & -\frac{3}{8} \\ -\frac{3}{8} & \frac{5}{8} - \lambda \end{vmatrix} = 0,$$

d.h. falls

$$\lambda^2 - \frac{10}{8}\lambda + \frac{1}{4} = 0.$$

Man findet die beiden Eigenwerte

$$\lambda_1 = \frac{1}{4} \quad \text{und} \quad \lambda_2 = 1.$$

Die Gleichungssysteme

$$\tilde{A} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad \text{und} \quad \tilde{A} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 1 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

liefern dazu die Eigenvektoren (hier auf Länge 1 normiert - jedes Vielfache hätte ebenso genommen werden können)

$$\underline{\mathbf{f}}^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \underline{\mathbf{f}}^{(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Die Eigenvektoren bilden eine Basis des  $\mathbb{R}^2$  und nach den Betrachtungen zu Beginn dieses Abschnittes ist die Matrix

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

zu bilden, wobei die **Orthonormalität der Eigenvektoren**

$$S^{-1} = S^T$$

liefert.

**Ergebnis.** Wird  $\underline{\mathbf{x}} \in \mathbb{R}^2$  bzgl. der neuen Orthonormalbasis  $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)})$  dargestellt,

$$\underline{\mathbf{x}} = \alpha_1 \underline{\mathbf{f}}^{(1)} + \alpha_2 \underline{\mathbf{f}}^{(2)}, \quad \alpha_1, \alpha_2 \in \mathbb{R},$$

so gilt wegen

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = S \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad \text{und wegen} \quad \underline{\mathbf{x}}^T = (x_1 \ x_2) = (\alpha_1 \ \alpha_2) S^T$$

die Beziehung

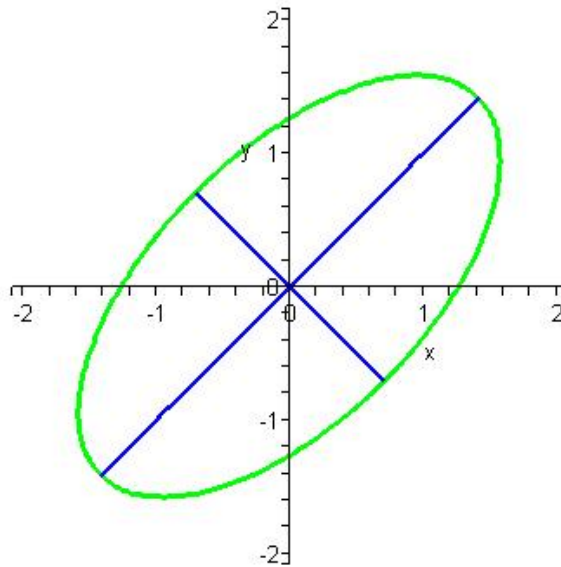
$$\underline{\mathbf{x}}^T \tilde{A} \underline{\mathbf{x}} = (\alpha_1 \ \alpha_2) \underbrace{S^T \tilde{A} S}_{=: B} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = (\alpha_1 \ \alpha_2) B \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}.$$

Aufgrund der Konstruktion hat  $B$  dabei automatisch **Diagonalgestalt**, d.h. auf der **Hauptdiagonalen** stehen die Eigenwerte und alle anderen Einträge verschwinden.

Es ist demnach (nachrechnen!)

$$B = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & 1 \end{pmatrix}.$$

- i) **Geometrisch wird die Menge  $\tilde{E}$  bzgl. der neuen Basis durch die Matrix  $B$  dargestellt**, die mit der darstellenden Matrix  $A$  des ersten Beispiels bzgl. der kanonischen Basis überstimmt.
- ii) Die neue Basis  $(\tilde{f}^{(1)}, \tilde{f}^{(2)})$  entsteht aus der kanonischen Basis durch eine **Rotation um  $\pi/4$**  und dementsprechend ist die Menge  $\tilde{E}$  nichts anderes als die um  $\pi/4$  rotierte Ellipse  $E$  (vgl. Abbildung 12.6).
- iii) Zeigen die Achsen von  $E$  in Richtung der Koordinatenachsen als Richtungen der Eigenvektoren von  $A$ , so zeigen die **Achsen von  $\tilde{E}$  in Richtung der Eigenvektoren von  $\tilde{A}$** .

Abbildung 12.6: Die Ellipse  $\tilde{E}$ .

### Zusammenfassung.

**Definition 12.5.** Es sei  $A \in M(n, n)$ .

- i) Ein Vektor  $\underline{\mathbf{v}} \neq \underline{\mathbf{0}}$  heißt Eigenvektor der Matrix  $A$  zum Eigenwert  $\lambda \in \mathbb{R}$ , falls

$$A\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}} .$$

- ii) Dabei berechnen sich die möglichen Eigenwerte als Lösungen der Gleichung

$$\det (A - \lambda I_n) = 0 .$$

**Bemerkung.** Die Anzahl und die Vielfachheit der reellen Eigenwerte und die Dimension der Eigenräume, d.h. der Unterräume aller Eigenvektoren zu jedem Eigenwert, ist a priori nicht direkt ersichtlich.

Als Beispiel berechne man alle Eigenwerte und Eigenvektoren der  $2 \times 2$  Matrizen

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} , \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} .$$

Im **Spezialfall symmetrischer Matrizen** ist die obige Konstruktion hingegen stets möglich.

Es gilt der **Satz über die Hauptachsentransformation**:

**Satz 12.6.** *Es sei  $A \in M(n, n)$  symmetrisch. d.h. es gelte  $A^T = A$ .*

*i) Dann gibt es eine **orthogonale Matrix**  $S$ , die  $A$  auf **Diagonalgestalt** transformiert*

$$S^T A S = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

*ii) Dabei sind die  $\lambda_i$ ,  $i = 1, \dots, n$ , **reelle Eigenwerte** von  $A$  (nicht notwendig verschieden) und die **Spaltenvektoren der Transformationsmatrix  $S$**  bilden eine **Orthonormalbasis aus Eigenvektoren**.*



# Kapitel 13

## Grenzwerte und Stetigkeit im $\mathbb{R}^n$

In diesem sehr kurzen Kapitel wird im Wesentlichen nur an die Kapitel 5 und 7 zur Konvergenz reeller Zahlenfolgen und zur Stetigkeit von Funktionen einer Veränderlichen erinnert.

Im Vergleich zu den genannten Kapiteln wird die **reelle Betragsfunktion**  $|\cdot|$  durch die **Euklidische Norm**  $\|\cdot\|$  im  $\mathbb{R}^n$  ersetzt.

### **Konvergente Folgen im $\mathbb{R}^n$ .**

Der Folgenbegriff aus Definition 5.1 ist bereits für beliebige Mengen  $A$  – also insbesondere im  $\mathbb{R}^n$  – definiert.

Konvergenz bedeutet (vgl. Definition 5.2):

**Definition 13.1.** Eine Folge  $\{\underline{\mathbf{a}}^{(k)}\}$ ,  $\underline{\mathbf{a}}^{(k)} \in \mathbb{R}^n$  für alle  $k \in \mathbb{N}$ , heißt konvergent, wenn es ein  $\underline{\mathbf{a}} \in \mathbb{R}^n$  gibt mit der Eigenschaft:

*Zu jedem* (noch so kleinen)  $\varepsilon > 0$  existiert eine (meist von  $\varepsilon$  abhängende) Zahl  $N = N(\varepsilon) \in \mathbb{N}$  mit

$$\|\underline{\mathbf{a}}^{(k)} - \underline{\mathbf{a}}\| < \varepsilon \quad \text{für alle } k \geq N(\varepsilon).$$

Dann heißt  $\underline{\mathbf{a}}$  der **Grenzwert** oder **Limes** der Folge.

*Notation:*

$$\lim_{k \rightarrow \infty} \underline{\mathbf{a}}^{(k)} = \underline{\mathbf{a}} \quad \text{oder} \quad \underline{\mathbf{a}}^{(k)} \rightarrow \underline{\mathbf{a}} \quad \text{für } n \rightarrow \infty.$$

Eine Folge, die nicht konvergiert, heißt **divergent**.

Die anschauliche Vorstellung ändert sich im Vergleich zu Kapitel 5 nur insoweit, dass **Abbildung 5.2 durch Abbildung 13.1 zu ersetzen ist.**

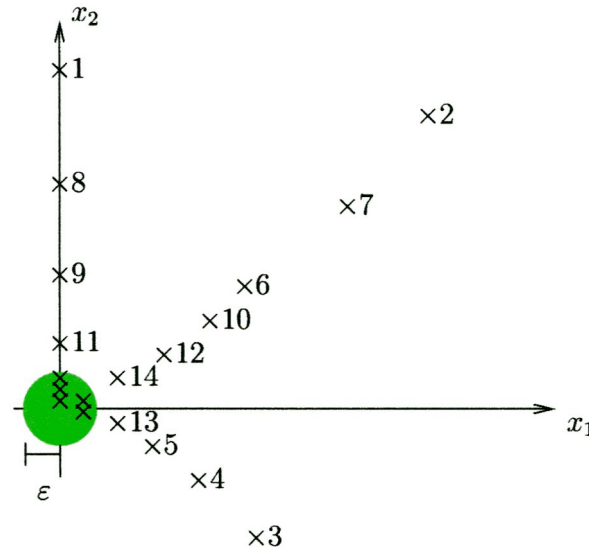


Abbildung 13.1: Für dieses  $\varepsilon$  sind ab  $k = 15$  alle Folgenglieder näher bei  $\underline{\mathbf{0}}$  als  $\varepsilon$ .

Wie der folgende Satz zeigt, **ändern sich die konkreten Rechnungen tatsächlich nicht:**

**Satz 13.1.** Es seien  $\underline{\mathbf{a}}^{(k)} = \begin{pmatrix} a_1^{(k)} \\ \vdots \\ a_n^{(k)} \end{pmatrix}$ ,  $k \in \mathbb{N}$ , und  $\underline{\mathbf{a}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ .

Die Folge  $\{\underline{\mathbf{a}}^{(k)}\}$  ist **genau dann** konvergent gegen  $\underline{\mathbf{a}}$ , falls für jedes feste  $i$ ,  $i = 1, \dots, n$ , gilt:

$$\lim_{k \rightarrow \infty} a_i^{(k)} = a_i .$$

Mit anderen Worten: Es gilt  $\lim_{k \rightarrow \infty} \underline{\mathbf{a}}^{(k)} = \underline{\mathbf{a}}$  genau dann, wenn **komponentenweise** Konvergenz (**für alle Komponenten!**) vorliegt.



**Beispiele.**

i) Es sei  $n = 3$  und

$$\underline{\mathbf{a}}^{(k)} = \begin{pmatrix} \frac{2k^3+1}{k^3+2k-1} \\ \left(1 + \frac{1}{k}\right)^k \\ \frac{k!}{k^k} \end{pmatrix} \quad \text{für alle } k \in \mathbb{N} .$$

Die Folge  $\{\underline{\mathbf{a}}^{(k)}\}$  konvergiert gegen  $\underline{\mathbf{a}} = \begin{pmatrix} 2 \\ e \\ 0 \end{pmatrix}$ .

ii) Ist für  $n = 2$  und für alle  $k \in \mathbb{N}$

$$\underline{\mathbf{a}}^{(k)} = \begin{pmatrix} 1/k \\ k \end{pmatrix} ,$$

so divergiert die Folge  $\{\underline{\mathbf{a}}^{(k)}\}$ .

**Stetigkeit von Funktionen mehrerer Veränderlicher.**

Wie der Grenzwertbegriff für Folgen überträgt sich der Stetigkeitsbegriff auf mehrere Veränderliche, wenn **reelle Beträge durch Euklidische Normen ersetzt werden**.

Bei der **geometrischen Vorstellung** gibt es allerdings einen großen Unterschied:

Bei Funktionen einer Veränderlichen muss **nur "eine Richtung"**, d.h. der Zahlenstrahl betrachtet werden.

Bei Funktionen mehrerer Veränderlicher ist zu beachten: Lläuft man **aus unterschiedlichen Richtungen** auf einen fixierten Punkt  $\underline{\mathbf{x}}_0$  zu, so kann die Funktion **völlig unterschiedliches Verhalten** zeigen.

Im Folgenden ist der Einfachheit halber  $U \subset \mathbb{R}^m$  **stets eine offene Menge**.

Das bedeutet, dass um jeden Punkt eine kleine Kugel gelegt werden kann, die ganz in der Menge liegt, und ist in gewissem Sinne eine

Verallgemeinerung des Intervallbegriffs  $(a, b)$  in  $\mathbb{R}$ .

Weiter ist  $\underline{\mathbf{f}}$  stets eine Funktion  $U \rightarrow \mathbb{R}^n$ .

**Beobachtung.** Ist  $\{\underline{\mathbf{x}}^{(k)}\}$  eine Folge von Elementen aus  $U \subset \mathbb{R}^m$ , so ist  $\{\underline{\mathbf{f}}(\underline{\mathbf{x}}^{(k)})\}$  eine Folge im  $\mathbb{R}^n$ , die auf Konvergenz untersucht werden kann.

**Definition 13.2.** i) Es seien  $\underline{\mathbf{x}}^{(0)} \in U$  und  $\underline{\mathbf{a}} \in \mathbb{R}^n$  fixiert.

Dann hat  $\underline{\mathbf{f}}$  an der Stelle  $\underline{\mathbf{x}}^{(0)}$  den Grenzwert  $\underline{\mathbf{a}}$ , falls für jede Folge  $\{\underline{\mathbf{x}}^{(k)}\}$  aus  $U$  mit  $\underline{\mathbf{x}}^{(k)} \neq \underline{\mathbf{x}}^{(0)}$  für alle  $k \in \mathbb{N}$  und mit  $\underline{\mathbf{x}}^{(k)} \rightarrow \underline{\mathbf{x}}^{(0)}$  für  $k \rightarrow \infty$  gilt

$$\lim_{k \rightarrow \infty} \underline{\mathbf{f}}(\underline{\mathbf{x}}^{(k)}) = \underline{\mathbf{a}}.$$

Notation:

$$\lim_{\underline{\mathbf{x}} \rightarrow \underline{\mathbf{x}}^{(0)}} \underline{\mathbf{f}}(\underline{\mathbf{x}}) = \underline{\mathbf{a}}.$$

ii) Die Funktion  $\underline{\mathbf{f}}$  heißt stetig in einem Punkt  $\underline{\mathbf{x}}^{(0)} \in U$ , falls

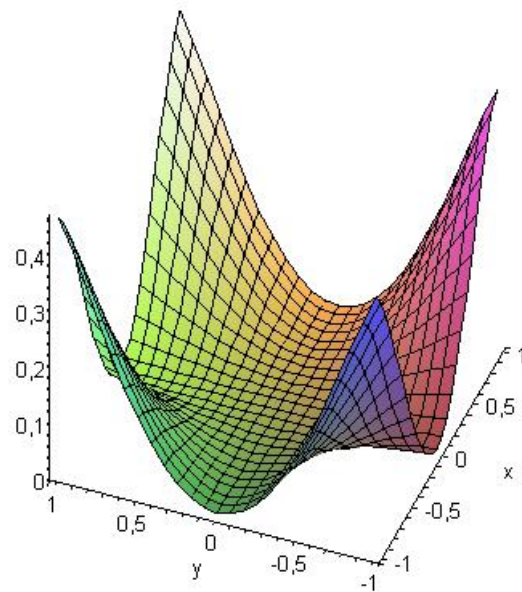
$$\lim_{\underline{\mathbf{x}} \rightarrow \underline{\mathbf{x}}^{(0)}} \underline{\mathbf{f}}(\underline{\mathbf{x}}) = \underline{\mathbf{f}}(\underline{\mathbf{x}}^{(0)}).$$

iii) Die Funktion heißt stetig auf  $U$ , wenn sie in jedem Punkt  $\underline{\mathbf{x}}^{(0)} \in U$  stetig ist.

Die Rechenregeln für stetige Funktionen aus Kapitel 7 übertragen sich identisch.

### Beispiele.

- i) Der Graph einer stetigen Funktion (intuitive Vorstellung) und der Graph der Betragsfunktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  sind in den Abbildungen 13.2, 13.3 dargestellt.
- ii) In Abbildung 13.4 erkennt man eine Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , die als Funktion der einen Variable stetig, als Funktion der anderen Variable unstetig ist.
- iii) Analog zur Abbildung 7.4 ist es anhand von Abbildung 13.5 schwierig zu entscheiden, ob es sich um eine stetige Funktion handelt.

Abbildung 13.2: Der Graph einer stetigen Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ .

iv) Es sei  $U = \mathbb{R}^2$  und  $f: U \rightarrow \mathbb{R}$  sei gegeben durch

$$f(\underline{\mathbf{x}}) = \begin{cases} \frac{x_1^2 x_2^2}{x_1^2 + x_2^2} & \text{für } \underline{\mathbf{x}} \neq \underline{\mathbf{0}}, \\ 0 & \text{für } \underline{\mathbf{x}} = \underline{\mathbf{0}}. \end{cases}$$

Für jede gegen  $\underline{\mathbf{0}}$  konvergente Folge  $\{\underline{\mathbf{x}}^{(k)}\}$  gilt für  $k \rightarrow \infty$

$$\begin{aligned} |f(\underline{\mathbf{x}}^{(k)}) - f(\underline{\mathbf{0}})| &= |f(\underline{\mathbf{x}}^{(k)})| = \left| \frac{(x_1^{(k)})^2 (x_2^{(k)})^2}{(x_1^{(k)})^2 + (x_2^{(k)})^2} \right| \\ &\leq \frac{\|\underline{\mathbf{x}}^{(k)}\|^4}{\|\underline{\mathbf{x}}^{(k)}\|^2} = \|\underline{\mathbf{x}}^{(k)}\|^2 \rightarrow 0, \end{aligned}$$

d.h.  $f$  ist im Nullpunkt stetig (vgl. Abbildung 13.6).

v) Es sei  $U = \mathbb{R}^2$  und  $f: U \rightarrow \mathbb{R}$  sei nun gegeben durch

$$f(\underline{\mathbf{x}}) = \begin{cases} \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2} & \text{für } \underline{\mathbf{x}} \neq \underline{\mathbf{0}}, \\ 0 & \text{für } \underline{\mathbf{x}} = \underline{\mathbf{0}}. \end{cases}$$

Für eine Folge  $\{\underline{\mathbf{x}}^{(k)}\}$  der Form

$$\underline{\mathbf{x}}^{(k)} = \begin{pmatrix} x_1^{(k)} \\ 0 \end{pmatrix}, \quad x_1^{(k)} \xrightarrow{k \rightarrow \infty} 0, \quad x_1^{(k)} \neq 0,$$

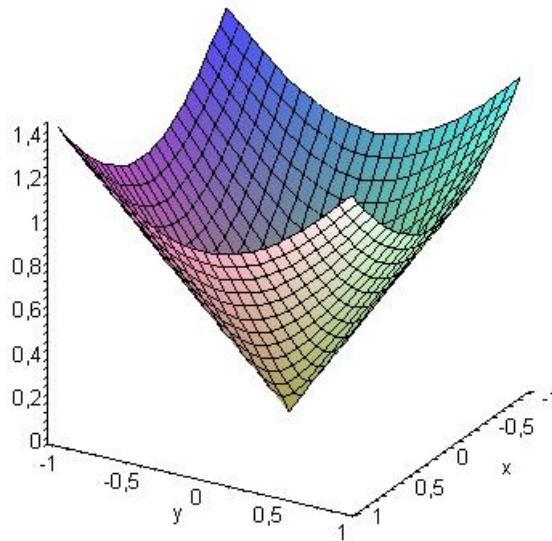


Abbildung 13.3: Der Graph der Betragsfunktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ .

gilt

$$\lim_{k \rightarrow \infty} f(\underline{\mathbf{x}}^{(k)}) = 1 .$$

Für eine Folge  $\tilde{\underline{\mathbf{x}}}^{(k)}$  der Form

$$\tilde{\underline{\mathbf{x}}}^{(k)} = \begin{pmatrix} 0 \\ \tilde{x}_2^{(k)} \end{pmatrix}, \quad \tilde{x}_2^{(k)} \rightarrow 0, \quad \tilde{x}_2^{(k)} \neq 0,$$

gilt hingegen

$$\lim_{k \rightarrow \infty} f(\tilde{\underline{\mathbf{x}}}^{(k)}) = -1 .$$

Die Funktion  $f$  ist an der Stelle  $\underline{\mathbf{x}}^{(0)}$  nicht stetig (vgl. Abbildung 13.7).

**Bemerkung.** Nach Satz 13.1 kann die Stetigkeit einer Funktion  $f: U \rightarrow \mathbb{R}^m$  **komponentenweise** untersucht werden.

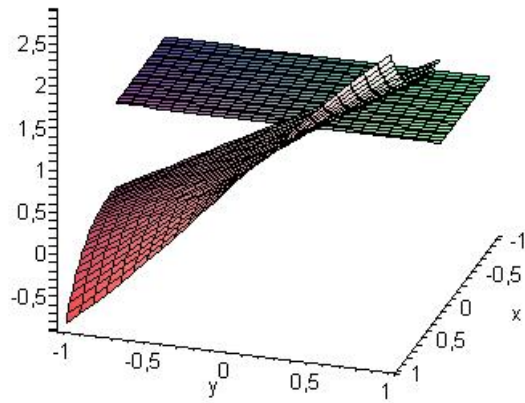


Abbildung 13.4: Der Graph einer unstetigen Funktion, die stetig in einer Richtung ist.

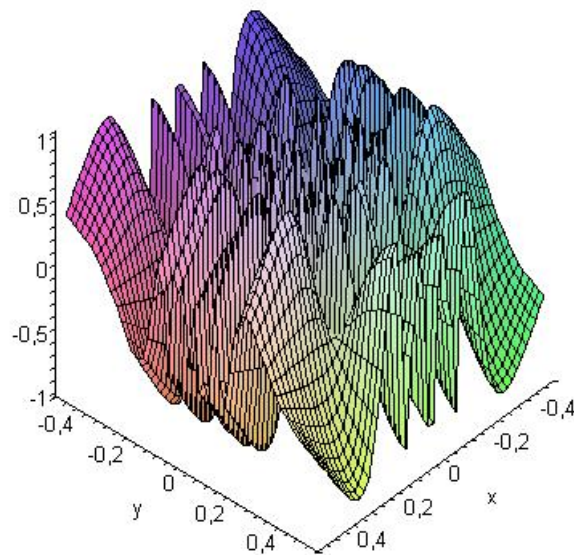
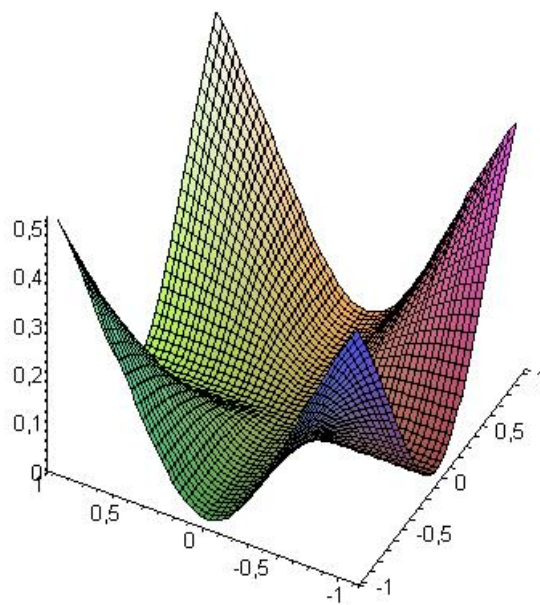
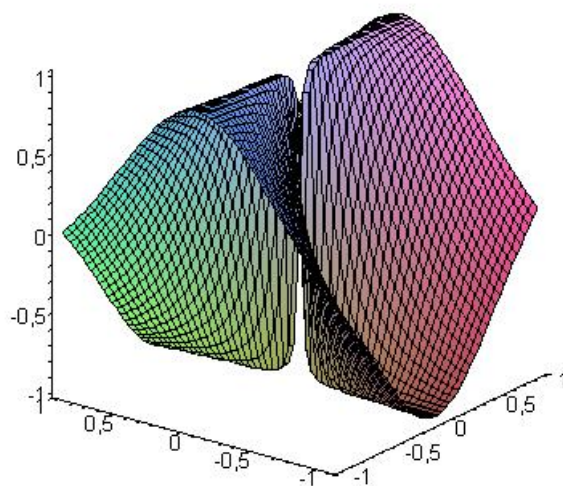


Abbildung 13.5: Diese Funktion scheint nicht stetig zu sein.

Abbildung 13.6: Zum Beispiel  $iv$ ).Abbildung 13.7: Zum Beispiel  $v$ ).

# Kapitel 14

## Differentialrechnung von Funktionen mehrerer Veränderlicher

**Vorbemerkung.** Ist eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  differenzierbar, so ist die **Ableitung nicht notwendigerweise selbst eine stetige Funktion**.

Dies kann man etwa anhand des Beispiels  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,

$$f(x) = \begin{cases} x^2 \sin(1/x) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0, \end{cases}$$

einsehen.

Funktionen deren Ableitungen stetig sind, heißen **stetig differenzierbar**.

Bei Funktionen mehrerer Veränderlicher  $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$  ist dieser **zunächst unscheinbare Aspekt für die Entwicklung der Theorie von herausragender Bedeutung**, da es verschiedene Zugänge zum Ableitungsbegriff gibt, die **nicht äquivalent sind**.

Um diese und andere technische Schwierigkeiten zu vermeiden, gelte:

### **Generalvoraussetzung.**

- i)* Alle im Folgenden auftretenden Ableitungen (insbesondere in Kapitel 14.2) werden, wenn nötig und ohne dass es explizit erwähnt wird, als stetige Funktionen angenommen.

- ii) Wird eine Funktion  $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$  in einem Punkt  $\underline{\mathbf{x}}^{(0)} \in U$  auf Differenzierbarkeit untersucht bzw. dort als differenzierbar vorausgesetzt, so ist damit immer auch gemeint, dass  $\underline{\mathbf{x}}^{(0)}$  ein sogenannter **innerer Punkt** von  $U$  ist. D.h. es existiert ein  $\varepsilon > 0$  mit

$$B_\varepsilon(\underline{\mathbf{x}}^{(0)}) = \{ \underline{\mathbf{x}} \in \mathbb{R}^m : \|\underline{\mathbf{x}} - \underline{\mathbf{x}}^{(0)}\| < \varepsilon \} \subset U$$

(vgl. Übungen).

Mit der Differentialrechnung von Funktionen mehrerer Veränderlicher geht es in diesem Kapitel um den vielleicht wichtigsten Teil der Analysis, da bei der Beschreibung realer Vorgänge in der Regel eindimensionale Betrachtungsweisen zu kurz greifen und neue geometrische und/oder abstrakte Konzepte eingeführt werden müssen.

## 14.1 Kurven im $\mathbb{R}^n$

Im allgemeinen Fall geht es in diesem Kapitel 14 um **vektorwertige Funktionen** ( $n \geq 1$ )  $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$  **mehrer Veränderlicher** ( $m \geq 1$ ).

Zunächst wird der Spezialfall  $m = 1$  von **Kurven im  $\mathbb{R}^n$  angesprochen**, d.h. der Definitionsbereich der betrachteten Funktionen ist eine Teilmenge der reellen Achse – d.h. im Fall von Kurven ein Intervall.

Dies ist einerseits eine wesentliche Vereinfachung, andererseits wird ein für das Folgende unverzichtbarer geometrischer Begriff vorgestellt.

Unter einer **Kurve im  $\mathbb{R}^n$** ,  $n \geq 2$ , versteht man eine **stetige** Abbildung  $\gamma: I \rightarrow \mathbb{R}^n$ ,

$$\gamma(t) = \begin{pmatrix} \gamma_1(t) \\ \vdots \\ \gamma_n(t) \end{pmatrix}, \quad t \in I,$$

wobei  $I \subset \mathbb{R}$  ein (evtl. unendliches) Intervall bezeichnet.



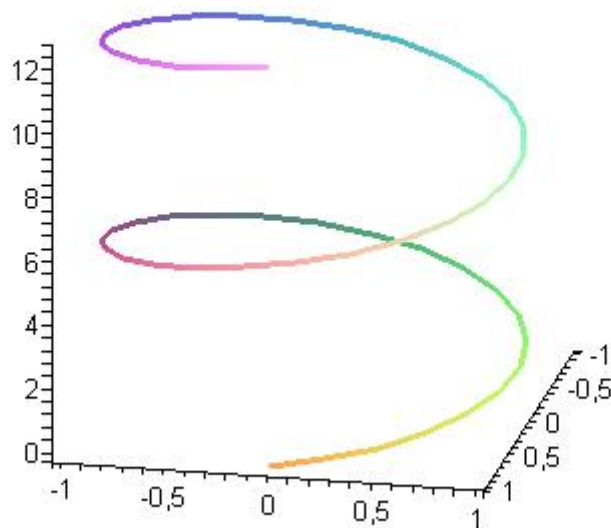


Abbildung 14.1: Eine Helix oder Schraubenlinie.

**Beispiel.** Eine der bekanntesten Kurven im  $\mathbb{R}^3$  ist die sogenannte **Helix** oder **Schraubenlinie**

$$\gamma(t) = \begin{pmatrix} a \cos(t) \\ a \sin(t) \\ bt \end{pmatrix}, \quad t \in \mathbb{R},$$

wobei  $a, b > 0$  fixierte Konstanten bezeichnen.

Die **Spur** der Helix, d.h.

$$\text{bild } \gamma = \{ \underline{\mathbf{x}} \in \mathbb{R}^n : \underline{\mathbf{x}} = \gamma(t) \text{ für ein } t \in I = \mathbb{R} \}$$

ist in Abbildung 14.1 dargestellt.

Zu beachten ist, dass **unterschiedliche Kurven dieselbe Spur** haben können und dass eine Kurve nicht notwendig injektiv sein muss, d.h. es können **Doppelpunkte** ( $\gamma(t_1) = \gamma(t_2)$  für  $t_1 \neq t_2 \in I$ ) existieren.

Zudem fehlt in einer Skizze wie in Abbildung 14.1 jegliche Information darüber, **mit welcher Geschwindigkeit und in welche Richtung die Kurve durchlaufen wird.**

Wie bereits in Abbildung 8.2 angedeutet, liefert der **Geschwindigkeits-** oder **Tangentenvektor** an die Kurve diese fehlenden Informationen:

**Definition 14.1.** *i) Eine Kurve  $\gamma$  heißt differenzierbar, wenn alle Funktionen (Komponenten)  $\gamma_k$ ,  $1 \leq k \leq n$ , differenzierbar (bzw. stetig differenzierbar) sind.*

*Ist  $\gamma$  differenzierbar, so heißt*

$$\gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \vdots \\ \gamma'_m(t) \end{pmatrix}, \quad t \in I,$$

*Tangentenvektor (oder Tangentialvektor) der Kurve  $\gamma$  zum Parameterwert  $t$ .*

*ii) Ist  $\gamma: I \rightarrow \mathbb{R}^n$  differenzierbar, so heißt die Kurve **regulär**, falls  $\gamma'(t) \neq \mathbf{0}$  für alle  $t \in I$ .*

*In diesem Fall  $\gamma'(t) \neq \mathbf{0}$  heißt  $\gamma'(t)/\|\gamma'(t)\|$  **Tangenteneinheitsvektor**.*

*Andernfalls heißt ein Parameterwert  $t \in I$  mit  $\gamma'(t) = \mathbf{0}$  **singulär**.*

## Beispiele.

*i) Die Kurve*

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi),$$

*hat als Spur die Einheitskreislinie im  $\mathbb{R}^2$ .*

*Die gleiche Spur (zweimal durchlaufen) hat die Kurve*

$$\tilde{\gamma}(t) = \begin{pmatrix} \cos(2t) \\ \sin(2t) \end{pmatrix}, \quad t \in [0, 2\pi).$$

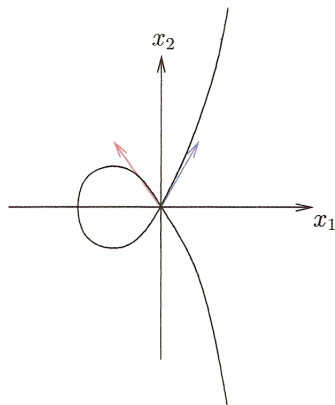


Abbildung 14.2: Eine reguläre Kurve mit Doppelpunkt.

- ii) Der Punkt  $\mathbf{0} = \gamma(1) = \gamma(-1)$  ist ein Doppelpunkt der Kurve (vgl. Abbildung 14.2)

$$\gamma(t) = \begin{pmatrix} t^2 - 1 \\ t^3 - t \end{pmatrix}, \quad t \in \mathbb{R}.$$

Trotzdem ist der Begriff “Tangentialvektoren zu den Parameterwerten  $t = 1$  und  $t = -1$ ” wohldefiniert.

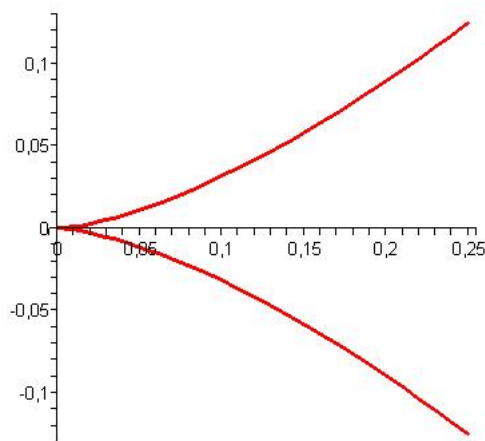


Abbildung 14.3: Die Neilsche Parabel.

- iii) Die **Neilsche Parabel** (vgl. Abbildung 14.3),

$$\gamma(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}, \quad t \in \mathbb{R},$$

ist im Punkt  $t = 0$  singulär.

iv) Übung: Man betrachte die Kurven  $\gamma, \tilde{\gamma}: \mathbb{R} \rightarrow \mathbb{R}^2$ ,

$$\gamma(t) = \begin{pmatrix} t \\ t \end{pmatrix}, \quad \tilde{\gamma}(t) = \begin{pmatrix} t^3 \\ t^3 \end{pmatrix},$$

und skizziere deren Spur. Sind die Kurven regulär?

Fazit: Ist eine Kurve “ungeschickt” parametrisiert, so sieht man ihrer Spur singuläre Punkte nicht an.

### Was ist die Länge einer Kurve?

**Idee.** Man approximiere eine Kurve mit **Polygonzügen**, d.h. Punkte auf der Spur der Kurve werden (affin) linear (mit Strecken) verbunden (vgl. Abbildung 14.4).

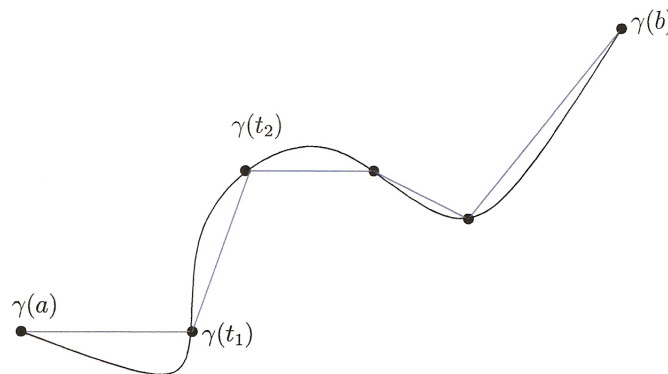


Abbildung 14.4: Zur sogenannten **Rektifizierbarkeit** einer Kurve.

**Satz 14.1.** Die Länge (auch **Bogenlänge** genannt) einer differenzierbaren Kurve  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  berechnet sich zu

$$L = \int_a^b \|\gamma'(t)\| dt .$$

### Beispiele.

- i) Wenn die Bogenlänge sinnvoll definiert ist, muss sie etwa auch bei der Kreislinie vom Radius  $r$  der elementargeometrisch bekannten entsprechen.

Es sei also  $r > 0$  fixiert und

$$\gamma(t) = \begin{pmatrix} r \cos(t) \\ r \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi], \quad \text{d.h.}$$

$$\gamma'(t) = \begin{pmatrix} -r \sin(t) \\ r \cos(t) \end{pmatrix}.$$

Wie erwartet folgt

$$L = \int_0^{2\pi} \sqrt{r^2 \sin^2(t) + r^2 \cos^2(t)} \, dt = \int_0^{2\pi} r \, dt = 2\pi r.$$

ii) Im  $\mathbb{R}^2$  sei ein **Graph** betrachtet, d.h. eine Kurve der Form

$$\gamma(t) = \begin{pmatrix} t \\ g(t) \end{pmatrix}, \quad t \in [a, b].$$

Ist  $\gamma$  stetig differenzierbar, so gilt

$$\gamma'(t) = \begin{pmatrix} 1 \\ g'(t) \end{pmatrix}, \quad \text{also}$$

$$L = \int_a^b \sqrt{1 + (g')^2(t)} \, dt.$$

### Haben Kurven mit gleicher Spur die gleiche Länge?

Ändert man an einer Kurve lediglich die Geschwindigkeit, mit der die Spur durchlaufen wird, so spricht man von einer **Parametertransformation**.

**Beispiel.** Die Kreislinie vom Radius  $r$  werde mit doppelter Geschwindigkeit durchlaufen, d.h. es sei  $r > 0$  fixiert und

$$\gamma(t) = \begin{pmatrix} r \cos(2t) \\ r \sin(2t) \end{pmatrix}, \quad t \in [0, \pi],$$

$$\gamma'(t) = \begin{pmatrix} -r2 \sin(2t) \\ r2 \cos(2t) \end{pmatrix}.$$

Es folgt

$$L = \int_0^\pi \sqrt{r^2 4 \sin^2(t) + r^2 4 \cos^2(t)} dt = \int_0^\pi 2r dt = 2\pi r ,$$

und wie erwartet ist die Länge invariant.

### Parametertransformationen für Kurven.

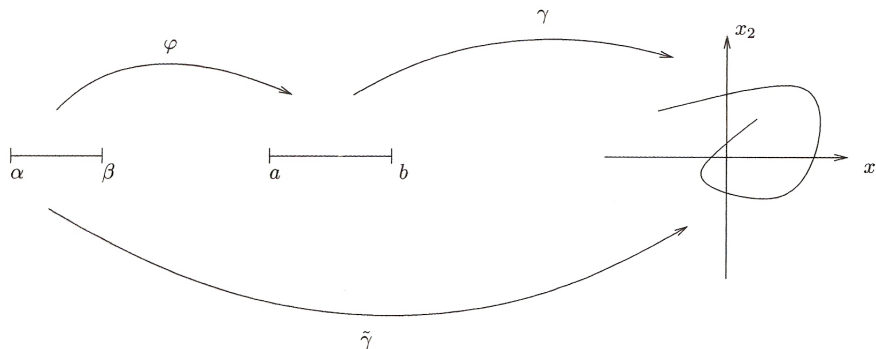


Abbildung 14.5: Eine Parametertransformation.

**Definition 14.2.** Es sei  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  eine Kurve,  $[\alpha, \beta] \subset \mathbb{R}$  ein weiteres Intervall und  $\varphi: [\alpha, \beta] \rightarrow [a, b]$  eine bijektive, stetige Abbildung.

i) Dann ist die zusammengesetzte Abbildung

$$\tilde{\gamma} := \gamma \circ \varphi : [\alpha, \beta] \rightarrow \mathbb{R}^n$$

wieder eine Kurve im  $\mathbb{R}^n$  (mit gleicher Spur und gleicher Länge),  $\tilde{\gamma}$  geht aus  $\gamma$  durch die Parametertransformation  $\varphi$  hervor.

ii)  $\varphi$  heißt *orientierungstreu* (bzw. *orientierungsumkehrend*), falls die Abbildung *monoton wächst* (bzw. *monoton fällt*).

iii) Sind sowohl  $\varphi$  als auch die Umkehrabbildung  $\varphi^{-1}: [a, b] \rightarrow [\alpha, \beta]$  stetig differenzierbar, so heißt  $\varphi$  eine  *$C^1$ -Parametertransformation*.

## 14.2 Ableitungen einer Funktion mehrerer Veränderlicher

Im letzten Paragraphen wurden Kurven im  $\mathbb{R}^n$  diskutiert, d.h. Abbildungen  $f: \mathbb{R} \rightarrow \mathbb{R}^n$ .

Im Gegensatz dazu sei jetzt zunächst der **Bildbereich reell**, also

$$f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}, \quad U \ni \underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto f(\underline{\mathbf{x}}) \in \mathbb{R}.$$

Es ist nicht evident, wie der Ableitungsbegriff auf den Fall mehrerer Veränderlicher verallgemeinert werden kann – hier kann **kein Differenzenquotient** betrachtet werden, da durch eine vektorwertige Größe nicht dividiert werden kann.

### Was erwartet man von einer Ableitung?

Betrachtet sei  $f: \mathbb{R}^m \rightarrow \mathbb{R}$ .

- i) Die Ableitung der Funktion  $f$  soll eine **lineare Abbildung**  $\mathbb{R}^m \rightarrow \mathbb{R}$  sein, die  $f$  in gewissem Sinne approximiert.

Der **Graph** dieser (affin) linearen Abbildung ist im Fall einer Veränderlichen dabei die Tangente in einem Punkt.

Hier muss analog die sogenannte **Tangentialebene** betrachtet werden (vgl. Abbildung 14.6).

- ii) Eine lineare Abbildung  $\mathbb{R}^m \rightarrow \mathbb{R}$  kann nach Kapitel 12.1 mit einer Matrix  $A \in M(1, m)$ , d.h. mit einem **Zeilenvektor** identifiziert werden, d.h. gesucht ist ein geeigneter Zeilenvektor als Repräsentant der Ableitung.

- iii) **Variiert man die unabhängige Variable  $\underline{\mathbf{x}}$  nur in eine Koordinatenrichtung  $\underline{\mathbf{e}}^{(i)}$  und hält die anderen Koordinaten fest**, so ist die Situation analog zu der einer Funktion nur einer Veränderlichen.

Geometrisch betrachtet man eine Kurve auf dem Graphen (vgl. Abbildung 14.7).

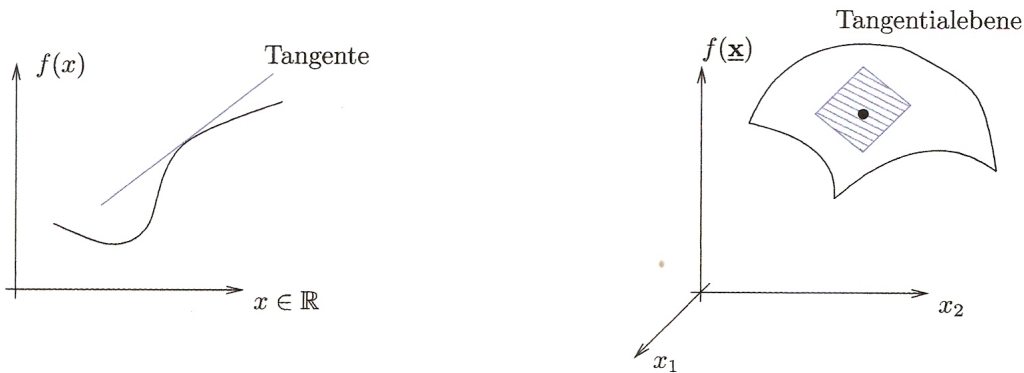


Abbildung 14.6: Approximation mit einer Tangentialebene.

- iv) Nach iii) sollten gewisse Arten eindimensionaler Ableitungen die Eintragungen des Zeilenvektors aus ii) sein.

**Definition 14.3.** Es sei  $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$ .

- i) Es heißt  $f$  im Punkt  $\underline{\mathbf{x}}^{(0)} \in U$  *partiell differenzierbar bzgl. der  $i^{\text{ten}}$  Koordinatenrichtung*, falls der Grenzwert

$$\begin{aligned} \frac{\partial f}{\partial x_i}(\underline{\mathbf{x}}^{(0)}) &:= D_i f(\underline{\mathbf{x}}^{(0)}) \\ &:= \lim_{h \rightarrow 0, h \neq 0} \frac{f(\underline{\mathbf{x}}^{(0)} + h \underline{\mathbf{e}}^{(i)}) - f(\underline{\mathbf{x}}^{(0)})}{h} \end{aligned}$$

existiert, wobei  $\underline{\mathbf{e}}^{(i)}$ ,  $1 \leq i \leq m$ , den  $i^{\text{ten}}$  Einheitsvektor der Standardbasis des  $\mathbb{R}^m$  bezeichne.

$D_i f(\underline{\mathbf{x}}^{(0)})$  heißt die  $i^{\text{te}}$  *partielle Ableitung* von  $f$  im Punkt  $\underline{\mathbf{x}}^{(0)}$ .

- ii) Existieren alle partiellen Ableitungen von  $f$  in allen Punkten  $\underline{\mathbf{x}}^{(0)} \in U$  (und sind als Funktion auf  $U$  stetig, siehe Generalvoraussetzung), so heißt

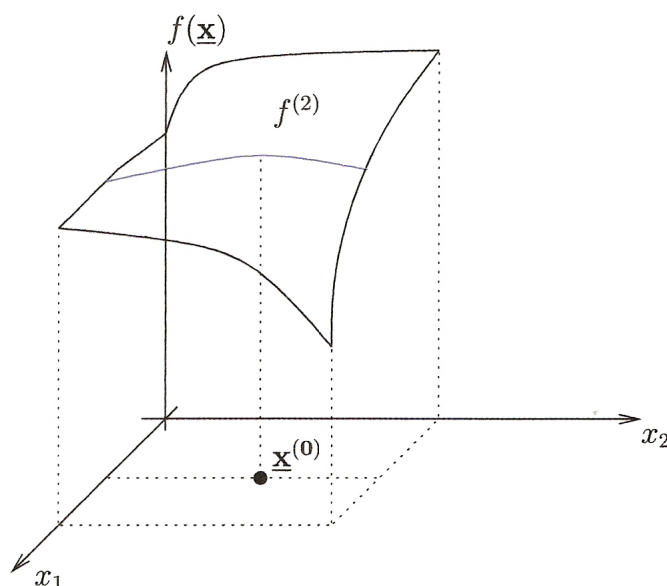
$$Df(\underline{\mathbf{x}}) := \text{grad} f(\underline{\mathbf{x}}) := \nabla f(\underline{\mathbf{x}}) := \left( \frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_m} \right) (\underline{\mathbf{x}})$$

die *Ableitung* oder der *Gradient* von  $f$  im Punkt  $\underline{\mathbf{x}} \in U$ .

In diesem Fall heißt  $f$  *einmal stetig differenzierbar*.

Notation:  $f \in C^1(U)$ .



Abbildung 14.7: Es wird nur in  $\underline{e}^{(1)}$ -Richtung variiert..**Bemerkungen.**

- i) An dieser Stelle sei an die Bedeutung der Generalvoraussetzung dieses Kapitels erinnert:

Ein konsistentes Konzept der Differenzierbarkeit sollte wie bei der Differentialrechnung in einer Variablen die Stetigkeit einer differenzierbaren Funktion beinhalten.

Allein aus der Existenz der partiellen Ableitungen folgt dies aber nicht, wie das folgende Beispiel zeigt:

Es sei  $f: \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $m \geq 2$ , definiert durch

$$f(\underline{x}) = \begin{cases} \frac{x_1 x_2 \dots x_m}{(x_1^2 + x_2^2 + \dots + x_m^2)^m} & \text{für } \underline{x} \neq \underline{0}, \\ 0 & \text{für } \underline{x} = \underline{0}. \end{cases}$$

Dann ist  $f$  zwar partiell differenzierbar,  $f$  ist im Nullpunkt aber nicht stetig.

Betrachtet man hingegen eine Funktion, deren partielle Ableitungen stetig sind (wie hier stets angenommen), so ist die Funktion selbst ebenfalls stetig.

ii) Bei der partiellen Ableitung einer Funktion wird also für ein fixiertes  $\underline{\mathbf{x}}^{(0)} \in U$  und für ein fixiertes  $1 \leq i \leq m$  die folgende Funktion einer Veränderlichen betrachtet:

$$t \mapsto f^{(i)}(t) := f \begin{pmatrix} x_1^{(0)} \\ \vdots \\ x_{i-1}^{(0)} \\ t \\ x_{i+1}^{(0)} \\ \vdots \\ x_m^{(0)} \end{pmatrix} .$$

Mit dieser Notation ist

$$D_i f(\underline{\mathbf{x}}^{(0)}) = \lim_{h \rightarrow 0, h \neq 0} \frac{f^{(i)}(x_i^{(0)} + h) - f^{(i)}(x_i^{(0)})}{h} = f^{(i)'}(x_i^{(0)}) .$$

Mit anderen Worten:

Man leite nach der  $i^{\text{ten}}$  Variablen ab, die restlichen  $n - 1$  Variablen sind fixiert (eingefroren) (vgl. Abbildung 14.7).

Deshalb übertragen sich die Rechenregeln für “d/dx” auf die Operatoren  $D_i$ .

iii) Völlig analog zu ii) kann auch die Abbildung  $\mathbb{R} \supset I \rightarrow \mathbb{R}$ ,

$$s \mapsto f(\underline{\mathbf{x}}^{(0)} + s\underline{\mathbf{e}}^{(i)}) ,$$

für fixiertes  $i = 1, \dots, m$  betrachtet werden. Dann gilt

$$D_i f(\underline{\mathbf{x}}^{(0)}) := \frac{d}{ds} f(\underline{\mathbf{x}}^{(0)} + s\underline{\mathbf{e}}^{(i)})|_{s=0} .$$

**Beispiel.** Man betrachte die Funktion

$$f(\underline{\mathbf{x}}) = \left( \sum_{i=1}^m x_i^2 \right)^{\frac{1}{2}} = \|\underline{\mathbf{x}}\| , \quad U = \mathbb{R}^m - \{\underline{\mathbf{0}}\} .$$

Ist  $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$  fixiert, so ist für alle  $i = 1, \dots, m$

$$t \mapsto \sqrt{x_1^2 + \dots + t^2 + \dots + x_m^2}$$

$\uparrow i$

als Funktion einer Variablen differenzierbar, also ist  $f$  partiell differenzierbar und es gilt

$$D_i f(\underline{\mathbf{x}}) = \frac{t}{\sqrt{x_1^2 + \dots + t^2 + \dots + x_m^2}} \Big|_{t=x_i} = \frac{x_i}{f(\underline{\mathbf{x}})} = \frac{x_i}{\|\underline{\mathbf{x}}\|}.$$

Die Funktion ist auf  $U$  differenzierbar, wobei wie im Fall der Betragsfunktion einer Veränderlichen der Nullpunkt nicht in der Menge  $U$  liegen kann, in der die Funktion differenzierbar ist.

### Richtungsableitungen und erste Interpretation des Gradienten.

Es gibt natürlich überhaupt keinen Grund dafür, nur Kurven auf dem Graphen in Richtung der Koordinatenachsen zu studieren.

Seien also  $\underline{\mathbf{x}}^{(0)} \in U$  und irgendeine Richtung  $\underline{\mathbf{v}} \in \mathbb{R}^m$  gegeben.

Die **Richtungsableitung** von  $f$  im Punkt  $\underline{\mathbf{x}}^{(0)}$  in Richtung  $\underline{\mathbf{v}}$  ist analog zu obiger Bemerkung (vgl. Abbildung 14.8)

$$D_{\underline{\mathbf{v}}} f(\underline{\mathbf{x}}^{(0)}) := \frac{d}{dt} f(\underline{\mathbf{x}}^{(0)} + t\underline{\mathbf{v}}) \Big|_{t=0} = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(\underline{\mathbf{x}}^{(0)}) v_i = Df(\underline{\mathbf{x}}^{(0)}) \underline{\mathbf{v}}.$$

Betrachtet man nun in dem fixierten Punkt  $\underline{\mathbf{x}}^{(0)}$  die Richtungsableitungen in alle möglichen Richtungen  $\underline{\mathbf{v}}$ , also den Ausdruck

$$\nabla f(\underline{\mathbf{x}}^{(0)}) \underline{\mathbf{v}}, \quad \|\underline{\mathbf{v}}\| = 1,$$

so erkennt man, dass dieser Ausdruck maximal wird, falls  $\underline{\mathbf{v}}$  in Richtung von  $\nabla f(\underline{\mathbf{x}}^{(0)})$  (eigentlich des transponierten Vektors) zeigt:

Als Spaltenvektor im  $\mathbb{R}^m$  zeigt  $\nabla f(\underline{\mathbf{x}}^{(0)})$  in die Richtung des stärksten Anstiegs von  $f$  im Punkt  $\underline{\mathbf{x}}^{(0)}$ .

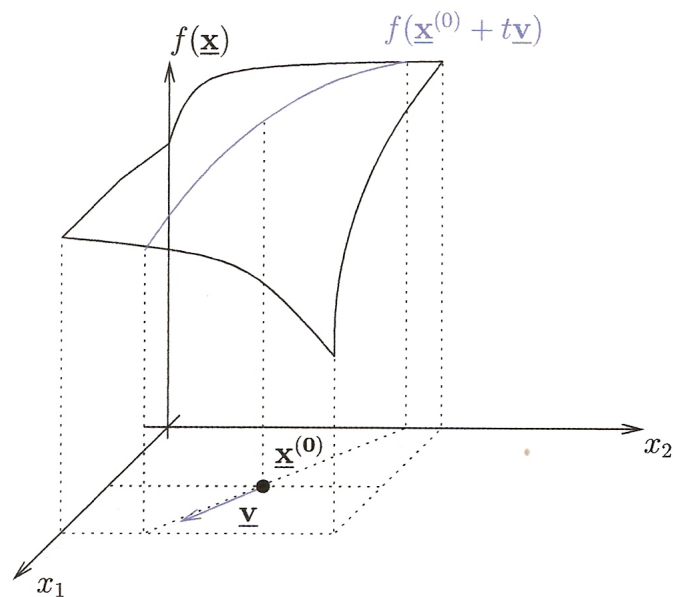


Abbildung 14.8: Zur Richtungsableitung.

Zu einer ähnlichen Interpretation gelangt man, wenn die Funktion  $f$  mit Hilfe von [Höhenlinien](#) diskutiert wird ( $\rightsquigarrow$  Übungen).

### Tangente und Tangentialebene.

**Erinnerung.** (Vgl. Abbildung 14.6) Im Fall  $m = 1$  ist der Graph von  $f$  die Spur der Kurve

$$\gamma(t) = \begin{pmatrix} t \\ f(t) \end{pmatrix},$$

im  $\mathbb{R}^2$ . Der **Richtungsvektor der Tangente** im Punkt  $t$  ist gegeben durch

$$\gamma'(t) = \begin{pmatrix} 1 \\ f'(t) \end{pmatrix}.$$

Sei nun beispielsweise  $m = 2$ ,  $f: \mathbb{R}^2 \supset U \rightarrow \mathbb{R}$  und  $\underline{x}^{(0)} \in U$  gegeben.

Dann betrachtet man die [Koordinatenkurven](#) durch den Punkt  $\underline{x}^{(0)}$  auf

dem Graphen von  $f$ ,

$$s \mapsto \begin{pmatrix} x_1^{(0)} + s \\ x_2^{(0)} \\ f(\underline{\mathbf{x}}^{(0)} + s\underline{\mathbf{e}}^{(1)}) \end{pmatrix} \quad \text{bzw.} \quad s \mapsto \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} + s \\ f(\underline{\mathbf{x}}^{(0)} + s\underline{\mathbf{e}}^{(2)}) \end{pmatrix} .$$

Die Tangentenvektoren im Punkt  $\underline{\mathbf{x}}^{(0)}$  sind in Verallgemeinerung des eindimensionalen Falls durch

$$\underline{\mathbf{u}} = \begin{pmatrix} 1 \\ 0 \\ D_1 f(\underline{\mathbf{x}}^{(0)}) \end{pmatrix} \quad \text{und} \quad \underline{\mathbf{v}} = \begin{pmatrix} 0 \\ 1 \\ D_2 f(\underline{\mathbf{x}}^{(0)}) \end{pmatrix}$$

gegeben.

Diese Vektoren sind stets linear unabhängig (warum?) und **spannen die Tangentialebene  $T_{\underline{\mathbf{x}}^{(0)}}$  an den Graphen von  $f$  im Punkt  $\underline{\mathbf{x}}^{(0)}$  auf.**

**Beispiel.** Man betrachte die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$f(\underline{\mathbf{x}}) = c + x_1^2 + x_2^2 ,$$

wobei  $c > 0$  eine positive Konstante bezeichne.

Wie in Abbildung 14.9 skizziert, handelt es sich hier um ein **Paraboloid**.

Mit obiger Notation ist

$$\underline{\mathbf{u}} = \begin{pmatrix} 1 \\ 0 \\ 2x_1 \end{pmatrix} , \quad \underline{\mathbf{v}} = \begin{pmatrix} 0 \\ 1 \\ 2x_2 \end{pmatrix} ,$$

im Punkt  $\underline{\mathbf{0}}$  wird die Tangentialebene  $T_{\underline{\mathbf{0}}}$  aufgespannt von den Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} ,$$

$T_{\underline{\mathbf{0}}}$  ist die  $\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}$ -Ebene im  $\mathbb{R}^3$ .

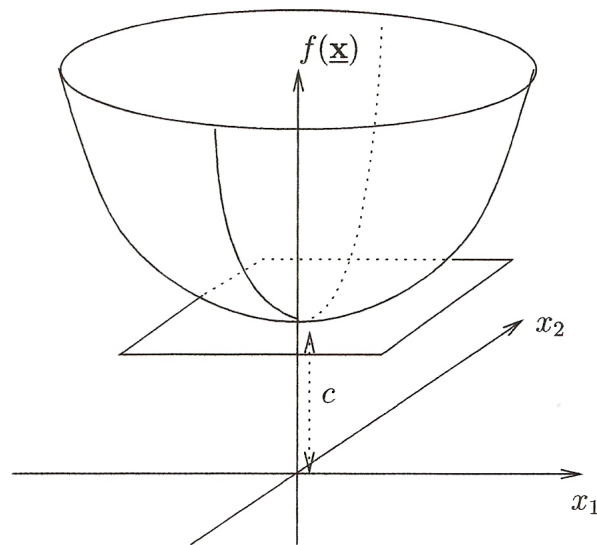


Abbildung 14.9: Ein Paraboloid.

### Vektorwertige Funktionen mehrerer Veränderlicher.

**Definition 14.4.** Es sei  $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$  von der Klasse  $C^1(U; \mathbb{R}^n)$ , d.h. *alle* partiellen Ableitungen *aller* Komponentenfunktionen existieren in *jedem* Punkt aus  $U$  und seien *stetige* Funktionen auf  $U$ .

Dann heißt die Matrix  $A = A(\underline{\mathbf{x}}^{(0)})$ , bestehend aus allen partiellen Ableitungen in einem Punkt  $\underline{\mathbf{x}}^{(0)} \in U$ ,

$$A := Df(\underline{\mathbf{x}}^{(0)}) := J_f(\underline{\mathbf{x}}^{(0)}) := \left( \frac{\partial f_i}{\partial x_j} \right)_{\substack{j=1, \dots, m \\ i=1, \dots, n}} \in M(n, m)$$

das *Differential* oder die *Jacobi-Matrix* oder die *Funktionalmatrix* von  $f$  in  $\underline{\mathbf{x}}^{(0)}$ .

### Beispiele.

i) Im Fall  $n = 1$  ist  $A(\underline{\mathbf{x}}^{(0)})$  der Gradient von  $f$  im Punkt  $\underline{\mathbf{x}}^{(0)}$ .

ii) Es sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ ,

$$f(\underline{\mathbf{x}}) = \begin{pmatrix} x_1^2 + x_2 \\ x_1 x_2 \\ e^{x_1} \end{pmatrix}.$$

Dann ist  $f$  von der Klasse  $C^1(\mathbb{R}^2; \mathbb{R}^3)$  und für jedes  $\underline{x} \in \mathbb{R}^2$  ist

$$A = A(\underline{x}) = \begin{pmatrix} 2x_1 & 1 \\ x_2 & x_1 \\ e^{x_1} & 0 \end{pmatrix}.$$

### Gibt es eine Verallgemeinerung der Kettenregel?

**Satz 14.2.** *Es seien  $U \subset \mathbb{R}^m$ ,  $V \subset \mathbb{R}^k$ ,  $g: U \rightarrow \mathbb{R}^k$ ,  $f: V \rightarrow \mathbb{R}^n$ , es gelte  $g(U) \subset V$  und die Abbildungen  $f, g$  seien von der Klasse  $C^1$ . Dann ist auch die Abbildung  $f \circ g: \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$  von der Klasse  $C^1$  und in jedem Punkt  $\underline{x}^{(0)} \in U$  gilt*

$$D(f \circ g)(\underline{x}^{(0)}) = Df(g(\underline{x}^{(0)}))Dg(\underline{x}^{(0)}).$$

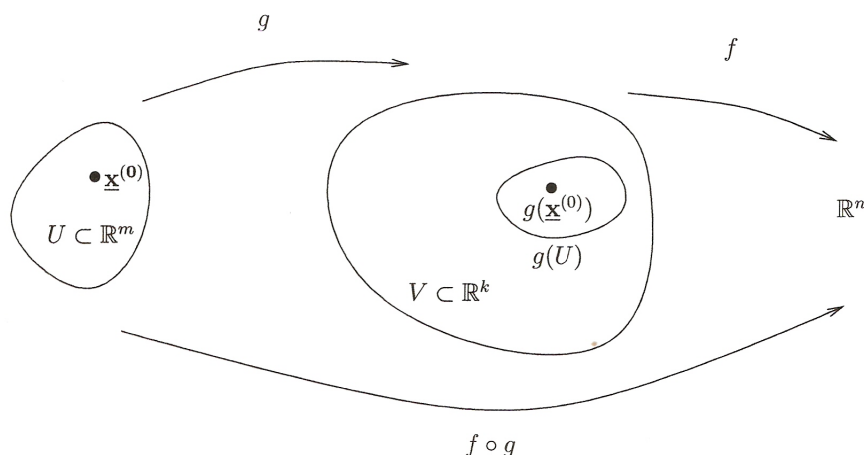


Abbildung 14.10: Zur Kettenregel.

### Beispiele.

i) Im Spezialfall  $f$  reellwertig ( $n = 1$ ) gilt ( $h := f \circ g$ )

$$\begin{aligned} Dh(\underline{x}^{(0)}) &= Df(g(\underline{x}^{(0)}))Dg(\underline{x}^{(0)}) \\ &= \left( \frac{\partial f}{\partial y_1}(g(\underline{x}^{(0)})) \dots \frac{\partial f}{\partial y_k}(g(\underline{x}^{(0)})) \right) \\ &\quad \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\underline{x}^{(0)}) & \dots & \frac{\partial g_1}{\partial x_m}(\underline{x}^{(0)}) \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_1}(\underline{x}^{(0)}) & \dots & \frac{\partial g_k}{\partial x_m}(\underline{x}^{(0)}) \end{pmatrix}. \end{aligned}$$

Insbesondere ist ( $i = 1, \dots, m$ )

$$\frac{\partial h}{\partial x_i}(\underline{\mathbf{x}}^{(0)}) = \sum_{j=1}^k \frac{\partial f}{\partial y_j}(g(\underline{\mathbf{x}}^{(0)})) \frac{\partial g_j}{\partial x_i}(\underline{\mathbf{x}}^{(0)}) .$$

ii) Es sei (**Polarkoordinaten**)

$$U = \left\{ \begin{pmatrix} r \\ \varphi \end{pmatrix} : 0 < r, 0 < \varphi < 2\pi \right\} ,$$

$$g : \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \underline{\mathbf{x}} = \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix} , \quad f : \mathbb{R}^2 \rightarrow \mathbb{R} .$$

Dann gilt für  $h = f \circ g$ :

$$\frac{\partial h}{\partial r} = \frac{\partial f}{\partial x_1} \cos(\varphi) + \frac{\partial f}{\partial x_2} \sin(\varphi) , \quad (1)$$

$$\frac{\partial h}{\partial \varphi} = \frac{\partial f}{\partial x_1} (-r \sin(\varphi)) + \frac{\partial f}{\partial x_2} r \cos(\varphi) . \quad (2)$$

Multipliziert man die Gleichung (1) mit  $r \sin(\varphi)$ , Gleichung (2) mit  $\cos(\varphi)$  und addiert sie anschließend, so folgt

$$\frac{\partial}{\partial x_2} f = \sin(\varphi) \frac{\partial}{\partial r} h + \frac{\cos(\varphi)}{r} \frac{\partial}{\partial \varphi} h . \quad (3)$$

Wird die Gleichung (3) wiederum in (1) eingesetzt, so folgt

$$\frac{\partial}{\partial x_1} f = \cos(\varphi) \frac{\partial}{\partial r} h - \frac{\sin(\varphi)}{r} \frac{\partial}{\partial \varphi} h . \quad (4)$$

Die Beziehungen (3) und (4) werden **symbolisch** geschrieben als

$$\frac{\partial}{\partial x_1} = \cos(\varphi) \frac{\partial}{\partial r} - \frac{\sin(\varphi)}{r} \frac{\partial}{\partial \varphi} ,$$

$$\frac{\partial}{\partial x_2} = \sin(\varphi) \frac{\partial}{\partial r} + \frac{\cos(\varphi)}{r} \frac{\partial}{\partial \varphi} ,$$

es handelt sich um den **Nabla-Operator in Polarkoordinaten**. Analoge Ausdrücke können auch für andere **krümmelige Koordinaten** hergeleitet werden.



## Spezielle Differentialoperatoren.

In den Anwendungen spielen neben dem Gradienten weitere spezielle Differentialoperatoren eine herausragende Rolle:

**Definition 14.5.** Ein *Vektorfeld* auf  $U$  ist eine Abbildung

$$F : \mathbb{R}^m \supset U \rightarrow \mathbb{R}^m ,$$

also eine Abbildung, die jedem  $\underline{\mathbf{x}} \in U$  wieder einen Vektor  $\underline{\mathbf{F}}(\underline{\mathbf{x}}) \in \mathbb{R}^m$  zuordnet. (Beispiel:  $\underline{\mathbf{x}} \mapsto \nabla f(\underline{\mathbf{x}})$  als Spaltenvektor interpretiert).

i) Es heißt

$$\operatorname{div} F(\underline{\mathbf{x}}) := \langle \nabla^T, F \rangle(\underline{\mathbf{x}}) := \sum_{i=1}^m \frac{\partial F_i}{\partial x_i}(\underline{\mathbf{x}}) \in \mathbb{R}$$

die *Divergenz* des Vektorfeldes  $F$  im Punkt  $\underline{\mathbf{x}} \in U$ .

ii) Ist speziell  $m = 3$  und, so heißt

$$\operatorname{rot} F(\underline{\mathbf{x}}) := (\nabla^T \times F)(\underline{\mathbf{x}}) := \begin{pmatrix} \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \\ \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{pmatrix}(\underline{\mathbf{x}}) \in \mathbb{R}^3$$

die *Rotation* des Vektorfeldes  $v$  im Punkt  $\underline{\mathbf{x}} \in U$ .

iii) Man setzt

$$\Delta f(\underline{\mathbf{x}}) := \operatorname{div} \operatorname{grad} f(\underline{\mathbf{x}}) = \sum_{i=1}^m \frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_i} \right)(\underline{\mathbf{x}}) ,$$

wobei  $\Delta$  *Laplace-Operator* heißt.

## Bemerkungen.

i) Die Divergenz ist ein Maß für “**Quellen und Senken**” eines Vektorfeldes.

In der Fluid-Mechanik bedeutet “div Geschwindigkeitsfeld = 0” die **Massenerhaltung**.

- ii) Die Rotation gibt Informationen über mögliche “Wirbel” eines Vektorfeldes.
- iii) Die partielle Differentialgleichung  $\Delta f = 0$  heißt die Potentialgleichung, die Lösungen sind sogenannte harmonische Funktionen.
- iv) Für Beziehungen wie  $\text{rot grad} = \mathbf{0}$  etc. sei auf die Übungen verwiesen.

### Beispiele.

i) Es sei  $F(\underline{\mathbf{x}}) = \underline{\mathbf{x}}$ :

$$\text{div } F(\underline{\mathbf{x}}) = \sum_{i=1}^m \frac{\partial F_i}{\partial x_i} = \sum_{i=1}^m \frac{\partial x_i}{\partial x_i} = m .$$

ii) Es sei  $f(\underline{\mathbf{x}}) = \left( \sum_{i=1}^m x_i^2 \right)^{1/2}$ ,  $U = \mathbb{R}^m - \mathbf{0}$ . Dann gilt ( $\nabla f$  als Spaltenvektor interpretiert)

$$\nabla f(\underline{\mathbf{x}}) = \frac{\underline{\mathbf{x}}}{f(\underline{\mathbf{x}})} \in \mathbb{R}^m ;$$

$$\begin{aligned} \Delta f(\underline{\mathbf{x}}) &= \text{div grad } f(\underline{\mathbf{x}}) = \sum_{i=1}^m \frac{\partial}{\partial x_i} \frac{x_i}{f(\underline{\mathbf{x}})} \\ &= \sum_{i=1}^m \left[ \frac{1}{f(\underline{\mathbf{x}})} - \frac{x_i}{f^2(\underline{\mathbf{x}})} \frac{x_i}{f(\underline{\mathbf{x}})} \right] \\ &= \left[ \sum_{i=1}^m \frac{1}{f(\underline{\mathbf{x}})} \right] - \frac{1}{f(\underline{\mathbf{x}})} \\ &= \frac{m-1}{f(\underline{\mathbf{x}})} . \end{aligned}$$

### Höhere Ableitungen.

Oben sind bereits partielle Ableitungen  $D_i f$  selbst als Abbildungen  $U \rightarrow \mathbb{R}^n$  aufgefasst worden.

So können auch **höhere Ableitungen** eingeführt werden, wobei hier der Einfachheit halber der skalare Fall  $n = 1$  betrachtet wird:

**Definition 14.6.** Es sei  $f: U \rightarrow \mathbb{R}$  von der Klasse  $C^1(U)$  und die **partiellen Ableitungen**  $D_i f: U \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$ , **seien selbst wieder partiell differenzierbar** (und nach der Generalvoraussetzung stetige Funktionen auf  $U$ ).

Dann heißt  $f$  **zweimal (stetig) differenzierbar mit partiellen Ableitungen zweiter Ordnung**  $D_j D_i f$ ,  $i = 1, \dots, m$ ,  $j = 1, \dots, m$ .

Notation:  $f \in C^2(U)$ .

**Bemerkung.** Induktiv werden **Ableitungen  $k^{\text{ter}}$  Ordnung**,  $k > 2$ , definiert.

**Beispiel.** (**Laplace-Operator in Polarkoordinaten**) Die zweimalige Anwendung von (3) und (4) liefert in Polarkoordinaten (wieder in der symbolischen Schreibweise) ( $\rightsquigarrow$  Übungen)

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{1}{r} \frac{\partial}{\partial r} .$$

**Wie hängt  $D_j D_i f$  mit  $D_i D_j f$  zusammen?**

**Idee.** Klarerweise sind die **Ableitungen vertauschbar**, falls

$$f(\underline{\mathbf{x}}) = x_1^a x_2^b, \quad a, b \in \mathbb{R}, \quad U = \mathbb{R}^2 .$$

**Satz 14.3.** Ist  $f: U \rightarrow \mathbb{R}$  von der Klasse  $C^2(U)$ , so gilt für alle  $\underline{\mathbf{x}} \in U$ ,  $i, j = 1, \dots, n$

$$\frac{\partial^2 f}{\partial x_j \partial x_i} := D_j D_i f(\underline{\mathbf{x}}) = D_i D_j f(\underline{\mathbf{x}}) .$$

**Vorsicht.** Ohne die Generalvoraussetzung der Stetigkeit aller Ableitungen ist Satz 14.3 **falsch**.

Beispiel (Übung!):

$$f(\underline{\mathbf{x}}) = \begin{cases} x_1 x_2 \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2} & \text{für } \underline{\mathbf{x}} \neq \underline{\mathbf{0}}, \\ 0 & \text{für } \underline{\mathbf{x}} = \underline{\mathbf{0}}. \end{cases}$$

### 14.3 Extremwertaufgaben in mehreren Veränderlichen

Da bei Extremwertaufgaben Funktionswerte verglichen werden und im  $\mathbb{R}^n$  für  $n > 1$  keine Relation “<” definiert ist, sind die folgenden Betrachtungen nur im skalaren Fall  $n = 1$  sinnvoll und können nicht auf die vektorwertige Situation verallgemeinert werden.

Wie im Fall einer Variablen wird zunächst eine **notwendige Bedingung** studiert (vgl. Satz 8.5).

Als Vorbereitung ist das Analogon zu Definition 8.3 festzuhalten:

**Definition 14.7.** Eine Funktion  $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$  hat im Punkt  $\underline{\mathbf{x}}^{(0)} \in U$  ein *lokales Maximum (lokales Minimum)*, falls ein  $r > 0$  existiert mit

$$f(\underline{\mathbf{x}}^{(0)}) \geq f(\underline{\mathbf{x}}) \quad (\text{lok. Max.}) \quad \text{bzw.} \quad f(\underline{\mathbf{x}}^{(0)}) \leq f(\underline{\mathbf{x}}) \quad (\text{lok. Min.})$$

für alle  $\underline{\mathbf{x}} \in \{\underline{\mathbf{y}} \in U : \|\underline{\mathbf{y}} - \underline{\mathbf{x}}^{(0)}\| \leq r\}$ .

Gilt jeweils die **strikte Ungleichung**, so spricht man von einem **strengen lokalen Maximum (Minimum)**.

**Bemerkungen.** Wie im Fall  $m = 1$  gilt:

- i) Lokale Maxima und lokale Minima heißen **lokale Extrema**.
- ii) Ein lokales Extremum ist von einem globalen (oder absoluten) Extremum zu unterscheiden.

**Satz 14.4.** Die Funktion  $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$  habe im Punkt  $\underline{\mathbf{x}}^{(0)} \in U$  ein **lokales Extremum** und sei dort differenzierbar.

Dann gilt

$$\nabla f(\underline{\mathbf{x}}^{(0)}) = \underline{\mathbf{0}}.$$

**Beispiel.** Betrachtet sei das Paraboloid aus Abbildung 14.9, d.h.  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(\underline{\mathbf{x}}) = c + x_1^2 + x_2^2$  für alle  $\underline{\mathbf{x}} \in \mathbb{R}^2$  und für eine Konstante  $c \in \mathbb{R}$ . Es ist

$$\nabla f(\underline{\mathbf{x}}) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

und einziger **Kandidat** für ein lokales Extremum ist der Punkt  $\underline{\mathbf{x}}^{(0)} = \underline{\mathbf{0}}$ .

### Bemerkungen.

i) Beim analogen Satz 8.5 wurden Punkte  $x_0$  aus einem offenen Intervall betrachtet.

Als Gegenstück greift hier der zweite Teil der Generalvoraussetzung dieses Kapitels, d.h. es werden nur innere Punkte von  $U$  betrachtet.

ii) Die Bedingung  $\nabla f(\underline{\mathbf{x}}^{(0)}) = \underline{\mathbf{0}}$  ist lediglich **notwendig, aber nicht hinreichend** dafür, dass in  $\underline{\mathbf{x}}^{(0)}$  ein lokales Extremum vorliegt (völlig analog zum Fall  $m = 1$ ).

In Abbildung 14.11 ist eine Funktion skizziert, die im Nullpunkt lediglich einen **Sattelpunkt (kritischen Punkt)** mit verschwindender Ableitung hat.

### Hinreichende Bedingungen für lokale Extrema?

Satz 8.8 hält hinreichende Bedingungen für die Existenz lokaler Extrema im Fall  $m = 1$  fest. Dort wird mit Hilfe des Vorzeichens der zweiten Ableitung argumentiert.

Im Fall mehrerer Veränderlicher ist zunächst nicht klar, wie eine Vorzeichenbedingung zu formulieren ist, um eine Verallgemeinerung des Satzes zu erhalten. Man setzt:

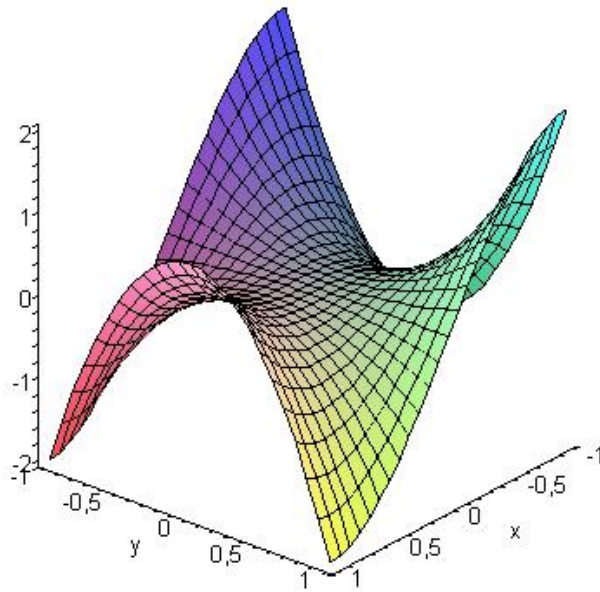


Abbildung 14.11: Ein Sattelpunkt.

**Definition 14.8.** Es sei  $A \in M(m, m)$  eine *symmetrische Matrix*.

i)  $A$  heißt *positiv definit*, falls

$$\langle \underline{\mathbf{v}}, A\underline{\mathbf{v}} \rangle > 0 \quad \text{für alle } \underline{\mathbf{v}} \in \mathbb{R}^m - \{\mathbf{0}\}.$$

ii)  $A$  heißt *positiv semidefinit*, falls

$$\langle \underline{\mathbf{v}}, A\underline{\mathbf{v}} \rangle \geq 0 \quad \text{für alle } \underline{\mathbf{v}} \in \mathbb{R}^m.$$

iii)  $A$  heißt *negativ definit* (bzw. *negativ semidefinit*), falls die Matrix  $-A$  positiv definit (bzw. positiv semidefinit) ist.

iv)  $A$  heißt *indefinit*, falls  $\underline{\mathbf{v}}, \underline{\mathbf{w}} \in \mathbb{R}^m$  existieren mit

$$\langle \underline{\mathbf{v}}, A\underline{\mathbf{v}} \rangle > 0, \quad \langle \underline{\mathbf{w}}, A\underline{\mathbf{w}} \rangle < 0.$$

**Beispiele.**

i) Die Matrix  $A = I_m \in M(m, m)$  ist positiv definit, da für alle  $\underline{\mathbf{v}} \in \mathbb{R}^m - \{\mathbf{0}\}$  gilt

$$\langle \underline{\mathbf{v}}, I_m \underline{\mathbf{v}} \rangle = \|\underline{\mathbf{v}}\|^2 > 0.$$

ii) Es sei

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in M(2, 2) .$$

Sind  $\underline{\mathbf{e}}^{(1)}$ ,  $\underline{\mathbf{e}}^{(2)}$  die kanonischen Basisvektoren des  $\mathbb{R}^2$ , so gilt

$$\langle \underline{\mathbf{e}}^{(1)}, A\underline{\mathbf{e}}^{(1)} \rangle = 1 , \quad \langle \underline{\mathbf{e}}^{(2)}, A\underline{\mathbf{e}}^{(2)} \rangle = -1 ,$$

die Matrix  $A$  ist indefinit.

**Bemerkung.** Der Zusammenhang mit dem Vorzeichen der Eigenwerte wird in den Übungen diskutiert.

Zur weiteren Behandlung von Extremwertaufgaben ist die folgende Matrix der zweiten Ableitungen zu bilden:

**Definition 14.9.** Es sei  $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$  zweimal stetig differenzierbar.

Die *Hessesche Matrix* von  $f$  im Punkt  $\underline{\mathbf{x}}^{(0)} \in U$  ist die *symmetrische*  $m \times m$ -Matrix

$$(\text{Hess } f)(\underline{\mathbf{x}}^{(0)}) := (D_j D_i f(\underline{\mathbf{x}}^{(0)}))_{\substack{1 \leq j \leq m \\ 1 \leq i \leq m}} \in M(m, m) .$$

Mit den Definitionen 14.8 und 14.9 kann schließlich formuliert werden:

**Satz 14.5.** Es sei  $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Existiert ein Punkt  $\underline{\mathbf{x}}^{(0)} \in U$  mit  $\nabla f(\underline{\mathbf{x}}^{(0)}) = \underline{\mathbf{0}}$ , so gilt:

- i) Ist  $(\text{Hess } f)(\underline{\mathbf{x}}^{(0)})$  *positiv definit*, so ist  $\underline{\mathbf{x}}^{(0)}$  ein *striker Minimierer* von  $f$ .
- ii) Ist  $(\text{Hess } f)(\underline{\mathbf{x}}^{(0)})$  *negativ definit*, so ist  $\underline{\mathbf{x}}^{(0)}$  ein *striker Maximierer* von  $f$ .
- iii) Ist  $(\text{Hess } f)(\underline{\mathbf{x}}^{(0)})$  *indefinit*, so besitzt  $f$  in  $\underline{\mathbf{x}}^{(0)}$  *kein lokales Extremum*.

**Beispiel.** Man betrachte wieder obiges Paraboloid. Die Hessesche Matrix

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

ist stets positiv definit,  $\underline{\mathbf{x}}^{(0)} = \underline{\mathbf{0}}$  ist also ein strikter Minimierer, wie es in Abbildung 14.9 angedeutet ist.

**Bemerkungen.**

i) Im Fall  $m = 2$  kann die Situation wie folgt zusammengefasst werden:

Ist in einem kritischen Punkt  $\underline{\mathbf{x}}^{(0)} \in U$

$$(\text{Hess } f)(\underline{\mathbf{x}}^{(0)}) = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

so ist

- $a > 0, ac - b^2 > 0$  **hinreichend für ein Minimum;**
- $a < 0, ac - b^2 > 0$  **hinreichend für ein Maximum;**
- $a \geq 0, ac - b^2 \geq 0$  **notwendig für ein Minimum;**
- $a \leq 0, ac - b^2 \geq 0$  **notwendig für ein Maximum.**

Ist hingegen  $ac - b^2 < 0$ , so kann  $\underline{\mathbf{x}}^{(0)}$  **keine Extremalstelle** sein.

ii) Zur systematischen Untersuchung des Falls  $m > 2$  sei auf die Literatur verwiesen.

**Beispiel.** Ist  $m = 2$  und

$$f(\underline{\mathbf{x}}) = 6x_1x_2 - 3x_2^2 - 2x_1^3,$$

so sind die einzigen kritischen Punkte

$$\underline{\mathbf{x}}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{x}}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die Hessesche Matrix ist

$$(\text{Hess } f)(\underline{\mathbf{x}}) = \begin{pmatrix} -12x_1 & 6 \\ 6 & -6 \end{pmatrix}.$$

Demnach kann  $\underline{\mathbf{x}}^{(0)}$  keine Extremalstelle sein, in  $\underline{\mathbf{x}}^{(1)}$  hat  $f$  ein lokales Maximum.



# Kapitel 15

## Kurvenintegrale

**Typisches Beispiel.** Man betrachte eine stetig differenzierbare Kurve

$$\gamma : I = [a, b] \rightarrow U \subset \mathbb{R}^m .$$

In einem **Kraftfeld**  $F: U \rightarrow \mathbb{R}^m$  bewege sich ein Massenpunkt **längs dieser Kurve**.

Die zu verrichtende Arbeit (“Kraft  $\times$  Weg”) ist nur vom **Anteil der Kraft in Kurvenrichtung** abhängig, wie es in Abbildung 15.1 angedeutet ist, d.h. es ist das **Skalarprodukt aus Kraft und Weg** zu betrachten.

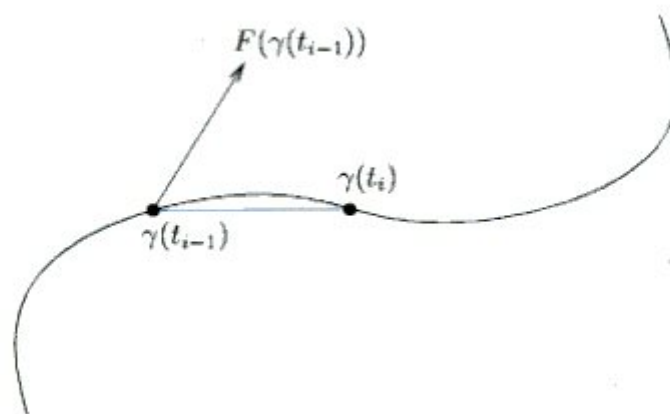


Abbildung 15.1: Zur Berechnung der verrichteten Arbeit.

**Definition 15.1.** Es sei  $U \subset \mathbb{R}^m$  offen,  $F: U \rightarrow \mathbb{R}^m$  sei ein stetiges **Vektorfeld** und  $\gamma: I = [a, b] \rightarrow U$  sei eine glatte Kurve.

Dann heißt

$$\int_{\gamma} \langle F, d\mathbf{x} \rangle := \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt =: \int_{\gamma} \left[ \sum_{i=1}^m F_i dx_i \right]$$

das *Kurvenintegral* des Vektorfeldes  $F$  längs der Kurve  $\gamma$ .

### Bemerkungen.

- i) Um auch **stückweise glatte Kurven**  $\gamma$  betrachten zu können, die sich wie in Abbildung 15.2 aus glatten “Teilkurven”  $\gamma^{(1)}, \dots, \gamma^{(k)}$  stetig zusammensetzen, definiert man

$$\int_{\gamma} \langle F, d\mathbf{x} \rangle := \sum_{i=1}^k \int_{\gamma^{(i)}} \langle F, d\mathbf{x} \rangle .$$

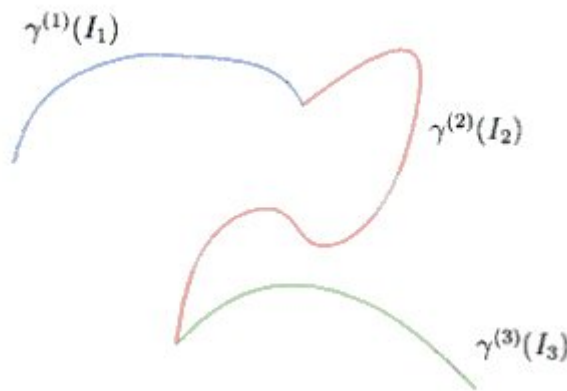


Abbildung 15.2: Eine Stückweise glatte Kurve (Spur).

- ii) Nach obiger Motivation des Kurvenintegrals über den Begriff der Arbeit sollte ein Kurvenintegral **invariant unter orientierungserhaltenden Parametertransformationen** sein (vgl. Definition 14.2).

Tatsächlich gilt für eine Kurve  $\gamma: [a, b] \rightarrow \mathbb{R}^m$ , eine glatte orientierungstreue Parametertransformation  $\varphi: [\alpha, \beta] \rightarrow [a, b]$  sowie  $\tilde{\gamma}$ :

$[\alpha, \beta] \rightarrow \mathbb{R}^m, \tilde{\gamma} := \gamma \circ \varphi:$

$$\begin{aligned} \int_{\tilde{\gamma}} \langle F, d\mathbf{x} \rangle &= \int_{\alpha}^{\beta} \langle F(\tilde{\gamma}(\tau)), \tilde{\gamma}'(\tau) \rangle d\tau \\ &= \int_{\alpha}^{\beta} \langle F \circ \gamma(\varphi(\tau)), \gamma'(\varphi(\tau)) \rangle \varphi'(\tau) d\tau \\ &\stackrel{\text{Subst.}}{=} \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt = \int_{\gamma} \langle F, d\mathbf{x} \rangle . \end{aligned}$$

iii) Ist  $\tilde{\gamma}$  eine Kurve, die mittels einer **orientierungsumkehrenden** Parametertransformation aus  $\gamma$  hervorgeht, so sieht man wie oben

$$\int_{\tilde{\gamma}} \langle F, d\mathbf{x} \rangle = - \int_{\gamma} \langle F, d\mathbf{x} \rangle .$$

Aus diesem Grunde wird manchmal  $-\gamma$  für einen in umgekehrter Orientierung durchlaufenen Weg geschrieben.

### Beispiele.

i) Es fließe ein konstanter Strom  $I$  durch einen unendlich langen Leiter im  $\mathbb{R}^3$ . Es wird ein Magnetfeld

$$B : U = \{ \mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 \neq 0 \} \rightarrow \mathbb{R}^3$$

aufgebaut mit

$$B(\mathbf{x}) = I \begin{pmatrix} -\frac{x_2}{x_1^2 + x_2^2} \\ \frac{x_1}{x_1^2 + x_2^2} \\ 0 \end{pmatrix} .$$

Berechnet werden soll nun das Kurvenintegral längs der Kurve  $\gamma$ ,

$$t \in [0, 2\pi) \mapsto \begin{pmatrix} r \cos(t) \\ r \sin(t) \\ konst. \end{pmatrix} , \quad r > 0 .$$

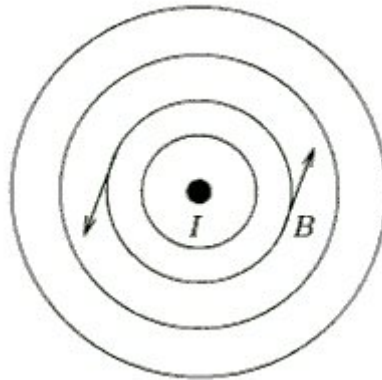


Abbildung 15.3: Magnetfeld um einen Leiter.

Es ist

$$\begin{aligned} \int_{\gamma} \langle B, d\underline{\mathbf{x}} \rangle &= I \int_0^{2\pi} \frac{1}{r^2} \left\langle \begin{pmatrix} -r \sin(t) \\ r \cos(t) \\ 0 \end{pmatrix}, \begin{pmatrix} -r \sin(t) \\ r \cos(t) \\ 0 \end{pmatrix} \right\rangle dt \\ &= I \int_0^{2\pi} 1 dt = 2\pi I . \end{aligned}$$

Auf dieses Beispiel wird in Kürze noch zurückgegriffen werden.

ii) Das Vektorfeld  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  sei gegeben durch

$$F(\underline{\mathbf{x}}) = \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 x_2 \end{pmatrix} .$$

Betrachtet seien weiter die Kurven

$$\begin{aligned} \gamma : t \in [0, 1] &\rightarrow \begin{pmatrix} t \\ t \end{pmatrix} \in \mathbb{R}^2 , \\ \tilde{\gamma} : t \in [0, 1] &\rightarrow \begin{pmatrix} t \\ t^2 \end{pmatrix} \in \mathbb{R}^2 . \end{aligned}$$

Zu beachten ist dabei: **Anfangs- und Endpunkt beider Kurven stimmen überein** (vgl. Abbildung 15.4).

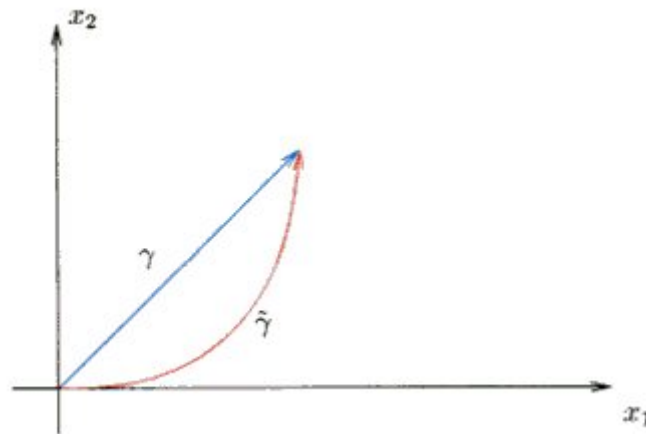


Abbildung 15.4: Zur Wegabhängigkeit des Kurvenintegrals.

Es gilt aber:

$$\int_{\gamma} \langle F, d\underline{x} \rangle = \int_0^1 \left\langle \begin{pmatrix} t^2 + t^2 \\ t^2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle dt = \int_0^1 3t^2 dt = 1,$$

$$\begin{aligned} \int_{\tilde{\gamma}} \langle F, d\underline{x} \rangle &= \int_0^1 \left\langle \begin{pmatrix} t^2 + t^4 \\ t^3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2t \end{pmatrix} \right\rangle dt = \int_0^1 (3t^4 + t^2) dt \\ &= \frac{14}{15}. \end{aligned}$$

Im Allgemeinen hängt das Kurvenintegral also nicht nur vom Anfangs- und Endpunkt der Kurve ab, es kommt auch auf die spezielle Wahl des Weges an.

### Wegunabhängigkeit des Kurvenintegrals?

**Frage.** In welchen Fällen ist das Kurvenintegral wegunabhängig, d.h. nur vom Anfangs- und vom Endpunkt abhängig?

Dabei ist die Wegunabhängigkeit äquivalent zu

$$\int_{\gamma} \langle F, d\underline{x} \rangle = 0 \quad \text{für jeden geschlossenen Weg,}$$

d.h. für jede geschlossen durchlaufene Kurve  $\gamma$ , d.h.  $\gamma(a) = \gamma(b)$ .

Dies sieht man anhand von Abbildung 15.5 ein, indem man den Weg  $\gamma = \gamma^{(1)} - \gamma^{(2)}$  betrachtet.

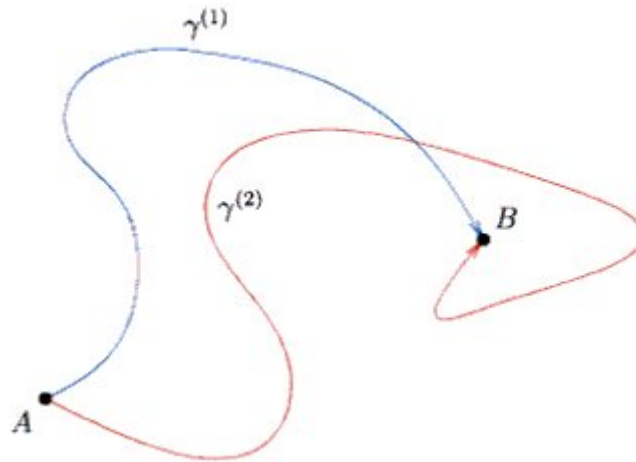


Abbildung 15.5: Wegabhängigkeit und geschlossene Wege.

**Beispiel.** Wird eine Masse im Gravitationsfeld angehoben, so ist die verrichtete Arbeit nur abhängig von der überwundenen Höhe, der Weg, auf dem das geschehen ist, spielt keine Rolle. Das Gravitationsfeld ist **konservativ**.

Werden hingegen Reibungsverluste berücksichtigt, so wird die spezielle Wahl des Weges eine wichtige Rolle spielen.

**Definition 15.2.** Es sei  $U \subset \mathbb{R}^m$  offen und  $F: U \rightarrow \mathbb{R}^m$  sei ein stetiges Vektorfeld.

Falls eine Funktion  $\varphi: U \rightarrow \mathbb{R}$  existiert mit

$$F(\underline{x}) = \nabla\varphi(\underline{x}) \quad \text{für alle } \underline{x} \in U ,$$

so heißt  $\varphi$  ein **Potential** von  $F$ . Das Vektorfeld  $F$  bezeichnet man in diesem Fall als **konservativ**.

**Satz 15.1.** Es sei  $U \subset \mathbb{R}^m$  offen und  $F: U \rightarrow \mathbb{R}^m$  sei ein stetiges, **konservatives** Vektorfeld.

Ist  $\gamma: [a, b] \rightarrow U$  eine (stückweise) glatte Kurve und ist  $\varphi$  ein **Potential** von  $F$ , so gilt

$$\int_{\gamma} \langle F, d\underline{x} \rangle = \varphi(\gamma(b)) - \varphi(\gamma(a)) .$$

**Beweisidee.** Man berechne  $\int_a^b \frac{d}{dt} \varphi(\gamma(t)) dt$  ( $\rightsquigarrow$  Übungen).

**Beispiele.**

i) Es sei  $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,

$$F(\underline{\mathbf{x}}) = \begin{pmatrix} x_1 + x_2^3 \\ 3x_1x_2^2 \\ 1 \end{pmatrix}.$$

Dann gilt  $F = \nabla \varphi$  mit

$$\varphi(\underline{\mathbf{x}}) = \frac{1}{2}x_1^2 + x_1x_2^3 + x_3.$$

Zu beachten ist hier (nachrechnen!)

$$\text{rot } F = \underline{\mathbf{0}}.$$

Allgemein wurde für beliebiges  $\varphi: \mathbb{R}^3 \subset U \rightarrow \mathbb{R}$  von der Klasse  $C^2$  in den Übungen nachgerechnet:

$$\text{rot}(\text{grad } \varphi) = \underline{\mathbf{0}}.$$

Hat demnach ein (glattes) Vektorfeld  $F: \mathbb{R}^3 \subset U \rightarrow \mathbb{R}^3$  ein Potential, so **muss rot  $F$  verschwinden** (als notwendige Bedingung für die Existenz eines Potentials).

ii) Zurück zum Beispiel des unendlich langen Leiters: Es sei wieder

$$B(\underline{\mathbf{x}}) = \begin{pmatrix} -\frac{x_2}{x_1^2 + x_2^2} \\ \frac{x_1}{x_1^2 + x_2^2} \\ 0 \end{pmatrix}.$$

Betrachtet man auf  $U = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1 > 0\}$  die Funktion

$$\varphi(\underline{\mathbf{x}}) = \arctan\left(\frac{x_2}{x_1}\right),$$

so erkennt man

$$\frac{\partial \varphi}{\partial x_1} = \frac{1}{1 + \frac{x_2^2}{x_1^2}} \left( -\frac{1}{x_1^2} x_2 \right) = -\frac{x_2}{x_1^2 + x_2^2},$$

$$\frac{\partial \varphi}{\partial x_2} = \frac{1}{1 + \frac{x_2^2}{x_1^2}} \frac{1}{x_1} = \frac{x_1}{x_1^2 + x_2^2},$$

$$\frac{\partial \varphi}{\partial x_3} = 0,$$

$\varphi$  ist also auf  $U$  ein Potential von  $B$ .

Nach Satz 15.1 kann  $B$  aber kein Potential auf  $\{\underline{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 \neq 0\}$  haben (vgl. auch Abbildung 15.6).

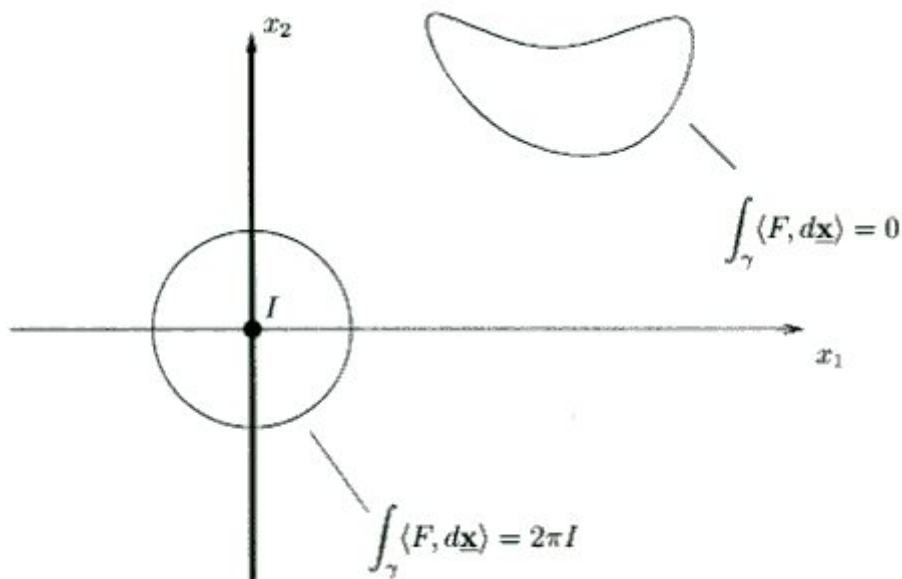


Abbildung 15.6:  $B$  kann kein "globales" Potential haben.

**Übung.** Man berechne  $\operatorname{rot} B$  auf  $\{\underline{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 \neq 0\}$  und interpretiere das Ergebnis.



# Kapitel 16

## Integrale im $\mathbb{R}^n$

### 16.1 Das Riemannsches Integral

**Idee.** Ist  $f: M \subset \mathbb{R}^n \rightarrow \mathbb{R}$  eine beschränkte Funktion, so sollte das Integral  $\int_M f \, dV$  ein Maß für das Volumen des vom Graphen und von  $M$  berandeten Körpers sein, so wie es in Abbildung 16.1 angedeutet ist.

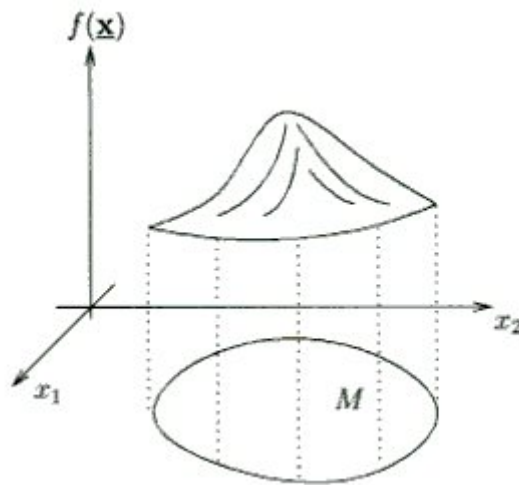


Abbildung 16.1: Zur Idee des Riemannsches Integrals.

**Bemerkung.** Die einführenden Bemerkungen zur Vorgehensweise übertragen sich natürlich sinngemäß aus Kapitel 9.

**Beispiel zur Definition des Integrals.** Es sei  $n = 2$  und es sei  $R = [a, b] \times [c, d] \subset \mathbb{R}^2$  ein Rechteck im  $\mathbb{R}^2$ .

Abbildung 16.2 zeigt, wie Unter- und Obersummen in diesem Fall völlig analog zur eindimensionalen Situation erklärt sind.

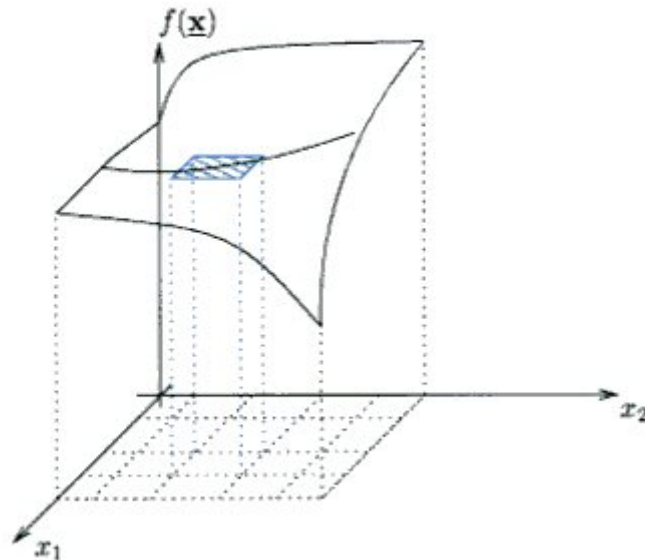


Abbildung 16.2: Unter- und Obersummen.

**Übung.** Es sei  $n \in \mathbb{N}$ ,  $I_1, I_2, \dots, I_n$  seien reelle Intervalle und es sei

$$C := I_1 \times I_2 \times \dots \times I_n \subset \mathbb{R}^n .$$

$C$  heißt dann eine **Zelle** im  $\mathbb{R}^n$ . Man definiere analog zum Fall  $n = 1$  den Begriff **integrierbare Funktion**  $f: C \rightarrow \mathbb{R}$ .

Notation:

$$\begin{aligned} \mathcal{I}(f) &= \int_C f \, dV = \int_C f(\underline{\mathbf{x}}) \, dV \\ &= \int \int \dots \int_C f(\underline{\mathbf{x}}) \, dx_1 \dots dx_n , \end{aligned}$$

$\mathcal{I}(f)$  heißt das **Riemannsche Integral** von  $f$  über die Zelle  $C$ ,  $dV$  heißt das **Volumenelement**.

**Bemerkung.** Die Bemerkungen und Integrabilitätskriterien aus dem Fall  $n = 1$  übertragen sich analog, beispielsweise gilt

**Satz 16.1.** Ist  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  *stetig* auf der Zelle  $C$ , so ist  $f$  integrierbar auf  $C$ .

### Wie berechnet man Integrale über Zellen im $\mathbb{R}^n$ ?

**Satz 16.2.** Es sei o.E.  $n = 2$  (für  $n \geq 3$  gelten analoge Aussagen). Es sei weiter  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  integrierbar auf der Zelle

$$C = \{\underline{\mathbf{x}} \in \mathbb{R}^2 : a \leq x_1 \leq b, c \leq x_2 \leq d\}.$$

i) Falls für jedes *fixierte*  $x_1 \in [a, b]$  das Integral  $\int_c^d f(\underline{\mathbf{x}}) dx_2$  existiert, so existiert das *“iterierte”* Integral

$$\int_a^b \left[ \int_c^d f(\underline{\mathbf{x}}) dx_2 \right] dx_1$$

und es gilt

$$\int_C f(\underline{\mathbf{x}}) dV = \int_a^b \left[ \int_c^d f(\underline{\mathbf{x}}) dx_2 \right] dx_1.$$

ii) Falls für jedes *fixierte*  $x_2 \in [c, d]$  das Integral  $\int_a^b f(\underline{\mathbf{x}}) dx_1$  existiert, so gilt die analoge Aussage.

### Bemerkungen.

i) Mit diesem Satz ist erst die Notation als **Mehrfachintegral**  $\int \int_C f(\underline{\mathbf{x}}) dx_1 dx_2$  gerechtfertigt.

ii) Die Berechnung eines Integrals im  $\mathbb{R}^n$  wird mithilfe von Satz 16.2 auf zwei nacheinander auszuführende “eindimensionale” Integrationen zurückgeführt.

Damit **übertragen sich** (unter den Voraussetzungen von Satz 16.2) **die bekannten Eigenschaften** aus der Integralrechnung einer Variablen (Linearität ...)

iii) Wie Beispiele zeigen, folgt allein aus der Existenz der iterierten Integrale in Satz 16.2 nicht die Existenz von  $\int_C f(\underline{x}) \, dV$ .

Alle im Folgenden auftretenden Funktionen sind jedoch als integrierbar an genommen.

iv) Es sei also  $f$  beispielsweise stetig auf  $C$ , so dass alle auftretenden Integrale existieren.

Dann gilt nach Satz 16.2 für das Zweifachintegral (mit der Notation  $f(\underline{x}) = f(x_1, x_2)$ )

$$(a) \int_C f(\underline{x}) \, dx = \int_a^b \left[ \int_c^d f(x_1, x_2) \, dx_2 \right] dx_1 ,$$

also wird für festes  $x_1$  erst das Einfachintegral

$$\int_c^d f(x_1, x_2) \, dx_2 =: g(x_1)$$

bestimmt, dann die Funktion  $g(x_1)$  bzgl.  $x_1$  integriert.

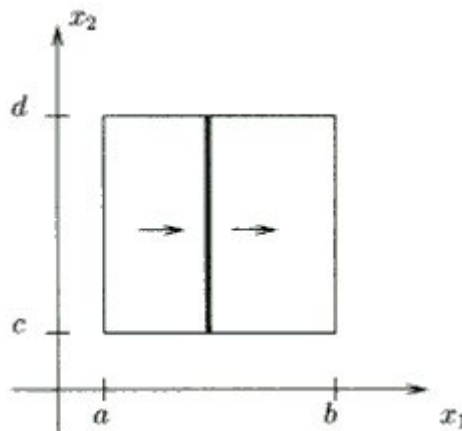


Abbildung 16.3: Eine Möglichkeit zur Berechnung eines Zweifachintegrals.

$$(b) \int_C f(\underline{x}) \, dx = \int_c^d \left[ \int_a^b f(x_1, x_2) \, dx_1 \right] dx_2 .$$

Hier ist die Situation umgekehrt (siehe Abbildung 16.4).

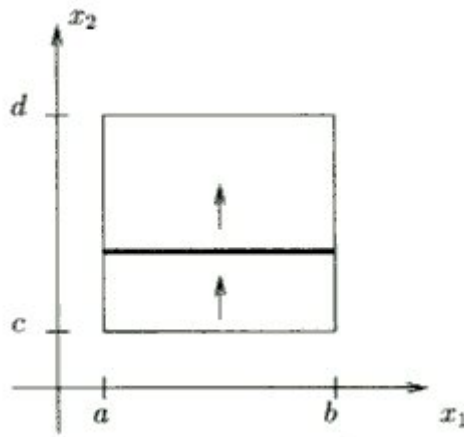


Abbildung 16.4: Eine zweite Möglichkeit zur Berechnung eines Zweifachintegrals.

**Beispiele.**

i) Es sei  $C = \{\underline{\mathbf{x}} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 1 \leq x_2 \leq 2\}$ ,  $f(x_1, x_2) = x_1^{x_2}$ .

Dann ist  $f$  stetig auf  $C$ , insbesondere sind die Voraussetzungen aus Satz 16.2 erfüllt, es gilt

$$\int_C x_1^{x_2} dV = \int_1^2 \left[ \int_0^1 x_1^{x_2} dx_1 \right] dx_2,$$

$$\int_0^1 x_1^{x_2} dx_1 = \frac{x_1^{x_2+1}}{x_2+1} \Big|_0^1 = \frac{1}{x_2+1}$$

und damit

$$\int_C x_1^{x_2} dV = \int_1^2 \frac{1}{x_2+1} dx_2 = \ln(1+x_2) \Big|_1^2 = \ln(3/2).$$

ii) Es sei jetzt  $n = 3$ ,

$$C = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 1, 2 \leq x_3 \leq 4\}.$$

Als stetiger Integrand  $f$  sei die Funktion

$$f(\underline{\mathbf{x}}) = x_1 + x_2 + x_3.$$

betrachtet. Mit der Notation

$$C_{x_2, x_3} = \left\{ \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} : 0 \leq x_2 \leq 1, 2 \leq x_3 \leq 4 \right\}$$

folgt

$$\int_C f(\underline{x}) \, dV = \int_{C_{x_2, x_3}} \left[ \underbrace{\int_0^2 (x_1 + x_2 + x_3) \, dx_1}_{=:g(x_2, x_3)} \right] dV(x_2, x_3),$$

wobei  $dV(x_2, x_3)$  das zweidimensionale Volumenelement (bzgl. der Variablen  $x_2, x_3$ ) bezeichne.

Eine nochmalige Anwendung von Satz 16.2 liefert

$$\int_{C_{x_2, x_3}} g(x_2, x_3) \, dV(x_2, x_3) = \int_2^4 \left[ \int_0^1 g(x_2, x_3) \, dx_2 \right] dx_3$$

und insgesamt:

$$\begin{aligned} \int_C f(\underline{x}) \, dV &= \int_2^4 \left[ \int_0^1 \left[ \int_0^2 (x_1 + x_2 + x_3) \, dx_1 \right] dx_2 \right] dx_3 \\ &= \int_2^4 \left[ \int_0^1 \left[ \frac{x_1^2}{2} + x_1(x_2 + x_3) \right]_0^2 dx_2 \right] dx_3 \\ &= \int_2^4 \left[ \int_0^1 [2 + 2x_2 + 2x_3] dx_2 \right] dx_3 \\ &= \int_2^4 [2x_2(1 + x_3) + x_2^2]_0^1 dx_3 \\ &= \int_2^4 [3 + 2x_3] dx_3 = [3x_3 + x_3^2]_2^4 = 18. \end{aligned}$$

### Man will aber nicht nur über Zellen integrieren!

Bei der Integration in einer Veränderlichen genügt es, Integrale über Intervalle definiert zu haben.

In mehreren Veränderlichen ist die Beschränkung auf Zellen hingegen viel zu restriktiv.

Das Riemannsche Integral kann etwa auf sogenannten **quadrierbaren Mengen** eingeführt werden, wozu hier lediglich auf die Literatur verwiesen sei.

Eine kleine aber für praktische Zwecke gut handbare Klasse ist:

**Definition 16.1.** i) Eine Menge  $M \subset \mathbb{R}^2$  heißt  *$x_2$ -projizierbar* oder *Normalbereich in  $x_2$ -Richtung*, falls  $M$  von der Form ist

$$M = \{\underline{\mathbf{x}} \in \mathbb{R}^2 : x_1 \in [a, b], \varphi_1(x_1) \leq x_2 \leq \varphi_2(x_1)\},$$

wobei  $\varphi_1, \varphi_2$  zwei auf  $[a, b]$  stetige Funktionen seien.

ii) Eine Menge  $M \subset \mathbb{R}^2$  heißt  *$x_1$ -projizierbar* oder *Normalbereich in  $x_1$ -Richtung*, falls  $M$  von der Form ist

$$M = \{\underline{\mathbf{x}} \in \mathbb{R}^2 : x_2 \in [c, d], \psi_1(x_2) \leq x_1 \leq \psi_2(x_2)\},$$

wobei  $\psi_1, \psi_2$  zwei auf  $[c, d]$  stetige Funktionen seien.

iii) Eine Menge  $M \subset \mathbb{R}^n$  heißt *Normalbereich in  $x_j$ -Richtung*,  $1 \leq j \leq n$ , wenn sie die Form

$$M = \{\underline{\mathbf{x}} \in \mathbb{R}^n : \underline{\mathbf{y}} \in K, \xi_1(\underline{\mathbf{y}}) \leq x_j \leq \xi_2(\underline{\mathbf{y}})\}$$

hat, wobei

$$\underline{\mathbf{y}} := \begin{pmatrix} x_1 \\ \vdots \\ x_{j-1} \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n-1}$$

und wobei  $K$  eine (präzise: kompakte, quadrierbare Menge) des  $\mathbb{R}^{n-1}$  bezeichne. Die Funktionen  $\xi_1, \xi_2: K \rightarrow \mathbb{R}$  seien stetig.

**Idee.** M.a.W. bedeutet etwa  $x_2$ -Projizierbarkeit, dass  $M$  von den Graphen der Funktionen  $\varphi_1(x_1)$  und  $\varphi_2(x_1)$  eingeschlossen ist.

Bei einer Integration in Verallgemeinerung von Satz 16.2 kann dann zunächst das Integral bzgl.  $x_2$  für fixiertes  $x_1$  zwischen den Grenzen  $\varphi_1(x_1)$  und  $\varphi_2(x_1)$  berechnet werden, anschließend wird das von  $x_1$  abhängende Ergebnis bzgl.  $x_1$  integriert.

### Beispiele zur Projizierbarkeit.

- i) Zweidimensionale Beispiele sind in den Abbildungen 16.5–16.8 dargestellt.

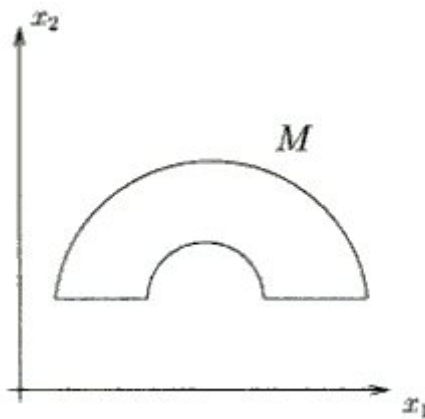


Abbildung 16.5:  $M$  ist  $x_2$  projizierbar, nicht  $x_1$  projizierbar.

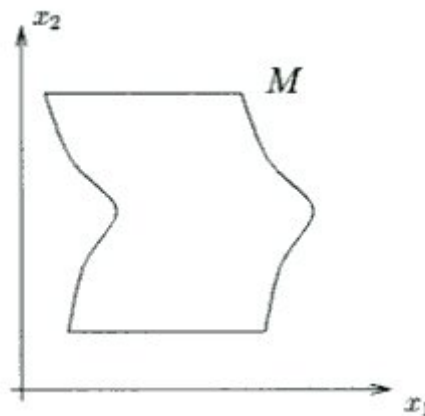
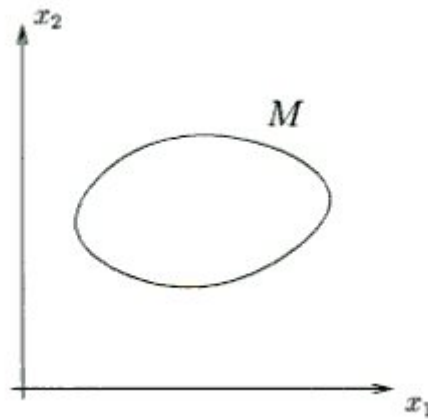
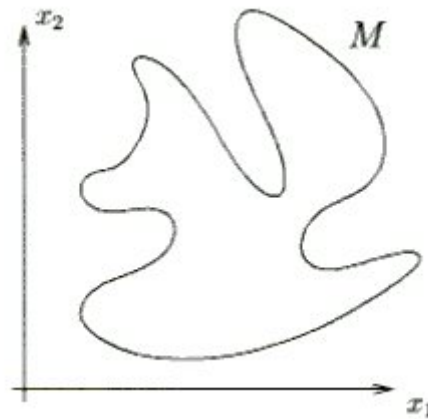


Abbildung 16.6:  $M$  ist  $x_1$  projizierbar, nicht  $x_2$  projizierbar.

- ii) Man betrachte obiges dreidimensionales Beispiel der Zelle

$$C = \{ \underline{\mathbf{x}} \in \mathbb{R}^3 : 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 1, 2 \leq x_3 \leq 4 \} .$$



Abbildung 16.7:  $M$  ist  $x_1$  projizierbar und  $x_1$  projizierbar.Abbildung 16.8:  $M$  ist weder  $x_1$  projizierbar noch  $x_1$  projizierbar.

Mit der Notation

$$\underline{\mathbf{y}} = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}, \quad \xi_1(\underline{\mathbf{y}}) \equiv 0, \quad \xi_2(\underline{\mathbf{y}}) \equiv 2,$$

ist

$$K = C_{x_2, x_3} = \{\underline{\mathbf{y}} : 0 \leq x_2 \leq 1, 2 \leq x_3 \leq 4\}$$

und

$$C = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{y}} \in K, \xi_1(\underline{\mathbf{y}}) \leq x_1 \leq \xi_2(\underline{\mathbf{y}})\},$$

die Zelle ist also ein Normalbereich in  $x_1$ -Richtung (selbstverständlich ebenso in  $x_2$ - und  $x_3$ -Richtung).

- iii) Es sei  $M \subset \mathbb{R}^3$  das ‘‘Raumstück’’, welches den beiden Zylindern  $\{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq 1\}$  und  $\{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1^2 + x_3^2 \leq 1\}$  gemeinsam

ist,

$$M = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq 1\} \cap \{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1^2 + x_3^2 \leq 1\} .$$

Dann ist  $M$  Normalbereich in  $x_3$ -Richtung (analog in  $x_2$ -Richtung), denn mit

$$\underline{\mathbf{y}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in K := \{\underline{\mathbf{y}} \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$$

und mit

$$\xi_1(\underline{\mathbf{y}}) = -\sqrt{1 - x_1^2}, \quad \xi_2(\underline{\mathbf{y}}) = \sqrt{1 - x_1^2} \quad \text{für alle } \underline{\mathbf{y}} \in K$$

ist

$$M = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{y}} \in K, \xi_1(\underline{\mathbf{y}}) \leq x_3 \leq \xi_2(\underline{\mathbf{y}})\} .$$

**Übung.** Fertigen Sie eine Skizze an.

*iv)* Es sei

$$M = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_1 + x_2 + x_3 \leq 1\} .$$

Wieder ist  $M$  Normalbereich in  $x_3$ -Richtung (andere Richtungen analog).

Die Menge  $K$  findet man wie folgt. Zunächst muss auf  $K$  gelten:  $x_1 \geq 0$  und  $x_2 \geq 0$ . Es ist weiter in  $K$

$$\underline{\mathbf{y}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \xi_1(\underline{\mathbf{y}}) = 0, \quad \xi_2(\underline{\mathbf{y}}) = 1 - x_1 - x_2 .$$

Da die Bedingung  $\xi_1(\underline{\mathbf{y}}) \leq \xi_2(\underline{\mathbf{y}})$  auf ganz  $K$  erfüllt sein muss, ergibt sich

$$K = \{\underline{\mathbf{y}} \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\} .$$

**Übung.** Fertigen Sie eine Skizze an.

Als Definition des Riemannsches Integrals über Normalbereiche und gleichzeitig als Satz zu deren Berechnung kann nun formuliert werden ([Cavalierisches Prinzip](#) zur Berechnung von Volumina für  $f \equiv 1$  oder der [Satz von Fubini](#))

**Satz 16.3.** *Es sei  $M \subset \mathbb{R}^n$  und  $f: M \rightarrow \mathbb{R}$  stetig. Dann gilt mit den obigen Bezeichnungen:*

i) *Ist  $n = 2$  und  $M$  ein Normalbereich in  $x_2$ -Richtung, so gilt*

$$\int_M f \, dV = \int_a^b \left[ \int_{\varphi_1(x_1)}^{\varphi_2(x_1)} f(x_1, x_2) \, dx_2 \right] dx_1 .$$

ii) *Ist  $n = 2$  und  $M$  ein Normalbereich in  $x_1$ -Richtung, so gilt*

$$\int_M f \, dV = \int_c^d \left[ \int_{\psi_1(x_2)}^{\psi_2(x_2)} f(x_1, x_2) \, dx_1 \right] dx_2 .$$

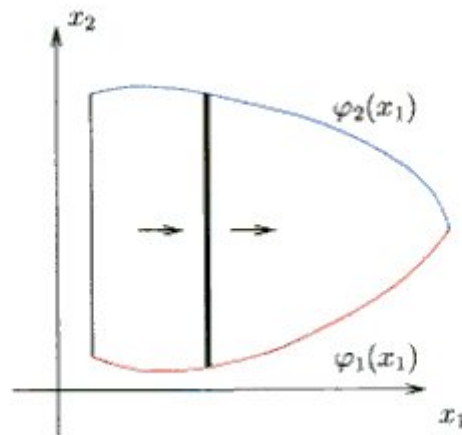
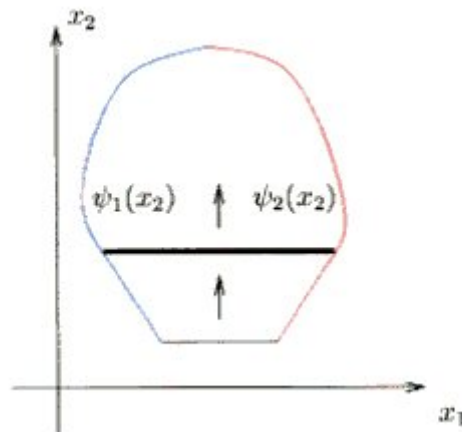
iii) *Im allgemeinen Fall  $n \geq 2$ ,  $M$  Normalbereich in eine Richtung  $x_j$ ,  $1 \leq j \leq n$ , gilt*

$$\int_M f \, dV = \int_K \left[ \int_{\xi_1(\underline{y})}^{\xi_2(\underline{y})} f(\underline{x}) \, dx_j \right] dV_{n-1} ,$$

wobei  $dV_{n-1}$  das  $(n - 1)$ -dimensionale Volumenelement auf  $K \subset \mathbb{R}^{n-1}$  bezeichne.

### Bemerkungen.

- i) Der Satz führt Integrale in  $\mathbb{R}^n$  als Mehrfachintegrale ein und ist damit auch ein wesentliches Hilfsmittel zur **konkreten Berechnung**. Analog zu Satz 16.2 ist die anschauliche Vorstellung in den Abbildungen 16.9 und 16.10 wiedergegeben.
- ii) Ist  $M$  in verschiedene Richtungen ein Normalbereich, so kann natürlich die "einfachste" Berechnungsmethode gewählt werden.
- iii) Ist  $M$  kein Normalbereich, so kann versucht werden,  **$M$  in endlich viele Normalbereiche zu zerlegen**, um Satz 16.3 anzuwenden.

Abbildung 16.9: Iterierte Integration für einen  $x_2$ -Normalbereich.Abbildung 16.10: Iterierte Integration für einen  $x_1$ -Normalbereich.

### Beispiele.

- i) Es sei  $R > 0$  fixiert und  $M := \{\underline{\mathbf{x}} \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq R^2, x_2 \geq 0\}$ .  
Der Integrand  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  sei gegeben durch

$$f(\underline{\mathbf{x}}) = x_1^2 x_2 .$$

1<sup>te</sup> Möglichkeit: Man rechne nach Satz 16.3, i) (vgl. Abbildung 16.12):

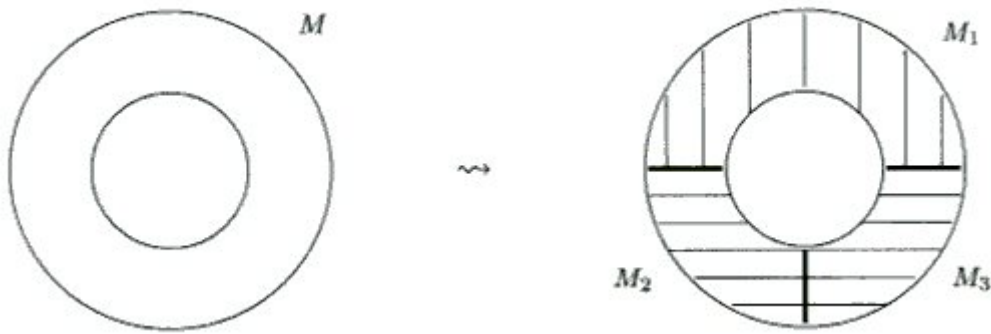


Abbildung 16.11: Eine Zerlegung in Normalbereiche.

$$\begin{aligned}
 \int_M x_1^2 x_2 \, dV &= \int_{-R}^R \left[ \int_0^{\sqrt{R^2-x_1^2}} x_1^2 x_2 \, dx_2 \right] dx_1 \\
 &= \int_{-R}^R \left[ \frac{x_1^2 x_2^2}{2} \right]_0^{\sqrt{R^2-x_1^2}} dx_1 \\
 &= \frac{1}{2} \int_{-R}^R x_1^2 (R^2 - x_1^2) dx_1 \\
 &= \frac{1}{2} \left[ \frac{x_1^3}{3} R^2 - \frac{x_1^5}{5} \right]_{-R}^R = \frac{2}{15} R^5.
 \end{aligned}$$

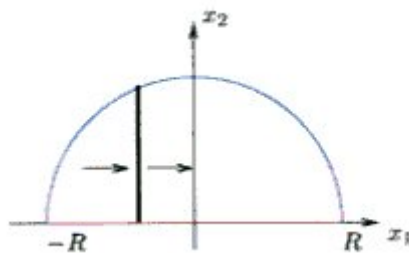
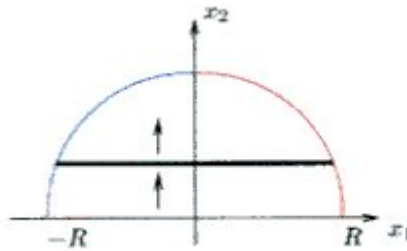


Abbildung 16.12: 1<sup>te</sup> Möglichkeit.

2<sup>te</sup> Möglichkeit: Man rechne nach Satz 16.3, *ii*) (vgl. Abbildung 16.13):

$$\begin{aligned}
\int_M x_1^2 x_2 \, dV &= \int_0^R \left[ \int_{-\sqrt{R^2-x_2^2}}^{\sqrt{R^2-x_2^2}} x_1^2 x_2 \, dx_1 \right] dx_2 \\
&= \int_0^R \left[ \frac{x_1^3}{3} x_2 \right]_{-\sqrt{R^2-x_2^2}}^{\sqrt{R^2-x_2^2}} dx_2 \\
&= \frac{1}{3} \int_0^R 2x_2 (R^2 - x_2^2)^{\frac{3}{2}} dx_2 \\
&= \frac{1}{3} \left[ \left( -\frac{2}{5} \right) (R^2 - x_2^2)^{\frac{5}{2}} \right]_0^R = \frac{2}{15} R^5.
\end{aligned}$$

Abbildung 16.13: 2<sup>te</sup> Möglichkeit.

ii) Gesucht sei das Volumen  $V$  des Tetraeders  $T$  mit den Ecken  $O = \mathbf{0}$ ,

$$A = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ b \\ 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}, \quad a, b, c > 0 \text{ fixiert,}$$

d.h. man berechne

$$\int_T 1 \, dV,$$

wobei  $dV$  das dreidimensionale Volumenelement bezeichne.

Um die richtigen Integrationsgrenzen zu finden, beachtet man, dass die Punkte  $A, B, C$  in der durch

$$\frac{x_1}{a} + \frac{x_2}{b} + \frac{x_3}{c} = 1$$

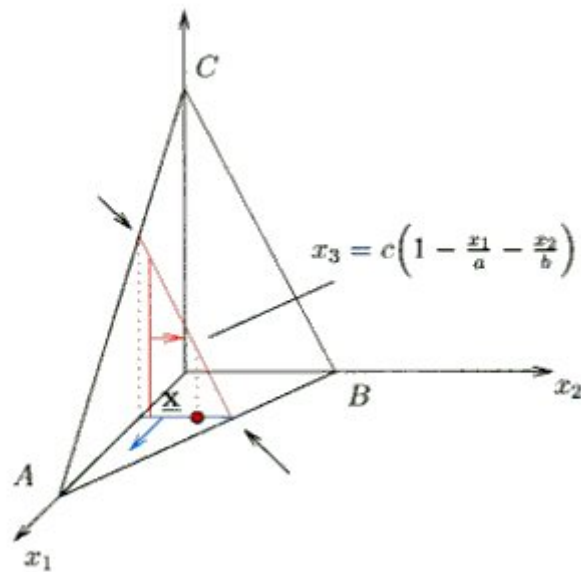


Abbildung 16.14: Zur Volumenberechnung des Tetraeders.

gegebenen Ebene liegen (vgl. 16.14). Das gesuchte Volumen berechnet sich damit zu

$$\begin{aligned}
 V &= \int_T 1 \, dV \\
 &= \int_0^a \left[ \int_0^{b(1-x_1/a)} \left[ \int_0^{c(1-x_1/a-x_2/b)} dx_3 \right] dx_2 \right] dx_1 \\
 &= \int_0^a \left[ \int_0^{b(1-x_1/a)} c \left(1 - \frac{x_1}{a} - \frac{x_2}{b}\right) dx_2 \right] dx_1 \\
 &= \int_0^a c \left[ \left(1 - \frac{x_1}{a}\right)x_2 - \frac{1}{2} \frac{x_2^2}{b} \right]_0^{b(1-x_1/a)} dx_1 \\
 &= \int_0^a \frac{bc}{2} \left(1 - \frac{x_1}{a}\right)^2 dx_1 = \frac{bc}{2} \left[ -\frac{a}{3} \left(1 - \frac{x_1}{a}\right) \right]_0^a \\
 &= \frac{abc}{6}.
 \end{aligned}$$

**Übung.** Man berechne das Integral von

$$f(x_1, x_2) = c \left( 1 - \frac{x_1}{a} - \frac{x_2}{b} \right)$$

über dem Dreieck mit den Ecken  $O, A, B$  in der  $(x_1, x_2)$ -Ebene (vgl. wieder der Abbildung 16.14) und interpretiere das Ergebnis.

## 16.2 Der Transformationssatz

Kann analog zum Fall einer Funktion einer Variablen die Integration in bestimmten Situationen durch eine (**Substitution**) **Transformation** (d.h. durch eine geeignete **Koordinatentransformation**) vereinfacht werden?

Betrachtet seien dazu zwei Mengen  $U, V \subset \mathbb{R}^n$  (präzise: mit nur inneren Punkten) und eine Abbildung  $g: U \rightarrow V$ ,

$$U \ni \underline{\mathbf{u}} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \mapsto g(\underline{\mathbf{u}}) = \underline{\mathbf{v}} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in V.$$

Die Abbildung  $g$  sei ein **Diffeomorphismus**, d.h. bijektiv und sowohl  $g$  als auch  $g^{-1}$  seien stetig differenzierbar.

Die sogenannte **Jacobische Funktionaldeterminante** ist für einen solchen Diffeomorphismus stets von Null verschieden, d.h. für alle  $\underline{\mathbf{u}} \in U$  gilt

$$\det Dg(\underline{\mathbf{u}}) = \left| \begin{pmatrix} \frac{\partial g_1}{\partial u_1} & \cdots & \frac{\partial g_1}{\partial u_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial u_1} & \cdots & \frac{\partial g_n}{\partial u_n} \end{pmatrix} \right| (\underline{\mathbf{u}}) \neq 0.$$

Weiterhin sei eine Teilmenge  $K \subset V$  gegeben (präzise:  $K$  wird als kompakt vorausgesetzt, – ein typisches Beispiel einer kompakten Menge im  $\mathbb{R}^n$  ist die in den Übungen diskutierte “Kugel mit Rand”  $\overline{B_\varepsilon(\underline{\mathbf{x}}^{(0)})}$ ).

**Unter diesen Voraussetzungen** gilt der



**Satz 16.4.** (*Transformationssatz*)

Ist  $f: V \rightarrow \mathbb{R}$  stetig, so gilt

$$\int_K f(\underline{v}) \, dV(\underline{v}) = \int_{g^{-1}(K)} f(g(\underline{u})) |\det Dg(\underline{u})| \, dV(\underline{u}).$$

Dabei ist

$$g^{-1}(K) = \{\underline{u} \in U : g(\underline{u}) \in K\}$$

und im Volumenelement ist angedeutet, bzgl. welcher Variablen zu integrieren ist.

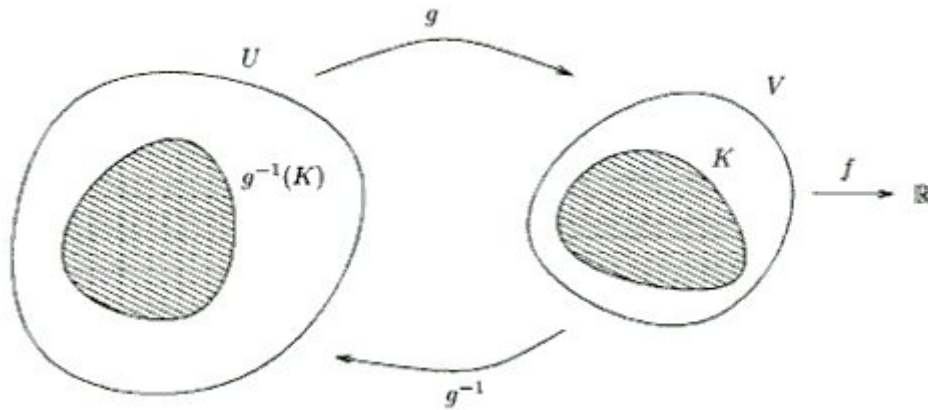


Abbildung 16.15: Zum Transformationssatz.

**Bemerkung.** Satz 16.4 hat formal die gleiche Gestalt wie die Substitutionsregel im Fall  $n = 1$ . Allerdings ist hier der **Betrag** der Jacobischen Funktionaldeterminante zu betrachten.

**Beispiele.**

i) Im  $\mathbb{R}^2$  seien Polarkoordinaten betrachtet: Es ist

$$\underline{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} r \\ \varphi \end{pmatrix}, \quad 0 < r < \infty, \quad 0 < \varphi < 2\pi,$$

sowie

$$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = g(r, \varphi) = \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix}.$$

Die Menge

$$U = \left\{ \begin{pmatrix} r \\ \varphi \end{pmatrix} : 0 < r < \infty, 0 < \varphi < 2\pi \right\}$$

ist offen, die Bildmenge

$$V = \{ \underline{x} \in \mathbb{R}^2 : x_2 \neq 0 \text{ falls } x_1 \geq 0 \}$$

ist die “aufgeschlitzte”  $(x_1, x_2)$ -Ebene.

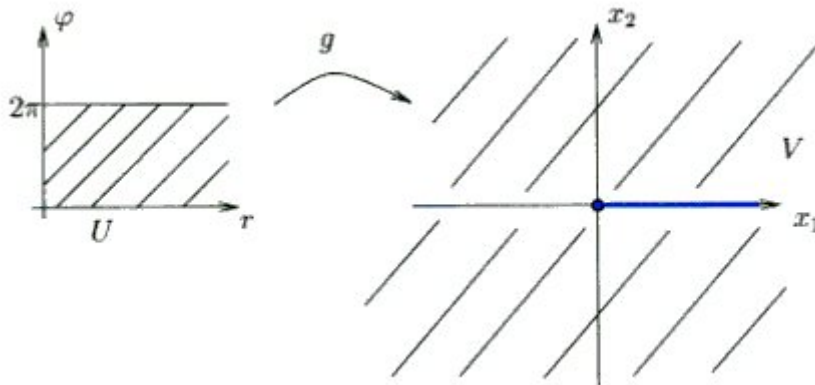


Abbildung 16.16: Polarkoordinaten im  $\mathbb{R}^2$ .

Die Funktionaldeterminante der Transformation  $g$  ist

$$|\det Dg| = \left| \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{pmatrix} \right| = r \neq 0.$$

ii) Es sei  $K = \{ \underline{x} \in \mathbb{R}^2 : 1 \leq x_1^2 + x_2^2 \leq 4 \}$ . Gesucht ist

$$\int_K \sqrt{x_1^2 + x_2^2} \, dV.$$

**Bemerkung.** Die Menge  $K$  ist nicht vollständig in der offenen Menge  $V$  aus i) enthalten, da die  $(x_1, x_2)$ -Ebene geschlitzt ist.

In  $V$  fehlt aber nur eine Menge mit “Volumen Null”, nämlich eine “Linie”, was die hier auftretenden Integrale nicht verändert und die folgenden Rechnungen rechtfertigt (vgl. Abbildung 16.17 für eine formal korrekte Argumentation).

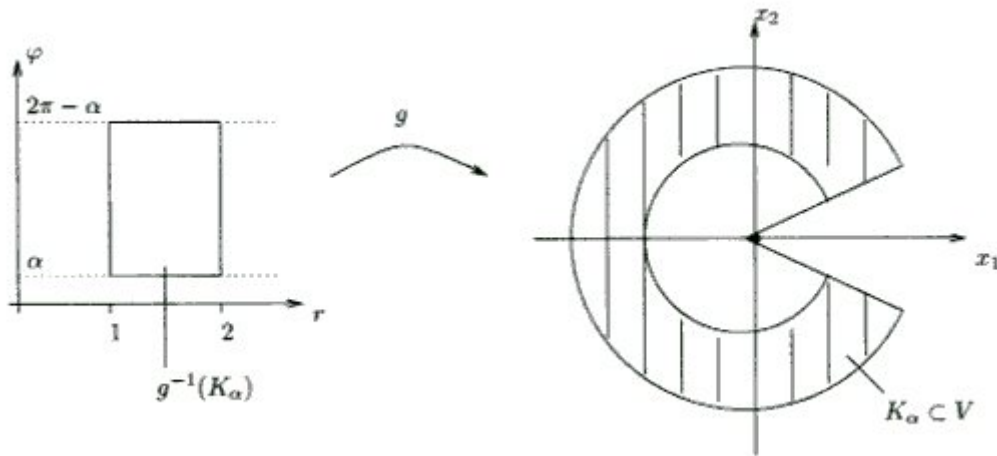


Abbildung 16.17: Formal muss die schraffierte Menge betrachtet und zum Grenzübergang  $\alpha \rightarrow 0$  übergegangen werden.

Als Abbildung  $g: g^{-1}(K) \rightarrow K$  werden Polarkoordinaten eingeführt, und aus Satz 16.4 folgt

$$\begin{aligned} \int_K \sqrt{x_1^2 + x_2^2} \, dV(\underline{\mathbf{x}}) &= \int_{g^{-1}(K)} \sqrt{r^2 \cos^2(\varphi) + r^2 \sin^2(\varphi)} \, r \, dV(r, \varphi) \\ &= \int_0^{2\pi} \left[ \int_1^2 r^2 \, dr \right] d\varphi = \int_0^{2\pi} \frac{7}{3} \, d\varphi = \frac{14}{3} \pi. \end{aligned}$$

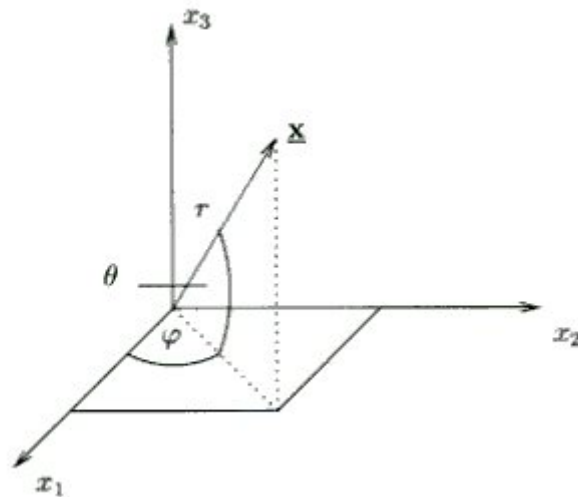
iii) [Kugelkoordinaten im  \$\mathbb{R}^3\$](#) .

Hier wird die Transformation

$$\underline{\mathbf{u}} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} r \\ \varphi \\ \theta \end{pmatrix} \xrightarrow{g} \underline{\mathbf{v}} = \underline{\mathbf{x}} = \begin{pmatrix} r \cos(\varphi) \cos(\theta) \\ r \sin(\varphi) \cos(\theta) \\ r \sin(\theta) \end{pmatrix}.$$

betrachtet. Es ist

$$U = \left\{ \begin{pmatrix} r \\ \varphi \\ \theta \end{pmatrix} : r > 0, 0 < \varphi < 2\pi, -\frac{\pi}{2} < \theta < \frac{\pi}{2} \right\},$$

Abbildung 16.18: Kugelkoordinaten im  $\mathbb{R}^3$ .

das Bild  $V$  ist der  $\mathbb{R}^3$  ohne die nicht-negative  $x_1$ -Achse und ohne die  $x_3$ -Achse.

Es gilt (Entwicklung nach der dritten Zeile, vgl. Definition 12.4)

$$\begin{aligned}
 & |\det Dg| \\
 &= \left| \begin{pmatrix} \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \cos(\theta) & -r \cos(\varphi) \sin(\theta) \\ \sin(\varphi) \cos(\theta) & r \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \sin(\theta) \\ \sin(\theta) & 0 & r \cos(\theta) \end{pmatrix} \right| \\
 &= \sin(\theta) (r^2 \sin^2(\varphi) \cos(\theta) \sin(\theta) + r^2 \cos^2(\varphi) \sin(\theta) \cos(\theta)) \\
 &\quad + r \cos(\theta) (r \cos^2(\varphi) \cos^2(\theta) + r \sin^2(\varphi) \cos^2(\theta)) \\
 &= \sin(\theta) r^2 \sin(\theta) \cos(\theta) + r \cos(\theta) r \cos^2(\theta) \\
 &= r^2 \cos(\theta) \neq 0 \quad \text{für } 0 < r, \quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}.
 \end{aligned}$$

*iv)* Man betrachte jetzt

$$K = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_1^2 + x_2^2 + x_3^2 \leq 1\}$$

und die Funktion

$$f(\underline{\mathbf{x}}) = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Mittels der Transformation auf Kugelkoordinaten folgt

$$\begin{aligned} \int_K \sqrt{x_1^2 + x_2^2 + x_3^2} \, dV(\underline{\mathbf{x}}) &= \int_0^1 \left[ \int_0^{\frac{\pi}{2}} \left[ \int_0^{\frac{\pi}{2}} r r^2 \cos(\theta) \, d\varphi \right] d\theta \right] dr \\ &= \frac{\pi}{2} \int_0^1 \left[ \int_0^{\frac{\pi}{2}} r^3 \cos(\theta) \, d\theta \right] dr \\ &= \frac{\pi}{2} \int_0^1 r^3 \, dr = \frac{\pi}{8}. \end{aligned}$$

**Bemerkung.** Oft sind etwa auch Zylinderkoordinaten im  $\mathbb{R}^3$  hilfreich (vgl. Übungen).



# Kapitel 17

## Der Gaußsche Integralsatz in der Ebene

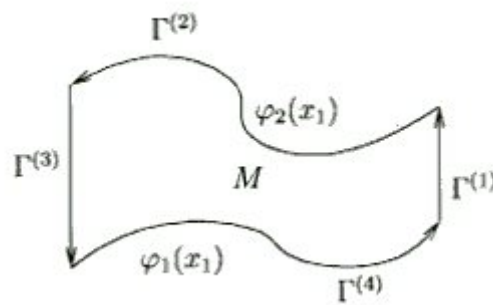
Der [Gaußsche Integralsatz](#) ist einer der bedeutendsten Sätze in der Analysis und deren Anwendungen.

Im zweidimensionalen Fall (in der Ebene) kann er mit den bereits bekannten Mitteln studiert werden, während zum Verständnis des Gaußschen Integralsatzes im  $\mathbb{R}^3$  weiterführende geometrische Betrachtungen nötig sind, die im Rahmen dieser Vorlesung nicht angestellt werden (ähnliches gilt für den [Stokesschen Satz](#)).

Mit dem Gaußschen Satz in der Ebene wird [ein zweidimensionales Integral über eine Divergenz \(Maß für Quellen und Senken\) als Integral über eine Randkurve](#) geschrieben, welches als [Flussintegral](#) zu interpretieren ist.

### Vorbemerkungen.

- i)* Ist  $M \subset \mathbb{R}^2$  ein Normalbereich etwa in  $x_2$ -Richtung (vgl. Definition 16.1), so kann [der Rand](#) von  $M$  wie in Abbildung 17.1 angedeutet durch eine stückweise glatte Kurve  $\gamma$  parametrisiert werden. Die Spur der Kurve ist  $\text{spur}\gamma = \Gamma^{(1)} \cup \Gamma^{(2)} \cup \Gamma^{(3)} \cup \Gamma^{(4)}$ .
- ii)* Es wird stets angenommen, dass diese Kurve regulär ist ( $\gamma' \neq 0$  auf den Teilkurven) und dass sie so orientiert ist, dass  $M$  beim Durchlaufen der Kurve zur Linken liegt, d.h.  [\$\partial M\$  ist positiv orientiert](#).

Abbildung 17.1: Der Rand  $\partial M$ .

Die Idee, den Gaußschen Integralsatz aus dem Satz von Fubini (Satz 16.3) und dem Hauptsatz der Differential- und Integralrechnung (Satz 9.4) abzuleiten, sei anhand des einfachsten Beispiels angedeutet.

**Beispiel.** Man betrachte das Quadrat

$$Q := \{\mathbf{x} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$$

und überlege sich, dass es von den Kurven (jeweils auf  $[0, 1]$  definiert)

$$\begin{aligned} \gamma^{(1)} : t &\mapsto \begin{pmatrix} t \\ 0 \end{pmatrix}, & \gamma^{(2)} : t &\mapsto \begin{pmatrix} 1 \\ t \end{pmatrix}, \\ \gamma^{(3)} : t &\mapsto \begin{pmatrix} 1-t \\ 1 \end{pmatrix}, & \gamma^{(4)} : t &\mapsto \begin{pmatrix} 0 \\ 1-t \end{pmatrix} \end{aligned}$$

positiv orientiert berandet wird.

Für ein glattes Vektorfeld  $F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  ist nach den oben genannten Sätzen (mit der Notation  $F(\mathbf{x}) = F(x_1, x_2)$ )

$$\begin{aligned} \int_Q \frac{\partial F_2}{\partial x_1} \, dV &= \int_0^1 \left[ \int_0^1 \frac{\partial F_2}{\partial x_1} \, dx_1 \right] dx_2 \\ &= \int_0^1 [F_2(1, x_2) - F_2(0, x_2)] \, dx_2 \\ &= \int_{\gamma^{(2)}} \langle F, d\mathbf{x} \rangle + \int_{\gamma^{(4)}} \langle F, d\mathbf{x} \rangle. \end{aligned}$$



Analog berechnet sich

$$-\int_Q \frac{\partial F_1}{\partial x_2} dV = \int_{\gamma^{(1)}} \langle F, d\mathbf{x} \rangle + \int_{\gamma^{(2)}} \langle F, d\mathbf{x} \rangle .$$

Bezeichnet  $\gamma$  die aus  $\gamma^{(1)}$  bis  $\gamma^{(4)}$  zusammengesetzte stückweise glatte Kurve, so ist gezeigt

$$\int_Q \left( \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dV = \int_{\gamma} \langle F, d\mathbf{x} \rangle .$$

Allgemein gilt:

**Satz 17.1.** (Gaußscher Integralsatz in der Ebene)

Es sei  $M \subset \mathbb{R}^2$  sowohl ein Normalbereich in  $x_1$ -Richtung als auch ein Normalbereich in  $x_2$ -Richtung.

Der Rand  $\partial M$  sei positiv orientiert, stückweise glatt und regulär.

Weiterhin sei  $F: U \rightarrow \mathbb{R}^2$  ein stetig differenzierbares Vektorfeld, definiert auf einer offenen Obermenge  $U \supset M$ .

Schreibt man  $F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$  und ist  $\gamma$  wie oben, so folgt

$$\int_M \left( \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dV = \int_{\gamma} \langle F, d\mathbf{x} \rangle .$$

**Bemerkungen.**

- i) Wie in Kapitel 15 gezeigt ist, hängt das Kurvenintegral auf der rechten Seite nicht von der speziellen Wahl der Parametrisierung ab.
- ii) Der Gaußsche Integralsatz in der Ebene gilt beispielsweise auch für sogenannte **Greensche Bereiche**, das sind endliche Vereinigungen von Normalbereichen wie oben.

Dabei ist **auf die Orientierung der einzelnen Randkurven zu achten** (vgl. Abbildung 17.2).

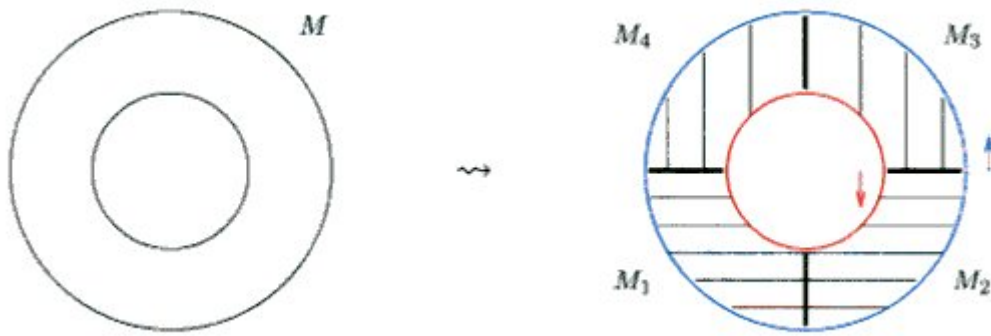


Abbildung 17.2: Ein Greenscher Bereich.

**Beispiele.**

i) In Satz 17.1 werde speziell

$$F(\underline{x}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

gewählt. Dann ist

$$\int_M \left( \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dV = 2 \int_M dV = 2A(M),$$

wobei  $A(M)$  den Flächeninhalt von  $M$  bezeichne.

Der Flächeninhalt  $A(M)$  kann also mittels

$$A(M) = \frac{1}{2} \int_{\gamma} \langle F, d\underline{x} \rangle, \quad F = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix},$$

berechnet werden.

ii) Ist speziell  $M$  die Ellipse

$$M = \left\{ \underline{x} \in \mathbb{R}^2 : \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1 \right\}, \quad 0 < a, b \in \mathbb{R},$$

so ist

$$\gamma(t) := \begin{pmatrix} a \cos(t) \\ b \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi],$$

eine Parametrisierung von  $\partial M$  wie in Satz 17.1 gefordert. Es folgt

$$\begin{aligned} A(M) &= \frac{1}{2} \int_0^{2\pi} \left\langle \begin{pmatrix} -b \sin(t) \\ a \cos(t) \end{pmatrix}, \begin{pmatrix} -a \sin(t) \\ b \cos(t) \end{pmatrix} \right\rangle dt \\ &= \frac{1}{2} \int_0^{2\pi} ab \, dt = \pi ab. \end{aligned}$$

### Interpretation als Fluss.

Es existiere eine reguläre  $C^1$ -Parametrisierung  $\gamma(t)$ ,  $t \in [a, b]$ , des positiv orientierten Randes  $\partial M$ . Dann gilt

$$\begin{pmatrix} \gamma'_2(t) \\ -\gamma'_1(t) \end{pmatrix} \perp \gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \gamma'_2(t) \end{pmatrix}$$

und man bezeichnet

$$\underline{\mathbf{n}}(t) = \frac{1}{\|\gamma'(t)\|} \begin{pmatrix} \gamma'_2(t) \\ -\gamma'_1(t) \end{pmatrix}$$

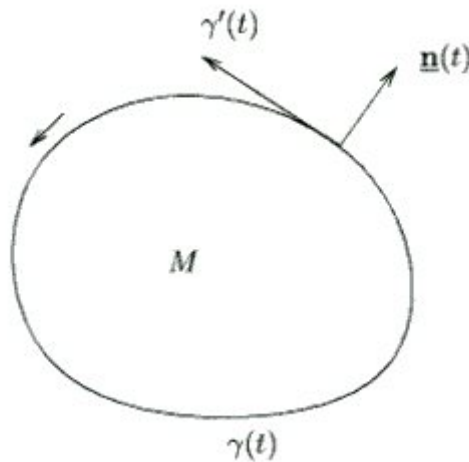


Abbildung 17.3: Die äußere Einheitsnormale an  $\partial M$ .

als die **äußere Einheitsnormale an  $\partial M$**  (zum Parameterwert  $t$ )

Ist  $F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$  ein stetig differenzierbares Vektorfeld und setzt man  $\tilde{F} = \begin{pmatrix} -F_2 \\ F_1 \end{pmatrix}$ , so folgt aus dem Gaußschen Satz für  $\tilde{F}$ :

$$\begin{aligned}
\int_M \operatorname{div} F \, dV &= \int_M \left( \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} \right) dV = \int_M \left( \frac{\partial \tilde{F}_2}{\partial x_1} - \frac{\partial \tilde{F}_1}{\partial x_2} \right) dV \\
&= \int_a^b \langle \tilde{F}(\gamma(t)), \gamma'(t) \rangle dt \\
&= \int_a^b \left( -F_2(\gamma(t))\gamma'_1(t) + F_1(\gamma(t))\gamma'_2(t) \right) dt \\
&= \int_a^b \langle F(\gamma(t)), \underline{\mathbf{n}}(t) \rangle \|\gamma'(t)\| dt .
\end{aligned}$$

### Bemerkungen.

- i)* Das Integral auf der rechten Seite heißt auch der **Fluss des Vektorfeldes  $F$  durch  $\partial M$** . Man beachte, dass hier nur die Komponente von  $F$  senkrecht zu  $\partial M$  zu berücksichtigen ist.
- ii)* In dieser Form lässt sich Satz 17.1 auf den  $\mathbb{R}^3$  verallgemeinern.

# Kapitel 18

## Elementares zur Wahrscheinlichkeitsrechnung und Statistik

### Zur Begriffsbildung.

- i) **Wahrscheinlichkeitstheorie/rechnung**: Studium von Gesetzmäßigkeiten zufälliger Ereignisse.
- ii) **Statistik**: Daten zufälliger Ereignisse werden gesammelt und in Bezug auf bestimmte Aspekte ausgewertet.

Dabei wird zwischen der **beschreibenden (deskriptiven) Statistik** (z.B. graphische oder tabellarische Darstellung relevanter Aspekte, wobei die Art der Darstellung durchaus schon ein Urteil beeinflussen kann) und der **beurteilenden (schließenden) Statistik** (aus statistischem Stichprobenmaterial werden Schlüsse auf die Grundgesamtheit gezogen, Test- und Schätzverfahren werden entwickelt) unterschieden.

- iii) **Stochastik**: Sammelbegriff aus Wahrscheinlichkeitstheorie und Statistik.

### 18.1 Grundbausteine der Kombinatorik

Die **Kombinatorik** ist die Lehre des Abzählens und als solche ein wichtiger Bestandteil der elementaren Wahrscheinlichkeitsrechnung und Statistik.

Das wesentliche Beweisprinzip für die Aussagen dieser Sektion ist die in Kapitel 2.1 vorgestellte vollständige Induktion.

### Der binomische Lehrsatz.

Von allgemeinem Interesse und zugleich wichtige Hilfsmittel an dieser Stelle sind der binomische Lehrsatz und damit verbunden die sogenannten [Binomialkoeffizienten](#):

**Satz 18.1.** (*Binomischer Lehrsatz*)

Für  $a, b \in \mathbb{R}$  und für alle  $n \in \mathbb{N}$  gilt

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k .$$

Dabei sind die Binomialkoeffizienten  $\binom{m}{l}$  definiert als

$$\binom{m}{l} := \frac{m!}{l!(m-l)!} = \frac{m(m-1)\dots(m-l+1)}{l!} ,$$

wobei  $m, l \in \mathbb{N}_0$ ,  $l \leq m$ ,  $\binom{m}{l} := 0$ , falls  $m < l$ ,  $0! = 1$ .

**Übung.** Rechnen mit Binomialkoeffizienten. Was ist das [Pascalsche Dreieck](#)?

Im Folgenden sei  $n \in \mathbb{N}$  und  $M$  eine Menge bestehend aus  $n$  Elementen, wobei in der Regel [ohne Einschränkung](#)  $M = \{1, \dots, n\}$  angenommen wird.

### Permutationen.

Zu  $k \in \mathbb{N}$  wird ein  $k$ -Tupel von Elementen aus  $M$  betrachtet, d.h.

$$(a_1, a_2, \dots, a_k) \in \underbrace{M \times M \times \dots \times M}_{k\text{-mal}} .$$

i) **Permutationen ohne Wiederholung.**

Sind in einem solchen Tupel alle Eintragungen verschieden (insbesondere muss dann  $k \leq n$  gelten), so spricht man von einer  **$k$ -Permutation aus  $M$  ohne Wiederholung**, die Menge aller derartigen Permutationen wird mit

$$\text{Per}_k^n(oW) := \{(a_1, \dots, a_k) : a_j \in \{1, \dots, n\}, 1 \leq j \leq k, \\ a_i \neq a_j \text{ für } 1 \leq i \neq j \leq k\}$$

bezeichnet.

Insgesamt gibt es

$$|\text{Per}_k^n(oW)| = n(n-1)(n-2) \dots (n-k+1)$$

verschiedene  $k$ -Permutationen aus  $M$  ohne Wiederholung.

**Beispiel Urnenmodell.** In einer Urne befinden sich  $n$  nummerierte, ansonsten gleichartige Kugeln. Aus der Urne werden  $k$  Kugeln gezogen.

Kommt es auf die **Reihenfolge** der Ziehungen an (wie bei den Eintragungen in einem Tupel) und werden die gezogenen Kugeln nicht wieder in die Urne zurückgelegt (wie in einem Tupel **ohne Wiederholung**), so gibt es

$$n(n-1) \dots (n-k+1) \text{ verschiedene Ziehungsmöglichkeiten.}$$

ii) **Permutationen mit Wiederholung.**

Müssen die Eintragungen in einem Tupel nicht verschieden sein, so spricht man von einer  **$k$ -Permutation aus  $M$  mit Wiederholung**, die Menge aller derartigen Permutationen wird mit

$$\text{Per}_k^n(mW) := \{(a_1, \dots, a_k) : a_j \in \{1, \dots, n\}, 1 \leq j \leq k\}$$

bezeichnet.

Insgesamt gibt es

$$|\text{Per}_k^n(mW)| = n^k$$

verschiedene  $k$ -Permutationen aus  $M$  mit Wiederholung.

**Beispiel Urnenmodell.** Kommt es wieder auf die Reihenfolge der Ziehungen an und werden die gezogenen Kugeln wieder in die Urne zurückgelegt (wie in einem Tupel mit Wiederholung), so gibt es

$n^k$  verschiedene Ziehungsmöglichkeiten.

### Kombinationen.

Kommt es nicht auf die Reihenfolge an, so können – wie bei den Lottozahlen – die Elemente eines Tupels einfach der Größe nach angeordnet werden,

$$(a_1, a_2, \dots, a_k) \in \underbrace{M \times M \times \dots \times M}_{k\text{-mal}}, \quad 1 \leq a_1 \leq a_2 \leq \dots \leq a_k \leq n.$$

i) **Kombinationen ohne Wiederholung.**

Sind in einem solchen der Größe nach geordneten Tupel alle Eintragungen verschieden (wieder muss in diesem Fall  $k \leq n$  gelten), so spricht man von einer  $k$ -Kombination aus  $M$  ohne Wiederholung, die Menge aller derartigen Kombinationen wird mit

$$\text{Kom}_k^n(oW) := \{(a_1, \dots, a_k) : a_j \in \{1, \dots, n\}, 1 \leq j \leq k, \\ 1 \leq a_1 < a_2 < \dots < a_k \leq n\}$$

bezeichnet.

Man beachte die strikte Ungleichungskette in der Definition, die die Bedingung “ohne Wiederholung” widerspiegelt.

**Bemerkung.** Eine  $k$ -Kombination aus  $M$  ohne Wiederholung entspricht genau einer  $k$ -elementigen Teilmenge von  $M$ .



Insgesamt gibt es

$$|\text{Kom}_k^n(oW)| = \binom{n}{k}$$

verschiedene  $k$ -Kombinationen aus  $M$  ohne Wiederholung.

**Beispiel Urnenmodell.** Kommt es **nicht auf die Reihenfolge** der Ziehungen an und werden die gezogenen Kugeln nicht wieder in die Urne zurückgelegt (**ohne Wiederholung**), so gibt es

$$\binom{n}{k} \text{ verschiedene Ziehungsmöglichkeiten.}$$

ii) **Kombinationen mit Wiederholung.**

Sind in einem solchen der Größe nach geordneten Tupel nicht notwendig alle Eintragungen verschieden, so spricht man von einer  **$k$ -Kombination aus  $M$  mit Wiederholung**, die Menge aller derartigen Kombinationen wird mit

$$\text{Kom}_k^n(mW) := \{(a_1, \dots, a_k) : a_j \in \{1, \dots, n\}, 1 \leq j \leq k, \\ 1 \leq a_1 \leq a_2 \leq \dots \leq a_k \leq n\}$$

bezeichnet.

Man beachte hier das “ $\leq$ ” in der Ungleichungskette, das die Bedingung “mit Wiederholung” widerspiegelt.

Insgesamt gibt es

$$|\text{Kom}_k^n(mW)| = \binom{n+k-1}{k}$$

verschiedene  $k$ -Kombinationen aus  $M$  mit Wiederholung.

**Beispiel Urnenmodell.** Kommt es **nicht auf die Reihenfolge** der Ziehungen an und werden die gezogenen Kugeln wieder in die Urne zurückgelegt (**mit Wiederholung**), so gibt es

$$\binom{n+k-1}{k} \text{ verschiedene Ziehungsmöglichkeiten.}$$

## 18.2 Wahrscheinlichkeitsrechnung

### Beispiele.

- i) Beim einmaligen Würfeln mit einem **fairen Würfel** gibt es insgesamt 6 Möglichkeiten und jedem der **Ereignisse**

$$\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$$

wird die **Wahrscheinlichkeit**  $1/6$  zugeordnet, nämlich bei diesem sogenannten **Laplace-Modell** die Wahrscheinlichkeit

$$P(A) = \frac{\text{Anzahl der für } A \text{ "günstigen" Fälle}}{\text{Anzahl aller möglichen Fälle}}$$

für das Eintreten des Ergebnisses  $A$ .

Das **Ereignis** "Wurf einer ungeraden Zahl" wird durch die Teilmenge

$$\{1, 3, 5\} \subset \{1, 2, 3, 4, 5, 6\} =: \Omega$$

repräsentiert und tritt nach dem oben Gesagten mit der Wahrscheinlichkeit  $1/2$  auf.

Die Menge  $\Omega$  fasst alle möglichen **Ergebnisse** eines **idealen Zufallsexperimentes** zusammen und heißt dementsprechend die **Ergebnismenge** oder die **Grundmenge** des Zufallsexperimentes.

Ein Ereignis  $A$  wiederum ist eine **Teilmenge des Grundraums** und **tritt ein**, wenn das Ergebnis eines Experimentes ein Element von  $A$  ist.

- ii) Bei einem **manipulierten Würfel** (trotzdem ideales Zufallsexperiment!) kann beispielsweise die Wahrscheinlichkeit, eine 6 zu würfeln, doppelt so groß sein wie die Wahrscheinlichkeit für die anderen Zahlen, d.h.

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = \frac{1}{7},$$

$$P(\{6\}) = \frac{2}{7}.$$

**Wahrscheinlichkeitsräume.**

**Definition 18.1.** Es sei  $\Omega$  eine endliche Menge  $\Omega = \{a_1, \dots, a_n\}$  oder eine abzählbar unendliche Menge  $\Omega = \{a_j : j \in \mathbb{N}\}$  und

$$\mathcal{P}(\Omega) = \{E \subset \Omega\}$$

die Menge aller Teilmengen von  $\Omega$ , genannt die Potenzmenge von  $\Omega$ . Eine diskrete Wahrscheinlichkeitsverteilung auf  $\Omega$  ist eine Funktion  $P: \mathcal{P}(\Omega) \rightarrow [0, 1]$  mit den Eigenschaften:

- i) Sind endlich viele oder abzählbar unendlich viele paarweise disjunkte Teilmengen  $E_j \subset \Omega$ ,  $E_j \cap E_i = \emptyset$  für  $i \neq j$ , gegeben, so gilt

$$P\left(\bigcup_i E_i\right) = \sum_i P(E_i).$$

- ii)  $P(\Omega) = 1$ .

Sprechweise:

- i)  $(\Omega, P)$  heißt diskreter Wahrscheinlichkeitsraum.  
 ii) Eine Teilmenge  $E \subset \Omega$  heißt Ereignis, ein einzelnes Element  $\omega \in \Omega$  heißt Elementarereignis.  
 iii) Für  $E \subset \Omega$  heißt  $P(E)$  die Wahrscheinlichkeit von  $E$ .

**Beispiel.** Die oben skizzierte Laplace-Verteilung als Modell für einen fairen Würfel ist eine Wahrscheinlichkeitsverteilung auf  $\{1, 2, 3, 4, 5, 6\}$ .

Elementar folgt aus Definition 18.1

**Satz 18.2.** Es sei  $(\Omega, P)$  ein diskreter Wahrscheinlichkeitsraum und  $E \subset \Omega$ . Dann gilt

i)  $P(E) = \sum_{\omega \in E} P(\{\omega\});$

- ii)  $P(\emptyset) = 0;$

$$iii) P(\Omega - E) = 1 - P(E).$$

**Beispiel.** Bei einem fairen Würfel ist

$$\frac{1}{2} = P(\{1, 3, 5\}) = P(\{1\}) + P(\{3\}) + P(\{5\}) = 1 - P(\{2, 4, 6\}).$$

**Beispiel.** Es sei  $\Omega = \{W, Z\}$  (Wappen oder Zahl) die Menge der Ergebnisse eines einmaligen Münzwurfes, d.h.

$$\Omega^n := \{(\omega_1, \omega_2, \dots, \omega_n) : \omega_i \in \{W, Z\}, i = 1, \dots, n\}$$

ist die Menge der möglichen Ergebnisse beim  $n$ -maligen Münzwurf.

Die Wahrscheinlichkeit bei einem Wurf Wappen zu werfen sei  $p \in (0, 1)$ , die Wahrscheinlichkeit Zahl zu werfen  $q = 1 - p \in (0, 1)$ .

Dann ist die Wahrscheinlichkeit bei  $n$  Würfeln in den ersten  $k$  Würfeln Wappen zu werfen und in den verbleibenden  $n - k$  Würfeln Zahl zu werfen

$$p^k q^{n-k}.$$

Die Wahrscheinlichkeit bei  $n$  Würfeln  $k$  mal Wappen und  $n - k$  mal Zahl zu werfen ist (Übung, vgl. Abschnitt 18.1)

$$\binom{n}{k} p^k q^{n-k}.$$

In dem Beispiel handelt es sich um einen Spezialfall von

**Definition 18.2.** Es sei  $\Omega = \{1, 2, \dots, n\}$ ,  $p \in (0, 1)$  und  $q = 1 - p$ . Dann ist die *Binomialverteilung mit den Parametern  $p$  und  $q$*  gegeben durch

$$B_{n,p}(k) := P(\{k\}) = \binom{n}{k} p^k q^{n-k}, \quad 0 \leq k \leq n.$$

**Übung.** Bei der Binomialverteilung handelt es sich um eine Wahrscheinlichkeitsverteilung im Sinne von Definition 18.1.

**Bemerkung.** Als Grenzwert der Binomialverteilung ( $n \rightarrow \infty$ ,  $np = \lambda = \text{konst}$ ) ergibt sich die sogenannte **Poissonverteilung** auf  $\mathbb{N}$ ,

$$P_\lambda(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \lambda > 0 \text{ fixiert.}$$

Schließlich ist noch zu definieren:

**Definition 18.3.** Es sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum und  $\emptyset \neq A, B \subset \Omega$  seien Ereignisse.

i) Die **bedingte Wahrscheinlichkeit** von  $A$  unter der Bedingung  $B$  ist

$$P(A|B) := \frac{P(A \cap B)}{P(B)}.$$

ii) Die Ereignisse  $A$  und  $B$  heißen **unabhängig**, wenn

$$P(A|B) = P(A)$$

gilt oder äquivalent ausgedrückt

$$P(A \cap B) = P(A)P(B).$$

## Zufallsvariable und deren Verteilungsfunktion.

**Definition 18.4.** Es sei  $(\Omega, P)$  ein diskreter Wahrscheinlichkeitsraum.

i) Eine **Funktion**  $X: \Omega \rightarrow \mathbb{R}$  heißt **Zufallsvariable** oder **diskrete Zufallsgröße**.

ii) Ist  $X: \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable, so heißt die Funktion  $F_X: \mathbb{R} \rightarrow [0, 1]$ ,

$$F_X(t) := P(\{\omega \in \Omega : X(\omega) \leq t\}) =: P(X \leq t)$$

die **Verteilung** oder **kumulative Verteilungsfunktion** der **Zufallsvariablen**  $X$ .

**Übung.** Die kumulative Verteilungsfunktion hat folgende Eigenschaften:

i)  $0 \leq F_X(t) \leq 1$  für alle  $t \in \mathbb{R}$ ;

ii) für  $a < b$  gilt

$$P(\{\omega \in \Omega : a < X(\omega) \leq b\}) =: P(a < X \leq b) = F_X(b) - F_X(a) ;$$

iii) die Verteilungsfunktion ist monoton wachsend;

iv)  $\lim_{t \rightarrow -\infty} F_X(t) = 0$ ,  $\lim_{t \rightarrow \infty} F_X(t) = 1$ .

**Beispiel.** Wie oben werde das  $n$ -malige Werfen einer Münze betrachtet, wobei wieder die Wahrscheinlichkeit für jedes einzelne Ereignis  $(\omega_1, \omega_2, \dots, \omega_n)$  mit genau  $k$  Wappen-Ergebnissen durch

$$P((\omega_1, \omega_2, \dots, \omega_n)) = p^k q^{n-k}$$

gegeben ist,  $p \in (0, 1)$ ,  $q = 1 - p$ .

Als Zufallsvariable betrachte man die Abbildung  $X$ , die einem Wurf mit  $k$  Treffern ( $k$  mal Wappen) den Wert  $k$  zuordnet.

Es ist bereits gezeigt, dass genau  $\binom{n}{k}$  Tupel mit  $k$  Treffern existieren, d.h.

$$\begin{aligned} P(X = k) &:= P(\{(\omega_1, \dots, \omega_n) \in \{W, T\}^n : X((\omega_1, \dots, \omega_n)) = k\}) \\ &= \binom{n}{k} p^k q^{n-k} = B_{n,p}(k) \end{aligned}$$

und

$$F_X(t) = P(X \leq t) = \sum_{k=0}^t \binom{n}{k} p^k q^{n-k} .$$

Die Zufallsvariable  $X$  ist binomial verteilt.

## Erwartungswert, Varianz und Standardabweichung einer Zufallsvariablen.

**Beispiel.** Im Falle einer endlichen Wertemenge und der Gleichverteilung einer Zufallsvariablen, d.h.  $X(\Omega) = \{x_1, x_2, \dots, x_n\}$  mit

$$P(X = x_i) = \frac{1}{n}, \quad 1 \leq i \leq n,$$

sollte der Erwartungswert mit dem arithmetischen Mittel übereinstimmen,

$$E(X) = \sum_{i=1}^n \frac{x_i}{n} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Allgemein setzt man

**Definition 18.5.** Es sei  $(\Omega, P)$  ein diskreter Wahrscheinlichkeitsraum und  $X: \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable mit Wertemenge  $X(\Omega) = \{x_1, x_2, \dots, x_n\}$  oder  $X(\Omega) = \{x_1, x_2, \dots\}$ .

i) Ist  $X(\Omega) = \{x_1, x_2, \dots, x_n\}$  oder ist die Reihe  $\sum_{j=1}^{\infty} x_j P(X = x_j)$  absolut konvergent, so heißt

$$E(X) := \sum_{x_j \in X(\Omega)} x_j P(X = x_j)$$

der Erwartungswert von  $X$ .

ii) Existiert der Erwartungswert  $E(X) = \mu$  und existiert der Erwartungswert für die Zufallsvariable

$$Y := (X - \mu)^2,$$

d.h. für die quadratische Abweichung vom Erwartungswert, so heißt

$$\sigma^2 := V(X) := E(Y)$$

die Varianz von  $X$ .

iii) Die Größe

$$\sigma = \sqrt{E(Y)}$$

heißt die Standardabweichung von  $X$ .

**Übung: Rechenregeln.** Es gilt

$$i) V(X) = E(X^2) - (E(X))^2;$$

$$ii) V(cX + d) = c^2V(X) \text{ für alle } c, d \in \mathbb{R};$$

iii) die normierte Zufallsgröße

$$\tilde{X} := \frac{1}{\sigma(X)}(X - \mu(X))$$

hat den Erwartungswert 0 und die Varianz 1.

**Beispiel.** Betrachtet sei die binomial verteilte Zufallsvariable  $X$  wie im obigen Beispiel des Münzwurfes. Dann gilt für den Erwartungswert per definitionem

$$E(X) = \sum_{k=0}^n kP(X = k) = \sum_{k=0}^n k \binom{n}{k} p^k q^{n-k},$$

und man erhält wegen (binomischer Lehrsatz)

$$\sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{(n-1)-j} = \underbrace{(p+q)^{n-1}}_{=1} = 1$$

schließlich

$$\begin{aligned} E(X) &= \sum_{k=0}^n k \frac{n!}{k!(n-k)!} p^k q^{n-k} \\ &= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p^{k-1} q^{(n-1)-(k-1)} \\ &= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{(n-1)-j} \\ &= np, \end{aligned}$$

d.h.: Wie zu erwarten (warum?) ist der Erwartungswert obiger binomial verteilter Zufallsvariablen  $X$  genau  $np$ .



## Normalverteilung und zentraler Grenzwertsatz.

**Beispielproblem.** Die Zufallsvariable  $X$  soll **ohne Kenntnis ihrer Verteilungsfunktion** anhand eines **Zufallsexperimentes** analysiert werden. Inwieweit hilft hier eine **häufige Wiederholung des gleichen Zufallsexperimentes**?

Zur Beantwortung dieser Frage benötigt man zunächst die **Gaußsche Glockenkurve** oder **Dichte der standardisierten Normalverteilung** (vgl. Abbildung 18.1)

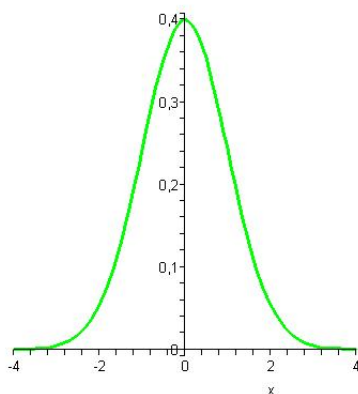


Abbildung 18.1: Die Gaußsche Glockenkurve.

$$\varphi(x) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x \in \mathbb{R},$$

wobei man die Eigenschaft “Verteilung” anhand von

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx = 1$$

(Beweis: Literatur) einsehen kann.

Die Wiederholung eines Zufallsexperimentes sei nun mit der Zufallsvariablen  $X_i$  beschrieben, wobei die  $X_i$  alle **unabhängig voneinander seien**, d.h. (vgl. Definition 18.3, ii)) für alle  $i, j \in \mathbb{N}$  und für alle  $y_i, y_j \in \mathbb{R}$  gelte (( $\Omega, P$ ) Wahrscheinlichkeitsraum,  $F_{X_i}$  Verteilung der Zufallsvariablen  $X_i$ )

$$P(\{\omega \in \Omega : X_i(\omega) \leq y_i, X_j(\omega) \leq y_j\}) = F_{X_i}(y_i)F_{X_j}(y_j).$$

Zu  $n \in \mathbb{N}$  bezeichne  $Y_n$  ferner den Mittelwert über  $n$ -Experimente,

$$Y_n := \frac{1}{n} \sum_{i=1}^n X_i .$$

**Satz 18.3.** (Zentraler Grenzwertsatz)

Es sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum und  $(X_i)_{i \in \mathbb{N}}$  eine Folge von unabhängigen Zufallsvariablen auf  $\Omega$ .

Der Erwartungswert  $\mu$  und die Varianz  $\sigma^2$  aller Zufallsvariablen sei gleich.

Weiterhin sei  $\tilde{Y}_n$  die *normierte Variable*

$$\tilde{Y}_n := \sqrt{n} \left( \frac{Y_n - \mu}{\sigma} \right) .$$

Dann gilt:

i) Der Erwartungswert des Mittelwertes ändert sich nicht, d.h.

$$E(Y_n) = \mu .$$

ii) Je größer  $n$  wird, desto geringer wird die Streuung des Mittelwertes im Sinne von

$$V(Y_n) = \left( \frac{\sigma}{\sqrt{n}} \right)^2 .$$

iii) Unabhängig von der Verteilungsfunktion für die Zufallsvariablen  $X_i$  gilt: *Der Grenzwert der Verteilungsfunktion für die normierte Variable  $\tilde{Y}_n$  ist die Normalverteilung, d.h.*

$$\lim_{n \rightarrow \infty} F_{\tilde{Y}_n}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{x^2}{2}\right) dx .$$

**Einige Symbole.**

Zahlenmengen	$\mathbb{N}$	Menge der natürlichen Zahlen
	$\mathbb{N}_0$	Menge der natürlichen Zahlen mit der 0
	$\mathbb{Z}$	Menge der ganzen Zahlen
	$\mathbb{Q}$	Menge der rationalen Zahlen
	$\mathbb{R}$	Menge der reellen Zahlen
	$\mathbb{R}^+$	Menge der positiven reellen Zahlen
	$\mathbb{R}_0^+$	Menge der positiven reellen Zahlen mit der 0
	$\mathbb{C}$	Menge der komplexen Zahlen
logische Operationen	$\wedge$	und (Konjunktion zweier Aussagen)
	$\vee$	oder (nicht-ausschließend, Disjunktion zweier Aussagen)
	$\Rightarrow$	Folgerung (Implikation)
	$\Leftrightarrow$	Äquivalenz
	$:\Leftrightarrow$	Seite des Doppelpunktes: per definitionem äquivalent
Quantoren	$\forall$	für alle
	$\exists$	es existiert
	$\exists_1$	es existiert genau ein (auch $\exists!$ )
Mengen	$\subset$	Teilmenge
	$\not\subset$	nicht Teilmenge
	$\cap$	Durchschnitt
	$\cup$	Vereinigung
	$-$	Differenz (auch $\setminus$ )
	$\emptyset$	leere Menge
	$\in$	Element von
	$\notin$	nicht Element von
	$:=$	Seite des Doppelpunktes wird definiert als
	$\sum$	Summenzeichen
	$\prod$	Produktzeichen
	$\pm$	plusminus
	$\mp$	minusplus
Intervalle	$[a, b]$	abgeschlossen
	$(a, b)$	offen
	$(a, b]$	halboffen
	$[a, b)$	halboffen

Funktionen	$\circ$ $f^{-1}$ $f \equiv$ $\text{grad } p$	Verkettung Umkehrfunktion $f$ ist identisch rechte Seite (für alle $x$ ) Grad eines Polynoms
Konstanten	$e$ $\pi$	Eulersche Zahl $e = 2.71828\dots$ Kreiszahl $\pi = 3.14159\dots$
komplexe Zahlen	$i$ $\bar{z}$	imaginäre Einheit konjugiert komplexe Zahl
Grenzwerte	$\lim_{n \rightarrow \infty}$ $\limsup_{n \rightarrow \infty}$ $\liminf_{n \rightarrow \infty}$ $\lim_{x \rightarrow x_0}$	Grenzwert (Limes) Limes superior Limes inferior Grenzwert einer Funktion
Ableitungen	$f'$ $f''$ $f^{(n)}$	erste Ableitung (auch $\frac{d}{dx}$ ) zweite Ableitung (auch $\frac{d}{dx^2}$ ) $n$ -te Ableitung
Integralrechnung	$\mathcal{Z}$ $\bar{\mathcal{I}}$ $\underline{\mathcal{I}}$ $\int_a^b f(x) dx$ $\int f(x) dx$ $\mathcal{R}(I)$	Zerlegung eines Intervalls Obersumme Untersumme bestimmtes Integral ( $\mathcal{I}(f)$ ) unbestimmtes Integral Klasse der Riemann integrierbaren Funktionen
Vektorrechnung	$\underline{\mathbf{x}}$	Vektor
sonstiges	$n!$ $\square$	Fakultät ( $1 \cdot 2 \cdot \dots \cdot n$ ) Beweisende (q.e.d: quod erat demonstrandum)