

Skript zur Vorlesung

# Mathematik B

Sommersemester 2021

Christian Wyss, 23. Mai 2022

initial in L<sup>A</sup>T<sub>E</sub>X gesetzt von Jannis Dohm, basierend auf dem  
Vorlesungsskript vom Sommersemester 2016  
vollständig überarbeitet zur Vorlesung im Sommersemester 2020



# Inhaltsverzeichnis

<b>1</b>	<b>Rekursion, Induktion, binomischer Satz</b>	<b>7</b>
1.1	Rekursion . . . . .	7
1.2	Summen und Produkte . . . . .	10
1.3	Permutationen . . . . .	11
1.4	Prinzip der vollständigen Induktion . . . . .	12
1.5	Binomialkoeffizienten . . . . .	14
1.6	Der allgemeine binomische Satz . . . . .	16
<b>2</b>	<b>Folgen und Grenzwerte</b>	<b>19</b>
2.1	Zahlenfolgen . . . . .	19
2.2	Grenzwerte . . . . .	21
2.3	Exakte Definition des Grenzwertes . . . . .	22
2.4	Eigenschaften konvergenter Folgen . . . . .	24
2.5	Rechenregeln für Grenzwerte . . . . .	27
2.6	Teilfolgen und Häufungswerte . . . . .	28
2.7	Bestimmte Divergenz . . . . .	30
2.8	Grenzwerte und Ungleichungen . . . . .	33
2.9	Grenzwerte und stetige Funktionen . . . . .	34
2.10	Die geometrische Folge . . . . .	35
2.11	Konvergenzkriterien . . . . .	37
2.12	Sandwich-Lemma . . . . .	37
2.13	Die Folgen $\sqrt[n]{n}$ und $\sqrt[n]{p}$ . . . . .	39
2.14	Monotoniekriterium . . . . .	42
2.15	Cauchy-kriterium . . . . .	43
<b>3</b>	<b>Reihen</b>	<b>45</b>
3.1	Unendliche Reihen . . . . .	45
3.2	Die geometrische Reihe . . . . .	47
3.3	Rechenregeln für konvergente Reihen . . . . .	48
3.4	Konvergenzkriterien für Reihen . . . . .	49
3.5	Reihen mit positiven Gliedern . . . . .	51
3.6	Absolute Konvergenz . . . . .	54
3.7	Majoranten- und Minorantenkriterium . . . . .	55
3.8	Quotienten- und Wurzelkriterium I . . . . .	57
3.9	Quotienten- und Wurzelkriterium II . . . . .	58
3.10	Integralkriterium . . . . .	62
3.11	Alternierende Reihen . . . . .	63
3.12	Umordnung von Reihen . . . . .	65

3.13	Die Cauchyproduktformel . . . . .	66
3.14	Die Exponentialreihe . . . . .	67
<b>4</b>	<b>Taylorentwicklung und Potenzreihen</b>	<b>69</b>
4.1	Taylorpolynome . . . . .	69
4.2	Approximationsfehler und Restglied . . . . .	71
4.3	Restgliedabschätzung . . . . .	73
4.4	Taylorreihen . . . . .	75
4.5	Potenzreihen . . . . .	78
4.6	Ableitung und Integration von Potenzreihen . . . . .	82
4.7	Taylorentwicklung des Logarithmus . . . . .	85
4.8	Grenzwertberechnung mit Potenzreihen . . . . .	87
4.9	Weitere Taylorentwicklungen . . . . .	88
4.10	Eine Funktion ohne Taylorentwicklung . . . . .	88
<b>5</b>	<b>Elementare Differentialgleichungen</b>	<b>91</b>
5.1	Einführung . . . . .	91
5.2	Allgemeine und spezielle Lösungen, Anfangswerte . . . . .	92
5.3	Differentialgleichungen in Anwendungen . . . . .	94
5.4	Dgl 1. Ordnung mit getrennten Variablen . . . . .	95
5.5	Direkte Lösung des AWP bei getrennten Variablen . . . . .	99
5.6	Lineare Differentialgleichungen 1. Ordnung . . . . .	100
5.7	Inhomogene lineare Differentialgleichungen 1. Ordnung . . . . .	102
5.8	Anwendung: RC-Reihenschaltung . . . . .	103
5.9	Das Richtungsfeld . . . . .	105
5.10	Lineare Differentialgleichungen 2. Ordnung . . . . .	107
5.11	Homogene Gleichungen 2. Ordnung . . . . .	109
5.12	Reelle und komplexe Fundamentalsysteme . . . . .	112
5.13	Inhomogene Gleichungen 2. Ordnung . . . . .	114
5.14	Lösung durch Übergang zum Komplexen . . . . .	117
5.15	Superposition partikulärer Lösungen . . . . .	118
5.16	Anwendung: gedämpfte Schwingungen . . . . .	119
	5.16.1 Freie gedämpfte Schwingungen . . . . .	120
	5.16.2 Erzwungene Schwingungen . . . . .	121
5.17	Variation der Konstanten für Dgl 2. Ordnung . . . . .	123
5.18	Zusammenfassung Lösungsmethoden . . . . .	124
5.19	Substitution . . . . .	124
5.20	Potenzreihenansatz . . . . .	126
<b>6</b>	<b>Eigenwerte und Eigenvektoren</b>	<b>129</b>
6.1	Einführung . . . . .	129
6.2	Das charakteristische Polynom einer Matrix . . . . .	131
6.3	Algebraische und geometrische Vielfachheiten . . . . .	134
6.4	Diagonalisierung . . . . .	138
6.5	Komplexe Eigenwerte und Eigenvektoren . . . . .	141

<b>7</b>	<b>Symmetrische Matrizen und Definitheit</b>	<b>143</b>
7.1	Symmetrische Matrizen . . . . .	143
7.2	Orthogonale Diagonalisierung . . . . .	145
7.3	Definitheit symmetrischer Matrizen . . . . .	147
7.4	Drehungen in der Ebene . . . . .	149
7.5	Drehungen im Raum . . . . .	151
<b>8</b>	<b>Differentialrechnung mehrerer Variablen</b>	<b>155</b>
8.1	Mengen in mehreren Raumdimensionen . . . . .	155
8.2	Funktionen mehrerer Variablen . . . . .	158
8.3	Grenzwerte und Stetigkeit . . . . .	160
8.4	Partielle Ableitung . . . . .	161
8.5	Ableitungen höherer Ordnung . . . . .	162
8.6	Totale Differenzierbarkeit und Gradient . . . . .	163
8.7	Anwendung: Fehlerrechnung . . . . .	167
8.8	Die Tangentialebene . . . . .	169
8.9	Richtungsableitung . . . . .	170
8.10	Geometrische Interpretation des Gradienten . . . . .	172
8.11	Niveaulinien . . . . .	173
8.12	Kurven im Raum . . . . .	175
8.13	Taylorentwicklung für mehrere Variablen . . . . .	177
8.14	Die Hessematrix . . . . .	180
8.15	Lokale Extrema, kritische Punkte . . . . .	182
8.16	Hinreichendes Kriterium für Extrema . . . . .	184
8.17	Extrema unter Nebenbedingungen . . . . .	186
8.18	Globale Extrema unter Nebenbedingungen . . . . .	189
8.19	Ableitung vektorwertiger Funktionen . . . . .	190
8.20	Anwendung: Ableitung in Polarkoordinaten . . . . .	193
<b>9</b>	<b>Mehrdimensionale Integration</b>	<b>197</b>
9.1	Zweidimensionale Integrale über Rechtecke . . . . .	197
9.2	Integrale über höherdimensionale Quader . . . . .	201
9.3	Integration über allgemeine Mengen . . . . .	202
9.4	Eigenschaften des Integrals . . . . .	203
9.5	Integration über Normalbereiche . . . . .	204
9.6	Vertauschen der Integrationsreihenfolge . . . . .	210
9.7	Dreidimensionale Normalbereiche . . . . .	213
9.8	Eine Anwendung des Integrals: Dichten . . . . .	215
9.9	Der Transformationssatz . . . . .	216
9.10	Transformationssatz für Polarkoordinaten . . . . .	220
9.11	Zylinderkoordinaten . . . . .	222
9.12	Rotationskörper . . . . .	223
9.13	Kugelkoordinaten . . . . .	224



# Kapitel 1

## Rekursion, Induktion, binomischer Satz

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik A, Kapitel 1.3, 1.4.

### 1.1 Rekursion

Wir beginnen in diesem Kapitel mit der Untersuchung von unendlichen Folgen von Zahlen, kurz *Folgen*. „Unendliche Folge“ bedeutet hier, dass es eine Folge von *unendlich vielen* Zahlen ist. Wie kann eine solche Folge von Zahlen eindeutig durch eine mathematische Formel beschrieben werden? Man kann ja schlecht alle – also unendlich viele – Zahlen aufschreiben...

Betrachten wir als erstes Beispiel diese Folge:

$$1, 2, 4, 8, 16, 32, \dots$$

Durch die Punkte ist angedeutet, dass die Folge „immer weiter geht“, also kein Ende hat; und hier ist auch klar, wie die nächsten Zahlen aussehen: 64, 128, ... Aber wie kann man diese Folge nun mathematisch exakt, und vorallem kurz beschreiben? Wir haben zwei prinzipielle Möglichkeiten:

#### 1. Zweierpotenzen

$$a_n = 2^n, \quad n \in \mathbb{N}.$$

(Erinnerung:  $\mathbb{N} = \{0, 1, 2, \dots\}$  bezeichnet die Menge der natürlichen Zahlen *mit* Null.)

Mit dieser Formel erhalten wir offenbar unsere Folge, denn

$$a_0 = 2^0 = 1, \quad a_1 = 2^1 = 2, \quad a_2 = 2^2 = 4, \quad a_3 = 2^3 = 8, \dots$$

**Beachte:** Die erste Zahl der Folge wird mit  $a_0$  bezeichnet. Die Folge beginnt also eigentlich mit der „nullten“ Zahl, was in der Mathematik aber nicht ungewöhnlich ist. Hier ergibt es sich ganz automatisch, weil unsere Folge mit  $1 = 2^0$  anfängt, also mit  $2^n$  mit  $n = 0$ .

## 2. „mal 2 rechnen“

$$\begin{aligned} a_0 &= 1 \\ a_{n+1} &= 2a_n \quad \text{für jedes } n \in \mathbb{N} \end{aligned}$$

Hier geben wir also zuerst das erste (bzw. nullte) Folgenglied explizit an ( $a_0 = 1$ ), und sagen dann, wie man von einem Folgenglied  $a_n$  zum nächsten Folgenglied  $a_{n+1}$  kommt: man rechnet „mal zwei“, also  $a_{n+1} = 2 \cdot a_n = 2a_n$ . Hiermit erhält man dann tatsächlich nacheinander alle Zahlen aus der Folge:

$$\begin{aligned} a_1 &= a_{0+1} = 2a_0 = 2 \cdot 1 = 2, \\ a_2 &= a_{1+1} = 2a_1 = 2 \cdot 2 = 4, \dots \end{aligned}$$

Diese beiden Möglichkeiten zur Definition einer Folge haben spezielle Namen:

- *explizite Definition:*  $a_n = 2^n$ ,  $n \in \mathbb{N}$
- *rekursive Definition:*

$$\begin{array}{ll} a_0 = 1 & \text{Startwert} \\ a_{n+1} = 2a_n, \quad n \in \mathbb{N} & \text{Rekursionsschritt} \end{array}$$

Der Name „explizite Definition“ deutet an, dass man für jedes  $n$  Zahl  $a_n$  der Folge direkt, in einem Schritt ausrechnen kann. „Rekursion“ bzw. „rekursiv“ bedeutet, aus einer Zahl mit immer derselben Formel nach und nach die folgenden Zahlen auszurechnen. Bei der rekursiven Definition sagt man statt „Rekursionsschritt“ auch kurz „Rekursion“ oder „Schritt von  $n$  nach  $n + 1$ “.

**Beachte:** Bei einer rekursiven Definition sind *beide* Teile wichtig, der Startwert und der Rekursionsschritt. Ohne den Startwert weiß man nicht, mit welcher Zahl man im ersten Rekursionsschritt beginnen soll. Und mit einem anderen Startwert erhält man offenbar ganz andere Zahlen in der Folge!

Wir betrachten ein paar weitere Beispiele für Folgen oder mathematische Ausdrücke, die man rekursiv definieren kann.

**Beispiel 1.1.1** (a) Rekursive Definition von Potenzen  $x^n$ 

$$\begin{array}{ll} x^0 = 1 & \text{Startwert} \\ x^{n+1} = x^n \cdot x & \text{Rekursion} \end{array}$$

(Beachte: für jede reelle Zahl  $x \in \mathbb{R}$  gilt  $x^0 = 1$ )

Eine andere mögliche Definition von  $x^n$  wäre

$$x^n = \underbrace{x \cdot x \cdot \dots \cdot x}_{n\text{-mal}}$$

Hierbei ist aber nicht wirklich klar, was die Formel für  $n = 0$  bedeuten soll.

(b) Fakultät  $n!$ 

Die *Fakultät* einer Zahl  $n$  ist definiert durch

$$n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n \quad \text{für } n \in \mathbb{N}^*. \quad (1.1)$$

Hier ist  $\mathbb{N}^* = \{1, 2, 3, \dots\}$  die Menge der natürlichen Zahlen *ohne* Null, also  $n \geq 1$ .

Die rekursive Definition der Fakultät ist

$$\begin{aligned} 1! &= 1 && \text{(Start)} \\ (n+1)! &= n! \cdot (n+1) && \text{(Rekursion)} \end{aligned}$$

Außerdem definiert man die Fakultät auch noch für  $n = 0$ , nämlich

$$0! = 1.$$

Das ergibt dann folgende erste Werte für die Fakultät:

$n$	0	1	2	3	4	5	6
$n!$	1	1	2	6	24	120	720

Die Tabellenwerte kann mit der Rekursionsformel einfach berechnen, z.B.

$$3! = (2+1)! = 2! \cdot (2+1) = 2! \cdot 3 = 2 \cdot 3 = 6,$$

$$4! = 3! \cdot 4 = 6 \cdot 4 = 24,$$

u.s.w.

Man sieht auch, dass die Werte von  $n!$  sehr schnell wachsen.

Beachten Sie, dass man den Wert  $0! = 1$  *nicht* aus der Formel (1.1) erhalten kann, es ist eine zusätzliche Definition! In die rekursive Definition kann man  $0! = 1$  jedoch einbauen, wenn man einfach  $0! = 1$  als Startwert nimmt, denn die Rekursionsformel funktioniert auch für  $n = 0$ :

$$1! = (0+1)! = 0! \cdot 1 = 1 \cdot 1 = 1$$

(sieht man auch in der Tabelle)

## (c) Fibonacci-Zahlen

Die *Fibonacci-Zahlen* sind rekursiv definiert durch

$$\begin{aligned} a_0 &= a_1 = 1 && \text{(Start)} \\ a_{n+1} &= a_n + a_{n-1} && \text{(Rekursion)} \end{aligned}$$

Der Rekursionsstart besteht hier also aus zwei Startwerten, der Rekursionsschritt benutzt auch den vorletzten Wert  $a_{n-1}$ . Die ersten Fibonacci-Zahlen sind

$n$	0	1	2	3	4	5	6	7	8
$a_n$	1	1	2	3	5	8	13	21	34

Beachten Sie wieder, wie sich die Tabelle aus der rekursiven Definition ergibt: für  $n = 0$  und  $n = 1$  hat man jeweils den Startwert 1, und die folgenden Werte berechnet man aus der Rekursionsformel  $a_{n+1} = a_n + a_{n-1}$ :  $1 + 1 = 2$ ,  $1 + 2 = 3$ ,  $2 + 3 = 5$ , u.s.w.

## 1.2 Summen und Produkte

Zur Vereinfachung der Schreibweise gibt es das Summen- und Produktzeichen:

$$\sum_{k=m}^n a_k = a_m + a_{m+1} + a_{m+2} + \dots + a_n \quad (n \geq m) \quad (1.2)$$

$$\prod_{k=m}^n a_k = a_m \cdot a_{m+1} \cdot a_{m+2} \cdot \dots \cdot a_n \quad (1.3)$$

Es steht für die Summe bzw. das Produkt der Zahlen  $a_m$  bis  $a_n$ . Dabei ist das  $a_k$  hinter dem Summen- und Produktzeichen  $\sum$  und  $\prod$  meistens eine Formel, in der die Variable  $k$  vorkommt. Und über und unter dem Summen- oder Produktzeichen steht dann, im welchem Bereich  $k$  „läuft“. In (1.2) und (1.3) ist das von  $k = m$  bis  $k = n$ , d.h. man setzt in Summe bzw. Produkt nacheinander für  $k$  die Werte  $m, m + 1, m + 2, \dots$  bis  $n$  ein.

**Beispiel 1.2.1** (a) Wir schreiben die Summe der ersten Quadratzahlen von  $1^2$  bis  $n^2$  mit dem Summenzeichen:

$$\sum_{k=1}^n k^2 = 1^2 + 2^2 + 3^2 + \dots + n^2$$

(b) Mit dem Produktzeichen kann man die Fakultät  $n!$ , also das Produkt der ersten  $n$  natürlichen Zahlen auch so schreiben:

$$\prod_{k=1}^n k = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n = n!$$

(c) Oder die Summe der ersten  $n$  natürlichen Zahlen:

$$\sum_{k=1}^n k = 1 + 2 + 3 + \dots + n$$

(d) Hier fängt die Summe mal nicht mit  $k = 1$  an:

$$\sum_{k=4}^7 \frac{1}{k+1} = \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}$$

Mit dem Summenzeichen können wir also z.B. die Summe der ersten  $n$  natürlichen Zahlen sehr schön kurz als  $\sum_{k=1}^n k$  schreiben. Wenn wir aber diese Summe tatsächlich *berechnen* wollen, hilft uns das noch nicht weiter. Eine kurze Formel zur Berechnung der Summe der ersten natürlichen Zahlen bekommt man mit dem sogenannten *Trick von Gauß*:

Wir schreiben die Summe  $\sum_{k=1}^n k$  zweimal untereinander, beim zweiten mal anders herum von  $n$  runter bis 1, und addieren dann die übereinander stehenden Zahlen:

$$\begin{aligned} 2 \cdot \sum_{k=1}^n k &= \begin{array}{cccccc} 1 & +2 & +3 & + \dots & +n-1 & +n \\ +n & +n-1 & +n-2 & + \dots & +2 & +1 \end{array} \\ &= \begin{array}{cccccc} n+1 & +n+1 & +n+1 & + \dots & +n+1 & +n+1 \end{array} \\ &= n(n+1) \end{aligned}$$

Im letzten Schritt wurde benutzt, dass man in jeder Spalte als Summe  $n + 1$  bekommt und man insgesamt  $n$  solche Spalten hat, nämlich jeweils eine Spalte für jede der Zahlen 1 bis  $n$ . Insgesamt also  $n$  mal  $n + 1$ . Wenn wir jetzt die ganze Gleichung noch durch 2 teilen, bekommen wir als Ergebnis

$$\sum_{k=1}^n k = \frac{n(n+1)}{2} \quad (1.4)$$

oder, wenn man die Summe links wieder lang ausschreibt,

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Für Summen gelten folgende allgemeine Rechenregeln:

$$\sum_{k=m}^n a_k = \sum_{i=m}^n a_i \quad (1.5)$$

$$\sum_{k=m}^n a_k = \sum_{k=m}^l a_k + \sum_{k=l+1}^n a_k \quad \text{wobei } m \leq l \leq n \quad (1.6)$$

$$\sum_{k=m}^n a_{k+j} = \sum_{k=m+j}^{n+j} a_k \quad (1.7)$$

Gleichung (1.5) bedeutet, dass man den Index  $k$  der Summation beliebig umbenennen kann, z.B. hier  $i$ . In (1.6) wurde die Summe einfach in einen ersten Teil von  $m$  bis  $l$  und einen zweiten Teil (den Rest) von  $l + 1$  bis  $n$  aufgeteilt. (1.7) nennt man auch *Index-Verschiebung* oder *-Shift*. Dass diese letzte Gleichung richtig ist, sieht man wenn man einfach die unteren und oberen Grenzen für  $k$  auf beiden Seiten einsetzt. Z.B. ist der erste Summand auf der linken Seite für  $k = m$  gleich  $a_{m+j}$  ( $k = m$  in  $a_{k+j}$  einsetzen!) und auch rechts erhält man  $a_{m+j}$  ( $a_k$  mit  $k = m + j$ ).

Die Summe  $\sum_{k=m}^n a_k$  kann man rekursiv definieren:

$$\begin{aligned} \sum_{k=m}^m a_k &= a_m && \text{Start } n = m \\ \sum_{k=m}^{n+1} a_k &= \left( \sum_{k=m}^n a_k \right) + a_{n+1} && \text{Schritt von } n \text{ nach } n + 1 \end{aligned}$$

Bei Produkten geht es entsprechend.

### 1.3 Permutationen

Eine Anordnung einer (endlichen) Menge von Zahlen in einer bestimmten Reihenfolge heißt *Permutation*. Schreibweise:

$$(a_1, a_2, \dots, a_n)$$

d.h. man schreibt die Zahlen in der gewählten Reihenfolge in runde Klammern.

**Beachte:**

- Menge  $\{a, b, c\}$ : Reihenfolge egal
- Permutation  $(a, b, c)$ : Reihenfolge festgelegt

**Beispiel 1.3.1** (a) Die Menge  $A = \{1, 2\}$  hat die beiden Permutationen

$$(1, 2) \quad \text{und} \quad (2, 1).$$

(b) Zwei Permutationen von  $A = \{1, 2, 3, 4\}$  sind (z.B.):

$$(1, 2, 3, 4) \quad \text{und} \quad (2, 4, 3, 1)$$

## 1.4 Prinzip der vollständigen Induktion

*Vollständige Induktion* ist ein Prinzip, um zu beweisen, dass eine bestimmte Aussage oder Formel für jede natürliche Zahl  $n$  richtig, also wahr ist. Es funktioniert so:

**Prinzip der vollständigen Induktion.** Man hat eine Aussage (oder Formel)  $A(n)$ , die von der natürlichen Zahl  $n \in \mathbb{N}$  abhängt. Wenn man weiß,

$$\text{I. } A(0) \text{ ist wahr,} \tag{IA}$$

$$\text{II. wenn } A(n) \text{ wahr ist, dann ist auch } A(n+1) \text{ wahr,} \tag{IS}$$

dann folgt:

$$A(n) \text{ ist wahr für jedes } n \in \mathbb{N}.$$

Man nennt I. auch *Induktionsanfang* (IA) und II. *Induktionsschritt* (IS) von  $n$  auf  $n+1$ . Wichtig ist zu verstehen, dass beim Induktionsschritt nicht vorausgesetzt oder bewiesen wird, dass  $A(n)$  wahr ist; und auch nicht, dass  $A(n+1)$  unbedingt wahr ist. Es wird nur eine Wenn-Dann Aussage gezeigt, also eine Implikation ( $\Rightarrow$ , „daraus folgt“): immer *wenn*  $A(n)$  wahr ist, *dann* ist auch  $A(n+1)$  wahr; oder kurz:  $A(n) \Rightarrow A(n+1)$ .

Man kann sich das sehr schön durch Dominosteine veranschaulichen:  $A(n)$  bedeutet, dass der  $n$ . Stein umfällt. Der Induktionsschritt (IS) sagt also aus: wenn der  $n$ . Stein umfällt, dann fällt auch der nächste Stein, der  $n+1$ . Stein, um. Damit allein ist noch nicht ausgesagt, dass der  $n$ . oder der  $n+1$ . Stein jemals umfällt. Erst (IA), der Induktionsanfang, setzt alles in Gang: (IA) heißt  $A(0)$  ist wahr, also der 0. Stein fällt um. Dann kommt (IS) (für  $n=0$ ): Wenn der 0. Stein umfällt, dann auch der 1. Stein. Da der 0. Stein wegen (IA) tatsächlich umgefallen ist, fällt also auch der 1. Stein. Jetzt kommt (IS) für  $n=1$ : Weil der 1. Stein umgefallen ist, fällt auch der 2. Stein, u.s.w. Es fallen alle Steine um, d.h. die Aussage  $A(n)$  ist für alle  $n \in \mathbb{N}$  wahr. Mathematisch kann man das alles so aufschreiben:

$$\begin{array}{l} \text{(IA)} \quad \Longrightarrow \quad A(0) \text{ ist wahr} \quad \xrightarrow{\text{(IS)}} \quad A(1) \text{ ist wahr} \\ \quad \quad \quad \xrightarrow{\text{(IS)}} \quad A(2) \text{ ist wahr} \quad \xrightarrow{\text{(IS)}} \quad \dots \end{array}$$

Variante: Die Induktion kann statt mit  $n = 0$  auch erst später bei einer Zahl  $n_0 \in \mathbb{N}$  starten. Der Induktionsanfang (IA) lautet dann „ $A(n_0)$  ist wahr“. Die vollständige Induktion liefert in diesem Fall als Ergebnis, dass  $A(n)$  wahr ist für alle  $n \geq n_0$ .

Wir schauen uns jetzt zwei Beweise mit vollständiger Induktion an:

**Beispiel 1.4.1** Wir zeigen, dass die Anzahl Permutationen einer Menge aus  $n$  Elementen gleich  $n!$  ist. (Hierbei ist  $n \geq 1$ .) Die Aussage  $A(n)$  ist dann:

$A(n)$  : jede Menge  $M$  mit  $n$  Elementen hat  $n!$  Permutationen

**Beweis. I. Induktionsanfang  $n = 1$ :**

Eine Menge aus einem Element  $M = \{a\}$  hat genau eine Permutation, nämlich  $(a)$ . Also  $M$  hat  $1 = 1!$  Permutationen. Also ist die Aussage für  $n = 1$  wahr, d.h.  $A(1)$  ist wahr.

**II. Induktionsschritt von  $n$  auf  $n + 1$ :**

Wir wollen zeigen:

wenn  $A(n)$  wahr ist, dann ist auch  $A(n + 1)$  wahr.

Nehmen wir also an,  $A(n)$  wäre für eine Zahl  $n \geq 1$  wahr. (Dies nennt man auch *Induktionsannahme*.) Um jetzt  $A(n + 1)$  zu zeigen, betrachten wir eine Menge mit  $n + 1$  Elementen  $M = \{a_1, \dots, a_{n+1}\}$ . Wir wählen ein  $a_j \in M$  und zählen alle Permutationen von  $M$  mit  $a_j$  am Anfang. Das sind Permutationen der Form  $(a_j, *, \dots, *)$  wobei  $(*, \dots, *)$  irgendeine Permutation der restlichen Elemente ist, also der Menge  $M \setminus \{a_j\}$ . Diese Menge  $M \setminus \{a_j\}$  hat  $n$  Elemente (denn  $M$  hat  $n + 1$ ). Nach der Induktionsannahme  $A(n)$  wissen wir nun, dass es  $n!$  Permutationen von  $M \setminus \{a_j\}$  gibt. Also gibt es auch genau  $n!$  Permutationen von  $M$  mit  $a_j$  am Anfang. Weil  $a_j$  irgendein Element aus  $M$  sein kann, haben wir  $n + 1$  Möglichkeiten,  $a_j$  zu wählen, und zu jeder dieser Wahlen gibt es immer  $n!$  Permutationen. Also hat  $M$  insgesamt  $(n + 1) \cdot n! = (n + 1)!$  Permutationen, d.h.  $A(n + 1)$  ist wahr.

Damit haben wir auch den Induktionsschritt  $A(n) \Rightarrow A(n + 1)$  gezeigt. Mit vollständiger Induktion folgt nun aus I. und II., dass  $A(n)$  wahr ist für alle  $n \geq 1$ .  $\square$

**Beispiel 1.4.2** Wir wollen mit vollständiger Induktion folgende Formel für alle  $n \geq 1$  zeigen:

$$A(n) : \quad \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

**Beweis. I. Induktionsanfang  $n = 1$ :**

Wir rechnen nacheinander die linke und rechte Seite der Formel für  $n = 1$  aus. Links ergibt das

$$\sum_{k=1}^1 k^2 = 1^2 = 1$$

und rechts

$$\frac{1(1+1)(2 \cdot 1 + 1)}{6} = \frac{2 \cdot 3}{6} = 1.$$

Also sind beide Seiten bei  $n = 1$  gleich, und damit ist  $A(1)$  wahr.

**II. Induktionsschritt  $n \rightarrow n + 1$ :**

Unsere Induktionsannahme ist wieder, dass die Formel  $A(n)$  für ein  $n \in \mathbb{N}$  wahr ist, und wir wollen zeigen, dass dann auch  $A(n + 1)$  wahr ist. D.h. wir wollen die Formel für  $n + 1$  herleiten. Wir fangen wieder auf der linken Seite an:

$$\begin{aligned} \sum_{k=1}^{n+1} k^2 &= \sum_{k=1}^n k^2 + (n+1)^2 \\ &\stackrel{A(n)}{=} \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\ &= \frac{(n+1)}{6} (n(2n+1) + 6(n+1)) = \frac{n+1}{6} (2n^2 + 7n + 6). \end{aligned}$$

Das  $A(n)$  über dem zweiten Gleichheitszeichen deutet an, dass wir an dieser Stelle die Induktionsannahme  $A(n)$  verwendet haben. Jetzt berechnen wir die rechte Seite für  $n + 1$  und ersetzen dafür rechts in der Formel von  $A(n)$  jedes  $n$  durch  $n + 1$ :

$$\begin{aligned} \frac{(n+1)(n+2)(2(n+1)+1)}{6} &= \frac{n+1}{6} (n+2)(2n+3) \\ &= \frac{n+1}{6} (2n^2 + 7n + 6). \end{aligned}$$

Linke und rechte Seite sind also gleich, d.h. es folgt

$$\sum_{k=1}^{n+1} k^2 = \frac{(n+1)(n+2)(2n+3)}{6}$$

und damit ist  $A(n + 1)$  wahr.  $\square$

**1.5 Binomialkoeffizienten**

Für natürliche Zahlen  $n, k \in \mathbb{N}$  mit  $k \leq n$  ist der *Binomialkoeffizient*  $\binom{n}{k}$  definiert durch die Formel

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \quad (1.8)$$

Für  $\binom{n}{k}$  sagt man auch „ $n$  über  $k$ “. Wenn wir die Fakultäten in der Definition ausschreiben, können wir einige Terme wegekürzen und bekommen damit eine andere Formel für den Binomialkoeffizient:

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \dots \cdot 3 \cdot 2 \cdot 1}{k! \cdot 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-k)} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!}$$

Zur Veranschaulichung machen wir das auch mit konkreten Zahlen, z.B.:

$$\binom{6}{2} = \frac{6!}{2! \cdot 4!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6}{1 \cdot 2 \cdot 1 \cdot 2 \cdot 3 \cdot 4} = \frac{5 \cdot 6}{1 \cdot 2} = 5 \cdot 3 = 15.$$

Wir sehen, dass wir den Nenner hier sogar komplett wegekürzen können, dass also der Binomialkoeffizient eine ganze Zahl ergibt (und keinen Bruch). Das ist tatsächlich immer so, wie wir später sehen werden.

Wir notieren einige Eigenschaften von  $\binom{n}{k}$ :

- Spezielle Werte:

$$\binom{n}{0} = \frac{n!}{0!(n-0)!} = \frac{n!}{n!} = 1, \quad \binom{n}{1} = \frac{n!}{1!(n-1)!} = \frac{n!}{(n-1)!} = n$$

- Symmetrie:

$$\binom{n}{n-k} = \frac{n!}{(n-k)!(n-(n-k))!} = \frac{n!}{(n-k)!k!} = \binom{n}{k}$$

Zum Beispiel ist  $\binom{6}{2} = \binom{6}{4}$  denn  $\binom{6}{2} = \binom{6}{6-2} = \binom{6}{4}$ . Genauso gilt

$$\binom{n}{n} = \binom{n}{n-0} = \binom{n}{0} = 1, \quad \binom{n}{n-1} = \binom{n}{1} = n.$$

- Rekursionsformel:

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1} \quad (1.9)$$

denn:

$$\begin{aligned} \binom{n}{k} + \binom{n}{k+1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-(k+1))!} \\ &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-k-1)!} \\ &= \frac{n!}{k!(n-k-1)!} \cdot \left( \frac{1}{n-k} + \frac{1}{k+1} \right) \\ &= \frac{n!}{k!(n-k-1)!} \cdot \frac{k+1+n-k}{(n-k)(k+1)} \\ &= \frac{n!}{k!(n-k-1)!} \cdot \frac{n+1}{(n-k)(k+1)} \\ &= \frac{(n+1)!}{(k+1)!(n-k)!} \\ &= \binom{n+1}{k+1} \end{aligned}$$

Man kann die Binomialkoeffizienten sehr schön im *Pascalschen Dreieck* darstellen:

$$\begin{array}{l} n = 0: \quad \binom{0}{0} = 1 \\ n = 1: \quad \binom{1}{0} = 1 \quad \binom{1}{1} = 1 \\ n = 2: \quad \binom{2}{0} = 1 \quad \binom{2}{1} = 2 \quad \binom{2}{2} = 1 \\ n = 3: \quad \binom{3}{0} = 1 \quad \binom{3}{1} = 3 \quad \binom{3}{2} = 3 \quad \binom{3}{3} = 1 \\ n = 4: \quad \binom{4}{0} = 1 \quad \binom{4}{1} = 4 \quad \binom{4}{2} = 6 \quad \binom{4}{3} = 4 \quad \binom{4}{4} = 1 \end{array}$$

u.s.w.

Wir schreiben das Pascalsche Dreieck nochmal auf, jetzt aber mit den Binomialkoeffizienten  $\binom{n}{k}$  nur am Rand.

$$\begin{array}{cccccccc}
 n = 0: & & & & & & & & \binom{0}{k} \\
 & & & & & & & & \\
 n = 1: & & & & & & & & \binom{1}{k} \\
 & & & & & & & & \\
 n = 2: & & & & & & & & \binom{2}{k} \\
 & & & & & & & & \\
 n = 3: & & & & & & & & \binom{3}{k} \\
 & & & & & & & & \\
 n = 4: & & & & & & & & \binom{4}{k} \\
 & & & & & & & & \\
 \text{u.s.w.} & & & & & & & & 
 \end{array}$$

Hier erkennt man jetzt das Bildungsgesetz des Pascalschen Dreiecks: In jeder neuen Zeile kommt eine weitere Zahl hinzu, wobei links und rechts am Rand immer Einsen stehen. Die Zahlen im Inneren des Dreiecks erhält man durch Addition der beiden Zahlen links und rechts darüber, was hier durch das  $\oplus$  Symbol angedeutet ist.

Im Pascalschen Dreieck tauchen die Eigenschaften von  $\binom{n}{k}$  wieder auf, die wir weiter oben schon gesehen hatten:

- Rand:  $\binom{n}{0} = \binom{n}{n} = 1$ , am Rand steht immer eine Eins.
- Symmetrie:  $\binom{n}{k} = \binom{n}{n-k}$ , z.B.  $\binom{4}{1} = \binom{4}{3}$ . Das Dreieck ist spiegelsymmetrisch zur Mittelachse des Dreiecks.
- Rekursionsformel (1.9): z.B.  $\binom{3}{1} + \binom{3}{2} = \binom{4}{2}$ . Das ist genau die  $\oplus$ -Addition im Inneren des Dreiecks.

Hier kann man jetzt auch sehen, dass jeder Binomialkoeffizient  $\binom{n}{k}$  stets eine ganze Zahl ergibt, dass sich also in (1.8) der Nenner letztlich immer komplett wegekürzt: Denn das Pascalschen Dreieck entsteht ja vollständig durch Hinzufügen von Einsen am Rand und Addition im Inneren, es können also nie Brüche entstehen.

## 1.6 Der allgemeine binomische Satz

Der allgemeine binomische Satz ist eine Verallgemeinerung der bekannten 1. binomischen Formel

$$(a + b)^2 = a^2 + 2ab + b^2$$

auf Exponenten größer als 2.

**Satz 1.6.1 (allgemeiner binomischer Satz)** Für alle  $n \in \mathbb{N}$  und  $a, b \in \mathbb{R}$  gilt

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k. \quad (1.10)$$

Wenn wir die Summe in (1.10) lang ausschreiben lautet der binomische Satz

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n.$$

Rechnen wir nach, was der binomische Satz für  $n = 2$  und  $n = 3$  ergibt:

$$\begin{aligned} (a+b)^2 &= \sum_{k=0}^2 \binom{2}{k} a^{2-k} b^k = \underbrace{\binom{2}{0}}_1 a^2 + \underbrace{\binom{2}{1}}_2 a^{2-1}b + \underbrace{\binom{2}{2}}_1 b^2 \\ &= a^2 + 2ab + b^2 \end{aligned}$$

Das ist (wie erwartet) die 1. binomische Formel  $(a+b)^2 = a^2 + 2ab + b^2$ . Für  $n = 3$  erhalten wir

$$\begin{aligned} (a+b)^3 &= \sum_{k=0}^3 \binom{3}{k} a^{3-k} b^k = \underbrace{\binom{3}{0}}_1 a^3 + \underbrace{\binom{3}{1}}_3 a^2b + \underbrace{\binom{3}{2}}_3 ab^2 + \underbrace{\binom{3}{3}}_1 b^3 \\ &= a^3 + 3a^2b + 3ab^2 + b^3. \end{aligned}$$

Man beachte, dass man die Binomialkoeffizienten in der Formel direkt aus der entsprechenden Zeile des Pascalschen Dreiecks ablesen kann! Für  $(a+b)^2$ , also  $n = 2$ , ist das 1, 2, 1 und damit  $1a^2 + 2ab + 1b^2$ . Für  $n = 3$  ist es 1, 3, 3, 1, also  $1a^3 + 3a^2b + 3ab^2 + 1b^3$ .

Warum ist die Formel für  $(a+b)^3$  richtig? Wir könnten das so nachrechnen:

$$\begin{aligned} (a+b)^3 &= (a+b)^2(a+b) \\ &= (a^2 + 2ab + b^2)(a+b) \\ &= a^3 + 2a^2b + ab^2 + a^2b + 2ab^2 + b^3 \\ &= a^3 + 3a^2b + 3ab^2 + b^3 \end{aligned}$$

Genauso könnten wir auch  $(a+b)^4$  ausrechnen, indem wir etwa  $(a+b)^2(a+b)^2$  ausmultiplizieren, oder  $(a+b)^3(a+b)$ , wobei wir für  $(a+b)^3$  das Ergebnis von eben verwenden. Es ist aber klar, dass das ziemlich aufwändig ist und man leicht Fehler beim Rechnen machen kann.

**Merke:** Es ist viel einfacher, schneller und sicherer, wenn man den allgemeinen binomischen Satz (1.10) benutzt!

Noch wissen wir nicht, warum die Formel (1.10) des allgemeinen binomischen Satzes für jedes  $n \in \mathbb{N}$  richtig ist. Wir haben das nur für  $n = 2$  und  $n = 3$  überprüft. Zum Abschluss beweisen wir deshalb jetzt den binomischen Satz 1.6.1. Das ist nochmal ein Beispiel für einen Beweis mit vollständiger Induktion.

$$A(n) : (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \quad n \in \mathbb{N}$$

**Beweis Satz 1.6.1.** Die Aussage  $A(n)$ , die wir zeigen wollen, ist Formel (1.10).

I. Induktionsanfang  $n = 0$ :

Wir rechnen nach, dass  $A(0)$  wahr ist, dass also die allgemeine binomische Formel (1.10) für  $n = 0$  richtig ist:

$$\sum_{k=0}^0 \binom{0}{k} a^{0-k} b^k = \binom{0}{0} a^0 b^0 = \binom{0}{0} = 1 = (a + b)^0.$$

II. Induktionsschritt  $n \rightarrow n + 1$ :

Induktionsannahme:  $A(n)$  ist wahr, d.h. (1.10) gilt für  $n$ . Wir rechnen nun die Formel für  $n + 1$  nach:

$$\begin{aligned} (a + b)^{n+1} &= (a + b)^n (a + b) \\ &= \left( \binom{n}{0} a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \dots + \binom{n}{n} b^n \right) (a + b) \\ &= \binom{n}{0} a^{n+1} + \binom{n}{1} a^n b + \binom{n}{2} a^{n-1} b^2 + \dots + \binom{n}{n} a b^n \\ &\quad + \binom{n}{0} a^n b + \binom{n}{1} a^{n-1} b^2 + \binom{n}{2} a^{n-2} b^3 + \dots + \binom{n}{n} b^{n+1} \\ &= a^{n+1} + \left( \binom{n}{1} + \binom{n}{0} \right) a^n b + \left( \binom{n}{2} + \binom{n}{1} \right) a^{n-1} b^2 \\ &\quad + \dots + \left( \binom{n}{n} + \binom{n}{n-1} \right) a b^n + b^{n+1} \\ &= a^{n+1} + \binom{n+1}{1} a^n b + \binom{n+1}{2} a^{n-1} b^2 \\ &\quad + \dots + \binom{n+1}{n} a b^n + b^{n+1} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k \end{aligned}$$

Im vorletzten Schritt wurde die Rekursionsformel (1.9) für Binomialkoeffizienten angewandt.  $A(n + 1)$  ist damit gezeigt und der Induktionsbeweis komplett.  $\square$

## Kapitel 2

# Folgen und Grenzwerte

In diesem Kapitel untersuchen und berechnen wir Grenzwerte von Zahlenfolgen. Grenzwerte sind wichtig für die folgenden Kapitel und überhaupt für große Teile der ganzen Mathematik (was wir in Mathematik A auch schon teilweise gesehen haben).

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik B, Kapitel 1.

### 2.1 Zahlenfolgen

Wie schon im ersten Kapitel betrachten wir wieder Zahlenfolgen. Zunächst etwas zur Schreibweise.

Eine *Folge* ist eine unendliche Folge reeller Zahlen

$$a_0, a_1, a_2, a_3, \dots \quad (a_n \in \mathbb{R})$$

Um eine Folge als Ganzes zu bezeichnen, schreibt man sie in runde Klammern und gibt dabei meistens nur das allgemeine Folgenglied  $a_n$  an:

$$(a_n)_{n \in \mathbb{N}} = (a_0, a_1, a_2, a_3, \dots) = (a_n)_{n \geq 0}$$

Der Index  $n \in \mathbb{N}$  bzw.  $n \geq 0$  unten an der schließenden Klammer gibt dabei den Bereich an, in dem  $n$  „läuft“. Die Folge kann auch erst mit  $a_{n_0}$  statt mit  $a_0$  beginnen; man schreibt dann  $(a_n)_{n \geq n_0}$ , z.B.

$$(a_n)_{n \geq 2} = (a_2, a_3, a_4, \dots).$$

Man kann auch einfach nur  $(a_n)$  für die Folge schreiben, wenn klar ist, welche  $n$  gemeint sind, oder wenn das keine Rolle spielt. Die Zahlen  $a_n$ , aus denen die Folge  $(a_n)$  besteht, heißen *Glieder* der Folge.

**Beispiel 2.1.1** (a) Die Folge mit den Folgengliedern  $a_n = n^2$  für  $n \geq 0$  sieht so aus:

$$(a_n)_{n \in \mathbb{N}} = (n^2)_{n \in \mathbb{N}} = (0^2, 1^2, 2^2, 3^2, \dots) = (0, 1, 4, 9, \dots)$$

(b) Die Folge mit den Gliedern  $a_n = \frac{1}{n}$  für  $n \geq 1$ :

$$(a_n)_{n \geq 1} = \left( \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \dots \right)$$

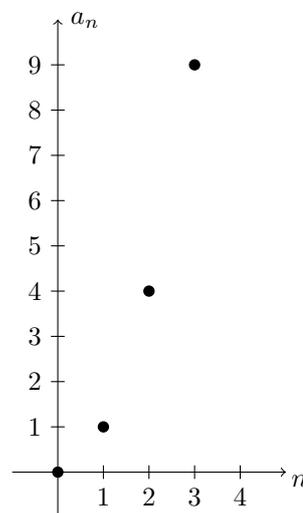
(c) Die Folge mit  $a_n = c$  für alle  $n \in \mathbb{N}$  ( $c \in \mathbb{R}$  konstant) heißt *konstante Folge*:

$$(a_n)_{n \in \mathbb{N}} = (c, c, c, \dots)$$

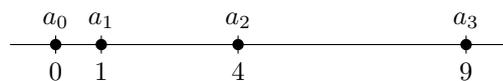
Zur Veranschaulichung kann man Folgen *graphisch darstellen*, entweder in einem Koordinatensystem oder auf dem Zahlenstrahl. Wir zeigen das für die obigen drei Beispielfolgen:

(a)  $a_n = n^2$

als Graph:

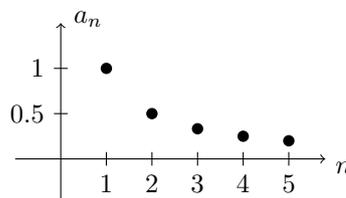


oder auf dem Zahlenstrahl:

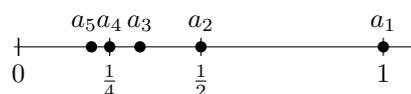


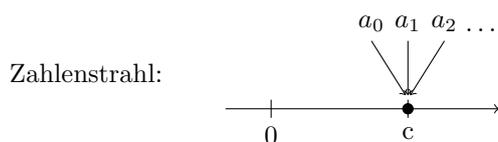
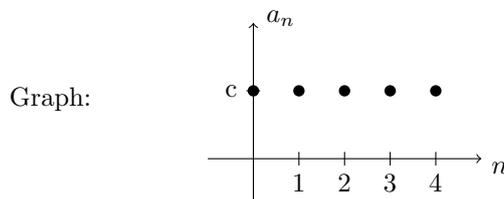
(b)  $a_n = \frac{1}{n}$

Graph:



Zahlenstrahl:



(c)  $a_n = c$ 

## 2.2 Grenzwerte

Eine Zahl  $a \in \mathbb{R}$  heißt *Grenzwert* der Folge  $(a_n)$ , wenn sich die Folgenglieder  $a_n$  bei wachsendem  $n$  „immer mehr und beliebig dicht“ an  $a$  annähern. In diesem Fall schreibt man

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

oder auch

$$a_n \rightarrow a \quad \text{bei} \quad n \rightarrow \infty.$$

Das Symbol  $\lim$  steht für *Limes* (lateinisch für Grenze), und für  $a_n \rightarrow a$  sagt man „ $a_n$  konvergiert (oder strebt) gegen  $a$ “.

Die obige Formulierung, dass die  $a_n$  sich „immer mehr und beliebig dicht“ an  $a$  annähern, ist mathematisch nicht exakt, gibt aber eine erste Anschauung dafür, was ein Grenzwert ist. Eine exakte Definition des Grenzwerts folgt im nächsten Abschnitt.

Zunächst schauen wir uns wieder unsere drei Beispielfolgen an und untersuchen sie auf Grenzwerte:

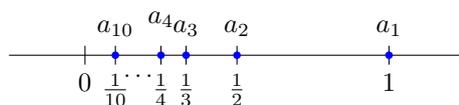
**Beispiel 2.2.1** (a)  $a_n = \frac{1}{n}$ ,  $n \geq 1$ . Die ersten Folgenglieder sind

$$(a_n)_{n \geq 1} = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots\right)$$

Man erkennt, dass die Glieder immer kleiner werden und sich dabei immer mehr an 0 annähern, je größer  $n$  wird, z.B.

$$a_{10} = \frac{1}{10} = 0.1, \quad a_{100} = \frac{1}{100} = 0.01, \quad a_{1000} = 0.001, \dots$$

Das sieht man auch bei der Darstellung auf dem Zahlenstrahl:



Da die Folgenglieder also „immer weiter gegen 0 streben“, hat die Folge  $(a_n)$  den Grenzwert  $a = 0$ . Als Formel

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0 \quad \text{oder} \quad \frac{1}{n} \rightarrow 0 \quad \text{für} \quad n \rightarrow \infty$$

(b)  $a_n = n^2$

Die Folgenglieder 0, 1, 4, 9, 16, ... wachsen immer schneller, und werden dabei größer als jede feste Zahl. Die Glieder nähern sich also bei wachsendem  $n$  keiner festen Zahl  $a$  beliebig nah an. Damit hat die Folge  $(n^2)_{n \in \mathbb{N}}$  keinen Grenzwert.

(c)  $a_n = c$  konstant

Diese Folge hat den Grenzwert  $a = c$ , d.h.  $\lim_{n \rightarrow \infty} c = c$ . Denn die Folgenglieder  $a_n$  kommen der Zahl  $a = c$  beliebig nah; sie sind ja sogar gleich  $c$ .

## 2.3 Exakte Definition des Grenzwertes

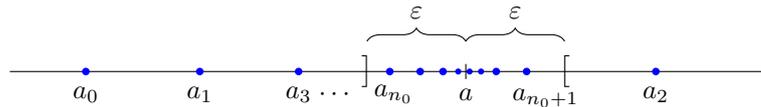
Wir schauen uns jetzt die mathematisch genaue Definition des Grenzwerts an.

**Definition 2.3.1** Eine Zahl  $a \in \mathbb{R}$  ist Grenzwert der Folge  $(a_n)$ , wenn es zu jedem  $\varepsilon > 0$  ein  $n_0 \in \mathbb{N}$  gibt, so dass gilt

$$|a_n - a| < \varepsilon \quad \text{für alle} \quad n \geq n_0. \quad (2.1)$$

Die Folge  $(a_n)$  heißt *konvergent* wenn sie einen Grenzwert hat, sonst *divergent*.

Der Ausdruck  $|a_n - a|$  ist genau der Abstand von  $a_n$  zu  $a$ . Damit bedeutet (2.1) also, dass ab dem Folgenglied  $a_{n_0}$  alle nachfolgenden Glieder  $a_n$  von  $a$  einen Abstand kleiner als  $\varepsilon$  haben. Graphisch sieht das dann so aus:



Ab  $a_{n_0}$  liegen *alle*  $a_n$  im Intervall zwischen den eckigen Klammern. Beachte:

- $|a_n - a| < \varepsilon \Leftrightarrow a - \varepsilon < a_n < a + \varepsilon \Leftrightarrow a_n \in ]a - \varepsilon, a + \varepsilon[$
- $n_0$  hängt von  $\varepsilon$  ab. Typischerweise wird  $n_0$  umso größer, je kleiner  $\varepsilon$  wird. Anschaulich: bei kleinerem Intervall muss man länger warten, bis die Folgenglieder  $a_n$  im Intervall liegen.

**Beispiel 2.3.2** (a) Betrachten wir die Folge

$$a_n = \frac{2n}{n+3}.$$

Einsetzen konkreter Zahlen für  $n$  liefert

$$a_1 = \frac{2}{4} = \frac{1}{2}, \quad a_2 = \frac{4}{5}, \quad a_3 = 1, \quad a_5 = \frac{5}{4}, \dots, \quad a_{10} = \frac{20}{13}, \dots, \quad a_{100} = \frac{200}{103}.$$

Die Folge scheint sich also immer mehr an 2 anzunähern, d.h. wir vermuten als Grenzwert

$$a = 2,$$

oder anders,

$$\lim_{n \rightarrow \infty} \frac{2n}{n+3} = 2.$$

Diese Vermutung werden wir nun exakt mit der Definition 2.3.1 nachrechnen. Zuerst berechnen wir den Betrag in (2.1):

$$|a_n - a| = \left| \frac{2n}{n+3} - 2 \right| = \left| \frac{2n - 2(n+3)}{n+3} \right| = \left| \frac{-6}{n+3} \right| = \frac{6}{n+3}$$

Um die Definition zu überprüfen, müssen wir jetzt für ein gegebenes  $\varepsilon > 0$  die Zahl  $n_0$  bestimmen, so dass (2.1) gilt. Machen wir das z.B. für  $\varepsilon = \frac{1}{10}$  indem wir unsere Werte für  $|a_n - a|$  und  $\varepsilon$  in (2.1) einsetzen und umformen:

$$\begin{aligned} |a_n - a| < \varepsilon &\Leftrightarrow \frac{6}{n+3} < \frac{1}{10} &\Leftrightarrow 60 < n+3 \\ &\Leftrightarrow n > 57 \end{aligned}$$

Wenn also  $n > 57$  ist, dann ist  $|a_n - a| < \varepsilon = \frac{1}{10}$ . Wenn wir jetzt  $n_0 = 58$  wählen, dann folgt aus  $n \geq n_0$  dass  $n > 57$  und damit ist (2.1) erfüllt. Bei  $\varepsilon = \frac{1}{100}$  erhalten wir

$$|a_n - a| = \frac{6}{n+3} < \frac{1}{100} \Leftrightarrow 600 < n+3 \Leftrightarrow n > 597,$$

d.h. in diesem Fall muss man  $n_0 = 598$  wählen damit (2.1) erfüllt ist.<sup>1</sup>

Das sind aber nur einzelne Beispiel für  $\varepsilon$ . Um die Definition des Grenzwerts wirklich zu überprüfen muss gezeigt werden, dass man zu *jedem*  $\varepsilon > 0$  so ein  $n_0$  finden kann. Das bedeutet, wir müssen unsere Rechnung für ein *allgemeines*  $\varepsilon > 0$  durchführen:

$$|a_n - a| = \frac{6}{n+3} < \varepsilon \Leftrightarrow \frac{6}{\varepsilon} < n+3 \Leftrightarrow n > \frac{6}{\varepsilon} - 3$$

Sei jetzt  $n_0$  die kleinste natürliche Zahl, die größer als  $\frac{6}{\varepsilon} - 3$  ist, für die also

$$n_0 > \frac{6}{\varepsilon} - 3$$

gilt. Dann gilt

$$n \geq n_0 \Rightarrow n > \frac{6}{\varepsilon} - 3 \Rightarrow |a_n - a| < \varepsilon$$

und damit (2.1). Damit haben wir zu jedem  $\varepsilon > 0$  ein  $n_0$  gefunden und Definition 2.3.1 ist überprüft. Somit haben wir bewiesen, dass  $a = 2$  der Grenzwert der Folge  $(a_n)$  ist, d.h.

$$\lim_{n \rightarrow \infty} \frac{2n}{n+3} = 2.$$

<sup>1</sup>Hier sieht man übrigens sehr schön, wie bei kleineren  $\varepsilon$  größere  $n_0$  nötig werden.

- (b) Die konstante Folge
- $a_n = c$
- hat den Grenzwert
- $a = c$
- .

Hier ist die Überprüfung von Definition 2.3.1 ganz einfach, denn

$$|a_n - a| = |c - c| = 0.$$

Daher gilt (2.1) automatisch für alle  $\varepsilon > 0$  und  $n \in \mathbb{N}$ . Also kann man (z.B.) immer  $n_0 = 0$  wählen.

## 2.4 Eigenschaften konvergenter Folgen

Hier sind einige grundlegende Eigenschaften konvergenter Folgen:

- (i) Die Folge
- $(a_n)_{n \in \mathbb{N}}$
- ist konvergent mit Grenzwert
- $a$
- genau dann, wenn die Folge
- $(a_n - a)_{n \in \mathbb{N}}$
- den Grenzwert 0 hat. Als Formel:

$$\lim_{n \rightarrow \infty} a_n = a \iff \lim_{n \rightarrow \infty} (a_n - a) = 0$$

Eine Folge, die gegen Null konvergiert, nennt man auch *Nullfolge*. Man kann also auch sagen:  $(a_n)_{n \in \mathbb{N}}$  konvergiert gegen  $a$  genau dann, wenn  $(a_n - a)_{n \in \mathbb{N}}$  eine Nullfolge ist.

Dies kann bei der Grenzwertberechnung helfen, denn oft ist es einfacher nachzurechnen, dass eine Folge gegen 0 konvergiert als gegen eine Zahl  $a \neq 0$ .

- (ii)
- $(a_n)_{n \in \mathbb{N}}$
- ist eine Nullfolge genau dann, wenn die Folge der Beträge
- $(|a_n|)_{n \in \mathbb{N}}$
- eine Nullfolge ist. Wieder als Formel:

$$\lim_{n \rightarrow \infty} a_n = 0 \iff \lim_{n \rightarrow \infty} |a_n| = 0$$

Anders gesagt: Bei der Frage, ob eine Folge eine Nullfolge ist oder nicht, spielt nur der Betrag der Folgenglieder  $a_n$  eine Rolle, das Vorzeichen ist egal.

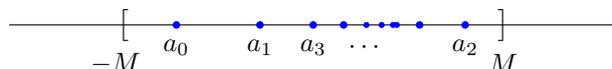
- (iii) Jede konvergente Folge
- $(a_n)$
- ist
- beschränkt*
- , d.h. es gibt eine Zahl
- $M \in \mathbb{R}$
- so dass gilt

$$|a_n| \leq M \quad \text{für alle } n \in \mathbb{N}.$$

Wegen  $|a_n| \leq M \iff -M \leq a_n \leq M$  bedeutet Beschränktheit also, dass die komplette Folge im Intervall  $[-M, M]$  liegt:

$$a_n \in [-M, M] \quad \text{für alle } n \in \mathbb{N}.$$

Graphisch:



- (iv) Jede konvergente Folge hat nur einen Grenzwert.

Alle diese Eigenschaften ergeben sich direkt aus der Definition des Grenzwerts, was wir hier kurz erläutern:

Die Definition besagt, dass  $|a_n - a|$  klein wird, wenn  $a$  Grenzwert ist. Bei einer Nullfolge ( $a = 0$ ) muss also  $|a_n|$  klein werden. Bei  $a \neq 0$ , wenn  $|a_n - a|$  klein wird, kann man daher äquivalent sagen, dass  $(a_n - a)$  eine Nullfolge ist; das ist Eigenschaft (i). Weil für eine Nullfolge der Betrag  $|a_n|$  klein werden muss, spielt das Vorzeichen von  $a_n$  dafür keine Rolle; das ist die Aussage von (ii).

Eigenschaft (iii) sieht man so: Wir haben oben gesehen, dass bei einer konvergenten Folge alle  $a_n$  ab dem Index  $n_0$  im Intervall  $]a - \varepsilon, a + \varepsilon[$  liegen. Anders gesagt liegen nur endlich viele  $a_n$ , nämlich höchstens die ersten von  $a_0$  bis  $a_{n_0-1}$ , außerhalb des Intervalls. Weil nur endlich viele Folgenglieder außerhalb liegen, kann man das Intervall so vergrößern, dass nun alle  $a_n$  im Intervall liegen, das Intervall aber immer noch beschränkt ist, d.h. eine endliche Länge hat. Also ist die Folge beschränkt.

Bei (iv) argumentiert man mit einem Widerspruch: Angenommen, eine konvergente Folge hätte zwei verschiedene Grenzwerte  $a$  und  $b$ . Wegen  $a \neq b$  kann man ein  $\varepsilon > 0$  finden, so dass die Intervalle  $I_a = ]a - \varepsilon, a + \varepsilon[$  und  $I_b = ]b - \varepsilon, b + \varepsilon[$  sich nicht überschneiden<sup>2</sup>, man muss die Intervalle – also das  $\varepsilon$  – nur klein genug machen. Weil  $a$  Grenzwert der Folge ist, liegen nach der Definition unendlich viele Folgenglieder (alle ab  $n_0$ ) im Intervall  $I_a$  um  $a$ , und nur endlich viele außerhalb von  $I_a$ . Weil  $b$  aber auch Grenzwert sein soll, sagt hier die Definition, dass unendlich viele Glieder im Intervall  $I_b$  um  $b$  liegen und damit außerhalb des Intervalls  $I_a$  (denn beide Intervalle schneiden sich nicht). Da wir gerade aber gesagt haben, dass nur endlich viele Glieder außerhalb von  $I_a$  liegen, haben wir einen Widerspruch. Daher ist die Annahme von zwei Grenzwerten falsch, d.h. eine konvergente Folge kann nur einen Grenzwert haben.

Wir benutzen jetzt die Eigenschaften (i), (ii) und (iii), um die Konvergenz bzw. Divergenz einiger Folgen zu begründen.

**Beispiel 2.4.1** (a) Wir betrachten die Folge

$$a_n = \frac{3n - 1}{n + 2}, \quad n \geq 0.$$

Einsetzen großer Zahlen für  $n$  liefert sofort die Vermutung, dass  $a = 3$  der Grenzwert von  $(a_n)$  ist. Wir benutzen jetzt (i) um das zu begründen. D.h. wir betrachten  $a_n - a$  und versuchen nachzurechnen, dass das eine Nullfolge ist.

$$a_n - a = \frac{3n - 1}{n + 2} - 3 = \frac{3n - 1 - 3(n + 2)}{n + 2} = \frac{-7}{n + 2}$$

Weil der Zähler im letzten Bruch konstant gleich  $-7$  ist, der Nenner  $n + 2$  aber beliebig groß wird, konvergiert der Bruch gegen Null,

$$\frac{-7}{n + 2} \rightarrow 0 \quad \text{bei} \quad n \rightarrow \infty,$$

d.h.  $(a_n - a)$  ist eine Nullfolge. Aus (i) schließen wir somit, dass  $a_n$  gegen  $a$  konvergiert, also

$$\lim_{n \rightarrow \infty} \frac{3n - 1}{n + 2} = 3.$$

<sup>2</sup>Der Schnitt der Intervalle ist also leer, d.h. die Intervalle sind *disjunkt*.

Noch eine Bemerkung zu diesem Beispiel: Um (i) benutzen zu können, mussten wir den Grenzwert  $a$  schon vorher wissen. (i) hilft also nicht wirklich, einen noch unbekanntes Grenzwert wirklich zu berechnen, sondern „nur“ um zu zeigen, dass  $a$  wirklich der Grenzwert ist.<sup>3</sup>

(b) Jetzt betrachten wir

$$a_n = \frac{(-1)^n}{n} \quad \text{für } n \geq 1.$$

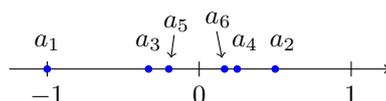
Der Term  $(-1)^n$  erzeugt ein sogenanntes *alternierendes*, also abwechselndes, Vorzeichen:

$$(-1)^1 = -1, \quad (-1)^2 = 1, \quad (-1)^3 = -1, \quad \text{u.s.w.}$$

Die Folge ist also

$$(a_n)_{n \geq 1} = \left( -1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, -\frac{1}{5}, \dots \right).$$

Das abwechselnde Vorzeichen bewirkt, dass die Folgenglieder abwechselnd kleiner und größer als Null sind, die Folge alterniert um Null:



In der Abbildung sieht man, dass sich die Folgenglieder immer weiter an Null annähern, und man vermutet, dass Null der Grenzwert ist. Um das zu zeigen, benutzen wir nun (ii) und betrachten den Betrag  $|a_n|$ :

$$|a_n| = \left| \frac{(-1)^n}{n} \right| = \frac{|(-1)^n|}{n} = \frac{1}{n} \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Dass  $\frac{1}{n}$  gegen Null konvergiert haben wir schon in Beispiel 2.2.1 gesehen. Wir haben also nachgerechnet, dass  $|a_n|$  gegen Null konvergiert, also eine Nullfolge ist, und damit ist nach (ii) auch  $a_n$  eine Nullfolge, d.h.

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

### Achtung!

Nur im Fall von Nullfolgen, also wenn der Grenzwert 0 ist, darf man zum Betrag  $|a_n|$  übergehen.

(c) Ein Beispiel für eine divergente Folge ist  $a_n = 2^n$ . Der Anfang der Folge lautet

$$\begin{aligned} (a_n)_{n \in \mathbb{N}} &= (2^0, 2^1, 2^3, 2^4, \dots) \\ &= (1, 2, 4, 8, 16, 32, \dots). \end{aligned}$$

Offenbar ist diese Folge nicht beschränkt (man sagt dann, sie ist *unbeschränkt*). Aus Eigenschaft (iii) folgt somit, dass die Folge nicht konvergent sein kann, denn sonst wäre sie ja beschränkt. Die Folge  $(2^n)_{n \in \mathbb{N}}$  ist also divergent.

<sup>3</sup>Man kann aber (i) auch für *beliebige* Zahlen  $a$  benutzen, d.h. ohne zu wissen, ob  $a$  der Grenzwert ist, oder nicht: wenn man zeigen kann, dass  $a_n - a$  nicht gegen Null konvergiert, ergibt sich, dass  $a$  nicht Grenzwert der Folge ist.

## 2.5 Rechenregeln für Grenzwerte

Mit den folgenden Regeln, auch *Grenzwertsätze* genannt, kann man Grenzwerte von Summen, Produkten und Quotienten konvergenter Folgen berechnen:

**Satz 2.5.1** Seien  $(a_n)$  und  $(b_n)$  konvergente Folgen mit  $\lim_{n \rightarrow \infty} a_n = a$  und  $\lim_{n \rightarrow \infty} b_n = b$ . Dann sind die Summen- und Produktfolgen  $(a_n \pm b_n)$  und  $(a_n b_n)$  ebenfalls konvergent und es gilt

$$(a) \quad \lim_{n \rightarrow \infty} (a_n \pm b_n) = a \pm b,$$

$$(b) \quad \lim_{n \rightarrow \infty} (a_n b_n) = ab.$$

Speziell gilt bei Multiplikation mit einer Konstanten

$$(b') \quad \lim_{n \rightarrow \infty} c \cdot a_n = c \cdot a \quad \text{für } c \in \mathbb{R}.$$

(c) Falls  $b \neq 0$ , dann ist auch die Quotientenfolge  $(\frac{a_n}{b_n})$  konvergent mit

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

**Beachte:** Obige Rechenregeln gelten nur für *endliche* Grenzwerte, d.h. für  $a, b \in \mathbb{R}$ . In Abschnitt 2.7 werden wir auch  $\pm\infty$  als Grenzwerte kennenlernen, sogenannte uneigentliche Grenzwerte. Bei uneigentlichen Grenzwerten gelten die Rechenregeln nur in bestimmten Fällen und man muss weitere Regeln beachten!

Wir benutzen jetzt die Grenzwertsätze, um Grenzwerte auszurechnen.

**Beispiel 2.5.2** (a) Aus der Konvergenz  $\frac{1}{n} \rightarrow 0$  können wir die Konvergenz der Folgen  $\frac{1}{n^k}$  für  $k = 2, 3, \dots$  folgern. Z.B. können wir  $\frac{1}{n^2}$  in das Produkt

$$\frac{1}{n^2} = \frac{1}{n} \cdot \frac{1}{n}$$

zerlegen. Da beide Faktoren  $\frac{1}{n}$  konvergieren, kann man Grenzwertsatz 2.5.1(b) anwenden und erhält als Grenzwert das Produkt der einzelnen Grenzwerte der Faktoren:

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} = \lim_{n \rightarrow \infty} \left( \frac{1}{n} \cdot \frac{1}{n} \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \lim_{n \rightarrow \infty} \frac{1}{n} = 0 \cdot 0 = 0.$$

Weil wir jetzt wissen, dass  $\frac{1}{n}$  und  $\frac{1}{n^2}$  konvergieren, können wir den Grenzwertsatz 2.5.1(b) erneut anwenden und erhalten

$$\lim_{n \rightarrow \infty} \frac{1}{n^3} = \lim_{n \rightarrow \infty} \left( \frac{1}{n^2} \cdot \frac{1}{n} \right) = \lim_{n \rightarrow \infty} \frac{1}{n^2} \cdot \lim_{n \rightarrow \infty} \frac{1}{n} = 0 \cdot 0 = 0,$$

u.s.w. Es gilt damit

$$\lim_{n \rightarrow \infty} \frac{1}{n^k} = 0 \quad \text{für jedes } k \in \mathbb{N}^* \quad (\text{d.h. } k \geq 1).$$

(b) Wir suchen den Grenzwert von

$$a_n = -3 - \frac{7}{n^2}.$$

Wir schreiben  $a_n$  geschickt als Summe bzw. Produkt von einfacheren Folgen, für die wir die Grenzwerte kennen:

$$\lim_{n \rightarrow \infty} \left( -3 - \frac{7}{n^2} \right) = \lim_{n \rightarrow \infty} \left( -3 - 7 \cdot \underbrace{\frac{1}{n^2}}_{\rightarrow 0} \right) = -3 - 7 \cdot 0 = -3.$$

Wir haben hier die Grenzwertsätze 2.5.1(a) und (b') benutzt, mit der konstanten Folge  $(-3)_{n \in \mathbb{N}}$ . Die genauen Zwischenschritte der Rechnung sind

$$\begin{aligned} \lim_{n \rightarrow \infty} \left( -3 - 7 \cdot \frac{1}{n^2} \right) &\stackrel{(a)}{=} \lim_{n \rightarrow \infty} (-3) - \lim_{n \rightarrow \infty} \left( 7 \cdot \frac{1}{n^2} \right) \\ &\stackrel{(b')}{=} \lim_{n \rightarrow \infty} (-3) - 7 \cdot \lim_{n \rightarrow \infty} \frac{1}{n^2} = -3 - 7 \cdot 0 = -3. \end{aligned}$$

(c) Wir wollen den Grenzwert

$$\lim_{n \rightarrow \infty} \frac{4n - 3n^2}{2n^2 + 1}$$

berechnen. Die Grenzwertsätze sind hier aber nicht direkt anwendbar, denn beim Zerlegen von Zähler und Nenner in die einzelnen Summanden bekommt man die Folgen  $n$  und  $n^2$ , die divergent sind.

Hier hilft ein Prinzip, das man häufig anwenden kann: Man sucht in Zähler und Nenner die höchste vorkommende Potenz, klammert diese aus und kürzt. Dadurch erhält man in Zähler und Nenner konvergente Folgen und die Grenzwertsätze sind anwendbar:

$$\lim_{n \rightarrow \infty} \frac{4n - 3n^2}{2n^2 + 1} = \lim_{n \rightarrow \infty} \frac{n^2 \left( \frac{4}{n} - 3 \right)}{n^2 \left( 2 + \frac{1}{n^2} \right)} = \lim_{n \rightarrow \infty} \frac{\frac{4}{n} - 3}{2 + \frac{1}{n^2}} = \frac{0 - 3}{2 + 0} = -\frac{3}{2}.$$

## 2.6 Teilfolgen und Häufungswerte

Im letzten Abschnitt haben wir gesehen, dass eine Folge maximal einen Grenzwert haben kann. Eine Folge kann aber mehrere sogenannte Häufungswerte besitzen, das sind Zahlen, denen unendlich viele Glieder der Folge beliebig nahe kommen. Häufungswerte sind eng mit Teilfolgen verbunden:

Eine *Teilfolge* entsteht aus einer Folge  $(a_n)$  durch Weglassen von Folgengliedern. Dabei können endlich viele oder unendlich viele Glieder der Ausgangsfolge weggelassen werden. In der Teilfolge müssen aber noch unendlich viele Glieder übrig bleiben, und auch die Reihenfolge darf nicht geändert werden.

Anders ausgedrückt entsteht eine Teilfolge, indem man aus der Folge  $(a_n)_{n \in \mathbb{N}}$  nacheinander Folgenglieder mit den Indizes  $n_0, n_1, n_2, \dots$  für die Teilfolge auswählt. Dabei muss  $n_0 < n_1 < n_2 < \dots$  gelten, d.h. die Reihenfolge der Glieder

wird nicht geändert und Glieder dürfen nicht mehrfach ausgewählt werden. Die Teilfolge ist dann

$$(a_{n_k})_{k \in \mathbb{N}} = (a_{n_0}, a_{n_1}, a_{n_2}, \dots). \quad (2.2)$$

Eine Zahl  $a \in \mathbb{R}$  ist nun ein *Häufungswert* der Folge  $(a_n)_{n \in \mathbb{N}}$  wenn es eine Teilfolge  $(a_{n_k})$  von  $(a_n)$  gibt, die gegen  $a$  konvergiert.

**Beispiel 2.6.1** Betrachten wir als Beispiel die Folge

$$a_n = (-1)^n + \frac{1}{n}, \quad n \geq 1.$$

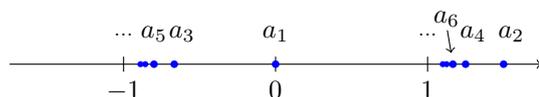
Wir berechnen die ersten Glieder der Folge. Den alternierenden Ausdruck  $(-1)^n$  haben wir schon gesehen; er ist gleich 1 für gerade  $n$  und  $-1$  für ungerade  $n$ .

$$\begin{array}{ll} a_1 = -1 + \frac{1}{1} = 0 & a_2 = 1 + \frac{1}{2} = \frac{3}{2} \\ a_3 = -1 + \frac{1}{3} = -\frac{2}{3} & a_4 = 1 + \frac{1}{4} = \frac{5}{4} \\ a_5 = -1 + \frac{1}{5} = -\frac{4}{5} & a_6 = 1 + \frac{1}{6} = \frac{7}{6} \\ a_7 = -\frac{6}{7} & a_8 = \frac{9}{8} \end{array}$$

u.s.w. Die Folge ist also

$$(a_n)_{n \geq 1} = \left( 0, \frac{3}{2}, -\frac{2}{3}, \frac{5}{4}, -\frac{4}{5}, \frac{7}{6}, \dots \right).$$

Auf dem Zahlenstrahl:



Man sieht, wie sich die geraden Folgenglieder immer mehr der 1 nähern, die ungeraden Glieder immer mehr der  $-1$ .

Schauen wir uns die Teilfolge *aller geraden Folgenglieder* an. Das ist die Folge

$$(a_{2k})_{k \geq 1} = (a_2, a_4, a_6, a_8, \dots) = \left( \frac{3}{2}, \frac{5}{4}, \frac{7}{6}, \frac{9}{8}, \dots \right).$$

Diese Teilfolge konvergiert offenbar gegen 1. Die Teilfolge *aller ungeraden Folgenglieder* ist

$$(a_{2k+1})_{k \geq 0} = (a_1, a_3, a_5, a_7, \dots) = \left( 0, -\frac{2}{3}, -\frac{4}{5}, -\frac{6}{7}, \dots \right);$$

sie konvergiert gegen  $-1$ .

Die Terme  $2k$  und  $2k + 1$  in der Kurzschreibweise der Teilfolgen erzeugen genau die geraden und die ungeraden natürlichen Zahlen, wenn  $k$  alle natürlichen Zahlen durchläuft. Das ist auch jeweils der Ausdruck für  $n_k$  in (2.2), der die Folgenglieder für die Teilfolge auswählt. Da unsere Ausgangsfolge mit  $n = 1$

beginnt, ist das erste gerade Folgenglied bei  $n = 2$ ; daher müssen wir für  $n_k = 2k$  bei  $k = 1$  starten:

$$n_k = 2k = 2, 4, 6, \dots \quad \text{für } k \geq 1$$

für die gerade Teilfolge. Für die ungerade Teilfolge ist das erste  $a_n$  bei  $n = 1$ . Das ist gleich  $2k + 1$  für  $k = 0$ , also

$$n_k = 2k + 1 = 1, 3, 5, \dots \quad \text{für } k \geq 0$$

für die ungerade Teilfolge.

Jetzt zu den Häufungswerten: Da 1 Grenzwert der Teilfolge der geraden Folgenglieder ist, und  $-1$  Grenzwert der Teilfolge der ungeraden Glieder, sind 1 und  $-1$  beides Häufungswerte der gesamten Folge  $(a_n)_{n \in \mathbb{N}}$ . Damit erhalten wir auch: die Gesamtfolge  $(a_n)$  hat keinen Grenzwert, sie ist divergent. Denn da sich unendlich viele Glieder (alle geraden) der Zahl 1 annähern, aber auch unendlich viele Glieder (alle ungeraden) der  $-1$ , kann es keine Zahl  $a \in \mathbb{R}$  geben, gegen die die *ganze* Folge  $(a_n)$  konvergiert.<sup>4</sup>

Allgemein gilt:

**Satz 2.6.2** (a) *Jede beschränkte Folge hat (mindestens) einen Häufungswert.*

(b) *Eine beschränkte Folge ist konvergent genau dann wenn sie nur einen Häufungswert hat. Dieser Häufungswert ist dann der Grenzwert.*

Satz 2.6.2(a) heißt auch **Satz von Bolzano-Weierstraß**. Den Häufungswert kann man folgendermaßen finden (das ist im Prinzip auch der Beweis des Satzes): Da die Folge  $(a_n)$  beschränkt ist, gibt es ein Intervall  $I_0 = [-M, M]$ , in dem alle Glieder  $a_n$  liegen. Wir halbieren das Intervall in der Mitte. Dann müssen in mindestens einer der beiden Hälften unendlich viele Glieder der Folge liegen (vielleicht auch in beiden). Wir wählen eine Hälfte aus, in der unendlich viele Folgenglieder liegen. Dieses Intervall nennen wir  $I_1$ . Jetzt halbieren wir  $I_1$ . Weil in  $I_1$  unendlich viele  $a_n$  liegen, müssen wieder in (mindestens) einer der Hälften unendlich viele  $a_n$  liegen. Diese Hälfte nennen wir  $I_2$ , u.s.w. Wir erhalten damit eine Folge von Intervallen  $I_0, I_1, I_2, \dots$ , die immer kleiner werden und dabei ineinander enthalten sind,

$$I_0 \supset I_1 \supset I_2 \supset \dots,$$

eine sogenannte *Intervallschachtelung*. Da wir die Intervalle immer halbiert haben, konvergiert die Länge der Intervalle gegen 0, d.h. die Intervallschachtelung zieht sich auf eine Zahl  $a \in \mathbb{R}$  zusammen, die in jedem Intervall  $I_k$  liegt. Da in jedem  $I_k$  unendlich viele  $a_n$  liegen, sieht man, dass unendlich viele Folgenglieder  $a_n$  der Zahl  $a$  beliebig nahe kommen. Und das heißt,  $a$  ist ein Häufungswert der Folge.

## 2.7 Bestimmte Divergenz

Wir untersuchen jetzt den Fall, dass eine Folge „gegen unendlich konvergiert“. In diesem Fall spricht man genauer von *bestimmter Divergenz*.

<sup>4</sup>Vergleiche dazu nochmal Eigenschaft (iv) in 2.4: Jede konvergente Folge hat nur einen Grenzwert.

**Definition 2.7.1** (a) Eine Folge  $(a_n)$  *divergiert bestimmt gegen*  $\infty$ , Schreibweise

$$\lim_{n \rightarrow \infty} a_n = \infty,$$

wenn es für jedes  $s \in \mathbb{R}$  ein  $n_0 \in \mathbb{N}$  gibt mit

$$a_n \geq s \quad \text{für alle } n \geq n_0. \quad (2.3)$$

(b) Analog *divergiert*  $(a_n)$  *bestimmt gegen*  $-\infty$ , geschrieben als

$$\lim_{n \rightarrow \infty} a_n = -\infty,$$

wenn es für jedes  $s \in \mathbb{R}$  ein  $n_0 \in \mathbb{N}$  gibt mit

$$a_n \leq s \quad \text{für alle } n \geq n_0. \quad (2.4)$$

Anstelle der Schreibweise mit dem Limes-Symbol kann man auch wieder

$$a_n \rightarrow \infty \quad (\text{bzw. } a_n \rightarrow -\infty) \quad \text{für } n \rightarrow \infty$$

schreiben. Statt „ $(a_n)$  divergiert bestimmt gegen  $\pm\infty$ “ sagt man auch „ $(a_n)$  *konvergiert uneigentlich* gegen  $\pm\infty$ “.

Die Ungleichung (2.3) bedeutet, dass die Folgenglieder  $a_n$  schließlich größer als die Zahl  $s$  werden, nämlich ab dem Index  $n_0$ . Das gilt aber für jedes  $s \in \mathbb{R}$ , egal wie groß. Dabei wird  $n_0$  größer, je größer  $s$  ist, d.h. „man muss länger warten“ bis  $a_n \geq s$  gilt. Da  $s$  also beliebig ist, wächst somit  $a_n$  *über jede Grenze*  $s$ , konvergiert also (uneigentlich) gegen  $\infty$ . Bei (2.4) muss man an große negative Werte für  $s$  denken, etwa  $s = -100, -1000, \dots$  Ungleichung (2.4) sagt daher, dass  $a_n$  kleiner wird als jede (große) negative Grenze  $s$ ; d.h.  $a_n$  konvergiert uneigentlich nach  $-\infty$ .

**Beachte:** Konvergiert eine Folge uneigentlich gegen  $\pm\infty$ , dann ist die Folge *nicht* konvergent, sondern divergent (nämlich bestimmt divergent). Bei konvergenten Folgen ist der Grenzwert immer eine endliche reelle Zahl!

**Beispiel 2.7.2** Einfache Beispiele für bestimmte Divergenz sind etwa

$$\lim_{n \rightarrow \infty} n^2 = \infty, \quad \lim_{n \rightarrow \infty} \sqrt{n} = \infty, \quad \lim_{n \rightarrow \infty} (-n) = -\infty.$$

Dass die Folge  $(n^2)_{n \in \mathbb{N}}$  über jede Grenze wächst, haben wir schon in Beispiel 2.2.1 gesehen. Bei  $\sqrt{n}$  muss man etwas genauer hinsehen: Die Wurzelfunktion wächst ja immer langsamer. Warum konvergiert  $\sqrt{n}$  trotzdem nach unendlich? Wir überprüfen dazu einfach die Ungleichung (2.3). Damit  $\sqrt{n} \geq s$  gilt (bei  $s \geq 0$ ), muss  $n \geq s^2$  sein:

$$a_n = \sqrt{n} \geq s \quad \Leftrightarrow \quad n \geq s^2.$$

D.h. in (2.3) kann man für  $n_0$  die kleinste natürliche Zahl größer (oder gleich)  $s^2$  wählen.

Wie bereits gesagt, sind die Rechenregeln für Grenzwerte (Satz 2.5.1) bei uneigentlicher Konvergenz nach  $\pm\infty$  nicht ohne weiteres anwendbar. Bei Grenzwerten mit  $\pm\infty$  gelten folgende Regeln:

**Satz 2.7.3** (a) Für eine Folge  $(a_n)$  mit  $a_n \neq 0$  gilt

$$\begin{aligned} |a_n| \rightarrow \infty &\iff \frac{1}{a_n} \rightarrow 0 \\ a_n \rightarrow 0 &\iff \left| \frac{1}{a_n} \right| \rightarrow \infty \end{aligned}$$

(b) Ist  $(a_n)$  bestimmt divergent nach  $\infty$  und  $(b_n)$  konvergent, d.h. gilt  $a_n \rightarrow \infty$  und  $b_n \rightarrow b \in \mathbb{R}$ , dann folgt

$$\begin{aligned} a_n + b_n &\rightarrow \infty \\ a_n b_n &\rightarrow \begin{cases} \infty, & \text{falls } b > 0 \\ -\infty, & \text{falls } b < 0 \end{cases} \end{aligned}$$

Im Fall  $b = 0$  ist für  $a_n b_n$  keine Aussage möglich.

(c) Wenn  $a_n \rightarrow \infty$  und  $b_n \rightarrow \infty$ , dann

$$a_n + b_n \rightarrow \infty \quad \text{und} \quad a_n b_n \rightarrow \infty.$$

Die Regeln in (a) ergeben sich aus der Eigenschaft, dass der Bruch  $\frac{1}{x}$  um so kleiner wird je größer der Betrag von  $x$  wird, und umgekehrt. Insbesondere folgt aus  $\frac{1}{a_n} \rightarrow 0$  erst einmal nur, dass  $|a_n| \rightarrow \infty$ . Ob dann auch  $a_n \rightarrow \infty$  oder  $a_n \rightarrow -\infty$  gilt, hängt vom Vorzeichen der  $a_n$  ab. Es kann auch vorkommen, dass zwar  $|a_n| \rightarrow \infty$  gilt, aber weder  $a_n \rightarrow \infty$  noch  $a_n \rightarrow -\infty$ ! Ein Beispiel dafür ist die Folge  $a_n = (-1)^n n$ .

Den Teil (b) kann man anschaulich folgendermaßen verstehen: Die Konvergenz  $b_n \rightarrow b \in \mathbb{R}$  impliziert, dass für große  $n$  die Folgenglieder  $b_n$  ungefähr gleich der Konstanten  $b$  sind. Bei  $a_n + b_n$  wird also zu  $a_n$  noch ungefähr die Konstante  $b$  addiert; im Fall  $b < 0$  heißt das übrigens, dass effektiv subtrahiert wird! Da aber  $b$  konstant ist und  $a_n$  über jede Grenze wächst, wächst in jedem Fall auch  $a_n + b_n$  über jede Grenze. Beim Produkt  $a_n b_n$  wird entsprechend für große  $n$  ungefähr mit der Konstanten  $b$  multipliziert.  $b$  ist ungleich 0, könnte aber sehr klein sein, etwa  $b = 10^{-6}$ . Da aber  $a_n$  über jede Grenze wächst, wächst schließlich auch  $a_n b_n$  über jede Grenze, jedenfalls der Betrag davon. Ist  $b > 0$ , so ist auch  $a_n b_n > 0$  (für große  $n$ ), denn  $a_n > 0$  weil  $a_n \rightarrow \infty$  und  $b_n > 0$  weil  $b_n \rightarrow b > 0$ . Damit folgt dann  $a_n b_n \rightarrow \infty$ . Für  $b < 0$  ist  $b_n$  für große  $n$  negativ, und damit auch  $a_n b_n < 0$  und es folgt  $a_n b_n \rightarrow -\infty$ .

Analog zu (b) und (c) gilt natürlich auch<sup>5</sup>

$$\begin{aligned} a_n \rightarrow -\infty \wedge b_n \rightarrow b \in \mathbb{R} &\implies a_n + b_n \rightarrow -\infty, \\ a_n \rightarrow -\infty \wedge b_n \rightarrow -\infty &\implies a_n + b_n \rightarrow -\infty. \end{aligned}$$

**Beispiel 2.7.4** Wir berechnen die Grenzwerte der beiden Folgen

$$a_n = n + \frac{n+1}{n-1}, \quad b_n = \frac{n^2}{1-n},$$

indem wir die Folgen wieder so geschickt zerlegen und kürzen, dass wir die Grenzwertsätze 2.5.1 und 2.7.3 benutzen können:

<sup>5</sup>Das logische Symbol  $\wedge$  bedeutet „und“.

(a)

$$\lim_{n \rightarrow \infty} \left( n + \frac{n+1}{n-1} \right) = \lim_{n \rightarrow \infty} \left( n + \underbrace{\frac{1 + \frac{1}{n}}{1 - \frac{1}{n}}}_{\rightarrow 1} \right) = \infty$$

Hier haben wir nach der Umformung also die Summe aus der Folge  $n$ , die nach  $\infty$  divergiert, und dem Bruch, der gegen 1 konvergiert. Wir können also 2.7.3(a) benutzen.

(b)

$$\lim_{n \rightarrow \infty} \frac{n^2}{1-n} = \lim_{n \rightarrow \infty} \frac{n^2}{n(\frac{1}{n} - 1)} = \lim_{n \rightarrow \infty} n \cdot \underbrace{\frac{1}{\frac{1}{n} - 1}}_{\rightarrow -1} = -\infty$$

Hier haben wir 2.7.3(b) mit  $b = -1 < 0$  benutzt.

**Achtung!**

Die Regeln 2.7.3(b) für  $a_n b_n$  bei  $a_n \rightarrow \infty$ ,  $b_n \rightarrow 0$  sowie (c) im Fall  $a_n \rightarrow \infty$ ,  $b_n \rightarrow -\infty$  liefern keine Aussagen. Anders gesagt, die Ausdrücke

$$\infty \cdot 0 \quad \text{und} \quad \infty - \infty$$

sind unbestimmt. Das Ergebnis kann  $\infty$ ,  $0$ ,  $-\infty$  oder irgendeine reelle Zahl sein! Beispiele für verschiedene Ergebnisse im Fall „ $\infty \cdot 0$ “ sind:

$$\begin{aligned} \lim_{n \rightarrow \infty} n \cdot \frac{1}{n^2} &= \lim_{n \rightarrow \infty} \frac{1}{n} = 0 \\ \lim_{n \rightarrow \infty} n^2 \cdot \frac{1}{n} &= \lim_{n \rightarrow \infty} n = \infty \\ \lim_{n \rightarrow \infty} 2n \cdot \frac{1}{n} &= \lim_{n \rightarrow \infty} 2 = 2 \end{aligned}$$

Beachten Sie, dass das Produkt vor dem Kürzen immer von der Art „ $\infty \cdot 0$ “ ist. Außerdem ist der Faktor 2 im letzten Beispiel beliebig, man kann hier auch jede andere Zahl ungleich Null nehmen.

## 2.8 Grenzwerte und Ungleichungen

Seien  $(a_n)$  und  $(b_n)$  konvergente Folgen. Gilt  $a_n \leq b_n$  für alle  $n \in \mathbb{N}$ , d.h. ist die Folge  $(a_n)$  immer kleiner als die Folge  $(b_n)$ , dann folgt

$$\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n. \quad (2.5)$$

**Beachte:**

- (a) Die Ungleichung (2.5) folgt auch dann, wenn  $a_n \leq b_n$  erst ab einem Index  $n_0 \in \mathbb{N}$  gilt, d.h.

$$a_n \leq b_n \text{ für alle } n \geq n_0 \quad \implies \quad \lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n$$

- (b) Die Aussage gilt *nicht* für „ $<$ “:

$$a_n < b_n \quad \not\Rightarrow \quad \lim_{n \rightarrow \infty} a_n < \lim_{n \rightarrow \infty} b_n$$

Das zeigt das nächste Beispiel.

**Beispiel 2.8.1** Sei

$$a_n = 1 - \frac{1}{n}, \quad b_n = 1 + \frac{1}{n} \quad (n \geq 1)$$

Es gilt dann  $a_n \leq b_n$  und

$$\lim_{n \rightarrow \infty} a_n = 1 \leq \lim_{n \rightarrow \infty} b_n = 1.$$

Zwar gilt auch  $a_n < b_n$ , jedoch ist

$$1 = \lim_{n \rightarrow \infty} a_n \not< \lim_{n \rightarrow \infty} b_n = 1.$$

## 2.9 Grenzwerte und stetige Funktionen

In diesem Abschnitt betrachten wir, wie man Grenzwerte in Funktionen auswerten kann.

Sei  $f : D \rightarrow \mathbb{R}$  eine Funktion mit Definitionsbereich  $D \subset \mathbb{R}$ . Im Kurs Mathematik A haben wir die Eigenschaft der *Stetigkeit* einer Funktion kennengelernt. Wir haben gesehen, dass Funktionen, die aus Summen, Produkten, Quotienten und Verkettungen der elementaren Funktionen wie  $x^n$ ,  $\sqrt{x}$ ,  $e^x$ ,  $\sin x$  zusammengesetzt sind, immer stetig sind. Anschaulich bedeutet Stetigkeit, dass der Graph von  $f$  außer an Definitionslücken keine Sprungstellen hat.

Grenzwerte und Stetigkeit hängen folgendermaßen zusammen:

**Satz 2.9.1** Die Funktion  $f$  ist stetig genau dann wenn, für jede konvergente Folge  $(a_n)$  in  $D$  (d.h.  $a_n \in D$ ) mit Grenzwert  $\lim_{n \rightarrow \infty} a_n = a \in D$  gilt

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

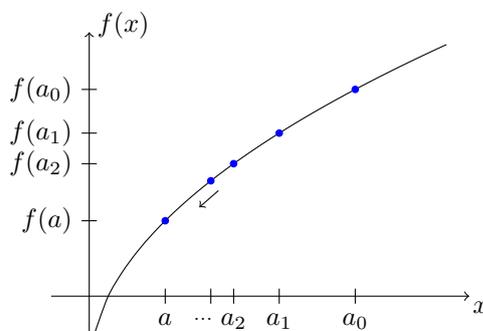
Wenn also  $a_n$  gegen  $a$  konvergiert und man die Folge in die Funktion  $f(x)$  einsetzt, so konvergiert die Folge der Funktionswerte  $f(a_n)$  gegen  $f(a)$ :

$$a_n \rightarrow a \implies f(a_n) \rightarrow f(a) \quad \text{für } n \rightarrow \infty,$$

oder auch

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

d.h. man darf den Grenzwert in die Funktion hineinziehen. Anschaulich bedeutet das:



Die Grafik zeigt, wie sich die Folgenglieder  $a_n$  immer mehr dem Grenzwert  $a$  annähern. Da der Graph von  $f$  keine Sprungstellen hat, müssen sich dann auch die Funktionswerte  $f(a_n)$  auf der Kurve immer mehr an  $f(a)$  annähern.

**Beispiel 2.9.2** Die Funktionen

$$\begin{aligned} f(x) &= e^x \quad \text{mit Definitionsbereich } D = \mathbb{R}, \\ f(x) &= \sqrt{x} \quad \text{mit } D = \{x \in \mathbb{R} \mid x \geq 0\} = [0, \infty[ \end{aligned}$$

sind beide stetig. Wir betrachten die Folge  $a_n = \frac{2n}{n-1}$ ,  $n \geq 2$ . Die Folge hat den Grenzwert

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{2n}{n-1} = 2.$$

Damit ergibt sich für die Folgen der Funktionswerte als Grenzwert

$$\lim_{n \rightarrow \infty} e^{\frac{2n}{n-1}} = e^2 \quad \text{und} \quad \lim_{n \rightarrow \infty} \sqrt{\frac{2n}{n-1}} = \sqrt{2}.$$

## 2.10 Die geometrische Folge

Sei  $q \in \mathbb{R}$  eine feste Zahl. Die Folge der Potenzen  $q^0, q^1, q^2, \dots$  heißt *geometrische Folge*:

$$(q^n)_{n \in \mathbb{N}} = (q^0, q^1, q^2, q^3, \dots) = (1, q, q^2, q^3, \dots).$$

Sie spielt in den folgenden Kapiteln (und überhaupt in der Mathematik!) eine sehr wichtige Rolle.

Hier interessiert uns zuerst einmal das *Konvergenzverhalten* der Folge, d.h. die Frage, ob und wann die Folge konvergiert. Für zwei einfache Beispiele können wir das sofort sagen: Die geometrische Folge mit  $q = 2$ , also

$$(2^n)_{n \in \mathbb{N}} = (1, 2, 4, 8, 16, \dots)$$

haben wir schon in Beispiel 2.4.1 untersucht und gesehen, dass Sie unbeschränkt und deswegen divergent ist. Da die Folge immer weiter wächst, divergiert sie also bestimmt nach Unendlich,

$$\lim_{n \rightarrow \infty} 2^n = \infty.$$

Dieses Ergebnis können wir jetzt benutzen, um auch den Grenzwert der geometrischen Folge mit  $q = \frac{1}{2}$  zu berechnen. Die Folge ist

$$\left(\left(\frac{1}{2}\right)^n\right)_{n \in \mathbb{N}} = \left(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots\right) = \left(\frac{1}{2^n}\right)_{n \in \mathbb{N}},$$

also genau der Kehrwert der Folge  $(2^n)$ . Weil  $2^n \rightarrow \infty$  folgt also (Satz 2.7.3(a))

$$\lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n = 0.$$

Genauso kann man natürlich argumentieren, wenn  $q = 3$ ,  $q = \frac{1}{3}$ ,  $q = 4$ ,  $q = \frac{1}{4}, \dots$  ist. Der Fall  $q = 1$  ist nicht sehr interessant, da die Folge dann natürlich

konstant gleich 1 ist und somit 1 auch der Grenzwert. Was passiert aber wenn  $q$  sehr nah bei 1 liegt, etwa  $q = 1.00001$  oder  $q = 0.99999$ ? (Probieren Sie es aus! Berechnen Sie die ersten Glieder der Folge  $q^n$  für diese Werte von  $q$ .) Die allgemeine Antwort gibt der folgende Satz.

**Satz 2.10.1** Die geometrische Folge  $(q_n)_{n \in \mathbb{N}}$  hat den Grenzwert

$$\lim_{n \rightarrow \infty} q^n = \begin{cases} 0 & \text{falls } -1 < q < 1, \\ 1 & \text{falls } q = 1, \\ \infty & \text{falls } q > 1. \end{cases}$$

Für  $q \leq -1$  ist  $(q^n)_{n \in \mathbb{N}}$  divergent.

**Beweis. 1. Fall  $q > 1$ :** Zu zeigen ist: für jedes  $s \in \mathbb{R}$  gibt es ein  $n_0 \in \mathbb{N}$  so dass

$$q^n \geq s \quad \text{wenn } n \geq n_0.$$

Wir müssen also die  $n$  berechnen, für die  $q^n \geq s$  gilt. Dazu formen wir die Ungleichung äquivalent um:

$$q^n \geq s \Leftrightarrow \ln(q^n) \geq \ln s \Leftrightarrow n \cdot \ln q \geq \ln s \Leftrightarrow n \geq \frac{\ln s}{\ln q}$$

Dabei haben wir im ersten Schritt benutzt, dass die Logarithmus-Funktion monoton wachsend ist, und im letzten Schritt, dass  $\ln q > 0$  weil  $q > 1$ ; beim Teilen durch  $\ln q$  dreht sich deswegen das Ungleichungszeichen nicht um! Wenn wir jetzt für  $n_0$  eine natürliche Zahl mit  $n_0 \geq \frac{\ln s}{\ln q}$  wählen, so folgt also tatsächlich  $n \geq n_0 \Rightarrow q^n \geq s$ , und damit gilt

$$\lim_{n \rightarrow \infty} q^n = \infty.$$

**2. Fall  $-1 < q < 1$ ,  $q \neq 0$ :** Diesen Fall können wir auf den ersten Fall zurückführen, indem wir den Kehrwert betrachten.  $-1 < q < 1$  bedeutet  $|q| < 1$  und damit gilt  $\frac{1}{|q|} > 1$ . Wir können daher auf die Folge  $(\frac{1}{|q|})^n$  den Fall 1 anwenden:

$$\left(\frac{1}{|q|}\right)^n \rightarrow \infty.$$

Es folgt dann

$$\left(\frac{1}{|q|}\right)^n = \frac{1}{|q^n|} \rightarrow \infty$$

und damit  $q^n \rightarrow 0$  (Satz 2.7.3(a)).

**3. Fall  $q = -1$ :** Die Folge

$$((-1)^n)_{n \in \mathbb{N}} = (1, -1, 1, -1, \dots)$$

ist divergent. (Sie hat die zwei Häufungswerte 1 und  $-1$ )

**4. Fall  $q < -1$ :** Nach Fall 1 gilt

$$|q^n| = |q|^n \rightarrow \infty$$

denn  $|q| > 1$ . Die Folge ist daher unbeschränkt und somit divergent.

Die restlichen Fälle  $q = 0$  und  $q = 1$  sind klar. □

**Beachte:** Bei  $q < -1$  ist  $(q^n)$  divergent, aber *nicht* bestimmt divergent (nach  $\infty$  oder  $-\infty$ ). Weil  $q$  negativ ist, hat nämlich  $q^n$  ein alternierendes Vorzeichen, z.B.

$$\left((-2)^n\right)_{n \in \mathbb{N}} = (1, -2, 4, -8, 16, -32, \dots).$$

**Beispiel 2.10.2** Auch bei Grenzwerten, die geometrische Folgen enthalten, hilft es oft, die größte Potenz auszuklammern. Im nächsten Beispiel ist die höchste Potenz im Nenner  $3^n$ .

$$\lim_{n \rightarrow \infty} \frac{3^{n+2}}{2^n + 3^n} = \lim_{n \rightarrow \infty} \frac{3^n \cdot 3^2}{3^n \left( \frac{2^n}{3^n} + 1 \right)} = \lim_{n \rightarrow \infty} \frac{3^2}{\underbrace{\left( \frac{2}{3} \right)^n}_{\rightarrow 0} + 1} = 9$$

Im letzten Schritt haben wir benutzt, dass  $\left(\frac{2}{3}\right)^n$  eine geometrische Folge mit  $q = \frac{2}{3} < 1$  ist, und daher gegen Null konvergiert.

## 2.11 Kovergenzkriterien

Bei unseren bisherigen Untersuchungen zu Folgen wussten wir immer konkret, was der Grenzwert einer gegebenen Folge ist: entweder konnten wir ihn mit Rechenregeln direkt ausrechnen, oder wir hatten schon vorher eine Vermutung, was der Grenzwert ist und konnten damit dann die Konvergenz zeigen. Jetzt lernen wir Kriterien kennen, mit denen man zeigen kann, dass eine Folge konvergiert *ohne* den Grenzwert schon vorab zu kennen oder auszurechnen, sogenannte *Konvergenzkriterien*.

Damit wird es dann möglich, Zahlen durch Grenzwerte von Folgen zu definieren. Z.B. kann man die Zahlen  $\pi$  oder  $e$  analytisch durch Grenzwerte definieren. Der Punkt ist hier: um eine Zahl als Grenzwert einer Folge definieren zu können, muss man wissen, dass die Folge konvergent ist, denn sonst hat sie ja keinen Grenzwert. Aber natürlich können wir die Konvergenz nicht durch Ausrechnen des Grenzwerts zeigen, denn dafür bräuchten wir ja die Zahl als Ergebnis schon, die wir aber durch den Grenzwert erst definieren wollen. Wir brauchen also „Konvergenz ohne (expliziten) Grenzwert“.

Als Konvergenzkriterien lernen wir das sogenannte „Sandwich-Lemma“, das Monotonie- und das Cauchy-kriterium kennen. In Beispiel 2.14.2 werden wir die Zahl  $e$  als Grenzwert einer Folge definieren.

## 2.12 Sandwich-Lemma

Der Name „Sandwich-Lemma“ deutet an, dass eine Folge  $(b_n)$  zwischen zwei anderen Folgen  $(a_n)$  und  $(c_n)$  „eingequetscht“ wird.<sup>6</sup>

**Satz 2.12.1 (Sandwich-Lemma)** *Seien  $(a_n), (b_n), (c_n)$  Folgen so dass gilt*

$$a_n \leq b_n \leq c_n \quad \text{für alle } n \in \mathbb{N}$$

<sup>6</sup>Beim Sandwich-Lemma erhält man zwar als Ergebnis auch den Grenzwert der Folge  $b_n$ , trotzdem zählen wir es zu den Konvergenzkriterien, da man nur die Grenzwerte der äußeren Folgen  $a_n$  und  $c_n$  vorab kennen muss.

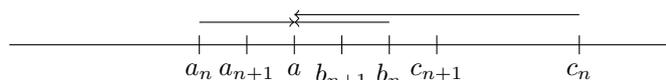
(oder alle  $n \geq n_0$  ab einem festen  $n_0 \in \mathbb{N}$ ). Sind die äußeren Folgen  $(a_n)$  und  $(c_n)$  konvergent und haben denselben Grenzwert, d.h.

$$\lim_{n \rightarrow \infty} a_n = a = \lim_{n \rightarrow \infty} c_n,$$

dann ist auch  $(b_n)$  konvergent und

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

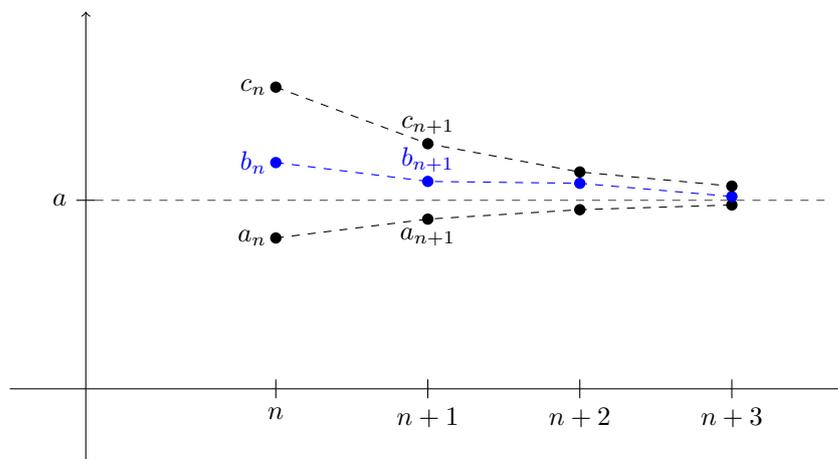
Man kann das folgendermaßen grafisch veranschaulichen:



$b_n$  liegt zwischen  $a_n$  und  $c_n$ ,  $b_{n+1}$  zwischen  $a_{n+1}$  und  $c_{n+1}$ , u.s.w. Wenn sich nun die äußeren Folgen  $a_n$  und  $c_n$  immer mehr an  $a$  annähern,  $b_n$  aber dazwischen liegt, bleibt natürlich der Folge  $b_n$  nicht anderes übrig, als sich ebenfalls immer mehr an  $a$  anzunähern. Schematisch kann man das auch so schreiben:

$$\begin{array}{ccc} a_n & \leq & b_n \leq c_n \\ \swarrow & & \searrow \\ n \rightarrow \infty & & \\ & & a \end{array}$$

Oder als Graphen im Koordinatensystem:



Als erste Folgerung aus dem Sandwich-Lemma erhalten wir eine sehr praktische Aussage über Nullfolgen:

**Satz 2.12.2** Ist die Folge  $(a_n)$  beschränkt und  $(b_n)$  eine Nullfolge (d.h.  $b_n \rightarrow 0$ ), dann ist auch die Produktfolge  $(a_n b_n)$  eine Nullfolge, d.h.  $a_n b_n \rightarrow 0$ . Anders gesagt: aus

$$|a_n| \leq M \text{ für alle } n \in \mathbb{N} \quad (M \in \mathbb{R} \text{ fest}) \quad \text{und} \quad \lim_{n \rightarrow \infty} b_n = 0$$

folgt

$$\lim_{n \rightarrow \infty} a_n b_n = 0.$$

**Beweis.** Es gilt

$$0 \leq |a_n b_n| = |a_n| \cdot |b_n| \leq M|b_n|,$$

die Folge  $|a_n b_n|$  liegt also zwischen 0 und der Folge  $M|b_n|$ . Wegen  $b_n \rightarrow 0$  gilt  $M|b_n| \rightarrow 0$ . Aus dem Sandwich-Lemma folgt dann  $|a_n b_n| \rightarrow 0$  und damit  $a_n b_n \rightarrow 0$ .  $\square$

**Beispiel 2.12.3** Den Grenzwert von

$$a_n = \frac{\cos(n)}{n}$$

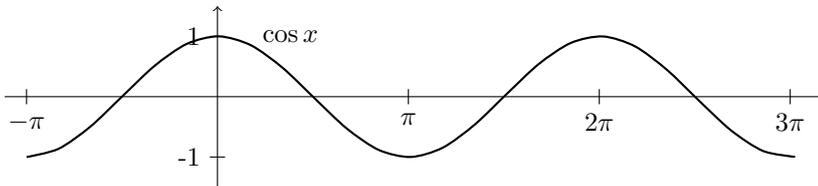
können wir mit dem Grenzwertsatz 2.5.1 nicht ausrechnen, denn die Folge  $\cos(n)$  konvergiert nicht. Da der Cosinus aber nur Werte zwischen  $-1$  und  $1$  annehmen kann, also beschränkt ist, und  $\frac{1}{n} \rightarrow 0$ , können wir  $a_n$  als Produkt einer beschränkten Folge mit einer Nullfolge schreiben und bekommen damit

$$\frac{\cos(n)}{n} = \underbrace{\cos(n)}_{\text{beschränkt}} \cdot \underbrace{\frac{1}{n}}_{\rightarrow 0} \rightarrow 0$$

also

$$\lim_{n \rightarrow \infty} \frac{\cos(n)}{n} = 0.$$

Die Beschränktheit sieht man am Graph der Cosinus-Funktion:



Es gilt

$$-1 \leq \cos x \leq 1$$

d.h.

$$|\cos x| \leq 1 \quad \text{für alle } x.$$

## 2.13 Die Folgen $\sqrt[n]{n}$ und $\sqrt[n]{p}$

Über die Grenzwerte der Folgen  $(\sqrt[n]{n})_{n \geq 1}$  und  $(\sqrt[n]{p})_{n \geq 1}$  für festes  $p > 0$  können wir bis jetzt nichts aussagen. Insbesondere  $\sqrt[n]{n} = n^{1/n}$  ist schwierig zu behandeln, da man den Grenzübergang  $n \rightarrow \infty$  für beide  $n$  zugleich machen muss, siehe auch die Warnung am Ende dieses Abschnitts.

Eine Berechnung der ersten Werte von  $\sqrt[n]{n}$  liefert

$$\begin{aligned} \sqrt[1]{1} &= 1, & \sqrt[2]{2} &= 1.4142\dots, & \sqrt[3]{3} &= 1.4422\dots, & \sqrt[4]{4} &= 1.4142\dots, \\ \sqrt[5]{5} &= 1.3797\dots, & \dots, & & \sqrt[10]{10} &= 1.2589\dots, & \dots \end{aligned}$$

zeigt, dass die Folge erst wächst, dann aber wieder langsam fällt. Wie geht es weiter? Wird die Folge sich wieder an 1 annähern oder schon vorher „stoppen“,

etwa bei 1.1 oder 1.2? Oder wird die Folge sogar kleiner als 1 werden? Um durch Rechnung eine Vermutung für den Grenzwert zu bekommen, muss man hier deutlich größere Werte für  $n$  einsetzen. (Probieren Sie es aus!)

Schließlich kommt man zur Vermutung  $\sqrt[n]{n} \rightarrow 1$ . Dass das tatsächlich richtig ist, werden wir jetzt mathematisch exakt mit dem Sandwich-Lemma beweisen. Den Grenzwert von  $\sqrt[p]{p}$  bekommen wir dann ebenfalls als Folgerung aus dem Sandwich-Lemma.

**Satz 2.13.1** *Es gilt*

$$(a) \lim_{n \rightarrow \infty} \sqrt[n]{n} = 1,$$

$$(b) \lim_{n \rightarrow \infty} \sqrt[p]{p} = 1 \text{ für jedes } p > 0.$$

**Beweis.**

(a) Zunächst gilt

$$\sqrt[n]{n} \geq 1 \quad \text{für alle } n \geq 1. \quad (2.6)$$

Denn  $x = \sqrt[n]{n}$  ist die positive Lösung der Gleichung  $x^n = n$ . Wäre  $x < 1$ , so auch  $x^n < 1$ , aber  $x^n = n \geq 1$ . Daher muss  $x = \sqrt[n]{n} \geq 1$  gelten.

Statt jetzt  $\sqrt[n]{n} \rightarrow 1$  zu zeigen, betrachten wir die Folge

$$a_n = \sqrt[n]{n} - 1$$

und beweisen dann  $a_n \rightarrow 0$ , siehe Eigenschaft (i) in Abschnitt 2.4. Es gilt

$$\sqrt[n]{n} = a_n + 1.$$

Durch Potenzieren folgt daraus mit dem binomischen Satz 1.6.1

$$n = (1 + a_n)^n = 1 + \binom{n}{1}a_n + \binom{n}{2}a_n^2 + \dots + a_n^n$$

Alle Terme auf der rechten Seite sind positiv, denn wegen (2.6) ist  $a_n \geq 0$ . Wir lassen nun alle Terme bis auf 1 und den Term mit  $a_n^2$  weg, wodurch die rechte Seite kleiner wird und wir

$$n \geq 1 + \binom{n}{2}a_n^2 = 1 + \frac{n(n-1)}{2}a_n^2$$

bekommen. Auflösen nach  $a_n$  liefert

$$n - 1 \geq \frac{n(n-1)}{2}a_n^2 \quad \Leftrightarrow \quad \frac{2}{n} \geq a_n^2 \quad \Leftrightarrow \quad a_n \leq \sqrt{\frac{2}{n}}$$

Es gilt somit

$$0 \leq a_n \leq \sqrt{\frac{2}{n}} \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

und aus dem Sandwich-Lemma folgt  $a_n \rightarrow 0$  also  $\sqrt[n]{n} \rightarrow 1$ .

(b) Sei zuerst  $p \geq 1$ . Für  $n \geq p$  gilt dann

$$1 \leq \sqrt[n]{p} \leq \sqrt[n]{n}.$$

Da  $\sqrt[n]{n} \rightarrow 1$  folgt mit dem Sandwich-Lemma  $\sqrt[n]{p} \rightarrow 1$ .

Sei jetzt  $0 < p < 1$ . Hier können wir wieder zum Kehrwert übergehen: Da  $\frac{1}{p} > 1$  ist, gilt nach dem ersten Fall ( $p > 1$ )

$$\frac{1}{\sqrt[n]{p}} = \sqrt[n]{\frac{1}{p}} \rightarrow 1$$

und somit  $\sqrt[n]{p} \rightarrow 1$ .

□

Eine andere Möglichkeit den Grenzwert  $\lim_{n \rightarrow \infty} \sqrt[n]{p}$  zu berechnen, ist es,  $\sqrt[n]{p} = p^{\frac{1}{n}}$  zu schreiben und dann den Grenzwert in der Funktion  $f(x) = p^x$  auszuwerten:

$$\sqrt[n]{p} = p^{\frac{1}{n}} \rightarrow p^0 = 1 \quad \text{für } n \rightarrow \infty.$$

Hier benutzt man dann, dass die Funktion  $f(x) = p^x$  für  $p > 0$  stetig ist.

**Beispiel 2.13.2** Wir wollen den Grenzwert  $\lim_{n \rightarrow \infty} \sqrt[n]{n^2 + n}$  berechnen. Zunächst klammert man dazu unter der Wurzel wieder die höchste Potenz aus:

$$\sqrt[n]{n^2 + n} = \sqrt[n]{n^2 \left(1 + \frac{1}{n}\right)} = \sqrt[n]{n^2} \cdot \sqrt[n]{1 + \frac{1}{n}} = (\sqrt[n]{n})^2 \cdot \sqrt[n]{1 + \frac{1}{n}}.$$

Der erste Faktor konvergiert gegen  $1^2$  und für den zweiten Faktor benutzt man das Sandwich-Kriterium:

$$1 = \sqrt[n]{1} \leq \sqrt[n]{1 + \frac{1}{n}} \leq \sqrt[n]{2} \rightarrow 1 \quad \implies \quad \sqrt[n]{1 + \frac{1}{n}} \rightarrow 1 \quad \text{bei } n \rightarrow \infty.$$

Damit folgt also insgesamt

$$\lim_{n \rightarrow \infty} \sqrt[n]{n^2 + n} = \lim_{n \rightarrow \infty} \underbrace{(\sqrt[n]{n})^2}_{\rightarrow 1^2} \cdot \underbrace{\sqrt[n]{1 + \frac{1}{n}}}_{\rightarrow 1} = 1.$$

### Achtung!

Bei der Berechnung von  $\lim_{n \rightarrow \infty} \sqrt[n]{n}$  oder ähnlichen Grenzwerten, wie etwa dem im letzten Beispiel, ist es nicht erlaubt, den Grenzübergang  $n \rightarrow \infty$  für einzelnen  $n$  in mehreren Schritten nacheinander auszuführen. D.h. die Rechnung  $\sqrt[n]{n} = n^{1/n} \rightarrow n^0 = 1$  ist falsch! Sie ergibt hier zwar das richtige Ergebnis, das ist aber nur Zufall. Entsprechend hätte man sonst auch  $n \cdot \frac{1}{n} \rightarrow n \cdot 0 = 0$  rechnen können, was offensichtlich falsch ist.

Auch die Rechnung  $\sqrt[n]{n} = n^{\frac{1}{n}} \rightarrow \infty^0$  ist nicht erlaubt und liefert nicht das Ergebnis, denn der Ausdruck  $\infty^0$  ist, genau wie  $\infty \cdot 0$ , unbestimmt.

## 2.14 Monotoniekriterium

Genau wie bei Funktionen kann man auch bei Folgen Monotonie definieren: Die Folge  $(a_n)_{n \in \mathbb{N}}$  heißt

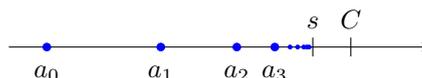
(i) *monoton wachsend* wenn  $a_n \leq a_{n+1}$  für alle  $n \in \mathbb{N}$ ,

(ii) *monoton fallend* wenn  $a_n \geq a_{n+1}$  für alle  $n \in \mathbb{N}$ .

Die Folge heißt *monoton*, wenn sie monoton wachsend oder fallend ist.

**Satz 2.14.1** *Ist die Folge  $(a_n)$  monoton und beschränkt, dann ist  $(a_n)$  konvergent.*

**Beweis.** Sei z.B.  $(a_n)$  monoton wachsend. Da  $(a_n)$  beschränkt ist, gibt es eine *obere Schranke*  $C$  von  $(a_n)$ , d.h. eine Zahl  $C \in \mathbb{R}$  so dass  $a_n \leq C$  für alle  $n \in \mathbb{N}$ . Es gibt dann viele (unendlich viele) obere Schranken von  $(a_n)$ , z.B. ist jede Zahl, die größer als  $C$  ist, auch eine obere Schranke. Genauso kann es aber auch obere Schranken von  $(a_n)$  geben, die kleiner als  $C$  sind. Sei nun  $s \in \mathbb{R}$  die *kleinste obere Schranke* von  $(a_n)$ . Anschaulich sieht das so aus:



Weil  $s \in \mathbb{R}$  die *kleinste* obere Schranke von  $(a_n)$  ist, muss nun  $(a_n)$  gegen  $s$  konvergieren: Denn einerseits bleiben alle  $a_n$  kleiner (oder gleich)  $s$ , da  $s$  obere Schranke ist. Andererseits müssen sich die  $a_n$  aber auch beliebig dicht an  $s$  annähern, denn sonst wäre „noch Platz“ zwischen den  $a_n$  und  $s$ , und dann wäre  $s$  nicht die kleinste obere Schranke.<sup>7</sup> Da die Folge  $(a_n)$  also wächst, sich  $s$  beliebig dicht annähert, und auch nicht größer werden kann, muss  $s$  der Grenzwert von  $(a_n)$  sein,  $\lim_{n \rightarrow \infty} a_n = s$ . Damit ist  $(a_n)$  konvergent.  $\square$

**Bemerkung:** Dass jede beschränkte Folge oder Menge in  $\mathbb{R}$  eine kleinste obere Schranke hat, ist eine grundlegende Eigenschaft der reellen Zahlen, genannt *Vollständigkeit*. Die rationalen Zahlen haben diese Eigenschaft nicht.

**Beispiel 2.14.2** Mithilfe des Monotoniekriteriums kann man die Eulersche Zahl  $e = 2.718281 \dots$  wie folgt als Grenzwert einer Folge definieren:

$$e := \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

Dazu zeigt man, dass die Folge

$$a_n = \left(1 + \frac{1}{n}\right)^n$$

monoton wachsend und beschränkt ist. (Das erfordert etwas Arbeit) Nach dem Monotoniekriterium ist die Folge dann konvergent, d.h. der Grenzwert  $\lim_{n \rightarrow \infty} a_n$  existiert, und man kann daher  $e$  als diesen Grenzwert definieren.

<sup>7</sup>Hier wird unser „Beweis“ etwas vage; für einen richtigen mathematischen Beweis würde man genauer argumentieren.

Warum die Folge  $a_n$  den richtigen Wert für  $e$  als Grenzwert ergibt, haben wir übrigens in Mathematik A als Beispiel für die Regel von l'Hospital zusammen mit Exponential- und Logarithmusfunktion gesehen. Hier in Mathematik B werden wir im Kapitel über Reihen noch einen anderen Weg kennenlernen, wie man  $e$  und die Exponentialfunktion definieren kann. Merke: In der Mathematik gibt es oft mehrere Wege, ein Ziel zu erreichen.

## 2.15 Cauchy Kriterium

Das Cauchy Kriterium ist das allgemeinste Kriterium, um die Konvergenz einer Folge zu zeigen ohne ihren Grenzwert zu kennen. Es ist sogar notwendig für die Konvergenz:

**Satz 2.15.1** *Die Folge  $(a_n)$  ist konvergent, genau dann wenn es zu jedem  $\varepsilon > 0$  ein  $n_0 \in \mathbb{N}$  gibt, sodass*

$$|a_n - a_m| < \varepsilon \quad \text{für alle } n, m \geq n_0. \quad (2.7)$$

Man beachte die Ähnlichkeit von (2.7) zur Abschätzung (2.1) bei der Definition des Grenzwerts. Dort bedeutet die Abschätzung, dass der Abstand der Folgenglieder  $a_n$  zum Grenzwert  $a$  beliebig klein wird (kleiner als jedes  $\varepsilon$ ). Hier sagt (2.7), dass der Abstand zwischen zwei Gliedern  $a_n$  und  $a_m$  beliebig klein wird ( $< \varepsilon$ ), wenn  $n$  und  $m$  groß werden (größer als  $n_0$ ).

Warum folgt nun aus (2.7) die Konvergenz der Folge  $(a_n)$ ? Dazu kann man sich zuerst überlegen, dass (2.7) die Beschränktheit der Folge impliziert. Der Satz 2.6.2(a) von Bolzano-Weierstraß sagt dann, dass die Folge mindestens einen Häufungswert hat. Jetzt benutzt man wieder (2.7) um zu zeigen, dass die Folge höchstens einen Häufungswert haben kann. Also hat sie genau einen Häufungswert und nach Satz 2.6.2(b) ist die Folge dann konvergent.



# Kapitel 3

## Reihen

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik B, Kapitel 2.

### 3.1 Unendliche Reihen

Eine *unendliche Reihe* ist eine unendliche Summe von Zahlen  $a_k \in \mathbb{R}$ :

$$\sum_{k=0}^{\infty} a_k = a_0 + a_1 + a_2 + a_3 + \dots \quad (3.1)$$

Um den Wert der Reihe, d.h. das Ergebnis der unendlichen Summation zu bestimmen, betrachtet man Teilsummen vom Anfang der Reihe:

$$\begin{aligned} s_0 &= a_0 \\ s_1 &= a_0 + a_1 \\ s_2 &= a_0 + a_1 + a_2 \\ &\vdots \\ s_n &= a_0 + a_1 + \dots + a_n = \sum_{k=0}^n a_k \end{aligned}$$

Je mehr Zahlen  $a_k$  man addiert, d.h. je größer  $n$  wird, desto mehr nähert man sich der Summe der ganzen Reihe. Formal erhält man dann den Wert der Reihe durch den Grenzübergang  $n \rightarrow \infty$ ,

$$\lim_{n \rightarrow \infty} s_n = a_0 + a_1 + a_2 + \dots = \sum_{k=0}^{\infty} a_k.$$

D.h. der Wert (bzw. die Summe) der Reihe  $\sum_{k=0}^{\infty} a_k$  ist der Grenzwert  $s$  der Teilsummen  $s_n$ ,

$$\sum_{k=0}^{\infty} a_k = s = \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k. \quad (3.2)$$

Die Reihe konvergiert, wenn die Folge  $(s_n)$  der Teilsumme konvergiert, d.h. wenn der Grenzwert  $s = \lim_{n \rightarrow \infty} s_n$  existiert. Man sagt dann auch, dass die Reihe gegen  $s$  konvergiert bzw. den Grenzwert  $s$  hat. Ist die Folge der Teilsummen  $(s_n)$  divergent, so *divergiert* die Reihe.

**Beachte:** Wichtig ist, zwischen der Reihe selbst und ihrem (Grenz-)Wert zu unterscheiden. Die Reihe selbst ist der Ausdruck der (formalen) unendlichen Summe in (3.1), der Wert der Reihe ist der Grenzwert  $s \in \mathbb{R}$  der Teilsummen in (3.2) (also eine Zahl). Für eine konvergente Reihe gilt (siehe (3.2))

$$\sum_{k=0}^{\infty} a_k = s,$$

d.h. der Grenzwert der Reihe wird wieder mit dem Symbol  $\sum_{k=0}^{\infty} a_k$  bezeichnet. Auch die Gleichung

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n a_k = \sum_{k=0}^{\infty} a_k,$$

gilt nur für eine konvergente Reihe (das ist wieder (3.2)). Bei einer divergenten Reihe existiert der Limes auf der linken Seite nicht und  $\sum_{k=0}^{\infty} a_k$  ist nur ein formaler Ausdruck, keine reelle Zahl.<sup>1</sup>

Wie bei Folgen ist auch bei Reihen der Start statt bei  $k = 0$  auch ab einem anderen Index  $k = k_0$  möglich, d.h.  $\sum_{k=k_0}^{\infty} a_k$ , etwa

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \dots$$

**Beispiel 3.1.1** Wir betrachten die Reihe  $\sum_{k=1}^{\infty} a_k$  mit  $a_k = \frac{1}{k^2}$ , also

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots$$

Berechnen wir zuerst die ersten Teilsummen:

$$\begin{aligned} s_1 &= a_1 = 1 && = 1 \\ s_2 &= a_1 + a_2 = 1 + \frac{1}{4} && = 1.25 \\ s_3 &= a_1 + a_2 + a_3 = 1 + \frac{1}{4} + \frac{1}{9} && = 1.361\dots \\ s_4 &= a_1 + a_2 + a_3 + a_4 = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} && = 1.423\dots \\ s_5 &= \sum_{k=1}^5 a_k = 1 + \frac{1}{4} + \dots + \frac{1}{25} && = 1.463\dots \\ s_6 &= \sum_{k=1}^6 a_k = 1 + \frac{1}{4} + \dots + \frac{1}{36} && = 1.491\dots \\ s_7 &= \sum_{k=1}^7 a_k = 1 + \frac{1}{4} + \dots + \frac{1}{49} && = 1.511\dots \end{aligned}$$

<sup>1</sup>Eine Ausnahme sind allerdings bestimmt divergente Reihen, siehe Abschnitt 3.5.

Man erkennt, dass die Folge der Teilsummen  $s_n$  monoton wachsend ist, denn es werden immer weitere positive Glieder  $\frac{1}{k^2}$  hinzu addiert. Gleichzeitig werden die Glieder mit größerem  $k$  aber schnell kleiner, d.h. es wird immer weniger hinzu addiert und  $s_n$  wächst immer langsamer. Die Frage ist nun: Wie weit wächst  $s_n$ ? Werden die  $s_n$  größer als 1.6, als 1.7, oder sogar größer als 2? Ist die Folge  $(s_n)$  vielleicht sogar unbeschränkt, d.h. werden die  $s_n$  schließlich größer als jede reelle Zahl?

Tatsächlich ist das nicht der Fall. Die Folge  $(s_n)$  ist beschränkt und damit dann auch konvergent. Im Grenzwert  $n \rightarrow \infty$  erhält man

$$\lim_{n \rightarrow \infty} s_n = \sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \dots = 1.644\dots$$

Da der Grenzwert  $\lim_{n \rightarrow \infty} s_n$  existiert, ist also die Reihe  $\sum_{k=1}^{\infty} \frac{1}{k^2}$  konvergent. Ihren genauen Grenzwert kann man mit der Theorie der Fourierreihen bestimmen<sup>2</sup>, es gilt

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \dots = \frac{\pi^2}{6}. \quad (3.3)$$

Wir werden im Abschnitt 3.5 zumindest die Beschränktheit der Teilfolgen  $s_n$ , und damit die Konvergenz der Reihe zeigen.

## 3.2 Die geometrische Reihe

Addiert man alle Glieder der geometrischen Folge  $(q^k)_{k \geq 0}$  auf, erhält man die *geometrische Reihe*

$$\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + q^3 + \dots$$

Wir untersuchen jetzt, für welche  $q \in \mathbb{R}$  die geometrische Reihe konvergiert und berechnen gleichzeitig den Wert der Reihe. Dazu betrachten wir wie im letzten Beispiel die Teilsummen  $s_n$ :

$$s_n = \sum_{k=0}^n q^k = 1 + q + \dots + q^n.$$

Wir können die Summe  $s_n$  mit einem Trick berechnen, der ein bisschen an den Trick von Gauß (Abschnitt 1.2) erinnert: Wir subtrahieren von  $s_n$  den Term

$$q \cdot s_n = q + q^2 + \dots + q^{n+1},$$

wodurch alle Potenzen  $q, \dots, q^n$  wegfallen:

$$s_n - qs_n = 1 + q + q^2 + \dots + q^n - (q + q^2 + \dots + q^n + q^{n+1}) = 1 - q^{n+1}.$$

Auf der linken Seite kann man nun  $s_n$  ausklammern und dann nach  $s_n$  auflösen:

$$\begin{aligned} (1 - q)s_n &= s_n - qs_n = 1 - q^{n+1} \\ \implies s_n &= \frac{1 - q^{n+1}}{1 - q} \quad \text{falls } q \neq 1. \end{aligned}$$

<sup>2</sup>Das ist ein Thema des Kurses Mathematik C.

Wir haben damit die wichtige Formel

$$\sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q} \quad \text{für } q \neq 1 \quad (3.4)$$

nachgerechnet. (Man beachte, dass die Formel nur für  $q \neq 1$  gilt, da man sonst durch 0 teilen würde!) Wir können jetzt den Grenzübergang  $n \rightarrow \infty$  durchführen: Wir wissen, dass die geometrische Folge  $q^n$  für  $|q| < 1$  (d.h.  $-1 < q < 1$ ) gegen Null konvergiert (Satz 2.10.1). Dann gilt auch  $q^{n+1} \rightarrow 0$  für  $n \rightarrow \infty$  und aus (3.4) folgt

$$\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{k=0}^n q^k = \lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}.$$

Für  $|q| < 1$  konvergiert also die geometrische Reihe und ihr Grenzwert ist

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1 - q} \quad \text{für } |q| < 1. \quad (3.5)$$

### 3.3 Rechenregeln für konvergente Reihen

Für konvergente Reihen gelten folgende einfache Regeln zur Addition zweier Reihen, Multiplikation mit einer Konstante und Änderung des Startindex der Reihe.

**Satz 3.3.1** Sind  $\sum_{k=0}^{\infty} a_k$  und  $\sum_{k=0}^{\infty} b_k$  konvergente Reihen, dann gilt:

$$\sum_{k=0}^{\infty} (a_k \pm b_k) = \sum_{k=0}^{\infty} a_k \pm \sum_{k=0}^{\infty} b_k \quad (3.6)$$

$$\sum_{k=0}^{\infty} (c \cdot a_k) = c \cdot \sum_{k=0}^{\infty} a_k \quad (c \in \mathbb{R} \text{ fest}) \quad (3.7)$$

$$\sum_{k=l}^{\infty} a_k = \left( \sum_{k=0}^{\infty} a_k \right) - a_0 - a_1 - \dots - a_{l-1} \quad (l \in \mathbb{N} \text{ fest}) \quad (3.8)$$

**Beweis.** Dies folgt sofort aus der Definition des Reihenwerts als Grenzwert der Teilsummen  $s_n$  und Anwendung der Grenzwertsätze 2.5.1 für Folgen. Hier ist zum Beispiel die Rechnung für (3.6):

$$\begin{aligned} \sum_{k=0}^{\infty} (a_k + b_k) &= \lim_{n \rightarrow \infty} \sum_{k=0}^n (a_k + b_k) = \lim_{n \rightarrow \infty} \left( \sum_{k=0}^n a_k + \sum_{k=0}^n b_k \right) \\ &\stackrel{2.5.1(a)}{=} \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k + \lim_{n \rightarrow \infty} \sum_{k=0}^n b_k = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k \end{aligned}$$

□

Wir erläutern die Regeln an einem Beispiel.

**Beispiel 3.3.2** Unser Ziel ist, die Reihe soweit umzuformen, dass wir die Formeln (3.3) und (3.5) für die Reihe  $\sum_{k=1}^{\infty} \frac{1}{k^2}$  und die geometrische Reihe anwenden können.

$$\begin{aligned} \sum_{k=1}^{\infty} \left( \frac{1}{4^k} - \frac{1}{\pi k^2} \right) &\stackrel{(3.6)}{=} \sum_{k=1}^{\infty} \frac{1}{4^k} - \sum_{k=1}^{\infty} \frac{1}{\pi k^2} \stackrel{(3.7)}{=} \sum_{k=1}^{\infty} \left( \frac{1}{4} \right)^k - \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{1}{k^2} \\ &\stackrel{(3.8)}{=} \sum_{k=0}^{\infty} \left( \frac{1}{4} \right)^k - \left( \frac{1}{4} \right)^0 - \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{1}{k^2} \\ &= \frac{1}{1 - \frac{1}{4}} - 1 - \frac{1}{\pi} \cdot \frac{\pi^2}{6} = \frac{4}{3} - 1 - \frac{\pi}{6} = \frac{1}{3} - \frac{\pi}{6} \end{aligned}$$

### 3.4 Konvergenzkriterien für Reihen

Wie schon bei Folgen gibt es auch bei Reihen verschiedene Kriterien, um die Konvergenz zu überprüfen, ohne dabei aber den Wert der Reihe berechnen zu müssen. Das folgende notwendige Kriterium besagt, dass für die Konvergenz der Reihe  $\sum_{k=0}^{\infty} a_k$  die Folge ihrer Glieder  $(a_k)$  eine Nullfolge sein muss.

**Satz 3.4.1 (Notwendiges Kriterium)** *Ist die Reihe  $\sum_{k=0}^{\infty} a_k$  konvergent, so gilt*

- (a)  $\lim_{k \rightarrow \infty} a_k = 0$ ,
- (b)  $\lim_{n \rightarrow \infty} \sum_{k=n+1}^{\infty} a_k = 0$ .

Hierbei ist

$$\sum_{k=n+1}^{\infty} a_k = a_{n+1} + a_{n+2} + a_{n+3} + \dots$$

der „Reihenrest“.

**Beweis.** Aus

$$a_n = \sum_{k=0}^n a_k - \sum_{k=0}^{n-1} a_k$$

folgt durch Grenzübergang  $n \rightarrow \infty$  sofort

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k - \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} a_k = \sum_{k=0}^{\infty} a_k - \sum_{k=0}^{\infty} a_k = 0.$$

Hierbei haben wir den Grenzwertsatz 2.5.1 auf die Differenz der beiden Summen angewandt, und dabei entscheidend die Konvergenz der Reihe benutzt, dass also die Teilsummen  $s_n = \sum_{k=0}^n a_k$  (und damit auch  $s_{n-1} = \sum_{k=0}^{n-1} a_k$ ) tatsächlich gegen  $\sum_{k=0}^{\infty} a_k$  konvergieren, vergleiche nochmal Abschnitt 3.1. Die zweite Aussage bekommt man ähnlich: Nach (3.8) gilt

$$\sum_{k=n+1}^{\infty} a_k = \sum_{k=0}^{\infty} a_k - \sum_{k=0}^n a_k$$

und bei  $n \rightarrow \infty$  folgt

$$\lim_{n \rightarrow \infty} \sum_{k=n+1}^{\infty} a_k = \sum_{k=0}^{\infty} a_k - \underbrace{\lim_{n \rightarrow \infty} \sum_{k=0}^n a_k}_{=\sum_{k=0}^{\infty} a_k} = 0.$$

□

Für eine konvergente Reihe haben wir jetzt eine Reihe von Konvergenzaussagen: Die Teilsummen  $s_n = \sum_{k=0}^n a_k$  konvergieren für  $n \rightarrow \infty$  gegen den Grenzwert der Reihe  $\sum_{k=0}^{\infty} a_k$ , die Glieder  $a_k$  der Reihe konvergieren gegen 0, und auch die Reihenreste  $\sum_{k=n+1}^{\infty} a_k$  konvergieren nach 0.

Das notwendige Kriterium 3.4.1(a) kann man benutzen, um die Divergenz einer Reihe zu zeigen: Konvergiert die Folge  $(a_k)$  nicht nach Null, dann kann die Reihe  $\sum_{k=0}^{\infty} a_k$  nicht konvergieren. Schauen wir uns das alles für konkrete Reihen an.

**Beispiel 3.4.2** (a) Wir betrachten die Reihe

$$\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k.$$

Das ist eine geometrische Reihe mit  $q = \frac{1}{2}$ . Da  $|q| = \frac{1}{2} < 1$ , ist die Reihe konvergent und mit (3.5) können wir ihren Grenzwert berechnen:

$$\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = \frac{1}{1 - \frac{1}{2}} = 2.$$

Die Folge der Glieder der Reihe ist  $a_n = \left(\frac{1}{2}\right)^n$ . Das ist die geometrische Folge mit  $q = \frac{1}{2} < 1$ , also eine Nullfolge (siehe Satz 2.10.1),

$$a_n = \left(\frac{1}{2}\right)^n \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Für die Teilsummen erhalten wir aus (3.4)

$$s_n = \sum_{k=0}^n \left(\frac{1}{2}\right)^k = \frac{1 - \left(\frac{1}{2}\right)^{n+1}}{1 - \frac{1}{2}} = 2 \left(1 - \left(\frac{1}{2}\right)^{n+1}\right) = 2 - \left(\frac{1}{2}\right)^n.$$

Für die ersten Wert von  $n$  erhalten wir:

$n$	0	1	2	3	4	...	$\rightarrow \infty$
$a_n$	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	...	$\rightarrow 0$
$s_n$	1	$\frac{3}{2}$	$\frac{7}{4}$	$\frac{15}{8}$	$\frac{31}{16}$	...	$\rightarrow 2$

Man sieht, wie die Reihenglieder  $a_n$  gegen 0, die Teilsummen  $s_n$  gegen 2, den Grenzwert der Reihe, konvergieren. Für den Reihenrest erhalten wir

$$\sum_{k=n+1}^{\infty} \left(\frac{1}{2}\right)^k = \underbrace{\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k}_2 - \underbrace{\sum_{k=0}^n \left(\frac{1}{2}\right)^k}_{2 - \left(\frac{1}{2}\right)^n} = \left(\frac{1}{2}\right)^n \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

der Reihenrest konvergiert gegen Null.

(b) Ist die Reihe

$$\sum_{k=1}^{\infty} \frac{k}{k+1}$$

konvergent oder divergent? Für die Glieder  $a_k$  der Reihe gilt

$$a_k = \frac{k}{k+1} \rightarrow 1 \neq 0 \quad \text{bei } k \rightarrow \infty.$$

Also ist  $(\frac{k}{k+1})$  keine Nullfolge und mit dem notwendigen Kriterium folgt, dass die Reihe  $\sum_{k=1}^{\infty} \frac{k}{k+1}$  divergent ist.

Als Anwendung des notwendigen Kriteriums untersuchen wir jetzt das komplette Konvergenzverhalten der geometrischen Reihe.

**Satz 3.4.3** Die geometrische Reihe  $\sum_{k=0}^{\infty} q^k$  ist konvergent falls  $|q| < 1$  und divergent für  $|q| \geq 1$ . Im konvergenten Fall gilt

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \quad (|q| < 1).$$

**Beweis.** Für  $|q| \geq 1$  ist  $(q^k)$  keine Nullfolge (Satz 2.10.1), also ist die Reihe divergent. Die Konvergenz und den Grenzwert  $\frac{1}{1-q}$  im Fall  $|q| < 1$  haben wir in 3.2 schon gezeigt.  $\square$

### 3.5 Reihen mit positiven Gliedern

Reihen mit positiven Gliedern, also Reihen  $\sum_{k=0}^{\infty} a_k$  mit  $a_k \geq 0$  für alle  $k \in \mathbb{N}$ , haben ein besonders einfaches Konvergenzverhalten: entweder ist die Reihe konvergent oder sie divergiert bestimmt nach  $\infty$ :

**Satz 3.5.1** Seien alle  $a_k \geq 0$ . Dann gilt:

(a)  $\sum_{k=0}^{\infty} a_k$  ist konvergent genau dann, wenn es ein  $M \in \mathbb{R}$  gibt mit

$$\sum_{k=0}^n a_k \leq M \quad \text{für alle } n \in \mathbb{N}, \quad (3.9)$$

d.h. wenn die Reihe beschränkt ist.

(b) Ist die Reihe nicht beschränkt, dann divergiert sie bestimmt nach  $\infty$ ,

$$\sum_{k=0}^{\infty} a_k = \infty.$$

**Beweis.** Aus  $a_k \geq 0$  folgt, dass die Folge der Teilsummen  $s_n = \sum_{k=0}^n a_k$  monoton wachsend ist. Nach dem Monotoniekriterium für Folgen 2.14.1 ist dann  $(s_n)$  und damit die Reihe konvergent, wenn  $(s_n)$  beschränkt ist, wenn also (3.9) gilt. Im anderen Fall ist  $(s_n)$  wachsend und unbeschränkt, und somit gilt  $s_n \rightarrow \infty$ , d.h. die Reihe divergiert bestimmt nach  $\infty$ .  $\square$

Zusammengefasst gilt für Reihen mit positiven Gliedern  $a_k \geq 0$ :

$$\begin{aligned} \sum_{k=0}^{\infty} a_k \text{ konvergent} &\iff \sum_{k=0}^{\infty} a_k < \infty \\ &\text{d.h. } \sum_{k=0}^{\infty} a_k \leq M \quad (\text{Reihe beschränkt}) \\ &\text{d.h. } \sum_{k=0}^n a_k \leq M \quad \text{für alle } n \quad (M \text{ fest}) \end{aligned}$$

**Beispiel 3.5.2** (a) Wir beweisen jetzt die Konvergenz der Reihe

$$\sum_{k=0}^{\infty} \frac{1}{k^2}$$

aus 3.1. Dazu wollen wir (3.9) zeigen, d.h. wir suchen eine obere Schranke  $M$  der Reihe. Wir schätzen dafür die Teilsummen  $s_n$  geschickt nach oben ab. Für  $k \geq 2$  gilt

$$\frac{1}{k^2} \leq \frac{1}{k(k-1)} = \frac{k - (k-1)}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}.$$

Setzen wir diese Abschätzung in die Teilsumme  $s_n$  ein, ergibt sich

$$\begin{aligned} s_n &= \sum_{k=1}^n \frac{1}{k^2} = 1 + \sum_{k=2}^n \frac{1}{k^2} \leq 1 + \sum_{k=2}^n \left( \frac{1}{k-1} - \frac{1}{k} \right) \\ &= 1 + \left( \frac{1}{1} - \frac{1}{2} \right) + \left( \frac{1}{2} - \frac{1}{3} \right) + \left( \frac{1}{3} - \frac{1}{4} \right) + \dots + \left( \frac{1}{n-1} - \frac{1}{n} \right) \\ &= 1 + \frac{1}{1} - \frac{1}{n} = 2 - \frac{1}{n} \leq 2. \end{aligned}$$

Das sich in obiger Rechnung alle benachbarten Terme der Klammern aufheben und nur der erste und letzte Term der Klammern stehenbleibt, nennt man auch „Teleskop-Summe“. Wir haben also gezeigt, dass  $s_n \leq 2$  für alle  $n$ , d.h.  $M = 2$  ist eine obere Schranke der Reihe. Damit ist die Reihe konvergent und es gilt außerdem

$$\sum_{k=0}^{\infty} \frac{1}{k^2} \leq 2.$$

Zur Erinnerung: In Abschnitt 3.1 haben wir erwähnt, dass  $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$  gilt. Tatsächlich ist  $\frac{\pi^2}{6} = 1.644\dots \leq 2$ , was also mit unserem Ergebnis übereinstimmt.

(b) Jetzt untersuchen wir die *harmonische Reihe*

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots \quad (3.10)$$

Wir zeigen, dass diese Reihe divergent ist, also

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

Wir gehen ähnlich vor wie eben, wollen aber jetzt zeigen, dass die Reihe unbeschränkt ist, dass also die Teilsummen  $s_n$  größer werden als jede Zahl. Dazu schätzen wir die Teilsumme für  $n = 2^m$  geschickt nach unten ab:

$$\begin{aligned} s_{2^m} &= 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{2^m} \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \dots + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^{m-1}+1} + \dots + \frac{1}{2^m}\right) \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \dots + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^m} + \dots + \frac{1}{2^m}\right) \\ &= 1 + \frac{1}{2} + \underbrace{2 \cdot \frac{1}{4}}_{\frac{1}{2}} + \underbrace{4 \cdot \frac{1}{8}}_{\frac{1}{2}} + \dots + \underbrace{2^{m-1} \cdot \frac{1}{2^m}}_{\frac{1}{2}} = 1 + m \cdot \frac{1}{2} \end{aligned}$$

Die Klammern laufen hier jeweils bis zur nächsten vollen Zweierpotenz im Nenner,  $4 = 2^2$ ,  $8 = 2^3$ , u.s.w. bis  $2^m$ . Die Anzahl Summanden pro Klammern ist dabei immer die nächst kleinere Zweierpotenz. Aus der Abschätzung folgt

$$s_{2^m} \geq 1 + \frac{m}{2} \rightarrow \infty \quad \text{bei } m \rightarrow \infty,$$

d.h. die Teilsummen  $s_n$  sind unbeschränkt, d.h. die Reihe divergiert. Man kann auch so schließen:

$$\sum_{k=1}^{\infty} \frac{1}{k} \geq \sum_{k=1}^{2^m} \frac{1}{k} \geq 1 + \frac{m}{2} \quad \text{für jedes } m \in \mathbb{N}.$$

Weil  $1 + \frac{m}{2}$  beliebig groß wird, folgt  $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$ .

**Beachte:** Nach dem notwendigen Kriterium Satz 3.4.1 gilt

$$\sum_{k=0}^{\infty} a_k \text{ konvergent} \Rightarrow a_k \rightarrow 0.$$

Die umgekehrte Implikation ist aber falsch, d.h.

$$a_k \rightarrow 0 \not\Rightarrow \sum_{k=0}^{\infty} a_k \text{ konvergent}.$$

Das Kriterium ist also nur notwendig und *nicht* hinreichend. Dies zeigt die harmonische Reihe, denn hier gilt

$$a_k = \frac{1}{k} \rightarrow 0, \quad \text{aber} \quad \sum_{k=1}^{\infty} \frac{1}{k} = \infty \text{ divergent}.$$

### 3.6 Absolute Konvergenz

Im Gegensatz zu Reihen mit positiven Gliedern, die wir im letzten Abschnitt betrachtet haben, ist die Frage der Konvergenz bei Reihen mit Gliedern  $a_k \in \mathbb{R}$ , die auch negativ sein können, im Allgemeinen schwieriger zu beantworten. Hier hilft die Eigenschaft der absoluten Konvergenz:

Ein Reihe  $\sum_{k=0}^{\infty} a_k$  mit  $a_k \in \mathbb{R}$  heißt *absolut konvergent*, wenn die Reihe der (Absolut-)Beträge  $\sum_{k=0}^{\infty} |a_k|$  konvergent ist. Da  $\sum_{k=0}^{\infty} |a_k|$  eine Reihe mit positiven Gliedern ist, ist also  $\sum_{k=0}^{\infty} a_k$  absolut konvergent genau dann, wenn die Reihe der Beträge beschränkt ist,

$$\sum_{k=0}^{\infty} |a_k| < \infty.$$

Absolute Konvergenz impliziert die Konvergenz der Reihe:

**Satz 3.6.1** *Ist die Reihe  $\sum_{k=0}^{\infty} a_k$  absolut konvergent, dann ist sie auch konvergent. Außerdem gilt dann*

$$\left| \sum_{k=0}^{\infty} a_k \right| \leq \sum_{k=0}^{\infty} |a_k|. \quad (3.11)$$

**Beispiel 3.6.2** Die Reihe

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

hat auf Grund des alternierenden Vorzeichens  $(-1)^k$  abwechselnd positive und negative Glieder. Die Reihe ist absolut konvergent, denn

$$\sum_{k=1}^{\infty} \left| \frac{(-1)^k}{k^2} \right| = \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty.$$

Damit ist die Reihe  $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2}$  insbesondere konvergent, und es gilt außerdem

$$\left| \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \right| \leq \sum_{k=1}^{\infty} \frac{1}{k^2} \leq 2.$$

Für den interessierten Leser geben wir hier noch den Beweis von Satz 3.6.1 an. Er beruht auf dem Cauchy Kriterium für Folgen.

**Beweis von Satz 3.6.1.** Um zu zeigen, dass die Folge der Teilsummen  $s_n = \sum_{k=0}^n a_k$  konvergiert, zeigen wir die Eigenschaft (2.7) des Cauchy Kriteriums für die Folge  $(s_n)$ . Für  $m > n$  gilt

$$|s_m - s_n| = \left| \sum_{k=0}^m a_k - \sum_{k=0}^n a_k \right| = \left| \sum_{k=n+1}^m a_k \right| \stackrel{(*)}{\leq} \sum_{k=n+1}^m |a_k| \leq \sum_{k=n+1}^{\infty} |a_k|.$$

Im Schritt (\*) haben wir mit der Dreiecksungleichung den Betrag in die Summe gezogen. Da  $\sum_{k=0}^{\infty} |a_k| < \infty$  konvergiert, konvergiert nach 3.4.1 der Reihenrest  $\sum_{k=n+1}^{\infty} |a_k|$  für  $n \rightarrow \infty$  gegen 0. Also gibt es zu  $\varepsilon > 0$  ein  $n_0$  mit

$$\sum_{k=n+1}^{\infty} |a_k| < \varepsilon \quad \text{für alle } n \geq n_0,$$

und daraus folgt dann

$$|s_m - s_n| < \varepsilon \quad \text{für alle } m \geq n \geq n_0.$$

Nach dem Cauchy Kriterium 2.15.1 konvergiert somit die Folge  $(s_n)$  und daher auch die Reihe  $\sum_{k=0}^{\infty} a_k$ . Die Abschätzung (3.11) folgt dann aus der Dreiecksungleichung für die endliche Summe

$$\left| \sum_{k=0}^n a_k \right| \leq \sum_{k=0}^n |a_k|.$$

beim Übergang  $n \rightarrow \infty$ . □

### 3.7 Majoranten- und Minorantenkriterium

Wir lernen jetzt mehrere Kriterien kennen, mit denen man die absolute Konvergenz einer Reihe überprüfen kann. Das erste ist das Majorantenkriterium, Teil (a) des nächsten Satzes; Teil (b) heißt Minorantenkriterium.

**Satz 3.7.1** (a) Gilt  $|a_k| \leq b_k$  für alle  $k \in \mathbb{N}$  und ist die Reihe  $\sum_{k=0}^{\infty} b_k$  konvergent, d.h.  $\sum_{k=0}^{\infty} b_k < \infty$ , dann ist  $\sum_{k=0}^{\infty} a_k$  absolut konvergent und es gilt

$$\left| \sum_{k=0}^{\infty} a_k \right| \leq \sum_{k=0}^{\infty} |a_k| \leq \sum_{k=0}^{\infty} b_k.$$

Die Reihe  $\sum_{k=0}^{\infty} b_k$  heißt Majorante der Reihe  $\sum_{k=0}^{\infty} a_k$ .

(b) Ist  $a_k \geq b_k \geq 0$  und  $\sum_{k=0}^{\infty} b_k = \infty$  divergent, dann ist auch  $\sum_{k=0}^{\infty} a_k = \infty$  divergent. In diesem Fall ist  $\sum_{k=0}^{\infty} b_k$  Minorante von  $\sum_{k=0}^{\infty} a_k$ .

**Beweis.** Im Fall (a) folgt aus  $|a_k| \leq b_k$  sofort

$$\sum_{k=0}^{\infty} |a_k| \leq \sum_{k=0}^{\infty} b_k < \infty$$

und damit die absolute Konvergenz von  $\sum_{k=0}^{\infty} a_k$ . Bei (b) gilt wegen  $a_k \geq b_k$

$$\sum_{k=0}^{\infty} a_k \geq \sum_{k=0}^{\infty} b_k = \infty,$$

und somit  $\sum_{k=0}^{\infty} a_k = \infty$ . □

**Beispiel 3.7.2** Ist die Reihe

$$\sum_{k=1}^{\infty} \frac{(-1)^k - \frac{1}{k}}{k^2 + 1} \tag{3.12}$$

konvergent oder divergent? Um eine Vermutung für die Antwort zu bekommen, betrachten wir zunächst das Verhalten der Reihenglieder

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

für große Werte von  $k$ .  $(-1)^k$  ist abwechselnd  $+1$  und  $-1$ , wozu dann  $\frac{1}{k}$  hinzuaddiert wird, was aber gegen  $0$  konvergiert, also für große  $k$  gegenüber  $\pm 1$  vernachlässigt werden kann. Genauso spielt für große  $k$  die Addition von  $1$  gegenüber  $k^2$  kaum eine Rolle. Für große  $k$  verhalten sich also die Reihenglieder wie  $\frac{(-1)^k}{k^2}$ . Da die Reihe der Beträge davon,  $\sum_{k=1}^{\infty} \frac{1}{k^2}$ , konvergiert, vermuten wir also, dass die Reihe (3.12) absolut konvergiert. Diese Vermutung zeigen wir nun mit dem Majorantenkriterium. Dazu betrachten wir die Reihe der Beträge von (3.12) und versuchen durch Abschätzung nach oben zu einer konvergenten Majorante zu gelangen:

$$\begin{aligned} \sum_{k=1}^{\infty} \left| \frac{(-1)^k - \frac{1}{k}}{k^2 + 1} \right| &= \sum_{k=1}^{\infty} \frac{|(-1)^k - \frac{1}{k}|}{k^2 + 1} \leq \sum_{k=1}^{\infty} \frac{|(-1)^k| + \frac{1}{k}}{k^2 + 1} = \sum_{k=1}^{\infty} \frac{1 + \frac{1}{k}}{k^2 + 1} \\ &\leq \sum_{k=1}^{\infty} \frac{2}{k^2 + 1} \leq \sum_{k=1}^{\infty} \frac{2}{k^2} = 2 \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty \end{aligned}$$

Also ist die Reihe  $\sum_{k=1}^{\infty} \frac{2}{k^2}$  eine konvergente Majorante. Damit ist  $\sum_{k=1}^{\infty} \frac{(-1)^k - \frac{1}{k}}{k^2 + 1}$  absolut konvergent, also auch konvergent.

Als weiteres Beispiel untersuchen wir jetzt die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \quad (\alpha > 0)$$

auf Konvergenz. Wir wissen schon (Beispiel 3.5.2), dass die Reihe für  $\alpha = 1$  divergiert (harmonische Reihe!) und für  $\alpha = 2$  konvergiert (Reihe mit  $\frac{1}{k^2}$ ). Mit Hilfe des Majoranten- und Minorantenkriteriums können wir daraus jetzt Aussagen für andere Werte von  $\alpha$  erhalten:

**Satz 3.7.3** Die Reihe  $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$  ist konvergent für  $\alpha > 1$  und divergent für  $0 < \alpha \leq 1$ . Es gilt also

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \begin{cases} < \infty & \text{falls } \alpha > 1, \\ = \infty & \text{falls } 0 < \alpha \leq 1. \end{cases}$$

**Beweis.** Wir betrachten zuerst den Fall  $0 < \alpha \leq 1$ : Aus  $\alpha \leq 1$  folgt

$$k^\alpha \leq k^1 = k \quad \text{für } k \geq 1.$$

Für  $\alpha = \frac{1}{2}$  zum Beispiel ist  $k^\alpha = k^{\frac{1}{2}} = \sqrt{k} \leq k$ . Für allgemeine Werte von  $\alpha$  folgt die Ungleichung aus der Monotonie der Exponentialfunktion. Aus  $k^\alpha \leq k$  folgt nun weiter

$$\frac{1}{k^\alpha} \geq \frac{1}{k}$$

und damit

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \geq \sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

Also ist nach dem Minorantenkriterium  $\sum_{k=1}^{\infty} \frac{1}{k^\alpha} = \infty$  divergent wenn  $\alpha \leq 1$ .

Im Fall  $\alpha \geq 2$  gilt  $k^\alpha \geq k^2$ . Dann gilt

$$\frac{1}{k^\alpha} \leq \frac{1}{k^2} \Rightarrow \sum_{k=1}^{\infty} \frac{1}{k^\alpha} \leq \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty.$$

Somit ist  $\sum_{k=1}^{\infty} \frac{1}{k^2}$  eine konvergente Majorante und  $\sum_{k=1}^{\infty} \frac{1}{k^\alpha} < \infty$  konvergent.

Den noch fehlende Fall  $1 < \alpha < 2$  können wir so nicht behandeln. Denn da wir nur die Konvergenz für  $\alpha = 2$  aus Beispiel 3.5.2(a) kennen, muss die Reihe mit  $\alpha = 2$  die Majorante sein, und  $\frac{1}{k^\alpha} \leq \frac{1}{k^2}$  gilt nur bei  $\alpha \geq 2$ . Mit dem Integralkriterium (Satz 3.10.1) werden wir später aber auch noch den fehlenden Fall beweisen.  $\square$

### 3.8 Quotienten- und Wurzelkriterium I

Die vielleicht wichtigsten Kriterien zur Überprüfung einer Reihe auf (absolute) Konvergenz, sind das Quotienten- und das Wurzelkriterium. In diesem Abschnitt schauen wir uns eine vereinfachte Version dieser Kriterien an. Die allgemeine Variante zusammen mit einem Beweis folgt in den nächsten Abschnitten.

**Satz 3.8.1 (Quotientenkriterium)** Sei  $\sum_{k=0}^{\infty} a_k$  eine Reihe mit  $a_k \neq 0$  für alle  $k$ . Existiert der Grenzwert

$$b = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right|$$

so gilt

$$(i) \quad b < 1 \Rightarrow \sum_{k=0}^{\infty} a_k \text{ absolut konvergent,}$$

$$(ii) \quad b > 1 \Rightarrow \sum_{k=0}^{\infty} a_k \text{ divergent.}$$

**Satz 3.8.2 (Wurzelkriterium)** Sei  $\sum_{k=0}^{\infty} a_k$  eine Reihe. Existiert der Grenzwert

$$b = \lim_{k \rightarrow \infty} \sqrt[k]{|a_k|}$$

so gilt

$$(i) \quad b < 1 \Rightarrow \sum_{k=0}^{\infty} a_k \text{ absolut konvergent,}$$

$$(ii) \quad b > 1 \Rightarrow \sum_{k=0}^{\infty} a_k \text{ divergent.}$$

**Beachte:** Bei beiden Kriterien ist im Fall  $b = 1$  keine allgemeine Aussage über die Konvergenz der Reihe möglich. D.h. abhängig von der konkreten Reihe kann diese konvergent oder auch divergent sein.

**Beispiel 3.8.3** Wir wollen die Reihe

$$\sum_{k=0}^{\infty} \frac{(-2)^k}{k!}$$

auf Konvergenz untersuchen und benutzen dazu das Quotientenkriterium. Das Reihenglied ist

$$a_k = \frac{(-2)^k}{k!}.$$

Für das Quotientenkriterium berechnen wir zunächst  $|\frac{a_{k+1}}{a_k}|$ :

$$\left| \frac{a_{k+1}}{a_k} \right| = \left| \frac{(-2)^{k+1}}{(k+1)!} \cdot \frac{k!}{(-2)^k} \right| = \left| (-2) \cdot \frac{k!}{(k+1)!} \right| = \left| \frac{-2}{k+1} \right| = \frac{2}{k+1}.$$

Für den Grenzwert erhält man dann

$$b = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = \lim_{k \rightarrow \infty} \frac{2}{k+1} = 0,$$

also  $b < 1$ . Somit folgt, dass die Reihe (absolut) konvergiert.

### 3.9 Quotienten- und Wurzelkriterium II

Beim Quotienten- und Wurzelkriterium in der Version des letzten Abschnitts berechnet man einen Grenzwert (Limes des Quotienten bzw. Limes der n. Wurzel). Das dieser Grenzwert existiert, war dabei eine Voraussetzung des Kriteriums. Wir wissen aber aus dem Kapitel über Folgen, dass Grenzwerte (auch uneigentliche) nicht unbedingt existieren müssen, z.B. bei Folgen mit mehreren Häufungswerten. In diesem Abschnitt lernen wir die allgemeine Variante des Quotienten- und Wurzelkriteriums kennen, die ohne Grenzwert auskommt.

**Satz 3.9.1 (Quotienten- und Wurzelkriterium allgemein)** *Gibt es Zahlen  $0 \leq q < 1$  und  $k_0 \in \mathbb{N}$  sodass*

$$(a) \ a_k \neq 0 \text{ und } \left| \frac{a_{k+1}}{a_k} \right| \leq q \text{ für alle } k \geq k_0, \text{ oder}$$

$$(b) \ \sqrt[k]{|a_k|} \leq q \text{ für alle } k \geq k_0,$$

dann ist  $\sum_{k=0}^{\infty} a_k$  absolut konvergent.

**Beweis.** Wir zeigen das Wurzelkriterium Teil (b), das Quotientenkriterium (a) beweist man ähnlich. Aus der Voraussetzung  $\sqrt[k]{|a_k|} \leq q$  folgt

$$|a_k| \leq q^k \quad \text{für } k \geq k_0$$

und damit

$$\sum_{k=k_0}^{\infty} |a_k| \leq \sum_{k=k_0}^{\infty} q^k \leq \sum_{k=0}^{\infty} q^k = \frac{1}{1-q} < \infty.$$

Beim Schritt vom Reihenanfang  $k = k_0$  zu  $k = 0$  haben wir  $q \geq 0$ , und dann im letzten Schritt  $q < 1$  für die Konvergenz der geometrischen Reihe benutzt. Es gilt also  $\sum_{k=k_0}^{\infty} |a_k| < \infty$ , woraus dann auch

$$\sum_{k=0}^{\infty} |a_k| < \infty$$

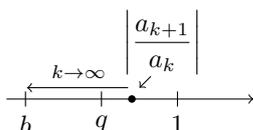
folgt, da hier ja nur endlich viele Summanden am Anfang der Reihe hinzukommen. Somit ist die Reihe  $\sum_{k=0}^{\infty} a_k$  absolut konvergent.  $\square$

**Beachte:**

- (a) Die „Limes“-Versionen des Quotienten- und Wurzelkriteriums (Satz 3.8.1 und 3.8.2) ergeben sich als Konsequenz aus der allgemeinen Variante. Wenn z.B.

$$\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = b < 1$$

gilt, dann kann man irgendein  $q$  zwischen  $b$  und  $1$  wählen, d.h.  $q \in ]b, 1[$ . Aus  $\left| \frac{a_{k+1}}{a_k} \right| \rightarrow b$  folgt dann  $\left| \frac{a_{k+1}}{a_k} \right| \leq q$  wenn  $k$  groß genug ist, also für alle  $k \geq k_0$  mit  $k_0$  hinreichend groß. Anschaulich:



- (b) Im Fall  $q = 1$  ist auch in der allgemeinen Variante wie zuvor in den Limes-Versionen keine Aussage möglich über die Konvergenz möglich. Der Beweis funktioniert für  $q = 1$  nicht, weil in dem Fall die geometrische Reihe nicht konvergiert.
- (c) Jedoch erhält man die Divergenz der Reihe wenn

$$\left| \frac{a_{k+1}}{a_k} \right| \geq 1 \quad \text{oder} \quad \sqrt[k]{|a_k|} \geq 1 \quad \text{für alle} \quad k \geq k_0,$$

denn dann kann  $(a_k)$  keine Nullfolge sein. (Bei der Divergenz ist der Fall 1 hier also erlaubt.)

Es stellt sich nun die Frage, welches der bisherigen Konvergenzkriterien für Reihen man im konkreten Fall benutzen soll. Dazu ein paar Hinweise:

- Quotienten- und Wurzelkriterium sind (relativ) einfach anzuwenden und führen oft zum Ziel.
- Das Quotientenkriterium bietet sich bei Reihen an, die Ausdrücke mit Fakultäten enthalten. Im Quotienten  $\frac{a_{k+1}}{a_k}$  kann man dann immer kürzen. Das Wurzelkriterium ist dagegen schlecht anzuwenden, da man Grenzwerte, die  $\sqrt[k]{k!}$  enthalten, nur schwer ausrechnen kann.
- Kommen in der Reihe Terme mit  $k^k$  oder ähnliches vor, aber keine Fakultäten, dann ist das Wurzelkriterium gut geeignet.
- Wenn das Quotientenkriterium keine Entscheidung über die Konvergenz liefert (Fall  $b = 1$ ), dann führt meist auch das Wurzelkriterium zu keiner Entscheidung, und umgekehrt. In diesem Fall ist häufig das Majoranten- oder Minorantenkriterium erfolgreich.
- Das notwendige Kriterium 3.4.1 zur Konvergenz ist sehr einfach anzuwenden (konvergiert  $a_k \rightarrow 0$ ?), kann aber nur die Divergenz zeigen: Wenn  $a_k \not\rightarrow 0$ , dann ist die Reihe divergent.

**Beispiel 3.9.2** Wir untersuchen folgende Reihen auf Konvergenz:

(a)

$$\sum_{k=1}^{\infty} \frac{6^k \cdot k^3}{(-k)^{3k}}$$

Der Term  $(-k)^{3k}$  legt das Wurzelkriterium nahe:

$$\begin{aligned} \sqrt[k]{|a_k|} &= \sqrt[k]{\left| \frac{6^k \cdot k^3}{(-k)^{3k}} \right|} = \sqrt[k]{\left| \frac{6^k \cdot k^3}{(-1)^{3k} \cdot k^{3k}} \right|} = \sqrt[k]{\frac{6^k \cdot k^3}{k^{3k}}} = \frac{\sqrt[k]{6^k} \cdot \sqrt[k]{k^3}}{\sqrt[k]{k^{3k}}} \\ &= \frac{6(\sqrt[k]{k})^3}{k^3} \rightarrow 0 < 1 \quad \text{für } k \rightarrow \infty \end{aligned}$$

denn  $\lim_{k \rightarrow \infty} \sqrt[k]{k} = 1$ . Die Reihe ist damit (absolut) konvergent.

(b)

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

Wir benutzen das Quotientenkriterium:

$$\begin{aligned} \left| \frac{a_{k+1}}{a_k} \right| &= \frac{2^k}{(k+2)^2} \cdot \frac{(k+1)^2}{2^{k-1}} = \frac{2(k+1)^2}{(k+2)^2} \\ &= 2 \left( \frac{k+1}{k+2} \right)^2 \xrightarrow{\rightarrow 1} 2 > 1 \quad \text{für } k \rightarrow \infty, \end{aligned}$$

die Reihe divergiert also. Mit dem Wurzelkriterium müsste man folgendermaßen argumentieren:

$$\sqrt[k]{|a_k|} = \sqrt[k]{\frac{2^{k-1}}{(k+1)^2}} = \sqrt[k]{\frac{2^k}{2(k+1)^2}} = \frac{2}{\sqrt[k]{2} (\sqrt[k]{k+1})^2}$$

Aus

$$1 \leq \sqrt[k]{k+1} \leq \sqrt[k]{2k} = \sqrt[k]{2} \cdot \sqrt[k]{k} \rightarrow 1 \quad \text{für } k \rightarrow \infty$$

folgt nach dem Sandwich-Lemma  $\sqrt[k]{k+1} \rightarrow 1$  und somit

$$\sqrt[k]{|a_k|} = \frac{2}{\sqrt[k]{2} (\sqrt[k]{k+1})^2} \rightarrow \frac{2}{1 \cdot 1^2} = 2 > 1,$$

also wieder Divergenz. Man sieht auch, dass hier das Wurzelkriterium etwas aufwändiger ist als das Quotientenkriterium.

(c)

$$\sum_{k=2}^{\infty} \frac{(-1)^k}{k^3 - k}$$

Wir versuchen zunächst das Quotientenkriterium:

$$\begin{aligned} \left| \frac{a_{k+1}}{a_k} \right| &= \left| \frac{(-1)^{k+1}}{(k+1)^3 - (k+1)} \cdot \frac{k^3 - k}{(-1)^k} \right| = \frac{k^3 - k}{(k+1)^3 - k - 1} \\ &= \frac{k^3 \left(1 - \frac{1}{k^2}\right)}{k^3 \left( \left(1 + \frac{1}{k}\right)^3 - \frac{1}{k^2} - \frac{1}{k^3} \right)} = \frac{1 - \frac{1}{k^2}}{\left(1 + \frac{1}{k}\right)^3 - \frac{1}{k^2} - \frac{1}{k^3}} \rightarrow 1, \end{aligned}$$

also

$$b = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = 1.$$

Das Quotientenkriterium liefert damit keine Aussage. Für das Wurzelkriterium ergibt sich zuerst

$$\sqrt[k]{|a_k|} = \frac{1}{\sqrt[k]{k^3 - k}}.$$

Dann gilt

$$\sqrt[k]{k^3 - k} = \sqrt[k]{k^3 \left(1 - \frac{1}{k^2}\right)} = \underbrace{(\sqrt[k]{k})^3}_{\rightarrow 1} \cdot \sqrt[k]{1 - \frac{1}{k^2}}.$$

Für den Grenzwert der zweiten Wurzel braucht man wieder „Sandwich“: Für  $k \geq 2$  gilt  $1 - \frac{1}{k^2} \geq \frac{3}{4}$  und damit

$$\underbrace{\sqrt[k]{\frac{3}{4}}}_{\rightarrow 1 \text{ bei } k \rightarrow \infty} \leq \sqrt[k]{1 - \frac{1}{k^2}} \leq \sqrt[k]{1} = 1.$$

Mit dem Sandwich-Lemma folgt also

$$\lim_{k \rightarrow \infty} \sqrt[k]{1 - \frac{1}{k^2}} = 1$$

und somit

$$\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|} = \lim_{k \rightarrow \infty} \frac{1}{(\sqrt[k]{k})^3 \cdot \sqrt[k]{1 - \frac{1}{k^2}}} = 1.$$

Also liefert auch das Wurzelkriterium keine Aussage. Wir können aber das Majorantenkriterium erfolgreich verwenden. Beachte dazu, dass sich für große  $k$  die Beträge der Reihenglieder  $\left| \frac{(-1)^k}{k^3 - k} \right|$  wie  $\frac{1}{k^3}$  verhalten. Da  $\sum \frac{1}{k^3}$  konvergent ist (Satz 3.7.3,  $\alpha = 3 > 1$ ), versuchen wir jetzt unsere Reihe nach oben gegen die Reihe  $\sum \frac{1}{k^3}$  abzuschätzen:

$$\left| \frac{(-1)^k}{k^3 - k} \right| = \frac{1}{k^3 - k} = \frac{1}{k^3 \underbrace{\left(1 - \frac{1}{k^2}\right)}_{\geq \frac{3}{4} \text{ (} k \geq 2 \text{)}}} \leq \frac{1}{\frac{3}{4}k^3} = \frac{4}{3k^3}.$$

Daraus folgt

$$\sum_{k=2}^{\infty} \left| \frac{(-1)^k}{k^3 - k} \right| \leq \frac{4}{3} \cdot \sum_{k=2}^{\infty} \frac{1}{k^3} < \infty.$$

Also ist die Reihe nach dem Majorantenkriterium absolut konvergent.

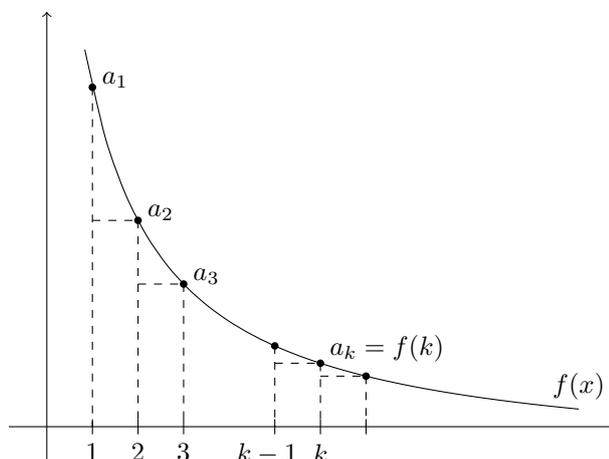
### 3.10 Integralkriterium

In diesem und dem nächsten Abschnitt folgen nun zwei Kriterien, die auf Reihen mit einer bestimmten Struktur anwendbar sind. Beim folgende Integralkriterium sind das Reihen mit positiven, monoton fallenden Gliedern.

**Satz 3.10.1** Sei  $a_k = f(k)$  für  $k \geq 1$  wobei die Funktion  $f : [1, \infty[ \rightarrow \mathbb{R}$  monoton fallend und positiv ist,  $f(x) \geq 0$ . Es gilt dann

$$\int_1^{\infty} f(x) dx < \infty \implies \sum_{k=1}^{\infty} a_k < \infty \text{ konvergent.}$$

**Beweis.** Wir betrachten Rechtecke der Breite 1 unterhalb des Graphen von  $f$ , die  $(k, a_k)$  als einen Eckpunkt haben:



Das Rechteck zwischen  $x = k - 1$  und  $x = k$  hat die Höhe  $a_k$ , somit ist sein Flächeninhalt gleich  $1 \cdot a_k = a_k = f(k)$ . Die Summe aller Rechtecksflächen für  $k \geq 2$  ist dann  $\sum_{k=2}^{\infty} a_k$ . Diese Summe ist kleiner als die Fläche unterhalb des Graphen von  $f$  im Intervall  $[1, \infty[$ , also

$$\sum_{k=2}^{\infty} a_k \leq \int_1^{\infty} f(x) dx < \infty.$$

Damit gilt auch  $\sum_{k=1}^{\infty} a_k < \infty$ . □

Als Beispiel für das Integralkriterium zeigen wir jetzt die Konvergenz der Reihe  $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$  für alle  $\alpha > 1$ , vergleiche Satz 3.7.3. (Dort hat uns noch der Fall  $1 < \alpha < 2$  gefehlt.)

**Beispiel 3.10.2** Wir betrachten die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \text{ für } \alpha > 1.$$

Hier ist also  $a_k = \frac{1}{k^\alpha}$ . Mit der Funktion

$$f(x) = \frac{1}{x^\alpha}$$

gilt dann  $a_k = f(k)$ . Die Funktion  $f$  ist monoton fallend und es gilt  $f(x) > 0$ . Für das Integral berechnen wir

$$\begin{aligned} \int_1^\infty \frac{1}{x^\alpha} dx &= \int_1^\infty x^{-\alpha} dx = \left[ \frac{1}{-\alpha+1} x^{-\alpha+1} \right]_1^\infty = \left[ \frac{-1}{\alpha-1} \cdot \frac{1}{x^{\alpha-1}} \right]_1^\infty \\ &= \lim_{x \rightarrow \infty} \left( \frac{-1}{\alpha-1} \cdot \underbrace{\frac{1}{x^{\alpha-1}}}_{\rightarrow 0} - \frac{-1}{\alpha-1} \right) = \frac{1}{\alpha-1} < \infty. \end{aligned}$$

Beim Grenzwert  $x \rightarrow \infty$  im letzten Schritt haben wir

$$\alpha > 1 \quad \Rightarrow \quad \alpha - 1 > 0 \quad \Rightarrow \quad x^{\alpha-1} \rightarrow \infty \quad \Rightarrow \quad \frac{1}{x^{\alpha-1}} \rightarrow 0$$

benutzt. Mit Integralkriterium erhalten wir also die Konvergenz  $\sum_{k=1}^\infty \frac{1}{k^\alpha} < \infty$  der Reihe.

### 3.11 Alternierende Reihen

Eine Reihe mit abwechselnd positiven und negative Gliedern heißt *alternierend*. Die Reihe

$$\sum_{k=0}^{\infty} (-1)^k a_k = a_0 - a_1 + a_2 - a_3 + a_4 - \dots$$

mit positiven Zahlen  $a_k \geq 0$  ist also alternierend.

**Satz 3.11.1 (Leibniz-Kriterium)** *Ist die Folge  $(a_k)_{k \in \mathbb{N}}$  eine monoton fallende Nullfolge, dann ist die alternierende Reihe  $\sum_{k=0}^{\infty} (-1)^k a_k$  konvergent.*

**Beachte:** Für das Leibniz-Kriterium spielt es keine Rolle, ob die Reihe mit einem positiven oder einem negativen Glied beginnt. Denn  $\sum_{k=0}^{\infty} (-1)^k a_k$  konvergiert genau dann, wenn

$$\sum_{k=0}^{\infty} (-1)^{k+1} a_k = - \sum_{k=0}^{\infty} (-1)^k a_k = -a_0 + a_1 - a_2 + a_3 - \dots$$

konvergiert.

**Beispiel 3.11.2** Ein wichtiges Beispiel einer alternierenden Reihe ist die *alternierende harmonische Reihe*

$$\sum_{k=1}^{\infty} (-1)^{k+1} \cdot \frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots \quad (3.13)$$

Die zugehörige Folge ist hier also

$$a_k = \frac{1}{k}.$$

Weil  $(\frac{1}{k})_{k \geq 1}$  monoton fallend und eine Nullfolge ist ( $\frac{1}{k} \rightarrow 0$ ), ist die alternierende harmonische Reihe (3.13) also nach dem Leibniz-Kriterium konvergent. Ihren Grenzwert werden wir im Kapitel zu Taylorreihen berechnen.

**Bemerkung:** Die alternierende harmonische Reihe ist übrigens auch ein Beispiel für eine konvergente Reihe, die nicht absolut konvergent ist. Denn die Reihe der Beträge ist hier die harmonische Reihe und damit divergent:

$$\sum_{k=1}^{\infty} \left| (-1)^{k+1} \cdot \frac{1}{k} \right| = \sum_{k=1}^{\infty} \frac{1}{k} = \infty \quad (\text{harmonische Reihe})$$

Zum Abschluss geben wir hier noch eine Begründung dafür, *warum* die alternierende harmonische Reihe konvergiert.<sup>3</sup> Wir betrachten die Teilsummen der Reihe, also

$$s_n = \sum_{k=1}^n (-1)^{k+1} \frac{1}{k},$$

abwechselnd für gerade und ungerade  $n$ .

$$\begin{aligned} s_1 &= 1 & s_2 &= 1 - \frac{1}{2} \\ s_3 &= 1 + \left( -\frac{1}{2} + \frac{1}{3} \right) & s_4 &= \left( 1 - \frac{1}{2} \right) + \left( \frac{1}{3} - \frac{1}{4} \right) \\ s_5 &= 1 + \left( -\frac{1}{2} + \frac{1}{3} \right) + \left( -\frac{1}{4} + \frac{1}{5} \right) & s_6 &= \left( 1 - \frac{1}{2} \right) + \left( \frac{1}{3} - \frac{1}{4} \right) + \left( \frac{1}{5} - \frac{1}{6} \right) \end{aligned}$$

Die Klammern haben wir hier zur Veranschaulichung eingefügt: Für die  $s_n$  mit ungeradem  $n$  auf der linken Seite ist jede Klammer negativ (weil die  $\frac{1}{k}$  monoton fallen!) Jedes nächste ungerade  $s_n$  ist also kleiner als das vorangehende,

$$s_1 \geq s_3 \geq s_5 \geq \dots$$

Bei den  $s_n$  mit geradem  $n$  rechts ist jede Klammer positiv, also wachsen die geraden  $s_n$ ,

$$s_2 \leq s_4 \leq s_6 \leq \dots$$

Außerdem ist jedes gerade  $s_n$  kleiner als das vorangehende ungerade  $s_{n-1}$ , denn es kommt ein negatives Glied hinzu. Z.B. ist  $s_6 < s_5$ . Zusammen erhalten wir für die Glieder  $s_1$  bis  $s_6$  also

$$s_2 \leq s_4 \leq s_6 < s_5 \leq s_3 \leq s_1.$$

Da aber *jedes* gerade  $s_{2m}$  kleiner als das vorangehende ungerade  $s_{2m-1}$  ist, und  $m$  beliebig groß werden kann, haben wir tatsächlich

$$s_2 \leq s_4 \leq s_6 \leq \dots \leq s_{2m} \leq \dots < \dots \leq s_{2m-1} \leq \dots \leq s_5 \leq s_3 \leq s_1.$$

D.h., die geraden  $s_{2m}$  wachsen, sind dabei aber immer kleiner als die ungeraden  $s_{2m-1}$ , die wiederum fallen. Jetzt schauen wir uns den Abstand zwischen  $s_{2m}$  und  $s_{2m-1}$  an, er ist

$$\begin{aligned} s_{2m-1} - s_{2m} &= 1 - \frac{1}{2} + \dots + \frac{1}{2m-1} - \left( 1 - \frac{1}{2} + \dots + \frac{1}{2m-1} - \frac{1}{2m} \right) \\ &= \frac{1}{2m}. \end{aligned}$$

<sup>3</sup>Der allgemeine Beweis des Leibniz-Kriteriums läuft ganz analog.

Der Abstand konvergiert also gegen 0 (denn  $\frac{1}{k}$  ist eine Nullfolge!) Damit ziehen sich also die geraden und ungeraden  $s_n$  auf einen Punkt  $s$  zusammen. (Das ist wieder eine Intervallschachtelung, wie schon bei Satz 2.6.2 beschrieben.) Also konvergieren die Teilsummen  $s_n$  gegen  $s$ , d.h. die Reihe konvergiert und ihr Grenzwert ist  $s$ . Insgesamt ist die Anordnung der Teilsummen  $s_n$  und des Grenzwerts  $s$  damit

$$s_2 \leq s_4 \leq s_6 \leq \dots \leq s_{2m} \leq \dots < s < \dots \leq s_{2m-1} \leq \dots \leq s_5 \leq s_3 \leq s_1.$$

## 3.12 Umordnung von Reihen

Eine *Umordnung einer Reihe* bedeutet die Änderung der Reihenfolge ihrer Summanden. D.h. die umgeordnete Reihe enthält die gleichen Glieder wie die Ausgangsreihe, nur ihre Anordnung innerhalb der Reihe ist anders. Eine mögliche Umordnung der Reihe

$$\sum_{k=0}^{\infty} a_k = a_0 + a_1 + a_2 + a_3 + \dots$$

ist etwa

$$a_0 + a_1 + a_3 + a_2 + a_5 + a_7 + a_9 + a_4 + a_{11} + a_{13} + a_{15} + a_{17} + a_6 + \dots$$

Hier folgt also auf das erste Reihenglied mit geradem Index  $a_0$  die ersten zwei Glieder mit ungeradem Index  $a_1$  und  $a_3$ , dann das nächste gerade Glied  $a_2$ , dann die nächsten drei ungeraden Glieder  $a_5$ ,  $a_7$ ,  $a_9$ , dann das nächste gerade  $a_4$ , dann die nächsten vier ungeraden  $a_{11}$  bis  $a_{17}$ , u.s.w.

Spielt die Reihenfolge der Glieder einer Reihe eine Rolle. Oder kann sich bei einer Umordnung die Konvergenz oder der Grenzwert einer Reihe ändern? Naiv wird man vermuten, dass das nicht sein kann. Schließlich ist die Addition kommutativ, d.h.  $a + b = b + a$ , die Reihenfolge spielt bei der Addition keine Rolle. Das ist richtig für die Addition *endlich* vieler Zahlen. Bei Reihen haben wir aber die Addition von *unendlich* vielen Zahlen, und in diesem Fall kann sich der Grenzwert der Reihe bei einer Umordnung tatsächlich ändern!

Die absolute Konvergenz einer Reihe entscheidet dabei darüber, ob sich der Wert der Reihe bei Umordnung ändert, oder nicht. Es gilt:

- Ist die Reihe  $\sum_{k=0}^{\infty} a_k$  absolut konvergent, so bleibt der (Grenz-)Wert der Reihe bei jeder Umordnung gleich.
- Ist die Reihe  $\sum_{k=0}^{\infty} a_k$  konvergent, aber nicht absolut konvergent, dann ändert sich der Reihenwert bei einer Umordnung im Allgemeinen. Die Reihe kann durch die Umordnung sogar divergent werden.

Dieses überraschende Phänomen illustrieren wir nun an der alternierenden harmonischen Reihe. Wie wir im letzten Abschnitt gesehen haben, ist diese Reihe konvergent, aber nicht absolut konvergent.

**Beispiel 3.12.1** In der ursprünglichen Anordnung von (3.13) ergibt sich für den Wert der alternierenden harmonischen Reihe

$$\sum_{k=0}^{\infty} (-1)^{k+1} \frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots = 0.69314\dots \quad (3.14)$$

Wir betrachten jetzt folgende Umordnung der Reihe: Wir addieren zuerst die beiden positiven Glieder 1 und  $\frac{1}{3}$ , dann das erste negative Glied  $-\frac{1}{2}$ , dann die nächsten beiden positiven, dann das nächste negative Glied, u.s.w. Die umgeordnete Reihe und ihr Grenzwert ist dann

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \dots = 1.0397\dots \quad (3.15)$$

Durch direktes Ausrechnen der Summen in (3.14) und (3.15) lassen sich die beiden verschiedenen Grenzwerte leicht bestätigen. (Versuchen Sie es!) Eine theoretische Begründung der Änderung des Grenzwerts bei Umordnungen ist schwieriger. Sie beruht darauf, dass die Reihen über die positiven bzw. negativen Glieder für sich allein jeweils divergieren,

$$1 + \frac{1}{3} + \frac{1}{5} + \dots = \infty, \quad -\frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \dots = -\infty.$$

Damit lässt sich z.B., egal wieviele positive Glieder bis zu einem bestimmten Punkt der umgeordneten Reihe schon benutzt wurden, durch die restlichen positiven Glieder in der Summe immer noch ein beliebig hoher positiver Wert erreichen. Wie hoch dieser Wert dann letztlich wird, hängt davon ab, wie „schnell“ die negativen Werte in der Umordnung folgen.

### 3.13 Die Cauchyproduktformel

Für das Produkt zweier Reihen gilt die folgende sogenannte *Cauchyproduktformel*. Voraussetzung für die Formel ist dabei die absolute Konvergenz der Reihen.

**Satz 3.13.1** Sind  $\sum_{k=0}^{\infty} a_k$  und  $\sum_{k=0}^{\infty} b_k$  absolut konvergent, dann gilt

$$\left( \sum_{k=0}^{\infty} a_k \right) \cdot \left( \sum_{k=0}^{\infty} b_k \right) = \sum_{n=0}^{\infty} c_n \quad \text{mit} \quad c_n = \sum_{k=0}^n a_{n-k} b_k.$$

Die Produktreihe  $\sum_{k=0}^{\infty} c_k$  ist dabei wieder absolut konvergent.

Eingesetzt gilt also für das Produkt zweier absolut konvergenter Reihen die Formel

$$\left( \sum_{k=0}^{\infty} a_k \right) \cdot \left( \sum_{k=0}^{\infty} b_k \right) = \sum_{n=0}^{\infty} \sum_{k=0}^n a_{n-k} b_k.$$

Die Cauchyproduktformel kommt folgendermaßen zustande. Bildet man formal das Produkt der beiden unendlichen Reihen, so erhält man zunächst

$$\begin{aligned} \left( \sum_{k=0}^{\infty} a_k \right) \cdot \left( \sum_{k=0}^{\infty} b_k \right) &= a_0 b_0 + a_0 b_1 + a_0 b_2 + a_0 b_3 + \dots \\ &\quad + a_1 b_0 + a_1 b_1 + a_1 b_2 + a_1 b_3 + \dots \\ &\quad + a_2 b_0 + a_2 b_1 + a_2 b_2 + a_2 b_3 + \dots \\ &\quad + a_3 b_0 + a_3 b_1 + \dots \\ &\quad \vdots \quad \quad \quad \vdots \end{aligned}$$

In der entstehenden unendlichen Summe „in Matrixform“ betrachten wir nun nach rechts aufsteigende Diagonalen von der linken Spalte zur oberen Zeile, beginnend mit dem einzelnen Summanden  $a_0b_0$  in der linken oberen Ecke:

$$\begin{aligned}c_0 &= a_0b_0, \\c_1 &= a_1b_0 + a_0b_1, \\c_2 &= a_2b_0 + a_1b_1 + a_0b_2, \\c_3 &= a_3b_0 + a_2b_1 + a_1b_2 + a_0b_3, \quad \dots\end{aligned}$$

allgemein

$$c_n = a_nb_0 + a_{n-1}b_1 + \dots + a_0b_n = \sum_{k=0}^n a_{n-k}b_k.$$

Die unendliche „Matrixsumme“ lässt sich somit als  $\sum_{n=0}^{\infty} c_n$  schreiben, wobei wir die ursprüngliche Reihenfolge der Summation geändert haben, was aber aufgrund der absoluten Konvergenz erlaubt ist. Wir erhalten damit

$$\left( \sum_{k=0}^{\infty} a_k \right) \cdot \left( \sum_{k=0}^{\infty} b_k \right) = \sum_{n=0}^{\infty} c_n,$$

also genau die Produktformel aus Satz 3.13.1.

## 3.14 Die Exponentialreihe

Wir lernen jetzt eine Reihe kennen, über die man die Exponentialfunktion  $e^x$  definieren kann.

Für  $x \in \mathbb{R}$  heißt

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

*Exponentialreihe.* Diese Reihe hängt also von der Zahl  $x \in \mathbb{R}$  als Parameter ab, ähnlich wie die geometrisch Reihe von  $q$  abhängt.

Wir untersuchen die Reihe zuerst auf Konvergenz und benutzen dazu das Quotientenkriterium:

$$\left| \frac{a_{k+1}}{a_k} \right| = \left| \frac{x^{k+1}}{(k+1)!} \cdot \frac{k!}{x^k} \right| = \left| \frac{x}{k+1} \right| = \frac{|x|}{k+1} \rightarrow 0 < 1 \quad \text{für } k \rightarrow \infty.$$

Die Exponentialreihe  $\sum_{k=0}^{\infty} \frac{x^k}{k!}$  ist also immer absolut konvergent, unabhängig vom Wert von  $x \in \mathbb{R}$ . (Man beachte den Unterschied zur geometrischen Reihe, wo die Konvergenz von  $q$  abhängt.)

Im nächsten Kapitel über Taylorreihen werden wir sehen, dass die Exponentialreihe  $\exp(x)$  genau die Exponentialfunktion  $e^x$  darstellt, d.h. es gilt  $\exp(x) = e^x$ . Damit gilt speziell

$$e = \exp(1) = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots$$

Man kann die Eulerzahl  $e$  also auch durch eine Reihe darstellen – eine Folge, die gegen  $e$  konvergiert, hatten wir in Beispiel 2.14.2 kennengelernt.

**Anwendung:** Häufig *definiert* man  $e^x$  durch die Exponentialreihe  $\exp(x)$ . Man muss dann mit Hilfe der Reihe alle Eigenschaften der Exponentialfunktion  $e^x$  beweisen. Als Beispiel machen wir das für die Potenzrechenregel  $e^{x+y} = e^x \cdot e^y$ . Wir benutzen dazu die Cauchyproduktformel:

$$\begin{aligned}
 e^x \cdot e^y &= \exp(x) \cdot \exp(y) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \cdot \sum_{k=0}^{\infty} \frac{y^k}{k!} \\
 &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{x^{n-k}}{(n-k)!} \cdot \frac{y^k}{k!} = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{k!(n-k)!} x^{n-k} y^k \\
 &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{n!} \binom{n}{k} x^{n-k} y^k = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \\
 &= \sum_{n=0}^{\infty} \frac{1}{n!} (x+y)^n = \exp(x+y) = e^{x+y}
 \end{aligned}$$

Hier haben wir die Definition des Binomialkoeffizienten  $\binom{n}{k}$  aus (1.8) sowie den allgemeinen binomische Satz 1.6.1 benutzt.

## Kapitel 4

# Taylorentwicklung und Potenzreihen

Bei der Taylorentwicklung wird eine beliebige gegebene Funktion durch einfache Funktionen, die Taylorpolynome, angenähert, also approximiert. Geht man von Taylorpolynomen zur Taylorreihe über, kann die gegebene Funktion sogar (in einem gewissen Bereich) exakt dargestellt werden. Taylorreihen sind eine spezielle Form von Potenzreihen, das sind Reihen, die Potenzen  $x^n$  beliebig hohen Grades enthalten.

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik B, Kapitel 4.

### 4.1 Taylorpolynome

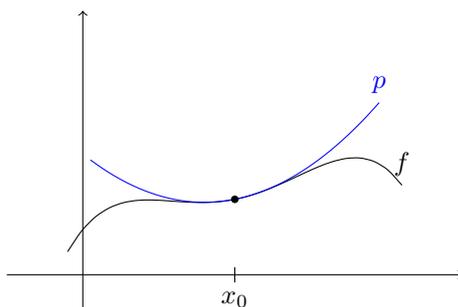
Es sei eine differenzierbare Funktion

$$f : ]a, b[ \rightarrow \mathbb{R}$$

gegeben. Die Funktion  $f(x)$  ist also definiert für  $x \in ]a, b[$ , d.h. ihr Definitionsbereich ist das offene Intervall  $]a, b[$ ,

$$D_f = ]a, b[ = \{x \in \mathbb{R} \mid a < x < b\}.$$

Unser Ziel ist, die Funktion  $f$  in der Nähe eines Punktes  $x_0 \in ]a, b[$  möglichst gut durch ein Polynom  $p$  anzunähern (zu approximieren), siehe Abbildung 4.1. Um das zu erreichen, sollte  $p$  im Punkt  $x_0$  denselben Funktionswert wie  $f$  haben, also  $p(x_0) = f(x_0)$ . Außerdem sollte die Steigung von  $p$  und  $f$  in  $x_0$  gleich sein, d.h.  $p'(x_0) = f'(x_0)$ . Um eine noch bessere Approximation zu erhalten, sollte auch die Krümmung beider Funktionen in  $x_0$  gleich sein; da die Krümmung durch die zweite Ableitung gegeben ist, sollte also auch  $p''(x_0) = f''(x_0)$  gelten. Man kann so immer höhere Ableitungen einbeziehen und erhält damit eine

Abbildung 4.1: Approximation von  $f$  durch Taylorpolynom  $p$ 

immer bessere Annäherung:

$$\begin{aligned} p(x_0) &= f(x_0) \\ p'(x_0) &= f'(x_0) \\ p''(x_0) &= f''(x_0) \\ &\vdots \\ p^{(n)}(x_0) &= f^{(n)}(x_0) \quad \leftarrow n. \text{ Ableitung} \end{aligned}$$

Dies sind  $n + 1$  Bedingungen an das Polynom  $p$ , kurz

$$p^{(k)}(x_0) = f^{(k)}(x_0) \quad \text{für } k = 0, 1, \dots, n. \quad (4.1)$$

Dabei ist  $f^{(0)}(x) = f(x)$ , sozusagen die „0. Ableitung“. Da ein Polynom  $n$ . Grades genau  $n + 1$  Parameter hat, kann man durch die obigen Bedingungen ein Polynom  $n$ . Grades bestimmen. Dazu machen wir für  $p$  den *Ansatz*:

$$p(x) = \sum_{k=0}^n a_k (x - x_0)^k = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n.$$

Durch Einsetzen der  $n + 1$  Bedingung bestimmen wir daraus jetzt die Parameter  $a_0, \dots, a_n$  von  $p$ .

$$\begin{aligned} p(x_0) &= a_0 \quad \Rightarrow \quad a_0 = f(x_0) \\ p'(x) &= \sum_{k=1}^n k a_k (x - x_0)^{k-1} = a_1 + 2a_2(x - x_0) + \dots + n a_n (x - x_0)^{n-1} \\ \Rightarrow p'(x_0) &= a_1 \quad \Rightarrow \quad a_1 = f'(x_0) \\ p''(x) &= \sum_{k=2}^n k(k-1) a_k (x - x_0)^{k-2} = 2a_2 + \dots + n(n-1)(x - x_0)^{n-2} \\ \Rightarrow p''(x_0) &= 2a_2 = f''(x_0) \quad \Rightarrow \quad a_2 = \frac{f''(x_0)}{2} \end{aligned}$$

u.s.w. Allgemein erhält man

$$a_k = \frac{f^{(k)}(x_0)}{k!}, \quad k = 0, 1, 2, \dots, n.$$

Setzen wir dieses Ergebnis in unseren Ansatz für  $p$  ein, ergibt sich

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

Dieses Polynom erfüllt also die Bedingungen (4.1); es ist das sogenannte Taylorpolynom:

**Definition 4.1.1** Das *Taylorpolynom*  $n$ . Grades von  $f$  im Entwicklungspunkt  $x_0$  ist

$$T_n(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k. \quad (4.2)$$

Ausgeschrieben ist das

$$T_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n.$$

**Beispiel 4.1.2** Wir berechnen das Taylorpolynom  $T_n(x)$  der Funktion  $f(x) = e^x$  im Entwicklungspunkt  $x_0 = 0$ . Die Ableitungen von  $f$  sind

$$f'(x) = e^x = f''(x) = \cdots = f^{(n)}(x).$$

Im Entwicklungspunkt  $x_0 = 0$  ist das

$$f(0) = f'(0) = \cdots = f^{(n)}(0) = e^0 = 1.$$

Einsetzen in (4.2) ergibt dann

$$\begin{aligned} T_n(x) &= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n \\ &= 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n \end{aligned}$$

oder

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

## 4.2 Approximationsfehler und Restglied

Unsere Herleitung der Formel für das Taylorpolynom suggeriert, dass

$$f(x) \approx T_n(x) \quad \text{wenn } x \text{ „nah“ bei } x_0 \text{ ist,}$$

siehe auch Abbildung 4.1. Die Frage ist dann: Wie gut ist die Approximation von  $f(x)$  durch das Taylorpolynom  $T_n(x)$  genau? Oder anders gefragt: Wie groß ist der Fehler?

Dazu definieren wir zuerst den *Approximationsfehler*  $R_n(x)$ ,

$$R_n(x) = f(x) - T_n(x),$$

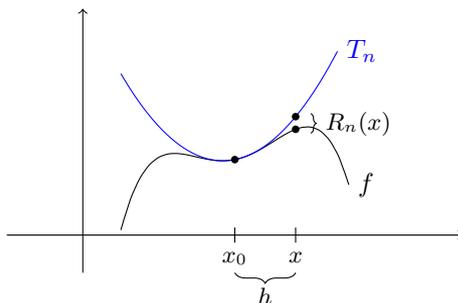


Abbildung 4.2: Approximationsfehler der Taylorentwicklung

siehe auch Abbildung 4.2. Anders gesagt ergibt sich also die Funktion  $f(x)$  als Summe aus Taylorpolynom und Approximationsfehler,

$$f(x) = T_n(x) + R_n(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + R_n(x). \quad (4.3)$$

Man nennt (4.3) auch *Taylorentwicklung* von  $f$  bis zum  $n$ . Grad und  $R_n(x)$  das *Restglied* der Entwicklung.

Zur Illustration betrachten wir das Taylorpolynom der Exponentialfunktion vom letzten Beispiel.

**Beispiel 4.2.1** Sei  $f(x) = e^x$  und  $x_0 = 0$  der Entwicklungspunkt. Das Taylorpolynom 2. Grades von  $f$  ist dann

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

Für zwei Werte von  $x$  „nahe“ bei  $x_0 = 0$  berechnen wir den Approximationsfehler:

$$\begin{aligned} x = 0.1 : \quad & f(0.1) = 1.1051709\dots \\ & T_2(0.1) = 1.105 \\ \Rightarrow \text{Fehler } R_2(0.1) & \approx 1.7 \cdot 10^{-4} \end{aligned}$$

$$\begin{aligned} x = 0.05 : \quad & f(0.05) = 1.0512710\dots \\ & T_2(0.05) = 1.05125 \\ \Rightarrow \text{Fehler } R_2(0.05) & \approx 2.1 \cdot 10^{-5} \end{aligned}$$

Wir sehen, dass der Fehler tatsächlich recht klein ist, und bei Annäherung an den Entwicklungspunkt  $x_0$  kleiner wird. Das Taylorpolynom approximiert die Funktion also gut.

Wir leiten jetzt eine allgemeine Formel für das Restglied  $R_n(x)$  her. Damit werden wir im nächsten Abschnitt die Größe des Approximationsfehlers abschätzen, ohne wie eben Funktionswerte ausrechnen zu müssen.

**Satz 4.2.2 (Restglied in Lagrange-Form)** *Es gilt*

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}. \quad (4.4)$$

Die Zahl  $\xi$  ist dabei abhängig von  $x$ ,  $x_0$  und  $n$ , und es gilt

$$\begin{aligned} \xi &\in [x_0, x] && \text{falls } x > x_0, \\ \xi &\in [x, x_0] && \text{falls } x < x_0. \end{aligned}$$

**Beweis.** Wir beweisen die Formel nur für den einfachsten Fall  $n = 0$ . Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$\int_{x_0}^x f'(t) dt = f(x) - f(x_0) = f(x) - T_0(x) = R_0(x).$$

Aus dem Mittelwertsatz der Integralrechnung (Mathematik A) folgt dann

$$R_0(x) = \int_{x_0}^x f'(t) dt = (x - x_0) \cdot f'(\xi) = \frac{f'(\xi)}{1!} (x - x_0)$$

mit einem  $\xi$  zwischen  $x$  und  $x_0$ . Dies ist genau (4.4) für  $n = 0$ . Den allgemeine Fall beweist man ähnlich mit partieller Integration und vollständiger Induktion.  $\square$

### 4.3 Restgliedabschätzung

Wir schauen uns jetzt an, wie man den Fehler bei der Taylorapproximation abschätzen kann. Das macht es dann z.B. möglich, aus dem Taylorpolynom  $T_n(x)$  einen Näherungswert für die Funktion  $f(x)$  zu berechnen.

In der Lagrange-Form (4.4) des Restglieds steht der Ausdruck  $f^{(n+1)}(\xi)$ . Da von  $\xi$  nur bekannt ist, dass es zwischen  $x$  und  $x_0$  liegt, braucht man eine allgemeine Abschätzung für die Ableitung  $f^{(n+1)}(x)$ . Wir suchen also eine Konstante  $M > 0$  für die gilt

$$|f^{(n+1)}(\xi)| \leq M \quad \text{für alle} \quad \begin{cases} \xi \in [x_0, x] & \text{falls } x > x_0, \\ \xi \in [x, x_0] & \text{falls } x < x_0. \end{cases}$$

Bezeichnen wir mit  $h = |x - x_0|$  den Abstand von  $x$  zu  $x_0$ , dann folgt aus (4.4)

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

Für den Approximationsfehler der Taylorentwicklung (siehe Abbildung 4.2) gilt also

$$|f(x) - T_n(x)| = |R_n(x)| \leq \frac{M}{(n+1)!} \cdot h^{n+1}.$$

**Beispiel 4.3.1** Wir betrachten wieder die Exponentialfunktion  $f(x) = e^x$  und den Entwicklungspunkt  $x_0 = 0$ . Das zugehörige Taylorpolynom  $n$ . Grades ist

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

siehe Beispiel 4.1.2, und für das Restglied in Lagrange-Form erhalten wir

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

Wir berechnen jetzt eine Restgliedabschätzung für den Punkt  $x = \frac{1}{10}$ . Das Restglied ist dann

$$R_n\left(\frac{1}{10}\right) = \frac{e^\xi}{(n+1)!} \left(\frac{1}{10}\right)^{n+1} \quad \text{mit } \xi \in [x_0, x] = \left[0, \frac{1}{10}\right].$$

Wir müssen damit  $|e^\xi|$  für  $\xi \in [0, \frac{1}{10}]$  abschätzen:

$$|e^\xi| = e^\xi \leq e^{\frac{1}{10}} \leq e^{\frac{1}{2}} = \sqrt{e} \leq \sqrt{4} = 2.$$

Hier haben wir benutzt, dass  $e^x$  monoton wachsend ist. Für den maximalen Fehler  $|R_n(\frac{1}{10})|$  erhalten wir damit

$$|R_n\left(\frac{1}{10}\right)| = \frac{|e^\xi|}{(n+1)!} \cdot \frac{1}{10^{n+1}} \leq \frac{2}{(n+1)! \cdot 10^{(n+1)}}.$$

Für die ersten Werte von  $n$  ergeben sich folgende Abschätzung für den Fehler:

$n$	1	2	3	4
$ R_n(\frac{1}{10})  \leq \dots$	$\frac{2}{2! \cdot 10^2} = 10^{-2}$	$\frac{2}{3! \cdot 10^3} \approx 3.3 \cdot 10^{-4}$	$8.3 \cdot 10^{-6}$	$1.7 \cdot 10^{-7}$

Damit kann man jetzt Näherungswerte für  $e^{\frac{1}{10}}$  berechnen: Für  $n = 2$  ist

$$T_2\left(\frac{1}{10}\right) = 1 + \frac{1}{10} + \frac{1}{2!} \left(\frac{1}{10}\right)^2 = \frac{221}{200} = 1.105$$

Mit der Abschätzung für den Fehler folgt damit

$$|e^{\frac{1}{10}} - 1.105| = |R_2\left(\frac{1}{10}\right)| \leq \frac{2}{3! \cdot 10^3} = \frac{1}{3} \cdot 10^{-3} = 0.00033 \dots$$

also

$$\begin{aligned} 1.105 - 0.00033 \dots &\leq e^{\frac{1}{10}} \leq 1.105 + 0.00033 \dots \\ \Leftrightarrow 1.10466 \dots &\leq e^{\frac{1}{10}} \leq 1.10533 \dots \end{aligned}$$

Für  $n = 3$  ist

$$\begin{aligned} T_3\left(\frac{1}{10}\right) &= \frac{6631}{6000} = 1.105166 \dots, \\ |R_3\left(\frac{1}{10}\right)| &\leq \frac{2}{4! \cdot 10^4} \approx 8.3 \cdot 10^{-6} = 0.0000083 \\ \Rightarrow 1.1051583 \dots &\leq e^{\frac{1}{10}} \leq 1.1051749 \dots \end{aligned}$$

Der Wert von  $e^{\frac{1}{10}}$  bis zur vierten Nachkommastelle ist also 1.1051. Auf diese Weise kann man, mit höheren  $n$ , den Wert von  $e^x$  beliebig genau berechnen. Beachten Sie, dass wir an keiner Stelle des Beispiels konkrete Werte von  $e^x$  als gegeben benutzt haben! Wir haben letztlich nur die Grundrechenarten Addition, Subtraktion, Multiplikation und Division verwendet.

Mit dieser Methode (und ein paar weiteren Tricks) werden in Taschenrechnern und Computern übrigens tatsächlich Funktionen wie  $e^x$  oder  $\ln x$ ,  $\sin x$ ,  $\cos x$  berechnet.

## 4.4 Taylorreihen

Im letzten Beispiel haben wir gesehen, wie der Fehler bei der Approximation von  $f(x)$  durch Taylorpolynome immer kleiner wird, je größer der Grad  $n$  des Taylorpolynoms ist. Und dies war ja auch eine der Ideen bei der Herleitung der Formel für das Taylorpolynom in Abschnitt 4.1. Wenn für  $n \rightarrow \infty$  der Fehler gegen Null konvergiert, sollten wir also  $f(x)$  exakt zurückerhalten. Beim Grenzübergang  $n \rightarrow \infty$  wird aus dem Taylorpolynom die Taylorreihe:

**Definition 4.4.1** Die Funktion  $f$  sei unendlich oft differenzierbar. Die Reihe

$$T_f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \quad (4.5)$$

heißt *Taylorreihe* von  $f$  im Entwicklungspunkt  $x_0$ .

Es gilt also (formal)

$$\lim_{n \rightarrow \infty} T_n(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = T_f(x).$$

**Beispiel 4.4.2** (a) Taylorreihe von  $f(x) = e^x$  im Entwicklungspunkt  $x_0 = 0$ :  
Wir haben in Beispiel 4.1.2 das Taylorpolynom  $n$ . Grades schon ausgerechnet. Bei  $n \rightarrow \infty$  ergibt sich so die Taylorreihe

$$T_{\exp}(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots = \sum_{k=0}^{\infty} \frac{1}{k!}x^k.$$

Diese Reihe haben wir schon untersucht; es ist die Exponentialreihe aus 3.14.

(b) Taylorreihe von  $\sin(x)$  in  $x_0 = 0$ . Zuerst berechnen wir die Ableitungen im Entwicklungspunkt  $x_0$ :

$$\begin{aligned} f(x) = \sin(x) &\Rightarrow f(0) = 0 \\ f'(x) = \cos(x) &\Rightarrow f'(0) = 1 \\ f''(x) = -\sin(x) &\Rightarrow f''(0) = 0 \\ f'''(x) = -\cos(x) &\Rightarrow f'''(0) = -1 \\ f^{(4)}(x) = \sin(x) &\Rightarrow f^{(4)}(0) = 0 \end{aligned}$$

u.s.w. Bei der 4. Ableitung sind wir wieder bei der Anfangsfunktion  $\sin(x)$  und die Ergebnisse wiederholen sich. Einsetzen in (4.5) ergibt dann

$$T_{\sin}(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{2k+1}.$$

Die Taylorreihe von Sinus enthält also nur die ungeraden Potenzen von  $x$ . Der Term  $2k+1$  erzeugt für  $k \geq 0$  die Zahlen  $1, 3, 5, \dots$ . Zusammen mit  $(-1)^k$  und dem Start der Reihe mit  $k=0$  ergibt das genau die richtigen Ausdrücke inklusive Vorzeichen.

Für jede Taylorreihe stellen sich zwei prinzipielle Fragen:

- Wann, d.h. für welche  $x$  konvergiert die Taylorreihe  $T_f(x)$ ?
- Wenn die Reihe konvergiert, gilt dann  $f(x) = T_f(x)$ , d.h. wird die Funktion durch die Reihe dargestellt? Oder: Für welche  $x$  gilt diese Gleichung?

**Satz 4.4.3** *Es gilt  $f(x) = T_f(x)$ , also*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \quad (4.6)$$

genau dann wenn  $\lim_{n \rightarrow \infty} R_n(x) = 0$ . Dabei ist  $R_n(x)$  das Restglied aus (4.3). Die Taylorreihe  $T_f(x)$  konvergiert also genau für die  $x$  gegen  $f(x)$ , für die  $\lim_{n \rightarrow \infty} R_n(x) = 0$  gilt.

Gleichung (4.6) nennt man auch *Taylorentwicklung* der Funktion  $f(x)$ .

**Beweis von Satz 4.4.3.** Aus (4.3)

$$f(x) = T_n(x) + R_n(x)$$

folgt bei  $n \rightarrow \infty$  sofort

$$f(x) = \lim_{n \rightarrow \infty} T_n(x) \iff \lim_{n \rightarrow \infty} R_n(x) = 0,$$

und es gilt

$$\lim_{n \rightarrow \infty} T_n(x) = T_f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k.$$

□

Für die Funktionen  $e^x$ ,  $\sin x$ ,  $\cos x$  gelten folgende Taylorentwicklungen:

**Satz 4.4.4** *Für alle  $x \in \mathbb{R}$  gilt*

$$\begin{aligned} e^x &= 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots = \sum_{k=0}^{\infty} \frac{1}{k!}x^k \\ \sin x &= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{2k+1} \\ \cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}x^{2k} \end{aligned}$$

$e^x$ ,  $\sin x$ ,  $\cos x$  werden also tatsächlich für alle  $x \in \mathbb{R}$  durch ihre Taylorreihen (mit  $x_0 = 0$ ) dargestellt.

**Beachte:** Die Taylorreihe von  $\sin x$  enthält genau die *ungeraden* Terme der Taylorreihe von  $e^x$  (also der Exponentialreihe), jeweils mit einem alternierenden Vorzeichen. Die Taylorreihe von  $\cos x$  dagegen enthält alle *geraden* Terme der  $e^x$ -Reihe, mit alternierendem Vorzeichen. Dies ist auch gekoppelt mit der Symmetrie von Sinus und Cosinus: Da die Sinusreihe nur ungerade Potenzen von  $x$  enthält, gilt offenbar  $\sin(-x) = -\sin x$ , d.h.  $\sin$  ist eine ungerade Funktion! Entsprechend gilt  $\cos(-x) = \cos x$  ( $\cos$  ist eine gerade Funktion), denn die Cosinusreihe enthält nur gerade Potenzen von  $x$ .

**Beweis Satz 4.4.4.**

(a)  $f(x) = e^x$ . Wir wissen schon, dass

$$T_{\text{exp}}(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

die Taylorreihe von  $e^x$  ist. Um jetzt  $e^x = T_{\text{exp}}(x)$  zu beweisen, müssen wir nach Satz 4.4.3  $\lim_{n \rightarrow \infty} R_n(x) = 0$  zeigen. Dazu benutzen wir eine Restgliedabschätzung wie in Beispiel 4.3.1. Für die Lagrange-Form des Restglieds gilt

$$|R_n(x)| = \left| \frac{e^{\xi}}{(n+1)!} x^{n+1} \right| = \frac{e^{\xi}}{(n+1)!} |x|^{n+1} \leq \frac{e^{|x|}}{(n+1)!} |x|^{n+1},$$

denn da  $\xi$  zwischen  $x_0 = 0$  und  $x$  liegt, gilt  $\xi \leq |\xi| \leq |x|$  und aus der Monotonie von  $e^x$  folgt  $e^{\xi} \leq e^{|x|}$ . Aus 3.14 wissen wir, dass die Exponentialreihe absolut konvergiert, d.h.

$$\sum_{k=0}^{\infty} \frac{|x|^k}{k!} < \infty$$

ist konvergent. Mit dem notwendigen Kriterium Satz 3.4.1 für die Konvergenz von Reihen folgt daraus

$$\lim_{n \rightarrow \infty} \frac{|x|^n}{n!} = 0 \quad (4.7)$$

und damit auch

$$\lim_{n \rightarrow \infty} e^{|x|} \cdot \underbrace{\frac{|x|^{n+1}}{(n+1)!}}_{\rightarrow 0} = 0,$$

also in der Tat  $\lim_{n \rightarrow \infty} R_n(x) = 0$  für jedes  $x \in \mathbb{R}$ .

(b)  $f(x) = \cos(x)$ . Wir berechnen zuerst die Taylorreihe für den Entwicklungspunkt  $x_0 = 0$ :

$$\begin{aligned} f(x) = \cos(x) &\Rightarrow f(0) = 1 \\ f'(x) = -\sin(x) &\Rightarrow f'(0) = 0 \\ f''(x) = -\cos(x) &\Rightarrow f''(0) = -1 \\ f'''(x) = \sin(x) &\Rightarrow f'''(0) = 0 \\ f^{(4)}(x) = \cos(x) &\Rightarrow f^{(4)}(0) = 1 \\ &\vdots \qquad \qquad \qquad \vdots \end{aligned}$$

Die Taylorreihe ist also

$$\begin{aligned} T_{\cos}(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \\ &= 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \frac{1}{6!} x^6 + \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}. \end{aligned}$$

Jetzt zeigen wir wieder  $\lim_{n \rightarrow \infty} R_n(x) = 0$ . Für das Restglied erhalten wir

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} \quad \text{mit} \quad f^{(n+1)}(\xi) = \begin{cases} \sin \xi & \text{für } n+1 \text{ ungerade,} \\ \cos \xi & \text{für } n+1 \text{ gerade.} \end{cases}$$

Da  $|\sin \xi| \leq 1$  und  $|\cos \xi| \leq 1$  für jedes  $\xi$ , folgt zusammen mit (4.7)

$$|R_n(x)| \leq \frac{1}{(n+1)!} |x|^{n+1} \rightarrow 0 \quad \text{bei } n \rightarrow \infty$$

für jedes  $x \in \mathbb{R}$ . Nach Satz 4.4.3 gilt also  $\cos(x) = T_{\cos}(x)$ . Der Beweis für die Taylorentwicklung von  $\sin(x)$  läuft analog.

□

## 4.5 Potenzreihen

Potenzreihen sind Reihen, die von einer Variablen  $x$  abhängen, wobei  $x$  als Potenzen  $x^k$  in der Reihe vorkommt.

**Definition 4.5.1** Die Reihe

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k \tag{4.8}$$

heißt *Potenzreihe* mit Entwicklungspunkt  $x_0$  und Koeffizienten  $a_k$  ( $x_0, a_k \in \mathbb{R}$ ).

**Beispiel 4.5.2** Folgende Reihen, die wir schon kennengelernt haben, sind Potenzreihen:

(a) Die Exponentialreihe

$$\sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

ist eine Potenzreihe mit Koeffizienten  $a_k = \frac{1}{k!}$  und Entwicklungspunkt  $x_0 = 0$ .

(b) Die Taylorreihe des Cosinus

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}$$

ist eine Potenzreihe mit Entwicklungspunkt  $x_0 = 0$ , die nur gerade Potenzen von  $x$  enthält. Wir können die Reihe in der Form (4.8) schreiben mit den Koeffizienten

$$a_{2k} = \frac{(-1)^k}{(2k)!}, \quad a_{2k+1} = 0 \quad \text{für } k \in \mathbb{N}.$$

(c) Jede Taylorreihe ist eine Potenzreihe wobei

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

Wie bei jeder Reihe stellt sich wieder die Frage nach der Konvergenz. Hier wird die Konvergenz von  $x$  abhängig sein; die Frage ist daher: Für welche  $x \in \mathbb{R}$  konvergiert die Potenzreihe?

Zur Beantwortung der Frage gehen wir in zwei Schritten vor:

1. Substitution  $\tilde{x} = x - x_0$ . Damit gilt

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k = \sum_{k=0}^{\infty} a_k \tilde{x}^k.$$

Die allgemeine Potenzreihe (4.8) mit der Variablen  $x$  entspricht also einer Potenzreihe mit der Variablen  $\tilde{x}$  und Entwicklungspunkt 0. Bei der Untersuchung des Konvergenzverhaltens von Potenzreihen langt es deshalb, nur den Entwicklungspunkt  $x_0 = 0$  zu betrachten. (Anders gesagt schreiben wir für  $\tilde{x}$  wieder  $x$ .)

2. Quotientenkriterium. Auf die Potenzreihen mit  $x_0 = 0$

$$\sum_{k=0}^{\infty} a_k x^k$$

wenden wir jetzt das Quotientenkriterium an: Sei

$$b = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1} x^{k+1}}{a_k x^k} \right| = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| \cdot |x|,$$

falls der Grenzwert existiert. Nach dem Quotientenkriterium gilt also

$$\begin{cases} b < 1 & \Rightarrow \text{Reihe konvergiert absolut} \\ b > 1 & \Rightarrow \text{Reihe divergiert} \end{cases}$$

Wir setzen jetzt

$$r = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| \quad (4.9)$$

(falls Grenzwert existiert). Damit gilt (falls  $r \neq 0$ )

$$b = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| \cdot |x| = \frac{1}{r} \cdot |x|,$$

also

$$b = \frac{|x|}{r} \begin{cases} < 1 & \Leftrightarrow |x| < r \\ > 1 & \Leftrightarrow |x| > r \end{cases}$$

Wir erhalten damit:

$$\sum_{k=0}^{\infty} a_k x^k \begin{cases} \text{konvergiert absolut für} & |x| < r & (\Leftrightarrow -r < x < r) \\ \text{divergiert für} & |x| > r & (\Leftrightarrow x < -r \vee x > r) \end{cases}$$

Man kann nun zeigen, dass es für jede Potenzreihe solch eine Zahl  $r$  gibt – selbst wenn der Grenzwert in (4.9) nicht existiert. (Gleichung (4.9) ist in diesem Fall nicht richtig.) Die Zahl  $r$  heißt *Konvergenzradius* der Potenzreihe.

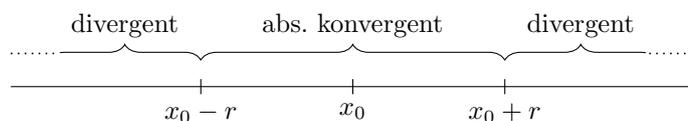


Abbildung 4.3: Konvergenzradius einer Potenzreihe

**Satz 4.5.3 (Konvergenzradius)** *Zu jeder Potenzreihe*

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

existiert eine Zahl  $r \geq 0$  (der Konvergenzradius) sodass die Reihe

- absolut konvergiert für  $|x - x_0| < r$ ,
- divergiert für  $|x - x_0| > r$ .

Bei  $|x - x_0| = r$  gibt es keine allgemeine Aussage über die Konvergenz der Reihe. Für den Konvergenzradius ist auch  $r = \infty$  möglich. In diesem Fall ist die Reihe absolute konvergent für alle  $x \in \mathbb{R}$ .

Die Bereiche von Konvergenz und Divergenz einer Potenzreihe sind also

$$\begin{aligned} \text{Konvergenz: } |x - x_0| < r &\Leftrightarrow x_0 - r < x < x_0 + r \\ &\Leftrightarrow x \in ]x_0 - r, x_0 + r[ \\ \text{Divergenz: } |x - x_0| > r &\Leftrightarrow x < x_0 - r \vee x > x_0 + r \\ &\Leftrightarrow x \in ]-\infty, x_0 - r[ \cup ]x_0 + r, \infty[ \end{aligned}$$

Graphisch ist das in Abbildung 4.3 veranschaulicht.

Zur Berechnung des Konvergenzradius stehen folgende Formeln zur Verfügung:

**Satz 4.5.4 (Berechnung des Konvergenzradius)** *Für den Konvergenzradius  $r$  gilt (bei  $a_k \neq 0$ )*

$$r = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| \quad \text{und} \quad r = \lim_{k \rightarrow \infty} \frac{1}{\sqrt[k]{|a_k|}} \quad (4.10)$$

sofern der jeweilige Grenzwert existiert.

**Beweis.** Die erste Formel haben wir in (4.9) bereits aus dem Quotientenkriterium hergeleitet. Die zweite Formel erhält man ähnlich aus dem Wurzelkriterium.  $\square$

**Bemerkung:**

- (a) Jede Potenzreihe hat einen Konvergenzradius  $r$ . Wenn einer (oder beide) der Grenzwerte in (4.10) nicht existieren, heißt das nur, dass man mit dieser Formel  $r$  nicht bestimmen kann.

(b) Die zweite Formel in (4.10) ist ein Spezialfall der *Formel von Hadamard*:

$$r = \liminf_{k \rightarrow \infty} \frac{1}{\sqrt[k]{|a_k|}}.$$

Der Limes in dieser Formel ist der „Limes Inferior“: er existiert immer und ist der kleinste Häufungswert einer Folge.

**Beispiel 4.5.5** Wir berechnen den Konvergenzradius der folgenden Potenzreihen:

(a) Für die Potenzreihe

$$\sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

gilt

$$a_k = \frac{1}{k!}.$$

Mit der ersten Formel in (4.10) folgt dann

$$r = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| = \lim_{k \rightarrow \infty} \frac{1}{k!} \cdot \frac{(k+1)!}{1} = \lim_{k \rightarrow \infty} (k+1) = \infty.$$

Der Konvergenzradius ist also  $r = \infty$ , d.h. die Reihe konvergiert (absolut) für alle  $x \in \mathbb{R}$ .

Dieses Ergebnis hätten wir alternativ auch wie folgt bekommen können: Die Potenzreihe ist die Exponentialreihe, von der wir aus 3.14 schon wissen, dass sie für alle  $x \in \mathbb{R}$  konvergiert. Also muss der Konvergenzradius  $r = \infty$  sein.

(b) Für die Potenzreihe

$$\sum_{k=1}^{\infty} \frac{-1}{k} x^k$$

ist

$$a_k = \frac{-1}{k}.$$

Damit folgt

$$\left| \frac{a_k}{a_{k+1}} \right| = \left| \frac{-1}{k} \cdot \frac{k+1}{-1} \right| = \frac{k+1}{k} \rightarrow 1 \quad \text{bei } k \rightarrow \infty.$$

Also ist der Konvergenzradius  $r = 1$ . D.h. die Reihe konvergiert für  $|x| < 1$  ( $\Leftrightarrow -1 < x < 1$ ) und divergiert für  $|x| > 1$ .

Die Formeln (4.10) zur Berechnung des Konvergenzradius kann man nur benutzen, wenn  $a_k \neq 0$  für alle Koeffizienten  $a_k$  gilt (oder alle mit  $k \geq k_0$ ). Bei Potenzreihen der Form

$$\sum_{k=0}^{\infty} a_{2k} x^{2k} \quad \text{oder} \quad \sum_{k=0}^{\infty} a_{2k+1} x^{2k+1} \quad (4.11)$$

sind aber alle ungeraden bzw. alle geraden Koeffizienten gleich Null – man kann deshalb (4.10) nicht direkt anwenden. Zur Berechnung des Konvergenzradius macht man hier zunächst die Substitution  $y = x^2$  und erhält so eine Potenzreihe in der neuen Variablen  $y$ :

$$\begin{aligned}\sum_{k=0}^{\infty} a_{2k} x^{2k} &= \sum_{k=0}^{\infty} a_{2k} (x^2)^k = \sum_{k=0}^{\infty} b_k y^k \\ \sum_{k=0}^{\infty} a_{2k+1} x^{2k+1} &= x \cdot \sum_{k=0}^{\infty} a_{2k+1} (x^2)^k = x \cdot \sum_{k=0}^{\infty} b_k y^k\end{aligned}$$

Die neuen Koeffizienten sind

$$b_k = a_{2k} \quad \text{bzw.} \quad b_k = a_{2k+1}.$$

Man berechnet nun den Konvergenzradius  $r_y$  der Reihe  $\sum_{k=0}^{\infty} b_k y^k$ , also bezüglich der Variablen  $y$ , z.B. mit (4.10). Dann bestimmt man den Konvergenzbe-  
reich bezüglich  $x$  durch Rücksubstitution von  $y = x^2$  in die Konvergenzbedin-  
gung für  $y$ : Die Reihe konvergiert für

$$|y| < r_y \iff x^2 < r_y \iff |x| < \sqrt{r_y}.$$

Damit ist  $r_x = \sqrt{r_y}$  der Konvergenzradius der Reihen (4.11).

**Beispiel 4.5.6** Wir bestimmen den Konvergenzradius von

$$\sum_{k=0}^{\infty} 3^k x^{2k+1}.$$

Mit der Substitution  $y = x^2$  ist

$$\sum_{k=0}^{\infty} 3^k x^{2k+1} = x \cdot \sum_{k=0}^{\infty} 3^k x^{2k} = x \cdot \sum_{k=0}^{\infty} \underbrace{3^k}_{b_k} y^k.$$

Für den Konvergenzradius bezüglich  $y$  erhalten wir dann

$$r_y = \lim_{k \rightarrow \infty} \left| \frac{b_k}{b_{k+1}} \right| = \lim_{k \rightarrow \infty} \frac{3^k}{3^{k+1}} = \frac{1}{3}.$$

Die Reihe ist somit konvergent für

$$|y| < \frac{1}{3} \iff x^2 < \frac{1}{3} \iff |x| < \sqrt{\frac{1}{3}}.$$

Also ist  $r_x = \sqrt{\frac{1}{3}}$  der gesuchte Konvergenzradius der Reihe bezüglich  $x$ .

## 4.6 Ableitung und Integration von Potenzreihen

Potenzreihen liefern einen von der Variablen  $x$  abhängigen Reihenwert. Man kann sie daher als Funktion

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \tag{4.12}$$

auffassen. Die Funktion ist dann für die  $x$  definiert, für die die Reihe konvergiert, d.h. innerhalb des Konvergenzradius, für  $|x| < r$ . Der Definitionsbereich ist also

$$D_f = ] - r, r[.$$

Als Funktion macht es Sinn, die Ableitung und das Integral einer Potenzreihe zu untersuchen.

**Satz 4.6.1** Sei  $\sum_{k=0}^{\infty} a_k x^k$  eine Potenzreihe mit Konvergenzradius  $r$ . Dann ist die in (4.12) definierte Funktion  $f$  unendlich oft differenzierbar mit den Ableitungen

$$\begin{aligned} f'(x) &= \sum_{k=1}^{\infty} k a_k x^{k-1} \\ f''(x) &= \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2} \\ &\vdots \end{aligned}$$

Alle Ableitungen sind wieder Potenzreihen mit dem gleichen Konvergenzradius  $r$  wie die Ausgangsreihe. Für  $a, b \in ] - r, r[$  ist das Integral von  $f$

$$\int_a^b \sum_{k=0}^{\infty} a_k x^k dx = \sum_{k=0}^{\infty} a_k \int_a^b x^k dx$$

Alle Aussagen gelten entsprechend für Reihen  $\sum_{k=0}^{\infty} a_k (x - x_0)^k$  mit Entwicklungspunkt  $x_0 \neq 0$ .

Kurz gesagt darf man also Potenzreihen *gliedweise ableiten und integrieren*.

Jede Taylorreihe ist eine Potenzreihe. Andererseits definiert nach dem letzten Satz jede Potenzreihe eine unendlich oft differenzierbare Funktion, zu der wir dann wiederum eine Taylorreihe berechnen können. Der nächste Satz sagt, dass diese beiden Reihen identisch sind.

**Satz 4.6.2** Ist  $f$  durch eine Potenzreihe mit Entwicklungspunkt  $x_0$  dargestellt, d.h. gilt

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k \quad \text{für} \quad |x - x_0| < r,$$

dann ist diese Reihe die Taylorreihe von  $f$  in  $x_0$ , d.h. es gilt

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

**Beweis.** Nach Satz 4.6.1 dürfen wir die Potenzreihe gliedweise ableiten:

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} a_k (x - x_0)^k = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + \dots \\ f'(x) &= \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1} = a_1 + 2a_2(x - x_0) + 3a_3(x - x_0)^2 + \dots \\ f''(x) &= \sum_{k=2}^{\infty} k(k-1) a_k (x - x_0)^{k-2} = 2a_2 + 3 \cdot 2a_3(x - x_0) + \dots \\ f'''(x) &= \sum_{k=3}^{\infty} k(k-1)(k-2) a_k (x - x_0)^{k-3} = 3 \cdot 2a_3 + \dots \\ &\vdots \end{aligned}$$

Einsetzen von  $x_0$  liefert

$$f(x_0) = a_0, \quad f'(x_0) = a_1, \quad f''(x_0) = 2a_2, \quad f'''(x_0) = 3 \cdot 2a_3, \dots$$

allgemein

$$f^{(k)}(x_0) = k! \cdot a_k$$

und damit

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

Also sind die Koeffizienten der Potenzreihe genau die Taylorkoeffizienten, d.h. die Reihe ist die Taylorreihe (4.5).  $\square$

Wir erläutern an zwei Beispielen das *Rechnen mit Potenzreihen*, speziell Differenzieren und Integrieren.

**Beispiel 4.6.3** (a) Wir wissen für die geometrische Reihe (siehe 3.2)

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \dots \quad \text{für } |x| < 1. \quad (4.13)$$

Diese Reihe ist offenbar eine Potenzreihe (mit  $a_k = 1$ ). Die Reihe konvergiert für  $|x| < 1$  und divergiert für  $|x| \geq 1$  (Satz 3.4.3). Damit ist der Konvergenzradius also  $r = 1$ . Wir leiten jetzt (4.13) ab. Zunächst gilt mit der Kettenregel

$$\begin{aligned} \left( \frac{1}{1-x} \right)' &= ((1-x)^{-1})' = (-1) \cdot (1-x)^{-2} \cdot (-1) \\ &= (1-x)^{-2} = \frac{1}{(1-x)^2} \end{aligned}$$

Also folgt aus (4.13)

$$\left( \frac{1}{1-x} \right)' = \frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} k x^{k-1} = 1 + 2x + 3x^2 + \dots, \quad |x| < 1.$$

Daraus können wir jetzt (zum Beispiel) durch Multiplikation mit  $x$  die Formel

$$\sum_{k=1}^{\infty} kx^k = \frac{x}{(1-x)^2}, \quad |x| < 1,$$

erhalten. Wir haben somit den Wert der Reihe  $\sum_{k=1}^{\infty} kx^k$  berechnet.

(b) Wir betrachten die Exponentialreihe

$$\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k,$$

siehe Abschnitt 3.14; ihr Konvergenzradius ist  $r = \infty$  (Beispiel 4.5.5). Nach Satz 4.6.1 ist die Funktion  $\exp(x)$  unendlich oft differenzierbar. Wir berechnen die Ableitung:

$$\begin{aligned} \exp(x) &= 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots \\ \Rightarrow \exp'(x) &= 0 + 1 + \frac{2}{2!}x + \frac{3}{3!}x^2 + \dots = 1 + x + \frac{1}{2!}x^2 + \dots = \exp(x) \end{aligned}$$

Die Ableitung von  $\exp$  ist also wieder  $\exp$ . Weiter gilt

$$\exp(0) = 1 + 0 + 0 + \dots = 1.$$

Die Funktion  $f(x) = \exp(x)$  erfüllt also, genau wie  $e^x$ , die Eigenschaften  $f'(x) = f(x)$  und  $f(0) = 1$ . Weil diese Eigenschaften  $e^x$  schon eindeutig festlegen, kann man die Exponentialfunktion  $e^x$  durch die Exponentialreihe  $\exp(x)$  definieren:

$$e^x := \exp(x).$$

## 4.7 Taylorentwicklung des Logarithmus

Als weiteres Beispiel für das Rechnen mit Potenzreihen bestimmen wir jetzt eine Potenzreihe, die den Logarithmus  $\ln x$  darstellt; wir erhalten damit die Taylorentwicklung des Logarithmus.

Wir gehen wieder von der geometrischen Reihe aus:

$$\frac{1}{1-q} = \sum_{k=0}^{\infty} q^k, \quad |q| < 1.$$

Ersetzen wir hier  $q = -x$ , ergibt sich

$$\frac{1}{1+x} = \sum_{k=0}^{\infty} (-x)^k = \sum_{k=0}^{\infty} (-1)^k x^k, \quad \text{für } |x| < 1. \quad (4.14)$$

Jetzt integrieren wir diese Gleichung über das Intervall  $[0, x]$  wobei  $|x| < 1$ . Da die rechte Seite eine Potenzreihe ist, können wir nach Satz 4.6.1 gliedweise integrieren und erhalten damit

$$\begin{aligned} \int_0^x \frac{1}{1+t} dt &= \sum_{k=0}^{\infty} \int_0^x (-1)^k t^k dt = \sum_{k=0}^{\infty} \left[ \frac{(-1)^k}{k+1} t^{k+1} \right]_0^x = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} \\ &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k \end{aligned}$$

Im letzten Schritt haben wir benutzt, dass  $(-1)^{k-1} = \pm 1$  ist, abhängig davon, ob  $k-1$  gerade oder ungerade ist. Da  $k-1$  und  $k+1$  gleichzeitig gerade bzw. ungerade sind, gilt also  $(-1)^{k-1} = (-1)^{k+1}$ . Das Integral links ist

$$\int_0^x \frac{1}{1+t} dt = [\ln(1+t)]_0^x = \ln(1+x)$$

und somit folgt

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \dots \quad \text{für } |x| < 1. \quad (4.15)$$

Dies ist die Potenzreihendarstellung des Logarithmus. Nach Satz 4.6.2 ist diese Reihe auch die Taylorreihe der Funktion  $f(x) = \ln(1+x)$ ,

$$T_{\ln(1+x)}(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k,$$

und (4.15) sagt somit, dass die Taylorentwicklung  $\ln(1+x) = T_{\ln(1+x)}(x)$  für alle  $|x| < 1$  gilt.

### Achtung!

Obwohl  $\ln(1+x)$  für  $x > 1$  definiert ist, gilt (4.15) und die Taylorentwicklung  $\ln(1+x) = T_{\ln(1+x)}(x)$  nicht für  $x > 1$ , denn die Reihe in (4.15) hat den Konvergenzradius  $r = 1$  und ist deshalb für  $x > 1$  divergent! ( $x$  liegt außerhalb des Konvergenzradius.)

**Bemerkung:** Das  $r = 1$  tatsächlich der Konvergenzradius der Reihe in (4.15) ist, kann man entweder direkt mit einer der Formeln (4.10) für den Konvergenzradius nachrechnen, oder man benutzt, dass die Ableitung von (4.15) wieder (4.14) ist. Diese beide Reihen haben also nach Satz 4.6.1 denselben Konvergenzradius, und für die Reihe in (4.14) ist  $r = 1$ , da es die geometrische Reihe mit  $q = -x$  ist, die genau für  $|q| = |x| < 1$  konvergiert.

Die Gleichung (4.15) gilt also für  $|x| < 1$ , aber nicht für  $x > 1$ . Für  $x \leq -1$  macht sie keinen Sinn, denn dann ist  $\ln(1+x)$  nicht definiert. Es bleibt die Frage: Was ist bei  $x = 1$ , am rechten Rand des Konvergenzbereichs?

Die Potenzreihe in (4.15) ist für  $x = 1$  gleich

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \dots,$$

also die alternierende harmonische Reihe; sie ist konvergent nach dem Leibniz-Kriterium. Beide Seiten in (4.15) sind also für  $x = 1$  definiert. Sind Sie auch gleich, d.h. gilt  $\ln(2) = T_{\ln(1+x)}(1)$ ? Nach Satz 4.4.3 gilt das genau dann wenn

$$\lim_{n \rightarrow \infty} R_n(1) = 0.$$

Wir betrachten daher jetzt das Restglied

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

für  $f(x) = \ln(1+x)$ . Die Ableitungen sind

$$f'(x) = \frac{1}{1+x}, \quad f''(x) = \frac{-1}{(1+x)^2}, \quad f'''(x) = \frac{2}{(1+x)^3}, \dots$$

bis

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

Damit erhalten wir für das Restglied

$$|R_n(1)| = \left| \frac{(-1)^n n!}{(n+1)!(1+\xi)^{n+1}} \cdot 1^{n+1} \right| = \frac{1}{(n+1)(1+\xi)^{n+1}}$$

Da  $\xi$  zwischen  $x_0 = 0$  und  $x = 1$  liegt, also  $\xi \in [0, 1]$ , ist  $1 + \xi \geq 1$ . Daraus folgt

$$\frac{1}{(1+\xi)^{n+1}} \leq 1$$

und somit

$$|R_n(1)| = \frac{1}{(n+1)(1+\xi)^{n+1}} \leq \frac{1}{n+1} \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Nach Satz 4.4.3 gilt also die Taylorentwicklung des Logarithmus (4.15) auch für  $x = 1$ , d.h.

$$\ln(2) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \dots$$

Anders gesagt haben wir gerade den Wert der alternierenden harmonischen Reihe  $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$  berechnet!

## 4.8 Grenzwertberechnung mit Potenzreihen

Bei Grenzwerten, die auf den Ausdruck " $\frac{0}{0}$ " führen, kann man die normalen Grenzwertrechenregeln, etwa Satz 2.5.1, nicht verwenden. In solchen Fällen kann man oft die Regel von l'Hospital (Mathematik A) benutzen. Eine andere Möglichkeit, die manchmal schneller ist, ist die Verwendung von Potenzreihenentwicklungen. Wir machen das hier an einem einfachen Beispiel:

**Beispiel 4.8.1** Der Grenzwert

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

führt auf " $\frac{0}{0}$ ". Mit der Regel von l'Hospital berechnet man den Grenzwert so:

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

Alternativ können wir aber auch die Taylorentwicklung von  $\sin x$  einsetzen und dann kürzen:

$$\begin{aligned}\lim_{x \rightarrow 0} \frac{\sin x}{x} &= \lim_{x \rightarrow 0} \frac{x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \dots}{x} \\ &= \lim_{x \rightarrow 0} \frac{x \cdot \left(1 - \frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \dots\right)}{x} \\ &= \lim_{x \rightarrow 0} \left(1 - \frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \dots\right) \\ &= 1 - 0 + 0 - \dots = 1\end{aligned}$$

Im letzten Schritt haben wir übrigens die Stetigkeit von Potenzreihen benutzt, da wir an dieser Stelle genau genommen zwei Grenzwerte vertauschen müssen (nämlich  $x \rightarrow 0$  und  $n \rightarrow \infty$ ; letzteres ergibt sich aus der Reihe  $1 - \frac{1}{3!}x^2 + \dots$  als Grenzwert von Teilsommen). Das Vertauschen zweier Grenzwerte ist nur in bestimmten Fällen erlaubt.

**Hinweis:** Im letzten Beispiel ist die Berechnung des Grenzwerts mit l'Hospital schneller. In komplizierteren Fällen jedoch, in denen man l'Hospital mehrfach anwenden müsste, kann die Methode der Taylorentwicklung tatsächlich einfacher und schneller sein!

## 4.9 Weitere Taylorentwicklungen

Hier sind Taylorentwicklungen einiger weiterer Funktionen. Jeweils mit angegeben ist der Bereich der  $x$ , für die die Entwicklung gilt.

$$\cosh x = \frac{1}{2}(e^x + e^{-x}) = \sum_{k=0}^{\infty} \frac{1}{(2k)!} x^{2k} \quad , \quad x \in \mathbb{R}$$

$$\sinh x = \frac{1}{2}(e^x - e^{-x}) = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} x^{2k+1} \quad , \quad x \in \mathbb{R}$$

$$\arctan x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} \quad , \quad |x| < 1$$

$$\arcsin x = x + \frac{1}{2 \cdot 3} x^3 + \frac{1 \cdot 3}{2 \cdot 4 \cdot 5} x^5 + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 7} x^7 + \dots \quad , \quad |x| < 1$$

$$\arccos x = \frac{\pi}{2} - \arcsin x = \frac{\pi}{2} - x - \frac{1}{2 \cdot 3} x^3 - \dots \quad , \quad |x| < 1$$

## 4.10 Eine Funktion ohne Taylorentwicklung

Für jede unendlich oft differenzierbare Funktion  $f(x)$  kann man die Taylorreihe  $T_f(x)$  berechnen. Ob und für welche  $x$  diese Taylorreihe aber die Funktion darstellt, d.h. wann  $f(x) = T_f(x)$  gilt, ist abhängig von der konkreten Funktion. Zum Beispiel wird  $e^x$  für alle  $x \in \mathbb{R}$  durch die Taylorreihe dargestellt (Satz 4.4.4). Für die Funktion  $f(x) = \frac{1}{1-x}$  gilt dagegen die (Taylor-) Entwicklung

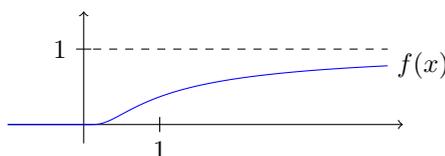
$$f(x) = \frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots \quad \text{nur für } |x| < 1,$$

für alle anderen  $x$  divergiert die Reihe. (geometrische Reihe!) Hier wird also  $f(x)$  für  $|x| > 1$  nicht durch die Taylorreihe dargestellt – obwohl  $f(x)$  dort definiert ist. Zumindest aber haben wir die Darstellung auf einer *Umgebung* des Entwicklungspunktes, nämlich auf  $] -1, 1[$  ( $|x| < 1$ , Entwicklungspunkt  $x_0 = 0$ ).

Es gibt allerdings Funktionen, die auf *keiner* Umgebung von  $x_0$  durch ihre Taylorreihe dargestellt werden. Ein klassisches Beispiel hierfür ist die Funktion

$$f(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0, \\ 0, & x \leq 0. \end{cases} \quad (4.16)$$

Plotet man den Graphen dieser Funktion, so erhält man:



Man beachte, dass  $f(x) > 0$  für  $x > 0$  (denn dann  $f(x) = e^{-1/x}$  und die Exponentialfunktion ist stets positiv und nie Null). Gleichzeitig aber ist  $f(x)$  sehr klein wenn  $x$  nah bei 0. Z.B. ist  $f(\frac{1}{10}) = e^{-10} \approx 4.5 \cdot 10^{-5}$ . Im Graphen erscheint  $f$  in einer kleinen Umgebung von  $x = 0$  (auch nach rechts) fast konstant.

Wir werden jetzt die Taylorreihe dieser Funktion im Entwicklungspunkt  $x_0 = 0$  bestimmen. Dazu berechnen wir alle Ableitungen  $f^{(k)}(0)$ . Die Ergebnisse werden uns auch zeigen, dass der Graph von  $f$  in der Nähe von Null tatsächlich so aussieht wie angegeben. Da  $f$  abschnittsweise für  $x > 0$  und  $x \leq 0$  definiert ist, müssen wir einseitige Ableitungen und Grenzwerte in 0 betrachten. Zuerst bestimmen wir den links- und den rechtsseitigen Grenzwert von  $f(x)$ :

$$\begin{aligned} \lim_{\substack{x \rightarrow 0 \\ x > 0}} f(x) &= \lim_{\substack{x \rightarrow 0 \\ x > 0}} e^{-\frac{1}{x}} = 0 && \text{(denn } -\frac{1}{x} \rightarrow -\infty \text{ bei } x \rightarrow 0, x > 0) \\ \lim_{\substack{x \rightarrow 0 \\ x < 0}} f(x) &= \lim_{x \rightarrow 0} 0 = 0 \end{aligned}$$

Da die einseitigen Grenzwerte also gleich sind und mit dem Funktionswert  $f(0) = 0$  übereinstimmen, folgt, dass  $f$  in  $x = 0$  stetig ist, also keine Sprungstelle hat.

Jetzt die erste Ableitung. Zur Berechnung der einseitigen Ableitungen in 0 bestimmen wir zuerst  $f'(x)$  für  $x > 0$  und für  $x < 0$ :

$$\begin{aligned} x > 0: \quad f'(x) &= e^{-\frac{1}{x}} \cdot \frac{1}{x^2} && \text{(Kettenregel!)} \\ x < 0: \quad f'(x) &= 0 \end{aligned}$$

Für die rechtsseitige Ableitung in 0 erhalten wir dann mit der Substitution  $y = \frac{1}{x}$

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} f'(x) = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{e^{-\frac{1}{x}}}{x^2} = \lim_{y \rightarrow \infty} y^2 e^{-y} = \lim_{y \rightarrow \infty} \frac{y^2}{e^y} \stackrel{\text{„}\infty\text{“}}{=} \lim_{y \rightarrow \infty} \frac{2y}{e^y} \stackrel{\text{„}\infty\text{“}}{=} \lim_{y \rightarrow \infty} \frac{2}{e^y} = 0.$$

Dabei haben wir am Schluss zweimal die Regel von l'Hospital angewandt. Die linksseitige Ableitung in 0 ist

$$\lim_{\substack{x \rightarrow 0 \\ x < 0}} f'(x) = 0$$

weil  $f'(x) = 0$  für  $x < 0$ . Da links- und rechtsseitige Ableitung übereinstimmen, ist  $f$  in  $x = 0$  differenzierbar mit Ableitung

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

Geometrisch bedeutet das eine waagerechte Tangente in  $x = 0$ , was mit dem Graphen übereinstimmt. Außerdem ist  $f'(x)$  auch stetig in 0, denn  $f'(0) = 0$  ist der Grenzwert von  $f'(x)$  bei  $x \rightarrow 0$ .

Genauso berechnen wir jetzt die zweite Ableitung in 0:

$$x > 0: \quad f''(x) = e^{-\frac{1}{x}} \cdot \left( \frac{1}{x^4} - \frac{2}{x^3} \right) \quad (\text{Ketten- und Produktregel})$$

$$x < 0: \quad f''(x) = 0$$

Daraus folgt (wieder mit  $y = \frac{1}{x}$ )

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} e^{-\frac{1}{x}} \left( \frac{1}{x^4} - \frac{2}{x^3} \right) = \lim_{y \rightarrow \infty} e^{-y} (y^4 - 2y^3) = \lim_{y \rightarrow \infty} \frac{y^4 - 2y^3}{e^y} \stackrel{\text{„}\infty\text{“}}{=} \dots = 0.$$

Hier ist am Schluss mehrfach l'Hospital angewandt worden. Da die linksseitige zweite Ableitung auch 0 ist, ist  $f$  in  $x = 0$  zweimal differenzierbar mit

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

Geometrisch heißt das: in  $x = 0$  ist die Krümmung von  $f$  Null.

Man kann auf diese Art auch die folgenden Ableitungen  $f'''$ ,  $f^{(4)}$ , u.s.w. ausrechnen<sup>1</sup>, und erhält schließlich als Ergebnis:

$f$  ist unendlich oft differenzierbar mit  $f^{(n)}(0) = 0$  für alle  $n \in \mathbb{N}$ .

Daraus ergibt sich nun für die Taylorreihe in  $x_0 = 0$ :

$$T_f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=0}^{\infty} 0 \cdot x^k = 0.$$

Die Taylorreihe  $T_f(x)$  ist also konstant gleich Null für alle  $x \in \mathbb{R}$ . Es folgt

$$f(x) \neq T_f(x) \quad \text{für jedes } x > 0.$$

D.h.  $f$  wird in keiner Umgebung von 0 durch die Taylorreihe dargestellt!

<sup>1</sup>Das würde man mit vollständiger Induktion machen.

# Kapitel 5

## Elementare Differentialgleichungen

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik B, Kapitel 5.

### 5.1 Einführung

Eine *Differentialgleichung* ist eine Gleichung für eine gesuchte Funktion  $y = y(x)$ , die neben den Variablen  $x$  und  $y$  auch noch Ableitungen von  $y(x)$  enthält. Die *Ordnung* einer Differentialgleichung ist die Ordnung der höchsten vorkommenden Ableitung. Abkürzend schreibt man für eine Differentialgleichung oft „Dgl“.

Bei Differentialgleichungen lässt man die Angabe der Stelle  $x$  bei der Funktion  $y(x)$  und ihren Ableitungen oft weg, d.h. man schreibt einfach  $y, y'$ . Man nennt

- $x$  die *unabhängige Variable* und
- $y$  die *abhängige Variable*

der Differentialgleichung. (Denn  $y = y(x)$  ist von  $x$  abhängig.)

**Beispiel 5.1.1** (a) Die Gleichung

$$y'(x) = -2x \cdot y(x)$$

ist eine Differentialgleichung 1. Ordnung, denn sie enthält als höchste Ableitung von  $y(x)$  die erste Ableitung  $y'(x)$ . In der Kurzschreibweise:

$$y' = -2xy.$$

(b)

$$y''(x) = x^3 \cdot y(x) + y'(x) - \sin(x)$$

oder kurz

$$y'' = x^3 y + y' - \sin(x)$$

ist eine Differentialgleichung 2. Ordnung, denn  $y''$  ist die höchste Ableitung.

## 5.2 Allgemeine und spezielle Lösungen, Anfangswerte

Wir besprechen zunächst grundsätzliche Lösungsbegriffe für Differentialgleichungen. Der Frage, wie man eine Lösung ausrechnen kann und welche Lösungsmethoden es gibt, wenden wir uns in den folgenden Abschnitten zu.

- Eine Funktion  $y(x)$  ist *Lösung* der Differentialgleichung, wenn sie mit ihren Ableitungen die Differentialgleichung erfüllt. Damit kann man für jede gegebene Funktion durch Einsetzen in die Differentialgleichung überprüfen, ob sie eine Lösung ist oder nicht.
- Eine Lösung, die noch freie Parameter (meist Integrationskonstanten) enthält, heißt *allgemeine Lösung* der Differentialgleichung. Bei einer Dgl.  $n$ . Ordnung sind es in der Regel  $n$  Parameter.
- Eine Lösung ohne freie Parameter heißt *spezielle Lösung*.
- Aus einer allgemeinen Lösung erhält man durch Wahl der Parameter (viele) neue spezielle Lösungen.

Im Gegensatz zu einer „normalen“ Gleichung ohne Ableitungen (auch algebraische Gleichung genannt), die typischerweise eine oder endliche viele feste Werte als Lösungen hat, hat eine Differentialgleichung in der Regel unendlich viele Funktionen als Lösung (die man durch Wahl der Parameter in der allgemeinen Lösung erhalten kann).

**Beispiel 5.2.1** Wir betrachten die Differentialgleichung

$$y' = -2xy. \quad (5.1)$$

- (a) Die Funktion  $y(x) = e^{-x^2}$  ist eine spezielle Lösung von (5.1), was wir leicht durch Einsetzen überprüfen können: Nach der Kettenregel ist die Ableitung von  $y(x)$

$$y'(x) = e^{-x^2} \cdot (-2x)$$

und somit

$$y'(x) = -2x \cdot \underbrace{e^{-x^2}}_{y(x)} = -2xy(x).$$

D.h.  $y(x) = e^{-x^2}$  erfüllt die Differentialgleichung  $y' = -2xy$ , ist also eine Lösung von (5.1).

- (b) Für jede Konstante  $c \in \mathbb{R}$  ist

$$y(x) = c \cdot e^{-x^2} \quad (5.2)$$

auch eine Lösung von (5.1). Das überprüfen wir wieder durch Einsetzen:

$$y'(x) = ce^{-x^2} \cdot (-2x) = -2x \cdot \underbrace{ce^{-x^2}}_{y(x)} = -2xy(x).$$

Die Funktion  $y(x) = ce^{-x^2}$  erfüllt also auch die Dgl.  $y' = -2xy$ . Damit ist (5.2) eine allgemeine Lösung von (5.1), denn sie enthält noch den freien Parameter  $c$ .

(c) Aus der allgemeinen Lösung (5.2) können wir durch (beliebige) Wahl des Parameters  $c$  weitere spezielle Lösungen erhalten:

$$\begin{aligned} c = 1 &\Rightarrow y(x) = e^{-x^2} && \text{(das ist wieder die Lsg aus (a))} \\ c = -17 &\Rightarrow y(x) = -17e^{-x^2} \\ c = 0 &\Rightarrow y(x) = 0 \end{aligned}$$

**Beachte:** Dass man in Beispiel 5.2.1 durch Multiplikation der speziellen Lösung  $e^{-x^2}$  mit der Konstante  $c$  wieder eine Lösung erhält, liegt an der speziellen Struktur der Differentialgleichung (5.1). (Es ist eine lineare homogene Dgl, siehe 5.6.) Im Allgemeinen erhält man durch Multiplikation einer Lösung mit einer Konstante *nicht* wieder eine Lösung.

Neben der Suche einer allgemeinen Lösung betrachtet man zu einer gegebenen Differentialgleichung  $n$ . Ordnung oft auch eine sogenanntes *Anfangswertproblem* (abgekürzt AWP):

Suche eine spezielle Lösung der Differentialgleichung mit vorgegebenen *Anfangswerten* bei  $x = x_0$ :

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \dots, \quad y^{(n-1)}(x_0) = y_{n-1}.$$

Es werden also die Werte in  $x_0$  für  $y(x)$  und alle Ableitungen  $y^{(k)}(x)$  bis  $k = n-1$  vorgegeben.

**Beachte:** Die Anzahl Anfangswerte ist immer gleich der Ordnung  $n$  der Differentialgleichung.

**Beispiel 5.2.2** Wir betrachten folgendes Anfangswertproblem:

$$y' = -2xy, \quad y(1) = 2. \tag{5.3}$$

Es ist also die Dgl (5.1) zusammen mit dem Anfangswert  $y(1) = 2$ . Aus Beispiel 5.2.1 kennen wir schon die allgemeine Lösung:

$$y(x) = ce^{-x^2}.$$

Durch Einsetzen des Anfangswerts können wir daraus die Lösung des Anfangswertproblems berechnen:

$$y(1) = ce^{-1} \stackrel{!}{=} 2 \quad \Leftrightarrow \quad \frac{c}{e} = 2 \quad \Leftrightarrow \quad c = 2e$$

Die allgemeine Lösung erfüllt also die Anfangsbedingung  $y(1) = 2$  genau für  $c = 2e$ . Somit ist

$$y(x) = 2e \cdot e^{-x^2}$$

die Lösung des Anfangswertproblems (5.3).

### 5.3 Differentialgleichungen in Anwendungen

Differentialgleichungen werden sehr häufig benutzt, um Phänomene der Natur mit einer dynamischen Zeitentwicklung zu beschreiben, z.B. in der Physik, Chemie oder Biologie. Dynamisch bedeutet hier, dass die zeitliche Änderung in gesetzmäßiger Weise vom aktuellen Zustand abhängt.

Die Zeit  $t$  ist in diesem Fall dann die unabhängige Variable, die abhängige Variable  $y = y(t)$  beschreibt den Zustand zur Zeit  $t$ . Im einfachsten Fall, den wir hier betrachten, ist  $y \in \mathbb{R}$ , d.h. der betrachtete Zustand wird einfach durch eine Zahl beschrieben. Weil die unabhängige Variable die Zeit  $t$  ist, sind die Ableitungen von  $y(t)$ , die in der Differentialgleichung vorkommen, Zeitableitungen. Für Zeitableitungen benutzt man häufig eine alternative Schreibweise mit einem Ableitungspunkt anstelle des Strichs:

$$\dot{y}(t) = y'(t), \quad \ddot{y}(t) = y''(t), \quad \dots$$

**Beispiel 5.3.1** Ein einfaches Beispiel für die dynamische Änderung eines Zustands ist *exponentielles Wachstum*. Das kann z.B. das exponentielle Wachstum einer Population von Individuen sein, etwa die Vermehrung von Kaninchen oder Zellteilung. Der Zustand  $y(t)$  ist dann die Anzahl von Individuen (Kaninchen, Zellen) zur Zeit  $t$ . Die Ableitung  $\dot{y}(t)$  ist die zeitliche Änderung der Anzahl, das momentane Wachstum pro Zeiteinheit der Population zur Zeit  $t$ . Je nach Art des Wachstums sind verschiedene Zusammenhänge zwischen  $\dot{y}(t)$  und  $y(t)$  möglich. Bei exponentiellem Wachstum ist  $\dot{y}(t)$  proportional zu  $y(t)$ , d.h.

$$\dot{y}(t) = a \cdot y(t). \tag{5.4}$$

Doppelt so viele Individuen (Kaninchen, Zellen) ergeben in einer festen Zeit doppelt so viele Nachkommen (neue Kaninchen, Zellen). Die Proportionalitätskonstante  $a$  ist die Wachstumsrate. (Für Wachstum ist  $a > 0$ .) Gleichung (5.4) ist die Differentialgleichung des exponentiellen Wachstums, in unserer Kurzschreibweise einfach

$$\dot{y} = ay. \tag{5.5}$$

Die allgemeine Lösung dieser Dgl ist

$$y(t) = c \cdot e^{at} \tag{5.6}$$

mit dem freien Parameter  $c \in \mathbb{R}$ . Dass dies tatsächlich die Lösung ist, kann man wieder einfach durch Ableiten nachprüfen:

$$\dot{y}(t) = ce^{at} \cdot a = a \cdot y(t)$$

Wir können zum Abschluss noch einen Anfangswert festlegen. Zur Anfangszeit  $t = 0$  bestehe die Population aus  $y_0$  Individuen, also

$$y(0) = y_0. \tag{5.7}$$

Der Zustand  $y(t)$  wird damit zum Anfang  $t = 0$  des betrachteten Zeitraums festgelegt. Danach bestimmt die Dgl die weitere Entwicklung des Zustands. Aus der Anfangsbedingung können wir wieder den freien Parameter bestimmen:

$$y(0) = ce^{a \cdot 0} = c \stackrel{!}{=} y_0.$$

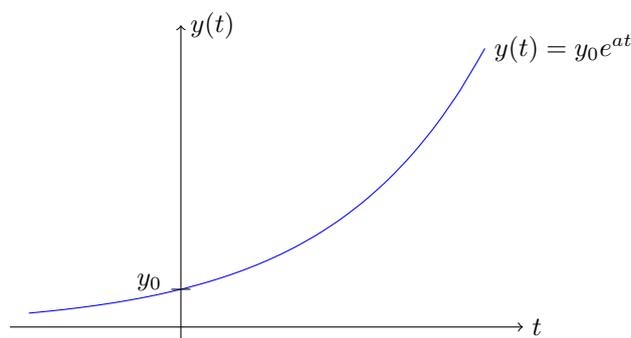


Abbildung 5.1: Exponentielles Wachstum

Damit ist

$$y(t) = y_0 e^{at}$$

die Lösung des Anfangswertproblems aus (5.5) und (5.7). Die Größe der Population wächst also ausgehend vom Anfangswert  $y_0$  exponentiell (vgl. Abbildung 5.1), daher der Name exponentielles Wachstum.

**Bemerkung:** Die Differentialgleichung (5.5) des exponentiellen Wachstums beschreibt auch andere Phänomene, bei der die Änderung proportional zum Zustand ist, z.B. den exponentiellen Zerfall einer radioaktiven Substanz. Hierbei ist dann  $a < 0$  (Zerfall!).

## 5.4 Differentialgleichungen 1. Ordnung mit getrennten Variablen

Wir betrachten jetzt verschiedene Lösungsmethoden für Differentialgleichungen 1. und 2. Ordnung. Jede Lösungsmethode ist dabei auf einen bestimmten Typ von Differentialgleichungen anwendbar. Um zu wissen, welche Methode anwendbar ist, muss man also zuerst den Typ der Differentialgleichung erkennen.

Wir beginnen mit der *Methode des Trennens der Variablen*. Sie ist anwendbar auf Differentialgleichungen der Form

$$y' = f(x) \cdot g(y). \quad (5.8)$$

In dieser Dgl sind die Variablen  $x$ ,  $y$  auf der rechten Seite als Produkt von zwei Faktoren getrennt: der  $f$ -Teil enthält nur die Variable  $x$ , der  $g$ -Teil nur  $y$ . Die Dgl (5.8) ist wie folgt lösbar durch *Trennen der Variablen*: Man schreibt die Ableitung  $y'$  in der Leibniz-Schreibweise

$$y' = y'(x) = \frac{dy}{dx},$$

bringt alle Faktoren in (5.8) mit  $y$  auf die eine, alle Faktoren mit  $x$  auf die

andere Seite, integriert dann und löst nach  $y$  auf:

$$\begin{aligned} \frac{dy}{dx} &= y' = f(x) \cdot g(y) \\ \Leftrightarrow \frac{dy}{g(y)} &= f(x) dx && (\text{für } g(y) \neq 0) \\ \Rightarrow \int \frac{dy}{g(y)} &= \int f(x) dx \\ \Leftrightarrow H(y) &= F(x) + c && (c \in \mathbb{R}) \\ \Leftrightarrow y &= H^{-1}(F(x) + c) \end{aligned}$$

Die Funktionen  $H(y)$  und  $F(x)$  sind die Stammfunktionen von  $1/g(y)$  und  $f(x)$ , die Konstante  $c$  ist ein freier Parameter. Die berechnete Lösung ist damit die *allgemeine Lösung* der Differentialgleichung (5.8).

### Achtung!

Für  $g(y) = 0$  erhält man eine oder mehrere weitere spezielle Lösungen, die in der allgemeinen Lösung nicht enthalten sind: Ist  $g(y_0) = 0$ , dann ist die konstante Funktion

$$y(x) = y_0$$

auch eine Lösung der Differentialgleichung (5.8), denn dann ist

$$y'(x) = 0 = f(x) \cdot \underbrace{g(y_0)}_0 = f(x)g(y(x)).$$

In der Herleitung der allgemeinen Lösung fehlt diese spezielle Lösung, da beim Teilen durch  $g(y)$  implizit  $g(y) \neq 0$  vorausgesetzt wird.

**Beispiel 5.4.1** (a) Wir suchen die allgemeine Lösung von

$$y' = x \cdot (y - 1)^2. \quad (5.9)$$

Die Differentialgleichung ist von der Form (5.8), da  $x$  und  $y$  als Produkt  $f(x) \cdot g(y)$  getrennt sind ( $f(x) = x$  und  $g(y) = (y - 1)^2$ ). Trennen der Variablen liefert

$$\begin{aligned} \frac{dy}{dx} &= x \cdot (y - 1)^2 \\ \Leftrightarrow \frac{dy}{(y - 1)^2} &= x dx && \text{falls } (y - 1)^2 \neq 0 \\ \Leftrightarrow \int \frac{dy}{(y - 1)^2} &= \int x dx \end{aligned}$$

Die Integrale sind

$$\int \frac{dy}{(y - 1)^2} = \int (y - 1)^{-2} dy = -(y - 1)^{-1} + c_1 = -\frac{1}{y - 1} + c_1$$

und

$$\int x dx = \frac{1}{2}x^2 + c_2.$$

Somit erhalten wir für die allgemeine Lösung

$$\begin{aligned} & \int \frac{dy}{(y-1)^2} = \int x \, dx \\ \Leftrightarrow & \quad -\frac{1}{y-1} = \frac{1}{2}x^2 + c \quad , \quad c \in \mathbb{R} \\ \Leftrightarrow & \quad y-1 = \frac{-1}{\frac{1}{2}x^2 + c} \\ \Leftrightarrow & \quad y = 1 - \frac{1}{\frac{1}{2}x^2 + c} \end{aligned}$$

(Wir haben hier die Integrationskonstanten  $c_1$  und  $c_2$  der beiden Integrale zu  $c = c_2 - c_1$  zusammengefasst.)

Beim Trennen der Variablen haben wir durch  $(y-1)^2$  geteilt und dabei  $(y-1)^2 \neq 0$  vorausgesetzt. Wir erhalten damit eine weitere Lösung für  $(y-1)^2 = 0$ , d.h. wenn  $y = 1$  konstant.

Zusammengefasst haben wir damit für (5.9) die allgemeine Lösung

$$y(x) = 1 - \frac{1}{\frac{1}{2}x^2 + c} \quad \text{mit Parameter } c \in \mathbb{R} \quad (5.10)$$

sowie die zusätzliche spezielle Lösung

$$y(x) = 1 \quad \text{konstant.} \quad (5.11)$$

Man beachte, dass die zusätzliche spezielle Lösung *nicht* in der allgemeinen Lösung enthalten ist.

### Achtung!

Ist ein Anfangswertproblem zu lösen, muss die allgemeine *und* die zusätzliche spezielle Lösung betrachtet werden:

Für den Anfangswert  $y(0) = 2$  zum Beispiel ist  $y(0) \neq 1$ , daher ergibt sich hier die Lösung zum Anfangswert aus der allgemeinen Lösung (5.10) (denn für die allg. Lsg ist  $y(x) \neq 1$ ). Einsetzen der Anfangsbedingung in (5.10) liefert

$$y(0) = 1 - \frac{1}{\frac{1}{2} \cdot 0^2 + c} = 1 - \frac{1}{c} \stackrel{!}{=} 2$$

also  $c = -1$ . Damit ist

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

die Lösung zum Anfangswert  $y(0) = 2$ .

Für den Anfangswert  $y(0) = 1$  dagegen ist die spezielle Lösung (5.11) die Lösung des Anfangswertproblems, denn  $y(x) = 1$  erfüllt natürlich  $y(0) = 1$ .

(b) Wir betrachten die Differentialgleichung

$$y' = ay$$

für exponentielles Wachstum und leiten jetzt die Formel der allgemeinen Lösung her, die wir in Beispiel 5.3.1 schon kennengelernt (und durch einsetzen verifiziert!) hatten. Auch hier ist „Trennen der Variablen“ anwendbar ( $f(x) = a$ ,  $g(y) = y$ ):

$$\begin{aligned} \frac{dy}{dx} &= a \cdot y \quad \Leftrightarrow \quad \frac{dy}{y} = a \, dx \quad \text{falls } y \neq 0 \\ \Rightarrow \int \frac{dy}{y} &= \int a \, dx \quad \Leftrightarrow \quad \ln |y| = ax + c \quad , \quad c \in \mathbb{R} \\ \Leftrightarrow |y| &= e^{ax+c} \\ \Leftrightarrow y &= \pm e^{ax+c} = \pm e^c e^{ax} = c_1 e^{ax} \end{aligned}$$

Im letzten Schritt haben wir die neue Konstante  $c_1 = \pm e^c$  eingeführt. Welche möglichen Werte kann  $c_1$  annehmen? Da  $c \in \mathbb{R}$  beliebig ist, ist  $e^c \in ]0, \infty[$  beliebig. Damit ist dann  $c_1 = \pm e^c \in \mathbb{R} \setminus \{0\}$  beliebig. Wir haben also zunächst die allgemeine Lösung

$$y(x) = c_1 e^{ax}, \quad c_1 \in \mathbb{R} \setminus \{0\}. \quad (5.12)$$

Beim Trennen der Variablen haben wir (beim Teilen durch  $y$ )  $y \neq 0$  vorausgesetzt. Wir bekommen somit eine weitere spezielle Lösung für  $y = 0$ :

$$y(x) = 0 \quad \text{konstant.}$$

Hier können wir (im Gegensatz zum Beispiel (a)) die spezielle Lösung mit der Formel (5.12) der allgemeinen Lösung zu einer gemeinsamen Lösungsformel kombinieren, indem wir in (5.12) für  $c_1$  auch den Wert 0 zulassen. Wir haben damit schließlich die allgemeine Lösung

$$y(x) = c_0 e^{ax}, \quad c_0 \in \mathbb{R}.$$

(Sie enthält (5.12) für  $c_0 = c_1 \in \mathbb{R} \setminus \{0\}$  und  $y(x) = 0$  für  $c_0 = 0$ .)

(c) Auch

$$y' = -\frac{x}{y} \quad (y \neq 0)$$

kann man mit Trennen der Variablen lösen:

$$\begin{aligned} \frac{dy}{dx} &= -\frac{x}{y} \quad \Leftrightarrow \quad y \, dy = -x \, dx \\ \Rightarrow \int y \, dy &= \int -x \, dx \\ \Leftrightarrow \frac{1}{2} y^2 &= -\frac{1}{2} x^2 + c \quad , \quad c \in \mathbb{R} \\ \Leftrightarrow y^2 &= -x^2 + 2c \\ \Leftrightarrow y &= \pm \sqrt{-x^2 + 2c} \end{aligned}$$

Wir haben hier also eigentlich zwei allgemeine Lösungen:

$$y(x) = \sqrt{-x^2 + 2c}, \quad c \in \mathbb{R},$$

mit Werten  $y \geq 0$ , und

$$y(x) = -\sqrt{-x^2 + 2c}, \quad c \in \mathbb{R},$$

mit Werten  $y \leq 0$ .

Betrachten wir jetzt das Anfangswertproblem

$$y' = -\frac{x}{y}, \quad y(1) = -1.$$

Da für den Anfangswert  $y(1) = -1 < 0$  gilt, kann die Anfangsbedingung nur durch die Lösung mit  $y \leq 0$ , also  $y(x) = -\sqrt{-x^2 + 2c}$  erfüllt werden. Einsetzen der Anfangsbedingung liefert

$$y(1) = -\sqrt{-1 + 2c} \stackrel{!}{=} -1 \quad \Rightarrow \quad -1 + 2c = 1 \quad \Rightarrow \quad c = 1.$$

Also ist  $y(x) = -\sqrt{-x^2 + 2}$  die Lösung des Anfangswertproblems.

## 5.5 Direkte Lösung des Anfangswertproblems bei getrennten Variablen

Wir betrachten das Anfangswertproblem für eine Differentialgleichung 1. Ordnung mit getrennten Variablen

$$y' = f(x)g(y) \quad y(x_0) = y_0,$$

wobei der Anfangswert die Bedingung  $g(y_0) \neq 0$  erfüllt. In diesem Fall kann man die Lösungsmethode des letzten Abschnitts so modifizieren, dass direkt die Lösung des Anfangswertproblems berechnet wird. Hierzu statt dem unbestimmten Integral das bestimmte Integral von  $x_0$  bis  $x$  gebildet:

$$\begin{aligned} y' = f(x)g(y) &\Leftrightarrow \frac{y'(x)}{g(y(x))} = f(x) \\ \Rightarrow \int_{x_0}^x \frac{y'(t)}{g(y(t))} dt &= \int_{x_0}^x f(t) dt \end{aligned}$$

Im linken Integral benutzen wir die Substitution  $u = y(t)$ . Dafür ist

$$\frac{du}{dt} = y'(t) \quad \Rightarrow \quad du = y'(t)dt$$

und Substitution der Grenzen ergibt

$$\begin{aligned} t = x &\Rightarrow u = y(x) = y, \\ t = x_0 &\Rightarrow u = y(x_0) = y_0. \end{aligned}$$

Hier haben wir im letzten Schritt die Anfangsbedingung verwendet. Einsetzen der Substitution ergibt nun

$$\int_{y_0}^y \frac{1}{g(u)} du = \int_{x_0}^x f(t) dt. \quad (5.13)$$

Aus dieser Gleichung erhält man durch Ausrechnen der Integrale und Auflösen nach  $y$  die Lösung des Anfangswertproblems.

**Beispiel 5.5.1** Wir lösen das Anfangswertproblem

$$y' = -3y, \quad y(0) = 10.$$

Trennen der Variablen liefert

$$\frac{dy}{dx} = -3y \Leftrightarrow \frac{dy}{y} = -3 dx$$

falls  $y \neq 0$ . Da für den Anfangswert  $y_0 = 10 \neq 0$  gilt, können wir jetzt (5.13) anwenden und damit die Lösung des AWP bestimmen:

$$\begin{aligned} \int_{10}^y \frac{dy}{y} &= \int_0^x -3 dx \Leftrightarrow [\ln y]_{10}^y = [-3x]_0^x \\ \Leftrightarrow \ln y - \ln 10 &= -3x \\ \Leftrightarrow \ln y &= -3x + \ln 10 \\ \Leftrightarrow y &= e^{-3x + \ln 10} = 10e^{-3x} \end{aligned}$$

Die Lösung des Anfangswertproblems ist damit  $y(x) = 10e^{-3x}$ .

Man beachte, dass die Stammfunktion von  $1/y$  im Allgemeinen  $\ln |y|$  ist. Wir konnten hier den Betrag aber weglassen, da für den Anfangswert  $y(0) = 10 > 0$  gilt und wegen der Bedingung  $y \neq 0$  für die Lösung dann  $y = y(x) > 0$  gelten muss.

## 5.6 Lineare Differentialgleichungen 1. Ordnung

Eine Differentialgleichung der Form

$$y' + a(x)y = g(x) \tag{5.14}$$

heißt *lineare* Differentialgleichung 1. Ordnung. Sie ist *homogen* wenn  $g(x) = 0$ , sonst *inhomogen*.

Der Ausdruck „linear“ bezieht sich hier auf die abhängige Variable  $y$ , die in (5.14) nur linear vorkommt<sup>1</sup>:  $y$  kommt nur in der ersten Potenz vor, multipliziert mit einem Faktor, der nur von  $x$  abhängt. Es gibt keine quadratischen oder höheren Potenzen  $y^k$ , keine Quotienten mit  $y$  im Nenner, und auch keine Funktionen, in die  $y$  eingesetzt wurde.

Für die homogene Differentialgleichung, also bei  $g(x) = 0$ , kann man die allgemeine Lösung leicht berechnen:

**Satz 5.6.1** Die allgemeine Lösung von

$$y' + a(x)y = 0 \tag{5.15}$$

ist

$$y(x) = c \cdot e^{-A(x)}, \quad c \in \mathbb{R}, \tag{5.16}$$

wobei  $A(x)$  eine Stammfunktion von  $a(x)$  ist.

<sup>1</sup>wie bei einer linearen Funktion mit  $y$  als Variable:  $my + b$

Äquivalent kann man die Lösung (5.16) auch als

$$y(x) = c \cdot e^{-\int a(x) dx}$$

schreiben, denn das unbestimmte Integral  $\int a(x) dx$  liefert ja eine Stammfunktion von  $a(x)$ . Zu beachten ist dabei, dass für das Integral keine Integrationskonstante nötig ist, da diese schon in der Konstante  $c$  der Lösung enthalten ist. Das ergibt sich aus der Herleitung Lösung im folgenden Beweis des Satzes:

**Beweis.** Durch Umformen erkennt man, dass die homogene Dgl durch Trennen der Variablen lösbar ist:

$$\begin{aligned} y' + a(x)y = 0 &\Leftrightarrow \frac{dy}{dx} = y' = -a(x)y \\ \Leftrightarrow \frac{dy}{y} = -a(x) dx &\quad (y \neq 0) \\ \Rightarrow \int \frac{dy}{y} = - \int a(x) dx & \\ \Leftrightarrow \ln |y| = -A(x) + c_0 &\quad , \quad c_0 \in \mathbb{R} \\ \Leftrightarrow y = \pm e^{-A(x)+c_0} = \pm e^{c_0} e^{-A(x)} = ce^{-A(x)} &\quad , \quad c \in \mathbb{R} \setminus \{0\} \end{aligned}$$

Der Fall  $y = 0$  liefert eine weitere, spezielle Lösung, die in der allgemeinen Lösung für  $c = 0$  enthalten ist. Damit ist die allgemeine Lösung  $y(x) = ce^{-A(x)}$  mit  $c \in \mathbb{R}$ .  $\square$

**Beispiel 5.6.2** (a) Die Differentialgleichung des exponentiellen Wachstums

$$y' = ay$$

ist eine homogene lineare Differentialgleichung 1. Ordnung. Um die Lösungsformel aus Satz 5.6.1 anwenden zu können, müssen wir die Dgl zuerst auf die Form (5.14) bringen:

$$y' = ay \Leftrightarrow y' - ay = 0$$

Der Vergleich mit (5.14) zeigt  $a(x) = -a$ . Die Stammfunktion dazu ist  $A(x) = -ax$ . Damit erhalten wir als allgemeine Lösung (man achte auf die Vorzeichen!)

$$y(x) = ce^{-A(x)} = ce^{ax}.$$

Das ist (natürlich) wieder die Lösung, die wir auch schon in Beispiel 5.4.1 berechnet hatten.

(b) Die Differentialgleichung

$$y' + x^3 y = 0$$

ist linear homogen 1. Ordnung mit  $a(x) = x^3$ . Eine Stammfunktion davon ist  $A(x) = \frac{1}{4}x^4$ . Also ist die allgemeine Lösung

$$y(x) = ce^{-\frac{1}{4}x^4}.$$

**Beachte:** Wichtig bei der Anwendung der Lösungsformel (5.16) sind die korrekten Vorzeichen für  $a(x)$  und  $A(x)$ . In der Lösungsformel steht in der Exponentialfunktion ein Minus vor  $A(x)$ . Und für das richtige Vorzeichen von  $a(x)$  muss die Dgl auf die Form (5.14) umgeformt sein, d.h. der  $y$ -Term muss auf der linken Seite der Gleichung bei  $y'$  stehen.

## 5.7 Inhomogene lineare Differentialgleichungen

### 1. Ordnung

Auch für die inhomogene lineare Differentialgleichung 1. Ordnung gibt es eine direkte Lösungsformel:

**Satz 5.7.1** *Die inhomogene lineare Differentialgleichung*

$$y' + a(x)y = g(x) \quad (5.17)$$

hat die allgemeine Lösung

$$y(x) = e^{-A(x)} \cdot \left( c + \int e^{A(x)} g(x) dx \right), \quad c \in \mathbb{R}, \quad (5.18)$$

Dabei ist  $A(x)$  eine Stammfunktion von  $a(x)$ .

Gleichung (5.18) heißt auch *Variation der Konstanten Formel*. Für das unbestimmte Integral  $\int e^{A(x)} g(x) dx$  ist keine zusätzliche Integrationskonstante nötig, da sie in (5.18) ja schon explizit steht.

**Beweis.** Da die inhomogene Gleichung (5.17) eine ähnliche Form wie die homogene Gleichung (5.15) hat, kann man vermuten, dass auch die Lösung eine ähnliche Form wie die homogene Lösung (5.16) hat. Wir machen daher für die Lösung den *Ansatz*

$$y(x) = c(x) \cdot e^{-A(x)}.$$

Die Konstante  $c$  der homogenen Lösung darf also jetzt von  $x$  abhängen; das ist die „Variation der Konstanten“. Um zu sehen, ob der Ansatz tatsächlich eine Lösung liefert, setzen wir ihn in die Dgl ein (genau wie schon im Beispiel 5.2.1). Mit Produkt- und Kettenregel erhalten wir

$$\begin{aligned} y'(x) &= c'(x) \cdot e^{-A(x)} + c(x) \cdot e^{-A(x)}(-a(x)) \\ &= c'(x)e^{-A(x)} - a(x)y(x) \\ \Leftrightarrow y'(x) + a(x)y(x) &= c'(x)e^{-A(x)} \end{aligned}$$

Also ist

$$\begin{aligned} y'(x) + a(x)y(x) &= g(x) \\ \Leftrightarrow c'(x)e^{-A(x)} &= g(x) \\ \Leftrightarrow c'(x) &= e^{A(x)}g(x) \\ \Rightarrow c(x) &= \int e^{A(x)}g(x) dx + c \end{aligned}$$

Setzt man dies in den Lösungsansatz ein, bekommt man (5.18).  $\square$

**Beispiel 5.7.2** Die Differentialgleichung

$$y' + xy = 3x$$

ist linear, inhomogen mit Ordnung 1. Es ist  $a(x) = x$ , also  $A(x) = \frac{1}{2}x^2$ . Mit der Variation der Konstanten Formel folgt

$$y(x) = e^{-\frac{1}{2}x^2} \cdot \left( c + \int e^{\frac{1}{2}x^2} \cdot 3x \, dx \right)$$

Zur Berechnung des Integrals substituieren wir  $u = \frac{1}{2}x^2$ . Dann ist  $\frac{du}{dx} = x$ , also  $du = x \, dx$  und damit

$$\int e^{\frac{1}{2}x^2} \cdot 3x \, dx = \int 3e^u \, du = 3e^u = 3e^{\frac{1}{2}x^2}.$$

Damit ist die allgemeine Lösung der Dgl

$$y(x) = e^{-\frac{1}{2}x^2} \cdot \left( c + 3e^{\frac{1}{2}x^2} \right) = ce^{-\frac{1}{2}x^2} + 3 \quad \text{mit } c \in \mathbb{R}.$$

Wie schon bei der Methode des Trennens der Variablen in 5.5 gibt es auch bei der Variation der Konstanten eine Version um die Lösung eines Anfangswertproblems direkt, ohne „Umweg“ über die allgemeine Lösung zu berechnen:

**Satz 5.7.3** *Das Anfangswertproblem*

$$y' + a(x)y = g(x) \quad , \quad y(x_0) = y_0$$

hat die Lösung

$$y(x) = e^{-A(x)} \cdot \left( y_0 + \int_{x_0}^x e^{A(t)} g(t) \, dt \right) \quad (5.19)$$

mit

$$A(x) = \int_{x_0}^x a(t) \, dt.$$

**Beweis.** Da  $A(x) = \int_{x_0}^x a(t) \, dt$  eine Stammfunktion von  $a(x)$  ist, ist (5.19) ein Spezialfall von (5.18) und somit eine Lösung der Dgl. Die Anfangsbedingung ist ebenfalls erfüllt, denn es gilt  $A(x_0) = \int_{x_0}^{x_0} a(t) \, dt = 0$  und damit

$$y(x_0) = e^{-A(x_0)} \left( y_0 + \int_{x_0}^{x_0} e^{A(t)} g(t) \, dt \right) = e^0 (y_0 + 0) = y_0.$$

□

## 5.8 Anwendung: RC-Reihenschaltung

Ein einfaches Anwendungsbeispiel einer linearen inhomogenen Differentialgleichung 1. Ordnung ist die RC-Reihenschaltung. Ein Widerstand  $R$  und ein Kondensator  $C$  sind mit einer Spannungsquelle  $U$  in Reihe geschaltet, siehe Abbildung 5.2. Der Zustand der Schaltung zu einem Zeitpunkt  $t$  ist bestimmt durch

- die Kondensatorladung  $Q(t)$ ,
- die Stromstärke  $I(t)$ ,
- die angelegte Spannung  $U(t)$ .

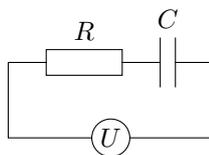


Abbildung 5.2: RC-Reihenschaltung

Für die an den Bauteilen der Schaltung abfallenden Spannungen gilt

$$U_R(t) + U_C(t) = U(t)$$

wobei

$$U_R(t) = RI(t), \quad U_C(t) = \frac{1}{C}Q(t)$$

die Spannungen am Widerstand  $R$  und Kondensator  $C$  sind. Außerdem gilt  $I(t) = \dot{Q}(t)$ , die Stromstärke ist die zeitliche Änderung der Ladung im Kondensator. Setzt man alles zusammen ergibt sich

$$\begin{aligned} RI(t) + \frac{1}{C}Q(t) &= U(t) \\ \Rightarrow \dot{Q}(t) + \frac{1}{RC}Q(t) &= \frac{1}{R}U(t) \end{aligned}$$

Dies ist eine inhomogene lineare Differentialgleichung 1. Ordnung für  $Q(t)$ . In der Schreibweise (5.17) mit  $y$  für die abhängige Variable  $Q(t)$ :

$$y' + \frac{1}{RC}y = \frac{1}{R}U(t),$$

d.h.  $a(t) = \frac{1}{RC}$  und  $g(t) = \frac{1}{R}U(t)$ .

Bei konkret gegebener angelegter äußerer Spannung  $U(t)$  kann die Lösung der Dgl mit der Variation der Konstanten Formel berechnet werden.

**Beispiel 5.8.1** Aufladen des Kondensators bei konstanter äußerer Spannung.

Wir legen eine konstante äußere Spannung  $U(t) = U_a$  an und beginnen zur Zeit  $t = 0$  mit einem vollständig entladenen Kondensator, also  $Q(0) = 0$ . Wir haben somit das Anfangswertproblem

$$\dot{Q} + \frac{1}{RC}Q = \frac{U_a}{R}, \quad Q(0) = 0. \quad (5.20)$$

Aus  $a(t) = \frac{1}{RC}$  folgt  $A(t) = \frac{1}{RC}t$ . Einsetzen in die Variation der Konstanten Formel (5.18) ergibt

$$Q(t) = e^{-A(t)} \cdot \left( c_0 + \int e^{A(t)} g(t) dt \right) = e^{-\frac{1}{RC}t} \cdot \left( c_0 + \int e^{\frac{t}{RC}} \cdot \frac{U_a}{R} dt \right)$$

Das Integral ist

$$\int e^{\frac{t}{RC}} \frac{U_a}{R} dt = e^{\frac{t}{RC}} \cdot RC \cdot \frac{U_a}{R} = U_a C e^{\frac{t}{RC}},$$

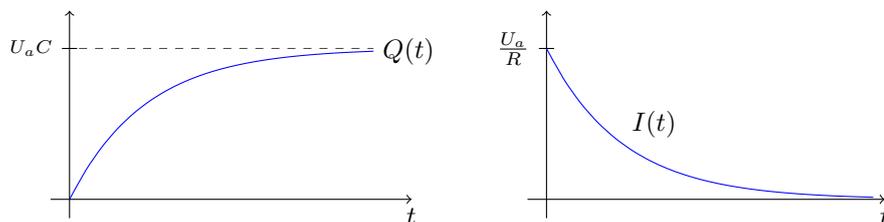


Abbildung 5.3: Ladungs- und Stromverlauf beim Aufladen eines Kondensators

so dass für die Lösung folgt

$$Q(t) = e^{-\frac{1}{RC}t} \cdot \left( c_0 + U_a C e^{\frac{1}{RC}t} \right) = c_0 e^{-\frac{1}{RC}t} + U_a C.$$

Dies ist noch die allgemeine Lösung mit dem freien Parameter  $c_0 \in \mathbb{R}$ . Diesen können wir jetzt aus dem Anfangswert berechnen:

$$Q(0) = c_0 + U_a C \stackrel{!}{=} 0 \quad \Rightarrow \quad c_0 = -U_a C.$$

Damit ist

$$Q(t) = -U_a C e^{-\frac{1}{RC}t} + U_a C = U_a C \left( 1 - e^{-\frac{1}{RC}t} \right) \quad (5.21)$$

die Lösung des Anfangswertproblems (5.20). Diese Funktion beschreibt also den zeitlichen Verlauf der Kondensatorladung beim Aufladen des Kondensators. Man erkennt, dass  $Q(t)$  tatsächlich monoton wachsend ist, denn  $e^{-\frac{1}{RC}t}$  ist monoton fallend und daher  $1 - e^{-\frac{1}{RC}t}$  monoton wachsend. Außerdem nähert sich die Ladung  $Q(t)$  für große  $t$  asymptotisch dem Wert  $U_a C$  an, denn  $e^{-\frac{1}{RC}t} \rightarrow 0$  für  $t \rightarrow \infty$ . Den Ladestrom erhält man aus (5.21) durch Ableiten:

$$I(t) = \dot{Q}(t) = U_a C \cdot \frac{1}{RC} e^{-\frac{1}{RC}t} = \frac{U_a}{R} e^{-\frac{1}{RC}t}$$

Ladungs- und Stromverlauf sind in Abbildung 5.3 dargestellt.

## 5.9 Das Richtungsfeld einer Differentialgleichung 1. Ordnung

Zu einer Differentialgleichung 1. Ordnung kann man das sogenannte *Richtungsfeld* zeichnen. Dieses Richtungsfeld gibt eine graphische Veranschaulichung der Differentialgleichung und zeigt dabei auch den qualitativen Verlauf der Lösungen der Dgl.

Wir betrachten eine Differentialgleichung 1. Ordnung in ganz allgemeiner Form<sup>2</sup>,

$$y' = f(x, y). \quad (5.22)$$

Für eine Lösung  $y = y(x)$  dieser Gleichung gilt also

$$y'(x) = f(x, y(x)).$$

<sup>2</sup>Dies ist die explizite Form. Das bedeutet, dass die Gleichung nach  $y'$  aufgelöst ist.

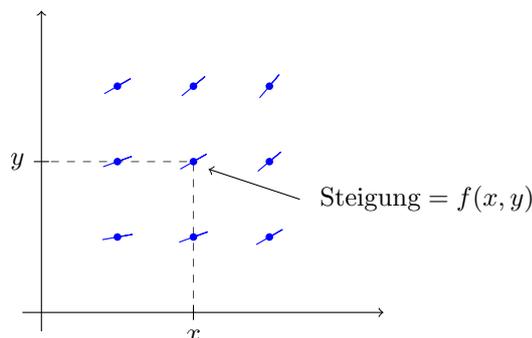


Abbildung 5.4: Das Richtungsfeld einer Differentialgleichung

Das bedeutet, dass die Lösung  $y(x)$  an der Stelle  $x$ , also im Punkt  $(x, y)$  mit  $y = y(x)$ , die Steigung  $y'(x) = f(x, y)$  hat. Kurz:

$$f(x, y) = \text{Steigung einer Lösung im Punkt } (x, y).$$

Das wichtige ist hier, dass wir die Steigung einer Lösung durch den Punkt  $(x, y)$  berechnen können *ohne* die Lösungsfunktion  $y(x)$  explizit zu kennen, denn wir brauchen in die rechte Seite  $f(x, y)$  der Differentialgleichung ja nur die Koordinaten des Punktes  $(x, y)$  einzusetzen.

Das Richtungsfeld der Differentialgleichung (5.22) erhält man nun, indem man an vielen Punkten  $(x, y)$  jeweils die entsprechende Steigung einzeichnet, siehe Abbildung 5.9.

Das Richtungsfeld gibt die Steigung der Lösungen der Differentialgleichung vor. Anders ausgedrückt: Jede Lösung der Differentialgleichung läuft an jeder Stelle tangential zum Richtungsfeld. Wir illustrieren das am nächsten Beispiel.

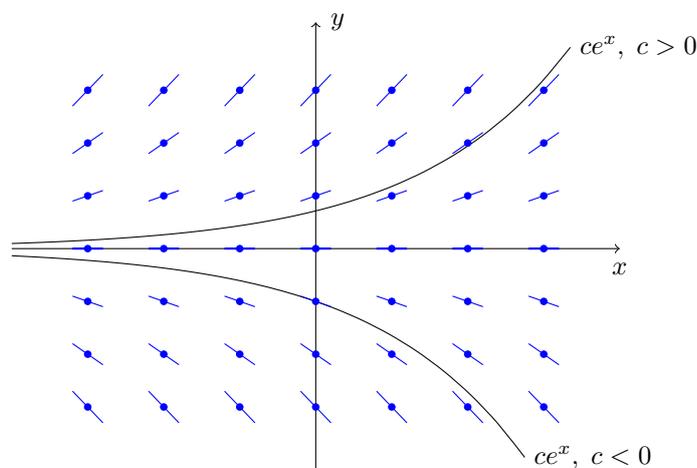
**Beispiel 5.9.1** Wir betrachten die ganz einfache Differentialgleichung

$$y' = y \tag{5.23}$$

also den Fall  $f(x, y) = y$ . Wir konstruieren das Richtungsfeld der Dgl indem wir für eine Reihe von Punkten  $(x, y)$  im Koordinatensystem die entsprechende Steigung einzeichnen; die Steigung ist gegeben durch  $f(x, y)$ , also hier gleich  $y$ . Damit haben wir hier also die gleiche Steigung für alle Punkte mit gleichem  $y$ -Wert, d.h. für Punkte auf Geraden parallel zur  $x$ -Achse, siehe Abbildung 5.5. Für Geraden oberhalb der  $x$ -Achse ist  $y > 0$ , die Steigung ist also positiv; und je höher die Gerade, also je größer  $y$ , desto größer auch die Steigung. Auf Geraden unterhalb der  $x$ -Achse dagegen ist die Steigung negativ (denn dann ist  $y < 0$ ), und sie wird umso stärker negativ (d.h. ihr Betrag wird größer) je weiter unterhalb der  $x$ -Achse die Gerade liegt. Auf der  $x$ -Achse selbst ist  $y = 0$  und damit die Steigung des Richtungsfelds Null.

Man kann nun das Richtungsfeld benutzen, um den Graphen von Lösungen der Dgl zu skizzieren. Da die Lösungskurve in jedem Punkt tangential, also parallel zum Richtungsfeld ist, folgt somit die Lösung von einem Startpunkt ausgehend immer dem Richtungsfeld. Auf diese Art erhält man in Abbildung 5.5 tatsächlich die Graphen der allgemeine Lösung von (5.23)

$$y(x) = ce^x, \quad c \in \mathbb{R},$$

Abbildung 5.5: Richtungsfeld von  $y' = y$ 

die wir in Beispiel 5.4.1(b) schon berechnet haben.

**Bemerkung:** Differentialgleichungen 2. Ordnung haben kein Richtungsfeld. Eine allgemeine Dgl 2. Ordnung ist von der Form

$$y'' = f(x, y, y')$$

und legt somit nicht die Steigung  $y'$  sondern die Krümmung  $y''$  fest. Ist die Dgl nicht explizit von  $x$  abhängig (*autonom*),

$$y'' = f(y, y'),$$

so kann man aber ein sogenanntes *Phasenporträt* anfertigen, das den Verlauf von Lösungen in einem  $(y, y')$ -Koordinatensystem darstellt.

## 5.10 Lineare Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten

Wir wenden uns jetzt den Differentialgleichungen 2. Ordnung zu. Dabei betrachten wir vor allem lineare Differentialgleichung, mit konstanten Koeffizienten, für die wir die Lösung in sehr allgemeinen Fällen angeben können. Die Gleichung

$$y'' + ay' + by = g(x) \quad (5.24)$$

heißt *lineare Differentialgleichung 2. Ordnung*, die Koeffizienten  $a, b \in \mathbb{R}$  sind dabei konstant. Wie schon bei Gleichungen 1. Ordnung ist (5.24) *homogen* wenn  $g(x) = 0$  und sonst *inhomogen*.

Für lineare Differentialgleichung gilt das *Superpositionsprinzip*, mit dem man aus schon bekannten Lösungen neue Lösungen konstruieren kann. Dabei werden Lösungen von (5.24) zu möglicherweise verschiedenen rechten Seiten  $g(x)$  miteinander kombiniert:

**Superpositionsprinzip (Linearität).** Ist  $y_1(x)$  eine Lösung von

$$y'' + ay' + by = g_1(x) \quad (5.25)$$

und  $y_2(x)$  eine Lösung von

$$y'' + ay' + by = g_2(x) \quad (5.26)$$

dann ist die Linearkombination  $y(x) = c_1y_1(x) + c_2y_2(x)$  ( $c_1, c_2 \in \mathbb{R}$ ) eine Lösung von

$$y'' + ay' + by = c_1g_1(x) + c_2g_2(x). \quad (5.27)$$

**Beweis.** Aus  $y(x) = c_1y_1(x) + c_2y_2(x)$ , oder kurz  $y = c_1y_1 + c_2y_2$ , folgt

$$\begin{aligned} y'' + ay' + by &= c_1y_1'' + c_2y_2'' + a(c_1y_1' + c_2y_2') + b(c_1y_1 + c_2y_2) \\ &= c_1 \cdot (y_1'' + ay_1' + by_1) + c_2 \cdot (y_2'' + ay_2' + by_2) \\ &= c_1g_1(x) + c_2g_2(x) \end{aligned}$$

Dabei haben wir im letzten Schritt verwendet, dass  $y_1(x)$  Lösung von (5.25) ist, also  $y_1'' + ay_1' + by_1 = g_1(x)$ , und  $y_2(x)$  Lösung von (5.26), d.h.  $y_2'' + ay_2' + by_2 = g_2(x)$ .  $\square$

Die folgenden beiden Spezialfälle des Superpositionsprinzips sind besonders wichtig:

- (a) Sind  $y_1$  und  $y_2$  beides Lösungen der homogenen Differentialgleichung

$$y'' + ay' + by = 0 \quad (5.28)$$

dann ist auch  $c_1y_1 + c_2y_2$  Lösung der homogenen Gleichung (5.28). (Das ist der Fall  $g_1 = g_2 = 0$ .)

- (b) Ist  $y_h$  Lösung der homogenen Differentialgleichung (5.28) und  $y_p$  Lösung der inhomogenen Differentialgleichung

$$y'' + ay' + by = g(x) \quad (5.29)$$

dann ist auch  $y_h + y_p$  Lösung der inhomogenen Gleichung (5.29). (Das ist der Fall  $c_1 = c_2 = 1, g_1 = 0$ .)

**Bemerkung:** Das Superpositionsprinzip spiegelt die Eigenschaft der *Linearität* der Differentialgleichung wider. Das ist genau die gleiche Eigenschaft, die wir auch schon in Mathematik A bei linearen Gleichungssystemen kennengelernt haben: Für die homogene Gleichung sind Summen und konstante skalare Vielfache von Lösungen wieder Lösungen. Und die (allgemeine) Lösung der inhomogenen Gleichung ist die Summe aus der (allgemeinen) Lösung der homogenen Gleichung und einer speziellen (*partikulären*) Lösung der inhomogenen Gleichung.

## 5.11 Homogene Gleichungen 2. Ordnung und das charakteristische Polynom

Für die Konstruktion von Lösungen linearer Differentialgleichungen 2. Ordnung betrachten wir, wie schon im Fall 1. Ordnung, zuerst die homogene Gleichung

$$y'' + ay' + by = 0. \quad (5.30)$$

Um den Ansatz für die Lösung zu motivieren, erinnern wir uns daran, dass die ähnliche homogene lineare Dgl 1. Ordnung  $y' + ay = 0$  die Lösung  $y(x) = e^{-ax}$  hat (Satz 5.6.1 mit  $a(x) = a \Rightarrow A(x) = ax$ ). Die Idee ist, dass eine Lösung der Gleichung 2. Ordnung (5.30) auch wieder eine Exponentialfunktion ist. Wir machen daher den *Ansatz*

$$y(x) = e^{\lambda x}$$

mit einer noch unbekanntem Zahl  $\lambda$ . Wir müssen nun  $\lambda$  so bestimmen, dass unser Ansatz tatsächlich eine Lösung von (5.30) ist. Dazu setzen wir  $y(x) = e^{\lambda x}$  in die Dgl ein. Wegen

$$y'(x) = \lambda e^{\lambda x}, \quad y''(x) = \lambda^2 e^{\lambda x}$$

gilt

$$\begin{aligned} y''(x) + ay'(x) + by(x) &= \lambda^2 e^{\lambda x} + a\lambda e^{\lambda x} + be^{\lambda x} \\ &= \underbrace{e^{\lambda x}}_{\neq 0} \cdot (\lambda^2 + a\lambda + b) \stackrel{!}{=} 0 \\ \iff \lambda^2 + a\lambda + b &= 0 \end{aligned}$$

Damit ist  $y(x) = e^{\lambda x}$  eine Lösung genau dann, wenn  $\lambda$  eine Nullstelle des Polynoms

$$p(\lambda) = \lambda^2 + a\lambda + b \quad (5.31)$$

ist. Dieses Polynom nennt man *charakteristisches Polynom* der Differentialgleichung (5.30). Jede Nullstelle von  $p(\lambda)$  liefert also eine Lösung  $e^{\lambda x}$  von (5.30) und nach dem Superpositionsprinzip des letzten Abschnitts ist jede Linearkombination davon wieder eine Lösung. Auf diese Weise erhalten wir tatsächlich die allgemeine Lösung von (5.30):

**Satz 5.11.1** Die allgemeine Lösung von

$$y'' + ay' + by = 0$$

ist

$$y(x) = c_1 y_1(x) + c_2 y_2(x), \quad c_1, c_2 \in \mathbb{R}. \quad (5.32)$$

Dabei ergeben sich die Funktionen  $y_1, y_2$  wie folgt aus den Nullstellen des charakteristischen Polynoms  $p(\lambda) = \lambda^2 + a\lambda + b$ :

(a) bei zwei Nullstellen  $\lambda_1 \neq \lambda_2$  ist

$$y_1(x) = e^{\lambda_1 x}, \quad y_2(x) = e^{\lambda_2 x};$$

(b) bei einer doppelten Nullstelle  $\lambda_1$  ist

$$y_1(x) = e^{\lambda_1 x}, \quad y_2(x) = x e^{\lambda_1 x}.$$

Die Funktionen  $y_1$  und  $y_2$ , aus denen die allgemeine Lösung (5.32) zusammengesetzt ist, nennt man *Fundamentalsystem* der Differentialgleichung.

**Beweis.**

(a) Fall von zwei Nullstellen  $\lambda_1 \neq \lambda_2$ : Wir haben oben schon gesehen, dass die Nullstellen  $\lambda_1, \lambda_2$  des charakteristischen Polynoms die Lösungen

$$y_1(x) = e^{\lambda_1 x}, \quad y_2(x) = e^{\lambda_2 x}$$

ergeben. Nach dem Superpositionsprinzip in 5.10 ist damit (5.32) eine Lösung der homogenen Dgl. Um zu zeigen, dass dies auch die allgemeine Lösung ist, muss man noch überprüfen, dass damit ein beliebiges Anfangswertproblem lösbar ist. Da die Differentialgleichung von 2. Ordnung ist, haben wir zwei Anfangswerte

$$y(x_0) = u, \quad y'(x_0) = v.$$

Einsetzen von (5.32) liefert das lineare Gleichungssystem

$$\begin{cases} y(x_0) = c_1 y_1(x_0) + c_2 y_2(x_0) \stackrel{!}{=} u \\ y'(x_0) = c_1 y_1'(x_0) + c_2 y_2'(x_0) \stackrel{!}{=} v \end{cases} \\ \Leftrightarrow \begin{pmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$$

für die Parameter  $c_1, c_2$ . Das Gleichungssystem ist lösbar (für beliebige Anfangswerte  $u, v$ ) genau dann, wenn

$$\det \begin{pmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{pmatrix} \neq 0.$$

Es gilt

$$\begin{aligned} \det \begin{pmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{pmatrix} &= \det \begin{pmatrix} e^{\lambda_1 x_0} & e^{\lambda_2 x_0} \\ \lambda_1 e^{\lambda_1 x_0} & \lambda_2 e^{\lambda_2 x_0} \end{pmatrix} \\ &= (\lambda_2 - \lambda_1) e^{\lambda_1 x_0} e^{\lambda_2 x_0} \neq 0 \end{aligned}$$

weil  $\lambda_1 \neq \lambda_2$ . Also ist (5.32) die allgemeine Lösung.

(b) Fall einer doppelten Nullstelle  $\lambda_1$ : Wir wissen wieder, dass  $y_1(x) = e^{\lambda_1 x}$  eine Lösung ist. Für  $y_2(x) = x e^{\lambda_1 x}$  müssen wir das noch nachrechnen: Nach der Produktregel ist

$$y_2'(x) = e^{\lambda_1 x} + \lambda_1 x e^{\lambda_1 x}, \quad y_2''(x) = 2\lambda_1 e^{\lambda_1 x} + \lambda_1^2 x e^{\lambda_1 x},$$

und damit

$$y_2''(x) + a y_2'(x) + b y_2(x) = x e^{\lambda_1 x} (\lambda_1^2 + a \lambda_1 + b) + e^{\lambda_1 x} (2\lambda_1 + a) = e^{\lambda_1 x} (2\lambda_1 + a),$$

wobei wir benutzt haben, dass  $\lambda_1$  Nullstelle von  $p$  ist, also  $p(\lambda_1) = \lambda_1^2 + a\lambda_1 + b = 0$ . Da  $\lambda_1$  eine doppelte Nullstelle ist, hat das charakteristische Polynom die Faktorisierung

$$p(\lambda) = (\lambda - \lambda_1)^2 = \lambda^2 - 2\lambda_1\lambda + \lambda_1^2.$$

Koeffizientenvergleich mit  $p(\lambda) = \lambda^2 + a\lambda + b$  ergibt  $a = -2\lambda_1$ , also  $2\lambda_1 + a = 0$ . Somit folgt  $y_2''(x) + ay_2'(x) + by_2(x) = 0$ , d.h. auch  $y_2$  ist eine Lösung der homogenen Dgl. Wie in (a) bleibt noch die Frage, ob (5.32) die allgemeine Lösung liefert. Hier gilt

$$\det \begin{pmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{pmatrix} = \det \begin{pmatrix} e^{\lambda_1 x_0} & x_0 e^{\lambda_1 x_0} \\ \lambda_1 e^{\lambda_1 x_0} & e^{\lambda_1 x_0} + \lambda_1 x_0 e^{\lambda_1 x_0} \end{pmatrix} = e^{2\lambda_1 x_0} \neq 0,$$

also ist (5.32) auch in diesem Fall die allgemeine Lösung.

□

**Beispiel 5.11.2** (a) Wir berechnen die allgemeine Lösung der homogenen Differentialgleichung

$$y'' - 2y = 0.$$

Das charakteristische Polynom dazu ist

$$p(\lambda) = \lambda^2 - 2.$$

Für die Nullstellen erhalten wir

$$\lambda^2 - 2 = 0 \quad \Leftrightarrow \quad \lambda = \pm\sqrt{2},$$

also zwei verschiedenen Nullstellen

$$\lambda_1 = \sqrt{2}, \quad \lambda_2 = -\sqrt{2}.$$

Nach Satz 5.11.1 (Fall  $\lambda_1 \neq \lambda_2$ ) ist damit

$$y_1(x) = e^{\sqrt{2}x}, \quad y_2(x) = e^{-\sqrt{2}x}$$

ein Fundamentalsystem der Dgl, und die allgemeine Lösung ist

$$y(x) = c_1 e^{\sqrt{2}x} + c_2 e^{-\sqrt{2}x}, \quad c_1, c_2 \in \mathbb{R}.$$

(b) Wir berechnen die Lösung des Anfangswertproblems

$$y'' - 2y' - 3y = 0, \quad y(0) = 1, \quad y'(0) = -2.$$

Das charakteristische Polynom ist  $p(\lambda) = \lambda^2 - 2\lambda - 3$ . Nullstellen:

$$\begin{aligned} p(\lambda) = \lambda^2 - 2\lambda - 3 = 0 &\Leftrightarrow (\lambda - 1)^2 = 4 \Leftrightarrow \lambda = 1 \pm 2 \\ &\implies \lambda_1 = -1, \quad \lambda_2 = 3 \end{aligned}$$

Die allgemeine Lösung der Dgl ist also

$$y(x) = c_1 e^{-x} + c_2 e^{3x}.$$

Für die Lösung des Anfangswertproblems berechnen wir nun zuerst  $y'(x)$ :

$$y'(x) = -c_1 e^{-x} + 3c_2 e^{3x}.$$

Die Anfangsbedingungen lauten damit

$$\begin{aligned} y(0) &= c_1 \cdot e^0 + c_2 \cdot e^0 = c_1 + c_2 \stackrel{!}{=} 1 \\ y'(0) &= -c_1 e^0 + 3c_2 e^0 = -c_1 + 3c_2 \stackrel{!}{=} -2 \end{aligned}$$

Dies ist das Gleichungssystem

$$\left| \begin{array}{cc|c} c_1 & +c_2 & = 1 & \text{I} \\ -c_1 & +3c_2 & = -2 & \text{II} \end{array} \right|$$

Addition der Gleichungen I+II liefert

$$4c_2 = -1 \Leftrightarrow c_2 = -\frac{1}{4} \xrightarrow{\text{I}} c_1 = 1 - c_2 = 1 + \frac{1}{4} = \frac{5}{4}$$

Somit ist

$$y(x) = \frac{5}{4} e^{-x} - \frac{1}{4} e^{3x}$$

die Lösung des Anfangswertproblems.

**Bemerkung:** Die Menge  $L$  aller Lösungen der homogenen linearen Differentialgleichung 2. Ordnung (5.30) ist ein *Vektorraum* (vgl. Mathematik A). Denn nach dem Superpositionsprinzip gilt

$$y, \tilde{y} \in L \implies y + \tilde{y} \in L, \quad c \cdot y \in L \quad (c \in \mathbb{R}).$$

Jedes  $y \in L$  hat eine Darstellung  $y = c_1 y_1 + c_2 y_2$  im Fundamentalsystem  $y_1, y_2$  (Satz 5.11.1). Diese Darstellung ist eindeutig, denn  $c_1, c_2$  ergeben sich als Lösungen des Gleichungssystems aus den Anfangswerten von  $y$ . Da *jedes*  $y \in L$  also eine *eindeutige* Darstellung im Fundamentalsystem hat, folgt, dass das System  $(y_1, y_2)$  eine *Basis* von  $L$  ist. Da die Basis aus zwei Funktionen besteht, hat der Vektorraum  $L$  somit die Dimension 2:

$$\dim L = 2.$$

Die Dimension ist also gleich der Ordnung der Differentialgleichung! Dies gilt tatsächlich für lineare Differentialgleichungen jeder Ordnung: Die Menge der Lösungen einer homogenen linearen Gleichung  $n$ . Ordnung hat die Dimension  $n$ , d.h. es gibt ein Fundamentalsystem aus  $n$  Funktionen. Für die homogene lineare Dgl 1. Ordnung haben wir das übrigens in Satz 5.6.1 gesehen.

## 5.12 Reelle und komplexe Fundamentalsysteme

Nach Satz 5.11.1 hat die homogenen linearen Dgl 2. Ordnung (5.30) die allgemeine Lösung

$$y(x) = c_1 y_1(x) + c_2 y_2(x).$$

Das Fundamentalsystem  $y_1, y_2$  erhält man dabei aus den Nullstellen des charakteristischen Polynoms  $p(\lambda)$ . Hat  $p$  zwei verschiedene *komplexe* Nullstellen

$\lambda_1, \lambda_2 \in \mathbb{C}$ , dann ist das zugehörige Fundamentalsystem  $y_1(x) = e^{\lambda_1 x}$ ,  $y_2(x) = e^{\lambda_2 x}$  ebenfalls komplex und damit auch die allgemeine Lösung  $y(x)$ .

Da das charakteristische Polynom  $p(\lambda) = \lambda^2 + a\lambda + b$  reell ist<sup>3</sup>, d.h.  $a, b \in \mathbb{R}$ , sind komplexe Nullstellen von  $p$  zueinander komplex konjugiert,  $\lambda_2 = \bar{\lambda}_1$ , also

$$\lambda_1 = \alpha + j\beta, \quad \lambda_2 = \alpha - j\beta.$$

Damit sind auch die Funktionen im Fundamentalsystem zueinander komplex konjugiert, denn

$$y_2(x) = e^{\lambda_2 x} = e^{\bar{\lambda}_1 x} = \overline{e^{\lambda_1 x}} = \overline{y_1(x)}.$$

Wir können nun diese Eigenschaft benutzen, um aus dem komplexen ein äquivalentes *reelles* Fundamentalsystem zu konstruieren. Damit erhalten wir dann auch die allgemeine Lösung in reeller Form. Wir benutzen zuerst die Eulerformel für die komplexe Exponentialfunktion:

$$\begin{aligned} y_1(x) &= e^{(\alpha+j\beta)x} = e^{\alpha x} e^{j\beta x} = e^{\alpha x} (\cos(\beta x) + j \sin(\beta x)) \\ &= e^{\alpha x} \cos(\beta x) + j e^{\alpha x} \sin(\beta x). \end{aligned}$$

Nach dem Superpositionsprinzip ist  $u_1(x) = \frac{1}{2}y_1(x) + \frac{1}{2}y_2(x)$  auch eine Lösung der homogenen Dgl. Aus  $y_2(x) = \overline{y_1(x)}$  folgt dann

$$u_1(x) = \frac{1}{2}(y_1(x) + \overline{y_1(x)}) = \operatorname{Re} y_1(x) = e^{\alpha x} \cos(\beta x).$$

In analoger Weise ist auch

$$u_2(x) = \frac{1}{2j}(y_1(x) - \overline{y_1(x)}) = \operatorname{Im} y_1(x) = e^{\alpha x} \sin(\beta x)$$

eine Lösung. Da man aus den neuen reellen Lösungen  $u_1(x), u_2(x)$  das ursprüngliche Fundamentalsystem durch  $y_1 = u_1 + ju_2$  und  $y_2 = u_1 - ju_2$  zurückerhalten kann, ist auch

$$u_1(x) = e^{\alpha x} \cos(\beta x), \quad u_2(x) = e^{\alpha x} \sin(\beta x) \quad (5.33)$$

ein Fundamentalsystem, das *reelle Fundamentalsystem* zu den komplex konjugierten Nullstellen  $\lambda_{1/2} = \alpha \pm j\beta$  von  $p(\lambda)$ . Linearkombination von  $u_1$  und  $u_2$  liefert dann die allgemeine Lösung von (5.30) in reeller Form

$$y(x) = c_1 u_1(x) + c_2 u_2(x), \quad c_1, c_2 \in \mathbb{R}. \quad (5.34)$$

**Beispiel 5.12.1** Wir berechnen Fundamentalsysteme und die allgemeine Lösung von

$$y'' - 2y' + 2y = 0.$$

Zuerst die Nullstellen des charakteristischen Polynoms:

$$p(\lambda) = \lambda^2 - 2\lambda + 2 = 0 \Leftrightarrow (\lambda - 1)^2 = -1 \Leftrightarrow \lambda = 1 \pm j$$

Wir haben hier also den Fall konjugiert komplexer Nullstellen

$$\lambda_1 = 1 + j, \quad \lambda_2 = 1 - j,$$

<sup>3</sup>Wir gehen hier von *reellen* Differentialgleichungen aus, also  $y'' + ay' + by = 0$  mit  $a, b \in \mathbb{R}$ .

(also  $\alpha = 1$  und  $\beta = 1$ .) Das zugehörige komplexe Fundamentalsystem ist dann

$$y_1(x) = e^{(1+j)x}, \quad y_2(x) = e^{(1-j)x}$$

und liefert die allgemeine Lösung in komplexer Form

$$y(x) = c_1 e^{(1+j)x} + c_2 e^{(1-j)x}.$$

Das reelle Fundamentalsystem ist

$$u_1(x) = e^x \cos x, \quad u_2(x) = e^x \sin x$$

mit entsprechender allgemeinen (reellen) Lösung

$$y(x) = c_1 e^x \cos x + c_2 e^x \sin x.$$

**Beachte:** Bei konjugiert komplexen Nullstellen kann man die allgemeine Lösung der Dgl wahlweise über das komplexe oder das reelle Fundamentalsystem berechnen. Obwohl sich dadurch zunächst verschiedene Formeln für die allgemeine Lösung ergeben, (5.32) und (5.34), erhält man aus beiden Formeln dieselben Lösungen (vorausgesetzt man lässt für  $c_1, c_2$  auch komplexe Zahlen zu). Beide Formeln für die allgemeine Lösung sind also äquivalent. Speziell kann man auch beide Formen verwenden, um ein Anfangswertproblem zu lösen und erhält auf beiden Wegen dieselbe Lösung. (Die Lösung des Anfangswertproblems ist eindeutig.) In praktischer Hinsicht ist der Vorteil des reellen Fundamentalsystems, dass damit das Rechnen mit komplexen Zahlen vermieden werden kann, z.B. bei der Lösung von Anfangswertproblemen (mit reellen Anfangswerten).

## 5.13 Inhomogene lineare Differentialgleichungen 2. Ordnung

Wir untersuchen jetzt die Lösungen der inhomogenen linearen Differentialgleichung 2. Ordnung

$$y'' + ay' + by = g(x). \quad (5.35)$$

Nach dem Superpositionsprinzip aus 5.10 ist die Summe einer Lösung  $y_h$  der homogenen Dgl und einer Lösung  $y_p$  der inhomogenen Dgl (5.35) wieder eine Lösung von (5.35). Benutzen wir für  $y_h$  die allgemeine Lösung der homogenen Gleichung, so erhalten wir die allgemeine Lösung von (5.35):

**Satz 5.13.1** *Ist  $y_p(x)$  eine spezielle (partikuläre) Lösung der inhomogenen Differentialgleichung (5.35) und ist  $y_h(x)$  die allgemeine Lösung der zugehörigen homogenen Differentialgleichung (5.30) dann ist*

$$y(x) = y_h(x) + y_p(x)$$

*die allgemeine Lösung der inhomogenen Differentialgleichung. Ist  $y_1, y_2$  ein Fundamentalsystem der homogenen Gleichung, so gilt  $y_h(x) = c_1 y_1(x) + c_2 y_2(x)$  und damit für die allgemeine inhomogene Lösung*

$$y(x) = c_1 y_1(x) + c_2 y_2(x) + y_p(x) \quad (c_1, c_2 \in \mathbb{R}).$$

$g(x)$	Ansatz für $y_p(x)$	Bedingung
$P(x)$	$Q(x)$	$p(0) \neq 0$
	$xQ(x)$	$p(0) = 0$
$P(x)e^{\alpha x}$	$Q(x)e^{\alpha x}$	$p(\alpha) \neq 0$
	$xQ(x)e^{\alpha x}$	$p(\alpha) = 0$ (einfach)
	$x^2Q(x)e^{\alpha x}$	$p(\alpha) = 0$ (doppelt)
$A_1 \cos(\omega x) + A_2 \sin(\omega x)$	$B_1 \cos(\omega x) + B_2 \sin(\omega x)$	$p(j\omega) \neq 0$
	$x(B_1 \cos(\omega x) + B_2 \sin(\omega x))$	$p(j\omega) = 0$

Tabelle 5.1: Ansätze für die partikuläre Lösung

**Beweis.** Nach dem Superpositionsprinzip ist  $y_h + y_p$  eine Lösung von (5.35). Um zu zeigen, dass dies die allgemeine Lösung ist, zeigen wir wieder, dass sich damit jedes Anfangswertproblem lösen lässt. Die Bedingungen

$$y(x_0) = y_h(x_0) + y_p(x_0) \stackrel{!}{=} u$$

$$y'(x_0) = y_h'(x_0) + y_p'(x_0) \stackrel{!}{=} v$$

ergeben nach Einsetzen von  $y_h = c_1 y_1 + c_2 y_2$  wie im Beweis von Satz 5.11.1 ein lineares Gleichungssystem für  $c_1, c_2$ , das wie dort gezeigt eine eindeutige Lösung hat.  $\square$

Nach Satz 5.13.1 langt es also für die allgemeine Lösung von (5.35) eine spezielle Lösung von (5.35) zu finden. Diese spezielle Lösung wird dann auch *partikuläre* Lösung genannt. Bleibt die Frage: Wie findet man so eine partikuläre Lösung?

Eine Methode ist, einen bestimmten Ansatz zu machen, der von der Form der rechten Seite  $g(x)$  (der „Inhomogenität“) abhängt. Dies funktioniert für Polynome, Exponentialfunktionen sowie Sinus und Cosinus-Funktionen als  $g(x)$  und ist in Tabelle 5.1 dargestellt. Hierbei sind  $P(x)$  und  $Q(x)$  Polynome vom selben Grad, und das Polynom  $Q(x)$  enthält für den Ansatz zunächst allgemeine Koeffizienten, die dann so zu bestimmen sind, dass sich eine Lösung der inhomogenen Dgl ergibt. Kommt etwa in der Inhomogenität  $g(x)$  das Polynom  $P(x) = x^2 + 4x$  vor, so ist der Ansatz für  $Q(x)$  ein allgemeines Polynom 2. Grades, also  $Q(x) = Ax^2 + Bx + C$ .

Der zu wählende Ansatz ist noch von einer Bedingung an die Nullstellen des charakteristischen Polynoms  $p(\lambda)$  abhängig (letzte Spalte der Tabelle). Ist – je nach Inhomogenität – eine bestimmte Zahl eine Nullstelle von  $p(\lambda)$ , so muss der Ansatz noch mit einem zusätzlichen Faktor  $x$  versehen werden, bei einer doppelten Nullstelle mit  $x^2$ . (Die doppelte Nullstelle ist nur bei  $g(x) = P(x)e^{\alpha x}$  möglich.)

**Beispiel 5.13.2** Wir berechnen zunächst eine partikuläre und damit dann die allgemeine Lösung der Differentialgleichung

$$y'' + y = e^{2x}.$$

Die Inhomogenität ist  $g(x) = e^{2x}$ . Ein Vergleich mit Tabelle 5.1 zeigt

$$g(x) = e^{2x} = P(x)e^{\alpha x} \quad \text{mit} \quad P(x) = 1, \quad \alpha = 2.$$

Das Polynom  $P(x) = 1$  hat den Grad 0. Also wählen wir für  $Q(x)$  das allgemeine Polynom 0. Grades, das ist das konstante Polynom  $Q(x) = A$ . Jetzt überprüfen wir noch die Bedingung: Das charakteristische Polynom ist  $p(\lambda) = \lambda^2 + 1$ . Damit ist

$$p(\alpha) = p(2) = 5 \neq 0.$$

Unser Ansatz für die partikuläre Lösung ist damit

$$y_p(x) = Q(x)e^{2x} = Ae^{2x}$$

mit der noch zu bestimmenden Konstante  $A$ . Zur Berechnung von  $A$  setzen wir den Ansatz nun in die Dgl ein:

$$\begin{aligned} y_p'(x) &= 2Ae^{2x}, & y_p''(x) &= 4Ae^{2x} \\ \Rightarrow y_p''(x) + y_p(x) &= 4Ae^{2x} + Ae^{2x} = 5Ae^{2x} \stackrel{!}{=} e^{2x} \end{aligned}$$

Der Ansatz erfüllt also die Dgl genau dann, wenn

$$5A = 1 \quad \Rightarrow \quad A = \frac{1}{5}.$$

Somit ist die partikuläre Lösung

$$y_p(x) = \frac{1}{5}e^{2x}.$$

Jetzt zur allgemeinen Lösung. Dazu bestimmen wir die allgemeine homogene Lösung aus den Nullstellen des charakteristischen Polynoms:

$$p(\lambda) = \lambda^2 + 1 = 0 \quad \Leftrightarrow \quad \lambda^2 = -1 \quad \Leftrightarrow \quad \lambda = \pm j.$$

Dies ist der Fall konjugiert komplexer Nullstellen  $\lambda_1 = j$ ,  $\lambda_2 = -j$ . Das reelle Fundamentalsystem ist dann ( $\alpha = 0$ ,  $\beta = 1$ )

$$y_1(x) = \cos(x), \quad y_2(x) = \sin(x)$$

und damit die allgemeine homogene Lösung

$$y_h(x) = c_1 \cos(x) + c_2 \sin(x), \quad c_1, c_2 \in \mathbb{R}.$$

Somit ist schließlich die gesuchte allgemeine Lösung der inhomogenen Differentialgleichung

$$y(x) = y_h(x) + y_p(x) = c_1 \cos(x) + c_2 \sin(x) + \frac{1}{5}e^{2x}, \quad c_1, c_2 \in \mathbb{R}.$$

**Hinweis:** Man kann das Überprüfen der Bedingung in Tabelle 5.1 weglassen und zunächst mit dem Ansatz ohne zusätzlichen Faktor  $x$  starten (also den Fall  $p(\dots) \neq 0$  annehmen). Wenn der Ansatz eine Lösung liefert, ist man fertig. Wenn der Ansatz aber nicht funktioniert und einen Widerspruch ergibt, dann weiß man, dass man doch im Fall  $p(\dots) = 0$  ist und somit zum Ansatz noch einen Faktor  $x$  hinzufügen muss.

## 5.14 Lösung durch Übergang zum Komplexen

Für bestimmte Arten von Inhomogenitäten  $g(x)$  kann man eine partikuläre Lösung dadurch bestimmen, dass man zunächst zu einer komplexen Differentialgleichung übergeht und dafür eine partikuläre Lösung berechnet. Betrachten wir etwa eine Gleichung der Form

$$y'' + ay' + by = e^{\alpha x} \cos(\beta x). \quad (5.36)$$

Die rechte Seite  $g(x)$  ist keine der Inhomogenitäten aus Tabelle 5.1. Die Idee ist nun,  $g(x)$  mit Hilfe der Eulerformel als Realteil einer komplexen Exponentialfunktion zu schreiben:

$$\begin{aligned} e^{(\alpha+j\beta)x} &= e^{\alpha x} e^{j\beta x} = e^{\alpha x} (\cos(\beta x) + j \sin(\beta x)) \\ &= e^{\alpha x} \cos(\beta x) + j e^{\alpha x} \sin(\beta x) \end{aligned}$$

Es ist also  $e^{\alpha x} \cos(\beta x) = \operatorname{Re}(e^{(\alpha+j\beta)x})$ . Wir betrachten daher jetzt die *komplexe* Differentialgleichung mit der Exponentialfunktion  $e^{(\alpha+j\beta)x}$  als rechter Seite:

$$y'' + ay' + by = e^{(\alpha+j\beta)x}. \quad (5.37)$$

Mit Tabelle 5.1 können wir hierfür eine (komplexe) Lösung  $y(x)$  bestimmen. Ihr Realteil  $u(x) = \operatorname{Re} y(x)$  ist dann eine Lösung von (5.36), denn diese Dgl erhält man genau als Realteil der komplexen Dgl (5.37):

$$u'' + au' + bu = \operatorname{Re}(y'' + ay' + by) = \operatorname{Re}(e^{(\alpha+j\beta)x}) = e^{\alpha x} \cos(\beta x).$$

Analog ist der Imaginärteil  $v(x) = \operatorname{Im} y(x)$  eine Lösung von

$$v'' + av' + bv = \operatorname{Im}(e^{(\alpha+j\beta)x}) = e^{\alpha x} \sin(\beta x).$$

**Beispiel 5.14.1** Gesucht ist eine partikuläre Lösung von

$$y'' + y' = e^{-x} \sin x. \quad (5.38)$$

Wegen  $\operatorname{Im}(e^{(-1+j)x}) = e^{-x} \sin x$  ist die zugehörige komplexe Dgl

$$y'' + y' = e^{(-1+j)x}. \quad (5.39)$$

Gemäß Tabelle 5.1 ist ein Ansatz für eine partikuläre Lösung von (5.39)

$$y_p(x) = Ae^{(-1+j)x}.$$

(Es ist  $p(\lambda) = \lambda^2 + \lambda$  und damit  $p(-1+j) \neq 0$ .) Damit erhalten wir

$$\begin{aligned} y_p'(x) &= (-1+j)Ae^{(-1+j)x}, \\ y_p''(x) &= (-1+j)^2 Ae^{(-1+j)x} = -2jAe^{(-1+j)x} \\ \Rightarrow y_p'' + y_p' &= (-2j + (-1+j))Ae^{(-1+j)x} = (-1-j)Ae^{(-1+j)x} \stackrel{!}{=} e^{(-1+j)x} \\ \Leftrightarrow (-1-j)A &= 1 \\ \Leftrightarrow A &= \frac{1}{-1-j} = \frac{-1+j}{(-1-j)(-1+j)} = \frac{-1+j}{2} = -\frac{1}{2} + \frac{1}{2}j \end{aligned}$$

Also ist

$$y_p(x) = \left(-\frac{1}{2} + \frac{1}{2}j\right)e^{(-1+j)x}$$

eine partikuläre Lösung der komplexen Dgl (5.39). Aufgrund von  $\operatorname{Im}(e^{(-1+j)x}) = e^{-x} \sin x$  ist dann der Imaginärteil  $v_p(x) = \operatorname{Im} y_p(x)$  eine partikuläre Lösung der reellen Dgl (5.38). Also ist

$$\begin{aligned} v_p(x) &= \operatorname{Im} y_p(x) = \operatorname{Im} \left( \left(-\frac{1}{2} + \frac{1}{2}j\right)e^{(-1+j)x} \right) \\ &= \operatorname{Im} \left( \left(-\frac{1}{2} + \frac{1}{2}j\right)(e^{-x} \cos x + je^{-x} \sin x) \right) \\ &= \operatorname{Im} \left( -\frac{1}{2}e^{-x} \cos x - \frac{1}{2}je^{-x} \sin x + \frac{1}{2}je^{-x} \cos x - \frac{1}{2}e^{-x} \sin x \right) \\ &= -\frac{1}{2}e^{-x} \sin x + \frac{1}{2}e^{-x} \cos x \end{aligned}$$

eine partikuläre Lösung von (5.38).

## 5.15 Superposition partikulärer Lösungen

Ist die Inhomogenität  $g(x)$  eine Summe von Termen, für die man getrennt partikuläre Lösungen berechnen kann, so erhält man die gesamte partikuläre Lösung für  $g(x)$  als Superposition der einzelnen partikulären Lösungen:

Ist  $y_{p1}(x)$  eine partikuläre Lösung von

$$y'' + ay' + by = g_1(x)$$

und  $y_{p2}(x)$  eine partikuläre Lösung von

$$y'' + ay' + by = g_2(x),$$

dann ist die Superposition  $cy_{p1} + dy_{p2}$  eine partikuläre Lösung von

$$y'' + ay' + by = cg_1(x) + dg_2(x).$$

Das ist genau die Aussage des Superpositionsprinzips von Abschnitt 5.10. Es gilt übrigens genauso für Summen mit drei oder mehr Summanden.

**Beispiel 5.15.1** Wir suchen eine partikuläre Lösung von

$$y'' + y' = -3x + 5e^x. \quad (5.40)$$

Die Inhomogenität ist also eine Summe

$$g(x) = -3x + 5e^x = cg_1(x) + dg_2(x)$$

mit  $c = -3$ ,  $g_1(x) = x$ ,  $d = 5$  und  $g_2(x) = e^x$ . Wir bestimmen daher zunächst partikuläre Lösungen von

$$y'' + y' = g_1(x) = x \quad \text{und} \quad y'' + y' = g_2(x) = e^x.$$

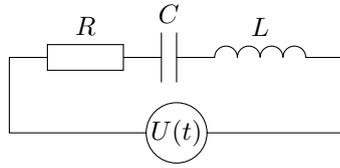


Abbildung 5.6: RCL-Schwingkreis

Es ist  $p(\lambda) = \lambda^2 + \lambda$ . Die Bedingung gemäß Tabelle 5.1 für die Dgl  $y'' + y' = x$  ist  $p(0) = 0$  und daher machen wir den Ansatz  $y_{p1}(x) = x \cdot (Ax + B) = Ax^2 + Bx$ . Einsetzen in die Dgl ergibt nach kurzer Rechnung  $A = \frac{1}{2}$  und  $B = -1$ , also

$$y_{p1}(x) = \frac{1}{2}x^2 - x.$$

Für  $y'' + y' = e^x$  ist die Bedingung  $p(1) \neq 0$  und der Ansatz  $y_{p2}(x) = Ae^x$  führt auf

$$y_{p2}(x) = \frac{1}{2}e^x.$$

(Die vollständigen Rechnungen überlassen wir dem Leser.) Es folgt, dass

$$y_p(x) = -3y_{p1}(x) + 5y_{p2}(x) = -\frac{3}{2}x^2 + 3x + \frac{5}{2}e^x$$

eine partikuläre Lösung von (5.40) ist.

## 5.16 Anwendung: gedämpfte lineare Schwingungen

Eine Anwendungsbeispiel für lineare Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten ist die Gleichung

$$\ddot{y} + 2\delta\dot{y} + \omega_0^2 y = g(t). \quad (5.41)$$

Sie beschreibt lineare gedämpfte Schwingungen. Die Koeffizienten der Gleichung sind physikalisch motiviert:

- $\delta \geq 0$  ist der *Dämpfungsfaktor*;
- $\omega_0 > 0$  ist die *Eigenfrequenz* des ungedämpften Systems (genauer die *Kreisfrequenz*);
- $g(t)$  ist die *Anregung* des Systems.

Gedämpfte Schwingungen treten z.B. in der Mechanik oder bei elektrischen Schaltungen auf. Ein konkretes Beispiel für letzteres ist der RCL-Schwingkreis aus Widerstand  $R$ , Kapazität  $C$  und Induktivität  $L$ , siehe Abbildung 5.6. Ist  $Q(t)$  die Kondensatorladung, so wird die zeitliche Entwicklung des Systems durch folgende Differentialgleichung beschrieben:

$$\begin{aligned} L\ddot{Q} + R\dot{Q} + \frac{1}{C}Q &= U(t) \\ \Leftrightarrow \ddot{Q} + \frac{R}{L}\dot{Q} + \frac{1}{LC}Q &= \frac{1}{L}U(t) \end{aligned}$$

Dies ist also eine lineare gedämpfte Schwingung (5.41) mit

$$\delta = \frac{R}{2L}, \quad \omega_0 = \frac{1}{\sqrt{LC}}, \quad g(t) = \frac{1}{L}U(t).$$

Wir berechnen jetzt die Lösungen von (5.41), zuerst für den freien (homogenen) Fall  $g(t) = 0$ .

### 5.16.1 Freie gedämpfte Schwingungen

Bei freien Schwingungen ist die Anregung  $g(t) = 0$ . Wir haben also die homogene lineare Dgl

$$\ddot{y} + 2\delta\dot{y} + \omega_0^2 y = 0. \quad (5.42)$$

Für die Nullstellen des charakteristischen Polynoms berechnen wir

$$\begin{aligned} p(\lambda) = \lambda^2 + 2\delta\lambda + \omega_0^2 = 0 &\Leftrightarrow (\lambda + \delta)^2 = \delta^2 - \omega_0^2 \\ &\Leftrightarrow \lambda = -\delta \pm \sqrt{\delta^2 - \omega_0^2} \end{aligned}$$

Je nach Vorzeichen des Terms  $\delta^2 - \omega_0^2$  unter der Wurzel erhalten wir verschiedene Lösungen:

- (a) Fall  $\delta > \omega_0$  (starke Dämpfung)

In diesem Fall ist  $\delta^2 - \omega_0^2 > 0$  und damit sind  $\lambda_{1,2} = -\delta \pm \sqrt{\delta^2 - \omega_0^2}$  zwei (verschiedene) reelle Nullstellen von  $p(\lambda)$ . Die allgemeine Lösung von (5.42) ist damit

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}.$$

Da  $\sqrt{\delta^2 - \omega_0^2} < \delta$  ist, ist  $\lambda_{1,2} < 0$ , die Lösung fällt also exponentiell.

- (b)  $\delta = \omega_0$  (kritische Dämpfung)

Hier ist  $\lambda_1 = -\delta$  eine doppelte Nullstelle von  $p(\lambda)$ . Die allgemeine Lösung ist somit

$$y(t) = c_1 e^{-\delta t} + c_2 t e^{-\delta t}.$$

Die Lösung fällt wieder exponentiell. (Beachte: auch  $t e^{-\delta t}$  fällt für große  $t$  exponentiell.)

- (c)  $0 < \delta < \omega_0$  (schwache Dämpfung)

Jetzt ist  $\delta^2 - \omega_0^2 < 0$  und damit  $\lambda_{1,2} = -\delta \pm j\sqrt{\omega_0^2 - \delta^2}$  komplex konjugierte Nullstellen. Die allgemeine Lösung ist somit

$$y(t) = c_1 e^{-\delta t} \cos(\omega_1 t) + c_2 e^{-\delta t} \sin(\omega_1 t)$$

mit  $\omega_1 = \sqrt{\omega_0^2 - \delta^2}$ . Man nennt  $\omega_1$  die Eigenfrequenz des gedämpften Systems.

In Abbildung 5.7 ist der typische Verlauf der Lösung für die drei Fälle skizziert. Die Fälle (a) und (b) mit starker bzw. kritischer Dämpfung zeigen keine Schwingung (sogenannte *aperiodische* Fälle). Im Fall (c) einer schwachen Dämpfung ist die Lösung dagegen eine exponentiell abklingende Schwingung (periodischer Fall).

Schließlich gibt es noch den *Sonderfall*

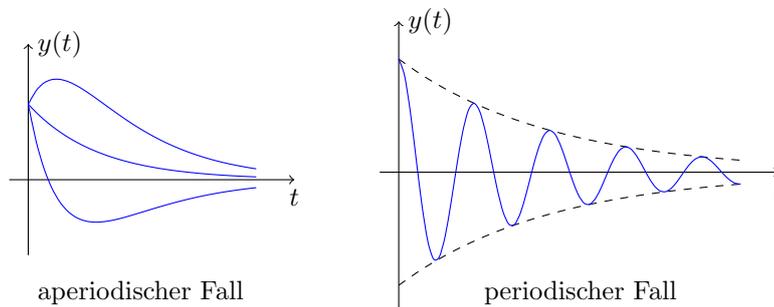


Abbildung 5.7: Freie gedämpfte Schwingung: Lösungen

(d)  $\delta = 0$  (keine Dämpfung)

Dann ist  $\lambda_{1/2} = \pm j\omega_0$  und die Lösung der Dgl ist

$$y(t) = c_1 \cos(\omega_0 t) + c_2 \sin(\omega_0 t).$$

Wir erhalten also Schwingungen in der ungedämpften Eigenfrequenz  $\omega_0$  mit konstanter Amplitude.

### 5.16.2 Erzwungene Schwingungen

Jetzt kommen wir zu gedämpfte Schwingungen mit äußerer Anregung  $g(t) \neq 0$ , sogenannte erzwungene Schwingungen. Wir betrachten hier speziell den Fall einer harmonischen Anregung mit Frequenz  $\omega > 0$ , d.h.

$$\ddot{y} + 2\delta\dot{y} + \omega_0^2 y = A \cos(\omega t). \quad (5.43)$$

Zur Bestimmung einer partikulären Lösung gehen wir zur zugehörigen komplexen Differentialgleichung über:

$$\ddot{y} + 2\delta\dot{y} + \omega_0^2 y = Ae^{j\omega t} \quad (\operatorname{Re}(e^{j\omega t}) = \cos(\omega t))$$

Der Ansatz für die (komplexe) partikuläre Lösung ist

$$y_p(t) = Be^{j\omega t}.$$

Einsetzen in die Dgl ergibt

$$\begin{aligned} \ddot{y}_p + 2\delta\dot{y}_p + \omega_0^2 y_p &= (j\omega)^2 B e^{j\omega t} + 2\delta \cdot j\omega B e^{j\omega t} + \omega_0^2 B e^{j\omega t} \\ &= (-\omega^2 + 2j\delta\omega + \omega_0^2) B e^{j\omega t} \stackrel{!}{=} A e^{j\omega t} \\ \Rightarrow B &= \frac{A}{\zeta} \quad \text{mit} \quad \zeta = \omega_0^2 - \omega^2 + 2\delta\omega j. \end{aligned}$$

Wir schreiben die Konstante  $\zeta$  in Polarform  $\zeta = |\zeta|e^{j\varphi}$ :

$$|\zeta| = \sqrt{(\omega_0^2 - \omega^2)^2 + 4\delta^2\omega^2}.$$

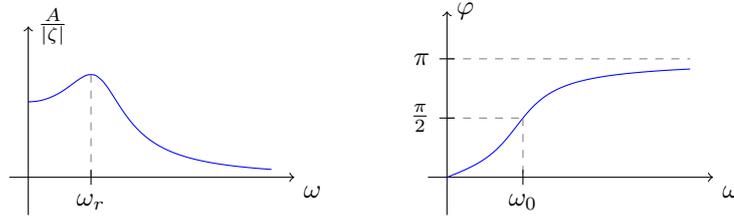


Abbildung 5.8: Erzwungene Schwingung: Amplitude und Phasenverschiebung

Für das Argument  $\varphi$  berechnen wir

$$\begin{aligned} \operatorname{Im} \zeta = 2\delta\omega \geq 0 &\Rightarrow \varphi = \arccos\left(\frac{\operatorname{Re} \zeta}{|\zeta|}\right) \\ &\Rightarrow \varphi = \arccos\left(\frac{\omega_0^2 - \omega^2}{|\zeta|}\right) \end{aligned}$$

Man beachte, dass  $\varphi \in [0, \pi]$  da  $\operatorname{Im} \zeta \geq 0$ . Es folgt nun  $B = \frac{A}{\zeta} = \frac{A}{|\zeta|} e^{-j\varphi}$  und damit ist die komplexe partikuläre Lösung

$$y_p(t) = B e^{j\omega t} = \frac{A}{|\zeta|} e^{-j\varphi} e^{j\omega t} = \frac{A}{|\zeta|} e^{j(\omega t - \varphi)}.$$

Die partikuläre Lösung der reellen Differentialgleichung ist (wegen  $\operatorname{Re}(e^{j\omega t}) = \cos(\omega t)$ ) der Realteil von  $y_p$ , also

$$y_{rp}(t) = \operatorname{Re} y_p(t) = \frac{A}{|\zeta|} \cos(\omega t - \varphi).$$

Die allgemeine Lösung von (5.43) ist somit

$$y(t) = y_h(t) + y_{rp}(t)$$

wobei  $y_h(t)$  die allgemeine homogene Lösung aus 5.16.1 ist.

Bei positiver Dämpfung  $\delta > 0$  gilt für die homogene Lösung  $\lim_{t \rightarrow \infty} y_h(t) = 0$ . Damit ist für große Werte von  $t$

$$y(t) \approx y_{rp}(t) = \frac{A}{|\zeta|} \cos(\omega t - \varphi), \quad (t \text{ groß})$$

d.h. nach einer gewissen Einschwingzeit schwingt das System in der Erregerfrequenz  $\omega$  mit der Phasenverschiebung  $\varphi$ .

Der Verlauf von Amplitude  $\frac{A}{|\zeta|}$  und Phasenverschiebung  $\varphi$  in Abhängigkeit von der Erregerfrequenz  $\omega$  sind in Abbildung 5.8 dargestellt. Die Amplitude hat ein Maximum bei der sogenannten *Resonanzfrequenz*

$$\omega_r = \sqrt{\omega_0^2 - 2\delta^2}.$$

Diese Resonanz des Systems tritt unter der Bedingung  $\delta < \frac{\omega_0}{\sqrt{2}}$  auf, also bei nicht zu starker Dämpfung.

Der Wert der Amplitude im Maximum  $\omega = \omega_r$  wird für kleinere  $\delta$  immer größer. Bei zu kleiner Dämpfung  $\delta$  wird die Amplitude schließlich so groß, dass in einem realen System die Schwingungen das System zerstören: das ist die *Resonanzkatastrophe*. Die physikalische Erklärung ist, dass bei sehr kleinem  $\delta$  die Resonanzfrequenz fast gleich der Eigenfrequenz  $\omega_0$  des ungedämpften Systems ist,  $\omega_r \approx \omega_0$ . Dem System wird dann durch die äußere Anregung mit  $\omega_r$  in optimaler Form Energie zugeführt, die durch die (zu kleine) Dämpfung nicht wieder abgebaut werden kann. Das System erhält so immer mehr Energie von außen bis es schließlich zerstört wird. Beispiele für Resonanzkatastrophen sind etwa Schwingungen von Brücken, ausgelöst von Menschenmengen, die sich in einer synchronen Weise über die Brücke bewegen, siehe z.B. den Fall der Londoner Millennium bridge.

## 5.17 Variation der Konstanten bei inhomogenen Gleichungen 2. Ordnung

Wie bei linearen Differentialgleichungen 1. Ordnung kann man auch im Fall 2. Ordnung für die Lösung der inhomogenen Gleichung

$$y'' + ay' + by = g(x)$$

eine Variation-der-Konstanten Formel herleiten. Wie schon im Fall 1. Ordnung, siehe Abschnitt 5.7, gehen wir von der allgemeinen Lösung  $y(x) = c_1 y_1(x) + c_2 y_2(x)$  der homogenen Gleichung 2. Ordnung aus und variieren darin die Konstanten  $c_1, c_2$ . Unser Ansatz für eine partikuläre Lösung der inhomogenen Differentialgleichung ist also

$$y_p(x) = c_1(x)y_1(x) + c_2(x)y_2(x),$$

wobei  $y_1, y_2$  ein Fundamentalsystem der homogenen Gleichung ist (Satz 5.11.1 oder (5.33)). Zur Bestimmung der unbekanntenen Funktionen  $c_1(x)$  und  $c_2(x)$  benötigt man noch eine Zusatzbedingung<sup>4</sup>; man wählt dafür

$$c_1'(x)y_1(x) + c_2'(x)y_2(x) = 0.$$

(Diese Bedingung ist so gewählt, dass die Rechnung vereinfacht wird.) Einsetzen des Ansatzes in die inhomogene Dgl liefert zusammen mit der Zusatzbedingung ein lineares Gleichungssystem für  $c_1'$  und  $c_2'$ . Daraus erhält man dann (die Details überlassen wir dem Leser)

$$c_1(x) = \int \frac{-y_2(x)g(x)}{W(x)} dx, \quad c_2(x) = \int \frac{y_1(x)g(x)}{W(x)} dx. \quad (5.44)$$

Hierbei ist

$$W(x) = \det \begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix}$$

die sogenannte *Wronski-Determinante*.

<sup>4</sup>Die Differentialgleichung liefert nur *eine* Bedingung für die *zwei* Unbekannten  $c_1(x)$  und  $c_2(x)$ .

## 5.18 Zusammenfassung: Typen von Differentialgleichungen und Lösungsmethoden

Wir fassen hier die bisher kennengelernten Typen von Differentialgleichungen und die zugehörigen Lösungsmethoden zusammen. Wie schon einmal gesagt, ist es für die Lösung entscheidend, zuerst den Typ der Differentialgleichung zu erkennen, damit man dann die korrekte Lösungsmethode anwenden kann.

Typ	Methode
$y' = f(x)g(y)$	$\int \frac{dy}{g(y)} = \int f(x) dx$ Trennen der Variablen
$y' + a(x)y = 0$ (homog. linear 1. Ord.)	$y(x) = ce^{-\int a(x) dx}$ (oder Trennen der Variablen)
$y' + a(x)y = g(x)$ (inhom. linear 1. Ord.)	$y(x) = e^{-A(x)} \cdot (c + \int e^{A(x)} g(x) dx)$ Variation der Konstanten
$y'' + ay' + by = 0$ (homog. linear 2. Ord., konst. Koeff.)	charak. Polynom $p(\lambda) = \lambda^2 + a\lambda + b$ → Fundamentalsystem $y_1, y_2$ $y(x) = c_1 y_1(x) + c_2 y_2(x)$
$y'' + ay' + by = g(x)$ (inhom. linear 2. Ord., konst. Koeff.)	$y(x) = y_h(x) + y_p(x)$ $y_h$ homogene Lsg, $y_p$ partikuläre Lsg, z.B. durch Ansatz

Eventuell muss eine Differentialgleichung erst umgeformt werden, um den Typ anhand der Tabelle ablesen zu können.

**Beispiel 5.18.1** Die Differentialgleichung

$$xy' - x^2y + e^x = 0$$

ist zunächst nicht von einem der Typen aus der Tabelle. Teilen wir aber durch  $x$  und bringen den Term ohne  $y$  auf die rechte Seite, so erhalten wir

$$y' - xy = -\frac{e^x}{x}.$$

Es handelt sich also um eine inhomogene lineare Differentialgleichung 1. Ordnung, die wir mit der Variation-der-Konstanten Formel lösen können.

## 5.19 Substitution

Zum Abschluss des Kapitels über Differentialgleichungen, stellen wir noch kurz zwei weitere Lösungsverfahren vor: Substitution sowie den Potenzreihenansatz.

Bei der Substitution wird die abhängige Variable  $y$  durch eine neue Variable  $u$  ersetzt mit dem Ziel, die Differentialgleichung so zu vereinfachen, dass sie lösbar wird. Die Schwierigkeit ist hier – wie schon bei der Substitutionsregel zur Berechnung von Integralen – zu erkennen, welche Substitution erfolgreich ist. Das hängt natürlich wieder von der Art der Differentialgleichung ab. Wir geben hier einige Beispiele an:

(a) Die Differentialgleichung

$$y' = f(ax + by + c)$$

ist in der Regel (abhängig von der konkreten Funktion  $f$ ) weder linear noch durch Trennen der Variablen lösbar. Da der Term  $ax + by + c$  in  $f$  eingesetzt wird, bietet es sich an, die Substitution  $u = ax + by + c$  zu probieren, also

$$u(x) = ax + by(x) + c.$$

Dann ist  $u'(x) = a + by'(x)$  und einsetzen der Dgl für  $y'$  ergibt

$$u' = a + by' = a + bf(ax + by + c) = a + bf(u).$$

Damit ist die neue Differentialgleichung für  $u = u(x)$  also

$$u' = a + bf(u).$$

Da hier  $x$  nicht mehr explizit vorkommt, ist diese Dgl durch Trennen der Variablen lösbar (jedenfalls prinzipiell).

(b) Für Differentialgleichungen der Form

$$y' = f\left(\frac{y}{x}\right)$$

kann man die Substitution

$$u(x) = \frac{y(x)}{x}$$

verwenden. Auflösen nach  $y$  und ableiten liefert

$$y = xu \quad \Rightarrow \quad y' = u + xu'$$

und mit Einsetzen der Dgl folgt

$$u + xu' = y' = f\left(\frac{y}{x}\right) = f(u).$$

Löst man das nach  $u'$  auf, erhält man die neue Differentialgleichung

$$u' = \frac{1}{x}(f(u) - u),$$

die durch Trennen der Variablen gelöst werden kann.

(c) Die Differentialgleichung zweiter Ordnung

$$y'' = f(x, y')$$

enthält die abhängige Variable  $y$  nicht selbst, sondern nur ihre erste und zweite Ableitung  $y'$  und  $y''$ . Durch die Substitution  $u = y'$  kann man die Dgl auf 1. Ordnung zurückführen: aus  $u = y'$  folgt  $u' = y''$  und damit

$$u' = f(x, u).$$

**Beispiel 5.19.1** Die Differentialgleichung

$$y' = \frac{x}{y} + \frac{y}{x}, \quad x, y > 0, \quad (5.45)$$

ist nicht linear und es ist auch kein Trennen der Variablen möglich. Sie ist aber von der Form in (b), sodass wir die Substitution

$$u = \frac{y}{x}$$

verwenden können. Die rechte Seite der Dgl wird dann zu

$$\frac{x}{y} + \frac{y}{x} = \frac{1}{u} + u = f(u).$$

Somit ist die neue Differentialgleichung

$$u' = \frac{1}{x}(f(u) - u) = \frac{1}{x} \cdot \frac{1}{u}.$$

Trennen der Variablen ergibt

$$\begin{aligned} \frac{du}{dx} = u' &= \frac{1}{x \cdot u} \\ \Rightarrow \int u \, du &= \int \frac{1}{x} \, dx \\ \Leftrightarrow \frac{1}{2}u^2 &= \ln(x) + c \quad (c \in \mathbb{R}) \end{aligned}$$

Wegen der Bedingung  $x > 0$  ist hierbei  $\ln(x)$  (statt  $\ln|x|$ ) die Stammfunktion von  $\frac{1}{x}$ . Weiter folgt aus  $x > 0$  und  $y > 0$  zusammen, dass  $u = \frac{y}{x} > 0$  und damit

$$\begin{aligned} \frac{1}{2}u^2 &= \ln(x) + c \\ \Leftrightarrow u^2 &= 2\ln(x) + 2c \\ \Leftrightarrow u &= \sqrt{2\ln(x) + 2c}. \end{aligned}$$

Damit ist

$$y(x) = xu = x\sqrt{2\ln(x) + 2c}, \quad c \in \mathbb{R},$$

die allgemeine Lösung von (5.45).

## 5.20 Potenzreihenansatz

Bei linearen Differentialgleichungen 2. Ordnung, deren Koeffizienten nicht konstant sind, also etwa

$$y'' + a(x)y' + b(x)y = 0$$

im homogenen Fall, kann man nicht das charakteristische Polynom und Satz 5.11.1 benutzen, um ein Fundamentalsystem zu bestimmen. Stattdessen kann man einen Potenzreihenansatz für die Lösung nutzen. Wir zeigen das am Beispiel der Dgl

$$y'' + xy' + y = 0. \quad (5.46)$$

Hier ist der Koeffizienten  $x$  vor  $y'$  nicht konstant. Wir machen für die Lösung  $y(x)$  den Ansatz einer Potenzreihe mit Entwicklungspunkt  $x_0 = 0$ ,

$$y(x) = \sum_{k=0}^{\infty} a_k x^k. \quad (5.47)$$

Durch Einsetzen in (5.46) bestimmen wir jetzt die Konstanten  $a_k$  der Potenzreihe:

$$\begin{aligned} y''(x) + xy'(x) + y(x) &= \sum_{k=2}^{\infty} k(k-1)a_k x^{k-2} + x \sum_{k=1}^{\infty} k a_k x^{k-1} + \sum_{k=0}^{\infty} a_k x^k \\ &= \sum_{k=0}^{\infty} (k+2)(k+1)a_{k+2} x^k + \sum_{k=1}^{\infty} k a_k x^k + \sum_{k=0}^{\infty} a_k x^k \\ &= \sum_{k=0}^{\infty} ((k+2)(k+1)a_{k+2} + k a_k + a_k) x^k \stackrel{!}{=} 0 \quad \text{für alle } x \end{aligned}$$

Das Ergebnis von  $y'' + xy' + y$  in der letzten Zeile ist wieder eine Potenzreihe; damit diese konstant gleich 0 ist, müssen alle Koeffizienten 0 sein:

$$\begin{aligned} (k+2)(k+1)a_{k+2} + (k+1)a_k &= 0 \quad \text{für alle } k \geq 0 \\ \Rightarrow a_{k+2} &= -\frac{1}{k+2} a_k, \quad k \geq 0. \end{aligned}$$

Daraus erhalten wir rekursiv

$$\begin{aligned} a_2 &= -\frac{1}{2} a_0 & a_3 &= -\frac{1}{3} a_1 \\ a_4 &= -\frac{1}{4} a_2 = \frac{1}{2 \cdot 4} a_0 & a_5 &= -\frac{1}{5} a_3 = \frac{1}{3 \cdot 5} a_1 \\ \vdots & & \vdots & \\ a_{2k} &= \frac{(-1)^k}{2 \cdot 4 \cdot \dots \cdot (2k)} a_0 & a_{2k+1} &= \frac{(-1)^k}{3 \cdot 5 \cdot \dots \cdot (2k+1)} a_1 \end{aligned}$$

Wir erhalten also Formeln für die  $a_k$  ab  $k = 2$ , wobei sich unterschiedliche allgemeine Formeln für die geraden und die ungeraden Koeffizienten  $a_{2k}$  bzw.  $a_{2k+1}$  ergeben. Speziell sind alle  $a_{2k}$  Vielfache von  $a_0$ , alle  $a_{2k+1}$  dagegen Vielfache von  $a_1$ . Für  $a_0$  und  $a_1$  selbst gibt es dagegen keine weiteren Bedingungen; es können beliebige reelle Zahlen sein. Durch Einsetzen dieser Ergebnisse in den Ansatz (5.47) erhalten wir

$$y(x) = a_0 \cdot \sum_{k=0}^{\infty} \frac{(-1)^k}{2 \cdot 4 \cdot \dots \cdot (2k)} x^{2k} + a_1 \cdot \sum_{k=0}^{\infty} \frac{(-1)^k}{3 \cdot 5 \cdot \dots \cdot (2k+1)} x^{2k+1}$$

wobei wir die Reihe in die geraden und ungeraden Terme aufgeteilt und  $a_0$  bzw.  $a_1$  ausgeklammert haben. Setzen wir jetzt als Abkürzung

$$y_1(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2 \cdot 4 \cdot \dots \cdot (2k)} x^{2k}, \quad y_2(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{3 \cdot 5 \cdot \dots \cdot (2k+1)} x^{2k+1},$$

so folgt

$$y(x) = a_0 y_1(x) + a_1 y_2(x), \quad a_0, a_1 \in \mathbb{R}.$$

Dies ist die allgemeine Lösung von (5.46); sie enthält die freien Parameter  $a_0, a_1$ . Wie in Satz 5.11.1 ist die allgemeine Lösung also eine Linearkombination eines Fundamentalsystems  $y_1, y_2$ . Das Fundamentalsystem ist hier aber von einer anderen Form als in Satz 5.11.1 oder Gleichung (5.33).

## Kapitel 6

# Eigenwerte und Eigenvektoren

In diesem und dem nächsten Kapitel verlassen wir kurz die Analysis – also die Untersuchung von Folgen, Reihen und Funktionen – und beschäftigen uns, gewissermaßen als Intermezzo, mit Matrizen, genauer mit quadratischen Matrizen. Wir bestimmen Eigenwerte und Eigenvektoren, die die interne Struktur einer quadratischen Matrix festlegen. Für viele Anwendungen ist das von großer Bedeutung.

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik A, Kapitel 8.2, 8.3.

### 6.1 Einführung

Sei  $A$  eine quadratische, reelle oder komplexe,  $n \times n$ -Matrix, kurz  $A \in \mathbb{K}^{n \times n}$ . Hier ist  $\mathbb{K} = \mathbb{R}$  im Fall einer reellen Matrix oder  $\mathbb{K} = \mathbb{C}$  bei einer komplexen Matrix, d.h. einer Matrix mit komplexen Einträgen. Die Matrix  $A$  kann mit einem Vektor  $x \in \mathbb{K}^n$ , also mit einem Vektor aus  $n$  Komponenten, multipliziert werden. Das Ergebnis ist wieder ein Vektor mit  $n$  Komponenten, d.h.  $Ax \in \mathbb{K}^n$ . Eine Matrix definiert auf diese Weise eine Transformation von Vektoren: aus dem gegebenen Vektor  $x$  wird der neue Vektor  $Ax$ .

Da jeder Vektor durch seine Länge und Richtung eindeutig bestimmt ist, verändert die Multiplikation mit der Matrix  $A$  im Allgemeinen Länge und Richtung des Vektors. Es gibt nun zu jeder Matrix gewisse Vektoren, bei der die Multiplikation nur die Länge, nicht aber die Richtung verändert. Wird bei einem Vektor nur die Länge verändert, entspricht das einer Multiplikation des Vektors mit einem Skalar (also einer Zahl)  $\lambda \in \mathbb{K}$ . Für Vektoren, bei denen die Matrixmultiplikation nur die Länge und nicht die Richtung ändert, gilt also  $Ax = \lambda x$ . Man nennt diese speziellen Vektoren *Eigenvektoren* der Matrix  $A$ . Die Zahlen  $\lambda$  nennt man *Eigenwerte*.

**Definition 6.1.1** Sei  $A \in \mathbb{K}^{n \times n}$ . Ein Vektor  $x \in \mathbb{C}^n$  ( $x \neq 0$ ) heißt Eigenvektor von  $A$  zum Eigenwert  $\lambda \in \mathbb{C}$  wenn gilt

$$Ax = \lambda x. \tag{6.1}$$

Der Vektor  $x = 0$  erfüllt (6.1) für jede Matrix  $A$  und jedes  $\lambda$ . Diese triviale Lösung ist aber nicht interessant (der Vektor  $x = 0$  legt ja auch keine Richtung fest), daher sind Eigenvektoren per Definition immer ungleich Null.

**Beispiel 6.1.2** Wir betrachten die reelle Matrix

$$A = \begin{pmatrix} 2 & 1 & 1 \\ -4 & -3 & -1 \\ 1 & 1 & 2 \end{pmatrix}$$

(also  $A \in \mathbb{R}^{3 \times 3}$ ). Für den Vektor

$$x = \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix}$$

berechnen wir das Matrix-Vektor-Produkt  $Ax$ :

$$Ax = \begin{pmatrix} 2 \cdot 1 + 1 \cdot (-5) + 1 \cdot 1 \\ -4 \cdot 1 - 3 \cdot (-5) - 1 \cdot 1 \\ 1 \cdot 1 + 1 \cdot (-5) + 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 2 - 5 + 1 \\ -4 + 15 - 1 \\ 1 - 5 + 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 10 \\ -2 \end{pmatrix}.$$

Wir sehen, dass der Ergebnisvektor eine Vielfaches des ursprünglichen Vektors  $x$  ist, nämlich das  $-2$ -fache:

$$Ax = \begin{pmatrix} -2 \\ 10 \\ -2 \end{pmatrix} = -2 \cdot \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix} = -2x.$$

Für diesen speziellen Vektor gilt also  $Ax = -2x$ . D.h.  $x$  ist ein Eigenvektor zum Eigenwert  $\lambda = -2$ . Jetzt probieren wir einen anderen Vektor, etwa

$$y = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Hierfür ist

$$Ay = \begin{pmatrix} 2 & 1 & 1 \\ -4 & -3 & -1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 - 1 \\ -3 + 1 \\ 1 - 2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix}.$$

Das ist offenbar kein Vielfaches des Ausgangsvektors  $y$ , also

$$Ay = \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix} \neq \lambda \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \lambda y \quad \text{für jedes } \lambda \in \mathbb{C}.$$

Denn sonst müsste für die zweite Komponente gelten  $-2 = \lambda \cdot 1 \Rightarrow \lambda = -2$ , für die dritte Komponente aber  $-1 = \lambda \cdot (-1) \Rightarrow \lambda = 1$ , was ein Widerspruch ist. Da also  $Ay \neq \lambda y$  für jedes  $\lambda \in \mathbb{C}$  gilt, ist  $y$  kein Eigenvektor von  $A$ .

Tatsächlich sind für eine gegebene Matrix die allermeisten Vektoren keine Eigenvektoren. Wir werden sehen, dass eine  $n \times n$ -Matrix typischerweise  $n$  wesentlich verschiedene Eigenvektoren besitzt (nämlich maximal  $n$  linear unabhängige Eigenvektoren).

**Beachte:**

- (a) Eigenvektoren sind in folgendem Sinn nicht eindeutig: Ist  $x$  ein Eigenvektor zum Eigenwert  $\lambda$ ,  $Ax = \lambda x$ , dann ist auch  $rx$  für jeden Skalar  $r \in \mathbb{C} \setminus \{0\}$  ein Eigenvektor zum gleichen Eigenwert, denn  $A(rx) = rAx = \lambda rx$ . Im Beispiel oben etwa ist auch  $v = -x$  ein Eigenvektor:

$$Av = \begin{pmatrix} 2 & 1 & 1 \\ -4 & -3 & -1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 5 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -10 \\ 2 \end{pmatrix} = -2 \begin{pmatrix} -1 \\ 5 \\ -1 \end{pmatrix} = -2v$$

- (b) Reelle Matrizen können auch komplexe Eigenvektoren und Eigenwerte haben.

**Anwendungen.** Wie schon erwähnt spielen Eigenwerte und Eigenvektoren eine wichtige Rolle in vielen Anwendungen, z.B.:

- (Eigen-)Schwingungen komplexer, d.h. zusammengesetzter Systeme (etwa gekoppelte mechanische Systeme oder elektrische Schaltungsnetzwerke). Mathematisch wird das durch Systeme von Differentialgleichungen beschrieben, das ist ein Thema in Mathematik C.
- Quantenmechanik (Eigenzustände, Messwerte)
- Drehung starrer Körper (Trägheitstensor, Hauptträgheitsachsen)

## 6.2 Berechnung von Eigenwerten und -vektoren: das charakteristische Polynom einer Matrix

Wie können wir Eigenvektoren einer Matrix berechnen? Dazu formen wir die Eigenvektorgleichung (6.1) um. Sei  $E$  die  $n \times n$  Einheitsmatrix, also

$$E = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix},$$

und für jeden Vektor  $x \in \mathbb{K}^n$  gilt  $Ex = x$ . Damit gilt

$$\begin{aligned} Ax = \lambda x &\Leftrightarrow Ax = \lambda Ex \\ &\Leftrightarrow (A - \lambda E)x = 0. \end{aligned}$$

Die letzte Gleichung ist ein homogenes lineares Gleichungssystem mit der Koeffizientenmatrix  $A - \lambda E$  und dem gesuchten Vektor  $x$ . Da Eigenvektoren immer ungleich Null sind, suchen wir also Lösungen  $x \neq 0$  eines homogenen Gleichungssystems. Aus Mathematik A wissen wir, dass ein quadratisches homogenes Gleichungssystem genau dann nicht-triviale Lösungen  $x \neq 0$  hat, wenn die Determinante der Matrix Null ist. Somit gilt

$$\begin{aligned} \lambda \text{ ist Eigenwert} &\Leftrightarrow \text{es gibt } x \neq 0 \text{ mit } Ax = \lambda x \\ &\Leftrightarrow (A - \lambda E)x = 0 \text{ hat eine Lösung } x \neq 0 \\ &\Leftrightarrow \det(A - \lambda E) = 0 \end{aligned}$$

Damit haben wir jetzt Gleichungen, mit denen wir Eigenwerte und Eigenvektoren einer Matrix berechnen können:

**Berechnung Eigenwerte und Eigenvektoren.**

- (a)  $\lambda$  ist Eigenwert der Matrix  $A \Leftrightarrow \det(A - \lambda E) = 0$   
 (b)  $x$  ist Eigenvektor zum Eigenwert  $\lambda$   
 $\Leftrightarrow x \neq 0$  ist Lösung des Gleichungssystems  $(A - \lambda E)x = 0$

Es zeigt sich, dass  $\det(A - \lambda E)$  ein Polynom in der Variablen  $\lambda$  ist. Die Nullstellen dieses Polynoms sind also die Eigenwerte von  $A$ .

**Definition 6.2.1** Das *charakteristische Polynom* der Matrix  $A \in \mathbb{K}^{n \times n}$  ist

$$p_A(\lambda) = \det(A - \lambda E).$$

**Satz 6.2.2** Sei  $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ . Dann ist  $p_A$  ein Polynom

$$p_A(\lambda) = c_n \lambda^n + c_{n-1} \lambda^{n-1} + \dots + c_1 \lambda + c_0$$

wobei

$$c_n = (-1)^n, \quad c_{n-1} = (-1)^{n-1} \text{Spur } A, \quad c_0 = \det A.$$

Die *Spur* einer Matrix  $A$  ist Summe ihrer Diagonaleinträge,

$$\text{Spur } A = a_{11} + a_{22} + \dots + a_{nn}.$$

**Beweis.** Wir rechnen die Aussage des Satzes für den  $2 \times 2$  Fall nach. Sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Dann

$$\begin{aligned} p_A(\lambda) &= \det(A - \underbrace{\lambda E}_{\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}}) = \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \\ &= (a - \lambda)(d - \lambda) - bc = \lambda^2 - a\lambda - d\lambda + ad - bc \\ &= \lambda^2 - \underbrace{(a + d)}_{\text{Spur } A} \lambda + \underbrace{ad - bc}_{\det A} = c_2 \lambda^2 + c_1 \lambda + c_0 \end{aligned}$$

mit

$$c_2 = 1 = (-1)^2, \quad c_1 = (-1) \cdot \text{Spur } A, \quad c_0 = \det A.$$

□

**Beispiel 6.2.3** Wir berechnen das charakteristische Polynom von

$$A = \begin{pmatrix} 2 & 1 \\ -2 & 7 \end{pmatrix}.$$

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda E) = \det \begin{pmatrix} 2 - \lambda & 1 \\ -2 & 7 - \lambda \end{pmatrix} \\ &= (2 - \lambda)(7 - \lambda) - (-2) \cdot 1 = \lambda^2 - 7\lambda - 2\lambda + 14 + 2 \\ &= \lambda^2 - 9\lambda + 16 \end{aligned}$$

Als Probe können wir überprüfen, dass

$$\text{Spur } A = 2 + 7 = 9, \quad \det A = 2 \cdot 7 - (-2) \cdot 1 = 16,$$

was mit unserem Ergebnis und der allgemeinen Formel aus Satz 6.2.2 übereinstimmt.

**Beispiel 6.2.4** Wir berechnen Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 1 & -3 \\ 10 & -10 \end{pmatrix}.$$

Zuerst das **charakteristische Polynom**:

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} 1-\lambda & -3 \\ 10 & -10-\lambda \end{pmatrix} = (1-\lambda)(-10-\lambda) + 30 \\ &= \lambda^2 + 10\lambda - \lambda - 10 + 30 = \lambda^2 + 9\lambda + 20 \end{aligned}$$

Als Probe können wir wieder verifizieren, dass die Koeffizienten des Polynoms die Spur (bis auf das Vorzeichen) und die Determinante der Matrix sind: Spur  $A = 1 - 10 = -9$ ,  $\det A = -10 + 30 = 20$ . Wir berechnen jetzt die Eigenwerte als Nullstellen des charakteristischen Polynoms:

$$\begin{aligned} p_A(\lambda) = 0 &\Leftrightarrow \lambda^2 + 9\lambda + 20 = 0 \\ &\Leftrightarrow \left(\lambda + \frac{9}{2}\right)^2 = \frac{81}{4} - 20 = \frac{1}{4} \\ &\Leftrightarrow \lambda = -\frac{9}{2} \pm \frac{1}{2} \end{aligned}$$

Also sind die **Eigenwerte von  $A$**

$$\lambda_1 = -\frac{9}{2} + \frac{1}{2} = -4, \quad \lambda_2 = -\frac{9}{2} - \frac{1}{2} = -5.$$

Jetzt berechnen wir zu jedem Eigenwert einen zugehörigen Eigenvektor.

Zuerst der **Eigenvektor zu  $\lambda_1 = -4$** : Wir suchen einen Vektor  $x \neq 0$  mit  $(A - \lambda_1 E)x = 0$ . Es ist

$$A - \lambda_1 E = A + 4E = \begin{pmatrix} 1 & -3 \\ 10 & -10 \end{pmatrix} + \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 5 & -3 \\ 10 & -6 \end{pmatrix}.$$

Wir suchen also eine Lösung  $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq 0$  des Gleichungssystems

$$\begin{pmatrix} 5 & -3 \\ 10 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir berechnen die Lösung mit dem Gaußverfahren (Mathematik A):

$$\left( \begin{array}{cc|c} 5 & -3 & 0 \\ 10 & -6 & 0 \end{array} \right) \xrightarrow{-2I} \left( \begin{array}{cc|c} 5 & -3 & 0 \\ 0 & 0 & 0 \end{array} \right) \quad (6.2)$$

Der Rang ist 1, also gibt es  $2 - 1 = 1$  freien Parameter. Wir wählen  $x_2 = s$  als freien Parameter. Die erste Zeile ergibt dann

$$5x_1 - 3x_2 = 0 \quad \Leftrightarrow \quad x_1 = \frac{3}{5}s.$$

Die (allgemeine) Lösung des Gleichungssystems ist damit

$$x = \begin{pmatrix} \frac{3}{5}s \\ s \end{pmatrix} = s \begin{pmatrix} \frac{3}{5} \\ 1 \end{pmatrix};$$

für jedes  $s \neq 0$  ist dies ein Eigenvektor zum Eigenwert  $\lambda_1 = -4$ . (Das ist die Nicht-Eindeutigkeit von Eigenvektoren, die wir in Abschnitt 6.1 erwähnt haben.) Wir wählen  $s$  so, dass sich ein möglichst einfacher Eigenvektor ergibt. Mit  $s = 5$  erhalten wir:

$$x = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } \lambda_1 = -4.$$

Dieses Ergebnis können wir auch etwas schneller wie folgt erhalten: Nach der Zeilenumformung (6.2) suchen wir einen Vektor  $x \neq 0$  mit

$$(A - \lambda_1 E)x = 0 \Leftrightarrow \begin{pmatrix} 5 & -3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Hieraus kann man den Eigenvektor direkt ablesen. Der Vektor soll multipliziert mit der Matrix  $\begin{pmatrix} 5 & -3 \\ 0 & 0 \end{pmatrix}$  Null ergeben. Offenbar gilt

$$\begin{pmatrix} 5 & -3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

und somit ist  $x = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$  der gesuchte Eigenvektor. (Der Vektor ergibt sich durch Vertauschen der Zahlen in der ersten Zeile der Matrix und ändern eines Vorzeichens; das funktioniert offenbar für beliebige Zahlen.)

Jetzt der **Eigenvektor zu  $\lambda_2 = -5$** : Es gilt

$$A - \lambda_2 E = A + 5E = \begin{pmatrix} 6 & -3 \\ 10 & -5 \end{pmatrix}.$$

Zeilenumformung ergibt

$$\left( \begin{array}{cc|c} 6 & -3 & 0 \\ 10 & -5 & 0 \end{array} \right) \cdot \frac{1}{3} \Leftrightarrow \left( \begin{array}{cc|c} 2 & -1 & 0 \\ 10 & -5 & 0 \end{array} \right) - 5I \Leftrightarrow \left( \begin{array}{cc|c} 2 & -1 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

Damit können wir den Eigenvektor wieder direkt ablesen: Es gilt

$$\begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

also ist

$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ der Eigenvektor zum Eigenwert } \lambda_2 = -5.$$

### 6.3 Algebraische und geometrische Vielfachheiten

Wir wissen, dass die Eigenwerte einer Matrix die Nullstellen des charakteristischen Polynoms sind. Andererseits kennen wir aus Mathematik A die vollständige Faktorisierung eines Polynoms  $p(x)$ , d.h. die Zerlegung in die Linearfaktoren

$x - x_i$  zu den Nullstellen  $x_i$  des Polynoms. Für eine mehrfache Nullstelle tritt der Linearfaktor dabei entsprechend oft auf:

$$p(x) = a_n(x - x_1)^{k_1} \dots (x - x_r)^{k_r}.$$

Die Nullstelle  $x_1$  hat die Vielfachheit  $k_1$ , u.s.w. Der Grad des Polynoms ist die Summe der Vielfachheiten  $n = k_1 + \dots + k_l$ . Ein Polynom  $n$ . Grades hat maximal  $n$  verschiedene Nullstellen (nämlich wenn alle Vielfachheiten  $k_i = 1$  sind). Wenden wir das auf das charakteristische Polynom an, so ergibt sich für die Eigenwerte:

**Satz 6.3.1** Sei  $A \in \mathbb{K}^{n \times n}$  und  $p_A(\lambda) = \det(A - \lambda E)$  das charakteristische Polynom.

- (a)  $\lambda$  ist Eigenwert von  $A \Leftrightarrow p_A(\lambda) = 0$ .
- (b)  $A$  hat maximal  $n$  verschiedene Eigenwerte.
- (c) Die vollständige Faktorisierung von  $p_A$  ist

$$p_A(\lambda) = (-1)^n (\lambda - \lambda_1)^{k_1} \dots (\lambda - \lambda_r)^{k_r};$$

dabei sind  $\lambda_1, \dots, \lambda_r$  die verschiedenen Eigenwerte von  $A$ .

$k_j$  heißt algebraische Vielfachheit des Eigenwerts  $\lambda_j$ . Es gilt:

$$\begin{aligned} k_1 + \dots + k_r &= n \\ \text{Spur } A &= k_1 \lambda_1 + \dots + k_r \lambda_r \\ \det A &= \lambda_1^{k_1} \dots \lambda_r^{k_r} \end{aligned}$$

**Beweis.** Die Formeln für Spur und Determinante ergeben sich durch Ausmultiplizieren der vollständigen Faktorisierung und Koeffizientenvergleich mit der Formel für  $p_A(\lambda)$  in Satz 6.2.2.  $\square$

Satz 6.3.1 beschreibt die Eigenwerte einer Matrix. Jetzt schauen wir uns die Struktur der Eigenvektoren an.

**Definition 6.3.2** Sei  $\lambda$  ein Eigenwert von  $A$ . Die Menge aller Eigenvektoren zu  $\lambda$

$$V(\lambda) = \{x \in \mathbb{C}^n \mid Ax = \lambda x\}$$

heißt *Eigenraum* zum Eigenwert  $\lambda$ .

Da die Eigenvektoren genau die Lösungen des homogenen Gleichungssystems  $(A - \lambda E)x = 0$  sind (siehe voriger Abschnitt), und die Lösungsmenge eines homogenen Gleichungssystems der Kern der Koeffizientenmatrix (siehe Mathematik A) ist, gilt

$$V(\lambda) = \{x \in \mathbb{C}^n \mid (A - \lambda E)x = 0\} = \text{Kern}(A - \lambda E).$$

Der Eigenraum ist als Kern ein Untervektorraum und hat damit eine Dimension. Diese Dimension

$$m = \dim V(\lambda) \tag{6.3}$$

heißt *geometrische Vielfachheit* des Eigenwerts  $\lambda$ .

**Satz 6.3.3** Sei  $k$  die algebraische und  $m = \dim V(\lambda)$  die geometrische Vielfachheit eines Eigenwerts  $\lambda$ . Es gilt dann:

(a)  $m \leq k$

(b)  $m =$  Anzahl linear unabhängiger Eigenvektoren zu  $\lambda$

**Beweis.** Teil (b) ist klar, da die Dimension eines Untervektorraums gleich der Anzahl der Vektoren in einer Basis ist; und die ist gleich der maximalen Anzahl linear unabhängiger Vektoren des Untervektorraums (Mathematik A). Teil (a) ist ein tiefergehendes Resultat, das wir hier nicht beweisen.  $\square$

Der letzte Satz ist wichtig für die Berechnung der Eigenvektoren: er sagt uns, wieviele wesentlich verschiedene (nämlich linear unabhängige) Eigenvektoren es zu einem Eigenwert gibt. Da der Eigenraum  $V(\lambda)$  ein Untervektorraum ist, enthält er unendlich viele Eigenvektoren. Aus Mathematik A wissen wir aber, dass sich jeder Vektor eines Unterraums als Linearkombination einer Basis schreiben lässt. Daher suchen wir zu jedem Eigenwert  $\lambda$  eine Basis des Eigenraums  $V(\lambda)$ , d.h.  $m$  linear unabhängige Eigenvektoren zu  $\lambda$ .

**Bemerkung:** Für ein lineares homogenes Gleichungssystem mit  $n$  Unbekannten ist die Dimension des Kerns gleich der Differenz  $n$  minus dem Rang des Gleichungssystems, was wiederum die Anzahl der freien Parameter der Lösung im Gauß-Algorithmus ist. Für die geometrische Vielfachheit  $m = \dim V(\lambda)$  gilt damit

$$\begin{aligned} m &= \dim \text{Kern}(A - \lambda E) \\ &= n - \text{Rang}(A - \lambda E) \\ &= \text{Anzahl freie Parameter des homog. Gl.sys. } (A - \lambda E)x = 0. \end{aligned}$$

Mit Satz 6.3.3 heißt das, dass die Anzahl linear unabhängiger Eigenvektoren also genau gleich der Anzahl freier Parameter im Gleichungssystem  $(A - \lambda E)x = 0$  ist.

**Beispiel 6.3.4** Wir berechnen Eigenwerte und Eigenvektoren von

$$A = \begin{pmatrix} 4 & 4 & 1 \\ -2 & -2 & -1 \\ 2 & 4 & 3 \end{pmatrix}.$$

Zuerst wieder das charakteristische Polynom. Die Determinante der  $3 \times 3$ -Matrix berechnen wir mit der Sarrus-Regel aus Mathematik A:

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda E) = \det \begin{pmatrix} 4 - \lambda & 4 & 1 \\ -2 & -2 - \lambda & -1 \\ 2 & 4 & 3 - \lambda \end{pmatrix} \\ &= (4 - \lambda)(-2 - \lambda)(3 - \lambda) - 8 - 8 - 2(-2 - \lambda) + 4(4 - \lambda) + 8(3 - \lambda) \\ &= (\lambda^2 - 2\lambda - 8)(3 - \lambda) + 28 - 10\lambda = -\lambda^3 + 5\lambda^2 - 8\lambda + 4 \end{aligned}$$

Eine Nullstelle des Polynoms ist  $\lambda_1 = 1$ . Polynomdivision mit dem Linearfaktor  $\lambda - 1$  liefert

$$-\lambda^3 + 5\lambda^2 - 8\lambda + 4 = (\lambda - 1)(-\lambda^2 + 4\lambda - 4).$$

Für die weiteren Nullstellen ergibt sich damit

$$-\lambda^2 + 4\lambda - 4 = 0 \Leftrightarrow \lambda^2 - 4\lambda + 4 = 0 \Leftrightarrow (\lambda - 2)^2 = 0 \Leftrightarrow \lambda = 2.$$

Diese quadratische Gleichung hat also nur die eine (doppelte) Nullstelle  $\lambda_2 = 2$ . Die Faktorisierung von  $p_A$  ist damit

$$p_A(\lambda) = -(\lambda - 1)(\lambda - 2)^2 \quad (6.4)$$

(Für die doppelte Nullstelle 2 kommt der Linearfaktor quadratisch vor,  $(\lambda - 2)^2$ , die Nullstelle 1 ist einfach,  $(\lambda - 1)^1$ .) Die Matrix  $A$  hat somit die Eigenwerte

$$\begin{aligned} \lambda_1 = 1 & \text{ mit algebraischer Vielfachheit } 1, \\ \lambda_2 = 2 & \text{ mit algebraischer Vielfachheit } 2. \end{aligned}$$

Mit Satz 6.3.3 folgt dann

$$\begin{aligned} \lambda_1 \text{ hat geometrische Vielfachheit } m \leq 1 & \Rightarrow m = 1, \\ \lambda_2 \text{ hat geometrische Vielfachheit } m \leq 2 & \Rightarrow m = 1 \vee m = 2. \end{aligned}$$

(Die geometrische Vielfachheit eines Eigenwertes ist immer  $\geq 1$ .)

**Eigenvektor zu  $\lambda_1 = 1$ :** Da für  $\lambda_1$  die geometrische Vielfachheit  $m = 1$  ist, wissen wir schon, dass wir genau *einen* Eigenvektor zu  $\lambda_1$  suchen. Wir lösen dazu wieder das Gleichungssystem  $(A - \lambda E)x = 0$  mit  $\lambda = \lambda_1 = 1$ :

$$\begin{aligned} A - \lambda_1 E = A - E &= \begin{pmatrix} 3 & 4 & 1 \\ -2 & -3 & -1 \\ 2 & 4 & 2 \end{pmatrix} \begin{array}{l} \leftrightarrow \text{III} \\ \\ \cdot \frac{1}{2}, \leftrightarrow \text{I} \end{array} \Leftrightarrow \begin{pmatrix} 1 & 2 & 1 \\ -2 & -3 & -1 \\ 3 & 4 & 1 \end{pmatrix} \begin{array}{l} \\ +2\text{I} \\ -3\text{I} \end{array} \\ &\Leftrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix} \begin{array}{l} \\ \\ +2\text{II} \end{array} \Leftrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Der Rang ist also 2 und es gibt  $3 - 2 = 1$  freie Parameter. (Beachte:  $m = 1 =$  Anzahl freie Parameter!) Wir wählen als freien Parameter  $x_3 = t$ . Rückwärtseinsetzen ergibt dann

$$\begin{aligned} 2. \text{ Zeile: } x_2 + x_3 = 0 &\Rightarrow x_2 = -x_3 = -t \\ 1. \text{ Zeile: } x_1 + 2x_2 + x_3 = 0 &\Rightarrow x_1 = -2x_2 - x_3 = t \\ \Rightarrow x &= \begin{pmatrix} t \\ -t \\ t \end{pmatrix}. \end{aligned}$$

Wie in Beispiel 6.2.4 wählen wir einen möglichst einfachen Eigenvektor: Für  $t = 1$  erhalten wir

$$x = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } \lambda_1 = 1.$$

**Eigenvektoren zu  $\lambda_2 = 2$ :** Hier ist die geometrische Vielfachheit  $m \leq 2$ , es kann also einen ( $m = 1$ ) oder zwei ( $m = 2$ ) linear unabhängige Eigenvektoren

zu  $\lambda_2$  geben. Der Gaußalgorithmus für das Gleichungssystem  $(A - \lambda_2 E)x = 0$  liefert:

$$A - \lambda_2 E = A - 2E = \begin{pmatrix} 2 & 4 & 1 \\ -2 & -4 & -1 \\ 2 & 4 & 1 \end{pmatrix} + I \Leftrightarrow \begin{pmatrix} 2 & 4 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Der Rang ist 1, also gibt es  $3 - 1 = 2$  freie Parameter. Damit ist auch

$$m = \dim \text{Kern}(A - 2E) = n - \text{Rang}(A - 2E) = 3 - 1 = 2,$$

wir suchen also zwei linear unabhängige Eigenvektoren. Wir wählen als Parameter  $x_2 = s$ ,  $x_3 = t$ . Rückwärtseinsetzen liefert

$$\begin{aligned} 1. \text{ Zeile: } 2x_1 + 4x_2 + x_3 = 0 &\Rightarrow 2x_1 = -4s - t \Rightarrow x_1 = -2s - \frac{1}{2}t \\ \Rightarrow x &= \begin{pmatrix} -2s - \frac{1}{2}t \\ s \\ t \end{pmatrix} \end{aligned}$$

Die zwei gesuchten linear unabhängigen Eigenvektoren erhalten wir nun, indem wir immer genau einen der beiden Parameter ungleich Null wählen:

$$\begin{aligned} s = 1, t = 0 &\Rightarrow x = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \\ s = 0, t = 2 &\Rightarrow x = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \end{aligned}$$

Also sind

$$v_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$$

linear unabhängige Eigenvektoren zum Eigenwert  $\lambda_2 = 2$ .

(Beachte:  $(v_1, v_2)$  linear unabhängig denn  $v_1 \neq rv_2$ .)

## 6.4 Diagonalisierung

Wir haben gesehen, dass man zu jedem Eigenwert eine bestimmte Anzahl linear unabhängiger Eigenvektoren bestimmen kann, nämlich genau so viele, wie die geometrische Vielfachheit angibt. Diese Vektoren sind dann jeweils eine Basis des Eigenraumes zum jeweiligen Eigenwert. Fasst man nun die Eigenvektoren für *alle* Eigenwerte zusammen, erhält man im optimalen Fall eine Basis des *ganzen* Raumes.

**Lemma 6.4.1** *Sind  $v_1, \dots, v_r$  Eigenvektoren zu verschiedenen Eigenwerten  $\lambda_i$  von  $A$ , dann ist  $(v_1, \dots, v_r)$  linear unabhängig.*

**Beweis.** Wir zeigen den Fall  $r = 2$ : Angenommen,  $(v_1, v_2)$  wäre linear abhängig. Dann wäre

$$v_1 = \alpha v_2 \Rightarrow Av_1 = \alpha Av_2 \Rightarrow \lambda_1 v_1 = \alpha \lambda_2 v_2 = \lambda_2 v_1 \Rightarrow (\lambda_1 - \lambda_2)v_1 = 0,$$

was wegen  $\lambda_1 \neq \lambda_2$  und  $v_1 \neq 0$  ein Widerspruch ist. Den Fall  $r > 2$  kann man ähnlich beweisen.  $\square$

**Satz 6.4.2** Sei  $A \in \mathbb{K}^{n \times n}$ . Es gilt dann:

für alle Eigenwerte von  $A$  ist geometrische = algebraische Vielfachheit  
 $\Leftrightarrow$  es gibt eine Basis  $(v_1, \dots, v_n)$  von  $\mathbb{K}^n$  aus Eigenvektoren von  $A$

**Beweis.** Sei  $k_i$  die algebraische Vielfachheit,  $m_i$  die geometrische Vielfachheit des Eigenwerts  $\lambda_i$ . Seien  $v_1^{(i)}, \dots, v_{m_i}^{(i)}$  linear unabhängige Eigenvektoren zu  $\lambda_i$ . Nach Lemma 6.4.1 ist dann

$$(v_1^{(1)}, \dots, v_{m_1}^{(1)}, v_1^{(2)}, \dots, v_{m_2}^{(2)}, \dots, v_1^{(r)}, \dots, v_{m_r}^{(r)})$$

linear unabhängig. Diese Vektoren bilden damit eine Basis, genau dann wenn ihre Anzahl gleich der Dimension des Raumes ist, wenn also

$$m_1 + \dots + m_r = \dim \mathbb{K}^n = n.$$

Da  $k_1 + \dots + k_r = n$  und  $m_i \leq k_i$  (Satz 6.3.1 und 6.3.3), folgt

$$m_1 + \dots + m_r = n \Leftrightarrow m_i = k_i \text{ für alle } i.$$

$\square$

**Beachte:** Die Basis im letzten Satz erhält man, indem man für jeden Eigenwert die maximale Anzahl linear unabhängiger Eigenvektoren ausrechnet (also soviele, wie die geometrische Vielfachheit angibt) und dann diese Eigenvektoren für alle Eigenwerte zusammenfasst.

**Beispiel 6.4.3** In Beispiel 6.3.4 haben wir für die Matrix

$$A = \begin{pmatrix} 4 & 4 & 1 \\ -2 & -2 & -1 \\ 2 & 4 & 3 \end{pmatrix}$$

die Eigenvektoren

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \text{ zum Eigenwert } \lambda = 1$$

und

$$v_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \text{ zum Eigenwert } \lambda = 2$$

berechnet (andere Bezeichnung als im alten Beispiel!). Da  $(v_2, v_3)$  linear unabhängig ist, ist nach Lemma 6.4.1 auch  $(v_1, v_2, v_3)$  linear unabhängig, also eine Basis von  $\mathbb{R}^3$  (3 Vektoren und  $\dim \mathbb{R}^3 = 3$ ). Damit ist  $(v_1, v_2, v_3)$  die Basis aus Eigenvektoren gemäß Satz 6.4.2.

Aus einer Basis aus Eigenvektoren von  $A$  erhält man eine invertierbare Matrix  $V$  mit der man die Matrix  $A$  *diagonalisieren* kann:

**Satz 6.4.4** Sei  $A \in \mathbb{K}^{n \times n}$  eine Matrix mit einer Basis  $(v_1, \dots, v_n)$  aus Eigenvektoren. Seien  $\lambda_1, \dots, \lambda_n$  die zugehörigen Eigenwerte, d.h.  $Av_i = \lambda_i v_i$ .

(a) Für  $x \in \mathbb{K}^n$  mit der Basisdarstellung  $x = \alpha_1 v_1 + \dots + \alpha_n v_n$  gilt

$$Ax = \alpha_1 \lambda_1 v_1 + \dots + \alpha_n \lambda_n v_n. \quad (6.5)$$

(b) Sei  $V = (v_1, \dots, v_n)_{n \times n} \in \mathbb{K}^{n \times n}$  die Matrix mit den Vektoren  $v_1, \dots, v_n$  als Spalten. Dann ist  $V$  invertierbar und es gilt

$$V^{-1}AV = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}. \quad (6.6)$$

Man nennt eine Matrix *diagonalisierbar*, wenn es eine Basis aus Eigenvektoren gibt. Gleichung (6.6) heißt dann auch *Diagonalisierung* von  $A$ .

**Beweis.**

(a) Folgt sofort aus  $Av_i = \lambda_i v_i$ :

$$Ax = \alpha_1 Av_1 + \dots + \alpha_n Av_n = \alpha_1 \lambda_1 v_1 + \dots + \alpha_n \lambda_n v_n$$

(b)  $V$  ist invertierbar, da  $(v_1, \dots, v_n)$  linear unabhängig ist. Nach den Regeln der Matrix-Vektor-Multiplikation gilt dann

$$x = \alpha_1 v_1 + \dots + \alpha_n v_n = V \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = V\alpha \quad \text{wobei} \quad \alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

und, zusammen mit (a),

$$Ax = \alpha_1 \lambda_1 v_1 + \dots + \alpha_n \lambda_n v_n = V \begin{pmatrix} \alpha_1 \lambda_1 \\ \vdots \\ \alpha_n \lambda_n \end{pmatrix} = V \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \alpha.$$

Daraus folgt

$$V^{-1}AV\alpha = V^{-1}Ax = V^{-1}V \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \alpha = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \alpha,$$

denn  $V^{-1}V = E$ . Da  $x$  und somit auch der Vektor  $\alpha$  beliebig ist, folgt daraus (6.6).

□

**Bemerkung:** Gleichungen (6.6) ist, durch Multiplikation von links mit  $V$  und von rechts mit  $V^{-1}$ , äquivalent zu

$$A = V \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} V^{-1}.$$

Das zeigt, wie im diagonalisierbaren Fall die Matrix komplett durch ihre Eigenvektoren und Eigenwerte festgelegt ist.

## 6.5 Komplexe Eigenwerte und Eigenvektoren

Die Eigenwerte einer reellen Matrix können komplex sein; in dem Fall sind dann auch die zugehörigen Eigenvektoren komplex. Wir zeigen das an einem simplen Beispiel:

Wir betrachten die reelle Matrix

$$A = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}.$$

Das charakteristische Polynom ist

$$p_A(\lambda) = (1 - \lambda)(-1 - \lambda) + 2 = \lambda^2 + 1,$$

und für die Nullstellen erhalten wir

$$p_A(\lambda) = \lambda^2 + 1 = 0 \Leftrightarrow \lambda = \pm j.$$

Die Matrix  $A$  hat also die beiden komplexen Eigenwerte  $\lambda_1 = j$  und  $\lambda_2 = -j$ . Die Eigenvektoren berechnet man wie gewohnt, da das Gleichungssystem jetzt aber komplex wird, ist die Rechnung etwas aufwändiger. Für den Eigenvektor zu  $\lambda_1 = j$  ergibt sich etwa

$$A - jE = \begin{pmatrix} 1-j & 1 \\ -2 & -1-j \end{pmatrix} \cdot (1+j) \Leftrightarrow \begin{pmatrix} 2 & 1+j \\ -2 & -1-j \end{pmatrix} + I \Leftrightarrow \begin{pmatrix} 2 & 1+j \\ 0 & 0 \end{pmatrix}$$

Damit ist  $x = \begin{pmatrix} 1+j \\ -2 \end{pmatrix}$  ein Eigenvektor zu  $\lambda_1 = j$ , denn

$$\begin{pmatrix} 2 & 1+j \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1+j \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$



## Kapitel 7

# Symmetrische Matrizen und Definitheit

In diesem Kapitel untersuchen wir Eigenwerte und Eigenvektoren symmetrischer Matrizen. Wir werden sehen, dass bei symmetrischen Matrizen keine komplexen Eigenwerte und -vektoren auftreten können. Das führt dann zur Eigenschaft der *Definitheit*.

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik A, Kapitel 8.4, 8.5, 8.6.

### 7.1 Symmetrische Matrizen

**Definition 7.1.1** Eine reelle quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *symmetrisch* wenn gilt  $A = A^T$ . (Dabei bezeichnet  $A^T$  die transponierte Matrix.)

Hat  $A$  die Einträge  $a_{ij}$ , so hat die transponierte Matrix  $A^T$  die Einträge  $a_{ji}$ . (Beim Transponieren werden die Spalten zu Zeilen und umgekehrt. D.h. die Einträge werden an der Diagonalen gespiegelt.) Es gilt also

$$A = (a_{ij}) \text{ symmetrisch} \quad \Leftrightarrow \quad a_{ij} = a_{ji}.$$

Bei einer symmetrischen Matrix liegen die Einträge damit symmetrisch zur Diagonalen. Zum Beispiel ist

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 1 & 0 & -2 \\ 5 & -2 & -4 \end{pmatrix}$$

eine symmetrische  $3 \times 3$  Matrix.

Die Symmetrie einer quadratischen Matrix ist eng mit dem *Skalarprodukt* von Vektoren verbunden. Zur Erinnerung: Für zwei Vektoren

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

ist das Skalarprodukt

$$\langle x, y \rangle = x \cdot y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Im Unterschied zu Mathematik A werden wir das Skalarprodukt hier immer mit  $\langle x, y \rangle$  bezeichnen; im Zusammenhang mit Matrizen ist das übersichtlicher. Wenn wir die allgemeine Regel für die Berechnung des Matrixprodukts benutzen („Zeile mal Spalte“), können wir das Skalarprodukt auch folgendermaßen schreiben:

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x^T y.$$

Sei nun  $A \in \mathbb{R}^{n \times n}$  eine beliebige quadratische Matrix. Es gilt dann

$$\langle Ax, y \rangle = (Ax)^T y = x^T A^T y = \langle x, A^T y \rangle,$$

wobei wir die Regel benutzt haben, dass sich beim Transponieren eines Produkts die Reihenfolge umdreht. Man kann also im Skalarprodukt eine Matrix durch Übergang zur transponierten Matrix von der einen auf die andere Seite bringen. Für eine *symmetrische* Matrix  $A$  folgt damit sofort die wichtige Gleichung

$$\langle Ax, y \rangle = \langle x, Ay \rangle. \quad (7.1)$$

Sie ist für die folgenden Rechnungen entscheidend.

**Satz 7.1.2** *Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch. Dann gilt:*

- (a) *Alle Eigenwerte und Eigenvektoren von  $A$  sind reell.*
- (b) *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal:*

$$Ax = \lambda x, Ay = \mu y, \lambda \neq \mu \implies \langle x, y \rangle = 0, \text{ d.h. } x \perp y.$$

- (c) *Es gibt eine Orthonormalbasis  $(v_1, \dots, v_n)$  aus Eigenvektoren von  $A$ :*

$$Av_i = \lambda_i v_i, \langle v_i, v_j \rangle = 0 \text{ für } i \neq j, \langle v_i, v_i \rangle = 1.$$

Erinnerung: In einer Orthonormalbasis (ONB, siehe Mathematik A) sind die Vektoren also paarweise orthogonal ( $\langle v_i, v_j \rangle = 0 \Leftrightarrow v_i \perp v_j$ ) und normiert auf Länge 1 ( $\langle v_i, v_i \rangle = |v_i|^2 = 1 \Rightarrow |v_i| = 1$ ).

**Beweis.**

- (a) Wir nehmen einen beliebigen Eigenwert und Eigenvektor, die also komplex sein dürfen, und zeigen dann, dass sie reell sind. Sei also

$$Ax = \lambda x \quad \text{mit} \quad \lambda \in \mathbb{C}, x \in \mathbb{C}^n \setminus \{0\}.$$

Da wir zunächst von einem komplexen Vektor ausgehen, benutzen wir auch das komplexe Skalarprodukt  $\langle x, y \rangle = x^T \bar{y}$ . Da  $A$  symmetrisch und reell ist, gilt auch hier

$$\langle Ax, y \rangle = x^T A^T \bar{y} = x^T A \bar{y} = x^T \overline{Ay} = \langle x, Ay \rangle.$$

Damit folgt für den Eigenvektor  $x$ :

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Ax, x \rangle = \langle x, Ax \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \langle x, x \rangle. \quad (7.2)$$

Da  $x \neq 0$ , ist auch

$$\langle x, x \rangle = x_1 \bar{x}_1 + \dots + x_n \bar{x}_n = |x_1|^2 + \dots + |x_n|^2 \neq 0.$$

Daher können wir in (7.2) durch  $\langle x, x \rangle$  teilen und erhalten  $\lambda = \bar{\lambda}$ , d.h.  $\lambda \in \mathbb{R}$ . Damit ist dann auch  $x \in \mathbb{R}^n$ .

(b) Es gilt

$$\begin{aligned} \lambda \langle x, y \rangle &= \langle \lambda x, y \rangle = \langle Ax, y \rangle = \langle x, Ay \rangle = \langle x, \mu y \rangle = \mu \langle x, y \rangle \\ \Rightarrow (\lambda - \mu) \langle x, y \rangle &= 0 \end{aligned}$$

Da die Eigenwerte  $\lambda$  und  $\mu$  verschieden sind, ist  $\lambda - \mu \neq 0$  und es folgt  $\langle x, y \rangle = 0$ .

(c) Der Beweis ist aufwändiger, wir lassen ihn hier aus.

□

## 7.2 Orthogonale Diagonalisierung

Da symmetrische Matrizen eine Basis aus Eigenvektoren besitzen, sind sie also diagonalisierbar (Satz 6.4.4). Da die Basis orthonormal ist (genauer: gewählt werden kann), hat die Diagonalisierung zusätzliche Eigenschaften:

**Satz 7.2.1** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch und  $(v_1, \dots, v_n)$  eine Orthonormalbasis aus Eigenvektoren von  $A$ . Sei  $V = (v_1, \dots, v_n)_{n \times n} \in \mathbb{R}^{n \times n}$ . Dann gilt

$$V^T V = E \quad (V \text{ ist orthogonale Matrix})$$

und

$$V^T A V = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad (\text{orthogonale Diagonalisierung}) \quad (7.3)$$

**Bemerkung und Beweis:** Eine Matrix mit  $V^T V = E$  heißt *orthogonal*. Sind  $v_1, \dots, v_n$  die Spalten von  $V$ , so sind  $v_1^T, \dots, v_n^T$  die Zeilen der transponierten Matrix  $V^T$ . Nach Definition des Matrixprodukts ist dann der Eintrag an der Stelle  $(i, j)$  von  $V^T V$  gleich  $v_i^T v_j = \langle v_i, v_j \rangle$ . Damit gilt

$$\begin{aligned} V^T V = E &\Leftrightarrow \langle v_i, v_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \\ &\Leftrightarrow (v_1, \dots, v_n) \text{ ist Orthonormalbasis} \end{aligned}$$

Daher ist die Matrix  $V$  in Satz 7.2.1 orthogonal.

Für eine orthogonale Matrix folgt aus  $V^T V = E$ , dass  $V$  invertierbar ist mit Inversen

$$V^{-1} = V^T.$$

Die orthogonale Diagonalisierung (7.3) folgt damit sofort aus (6.6). Bei einer orthogonalen Matrix kann man also die Inverse sehr einfach durch Transponieren berechnen und erspart sich damit die ansonsten eventuell sehr aufwändige Berechnung der Inversen! Das ist der Vorteil der orthogonalen Diagonalisierung und orthogonaler Matrizen.

Die orthogonale Diagonalisierung einer symmetrischen Matrix nennt man auch *Hauptachsentransformation*.

**Beispiel 7.2.2** Wir betrachten die symmetrische Matrix

$$A = \begin{pmatrix} 1 & 2\sqrt{2} \\ 2\sqrt{2} & -1 \end{pmatrix}.$$

Die Eigenwerte sind  $\lambda_1 = 3$  und  $\lambda_2 = -3$ , zugehörige Eigenvektoren sind

$$x_1 = \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix} \text{ zu } \lambda_1 = 3, \quad x_2 = \begin{pmatrix} -1 \\ \sqrt{2} \end{pmatrix} \text{ zu } \lambda_2 = -3.$$

Die Berechnung erfolgt wie in Kapitel 6, wir überlassen sie dem Leser ( $\rightarrow$  Übung!). Da die Eigenvektoren  $x_1$  und  $x_2$  zu verschiedenen Eigenwerten gehören, müssen sie nach Satz 7.1.2 orthogonal sein; wir überprüfen das:

$$\langle x_1, x_2 \rangle = \sqrt{2} \cdot (-1) + 1 \cdot \sqrt{2} = -\sqrt{2} + \sqrt{2} = 0,$$

also sind die Eigenvektoren tatsächlich orthogonal. Um eine Orthonormalbasis zu erhalten, normieren wir die Vektoren jetzt auf Länge Eins:

$$v_1 = \frac{1}{|x_1|} x_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix}, \quad v_2 = \frac{1}{|x_2|} x_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ \sqrt{2} \end{pmatrix}.$$

$(v_1, v_2)$  ist damit die Orthonormalbasis aus Eigenvektoren zur Matrix  $A$  (siehe wieder Satz 7.1.2). Wir berechnen jetzt die orthogonale Diagonalisierung nach Satz 7.2.1. Die Matrix  $V$  der Eigenvektoren aus der Orthonormalbasis ist

$$V = (v_1, v_2)_{2 \times 2} = \frac{1}{\sqrt{3}} \begin{pmatrix} \sqrt{2} & -1 \\ 1 & \sqrt{2} \end{pmatrix}$$

Wir wissen aus dem Satz, dass  $V$  orthogonal ist. Zur Veranschaulichung überprüfen wir das aber hier nochmal:

$$V^T V = \frac{1}{3} \begin{pmatrix} \sqrt{2} & 1 \\ -1 & \sqrt{2} \end{pmatrix} \begin{pmatrix} \sqrt{2} & -1 \\ 1 & \sqrt{2} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E,$$

also ist  $V$  orthogonal. Jetzt die Diagonalisierung von  $A$ :

$$\begin{aligned} V^T A V &= \frac{1}{3} \begin{pmatrix} \sqrt{2} & 1 \\ -1 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 2\sqrt{2} \\ 2\sqrt{2} & -1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & -1 \\ 1 & \sqrt{2} \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} \sqrt{2} & 1 \\ -1 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 3 \\ 3 & -3\sqrt{2} \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 9 & 0 \\ 0 & -9 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix}. \end{aligned}$$

Auch dieses Ergebnis kannten wir eigentlich schon vorher aus Satz 7.2.1, denn (7.3) ist hier

$$V^T AV = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix}.$$

### 7.3 Definitheit symmetrischer Matrizen

In Anwendungen ist für eine symmetrische Matrix  $A$  häufig das Vorzeichen des Skalarprodukts  $\langle Ax, x \rangle$  gesucht, wobei  $x$  alle Vektoren aus  $\mathbb{R}^n$  durchläuft. Den Ausdruck  $\langle Ax, x \rangle$  nennt man auch quadratische Form zur Matrix  $A$ . Je nach Vorzeichen der quadratischen Form hat die Matrix eine bestimmte Definitheit.

**Definition 7.3.1** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch.  $A$  heißt...

positiv definit	$\Leftrightarrow$	$\langle Ax, x \rangle > 0$ für alle $x \neq 0$
positiv semidefinit	$\Leftrightarrow$	$\langle Ax, x \rangle \geq 0$ für alle $x \neq 0$
negativ definit	$\Leftrightarrow$	$\langle Ax, x \rangle < 0$ für alle $x \neq 0$
negativ semidefinit	$\Leftrightarrow$	$\langle Ax, x \rangle \leq 0$ für alle $x \neq 0$
indefinit	$\Leftrightarrow$	$\langle Ax, x \rangle > 0$ und $< 0$ für verschiedene $x$

**Beachte:** Die definiten und semidefiniten Fälle sind nicht disjunkt: Jede positiv definite Matrix ist auch positiv semidefinit, jede negativ definite auch negativ semidefinit. Dies sieht man auch an den folgenden Kriterien für Definitheit.

Der folgende Satz beagt, dass die Definitheit einer Matrix vollständig durch die Vorzeichen der Eigenwerte bestimmt ist.

**Satz 7.3.2** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch,  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $A$ . Dann gilt:

$A$ positiv definit	$\Leftrightarrow$	alle $\lambda_i > 0$
$A$ positiv semidefinit	$\Leftrightarrow$	alle $\lambda_i \geq 0$
$A$ negativ definit	$\Leftrightarrow$	alle $\lambda_i < 0$
$A$ negativ semidefinit	$\Leftrightarrow$	alle $\lambda_i \leq 0$
$A$ indefinit	$\Leftrightarrow$	mindestens ein $\lambda_i > 0$ und ein $\lambda_i < 0$

**Beweis.** Sei  $(v_1, \dots, v_n)$  die Orthonormalbasis aus Eigenvektoren (Satz 7.1.2),  $Av_i = \lambda_i v_i$ . Sei  $x$  dargestellt in der Orthonormalbasis,  $x = \sum_{i=1}^n \alpha_i v_i$ . Nach Satz 6.4.4 ist dann  $Ax = \sum_{i=1}^n \alpha_i \lambda_i v_i$  und somit

$$\langle Ax, x \rangle = \left\langle \sum_{i=1}^n \alpha_i \lambda_i v_i, \sum_{j=1}^n \alpha_j v_j \right\rangle = \sum_{i,j=1}^n \alpha_i \alpha_j \lambda_i \langle v_i, v_j \rangle.$$

Aus

$$\langle v_i, v_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

folgt dann

$$\langle Ax, x \rangle = \sum_{i,j=1}^n \alpha_i \alpha_j \lambda_i \langle v_i, v_j \rangle = \sum_{i=1}^n \alpha_i^2 \lambda_i \langle v_i, v_i \rangle = \sum_{i=1}^n \alpha_i^2 \lambda_i.$$

Daraus ergeben sich alle Fälle. Z.B. ist

$$\begin{aligned} \langle Ax, x \rangle > 0 \text{ für alle } x \neq 0 &\Leftrightarrow \sum_{i=1}^n \alpha_i^2 \lambda_i > 0 \text{ für alle } \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \neq 0 \\ &\Leftrightarrow \lambda_1, \dots, \lambda_n > 0. \end{aligned}$$

□

Für  $2 \times 2$ -Matrizen gibt es ein einfaches Determinanten-Kriterium zur Überprüfung der Definitheit.

**Satz 7.3.3** Sei  $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ . Dann ist...

$$\begin{aligned} A \text{ positiv definit} &\Leftrightarrow \det A > 0 \wedge a > 0 \\ A \text{ negativ definit} &\Leftrightarrow \det A > 0 \wedge a < 0 \\ A \text{ indefinit} &\Leftrightarrow \det A < 0 \\ A \text{ positiv semidefinit} &\Leftrightarrow \det A \geq 0 \wedge a \geq 0 \wedge d \geq 0 \\ A \text{ negativ semidefinit} &\Leftrightarrow \det A \geq 0 \wedge a \leq 0 \wedge d \leq 0 \end{aligned}$$

**Beweis.** Nach Satz 6.3.1 ist

$$\det A = ad - b^2 = \lambda_1 \lambda_2, \quad \text{Spur } A = a + d = \lambda_1 + \lambda_2.$$

Damit gilt z.B. im Fall  $\det A > 0$ :

$$\begin{aligned} \det A = ad - b^2 > 0 &\Rightarrow ad > b^2 \geq 0 \\ &\Rightarrow a > 0 \wedge d > 0 \quad \text{oder} \quad a < 0 \wedge d < 0 \end{aligned}$$

Andererseits gilt

$$\det A = \lambda_1 \lambda_2 > 0 \quad \Rightarrow \quad \lambda_1 > 0 \wedge \lambda_2 > 0 \quad \text{oder} \quad \lambda_1 < 0 \wedge \lambda_2 < 0$$

Anders gesagt haben also die Zahlen  $a$  und  $d$  dasselbe Vorzeichen (und sind ungleich Null), genauso bei  $\lambda_1$  und  $\lambda_2$ . Zusammen mit  $\text{Spur } A = a + d = \lambda_1 + \lambda_2$  ergeben sich dann folgende Implikationen:

$$\begin{aligned} a > 0 &\Rightarrow d > 0 \Rightarrow \text{Spur } A > 0 \Rightarrow \lambda_1, \lambda_2 > 0 \Rightarrow A \text{ positiv definit} \\ a < 0 &\Rightarrow d < 0 \Rightarrow \text{Spur } A < 0 \Rightarrow \lambda_1, \lambda_2 < 0 \Rightarrow A \text{ negativ definit} \end{aligned}$$

Im Fall  $\det A < 0$  ergibt sich

$$\det A = \lambda_1 \lambda_2 < 0,$$

woraus folgt, dass  $\lambda_1$  und  $\lambda_2$  beide ungleich Null sind und unterschiedliche Vorzeichen haben. Damit ist  $A$  dann indefinit. Der Fall  $\det A = 0$  läuft ähnlich. □

**Beispiel 7.3.4** Wir untersuchen die Matrix

$$A = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}$$

auf Definitheit, zuerst mit Hilfe der Eigenwerte und dann mit dem Determinanten-Kriterium:

(a) mit den Eigenwerten:

$$p_A(\lambda) = \det \begin{pmatrix} 5-\lambda & -2 \\ -2 & 1-\lambda \end{pmatrix} = (5-\lambda)(1-\lambda) - 4 = \lambda^2 - 6\lambda + 1 = 0$$

$$\Leftrightarrow (\lambda - 3)^2 = 9 - 1 = 8 \quad \Leftrightarrow \lambda = 3 \pm \sqrt{8}$$

Die Eigenwerte sind also

$$\lambda_1 = 3 + \sqrt{8}, \quad \lambda_2 = 3 - \sqrt{8}.$$

Es ist  $\lambda_1 > 0$  und wegen  $\sqrt{8} < \sqrt{9} = 3$  ist auch  $\lambda_2 = 3 - \sqrt{8} > 0$ . Also ist  $A$  nach Satz 7.3.2 positiv definit.

(b) mit dem Determinanten-Kriterium:

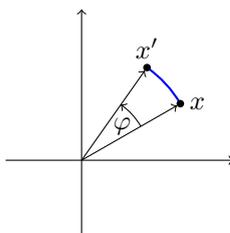
$$\det A = \det \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} = 5 - 4 = 1 > 0 \quad \text{und} \quad a = 5 > 0$$

Nach Satz 7.3.3 ist somit  $A$  positiv definit.

## 7.4 Drehungen in der Ebene

Zum Abschluss dieses Kapitels befassen wir uns noch mit der mathematischen Beschreibung von Drehungen in Ebene und Raum. Solche Drehung können durch Multiplikation mit speziellen Matrizen beschrieben werden, den *Drehmatrizen*.

Zur Herleitung der Formeln für die Drehung in der Ebene betrachten wir einen Punkt  $x \in \mathbb{R}^2$ , den wir um den Ursprung mit dem Winkel  $\varphi$  drehen wollen:



Was sind die neuen Koordinaten  $x'$  des Punktes nach der Drehung? Mit den Standard-Einheitsvektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

gilt für  $x = (x_1, x_2)$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 e_1 + x_2 e_2.$$

Die Drehung von  $x$  können wir jetzt durch Drehung der beiden Vektoren  $e_1$  und  $e_2$  erzeugen. Wir bezeichnen dafür die gedrehten Einheitsvektoren mit  $e'_1$  und  $e'_2$ . Dann hat der gedrehte Punkt  $x'$  bezüglich  $e'_1, e'_2$  dieselben Koordinaten  $x_1, x_2$  wie

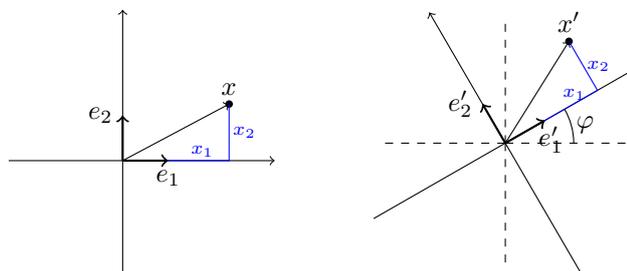
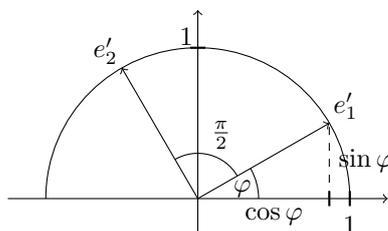
Abbildung 7.1: Konstruktion der Drehung um  $\varphi$ 

Abbildung 7.2: Drehung der Standardeinheitsvektoren

der ursprüngliche Punkt  $x$  bezüglich  $e_1, e_2$ . Graphisch ist das in Abbildung 7.1 veranschaulicht. Für den neuen Punkt gilt also

$$x' = x_1 e'_1 + x_2 e'_2.$$

Die gedrehten Einheitsvektoren können wir mittels Polarkoordinaten bestimmen, siehe Abbildung 7.2:

$$e'_1 = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \quad e'_2 = \begin{pmatrix} \cos(\varphi + \frac{\pi}{2}) \\ \sin(\varphi + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}.$$

Damit erhalten wir für den gedrehten Punkt  $x'$ :

$$\begin{aligned} x' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} &= x_1 \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} + x_2 \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos(\varphi) \cdot x_1 - \sin(\varphi) \cdot x_2 \\ \sin(\varphi) \cdot x_1 + \cos(\varphi) \cdot x_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned}$$

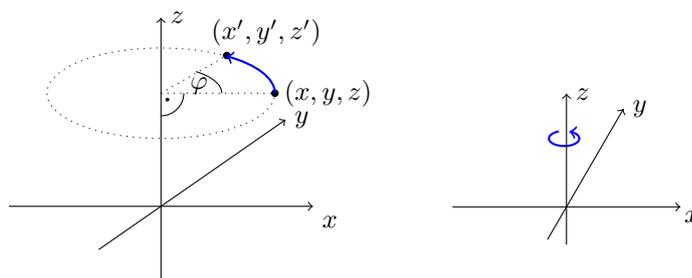
Die Drehung lässt sich also durch ein Matrix-Vektor-Produkt beschreiben, oder anders gesagt: Die Drehung um  $\varphi$  wird erzeugt durch die Matrix

$$D(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}. \quad (7.4)$$

Diese Matrix wird als *Drehmatrix* bezeichnet und beschreibt also die Drehung um den Winkel  $\varphi$  gegen den Uhrzeigersinn durch die Gleichung

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = D(\varphi) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Eigenschaften der Drehmatrix sind:

Abbildung 7.3: Drehung um  $z$ -Achse und Rechtsschraube um die  $z$ -Achse

- $D(\varphi)$  ist eine orthogonale Matrix, denn die Spalten  $e'_1, e'_2$  sind orthogonal und auf Länge 1 normiert, also eine Orthonormalbasis. Da  $D(\varphi)$  orthogonal ist, folgt

$$D(\varphi)^{-1} = D(\varphi)^T = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = D(-\varphi).$$

Die Umkehrung der Drehung um  $\varphi$  ist also – wie erwartet – die Drehung um  $-\varphi$ , also um denselben Winkel in die entgegengesetzte Richtung.

- $\det D(\varphi) = \cos^2 \varphi + \sin^2 \varphi = 1$ .

## 7.5 Drehungen im Raum

- Drehung um die  $z$ -Achse:

Bei der Drehung um die  $z$ -Achse wird der Punkt in einer Ebene parallel zur  $x$ - und  $y$ -Achse gedreht, vergleiche Abbildung 7.3 links. Die  $z$ -Komponente des Punktes bleibt also gleich, und die neuen  $x$ - und  $y$ -Komponenten ergeben sich durch Anwendung der 2-dimensionalen Drehmatrix der Ebene auf den Vektor  $(x, y)^T$ :

$$\begin{aligned} z' &= z \\ \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

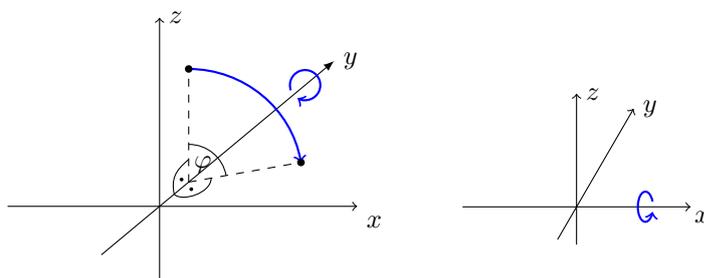
Das kann man wieder komplett als Matrix-mal-Vektor zusammenfassen:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = D_z(\varphi) \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Die Drehmatrix

$$D_z(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.5)$$

erzeugt also die Drehung von  $(x, y, z)$  um die  $z$ -Achse um den Winkel  $\varphi$ . Der Drehsinn ist dabei der gleiche wie in der Ebene: die positive  $x$ -Achse wird in Richtung der positiven  $y$ -Achse gedreht. Im Raum entspricht das einer „Rechtsschraube“ um die  $z$ -Achse, in Abbildung 7.3 rechts dargestellt.

Abbildung 7.4: Drehungen um  $y$ - sowie  $x$ -Achse

- Drehung um die  $y$ -Achse:

Hier wird in einer Ebene parallel zur  $x$ - und  $z$ -Achse gedreht,  $y$  bleibt konstant. Der Drehsinn ist jetzt entsprechend eine Rechtsschraube um die  $y$ -Achse. Das bedeutet, dass die positive  $z$ -Achse in Richtung der positiven  $x$ -Achse gedreht wird, siehe Abbildung 7.4 links. Damit gilt

$$y' = y$$

$$\begin{pmatrix} z' \\ x' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} z \\ x \end{pmatrix}.$$

Die „vertauschte“ Reihenfolge  $z, x$  beim Anwenden der 2-dimensionalen Drehmatrix ergibt sich daraus, dass die positive  $z$ -Achse auf die positive  $x$ -Achse gedreht wird (nicht andersherum). Sortiert man jetzt die Komponenten und schreibt wieder als Matrix-mal-Vektor, erhält man für die Drehung von  $(x, y, z)$  um die  $y$ -Achse

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = D_y(\varphi) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

mit der Drehmatrix

$$D_y(\varphi) = \begin{pmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{pmatrix}. \quad (7.6)$$

- Drehung um die  $x$ -Achse:

Der Drehsinn der Rechtsschraube bewirkt hier die Drehung der positiven  $y$ -Achse auf die positive  $z$ -Achse, siehe Abbildung 7.4 rechts. Damit gilt

$$x' = x$$

$$\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

und es folgt

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = D_x(\varphi) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

mit

$$D_x(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}. \quad (7.7)$$

Wie im Zweidimensionalen gilt: Alle Drehmatrizen  $D_x(\varphi)$ ,  $D_y(\varphi)$ ,  $D_z(\varphi)$  sind orthogonale Matrizen und haben Determinante 1.



## Kapitel 8

# Differentialrechnung von Funktionen mehrerer Variablen

In diesem Kapitel untersuchen wir Funktionen, die von mehreren Variablen abhängen. Wie schon bei Funktionen von einer Variablen benutzen wir wieder vor allem Ableitungen, um ihren Verlauf zu untersuchen. Z.B. berechnen wir Maxima und Minima und auch Taylorentwicklungen. Dass die Funktion von mehreren Variablen abhängt, führt zu neuen Begriffen und Situationen, die wir von Funktionen mit nur einer Variablen noch nicht kennen. Z.B. werden wir sehen, dass es mehrere Arten von Ableitungen gibt.

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik B, Kapitel 6.

### 8.1 Mengen in mehreren Raumdimensionen

Funktionen, die von mehreren Variablen abhängen, haben allgemein die Form

$$y = f(x_1, \dots, x_n),$$

d.h., die Funktion  $f$  hängt von den reellen Variablen  $x_1, \dots, x_n \in \mathbb{R}$  ab. Zur Vereinfachung macht es Sinn, die unabhängigen Variablen  $x_1, \dots, x_n$  als einen Punkt im Raum  $x = (x_1, \dots, x_n)$  zusammen zu fassen. Damit können wir die Funktion dann kurz als

$$y = f(x) \quad \text{mit} \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

schreiben. Im Fall von  $n = 2$  (Punkt in der Ebene) schreiben wir auch  $(x, y)$  statt  $(x_1, x_2)$  und entsprechend bei  $n = 3$  (Punkt im 3-dim. Raum)  $(x, y, z)$  statt  $(x_1, x_2, x_3)$ . Man kann den Punkt  $x$  auch als Vektor auffassen, d.h. schreiben

$$x = (x_1, \dots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n;$$

Abbildung 8.1:  $\varepsilon$ -Umgebung  $U_\varepsilon(a)$  eines Punktes und Rand  $\partial G$  einer Menge

manchmal ist das sinnvoll, obwohl in der Funktion  $y = f(x)$  das  $x$  eigentlich nicht die Rolle eines Vektors zur Angabe einer Länge und Richtung hat.

Da die Variable, von der die Funktion  $f$  jetzt abhängt, also ein Punkt im Raum ist, wird es wichtig sein, Begriffe und Eigenschaften von *Mengen von Punkten* im Raum zu beschreiben. Eine Menge von Punkten im Raum ist eine Teilmenge  $G \subset \mathbb{R}^n$ . Der Definitionsbereich von  $f$  etwa ist eine solche Menge von Punkten:

$$D_f = \{x \in \mathbb{R}^n \mid f(x) \text{ ist definiert}\},$$

also  $D_f \subset \mathbb{R}^n$ .

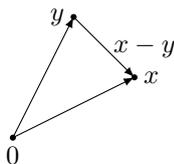
Wir beginnen mit der Formel für den Abstand zweier Punkte  $x, y \in \mathbb{R}^n$ :

$$|x - y| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}. \quad (8.1)$$

Das ist nichts anderes, als die Länge von  $x - y$ , wenn man  $x - y$  als Vektor betrachtet:

$$|x - y| = \left\| \begin{pmatrix} x_1 - y_1 \\ \vdots \\ x_n - y_n \end{pmatrix} \right\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

Man beachte, dass  $x - y$  als Vektor genau der Vektor vom Punkt  $y$  zum Punkt  $x$  ist,



Die  $\varepsilon$ -Umgebung eines Punktes  $a \in \mathbb{R}^n$ , ist die Menge aller Punkte, die von  $a$  einen Abstand kleiner als  $\varepsilon$  haben, die also „in der Nähe von  $a$ “ liegen, siehe Abbildung 8.1:

$$U_\varepsilon(a) = \{x \in \mathbb{R}^n \mid |x - a| < \varepsilon\}.$$

Die  $\varepsilon$ -Umgebung  $U_\varepsilon(a)$  ist also der Kreis bzw. die Kugel mit Radius  $\varepsilon$  um den Mittelpunkt  $a$  (wobei der Rand des Kreises/der Kugel nicht mehr dazu gehört, weil dort der Abstand *gleich*  $\varepsilon$  ist.)

Für den *Rand* einer Menge  $G \subset \mathbb{R}^n$  verwenden wir die Bezeichnung  $\partial G$ . Er ist die Menge aller Randpunkte von  $G$ ,

$$\partial G = \{\text{alle Randpunkte von } G\}.$$

Für einfache Menge ist geometrisch klar, was die Randpunkte sind, siehe Abbildung 8.1. Trotzdem geben wir hier noch die Definition an: Ein Punkt  $a \in \mathbb{R}^n$  ist *Randpunkt* von  $G$ , wenn jede Umgebung  $U_\varepsilon(a)$  Punkte aus  $G$  und aus  $\mathbb{R}^n \setminus G$  enthält.

Ausgehend vom Begriff des Randes, können wir jetzt die im folgenden wichtigen Begriffe offener und abgeschlossener Mengen definieren:

**Definition 8.1.1** Eine Menge  $G \subset \mathbb{R}^n$  heißt...

- (a) *offen*, wenn  $\partial G \cap G = \emptyset$ , d.h. wenn kein Randpunkt zu  $G$  gehört;
- (b) *abgeschlossen*, wenn  $\partial G \subset G$ , d.h. wenn alle Randpunkte zu  $G$  gehören.

**Beachte:** Offen und abgeschlossen sind nicht das Gegenteil von einander: eine Menge kann auch weder offen noch abgeschlossen sein, nämlich genau dann einige Randpunkte zur Menge gehören, andere aber nicht.

**Beispiel 8.1.2** (a) Die Menge

$$K = \left\{ (x, y) \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} < 1 \right\}$$

beschreibt einen Kreis um den Ursprung  $(0, 0)$  mit Radius 1. Denn  $r = \sqrt{x^2 + y^2}$  ist der Radius des Punkts  $(x, y)$  in Polarkoordinaten. Zur Menge gehören also alle Punkte mit  $r < 1$ . Alternativ kann man auch sagen

$$\sqrt{x^2 + y^2} = \left| \begin{pmatrix} x \\ y \end{pmatrix} \right| = \text{Länge von } \begin{pmatrix} x \\ y \end{pmatrix} = \text{Abstand } (x, y) \text{ zu } (0, 0).$$

Die Punkte  $(x, y)$  mit Abstand  $< 1$  gehören zu  $K$  dazu, die Punkte mit Abstand  $= 1$  aber nicht mehr. Daher ist  $K$  der Kreis mit Radius 1, aber *ohne* den Kreisrand, siehe auch Abbildung 8.2. Der Rand  $\partial K$  besteht damit genau aus den Punkten mit Abstand  $= 1$ , also

$$\partial K = \left\{ (x, y) \mid \sqrt{x^2 + y^2} = 1 \right\}. \quad (\text{Kreislinie})$$

Die Menge  $K$  ist offen, denn kein Randpunkt liegt in  $K$ . (Der Kreisrand gehört nicht zu  $K$  wegen dem „ $<$ “ in der Definition von  $K$ .)

(b) Die Menge

$$D = \left\{ (x, y) \mid \sqrt{x^2 + y^2} \leq 1 \right\}$$

beschreibt den Kreis mit Radius 1 *mit* Rand. Denn wegen dem „ $\leq$ “ gehören auch die Punkte mit Abstand  $= 1$  zur Menge dazu (Abbildung 8.2). Der Rand der Menge besteht wieder aus den Punkten mit Abstand  $= 1$ , d.h.

$$\partial D = \left\{ (x, y) \mid \sqrt{x^2 + y^2} = 1 \right\}.$$

Die Menge  $D$  ist abgeschlossen, denn alle Randpunkte liegen in  $D$  (wegen „ $\leq$ “).



Abbildung 8.2: Offene und abgeschlossene Kreisscheiben

- (c) Die Begriffe „offen“ und „abgeschlossen“ kennen wir schon für Intervalle in  $\mathbb{R}$  ( $n = 1$ ):

$$\begin{aligned} ]a, b[ &= \{x \in \mathbb{R} \mid a < x < b\} && \text{(offenes Intervall)} \\ [a, b] &= \{x \in \mathbb{R} \mid a \leq x \leq b\} && \text{(abgeschlossenes Intervall)} \end{aligned}$$

Die Randpunkte von  $]a, b[$  und  $[a, b]$  sind für beide Intervalle  $a$  und  $b$ , der linke und der rechte Randpunkt. Da die Randpunkte bei  $]a, b[$  nicht zum Intervall dazu gehören, ist das offene Intervall  $]a, b[$  tatsächlich eine offene Menge nach der Definition 8.1.1. Bei  $[a, b]$  gehören die Randpunkte dazu. Also ist das abgeschlossene Intervall  $[a, b]$  auch abgeschlossen. Die neue Definition stimmt mit den schon bekannten Begriffen überein.



## 8.2 Funktionen mehrerer Variablen

Für Funktionen von mehreren Variablen verwenden wir allgemein die Schreibweise

$$\begin{aligned} f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}, & & \text{oder} & & f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}, \\ y = f(x) & & & & x \mapsto f(x) \end{aligned}$$

Dabei steht wie schon erwähnt  $x$  für den Punkt  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ , die Funktion hängt also tatsächlich von  $n$  Variablen  $x_1, \dots, x_n$  ab. Die Menge  $G$  bezeichnet den Definitionsbereich von  $f$ , d.h.  $G = D_f$ . Man sagt auch „ $f$  ist eine Funktion von  $\mathbb{R}^n$  nach  $\mathbb{R}$ “ oder „ $f$  bildet von  $\mathbb{R}^n$  nach  $\mathbb{R}$  ab“.

**Beispiel 8.2.1** (a) Der Ausdruck

$$f(x_1, x_2, x_3) = \frac{x_1^2}{x_3} - x_2$$

definiert eine Funktion  $f : G \subset \mathbb{R}^3 \rightarrow \mathbb{R}$  mit Definitionsbereich

$$G = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 \neq 0\}.$$

(b) Der Ausdruck

$$f(x, y) = \sqrt{x^2 + y^2}$$

ist für alle Punkte  $(x, y) \in \mathbb{R}^2$  definiert (denn  $x^2 + y^2 \geq 0$ , sodass der Wurzelausdruck immer definiert ist). Hier ist also  $G = \mathbb{R}^2$ , und man lässt dann  $G$  in der Funktionsschreibweise weg,

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \sqrt{x^2 + y^2}.$$

So wie Funktionen von mehreren Variablen abhängen können, kann eine Funktion auch einen Vektor statt einer Zahl als Funktionswert liefern. Man spricht dann von *vektorwertigen Funktionen*. Im allgemeinen Fall ist eine von mehreren Variablen abhängige vektorwertige Funktion von der Form

$$f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$f(x_1, \dots, x_n) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$

was man wieder kurz als

$$f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad y = f(x)$$

schreiben kann, wobei  $x \in G \subset \mathbb{R}^n$  ist und  $y$  ein Vektor

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m.$$

Für  $m \geq 2$  haben wir hier tatsächlich eine vektorwertige Funktion; im Fall  $m = 1$  erhalten wir dagegen wieder  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , man nennt  $f$  dann auch *skalarwertig*.

**Beispiel 8.2.2** (a) Durch

$$f(x, y) = \begin{pmatrix} x + x^2y \\ 1 - xy \end{pmatrix}$$

erhalten wir eine Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Die erste Komponente der Funktion ist  $f_1(x, y) = x + x^2y$ , die zweite Komponente ist  $f_2(x, y) = 1 - xy$ .

(b) **Beispiele aus der Anwendung:** Die Temperaturverteilung ist eine skalarwertige Funktion

$$T : G \subset \mathbb{R}^3 \rightarrow \mathbb{R}, \quad T = T(x, y, z).$$

Sie beschreibt die (unterschiedliche) Temperatur in den Punkten eines Körpers. Der Definitionsbereich  $G$  hat hier die Rolle des Körpers:

der Körper  $G$  hat im Punkt  $(x, y, z) \in G$  die Temperatur  $T(x, y, z)$

Allgemeiner ist  $G$  das Gebiet, in dem die Funktion die Temperatur festlegt.

Ein Beispiel für eine vektorwertige Funktion ist das elektrische Feld

$$E : G \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3.$$

Im Punkt  $(x, y, z) \in G$  hat man die elektrische Feldstärke

$$\vec{E} = \begin{pmatrix} E_1(x, y, z) \\ E_2(x, y, z) \\ E_3(x, y, z) \end{pmatrix} = E(x, y, z).$$

### 8.3 Grenzwerte und Stetigkeit

Für die Analysis von Funktionen mehrerer Variablen benötigen wir Grenzwerte und Stetigkeit solcher Funktionen. Das funktioniert im Wesentlichen genauso wie bei Funktionen einer Variablen von  $\mathbb{R}$  nach  $\mathbb{R}$ :

(a) *Konvergenz* von  $x \in \mathbb{R}^n$  gegen  $a \in \mathbb{R}^n$ . Es gilt:

$$\begin{aligned} x \rightarrow a \quad \text{wobei} \quad x &= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \\ \Leftrightarrow x_1 \rightarrow a_1, \dots, x_n \rightarrow a_n &\quad (\text{komponentenweise}) \\ \Leftrightarrow |x - a| \rightarrow 0 & \end{aligned}$$

Die Konvergenz  $x \rightarrow a$  erfolgt also komponentenweise,  $x_i \rightarrow a_i$  für  $i = 1, \dots, n$ . Dies ist aber äquivalent zu  $|x - a| \rightarrow 0$  (siehe (8.1)), d.h.  $x \rightarrow a$  gilt genau dann, wenn der Abstand  $|x - a|$  beliebig klein wird. Die Richtung, aus der  $x$  sich an  $a$  annähert, ist dabei beliebig.

(b) *Grenzwerte* einer Funktion  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ : Wir schreiben

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

wenn aus der Konvergenz  $x \rightarrow a$  immer  $f(x) \rightarrow y$  folgt. (Das ist genau wie bei Funktionen von einer Variablen.)

(c)  $f$  ist *stetig* im Punkt  $a \in G$ , wenn

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

Es gilt (wie bei einer Variablen):

Summen, Produkte, Quotienten, Verkettungen der elementaren Funktionen sind stetig auf ihrem Definitionsbereich.

**Beispiel 8.3.1** Die Funktion

$$f(x_1, x_2, x_3) = \sqrt{x_1 x_2} + e^{x_3} \cdot (x_2 - x_3)$$

ist stetig auf dem Definitionsbereich  $G = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 x_2 \geq 0\}$ .

## 8.4 Partielle Ableitung

Für Funktionen von mehreren Variablen gibt es verschiedene Ableitungsbegriffe. Der erste ist der der partiellen Ableitung. Sei  $f$  eine skalarwertige Funktion

$$f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}, \quad y = f(x_1, \dots, x_n)$$

und der Definitionsbereich  $G$  eine offene Menge.

**Definition 8.4.1** Die *partielle Ableitung* von  $f$  nach der Variablen  $x_i$  im Punkt  $a = (a_1, \dots, a_n) \in G$  ist

$$\frac{\partial f}{\partial x_i}(a) = \lim_{x_i \rightarrow a_i} \frac{f(a_1, \dots, x_i, \dots, a_n) - f(a_1, \dots, a_i, \dots, a_n)}{x_i - a_i}.$$

$f$  heißt in  $a$  nach  $x_i$  partiell differenzierbar, wenn dieser Grenzwert existiert.

Mit der Abkürzung  $h = x_i - a_i$  können wir die partielle Ableitung auch als

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_i + h, \dots, a_n) - f(a_1, \dots, a_n)}{h}$$

schreiben oder noch kürzer als

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \rightarrow 0} \frac{f(a + he_i) - f(a)}{h}$$

wobei  $e_i \in \mathbb{R}^n$  der  $i$ . Einheitsvektor ist, also

$$a + he_i = a + h \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_i + h \\ \vdots \\ a_n \end{pmatrix}.$$

Auch die partielle Ableitung ist also der Grenzwert eines Differenzenquotienten. Da im Differenzenquotienten nur die  $i$ . Stelle von  $a_i$  zu  $x_i$  variiert wird, alle anderen Stellen aber konstant bleiben, ist  $\frac{\partial f}{\partial x_i}(a)$  nichts anderes als die gewöhnliche Ableitung der Funktion

$$x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$$

an der Stelle  $x_i = a_i$ . D.h. bei der partiellen Ableitung leitet man  $f(x_1, \dots, x_n)$  nach  $x_i$  ab wobei alle anderen Variablen  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$  konstant sind.

Wir haben für die Definition vorausgesetzt, dass die Menge  $G$  offen ist. Damit ist der Punkt  $a \in G$  automatisch kein Randpunkt des Definitionsbereichs, so dass hier keine einseitigen Ableitungen vorkommt. Das vereinfacht die Theorie.

Für die partielle Ableitung nach  $x_i$  gibt es viele Schreibweisen:

$$\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i} f = \partial_{x_i} f = \partial_i f = f_{x_i}.$$

**Beispiel 8.4.2** (a) Wir berechnen die partiellen Ableitungen von

$$f(x_1, x_2) = x_1^3 + x_1^2 x_2^2 + \sqrt{x_2}.$$

Für die Ableitung nach  $x_1$  behandeln wir die andere Variable  $x_2$  als konstant. Damit ergibt sich also

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 3x_1^2 + 2x_1 x_2^2.$$

Entsprechend ist beim Ableiten nach  $x_2$  die erste Variable  $x_1$  konstant:

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = x_1^2 \cdot 2x_2 + \frac{1}{2\sqrt{x_2}}.$$

(b) Für die partiellen Ableitungen von

$$f(x, y, z) = e^{xy} + x \cdot \sin(z)$$

erhalten wir

$$\frac{\partial f}{\partial x} = e^{xy} \cdot y + \sin(z) \quad (y \text{ und } z \text{ konstant})$$

$$\frac{\partial f}{\partial y} = e^{xy} \cdot x \quad (x \text{ und } z \text{ konstant})$$

$$\frac{\partial f}{\partial z} = x \cdot \cos(z) \quad (x \text{ und } y \text{ konstant})$$

## 8.5 Ableitungen höherer Ordnung

Durch mehrfaches partielles Ableiten ergeben sich partielle Ableitungen höherer Ordnung: Leiten wir  $\frac{\partial f}{\partial x_i}$  partiell nach  $x_k$  ab, so erhalten wir die partielle Ableitung 2. Ordnung

$$\frac{\partial^2 f}{\partial x_k \partial x_i} = \frac{\partial}{\partial x_k} \frac{\partial f}{\partial x_i} = \partial_{x_k} \partial_{x_i} f = f_{x_i x_k}.$$

Beim Ableiten nach denselben Variablen ( $k = i$ ) schreibt man

$$\frac{\partial^2}{\partial x_i^2} = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_i} = \partial_{x_i}^2 f.$$

Analog definiert und schreibt man partielle Ableitungen 3. und noch höherer Ordnung.

**Beispiel 8.5.1** Für die Funktion  $f(x, y) = x^2 \cdot \cos y$  ist die partielle Ableitung 1. Ordnung nach  $x$

$$\frac{\partial f}{\partial x} = 2x \cdot \cos y.$$

Leiten wir dies partiell nach  $x$  bzw.  $y$  ab, erhalten wir die zwei partiellen Ableitungen 2. Ordnung

$$\frac{\partial^2 f}{\partial x^2} = 2 \cdot \cos y, \quad \frac{\partial^2 f}{\partial y \partial x} = -2x \cdot \sin y.$$

Entsprechend erhält man aus  $\frac{\partial f}{\partial y}$  die Ableitungen 2. Ordnung  $\frac{\partial^2 f}{\partial x \partial y}$  und  $\frac{\partial^2 f}{\partial y^2}$ .

Es zeigt sich, dass gemischte partielle Ableitungen zweiter Ordnung in den allermeisten Fällen unabhängig von der Reihenfolge der Ableitungen sind. Im letzten Beispiel etwa liefert die Ableitung von  $\frac{\partial f}{\partial x}$  nach  $y$ , also  $\frac{\partial^2 f}{\partial y \partial x}$  dasselbe Ergebnis, wie die Ableitung von  $\frac{\partial f}{\partial y}$  nach  $x$ , also

Die Eigenschaft, die sicherstellt, dass die Reihenfolge bei partiellen Ableitungen egal ist, ist die *stetige partielle Differenzierbarkeit*:

**Definition 8.5.2** Sei  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $G$  offen.  $f$  heißt *stetig partiell differenzierbar* wenn  $f$  partiell nach jedem  $x_i$  differenzierbar ist und alle partiellen Ableitungen  $\frac{\partial f}{\partial x_i}$  stetig sind.  $f$  heißt *zweimal stetig partiell differenzierbar*, wenn  $f$  zweimal partiell differenzierbar ist und alle partiellen Ableitungen 2. Ordnung  $\frac{\partial^2 f}{\partial x_i \partial x_j}$  stetig sind.

Für zweimal stetig partiell differenzierbare Funktionen gilt der Satz von Schwarz über die Vertauschbarkeit partieller Ableitungen:

**Satz 8.5.3 (Satz von Schwarz)** Sei  $G \subset \mathbb{R}^n$  offen und  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  zweimal stetig partiell differenzierbar. Dann gilt:

$$\frac{\partial^2 f}{\partial x_k \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_k} \quad \text{für alle } i, k = 1, \dots, n.$$

**Beachte:** Nach dem Satz von Schwarz langt es für die Berechnung der partiellen Ableitungen 2. Ordnung die gemischten Ableitungen nur in einer Reihenfolge zu berechnen, da sich in der anderen Reihenfolge ja das gleiche ergibt. Man kann diese Eigenschaft aber auch *als Probe nutzen!* Man berechnet die gemischten Ableitungen auch in der jeweils anderen Reihenfolge und überprüft dann, dass man dasselbe Ergebnis erhält. Falls nicht, hat man einen Rechenfehler gemacht!

## 8.6 Totale Differenzierbarkeit und Gradient

Die partielle Differenzierbarkeit untersucht das Änderungsverhalten einer Funktion entlang einer festen Koordinatenrichtung  $x_i$ . Denn beim partiellen Ableiten nach  $x_i$  werden ja die übrigen Variablen  $x_k$  ( $k \neq i$ ) als konstant behandelt, d.h. man betrachtet nur Änderungen von  $x_i$ , d.h. entlang der Koordinatenrichtung  $x_i$ . Wir kommen jetzt zu einem Ableitungskonzept, das Änderungen in alle Raumrichtungen zugleich berücksichtigt.

Zur Motivation betrachten wir zunächst eine Funktion  $g : \mathbb{R} \rightarrow \mathbb{R}$ , die nur von einer Variablen abhängt. Die Gleichung der Tangente an  $g$  in  $x = a$  ist dann

$$t(x) = g(a) + g'(a) \cdot (x - a)$$

(Mathematik A). Die Tangente  $t$  nähert die Funktion  $g$  in der Nähe der Stelle  $a$  linear an,

$$g(x) \approx t(x) = g(a) + g'(a)(x - a) \quad \text{für } x \text{ „nah“ bei } a$$

(sogenannte lineare Approximation). Genauer gilt

$$g(x) = g(a) + g'(a)(x - a) + r(x - a) \quad (8.2)$$

wobei  $r(x-a)$  also genau der Fehler bei der Approximation von  $g(x)$  durch  $t(x)$  ist. Für diese Fehler gilt dann

$$\begin{aligned} r(x-a) &= g(x) - g(a) - g'(a)(x-a) \\ \Rightarrow \lim_{x \rightarrow a} \frac{r(x-a)}{x-a} &= \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x-a} - g'(a) = g'(a) - g'(a) = 0 \end{aligned}$$

Diese Grenzwerteigenschaft des Fehlers verwendet man jetzt, um zu definieren, wann eine Funktion mehrerer Variablen eine lineare Approximation besitzt. Da  $x-a \in \mathbb{R}^n$  dann ein Vektor ist, die Funktionswerte aber reelle Zahlen bleiben, muss der Term  $g'(a)(x-a)$  in (8.2) geändert werden. Man ersetzt ihn durch ein Skalarprodukt  $\langle d, x-a \rangle$ , wobei der Vektor  $d \in \mathbb{R}^n$  noch bestimmt werden muss.

**Definition 8.6.1** Die Funktion  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $G$  offen heißt in  $a \in G$  *total differenzierbar*, wenn es ein  $d \in \mathbb{R}^n$  gibt mit

$$f(x) = f(a) + \langle d, x-a \rangle + r(x-a)$$

und

$$\lim_{x \rightarrow a} \frac{r(x-a)}{|x-a|} = 0. \quad (8.3)$$

Totale Differenzierbarkeit ist eine stärkere Eigenschaft als partielle Differenzierbarkeit. Das zeigt der folgende Satz, in dem wir auch den noch unbekanntem Vektor  $d$  berechnen:

**Satz 8.6.2** *Ist  $f$  total differenzierbar, dann ist  $f$  auch partiell differenzierbar und der Vektor  $d$  aus Definition 8.6.1 ist*

$$d = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

**Beweis.** Sei  $f$  total differenzierbar und  $d = (d_1, \dots, d_n)^T$  der entsprechende Vektor aus Definition 8.6.1. Es gilt also

$$f(x) = f(a) + \langle d, x-a \rangle + r(x-a).$$

Sei  $x = a + he_i$  wobei  $h \in \mathbb{R}$  und  $e_i$  der  $i$ . Standard-Einheitsvektor ist, also

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

mit der 1 an der  $i$ . Stelle. Wir erhalten dann

$$f(a + he_i) = f(a) + \langle d, he_i \rangle + r(he_i) = f(a) + hd_i + r(he_i)$$

und damit

$$\frac{f(a + he_i) - f(a)}{h} = d_i + \frac{r(he_i)}{h}$$

Im Grenzwert  $h \rightarrow 0$  ergibt dies die partielle Ableitung nach  $x_i$  im Punkt  $a$ :

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \rightarrow 0} \frac{f(a + he_i) - f(a)}{h} = d_i + \lim_{h \rightarrow 0} \frac{r(he_i)}{h}. \quad (8.4)$$

Mit (8.3) folgt

$$\left| \frac{r(he_i)}{h} \right| = \frac{|r(he_i)|}{|he_i|} = \frac{|r(x-a)|}{|x-a|} \rightarrow 0 \quad \text{für } x \rightarrow a, \text{ d.h. } h \rightarrow 0.$$

Also ist

$$\lim_{h \rightarrow 0} \frac{r(he_i)}{h} = 0$$

und aus (8.4) folgt dann

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

Insbesondere existiert der Grenzwert in (8.4) und  $f$  ist daher partiell differenzierbar.  $\square$

**Definition 8.6.3** Sei  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Der *Gradient* von  $f$  ist

$$\text{grad } f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}.$$

Der Vektor  $d$  von Definition 8.6.1 ist also der Gradient an der Stelle  $a$ . Für eine total differenzierbare Funktion ergibt sich somit

$$f(x) = f(a) + \langle \text{grad } f(a), x - a \rangle + r(x - a) \quad (8.5)$$

wobei für den Fehler  $r(x-a)$  die Grenzwertbedingung (8.3) gilt. Diese Bedingung besagt, dass  $r(x-a)$  klein im Verhältnis zu  $x-a$  ist, wenn  $x$  nah bei  $a$  ist. Hieraus ergibt sich dann folgende Formel für die *lineare Approximation* einer Funktion mehrerer Variablen:

$$f(x) \approx f(a) + \langle \text{grad } f(a), x - a \rangle \quad \text{wenn } x \text{ nah bei } a. \quad (8.6)$$

Das ist sozusagen die mehrdimensionale Verallgemeinerung der Tangentengleichung. (Siehe dazu auch Abschnitt 8.8.) Wir können die Formel (8.6) der linearen Approximation noch verschieden umformen: Wenn wir das Skalarprodukt von (8.6) in Komponenten ausrechnen,

$$\begin{aligned} \langle \text{grad } f(a), x - a \rangle &= \left\langle \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}, \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{pmatrix} \right\rangle \\ &= \frac{\partial f}{\partial x_1}(a)(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n - a_n), \end{aligned}$$

erhalten wir

$$f(x) \approx f(a) + \frac{\partial f}{\partial x_1}(a)(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n - a_n) \quad (8.7)$$

für  $x$  nah bei  $a$ . Andererseits können wir mit den Abkürzungen

$$\Delta x = x - a, \quad \Delta f = f(x) - f(a)$$

schreiben

$$\Delta f \approx \langle \text{grad } f(a), \Delta x \rangle = \frac{\partial f}{\partial x_1}(a)\Delta x_1 + \cdots + \frac{\partial f}{\partial x_n}(a)\Delta x_n \quad (8.8)$$

für  $\Delta x$  klein.

Aus (8.8) können wir schließlich eine infinitesimale Version ableiten. Dabei werden die endlichen Differenzen  $\Delta x_i$  ersetzt durch die unendlich kleinen (infinitesimalen) Differentiale  $dx_i$  und entsprechend  $\Delta f$  durch  $df$ . Bei diesem Übergang verschwindet der Fehler in (8.8) (wegen der Bedingung (8.3)), so dass sich wieder eine Gleichheit ergibt:

$$df = \langle \text{grad } f, dx \rangle = \frac{\partial f}{\partial x_1}dx_1 + \cdots + \frac{\partial f}{\partial x_n}dx_n. \quad (8.9)$$

Dies ist das *totale Differential* von  $f$ .

Wann ist eine Funktion total differenzierbar? Der nächste Satz gibt dafür ein hinreichendes Kriterium:

**Satz 8.6.4** *Ist  $f$  stetig partiell differenzierbar (siehe Definition 8.5.2), dann ist  $f$  total differenzierbar.*

Zusammengefasst gilt also die Implikation

$$\begin{aligned} & f \text{ stetig partiell differenzierbar} \\ \Rightarrow & f \text{ total differenzierbar} \\ \Rightarrow & f \text{ partiell differenzierbar} \end{aligned}$$

Wir erläutern die neuen Begriffe anhand einer konkreten Funktion:

**Beispiel 8.6.5** Wir betrachten die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$f(x, y) = xe^{-y^2}.$$

Die partiellen Ableitungen sind

$$\frac{\partial f}{\partial x} = e^{-y^2}, \quad \text{und} \quad \frac{\partial f}{\partial y} = x \cdot e^{-y^2} \cdot (-2y) = -2xye^{-y^2}.$$

Somit ist der Gradient von  $f$

$$\text{grad } f(x, y) = \begin{pmatrix} e^{-y^2} \\ -2xye^{-y^2} \end{pmatrix}.$$

Da beide partiellen Ableitungen 1. Ordnung von  $f$  existieren und stetig sind (vergleiche Abschnitt 8.3), ist  $f$  also stetig partiell differenzierbar und damit dann auch total differenzierbar. Das totale Differential von  $f$  ist

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = e^{-y^2} dx - 2xye^{-y^2} dy.$$

Schließlich berechnen wir noch die lineare Approximation von  $f$  in einem bestimmten Punkt. Allgemein gilt

$$f(x, y) \approx f(a) + \frac{\partial f}{\partial x}(a)\Delta x + \frac{\partial f}{\partial y}(a)\Delta y.$$

Für den Punkt  $a = (\frac{1}{2}, 1)$  ergibt sich

$$\begin{aligned} f(a) &= f\left(\frac{1}{2}, 1\right) = \frac{1}{2}e^{-1}, \\ \frac{\partial f}{\partial x}\left(\frac{1}{2}, 1\right) &= e^{-1}, \quad \frac{\partial f}{\partial y}\left(\frac{1}{2}, 1\right) = -2 \cdot \frac{1}{2}e^{-1} = -e^{-1}, \\ \Delta x &= x - \frac{1}{2}, \quad \Delta y = y - 1. \end{aligned}$$

Damit ist die lineare Approximation von  $f$  in  $a$  also

$$f(x, y) \approx \frac{1}{2e} + \frac{1}{e} \left(x - \frac{1}{2}\right) - \frac{1}{e}(y - 1). \quad (8.10)$$

**Bemerkung:** Es gibt Funktionen, die zwar partiell, aber nicht total differenzierbar sind. Solche Funktionen sind dann nicht stetig partiell differenzierbar und typischerweise durch eine Fallunterscheidung definiert. Ein Beispiel einer solchen Funktion ist

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0). \end{cases}$$

Diese Funktion ist im Punkt  $(0, 0)$  partiell differenzierbar, aber nicht total differenzierbar.  $f$  ist übrigens in  $(0, 0)$  auch nicht stetig.

## 8.7 Anwendung: Fehlerrechnung

Eine Anwendung der linearen Approximation ist die Berechnung der Fortpflanzung von Messfehlern, auch *Fehlerrechnung* genannt. Gegeben sind mehrere physikalische Größen  $x_1, \dots, x_n$ , die gemessen werden. Diese Messgrößen sind jeweils mit Messfehlern  $\Delta x_1, \dots, \Delta x_n$  behaftet. Für eine weitere abhängige Größe  $y = f(x_1, \dots, x_n)$  ist dann der resultierende Fehler  $\Delta y$  gesucht. Die Funktion  $f$  ist dabei meist ein physikalisches Gesetz, das  $x_1, \dots, x_n$  mit  $y$  verknüpft. Wie plant sich also der Fehler  $\Delta x_1, \dots, \Delta x_n$  zu  $\Delta y$  fort?

Mit linearer Approximation von  $f$  gilt

$$\Delta y \approx \frac{\partial f}{\partial x_1} \Delta x_1 + \dots + \frac{\partial f}{\partial x_n} \Delta x_n.$$

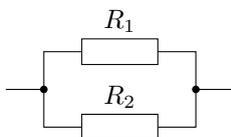


Abbildung 8.3: Parallelschaltung von Widerständen

Daraus erhält man mit Hilfe der Dreiecksungleichung als Abschätzung für den resultierenden Fehler

$$\begin{aligned} |\Delta y| &\approx \left| \frac{\partial f}{\partial x_1} \Delta x_1 + \dots + \frac{\partial f}{\partial x_n} \Delta x_n \right| \\ &\leq \left| \frac{\partial f}{\partial x_1} \right| \cdot |\Delta x_1| + \dots + \left| \frac{\partial f}{\partial x_n} \right| \cdot |\Delta x_n| \end{aligned}$$

**Beispiel 8.7.1** Wir betrachten die Parallelschaltung zweier Widerstände  $R_1$  und  $R_2$ , siehe Abbildung 8.3. Der resultierende Gesamtwiderstand  $R$  ist

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{R_1 + R_2}{R_1 R_2} \quad \Rightarrow \quad R = \frac{R_1 R_2}{R_1 + R_2}.$$

Die Werte der Einzelwiderstände sind durch Messung bekannt:

$$R_1 = (2 \pm 0.1) \text{ k}\Omega, \quad R_2 = (18 \pm 0.5) \text{ k}\Omega$$

Der Gesamtwiderstand ist damit

$$R = \frac{2 \cdot 18}{2 + 18} \text{ k}\Omega = \frac{36}{20} \text{ k}\Omega = 1.8 \text{ k}\Omega.$$

Wie groß ist der resultierende Fehler des Gesamtwiderstands? Die Funktion, die die gemessenen Größen  $R_1$  und  $R_2$  mit der resultierenden Größe  $R$  verknüpft, ist hier

$$R = f(R_1, R_2) = \frac{R_1 R_2}{R_1 + R_2}.$$

Damit folgt

$$|\Delta R| \leq \left| \frac{\partial f}{\partial R_1} \right| |\Delta R_1| + \left| \frac{\partial f}{\partial R_2} \right| |\Delta R_2|$$

Aus der Quotienten folgt für die beiden partiellen Ableitungen

$$\frac{\partial f}{\partial R_1} = \frac{R_2(R_1 + R_2) - R_1 R_2}{(R_1 + R_2)^2} = \frac{R_2^2}{(R_1 + R_2)^2}, \quad \frac{\partial f}{\partial R_2} = \frac{R_1^2}{(R_1 + R_2)^2}.$$

Einsetzen der Messwerte ergibt daraus

$$\frac{\partial f}{\partial R_1} = \frac{18^2}{20^2} = 0.9^2 = 0.81, \quad \frac{\partial f}{\partial R_2} = \frac{2^2}{20^2} = 0.1^2 = 0.01$$

sowie

$$|\Delta R_1| = 0.1 \text{ k}\Omega, \quad |\Delta R_2| = 0.5 \text{ k}\Omega.$$

Damit erhalten wir für den resultierenden Fehler

$$|\Delta R| \leq 0.81 \cdot 0.1 \text{ k}\Omega + 0.01 \cdot 0.5 \text{ k}\Omega = 0.081 \text{ k}\Omega + 0.005 \text{ k}\Omega = 0.086 \text{ k}\Omega$$

also insgesamt  $R = (1.8 \pm 0.086) \text{ k}\Omega$ .

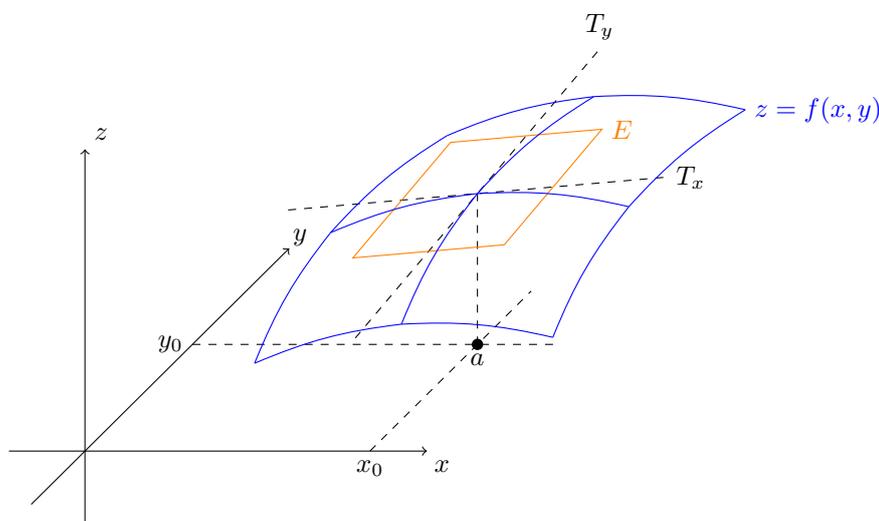


Abbildung 8.4: Die Tangentialebene an eine Funktion

## 8.8 Lineare Approximation graphisch: die Tangentialebene

Für eine Funktion von zwei Variablen  $f(x, y)$  lassen sich die partiellen Ableitungen und die lineare Approximation sehr schön graphisch veranschaulichen. Sei dazu  $a = (x_0, y_0)$  ein gegebener Punkt. Wie wir in Abschnitt 8.4 gesehen haben, ist die partielle Ableitung  $\frac{\partial f}{\partial x}$  im Punkt  $a$  die Ableitung der Funktion  $x \mapsto f(x, y_0)$  an der Stelle  $x = x_0$ , also eine gewöhnliche Ableitung einer Funktion von einer Variablen  $x$ . Wir wissen, dass die Ableitung gleich der Steigung der Tangente ist. Also ist  $\frac{\partial f}{\partial x}(a)$  die Steigung der Tangente in  $x = x_0$  an die Funktion  $x \mapsto f(x, y_0)$ , oder anders gesagt:

$$\frac{\partial f}{\partial x}(a) = \text{Steigung der Tangente } T_x \text{ an } f \text{ in } x\text{-Richtung im Punkt } a.$$

Entsprechend gilt

$$\frac{\partial f}{\partial y}(a) = \text{Steigung der Tangente } T_y \text{ an } f \text{ in } y\text{-Richtung im Punkt } a.$$

Die Tangenten  $T_x$  und  $T_y$  an  $f$  sind in Abbildung 8.4 skizziert. Der Graph der Funktion  $f$  ist dabei eine, im Allgemeinen gekrümmte, Fläche im Raum: die Fläche hat im Punkt  $(x, y)$  die Höhe  $z = f(x, y)$ .

Wir betrachten jetzt die lineare Approximation von  $f$  in  $a$ :

$$f(x, y) \approx f(a) + \frac{\partial f}{\partial x}(a)(x - x_0) + \frac{\partial f}{\partial y}(a)(y - y_0). \quad (8.11)$$

Die rechte Seite ist dabei nur abhängig von  $x$  und  $y$  (denn  $a = (x_0, y_0)$  ist fest); sie definiert eine Ebene in Koordinatenform:

$$E : z = f(a) + \frac{\partial f}{\partial x}(a)(x - x_0) + \frac{\partial f}{\partial y}(a)(y - y_0). \quad (8.12)$$

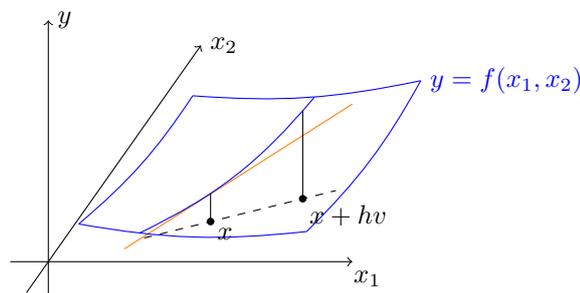


Abbildung 8.5: Zur Konstruktion der Richtungsableitung

Dies ist die Gleichung der *Tangentialebene an  $f$  in  $a$* . Die lineare Approximation (8.11) besagt dann, dass die Tangentialebene  $E$  den Graphen von  $f$  im Punkt  $a$  annähert.

Setzt man in (8.12)  $y = y_0$  bzw.  $x = x_0$ , so erhält man genau die Gleichungen der Tangenten  $T_x$  und  $T_y$ . Das heißt,  $T_x$  und  $T_y$  liegen in  $E$  und spannen daher  $E$  auf. Das ist ebenfalls in Abbildung 8.4 skizziert.

**Beispiel 8.8.1** Für das Beispiel 8.6.5 etwa können wir die Tangentialebene im Punkt  $a = (\frac{1}{2}, 1)$  direkt aus der linearen Approximation (8.10) ablesen: die Tangentialebene ist

$$E : z = \frac{1}{2e} + \frac{1}{e} \left( x - \frac{1}{2} \right) - \frac{1}{e} (y - 1).$$

**Beachte:** Die Veranschaulichung der linearen Approximation als Tangentialebene setzt eine Funktion von zwei Variablen voraus, also  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $n = 2$ . Die Tangentialebene ist eine 2-dimensionale Ebene im 3-dimensionalen Raum  $\mathbb{R}^3$ . Allgemein, d.h. bei beliebigem  $n$ , definiert die rechte Seite der linearen Approximation eine  $n$ -dimensionale „Fläche“ im  $n + 1$ -dimensionalen Raum  $\mathbb{R}^{n+1}$ .

## 8.9 Richtungsableitung

Die partielle Ableitung  $\frac{\partial f}{\partial x_i}$  gibt die Steigung einer Funktion in Richtung der Koordinate  $x_i$  an. Die Steigung der Funktion in Richtung eines *beliebigen* Vektors lässt sich mit der Richtungsableitung bestimmen.

**Definition 8.9.1** Für  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  und  $v \in \mathbb{R}^n$ ,  $v \neq 0$  heißt

$$\partial_v f(x) = \lim_{h \rightarrow 0} \frac{f(x + hv) - f(x)}{h}$$

die *Richtungsableitung* entlang (oder in Richtung von)  $v$ .

Die Richtungsableitung misst die Änderung der Funktionswerte von  $f$  entlang der Geraden von  $x$  nach  $x + hv$ , also des Kurvenstücks im Graphen von  $f$  oberhalb dieser Geraden, siehe Abbildung 8.5.

Für eine *normierte* Richtung  $|v| = 1$  ist die Richtungsableitung genau der Grenzwert des Differenzenquotienten

$$\frac{f(x + hv) - f(x)}{|(x + hv) - x|} = \frac{f(x + hv) - f(x)}{|hv|} = \frac{f(x + hv) - f(x)}{h},$$

und damit gilt für  $|v| = 1$ :

$$\partial_v f(x) = \text{Steigung von } f \text{ im Punkt } x \text{ in Richtung } v.$$

Speziell erhält man für  $v = e_i$  wieder die partielle Ableitung:  $\partial_{e_i} f(x) = \frac{\partial f}{\partial x_i}(x)$ . (Das folgt sofort durch Vergleich der Definitionen 8.4.1 und 8.9.1.)

Die praktische Berechnung der Richtungsableitung erfolgt mit durch Skalarprodukt der Richtung  $v$  mit dem Gradienten:

**Satz 8.9.2** *Ist  $f$  total differenzierbar, so gilt*

$$\partial_v f(x) = \langle \text{grad } f(x), v \rangle.$$

**Beweis.** Da  $f$  total differenzierbar ist, gilt

$$f(x + hv) = f(x) + \langle \text{grad } f(x), hv \rangle + r(hv)$$

und damit

$$\frac{f(x + hv) - f(x)}{h} = \langle \text{grad } f(x), v \rangle + \frac{r(hv)}{h}.$$

Im Grenzwert  $h \rightarrow 0$  gilt wegen (8.3)  $\frac{r(hv)}{h} \rightarrow 0$ , und es folgt

$$\partial_v f(x) = \lim_{h \rightarrow 0} \frac{f(x + hv) - f(x)}{h} = \langle \text{grad } f(x), v \rangle.$$

□

**Beachte:** Nur für normierte Richtungen  $|v| = 1$  gibt die Richtungsableitung die Steigung von  $f$  in Richtung  $v$  an. Ansonsten erhält man ein gewisses Vielfaches Steigung, denn nach dem letzten Satz gilt für  $r \in \mathbb{R} \setminus \{0\}$ :

$$\partial_{rv} f(x) = \langle \text{grad } f(x), rv \rangle = r \cdot \langle \text{grad } f(x), v \rangle = r \cdot \partial_v f(x).$$

**Beispiel 8.9.3** Wir berechnen die Richtungsableitung von

$$f(x, y) = x \cdot \ln(y) + x^2$$

im Punkt  $a = (1, 1)$  in Richtung  $v = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ . Der Gradient von  $f$  ist

$$\text{grad } f(x, y) = \begin{pmatrix} \ln(y) + 2x \\ \frac{x}{y} \end{pmatrix}.$$

Damit ist

$$\begin{aligned} \partial_v f(1, 1) &= \left\langle \text{grad } f(1, 1), \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} \ln(1) + 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\rangle = -2 + 3 = 1 \end{aligned}$$

Da  $|v| = \sqrt{10} \neq 1$ , ist dies nicht die Steigung von  $f$  in Richtung von  $v$ . Die Steigung erhält man mit der normierten Richtung (Einheitsvektor)  $e_v = \frac{1}{|v|}v$ :

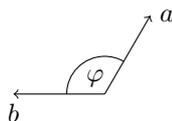
$$\partial_{\frac{v}{|v|}} f(1, 1) = \frac{1}{|v|} \cdot \partial_v f(1, 1) = \frac{1}{\sqrt{10}} \cdot 1 = \frac{1}{\sqrt{10}}.$$

## 8.10 Geometrische Interpretation des Gradienten

Der Gradient einer Funktion ist ein Vektor. Mit Hilfe von Satz 8.9.2 können wir jetzt sagen, was dieser Vektor geometrisch bedeutet. Wir erinnern zuerst an den Zusammenhang des Skalarprodukts mit dem Winkel zwischen Vektoren: Für zwei Vektoren  $a, b \in \mathbb{R}^n$  gilt

$$\langle a, b \rangle = a \cdot b = |a| \cdot |b| \cos \varphi$$

wobei  $\varphi = \varphi(a, b) \in [0, \pi]$  der Winkel zwischen den Vektoren  $a$  und  $b$  ist:



Aus Satz 8.9.2 folgt dann

$$\partial_v f(x) = \langle \text{grad } f(x), v \rangle = |\text{grad } f(x)| \cdot |v| \cos \varphi$$

wobei  $\varphi$  der Winkel zwischen  $\text{grad } f(x)$  und  $v$  ist. Für eine normierte Richtung  $|v| = 1$  erhalten wir somit

$$\partial_v f(x) = \text{Steigung von } f \text{ in Richtung } v = |\text{grad } f(x)| \cos \varphi$$

Da  $\cos \varphi$  nur Werte zwischen 1 und  $-1$  annimmt, folgt

$$-|\text{grad } f(x)| \leq \partial_v f(x) \leq |\text{grad } f(x)|$$

Der maximale Wert  $|\text{grad } f(x)|$  für  $\partial_v f(x)$  ergibt sich für  $\cos \varphi = 1$ , also  $\varphi = 0$ ; der minimale Wert wird für  $\cos \varphi = -1 \Leftrightarrow \varphi = \pi$  angenommen. (Man beachte  $\varphi \in [0, \pi]$ !) Weiter gilt  $\varphi = 0$  wenn  $v$  in Richtung von  $\text{grad } f(x)$  weist, und  $\varphi = \pi$  wenn die Richtung von  $v$  entgegengesetzt ist. Also:

$$\partial_v f(x) = \begin{cases} |\text{grad } f(x)| & \text{maximal} \Leftrightarrow \varphi = 0 \Leftrightarrow v \text{ in Richtung } \text{grad } f(x) \\ -|\text{grad } f(x)| & \text{minimal} \Leftrightarrow \varphi = \pi \Leftrightarrow v \text{ in Richtung } -\text{grad } f(x) \end{cases}$$

Die Steigung von  $f$  ist also am größten in Richtung von  $\text{grad } f(x)$  und am kleinsten (negativ!) in Richtung  $-\text{grad } f(x)$ . Anders ausgedrückt:

- $\text{grad } f(x)$  ist die Richtung des steilsten *Anstiegs* von  $f$  im Punkt  $x$
- $-\text{grad } f(x)$  ist die Richtung des steilsten *Abstiegs* von  $f$  im Punkt  $x$

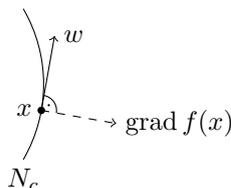


Abbildung 8.6: Der Gradient im Punkt einer Niveaumenge

## 8.11 Niveaulinien

Angesehen vom Graphen sind Niveaulinien oder – allgemeiner – Niveaumengen eine Möglichkeit zur graphischen Darstellung von Funktionen mehrerer Variablen. Für eine Funktion  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  und einen Wert  $c \in \mathbb{R}$  heißt

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

*Niveaumenge* von  $f$  zum Wert  $c$ . Die Niveaumenge  $N_c$  enthält also alle Punkte  $x$  mit dem gleichen Funktionswert  $f(x) = c$ .

Niveaumengen sind Teilmengen von  $\mathbb{R}^n$ . Bei einer Funktion zweier Variablen  $f(x, y)$ , d.h. für  $n = 2$ , ist  $N_c \subset \mathbb{R}^2$  eine Menge in der Ebene, in der Regel eine Kurve. Man spricht daher in diesem Fall von *Niveaulinien*. Der Graph ergibt eine dreidimensionale Darstellung der Funktion  $f(x, y)$ : für jeden Punkt  $(x, y)$  wird der Funktionswert  $z = f(x, y)$  als Höhe des Graphen eingezeichnet. Die Niveaulinie  $N_c$  enthält daher die Punkte gleicher Höhe der Funktion  $z = f(x, y) = c$ . (Höhenlinien, wie bei einer Landkarte! Wir illustrieren das im nächsten Beispiel.)

Gradient und Niveaumengen hängen allgemein wie folgt zusammen:

**Satz 8.11.1** *In jedem Punkt  $x \in N_c$  steht  $\text{grad } f(x)$  senkrecht auf der Niveaumenge  $N_c$ , d.h.*

$$x \in N_c \quad \Rightarrow \quad \text{grad } f(x) \perp N_c$$

**Beweis.** Sei  $x \in N_c$ , d.h.  $f(x) = c$ . Sei  $w \in \mathbb{R}^n$  tangential zu  $N_c$  im Punkt  $x$ , siehe Abbildung 8.6. Da auf der Niveaulinie  $N_c$  die Funktion  $f(x)$  konstant gleich  $c$  ist, gilt für die Richtungsableitung in  $x$  entlang  $N_c$ , also in Richtung  $w$ ,  $\partial_w f(x) = 0$ . Somit ist

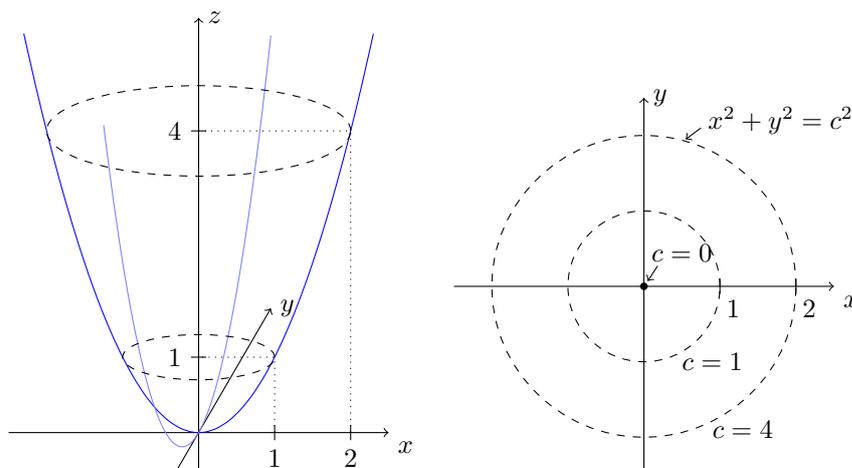
$$\partial_w f(x) = \langle \text{grad } f(x), w \rangle = 0,$$

und das heißt  $\text{grad } f(x) \perp w$ . □

**Beispiel 8.11.2** Wir untersuchen die Funktion

$$f(x, y) = x^2 + y^2.$$

Der Graph von  $f$  ist eine um die  $z$ -Achse rotierte Parabel. Denn für  $y = 0$ , also in Richtung der  $x$ -Achse, ist  $z = f(x, 0) = x^2$ ; das ist die Normalparabel in der  $(x, z)$ -Ebene. Entsprechend erhält man in Richtung der  $y$ -Achse  $z = f(0, y) =$

Abbildung 8.7: Graph und Niveaulinien von  $f(x, y) = x^2 + y^2$ 

$y^2$ , die Normalparabel in der  $(y, z)$ -Ebene. Das ist in Abbildung 8.7 dargestellt. Die Niveaulinien von  $f$  sind Kreise um den Ursprung. Zum Beispiel ist für  $c = 4$

$$N_4 = \{(x, y) \mid x^2 + y^2 = 4\} = \{(x, y) \mid \sqrt{x^2 + y^2} = 2\}$$

der Kreis mit Radius 2. Da die Niveaulinien den Linien gleicher Höhe im Graphen entsprechen, sieht man so, dass der Graph von  $f$  tatsächlich die um die  $z$ -Achse rotierte Parabel ist, siehe nochmal Abbildung 8.7. Für  $c = 0$  ergibt sich  $N_0 = \{(0, 0)\}$ ; die „Niveaulinie“ besteht hier also nur aus einem einzelnen Punkt. (Und bei  $c < 0$  ist  $N_c = \emptyset$ .)

Wir betrachten jetzt noch den Gradienten für dieses Beispiel. Es ist

$$\text{grad } f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

Wir untersuchen den Gradienten im Punkt  $a = (-1, 1)$ . Es gilt  $f(a) = f(-1, 1) = (-1)^2 + 1^2 = 2$ . Somit ist  $a \in N_2$ . Die Niveaumenge  $N_2$  ist gegeben durch

$$N_2 : f(x, y) = x^2 + y^2 = c = 2 \Leftrightarrow \sqrt{x^2 + y^2} = \sqrt{2},$$

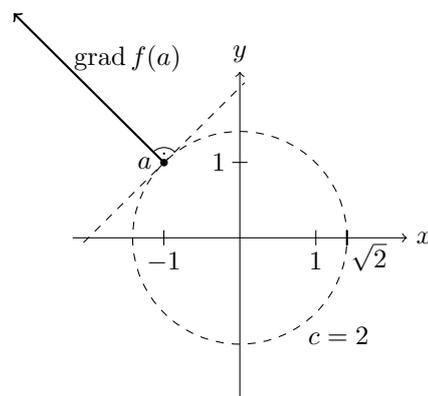
es ist also der Kreis mit Radius  $\sqrt{2}$ . Im Punkt  $a$  ist der Gradient

$$\text{grad } f(a) = \text{grad } f(-1, 1) = \begin{pmatrix} -2 \\ 2 \end{pmatrix}.$$

Niveaumenge und Gradient sind in Abbildung 8.8 skizziert; man sieht, wie der Gradient auf  $N_2$  senkrecht steht.

Der Gradient zeigt in Richtung des steilsten Anstiegs von  $f$ . Auch das sieht man in der Skizze 8.8, denn nach Außen hin liegen ja die Niveaulinien mit größerem Funktionswert  $c$ , nach Innen wird der  $c$  kleiner (vergleiche Abbildung 8.7). Die maximale Steigung von  $f$  im Punkt  $a$ , also die Steigung in Richtung des Gradienten ist

$$|\text{grad } f(a)| = \left| \begin{pmatrix} -2 \\ 2 \end{pmatrix} \right| = 2\sqrt{2}.$$

Abbildung 8.8: Gradient in  $a = (-1, 1)$  aus Beispiel 8.11.2

Schließlich ist die Richtung des steilsten *Abstiegs* in  $a$  gleich

$$-\text{grad } f(a) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

## 8.12 Kurven im Raum

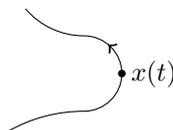
Für die weitere Untersuchung von Funktionen mehrerer Variablen benötigen wir als Hilfsmittel Kurven im Raum oder der Ebene. Kurven tauchen aber natürlich selbst in Anwendungen auf, zum Beispiel als Kurve, auf der ich ein Körper durch den Raum bewegt. Auf einer Kurve werden die Punkte nacheinander durchlaufen.

**Definition 8.12.1** Sei  $I \subset \mathbb{R}$  ein Intervall ( $I = \mathbb{R}$  ist erlaubt). Eine Funktion

$$x : I \rightarrow \mathbb{R}^n, \quad t \mapsto x(t),$$

heißt *Kurve* (oder auch *Weg*) im  $\mathbb{R}^n$ .

Die Funktion  $x(t)$  beschreibt also die Kurvenpunkte in Abhängigkeit vom Kurvenparameter  $t$ ; oft ist  $t$  die Zeit. Die Richtung, in der die Punkte der Kurve durchlaufen werden, ist die Richtung mit wachsendem  $t$ .



**Beispiel 8.12.2** (a) Die Kurve

$$x : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad x(t) = \begin{pmatrix} r \cos(t) \\ r \sin(t) \end{pmatrix}$$

ist der Kreis mit Radius  $r$  in der Ebene  $\mathbb{R}^2$ , gegen den Uhrzeigersinn durchlaufen.

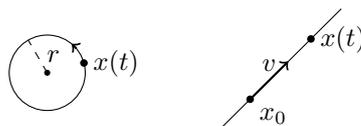


Abbildung 8.9: Die Kurven aus Beispiel 8.12.2

(b) Die Kurve

$$x : \mathbb{R} \rightarrow \mathbb{R}^n, \quad x(t) = x_0 + tv$$

ist die Gerade durch  $x_0 \in \mathbb{R}^n$  in Richtung  $v \in \mathbb{R}^n$ .

Beide Kurven sind in Abbildung 8.9 dargestellt.

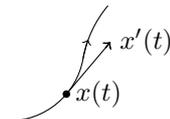
Die Ableitung einer Kurve ist wie immer durch den Grenzwert eines Differenzenquotienten definiert:

$$x'(t) = \lim_{h \rightarrow 0} \frac{1}{h} (x(t+h) - x(t)).$$

Die Berechnung erfolgt komponentenweise:

$$x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} \Rightarrow x'(t) = \begin{pmatrix} x'_1(t) \\ \vdots \\ x'_n(t) \end{pmatrix}.$$

Die Ableitung  $x'(t)$  ist ein Vektor, der im Punkt  $x(t)$  tangential zur Kurve in Laufrichtung liegt:



Physikalisch ist  $x'(t)$  der Geschwindigkeitsvektor beim Durchlaufen der Kurve.

Für eine Funktion  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  und eine Kurve  $x : I \rightarrow \mathbb{R}^n$  die im Definitionsbereich von  $f$  verläuft, d.h.  $x(t) \in G$ , kann man die Verkettung

$$\varphi = f \circ x : I \rightarrow \mathbb{R}, \quad \varphi(t) = f(x(t))$$

bilden. Dadurch wird die Funktion  $f$  sozusagen entlang der Kurve  $x(t)$  „abgefahren“. Die Ableitung  $\varphi'(t)$  beschreibt damit die Änderung von  $f(x)$  entlang der Kurve  $x(t)$ . Für die Ableitung der Verkettung  $\varphi = f \circ x$  gilt wie im Fall von Funktionen einer Variablen eine Kettenregel:

**Satz 8.12.3 (Kettenregel für Kurven)** Sind  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  und  $x : I \rightarrow \mathbb{R}^n$  differenzierbar<sup>1</sup> und gilt  $x(t) \in G$  für alle  $t \in I$ , dann ist  $\varphi = f \circ x : I \rightarrow \mathbb{R}$  differenzierbar mit Ableitung

$$\begin{aligned} \varphi'(t) &= \langle \text{grad } f(x(t)), x'(t) \rangle \\ &= \frac{\partial f}{\partial x_1}(x(t)) x'_1(t) + \dots + \frac{\partial f}{\partial x_n}(x(t)) x'_n(t). \end{aligned}$$

<sup>1</sup>Hier ist gemeint, dass  $f$  total differenzierbar ist. Bei einer Funktion mehrerer Variablen ist im folgenden mit differenzierbar immer (mindestens) total differenzierbar gemeint.

Äquivalent können wir das auch als

$$\frac{\partial \varphi}{\partial t} = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt}$$

schreiben, was an das totale Differential (8.9),  $df = \frac{\partial f}{\partial x_1} dx_1 + \dots + \frac{\partial f}{\partial x_n} dx_n$ , erinnert.

**Beweis.** Nach der Formel (8.5) für die totale Differenzierbarkeit von  $f$  in  $x(t_0)$  gilt

$$\frac{f(x(t)) - f(x(t_0))}{t - t_0} = \left\langle \text{grad } f(x(t_0)), \frac{x(t) - x(t_0)}{t - t_0} \right\rangle + \frac{r(x(t) - x(t_0))}{t - t_0}.$$

Für  $t \rightarrow t_0$  ist

$$\frac{f(x(t)) - f(x(t_0))}{t - t_0} \rightarrow \varphi'(t), \quad \frac{x(t) - x(t_0)}{t - t_0} \rightarrow x'(t_0)$$

sowie

$$\frac{r(x(t) - x(t_0))}{t - t_0} = \frac{r(x(t) - x(t_0))}{|x(t) - x(t_0)|} \cdot \frac{|x(t) - x(t_0)|}{t - t_0} \rightarrow 0 \cdot |x'(t_0)| = 0$$

und damit folgt  $\varphi'(t_0) = \langle \text{grad } f(x(t_0)), x'(t_0) \rangle$ .  $\square$

**Beispiel 8.12.4** Sei  $\varphi(t) = f(x(t))$  mit

$$f(x_1, x_2) = x_1 x_2^2, \quad x(t) = \begin{pmatrix} t \\ 1 - t^2 \end{pmatrix} = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

Wir berechnen  $\varphi'(t)$  mit der Kettenregel: Es gilt

$$\text{grad } f(x_1, x_2) = \begin{pmatrix} x_2^2 \\ 2x_1 x_2 \end{pmatrix}, \quad x'(t) = \begin{pmatrix} 1 \\ -2t \end{pmatrix}.$$

Einsetzen der Kurve  $x(t)$  in  $\text{grad } f(x_1, x_2)$ , also von  $x_1 = x_1(t) = t$  und  $x_2 = x_2(t) = 1 - t^2$  liefert

$$\text{grad } f(x(t)) = \begin{pmatrix} (1 - t^2)^2 \\ 2t(1 - t^2) \end{pmatrix}$$

und damit

$$\varphi'(t) = \langle \text{grad } f(x(t)), x'(t) \rangle = \left\langle \begin{pmatrix} (1 - t^2)^2 \\ 2t(1 - t^2) \end{pmatrix}, \begin{pmatrix} 1 \\ -2t \end{pmatrix} \right\rangle = (1 - t^2)^2 - 4t^2(1 - t^2),$$

was wir hier nicht weiter ausrechnen.

## 8.13 Taylorentwicklung für mehrere Variablen

Wir leiten jetzt Formeln für die Taylorentwicklung von Funktionen mehrerer Variablen her. Sei dazu  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$   $(m + 1)$ -mal stetig partiell differenzierbar und  $G$  offen. Wir betrachten den festen Entwicklungspunkt  $a \in G$  und einen verschobenen Punkt  $a + h$  gegeben durch den Vektor  $h \in \mathbb{R}^n$ , so dass  $a + h$  und das ganze Geradenstück von  $a$  nach  $a + h$  zu  $G$  gehören, d.h.

$$a + th \in G \quad \text{für alle } t \in [0, 1].$$

Das ist in Abbildung 8.10 illustriert.

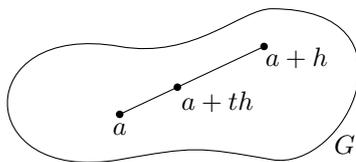


Abbildung 8.10: Setup zur Taylorentwicklung

**Satz 8.13.1** Sei  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$   $(m+1)$ -mal stetig partiell differenzierbar und  $a, h$  wie oben. Dann ist die Taylorentwicklung  $m$ . Ordnung von  $f$  im Punkt  $a$

$$f(a+h) = \sum_{k=0}^m \frac{1}{k!} \partial_h^k f(a) + R_m(h) \quad (8.13)$$

mit dem Restglied

$$R_m(h) = \frac{1}{(m+1)!} \partial_h^{m+1} f(a + \xi h) \quad \text{für ein } \xi \in [0, 1]. \quad (8.14)$$

Dabei ist  $\partial_h^k f$  die  $k$ -fache Richtungsableitung von  $f$  in Richtung  $h$ .

**Beweis.** Sei  $x(t) = a+th$  und  $\varphi(t) = f(x(t)) = f(a+th)$ . Die Taylorentwicklung (4.3) für  $\varphi$  in  $t_0 = 0$  ist

$$\varphi(t) = \sum_{k=0}^m \frac{1}{k!} \varphi^{(k)}(0) t^k + R_m(h),$$

für das Restglied in Lagrange-Form (4.4) erhält man

$$R_m(h) = \frac{1}{(m+1)!} \varphi^{(m+1)}(\xi) t^{m+1}, \quad 0 \leq \xi \leq t.$$

Wir berechnen die Ableitungen von  $\varphi(t)$  mit der Kettenregel 8.12.3:

$$\begin{aligned} \varphi'(t) &= \langle \text{grad } f(x(t)), x'(t) \rangle = \langle \text{grad } f(x(t)), h \rangle = \partial_h f(x(t)), \\ \varphi''(t) &= \langle \text{grad}(\partial_h f)(x(t)), h \rangle = \partial_h(\partial_h f)(x(t)) = \partial_h^2 f(x(t)), \quad \text{u.s.w.} \end{aligned}$$

Damit folgt dann

$$f(a+h) = \varphi(1) = \sum_{k=0}^m \frac{1}{k!} \partial_h^k f(x(0)) + R_m(h) = \sum_{k=0}^m \frac{1}{k!} \partial_h^k f(a) + R_m(h)$$

mit

$$R_m(h) = \frac{1}{(m+1)!} \partial_h^{m+1} f(x(\xi)) = \frac{1}{(m+1)!} \partial_h^{m+1} f(x(a+\xi h)), \quad 0 \leq \xi \leq 1.$$

□

In Anwendungen verwendet man oft folgende alternative Formel für die Taylorentwicklung:

$$f(a+h) = \sum_{k=0}^m \sum_{\substack{\alpha_1, \dots, \alpha_n \geq 0 \\ \alpha_1 + \dots + \alpha_n = k}} \frac{\partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n} f(a)}{\alpha_1! \dots \alpha_n!} h_1^{\alpha_1} \dots h_n^{\alpha_n} + R_m(h) \quad (8.15)$$

wobei

$$h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \quad \text{und} \quad \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n} f(a) = \left( \frac{\partial}{\partial x_1} \right)^{\alpha_1} \dots \left( \frac{\partial}{\partial x_n} \right)^{\alpha_n} f(a).$$

Die Formel (8.15) ergibt sich aus (8.13) durch Ausrechnen der mehrfachen Richtungsableitungen. Wir führen das hier für  $n = 2$  vor: aus

$$\partial_h f = \langle \text{grad } f, h \rangle = \left\langle \begin{pmatrix} \partial_{x_1} f \\ \partial_{x_2} f \end{pmatrix}, \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\rangle = h_1 \partial_{x_1} f + h_2 \partial_{x_2} f = (h_1 \partial_{x_1} + h_2 \partial_{x_2}) f$$

ergibt sich mit dem allgemeinen binomischen Satz 1.6.1:

$$\begin{aligned} \frac{1}{k!} \partial_h^k f(a) &= \frac{1}{k!} (h_1 \partial_{x_1} + h_2 \partial_{x_2})^k f(a) \\ &= \frac{1}{k!} \sum_{j=0}^k \frac{k!}{j!(k-j)!} (h_1 \partial_{x_1})^{k-j} (h_2 \partial_{x_2})^j f(a) \\ &= \sum_{\substack{j, l \geq 0 \\ j+l=k}} \frac{1}{j! l!} \partial_{x_1}^l \partial_{x_2}^j f(a) h_1^l h_2^j \quad (l = k - j) \end{aligned}$$

**Beispiel 8.13.2** Sei

$$f(x, y) = x^2 \sin y.$$

Wir berechnen die Taylorentwicklung von  $f$  bis zur Ordnung  $m = 3$  in  $a = (1, 0)$ :

$$\begin{aligned} f(a+h) &= f(a) \\ &+ \partial_x f(a) h_1 + \partial_y f(a) h_2 \\ &+ \frac{1}{2!} \partial_x^2 f(a) h_1^2 + \frac{1}{1!1!} \partial_x \partial_y f(a) h_1 h_2 + \frac{1}{2!} \partial_y^2 f(a) h_2^2 \\ &+ \frac{1}{3!} \partial_x^3 f(a) h_1^3 + \frac{1}{2!1!} \partial_x^2 \partial_y f(a) h_1^2 h_2 + \frac{1}{1!2!} \partial_x \partial_y^2 f(a) h_1 h_2^2 + \frac{1}{3!} \partial_y^3 f(a) h_2^3 \\ &+ R_3(h) \end{aligned}$$

Die partiellen Ableitungen sind

$$\begin{aligned} \partial_x f &= 2x \sin y, & \partial_y f &= x^2 \cos y, \\ \partial_x^2 f &= 2 \sin y, & \partial_x \partial_y f &= 2x \cos y, & \partial_y^2 f &= -x^2 \sin y, \\ \partial_x^3 f &= 0, & \partial_x^2 \partial_y f &= 2 \cos y, & \partial_x \partial_y^2 f &= -2x \sin y, & \partial_y^3 f &= -x^2 \cos y. \end{aligned}$$

Einsetzen von  $a = (1, 0)$ :

$$\begin{aligned} f(a) &= 1^2 \cdot \sin(0) = 0 \\ \partial_x f(a) &= 2 \cdot 1 \cdot \sin(0) = 0, \quad \partial_y f(a) = 1^2 \cdot \cos(0) = 1, \\ \partial_x^2 f(a) &= 0, \quad \partial_x \partial_y f(a) = 2, \quad \partial_y^2 f(a) = 0, \\ \partial_x^3 f(a) &= 0, \quad \partial_x^2 \partial_y f(a) = 2, \quad \partial_x \partial_y^2 f(a) = 0, \quad \partial_y^3 f(a) = -1 \end{aligned}$$

Somit ist die Taylorentwicklung

$$\begin{aligned} f(a+h) &= f(1+h_1, h_2) = h_2 + \frac{2}{1!1!} h_1 h_2 + \frac{2}{2!1!} h_1^2 h_2 + \frac{-1}{3!} h_2^3 + R_3(h) \\ &= h_2 + 2h_1 h_2 + h_1^2 h_2 - \frac{1}{6} h_2^3 + R_3(h). \end{aligned}$$

## 8.14 Die Hessematrix

So wie der Gradient alle partiellen Ableitungen 1. Ordnung zusammenfasst, fasst die Hessematrix die partiellen Ableitungen 2. Ordnung zusammen.

Für eine zweimal stetig partiell differenzierbare Funktion  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  heißt

$$H_f(x) = \begin{pmatrix} f_{x_1 x_1}(x) & f_{x_1 x_2}(x) & \cdots & f_{x_1 x_n}(x) \\ f_{x_2 x_1}(x) & f_{x_2 x_2}(x) & \cdots & f_{x_2 x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n x_1}(x) & f_{x_n x_2}(x) & \cdots & f_{x_n x_n}(x) \end{pmatrix} \quad (8.16)$$

die *Hessematrix* von  $f$ . Dabei benutzen wir die Abkürzung

$$f_{x_i x_k} = \partial_{x_k} \partial_{x_i} f = \frac{\partial}{\partial x_k} \frac{\partial f}{\partial x_i},$$

siehe Abschnitt 8.4 partielle Ableitung.

Nach dem Satz von Schwarz 8.5.3 gilt  $f_{x_i x_k} = f_{x_k x_i}$  und damit ist die Hessematrix  $H_f(x)$  symmetrisch:

$$H_f(x) = H_f(x)^T.$$

Die Hessematrix ist so aufgebaut, dass ihre Spalten nacheinander alle partiellen Ableitungen des Gradienten enthalten:

$$\begin{aligned} \text{grad } f(x) &= \begin{pmatrix} f_{x_1} \\ \vdots \\ f_{x_n} \end{pmatrix} \Rightarrow \frac{\partial}{\partial x_k} \text{grad } f(x) = \begin{pmatrix} f_{x_1 x_k} \\ \vdots \\ f_{x_n x_k} \end{pmatrix} \\ \Rightarrow H_f(x) &= \left( \frac{\partial}{\partial x_1} \text{grad } f, \frac{\partial}{\partial x_2} \text{grad } f, \dots, \frac{\partial}{\partial x_n} \text{grad } f \right) \end{aligned}$$

Das ist hilfreich bei der Berechnung von  $H_f(x)$ !

**Beispiel 8.14.1** Für  $f(x, y) = y^3 - xy^2$  ist

$$\text{grad } f(x, y) = \begin{pmatrix} -y^2 \\ 3y^2 - 2xy \end{pmatrix}.$$

Daraus folgt dann

$$H_f(x, y) = \begin{pmatrix} 0 & -2y \\ -2y & 6y - 2x \end{pmatrix}.$$

Die erste Spalte ergibt sich durch partielle Ableitung des Gradienten nach  $x$ , die zweite Spalte durch Ableitung von  $\text{grad } f$  nach  $y$ .

Mit dem Gradienten und der Hessematrix erhält man eine weitere alternative Formel für die Taylorentwicklung:

**Satz 8.14.2** Sei  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Die Taylorentwicklung 1. Ordnung ist

$$f(a + h) = f(a) + \langle \text{grad } f(a), h \rangle + R_1(h) \quad \text{mit} \quad R_1(h) = \frac{1}{2} \langle h, H_f(a + \xi h) h \rangle.$$

Für die Entwicklung 2. Ordnung gilt

$$f(a + h) = f(a) + \langle \text{grad } f(a), h \rangle + \frac{1}{2} \langle h, H_f(a) h \rangle + R_2(h).$$

**Beweis.** Mit  $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$  gilt

$$\partial_h f(x) = \langle \text{grad } f(x), h \rangle = (h_1 \partial_{x_1} + \cdots + h_n \partial_{x_n}) f(x)$$

$$\partial_h^2 f(x) = (h_1 \partial_{x_1} + \cdots + h_n \partial_{x_n})^2 f(x) = \sum_{i,k=1}^n h_i h_k \partial_{x_i} \partial_{x_k} f(x)$$

$$= \sum_{k=0}^n h_k (h_1 \partial_{x_1} \partial_{x_k} f(x) + \cdots + h_n \partial_{x_n} \partial_{x_k} f(x))$$

$$= \left\langle \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}, \begin{pmatrix} \vdots \\ h_1 f_{x_k x_1} + \cdots + h_n f_{x_k x_n} \\ \vdots \end{pmatrix} \right\rangle = \langle h, H_f(x) h \rangle = h^T H_f(x) h$$

□

**Beispiel 8.14.3** Für die Funktion des letzten Beispiels  $f(x, y) = y^3 - xy^2$  berechnen wir die Taylorentwicklung 2. Ordnung in  $a = (1, 1)$ : Es ist

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

$$\text{grad } f(x, y) = \begin{pmatrix} -y^2 \\ 3y^2 - 2xy \end{pmatrix} \Rightarrow \text{grad } f(a) = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

$$H_f(x, y) = \begin{pmatrix} 0 & -2y \\ -2y & 6y - 2x \end{pmatrix} \Rightarrow H_f(a) = \begin{pmatrix} 0 & -2 \\ -2 & 4 \end{pmatrix}.$$

Damit ist die Taylorentwicklung 2. Ordnung

$$f(a + h) = \left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\rangle + \frac{1}{2} \left\langle \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\rangle + R_2(h).$$

Wir überlassen die Berechnung der Matrix- und Skalarprodukte dem Leser.

## 8.15 Lokale Extrema, kritische Punkte

Wir benutzen jetzt unser Wissen über Differentialrechnung im  $\mathbb{R}^n$ , um Funktionen von mehreren Variablen auf lokale Extremwerte zu untersuchen.

**Definition 8.15.1** Sei  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Ein Punkt  $a \in G$  heißt

- *lokales Minimum* von  $f$  wenn es eine Umgebung  $U_\varepsilon(a)$  von  $a$  gibt so dass  $f(a) \leq f(x)$  für alle  $x \in U_\varepsilon(a)$ ;
- *lokales Maximum* von  $f$  wenn es eine Umgebung  $U_\varepsilon(a)$  von  $a$  gibt so dass  $f(a) \geq f(x)$  für alle  $x \in U_\varepsilon(a)$ .

Analog zum eindimensionalen Fall erhält man mit den partiellen Ableitungen 1. Ordnung ein notwendiges Kriterium für Extremwerte:

**Satz 8.15.2 (notwendiges Kriterium)** Für die differenzierbare Funktion  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  gilt

$$a \in G \text{ lokales Extremum} \Rightarrow \text{grad } f(a) = 0.$$

**Beweis.** Betrachte  $\varphi(t) = f(a + te_i) = f(x(t))$  mit  $x(t) = a + te_i$  wobei  $e_i$  der  $i$ . Standard-Einheitsvektor ist. Hat  $f$  in  $a$  ein lokales Extremum, so hat  $\varphi$  in  $t = 0$  ein lokales Extremum. Aus dem notwendigen Kriterium in einer Variablen und der mehrdimensionalen Kettenregel 8.12.3 folgt dann

$$0 = \varphi'(0) = \langle \text{grad } f(x(0)), x'(0) \rangle = \langle \text{grad } f(a), e_i \rangle = \frac{\partial f}{\partial x_i}(a)$$

Da  $i$  beliebig ist, folgt  $\text{grad } f(a) = 0$ . □

**Definition 8.15.3** Die Punkte  $a \in G$  mit  $\text{grad } f(a) = 0$  heißen *kritische Punkte* der Funktion  $f$ .

Die kritischen Punkte sind also die *möglichen* Extrema der Funktion.

Bei der Berechnung der kritischen Punkte ergibt sich ein – im Allgemeinen nicht-lineares – Gleichungssystem. Im Gegensatz zu linearen Gleichungssystemen gibt es hierfür leider kein allgemeingültiges Rechenschema. Es treten allerdings bestimmte typische Fälle auf.

**Beispiel 8.15.4** Wir berechnen die kritischen Punkte der Funktionen.

(a)

$$f(x, y) = \frac{1}{3}x^3 + xy^2 - x$$

Der Gradient ist

$$\text{grad } f(x, y) = \begin{pmatrix} x^2 + y^2 - 1 \\ 2xy \end{pmatrix}.$$

Zur Berechnung der kritischen Punkte setzen wir  $\text{grad } f$  gleich Null und versuchen, das entstehende Gleichungssystem zu lösen:

$$\text{grad } f(x, y) = \begin{pmatrix} x^2 + y^2 - 1 \\ 2xy \end{pmatrix} = 0 \Leftrightarrow \begin{cases} x^2 + y^2 - 1 = 0 & \text{(I)} \\ 2xy = 0 & \text{(II)} \end{cases}$$

Die zweite Gleichung ist ein Produkt, das Null ist. Daher muss einer der beiden Faktoren Null sein:

$$(II) \Leftrightarrow xy = 0 \Leftrightarrow x = 0 \vee y = 0$$

Wir haben also zwei Möglichkeiten ( $x = 0$  oder  $y = 0$ ) und führen daher eine *Fallunterscheidung* durch:

1. Fall  $x = 0$ . Einsetzen von  $x = 0$  in (I) gibt  $y^2 - 1 = 0 \Rightarrow y = \pm 1$ .  
 $\Rightarrow$  kritische Punkte  $(0, 1)$  und  $(0, -1)$
2. Fall  $y = 0$ . Einsetzen in (I) gibt  $x^2 - 1 = 0 \Rightarrow x = \pm 1$   
 $\Rightarrow$  kritische Punkte  $(1, 0)$  und  $(-1, 0)$

Insgesamt erhalten wir also:

$f$  hat die kritischen Punkte  $(1, 0)$ ,  $(-1, 0)$ ,  $(0, 1)$ ,  $(0, -1)$

(b)

$$f(x, y) = xy - y^2 - \frac{1}{2}x^4$$

$$\text{grad } f(x, y) = \begin{pmatrix} y - 2x^3 \\ x - 2y \end{pmatrix} = 0 \Rightarrow \begin{cases} y - 2x^3 = 0 & \text{(I)} \\ x - 2y = 0 & \text{(II)} \end{cases}$$

Hier können wir Gleichung (II) nach  $x$  auflösen und das dann in (I) einsetzen. Man könnte hier genauso gut (I) nach  $y$  auflösen und dies in (II) einsetzen. Wir beginnen aber mit (II) weil diese Gleichung einfacher ist. Oft macht es Sinn, mit der einfacheren Gleichung zu beginnen.

$$(II) \Leftrightarrow x = 2y \tag{8.17}$$

Einsetzen in (I):

$$y - 2x^3 = y - 2 \cdot (2y)^3 = y - 16y^3 = 0 \Rightarrow y(1 - 16y^2) = 0$$

Wir haben hier wieder ein Produkt, das Null ist. (Sowas ist immer gut!) Also gilt

$$y = 0 \vee 1 - 16y^2 = 0.$$

Die zweite dieser Gleichungen wiederum gibt

$$1 - 16y^2 = 0 \Leftrightarrow y^2 = \frac{1}{16} \Leftrightarrow y = \pm \frac{1}{4}$$

Wir erhalten also drei mögliche Werte (Fälle) für  $y$ . Durch Einsetzen in  $x = 2y$  erhalten wir dann die entsprechenden  $x$  Werte. Wichtig: Für jeden  $y$ -Wert ist hier nur *genau ein* Wert für  $x$  möglich. Denn (II) muss ja für die kritischen Punkte gelten, und das impliziert (8.17), was  $x$  festlegt.

- (a)  $y = 0 \xrightarrow{x=2y} x = 2 \cdot 0 = 0 \Rightarrow$  kritischer Punkt  $(0, 0)$
- (b)  $y = \frac{1}{4} \Rightarrow x = 2 \cdot \frac{1}{4} = \frac{1}{2} \Rightarrow$  kritischer Punkt  $(\frac{1}{2}, \frac{1}{4})$
- (c)  $y = -\frac{1}{4} \Rightarrow x = 2 \cdot (-\frac{1}{4}) = -\frac{1}{2} \Rightarrow$  kritischer Punkt  $(-\frac{1}{2}, -\frac{1}{4})$

Zusammen also:

$f$  hat die kritischen Punkte  $(0, 0)$ ,  $(\frac{1}{2}, \frac{1}{4})$ ,  $(-\frac{1}{2}, -\frac{1}{4})$

## 8.16 Hinreichendes Kriterium für Extrema

Wie bei Funktionen von einer Variablen erhalten wir über die zweite Ableitung ein hinreichendes Kriterium für Extrema. In der Bedingung tritt die Hessematrix auf, die ja alle partielle Ableitungen zweiter Ordnung zusammenfasst. Anders als im Fall einer Variablen treten hier im hinreichenden Kriterium nicht nur lokale Extrema auf, sondern auch Sattelpunkte.

**Satz 8.16.1** Sei  $a$  ein kritischer Punkt von  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Dann gilt:

$$\begin{aligned} H_f(a) \text{ positiv definit} &\Rightarrow a \text{ lokales Minimum} \\ H_f(a) \text{ negativ definit} &\Rightarrow a \text{ lokales Maximum} \\ H_f(a) \text{ indefinit} &\Rightarrow a \text{ Sattelpunkt} \end{aligned}$$

Dabei ist  $a$  ein Sattelpunkt, wenn in jeder Umgebung  $U_\varepsilon(a)$  Punkte  $u, v \in U_\varepsilon(a)$  existieren mit  $f(u) < f(a) < f(v)$ .

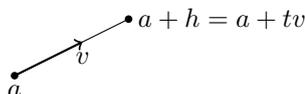
**Beweis.** Wir benutzen die Taylorentwicklung 2. Ordnung aus Satz 8.14.2 in  $a$ . Da  $a$  kritischer Punkt ist, also  $\text{grad } f(a) = 0$ , gilt

$$f(a+h) \approx f(a) + \frac{1}{2} \langle h, H_f(a)h \rangle$$

für  $h \in \mathbb{R}^n$  klein. Sei

$$h = tv$$

mit einem normierten Richtungsvektor  $|v| = 1$ . Der Punkt  $a+h$  in der Taylorentwicklung ist damit  $a+h = a+tv$ , ein Punkt auf der Geraden durch  $a$  in Richtung  $v$ :



Einsetzen in die Entwicklung ergibt

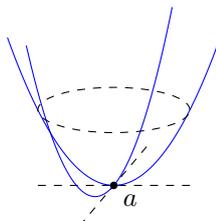
$$f(a+tv) \approx f(a) + \frac{1}{2} \langle tv, H_f(a)tv \rangle = f(a) + \frac{1}{2} \langle v, H_f(a)v \rangle \cdot t^2.$$

Es gilt also

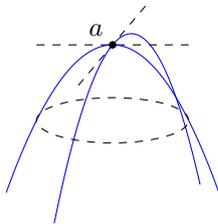
$$f(a+tv) \approx \alpha + \beta t^2$$

mit den Konstanten  $\alpha = f(a)$  und  $\beta = \frac{1}{2} \langle v, H_f(a)v \rangle$ , d.h. in Richtung  $v$  verläuft  $f$  ungefähr wie die Parabel  $\alpha + \beta t^2$ .

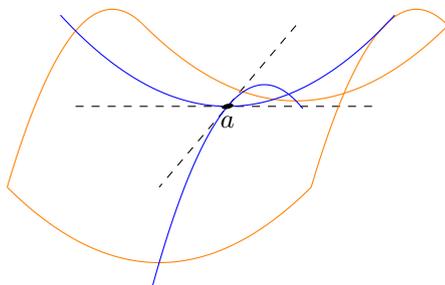
Ist nun  $H_f(a)$  positiv definit, so gilt  $\langle H_f(a)v, v \rangle > 0$  für alle  $v \neq 0$  (siehe Definition 7.3.1). Somit ist auch  $\beta > 0$  für jede Richtung  $v$ , d.h. die  $f$  annähernde Parabel  $\alpha + \beta t^2$  ist für jede Richtung nach oben geöffnet. Also ist  $a$  ein lokales Minimum.



Ist  $H_f(a)$  negativ definit, so ist entsprechend  $\langle H_f(a)v, v \rangle < 0$  für jedes  $v \neq 0$ , d.h.  $\beta < 0$ , also alle Parabeln nach unten geöffnet und damit  $a$  ein lokales Maximum.



Ist schließlich  $H_f(a)$  indefinit, so gibt es einen Vektor  $v$  mit  $\langle H_f(a)v, v \rangle > 0$  und einen Vektor  $\tilde{v}$  mit  $\langle H_f(a)\tilde{v}, \tilde{v} \rangle < 0$ . Also ist  $\beta > 0$  und  $\beta < 0$  für verschiedene Richtungen  $v$  und wir erhalten einen Sattelpunkt.



□

**Beispiel 8.16.2** Wir betrachten noch einmal die erste Funktion aus Beispiel 8.15.4:

$$f(x, y) = \frac{1}{3}x^3 + xy^2 - x.$$

Ihr Gradient war

$$\text{grad } f(x, y) = \begin{pmatrix} x^2 + y^2 - 1 \\ 2xy \end{pmatrix}$$

und als kritische Punkte hatten wir  $(1, 0)$ ,  $(-1, 0)$ ,  $(0, 1)$ ,  $(0, -1)$  berechnet. Die Hessematrix von  $f$  ist

$$H_f(x, y) = \begin{pmatrix} 2x & 2y \\ 2y & 2x \end{pmatrix}.$$

(Erinnerung: die erste Spalte ist die Ableitung des Gradienten nach  $x$ , die zweite Spalte die Ableitung nach  $y$ .) Mit dem hinreichenden Kriterium 8.16.1 für lokale Extrema untersuchen wir jetzt den Typ der kritischen Punkte. Dazu bestimmen wir die Definitheit der Hessematrix mit Satz 7.3.3.

- Punkt  $(1, 0)$ : Die Hessematrix in  $(1, 0)$  ist

$$H_f(1, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Es gilt

$$\det \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 4 > 0 \quad \text{und} \quad a = 2 > 0.$$

Nach Satz 7.3.3 ist somit  $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$  positiv definit. Also ist  $(1, 0)$  ein lokales Minimum.

- Punkt  $(-1, 0)$ : Hier ist

$$\begin{aligned} H_f(-1, 0) &= \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}, \\ \det \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} &= 4 > 0 \quad \text{und} \quad a = -2 < 0 \\ \Rightarrow \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} &\text{negativ definit} \quad \Rightarrow \quad (-1, 0) \text{ lokales Maximum.} \end{aligned}$$

- Punkt  $(0, 1)$ :  $H_f(0, 1) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$ .

$$\begin{aligned} \det \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} &= -4 < 0 \\ \Rightarrow \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} &\text{indefinit} \quad \Rightarrow \quad (0, 1) \text{ Sattelpunkt.} \end{aligned}$$

- Punkt  $(0, -1)$ :  $H_f(0, -1) = \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}$

$$\begin{aligned} \det \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix} &= -4 < 0 \\ \Rightarrow \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix} &\text{indefinit} \quad \Rightarrow \quad (0, -1) \text{ Sattelpunkt.} \end{aligned}$$

## 8.17 Extrema unter Nebenbedingungen

Oft sind in Anwendungen nicht einfach nur die Maxima oder Minima einer Funktion gesucht, sondern es sollen zugleich weitere Bedingungen erfüllt sein, die durch Gleichungen gegeben sind. Diese sogenannten *Nebenbedingungen* legen also fest, welche Punkte überhaupt nur auf Minima und Maxima untersucht werden. Allgemein kann man das so formulieren:

Gegeben sind Funktionen  $f, g_1, \dots, g_m : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Gesucht sind die Extrema von  $f$  unter den Nebenbedingungen

$$g_1(x) = c_1, \dots, g_m(x) = c_m \tag{8.18}$$

wobei die Konstanten  $c_1, \dots, c_m \in \mathbb{R}$  als Teil der Nebenbedingungen mit gegeben sind. Die Nebenbedingungen (8.18) definieren die Menge  $M$  der für die Extremwerte zulässigen Punkte  $x$ :

$$M = \{x \in G \mid g_1(x) = c_1, \dots, g_m(x) = c_m\}.$$

Anders gesagt: wir suchen die Extrema von  $f(x)$  auf der Menge  $M$ .

Geometrisch ist  $M$  eine Menge von Punkten im Raum  $\mathbb{R}^n$ . Folgende Fälle sind typisch:

- $m = 1, n = 2$  (d.h. eine Nebenbedingung, zwei Variablen  $x = (x_1, x_2)$ ):  $M$  ist eine Kurve in der Ebene  $\mathbb{R}^2$ .
- $m = 1, n = 3$  (eine Nebenbedingung, drei Variablen  $x = (x_1, x_2, x_3)$ ):  $M$  ist eine Fläche im dreidimensionalen Raum  $\mathbb{R}^3$ .
- $m = 2, n = 3$  (zwei Nebenbedingungen, drei Variablen):  $M$  ist eine Kurve im Raum  $\mathbb{R}^3$ .

Allgemein ist  $M$  eine „Fläche“ der Dimension  $n - m$  im  $n$ -dimensionalen Raum  $\mathbb{R}^n$ .

Wir geben ein einfaches Beispiel für eine Nebenbedingung und die entsprechende Menge  $M$ .

**Beispiel 8.17.1** Gesucht sind die Extrema einer Funktion  $f(x_1, x_2, x_3)$  unter der Nebenbedingung  $x_1^2 + x_2^2 + x_3^2 = 9$ . In unserer allgemeinen Schreibweise ist die Nebenbedingung

$$g_1(x) = c_1 \quad \text{wobei} \quad g_1(x) = x_1^2 + x_2^2 + x_3^2, \quad c_1 = 9.$$

Die Menge  $M$  ist

$$\begin{aligned} M &= \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 9\} \\ &= \left\{x \in \mathbb{R}^3 \mid |x| = \sqrt{x_1^2 + x_2^2 + x_3^2} = 3\right\} \end{aligned}$$

$M$  ist also die Kugeloberfläche mit Radius 3 und Mittelpunkt im Ursprung  $(0, 0, 0)$  (siehe auch Abschnitt 8.1). Gesucht sind damit die Extrema von  $f$  auf der Kugel  $M$ .

Zur Berechnung von Extrema unter Nebenbedingungen gibt es das folgende notwendige Kriterium:

**Satz 8.17.2 (Lagrangesche Multiplikatorenregel)** *Seien  $f, g_1, \dots, g_m$  stetig partiell differenzierbar und  $(\text{grad } g_1(x), \dots, \text{grad } g_m(x))$  linear unabhängig für alle  $x \in G$ . Ist  $x = a$  ein lokales Extremum von  $f$  unter den Nebenbedingungen  $g_1(x) = c_1, \dots, g_m(x) = c_m$  dann gilt*

$$\text{grad } f(a) = \sum_{i=1}^m \lambda_i \text{grad } g_i(a) \quad \text{mit} \quad \lambda_1, \dots, \lambda_m \in \mathbb{R}. \quad (8.19)$$

Die Zahlen  $\lambda_1, \dots, \lambda_m$  in (8.19) heißen auch *Lagrange-Multiplikatoren*. Aus dem Satz folgt, dass die möglichen Extrema genau die Lösungen des Gleichungssystems

$$\begin{cases} \text{grad } f(x) = \sum_{i=1}^m \lambda_i \text{grad } g_i(x) \\ g_1(x) = c_1, \dots, g_m(x) = c_m \end{cases} \quad (8.20)$$

sind.

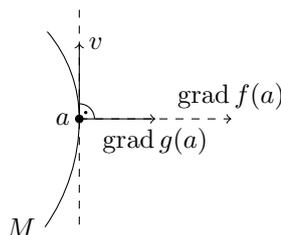


Abbildung 8.11: Zum Beweis der Langrangeschen Multiplikatorenregel

**Beweis Satz 8.17.2.** Wir geben den Beweis im Fall  $m = 1$  an. Sei also  $x = a$  ein lokales Extremum von  $f(x)$  unter der Nebenbedingung  $g(x) = c$ . Die Menge  $M$  ist in diesem Fall

$$M = \{x \in G \mid g(x) = c\}.$$

Das ist genau die Niveaumenge von  $g$  zum Wert  $c$ . Aus dem Abschnitt zu Niveaumengen wissen wir, dass  $\text{grad } g(a) \perp M$  im Punkt  $a$  (Satz 8.11.1), siehe Abbildung 8.11. Sei  $x(t)$  eine Kurve auf  $M$  durch den Punkt  $a$  in Richtung  $v$ , d.h.

$$x(0) = a, \quad x'(0) = v.$$

$v$  liegt dann tangential zu  $M$  (Abbildung 8.11). Wir betrachten die Funktion  $\varphi(t) = f(x(t))$ , also die Funktionswerte von  $f$  auf der Kurve  $x(t)$ . Da  $a$  ein lokales Extremum von  $f$  auf  $M$  ist, und die Kurve  $x(t)$  in  $M$  durch  $a$  läuft ( $x(0) = a$ ), hat  $\varphi(t) = f(x(t))$  also bei  $t = 0$  ein lokales Extremum. Somit gilt  $\varphi'(0) = 0$ , und mit der Kettenregel 8.12.3 folgt

$$0 = \varphi'(0) = \langle \text{grad } f(x(0)), x'(0) \rangle = \langle \text{grad } f(a), v \rangle \quad \Rightarrow \quad \text{grad } f(a) \perp v$$

Weil  $v$  eine beliebige Tangentialrichtung an  $M$  ist, ist  $\text{grad } f(a)$  senkrecht zu  $M$  im Punkt  $a$ . Also müssen  $\text{grad } f(a)$  und  $\text{grad } g(a)$  parallel liegen (Abbildung 8.11), d.h. es gilt

$$\text{grad } f(a) = \lambda \text{grad } g(a) \quad \text{für ein } \lambda \in \mathbb{R}.$$

Das ist genau (8.19) für  $m = 1$  (eine Nebenbedingung).  $\square$

**Hinweis:** Die Bedingung in Satz 8.17.2, dass die Gradienten von  $g_1, \dots, g_m$  linear unabhängig sind, stellt sicher, dass die Nebenbedingungen voneinander unabhängig sind, dass also z.B. nicht aus einer Nebenbedingung schon eine andere folgt. Nur damit ergibt sich oben  $n - m$  als Dimension von  $M$ . Und für  $m = 1$  bedeutet die lineare Unabhängigkeit einfach, dass  $\text{grad } g_1(x)$  stets ungleich Null ist.

**Beispiel 8.17.3** Wir suchen die Extrema von  $f(x, y) = 2x - 2y + 1$  unter der Nebenbedingung

$$g(x, y) = x^2 + y^2 = 4.$$

Die Lagrange-Multiplikatorenregel (8.19) liefert

$$\text{grad } f(x, y) = \begin{pmatrix} 2 \\ -2 \end{pmatrix} \stackrel{!}{=} \lambda \text{grad } g(x, y) = \lambda \cdot \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

Zusammen mit der Nebenbedingung ergibt sich das Gleichungssystem

$$\begin{cases} 2 = 2\lambda x & \text{(I)} \\ -2 = 2\lambda y & \text{(II)} \\ x^2 + y^2 = 4 & \text{(III)} \end{cases}$$

Aus Gleichung (I) und (II) erhalten wir

$$(I) \Leftrightarrow 1 = \lambda x \Leftrightarrow \lambda \neq 0, x = \frac{1}{\lambda} \quad (8.21)$$

$$(II) \Leftrightarrow -1 = \lambda y \Leftrightarrow y = -\frac{1}{\lambda} \quad (8.22)$$

Einsetzen in die Nebenbedingung (III) liefert dann

$$\begin{aligned} x^2 + y^2 &= \left(\frac{1}{\lambda}\right)^2 + \left(-\frac{1}{\lambda}\right)^2 = \frac{2}{\lambda^2} \stackrel{!}{=} 4 \\ \Rightarrow \lambda^2 &= \frac{1}{2} \Rightarrow \lambda = \pm \frac{1}{\sqrt{2}} \end{aligned}$$

Wir haben also zwei Fälle. Für jeden Wert von  $\lambda$  erhalten wir dann aus (8.21) und (8.22) den Wert für  $x$  und  $y$ :

- $\lambda = \frac{1}{\sqrt{2}} \Rightarrow x = \frac{1}{\lambda} = \sqrt{2}, y = -\frac{1}{\lambda} = -\sqrt{2}$
- $\lambda = -\frac{1}{\sqrt{2}} \Rightarrow x = -\sqrt{2}, y = \sqrt{2}$

Somit haben wir die Punkte  $(\sqrt{2}, -\sqrt{2})$  und  $(-\sqrt{2}, \sqrt{2})$  als mögliche (lokale) Extrema.

## 8.18 Globale Extrema unter Nebenbedingungen

Die Lagrangesche Multiplikatorenregel ist ein notwendiges Kriterium und liefert daher nur *möglich* Extrema. Bei Extrema unter Nebenbedingungen gibt es *kein* allgemeines hinreichendes Kriterium um zu sagen, ob ein mögliches Extremum auch tatsächlich ein Extremum unter den Nebenbedingungen ist, ein lokales Maximum oder Minimum. (Im Gegensatz hierzu konnten wir bei Extremwerten ohne Nebenbedingung Satz 8.16.1 und die Hessematrix benutzen.)

Anders sieht es aus, wenn wir nach *globalen* Extremwerten fragen. Hier gibt es ein hinreichendes Kriterium für den Fall, dass die durch die Nebenbedingungen gegebene Menge  $M$  *beschränkt* ist – ein typischer Fall in Anwendungen.

**Definition 8.18.1** Eine Menge  $M \subset \mathbb{R}^n$  ist *beschränkt* wenn es einen  $r \in \mathbb{R}$ ,  $r > 0$  gibt, so dass  $M$  in der Kugel mit Radius  $r$  um 0 enthalten ist, d.h.

$$M \subset \{x \in \mathbb{R}^n \mid |x| \leq r\}.$$

Die Menge  $M$  ist also beschränkt, wenn sie in eine endlich große Kugel passt oder, grob gesprochen, wenn sie „in keiner Richtung nach Unendlich läuft“.

Der folgende grundlegende Satz sichert allgemein die Existenz globaler Extrema.

**Satz 8.18.2** Ist  $M \subset \mathbb{R}^n$  beschränkt und abgeschlossen (der Rand gehört also zu  $M$ ), und ist  $f : M \rightarrow \mathbb{R}$  eine stetige Funktion, dann hat  $f$  auf  $M$  ein globales Minimum und ein globales Maximum. D.h. es gibt  $x_0, x_1 \in M$  so dass gilt

$$f(x_0) \leq f(x) \leq f(x_1) \quad \text{für alle } x \in M.$$

(Hinweis: Wie üblich bedeutet hier „ein“ globales Maximum bzw. Minimum, dass es *mindestens* ein Maximum und Minimum gibt!)

Diese allgemeine Aussage können wir jetzt auf Extrema unter Nebenbedingungen anwenden:

**Satz 8.18.3** Sind die Funktionen  $f, g_1, \dots, g_m : G \subset \mathbb{R}^n \rightarrow \mathbb{R}$  stetig und ist die durch die Nebenbedingungen (8.18) definierte Menge

$$M = \{x \in G \mid g_1(x) = c_1, \dots, g_m(x) = c_m\}$$

beschränkt, dann hat  $f$  auf  $M$  (also unter den Nebenbedingungen) ein globales Maximum und Minimum.

**Beweis.**  $M$  ist abgeschlossen (typischerweise ist hier  $\partial M = M$ ) und nach Voraussetzung beschränkt. Also gibt es nach Satz 8.18.2 Maximum und Minimum.  $\square$

Da die globalen Extrema auch lokal sind, sind sie für beschränktes  $M$  Lösungen von (8.20). Anders gesagt ist unter allen möglichen Extrema, die sich aus (8.20) ergeben, der Punkt (oder die Punkte) mit dem größten Wert von  $f$  das globale Maximum, der Punkt mit kleinstem Wert das globale Minimum.

**Beispiel 8.18.4** Wir setzen Beispiel 8.17.3 fort. Für die Funktion  $f(x, y) = 2x - 2y + 1$  unter der Nebenbedingung  $g(x, y) = x^2 + y^2 = 4$  hatten wir als mögliche Extrema  $(\sqrt{2}, -\sqrt{2})$  und  $(-\sqrt{2}, \sqrt{2})$  berechnet. Hier ist

$$M = \{(x, y) \mid x^2 + y^2 = 4\},$$

d.h.  $M$  ist der Kreis (Kreislinie) mit Radius 2. Also ist  $M$  beschränkt, und damit hat  $f$  auf  $M$  ein globales Maximum und Minimum. Es gilt

$$\begin{aligned} f(\sqrt{2}, -\sqrt{2}) &= 2\sqrt{2} - 2(-\sqrt{2}) + 1 = 4\sqrt{2} + 1 \\ f(-\sqrt{2}, \sqrt{2}) &= -2\sqrt{2} - 2\sqrt{2} + 1 = -4\sqrt{2} + 1 \\ \implies f(-\sqrt{2}, \sqrt{2}) &< f(\sqrt{2}, -\sqrt{2}) \end{aligned}$$

Also ist  $(\sqrt{2}, -\sqrt{2})$  das (globale) Maximum von  $f$  unter der Nebenbedingung,  $(-\sqrt{2}, \sqrt{2})$  ist das (globale) Minimum.

## 8.19 Ableitung vektorwertiger Funktionen

Wir kommen nun zu vektorwertigen Funktionen. Sei  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  mit  $G$  offen. Es ist also

$$y = f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}.$$

Die Funktion hat also  $m$  Komponenten  $f_1, \dots, f_m$ , die alle wiederum von den Variablen  $x_1, \dots, x_n$  abhängen.

Die Funktion  $f$  ist *partiell differenzierbar* wenn alle partiellen Ableitungen aller Komponenten  $\frac{\partial f_i}{\partial x_j}$  ( $i = 1, \dots, m, j = 1, \dots, n$ ) existieren. Die Matrix

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (8.23)$$

heißt *Jacobimatrix* von  $f$ . Hier ist die Anordnung wichtig: Die Komponenten  $f_1, \dots, f_m$  sind untereinander angeordnet, genau wie auch im Vektor der Funktion  $f$  selbst. Die partiellen Ableitungen nach  $x_1, \dots, x_n$  stehen hintereinander. Die Matrix hat daher  $m$  Zeilen und  $n$  Spalten. Die Jacobimatrix kann man auch wie folgt schreiben, was je nach Situation praktisch ist:

$$J_f(x) = \begin{pmatrix} \text{grad } f_1(x)^T \\ \vdots \\ \text{grad } f_m(x)^T \end{pmatrix}, \quad (8.24)$$

die  $i$ . Zeile von  $J_f(x)$  ist der (transponierte) Gradient der  $i$ . Komponente  $f_i$  von  $f$ . Oder

$$J_f(x) = \left( \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) \quad (8.25)$$

wobei

$$\frac{\partial f}{\partial x_j} = \begin{pmatrix} \frac{\partial f_1}{\partial x_j} \\ \vdots \\ \frac{\partial f_m}{\partial x_j} \end{pmatrix} \quad \text{da} \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}.$$

In diesem Sinne stehen also in der Jacobimatrix nacheinander die partiellen Ableitungen des Funktionsvektors  $f$ .

**Beachte:** Die Jacobimatrix ist auch für die extremen Fälle  $m = 1$  sowie  $n = 1$  definiert:

- $m = 1$ :  $f$  ist skalarwertig,  $J_f(x)$  hat nur eine Zeile.
- $n = 1$ :  $f$  hängt nur von einer Variablen ab,  $J_f(x)$  hat nur eine Spalte.

**Beispiel 8.19.1** (a) Wir berechnen die Jacobimatrix der Funktion

$$f(x, y) = \begin{pmatrix} x^2 \cos y \\ x^2 + y^2 \end{pmatrix}.$$

Wir leiten zunächst den Funktionsvektor partiell nach  $x$  und  $y$  ab:

$$\frac{\partial f}{\partial x} = \begin{pmatrix} 2x \cos y \\ 2x \end{pmatrix}, \quad \frac{\partial f}{\partial y} = \begin{pmatrix} -x^2 \sin y \\ 2y \end{pmatrix}.$$

Das sind genau die Spalten der Jacobimatrix,  $J_f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$ , siehe (8.25). Also

$$J_f(x, y) = \begin{pmatrix} 2x \cos y & -x^2 \sin y \\ 2x & 2y \end{pmatrix}$$

(b) Die Funktion  $g(x, y, z) = x + y^2 + xz$  ist skalarwertig. Die Jacobimatrix ist

$$J_g(x, y, z) = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}\right) = (1 + z, 2y, x).$$

Das ist natürlich der transponierte Gradient  $\text{grad } g(x, y, z)^T$ , vergleiche (8.24).

Wir verallgemeinern jetzt das totale Differential (8.9) auf den vektorwertigen Fall. Für eine stetig partiell differenzierbare Funktion  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist mit den naheliegenden Abkürzungen

$$df = \begin{pmatrix} df_1 \\ \vdots \\ df_m \end{pmatrix}, \quad dx = \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}$$

das *vektorwertige totale Differential* gegeben durch

$$df = J_f(x)dx. \quad (8.26)$$

Denn das totale Differential der  $i$ . Komponente von  $f$  ist nach (8.9)

$$df_i = \langle \text{grad } f_i(x), dx \rangle = \frac{\partial f_i}{\partial x_1} dx_1 + \cdots + \frac{\partial f_i}{\partial x_n} dx_n = \text{grad } f_i(x)^T dx,$$

und dies ist genau die  $i$ . Komponente von  $J_f(x)dx$ , da  $\text{grad } f_i(x)^T$  die  $i$ . Zeile der Jacobimatrix ist.

Schließlich betrachten wir noch die Kettenregel im allgemeinen vektorwertigen Fall. Seien  $f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $g : H \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$  zwei Funktionen so dass  $f(x) \in H$  für  $x \in G$  gilt. Der Funktionswert  $f(x)$  liegt also im Definitionsbereich  $H$  der Funktion  $g$ , und wir können damit die Verkettung beider Funktionen bilden:

$$g \circ f : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^p, \quad (g \circ f)(x) = g(f(x)).$$

Schreibt man dies in allen Komponenten aus, ist das

$$g(f(x)) = g(u) = \begin{pmatrix} g_1(u_1, \dots, u_m) \\ \vdots \\ g_p(u_1, \dots, u_m) \end{pmatrix} \quad \text{mit } u = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}.$$

Man beachte, dass die Anzahl der Komponenten von  $f$  (also  $f_1, \dots, f_m$ ) gleich der Anzahl der Variablen von  $g$  ist ( $u_1, \dots, u_m$ ), und man deshalb  $f(x)$  überhaupt in  $g(u)$  einsetzen kann. In dieser Situation gilt jetzt folgende *Kettenregel*:

**Satz 8.19.2 (allgemeine Kettenregel)** Sind  $f$  und  $g$  differenzierbar, so ist auch  $g \circ f$  differenzierbar und es gilt

$$J_{g \circ f}(x) = J_g(f(x))J_f(x).$$

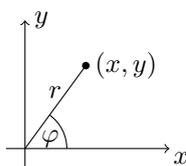


Abbildung 8.12: Polarkoordinaten

**Beweis.** Wir zeigen hier nur die Formel der Kettenregel, nicht die Differenzierbarkeit von  $g \circ f$ . Wir betrachten die totalen Differentiale von  $y = g(u)$  und  $u = f(x)$ :

$$dy = J_g(u)du, \quad du = J_f(x)dx$$

woraus

$$dy = J_g(u)J_f(x)dx$$

folgt. Andererseits ist das totale Differential von  $y = (g \circ f)(x)$

$$dy = J_{g \circ f}(x)dx.$$

Somit folgt  $J_{g \circ f}(x) = J_g(u)J_f(x) = J_g(f(x))J_f(x)$ .  $\square$

Man beachte noch, dass  $J_g(u) \in \mathbb{R}^{p \times m}$  und  $J_f(x) \in \mathbb{R}^{m \times n}$ . Die Matrizen passen also so zusammen, dass man das Matrixprodukt  $J_g(f(x))J_f(x)$  bilden kann; das Ergebnis ist dann in  $\mathbb{R}^{p \times n}$ , genau wie die Jacobimatrix  $J_{g \circ f}$  der Verkettung.

**Bemerkung:** Die Kettenregel

$$J_{g \circ f}(x) = J_g(f(x))J_f(x)$$

ist nicht nur die Verallgemeinerung der Kettenregel im eindimensionalen Fall,

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

sie hat auch genau die gleiche Struktur: „äußere“ mal „innere“ Ableitung, wobei in die äußere Ableitung wieder die innere Funktion einzusetzen ist. Aus der gewöhnlichen Ableitung wird hier im allgemeinen Fall jeweils die entsprechende Jacobimatrix.

Durch die Schreibweise mit der Jacobimatrix wird die Kettenregel sehr kurz und elegant. Man beachte, dass hier ein Matrixprodukt aus lauter partiellen Ableitungen zu berechnen ist: das Produkt hat  $p \cdot n$  Komponenten, von denen jede eine Summe von  $m$  Produkten jeweils einer partiellen Ableitung von  $g$  mit einer partiellen Ableitung von  $f$  ist!

## 8.20 Anwendung: Ableitung in Polarkoordinaten

In Polarkoordinaten wird ein Punkt  $(x, y)$  in der Ebene beschrieben durch seine Entfernung  $r$  zum Ursprung („Radius“) und den Winkel  $\varphi$  zwischen positiver  $x$ -Achse und der Gerade vom Ursprung zum Punkt, siehe Abbildung 8.12. Der

Zusammenhang zwischen kartesischen Koordinaten  $x, y$  und Polarkoordinaten  $r, \varphi$  ist<sup>2</sup>

$$\begin{aligned}x &= r \cos \varphi \\y &= r \sin \varphi\end{aligned}\tag{8.27}$$

Diese Koordinatentransformation kann man folgendermaßen als eine vektorwertige Funktion  $\Phi$  schreiben:

$$\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} = \Phi(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}.\tag{8.28}$$

Eine beliebige Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  kann man sowohl in kartesischen als auch in Polarkoordinaten darstellen. Zum Beispiel ist

$$f(x, y) = x + x^2 + y^2$$

in Polarkoordinaten von der Form

$$\tilde{f}(r, \varphi) = r \cos \varphi + r^2,$$

denn durch Einsetzen der Polarkoordinaten (8.27) ist

$$\begin{aligned}x + x^2 + y^2 &= r \cos \varphi + r^2 \cos^2 \varphi + r^2 \sin^2 \varphi \\&= r \cos \varphi + r^2 \underbrace{(\cos^2 \varphi + \sin^2 \varphi)}_1 = r \cos \varphi + r^2.\end{aligned}$$

Zur Unterscheidung der beiden Funktionsterme, einmal kartesisch, einmal polar dargestellt, haben wir hier die verschiedenen Bezeichnungen  $f$  beziehungsweise  $\tilde{f}$  benutzt:  $f$  ist die Funktion in kartesischen Koordinaten,  $\tilde{f}$  die Funktion in Polarkoordinaten. Es gilt dann allgemein der Zusammenhang

$$f(x, y) = f(r \cos \varphi, r \sin \varphi) = \tilde{f}(r, \varphi).$$

Mit der Transformationsfunktion  $\Phi$  aus (8.28) können wir das auch so schreiben:

$$\tilde{f}(r, \varphi) = f(r \cos \varphi, r \sin \varphi) = f(\Phi(r, \varphi)) = f \circ \Phi(r, \varphi),\tag{8.29}$$

kurz  $\tilde{f} = f \circ \Phi$  (Verkettung von  $f$  mit  $\Phi$ !)

Wir können nun partielle Ableitungen der Funktion  $f$  nach kartesischen Koordinaten  $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$  berechnen, oder nach Polarkoordinaten  $\frac{\partial}{\partial r}, \frac{\partial}{\partial \varphi}$ . Diese partiellen Ableitungen können mithilfe der Kettenregel ineinander umgerechnet werden: Aus (8.29) folgt mit der Kettenregel

$$J_{\tilde{f}} = J_f J_{\Phi}.$$

Die Jacobimatrix von  $\Phi$  ist

$$J_{\Phi}(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix},\tag{8.30}$$

und damit folgt

$$J_{\tilde{f}} = \begin{pmatrix} \frac{\partial \tilde{f}}{\partial r} & \frac{\partial \tilde{f}}{\partial \varphi} \end{pmatrix} = J_f J_{\Phi} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

<sup>2</sup>Polarkoordinaten werden ausführlicher im Skript von Mathematik A erklärt.

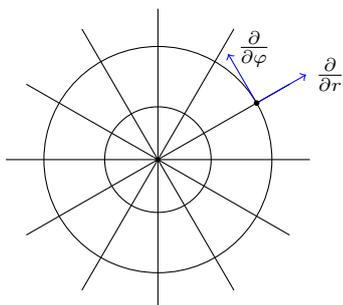


Abbildung 8.13: Partielle Ableitung nach Polarkoordinaten

Durch Ausrechnen des Matrixproduktes ergibt das

$$\frac{\partial \tilde{f}}{\partial r} = \cos \varphi \frac{\partial f}{\partial x} + \sin \varphi \frac{\partial f}{\partial y} \quad , \quad \frac{\partial \tilde{f}}{\partial \varphi} = -r \sin \varphi \frac{\partial f}{\partial x} + r \cos \varphi \frac{\partial f}{\partial y}. \quad (8.31)$$

Mit dieser Formel kann man also die partiellen Ableitungen  $\frac{\partial}{\partial r}$ ,  $\frac{\partial}{\partial \varphi}$  aus den kartesischen partiellen Ableitungen  $\frac{\partial}{\partial x}$ ,  $\frac{\partial}{\partial y}$  berechnen.

Die partiellen Ableitungen nach den Polarkoordinaten entsprechen Richtungsableitungen in die jeweilige Polarkoordinatenrichtung, vergleiche Abbildung 8.13:

$$\begin{aligned} \frac{\partial \tilde{f}}{\partial r} &= \text{Richtungsableitung von } f \text{ in Richtung } r \\ \frac{\partial \tilde{f}}{\partial \varphi} &= \text{Richtungsableitung von } f \text{ in Richtung } \varphi \end{aligned}$$

Denn nach der Definition der partielle Ableitung über den Differenzenquotient beschreibt zum Beispiel  $\frac{\partial}{\partial r}$  die Änderung der Funktionswerte bei einer Änderung von  $r$ , d.h. in Richtung der Koordinate  $r$ . Man kann aber auch (8.31) als Richtungsableitungen umschreiben; so ist etwa

$$\frac{\partial \tilde{f}}{\partial r} = \cos \varphi \frac{\partial f}{\partial x} + \sin \varphi \frac{\partial f}{\partial y} = \left\langle \text{grad } f, \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \right\rangle = \partial_{v_r} f \quad \text{mit} \quad v_r = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix},$$

also die Richtungsableitung von  $f$  in Richtung  $v_r$ . Und der Vektor  $v_r$  zeigt im Punkt mit den Polarkoordinaten  $r, \varphi$  genau vom Ursprung weg, in  $r$ -Richtung.

Wir berechnen jetzt noch die *Umkehrung* von (8.31), also eine Formel um die kartesischen partiellen Ableitungen aus den Ableitungen nach Polarkoordinaten zu berechnen. Mit der Inversen der Jacobimatrix  $J_{\Phi}^{-1}$  gilt

$$J_{\tilde{f}} = J_f J_{\Phi} \quad \Rightarrow \quad J_f = J_{\tilde{f}} J_{\Phi}^{-1}.$$

Aus der Formel für die Inverse einer  $2 \times 2$ -Matrix (siehe Mathe A),

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

und  $\det J_{\Phi} = r \cos^2 \varphi + r \sin^2 \varphi = r$  folgt

$$J_{\Phi}^{-1} = \frac{1}{r} \begin{pmatrix} r \cos \varphi & r \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}.$$

Damit ist

$$J_f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = \left( \frac{\partial \tilde{f}}{\partial r}, \frac{\partial \tilde{f}}{\partial \varphi} \right) \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{1}{r} \sin \varphi & \frac{1}{r} \cos \varphi \end{pmatrix}$$

und somit

$$\frac{\partial f}{\partial x} = \cos \varphi \frac{\partial \tilde{f}}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial \tilde{f}}{\partial \varphi} \quad , \quad \frac{\partial f}{\partial y} = \sin \varphi \frac{\partial \tilde{f}}{\partial r} + \frac{1}{r} \cos \varphi \frac{\partial \tilde{f}}{\partial \varphi}. \quad (8.32)$$

# Kapitel 9

## Mehrdimensionale Integration

In diesem Kapitel betrachten wir die mehrdimensionale Integration, also die Integration von Funktionen mehrerer Variablen. Während bei eindimensionalen (bestimmten) Integralen über ein eindimensionales Intervall integriert wird, wird bei mehrdimensionalen Integralen über eine höherdimensionale Menge (Bereich) integriert, die im einfachen Fall ein Rechteck oder Quader ist, im Allgemeinen aber verschiedene Formen haben kann. Die Berechnung der mehrdimensionalen Integrale erfolgt letztlich durch wiederholte eindimensionale Integration, je nach Form des Integrationsbereichs können aber noch zusätzliche Schritte nötig sein.

*Hinweis:* In den alten Mathematik A&B Skripten finden sie den Stoff dieses Kapitels im Skript Mathematik B, Kapitel 7.

### 9.1 Zweidimensionale Integrale über Rechtecke

Wir erinnern zunächst an die Definition des Integrals von einer Variablen: Die Integration erfolgt über ein Intervall  $[a, b]$ , das dazu in  $n$  Teilstücke unterteilt wird, mit den Teilungspunkten  $x_0 = a < x_1 < x_2 < \dots < x_n = b$ . Jedes einzelne Teilstück hat die Länge  $\Delta x_i$ . Das Integral ist dann definiert als Grenzwert von Riemann-Summen

$$\int_a^b f(x) dx = \lim_{\substack{n \rightarrow \infty \\ \Delta x_i \rightarrow 0}} \sum_{i=1}^n f(x_i) \Delta x_i.$$

Das Produkt  $f(x_i) \Delta x_i$  entspricht (bei positiver Funktion) dem Flächeninhalt des Rechteckstreifens am Punkt  $x_i$  von der  $x$ -Achse bis zur Funktion  $f(x)$ , der die Breite  $\Delta x_i$  und die Höhe  $f(x_i)$  hat, siehe Abbildung 9.1. Im Grenzwert  $n \rightarrow \infty$  wird die Anzahl  $n$  der Unterteilungen immer weiter erhöht wobei gleichzeitig die Streifenbreite  $\Delta x_i$  gegen Null geht.

Wir wollen nun diese Konstruktion auf eine Funktion von zwei Variablen  $f(x, y)$  übertragen. Wir wollen die Integration für  $x$  über das Intervall  $[a, b]$ , für  $y$  über das Intervall  $[c, d]$  ausführen, das heißt insgesamt über den Bereich

$$Q = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\},$$

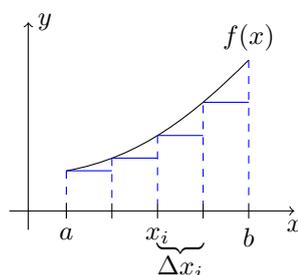
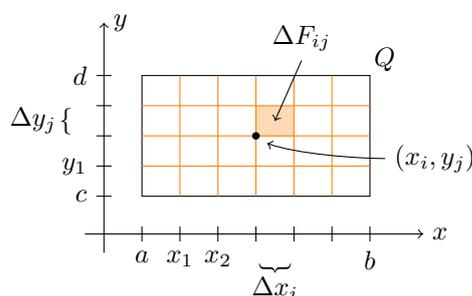


Abbildung 9.1: Konstruktion des eindimensionalen Integrals

Abbildung 9.2: Unterteilung des Rechtecks  $Q$ 

also ein Rechteck, siehe Abbildung 9.2. Wir können  $Q$  äquivalent (und etwas kürzer) als *kartesisches Produkt*

$$Q = [a, b] \times [c, d]$$

schreiben (vergleiche Mathematik A). Das ist nichts anderes als die Menge aller Punkte  $(x, y)$  mit  $x \in [a, b]$  und  $y \in [c, d]$ . Für die zu integrierende Funktion  $f$  nehmen wir an, dass sie (mindestens) auf  $Q$  definiert und dort stetig ist,

$$f : Q \subset \mathbb{R}^2 \rightarrow \mathbb{R} \quad \text{stetig.}$$

Analog zum eindimensionalen Fall zerlegen wir jetzt den Integrationsbereich  $Q$  in Teilstücke. Dazu unterteilen wir die Intervalle  $[a, b]$  und  $[c, d]$  durch

$$a = x_0 < x_1 < \dots < x_n = b, \quad c = y_0 < y_1 < \dots < y_n = d$$

und betrachten dann die entstehenden Flächenstücke

$$\Delta F_{ij} = \Delta x_i \Delta y_j.$$

Das ist in Abbildung 9.2 dargestellt. Wir bilde dann die Riemannsche Summe

$$\sum_{i,j=1}^n f(x_i, y_j) \Delta F_{ij} \tag{9.1}$$

wobei das Flächenstück  $\Delta F_{ij}$  im Punkt  $(x_i, y_j)$  mit dem Funktionswert in diesem Punkt multipliziert und dann über alle Teilflächen summiert wird. Der

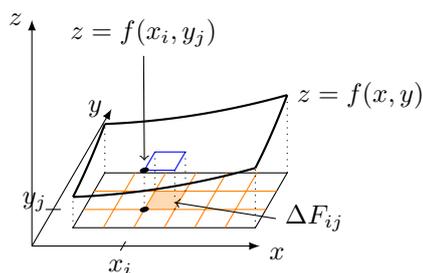


Abbildung 9.3: Zur Interpretation des Integrals als Volumen

Grenzwert  $n \rightarrow \infty$ ,  $\Delta x_i, \Delta y_j \rightarrow 0$  ergibt dann das Integral:

$$\int_Q f(x, y) dF = \lim_{\substack{n \rightarrow \infty \\ \Delta x_i, \Delta y_j \rightarrow 0}} \sum_{i,j=1}^n f(x_i, y_j) \Delta F_{ij} \quad (9.2)$$

Man bezeichnet  $\int_Q f(x, y) dF$  als das *Integral von  $f$  über  $Q$* . Da die Menge  $Q$ , über die integriert wird, eine Fläche ist, nennt man das Integral auch *Flächenintegral*. In der Schreibweise wird das durch „ $dF$ “ angedeutet.

Für eine positive Funktion einer Variablen  $f(x) \geq 0$  ergibt das Integral den Flächeninhalt zwischen dem Intervall  $[a, b]$  und der Funktion, siehe Abbildung 9.1. Für eine positive Funktion von zwei Variablen  $f(x, y) \geq 0$  hat das Integral ebenfalls eine geometrische Interpretation: In diesem Fall ist  $f(x_i, y_j) \Delta F_{ij}$  das Volumen des Quaders über der Rechteckfläche  $\Delta F_{ij}$  im Punkt  $(x_i, y_j)$  mit Höhe  $z = f(x_i, y_j)$ , vergleiche Abbildung 9.3. Die Riemannsche Summe (9.1) approximiert damit das Volumen über der Grundfläche  $Q$ , das nach oben durch die Funktion  $z = f(x, y)$  begrenzt wird. Je größer  $n$ , d.h. je feiner die Unterteilung von  $Q$ , desto genauer die Approximation. Im Grenzwert  $n \rightarrow \infty$  ergibt sich daher für das Integral:

$$f(x, y) \geq 0 \Rightarrow \int_Q f(x, y) dF = \text{Volumen des Körpers über } Q, \text{ der nach oben durch } z = f(x, y) \text{ begrenzt wird.}$$

Die Berechnung eines Integrals über ein Rechteck  $Q$  geschieht mit dem *Satz von Fubini* als Doppelintegral:

**Satz 9.1.1 (Fubini)** Für  $Q = [a, b] \times [c, d]$  und  $f : Q \rightarrow \mathbb{R}$  stetig gilt

$$\int_Q f(x, y) dF = \int_c^d \int_a^b f(x, y) dx dy = \int_a^b \int_c^d f(x, y) dy dx.$$

Speziell ist also in diesem Fall die Integrationsreihenfolge vertauschbar.

**Beweis.** Wir benutzen die Definition des Integrals in (9.2) und schreiben die Doppelsumme explizit als zwei einzelne Summen: und führen nacheinander erst den Grenzübergang  $\Delta x_i \rightarrow 0$  und dann  $\Delta y_j \rightarrow 0$  durch

$$\int_Q f(x, y) dF = \lim_{\Delta x_i, \Delta y_j \rightarrow 0} \sum_{j=1}^n \sum_{i=1}^n f(x_i, y_j) \Delta x_i \Delta y_j$$

Die innere Summe ist eine Riemannsche Summe für die Funktion  $f(x, y_j)$  mit konstantem  $y = y_j$ . Im Grenzwert  $\Delta x_i \rightarrow 0$  ergibt sich daraus das Integral über  $x$ ,

$$\lim_{\Delta x_i \rightarrow 0} \sum_{i=1}^n f(x_i, y_j) \Delta x_i = \int_a^b f(x, y_j) dx.$$

Also gilt

$$\begin{aligned} \int_Q f(x, y) dF &= \lim_{\Delta x_i, \Delta y_j \rightarrow 0} \sum_{j=1}^n \sum_{i=1}^n f(x_i, y_j) \Delta x_i \Delta y_j \\ &= \lim_{\Delta y_j \rightarrow 0} \sum_{j=1}^n \int_a^b f(x, y_j) dx \Delta y_j. \end{aligned}$$

Das ist noch eine Riemannsche Summe bezüglich der Variablen  $y$  und damit folgt dann

$$\int_Q f(x, y) dF = \lim_{\Delta y_j \rightarrow 0} \sum_{j=1}^n \int_a^b f(x, y_j) dx \Delta y_j = \int_c^d \int_a^b f(x, y) dx dy.$$

Durch Vertauschen der Summen über  $i$  und  $j$  erhält man genau so die Formel mit vertauschter Integrationsreihenfolge:

$$\begin{aligned} \int_Q f(x, y) dF &= \lim_{\Delta x_i, \Delta y_j \rightarrow 0} \sum_{i=1}^n \sum_{j=1}^n f(x_i, y_j) \Delta y_j \Delta x_i \\ &= \lim_{\Delta x_i \rightarrow 0} \sum_{i=1}^n \int_c^d f(x_i, y) dy \Delta x_i = \int_a^b \int_c^d f(x, y) dy dx \end{aligned}$$

□

**Beachte:** Das Doppelintegral im Satz von Fubini ist mit einer impliziten Klammerung zu verstehen, z.B.:

$$\int_a^b \int_c^d f(x, y) dy dx = \int_a^b \left( \int_c^d f(x, y) dy \right) dx$$

Das bedeutet, dass die inneren Integrationsgrenzen auch zur inneren Integrationsvariablen passen müssen. Hier ist die innere Integration nach  $y$  (weil  $dy$  links von  $dx$  steht), und damit sind auch die inneren Grenzen die von  $y$ , also von  $c$  bis  $d$ . Entsprechend ist die äußere Integration nach  $x$  mit den Grenzen  $a$  bis  $b$ . Beim Vertauschen der Integrationen vertauscht sich damit  $dx$  mit  $dy$  und auch die Grenzen  $a, b$  mit  $c, d$ .

**Beispiel 9.1.2** Wir integrieren  $f(x, y) = xy^2$  über das Rechteck

$$Q = [0, 1] \times [0, 2]$$

(d.h.  $x \in [0, 1]$ ,  $y \in [0, 2]$ ).

$$\begin{aligned}\int_Q xy^2 dF &= \int_0^2 \int_0^1 xy^2 dx dy = \int_0^2 \underbrace{\left[ \frac{1}{2}x^2y^2 \right]_{x=0}^{x=1}}_{\frac{1}{2}y^2-0} dy \\ &= \int_0^2 \frac{1}{2}y^2 dy = \left[ \frac{1}{6}y^3 \right]_0^2 = \frac{8}{6} - 0 = \frac{4}{3}\end{aligned}$$

Hier haben wir also zuerst (innen) nach  $x$  und dann nach  $y$  integriert. Genauso kann man anders herum integrieren, d.h. innen nach  $y$  und außen nach  $x$ :

$$\begin{aligned}\int_Q xy^2 dF &= \int_0^1 \int_0^2 xy^2 dy dx = \int_0^1 \left[ \frac{1}{3}xy^3 \right]_{y=0}^2 dx = \int_0^1 \frac{8}{3}x dx \\ &= \left[ \frac{4}{3}x^2 \right]_0^1 = \frac{4}{3}\end{aligned}$$

Nach dem Satz von Fubini ist das Ergebnis natürlich das gleiche.

## 9.2 Integrale über höherdimensionale Quader

Die Definition des Integrals vom letzten Abschnitt kann man ohne Problem auf höhere Dimensionen, also für Funktionen von drei oder mehr Variablen verallgemeinern. Nur die graphische Veranschaulichung wird dann natürlich schwierig.

Wir betrachten den *dreidimensionalen Fall* ausführlich: Wir haben eine von drei Variablen abhängige Funktion  $f(x, y, z)$ . Der Integrationsbereich  $Q$  ist wieder eine Menge, in der jede einzelne der Variablen  $x, y, z$  in einem Intervall läuft; das ergibt einen dreidimensionalen Quader:

$$\begin{aligned}Q &= \{(x, y, z) \mid a_1 \leq x \leq b_1, a_2 \leq y \leq b_2, a_3 \leq z \leq b_3\} \\ &= [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]\end{aligned}$$

Die zweite Schreibweise für  $Q$  ist wieder die eines kartesischen Produkts: Ein Punkt  $(x, y, z)$  gehört zu  $Q$  genau dann wenn  $x \in [a_1, b_1]$ ,  $y \in [a_2, b_2]$ ,  $z \in [a_3, b_3]$ .

Für die Integration gehen wir wieder von einer stetigen Funktion  $f: Q \rightarrow \mathbb{R}$  aus. Analog zum zweidimensionalen Fall im letzten Abschnitt zerlegen wir  $Q$  in Volumenstücke  $\Delta V_{ijk}$ : jedes der drei Intervalle wird unterteilt,

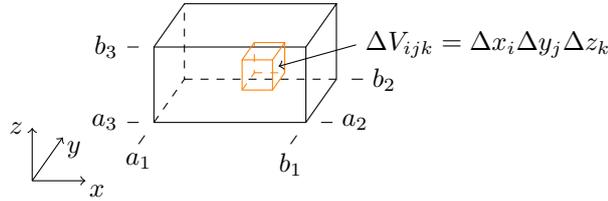
$$\begin{aligned}a_1 &= x_0 < x_1 < \dots < x_n = b_1, \\ a_2 &= y_0 < y_1 < \dots < y_n = b_2, \\ a_3 &= z_0 < z_1 < \dots < z_n = b_3,\end{aligned}$$

so dass sich die Volumenstücke

$$\Delta V_{ijk} = \Delta x_i \Delta y_j \Delta z_k$$

an den Punkten  $(x_i, y_j, z_k)$  ergeben, siehe Abbildung 9.4. Das Integral kann nun wieder als Grenzwert Riemannscher Summen definiert werden:

$$\int_Q f(x, y, z) dV = \lim_{n \rightarrow \infty} \sum_{i,j,k=1}^n f(x_i, y_j, z_k) \Delta V_{ijk}$$

Abbildung 9.4: Unterteilung eines Quaders  $Q$ 

Man nennt das Integral auch *Volumenintegral*, da über den dreidimensionalen Körper (Volumen)  $Q$  integriert wird. Eine andere Schreibweise für das Integral ist

$$\int_Q f(x, y, z) d(x, y, z).$$

Hierbei wird explizit gemacht, dass nach den Variablen  $x, y, z$  integriert wird.

Entsprechend definiert man Integrale nach  $n$  Variablen: Für den  $n$ -dimensionale Quader

$$Q = \{(x_1, \dots, x_n) \mid a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n\}$$

und die stetige Funktion  $f : Q \subset \mathbb{R}^n \rightarrow \mathbb{R}$  hat man das Integral

$$\int_Q f(x_1, \dots, x_n) d(x_1, \dots, x_n)$$

was man mit der schon bekannten Abkürzung  $x = (x_1, \dots, x_n)$  auch kurz als  $\int_Q f(x) dx$  schreiben kann.

Zur Berechnung der Integrale kann man wieder die Formel von *Fubini* benutzen:

$$\int_Q f(x) dx = \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f(x_1, \dots, x_n) dx_1 \dots dx_n \quad (9.3)$$

Genau wie in Satz 9.1.1 ist die Integrationsreihenfolge vertauschbar, und die inneren Integrationsgrenzen (hier  $a_1, b_1$ ) gehören zur inneren Integrationsvariablen (hier  $dx_1$ ), und so weiter von innen nach außen.

### 9.3 Integration über allgemeine Mengen

Nicht immer ist die Menge, über die man eine Funktion mehrerer Variablen integrieren will, ein Rechteck oder Quader. Z.B. könnte es ein Kreis oder dreidimensional eine Kugel sein, oder allgemeiner irgendein krummlinig begrenzter Bereich. Auch für diesen Fall definieren wir jetzt das Integral.

Sei  $f : B \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  stetig und die Menge  $B$  abgeschlossen und beschränkt. Wir wollen  $f$  über  $B$  integrieren und approximieren dazu  $B$  durch Rechteckstücke  $\Delta F_i$ , siehe Abbildung 9.5. Dabei betrachten wir nur die Rechteckstücke, die *ganz* in  $B$  liegen (in der Abbildung gefüllt dargestellt). In jedem solchen  $\Delta F_i$  wählen wir einen Punkt  $(x_i, y_i)$ . Das Integral ist dann wieder definiert als Grenzwert der Riemannschen Summen,

$$\int_B f(x, y) dF = \lim_{\substack{n \rightarrow \infty \\ \Delta F_i \rightarrow 0}} \sum_{i=1}^n f(x_i, y_i) \Delta F_i. \quad (9.4)$$

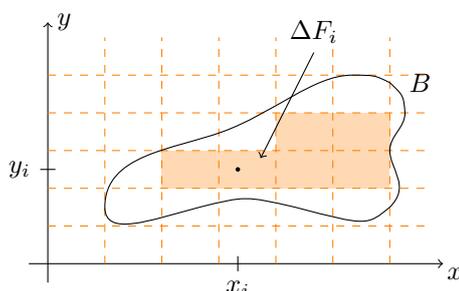


Abbildung 9.5: Unterteilung eines beliebigen Bereichs

Da die Summe nur über die  $\Delta F_i$  geht, die ganz in  $B$  liegen, wird natürlich die Fläche  $B$  nicht ganz ausgeschöpft. Dieser Fehler am Rand von  $B$  wird aber kleiner bei einer feineren Unterteilung, d.h. bei mehr und kleineren Rechtecken  $\Delta F_i$ . Im Grenzwert  $n \rightarrow \infty$ ,  $\Delta F_i \rightarrow 0$  verschwindet der Fehler dann ganz, falls der Rand von  $B$  hinreichend „regulär“ ist. Das Integral – d.h. der Grenzwert in (9.4) – existiert dann.

Analog werden  $n$ -dimensionale Integrale über Bereiche  $B \subset \mathbb{R}^n$  definiert.

**Bemerkung:** Bereiche  $B$  mit „schlechtem Rand“, für die der Fehler nicht verschwindet, sind z.B. Fraktale. Man kann das Integral dann nicht wie hier als Riemann-Integral, d.h. über Riemannsche Summen definieren.

## 9.4 Eigenschaften des Integrals

Wir stellen hier ein paar allgemeine Eigenschaften des Integrals zusammen. Einige davon sind schon von eindimensionalen Integralen aus Mathematik A bekannt.

Sei  $B \subset \mathbb{R}^n$  und  $f, g : B \rightarrow \mathbb{R}$  stetig.

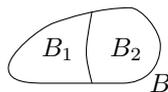
- (a) Das Integral ist *linear*, d.h. es gilt

$$\int_B (\alpha f(x) + \beta g(x)) dx = \alpha \int_B f(x) dx + \beta \int_B g(x) dx, \quad \alpha, \beta \in \mathbb{R}. \quad (9.5)$$

- (b) Das Integral ist *monoton*:

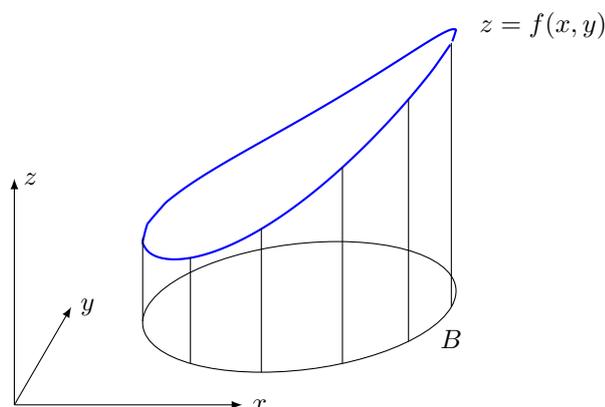
$$f(x) \leq g(x) \Rightarrow \int_B f(x) dx \leq \int_B g(x) dx \quad (9.6)$$

- (c) Das Integral ist *additiv*: Wird  $B$  zerlegt in Teilmengen  $B_1$  und  $B_2$ , also  $B = B_1 \cup B_2$  wobei sich  $B_1$  und  $B_2$  nicht überschneiden,



so gilt

$$\int_B f(x) dx = \int_{B_1} f(x) dx + \int_{B_2} f(x) dx. \quad (9.7)$$

Abbildung 9.6: Körper mit Grundfläche  $B$ , nach oben durch  $z = f(x, y)$  begrenzt

(d) Für eine Fläche (2-dimensionale Menge)  $B \subset \mathbb{R}^2$  gilt:

$$F = \int_B 1 \, dF \quad \text{Flächeninhalt von } B \quad (9.8)$$

$$f(x, y) \geq 0 \Rightarrow V = \int_B f(x, y) \, dF \quad \text{Volumen des Körper mit Grundfläche } B, \text{ der nach oben durch } z = f(x, y) \text{ begrenzt ist.} \quad (9.9)$$

(e) Für einen 3-dimensionalen Körper  $B \subset \mathbb{R}^3$  ist

$$V = \int_B 1 \, dV = \text{Volumen des Körpers } B. \quad (9.10)$$

Integriert man die konstante Funktion 1, so erhält man also bei einer Fläche als Integrationsbereich  $B$  ihren Flächeninhalt (9.8), bei einem 3-dimensionalen Körper sein Volumen (9.10).

## 9.5 Integration über Normalbereiche

Wir haben in Abschnitt 9.3 das Integral zwar für beliebig geformte Mengen  $B$  definiert, konkret berechnen können wir es bis jetzt aber nur für Rechtecke oder höherdimensionale Quader (Satz von Fubini 9.1.1 und (9.3)). Für die Integration über allgemeiner geformte Mengen werden wir im folgenden zwei Möglichkeiten kennenlernen. Die erste sind Normalbereiche.

Seien  $g_1, g_2 : [a, b] \rightarrow \mathbb{R}$  stetig, so dass  $g_1(x) \leq g_2(x)$ . Die Menge

$$B = \{(x, y) \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\} \quad (9.11)$$

heißt *Normalbereich in y-Richtung*. Die Menge enthält also alle Punkte, mit  $x$ -Koordinate zwischen  $a$  und  $b$ , und  $y$ -Koordinate zwischen  $g_1(x)$  und  $g_2(x)$  (abhängig von  $x$ ). Der Normalbereich  $B$  ist damit nach links und rechts durch

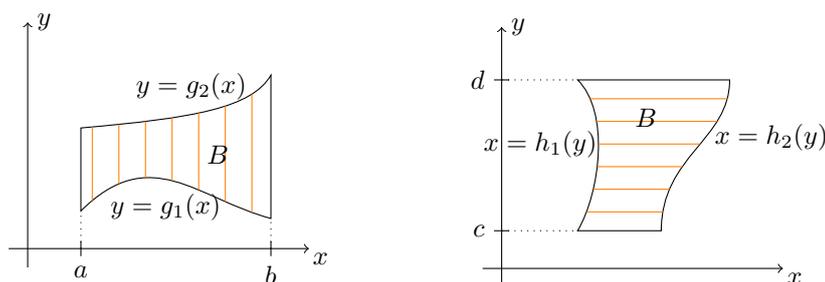
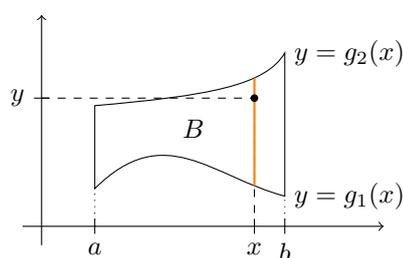
Abbildung 9.7: Normalbereich in  $y$ -Richtung bzw.  $x$ -Richtung

Abbildung 9.8: Zur Integration über Normalbereiche

$x = a$  und  $x = b$  begrenzt, nach unten und oben durch die beiden Funktionen  $y = g_1(x)$  und  $y = g_2(x)$ , siehe Abbildung 9.7 links.

Die Berechnung eines Integrals über den Normalbereich  $B$  geschieht wie beim Satz von Fubini als Doppelintegral:

$$\int_B f(x, y) dF = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) dy dx. \quad (9.12)$$

Anschaulich wird hier im inneren Integral entlang einer Gerade mit festem  $x$ -Wert aus  $[a, b]$  von  $y = g_1(x)$  bis  $y = g_2(x)$  integriert, siehe Abbildung 9.8. Danach wird im äußeren Integral dann von  $x = a$  bis  $b$  integriert, d.h. die Gerade wird sozusagen von  $x = a$  bis  $x = b$  verschoben, so dass von links nach rechts der komplette Bereich  $B$  ausgeschöpft wird.

**Beachte:** Bei der Integration über Normalbereiche ist die Reihenfolge der Integration nicht beliebig; in (9.12) muss die abhängige Grenze  $g_1(x)$ ,  $g_2(x)$  immer innen stehen! D.h. hier muss beim inneren Integral nach  $y$  integriert werden, und beim äußeren dann nach  $x$ .

Durch Vertauschen der Rollen von  $x$  und  $y$  in der Definition von  $B$  in (9.11) erhält man Normalbereiche, die nach links und rechts durch Funktionen begrenzt sind: Für stetige Funktionen  $h_1, h_2 : [c, d] \rightarrow \mathbb{R}$  heißt die Menge

$$B = \{(x, y) \mid c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\} \quad (9.13)$$

*Normalbereich in  $x$ -Richtung*, siehe Abbildung 9.7 rechts. Das Integral über solch eine Menge  $B$  ist dann entsprechend

$$\int_B f(x, y) dF = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y) dx dy. \quad (9.14)$$

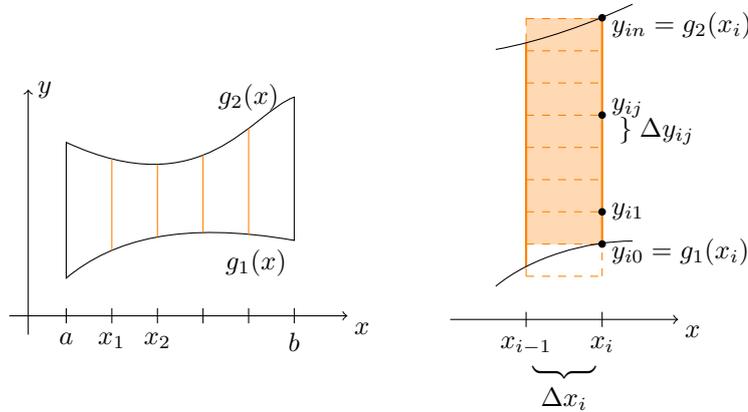


Abbildung 9.9: Unterteilung eines Normalbereichs

Da jetzt die abhängigen Grenzen  $x = h_1(y)$ ,  $x = h_2(y)$  sind, muss hier innen nach  $x$  und außen nach  $y$  integriert werden.

**Beweis zur Integration über Normalbereiche.** Wir zeigen Formel (9.12) zur Integration über einen Normalbereich in  $y$ -Richtung. Der Beweis von (9.14) für Normalbereiche in  $x$ -Richtung verläuft analog.

Wir wollen das Integral über den Normalbereich

$$B = \{(x, y) \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$$

berechnen und zerlegen dazu  $B$  an den Stellen

$$a = x_0 < x_1 < x_2 < \dots < x_n = b$$

in vertikale Streifen der Breite  $\Delta x_i = x_i - x_{i-1}$ . Das ist in Abbildung 9.9 links skizziert. Jeder Streifen wird dann weiter an den Punkten

$$g_1(x_i) = y_{i0} < y_{i1} < \dots < y_{in} = g_2(x_i)$$

in Rechteckstücke der Höhe  $\Delta y_{ij} = y_{ij} - y_{i,j-1}$  zerlegt, dargestellt in Abbildung 9.9 rechts. Wir erhalten so eine Unterteilung des gesamten Bereichs  $B$  in kleine Rechtecke, ähnlich wie in Abbildung 9.5 bei einer allgemeinen Fläche dargestellt. Genau wie dort machen wir hier am Rand von  $B$  einen Fehler, da wir oben und unten mit Rechtecken statt den Kurven  $y = g_1(x)$ ,  $y = g_2(x)$  arbeiten. Da die Randfunktionen  $g_1, g_2$  stetig sind, kann man aber zeigen, dass der Fehler für  $\Delta x_i \rightarrow 0$  gegen Null geht. Der Rand ist somit im Sinne von 9.3 regulär und das Integral existiert.

Wir berechnen nun das Integral mit der allgemeinen Definition (9.4): Die Fläche eines einzelnen Rechteckstücks im Punkt  $(x_i, y_{ij})$  ist

$$\Delta F_{ij} = \Delta x_i \Delta y_{ij}.$$

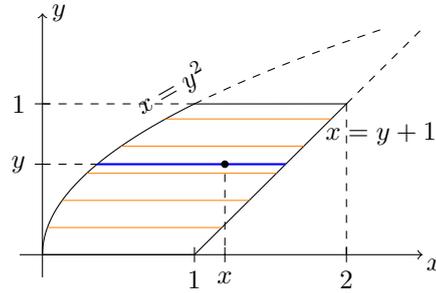


Abbildung 9.10: Die Menge  $B$  aus Beispiel 9.5.1

Damit folgt, analog zum Beweis vom Satz 9.1.1 von Fubini,

$$\begin{aligned} \int_B f(x, y) dF &= \lim_{\substack{n \rightarrow \infty \\ \Delta x_i \rightarrow 0, \Delta y_{ij} \rightarrow 0}} \sum_{i=1}^n \sum_{j=1}^n f(x_i, y_{ij}) \Delta y_{ij} \Delta x_i \\ &= \lim_{\substack{n \rightarrow \infty \\ \Delta x_i \rightarrow 0}} \sum_{i=1}^n \int_{g_1(x_i)}^{g_2(x_i)} f(x_i, y) dy \Delta x_i = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) dy dx. \end{aligned}$$

Hier ist die innere Summe eine Riemannsche Summe über  $y$  von  $y_{i0} = g_1(x_i)$  bis  $y_{in} = g_2(x_i)$  (bei konstantem  $x_i$ ), so dass

$$\lim_{\substack{n \rightarrow \infty \\ \Delta y_{ij} \rightarrow 0}} \sum_{j=1}^n f(x_i, y_{ij}) \Delta y_{ij} = \int_{g_1(x_i)}^{g_2(x_i)} f(x_i, y) dy.$$

Der Rest ist dann eine Riemannsche Summe von  $\int_{g_1(x)}^{g_2(x)} f(x, y) dy$  über die Variable  $x$ , was zum äußeren Integral führt.  $\square$

**Beispiel 9.5.1** Wir betrachten die Menge

$$B = \{(x, y) \mid 0 \leq y \leq 1, y^2 \leq x \leq y + 1\}$$

und wollen das Integral von  $f(x, y) = 1 - x$  über  $B$  berechnen.

Die Menge  $B$  ist ein Normalbereich in  $x$ -Richtung mit den Randfunktionen  $x = h_1(y) = y^2$  und  $x = h_2(y) = y + 1$ . Wir skizzieren zunächst  $B$ , siehe Abbildung 9.10:  $y$  läuft von 0 bis 1. Für jeden festen Wert von  $y$  läuft dann  $x$  von  $x = h_1(y) = y^2$  bis  $x = h_2(y) = y + 1$ . Um diese Randfunktionen zu zeichnen, lösen wir sie nach  $y$  auf:

$$\begin{aligned} x = y^2 &\Leftrightarrow y = \sqrt{x} \quad (+\sqrt{x} \text{ denn } y \geq 0 \text{ in } B!) \\ x = y + 1 &\Leftrightarrow y = x - 1 \end{aligned}$$

(Alternativ kann man beide Funktionen auch in der ursprünglichen Form mit  $y$  als unabhängiger Variablen zeichnen, wobei dann also die unabhängige Variable auf der Hochachse läuft, die abhängige dagegen auf der horizontalen Achse.)

Für jeden Wert von  $y$  erhält man so eine horizontale Gerade, auf der sich ein Punkt  $(x, y) \in B$  befinden kann:  $y^2 \leq x \leq y + 1$  bedeutet, dass  $x$  vom kleinsten

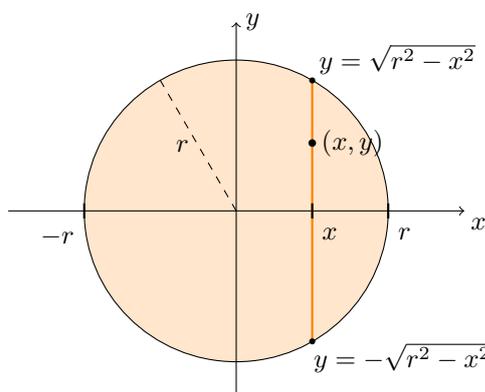


Abbildung 9.11: Ein Kreis als Normalbereich (Beispiel 9.5.2)

möglichen Wert  $x = y^2$  links bis zum größten möglichen Wert  $x = y + 1$  rechts läuft. Die Höhe  $y$  der Geraden läuft dabei von  $y = 0$  bis  $y = 1$ , entsprechend der Ungleichung  $0 \leq y \leq 1$  in  $B$ .

Jetzt berechnen wir noch das Integral von  $f(x, y) = 1 - x$ :

$$\begin{aligned} \int_B (1 - x) dF &= \int_0^1 \int_{y^2}^{y+1} (1 - x) dx dy = \int_0^1 \left[ x - \frac{1}{2}x^2 \right]_{x=y^2}^{y+1} dy \\ &= \int_0^1 \left( y + 1 - \frac{1}{2}(y + 1)^2 - \left( y^2 - \frac{1}{2}y^4 \right) \right) dy \\ &= \int_0^1 \left( y + 1 - \frac{1}{2}y^2 - y - \frac{1}{2} - y^2 + \frac{1}{2}y^4 \right) dy \\ &= \int_0^1 \left( \frac{1}{2}y^4 - \frac{3}{2}y^2 + \frac{1}{2} \right) dy \\ &= \left[ \frac{1}{10}y^5 - \frac{1}{2}y^4 + \frac{1}{2}y \right]_0^1 = \frac{1}{10} - \frac{1}{2} + \frac{1}{2} = \frac{1}{10} \end{aligned}$$

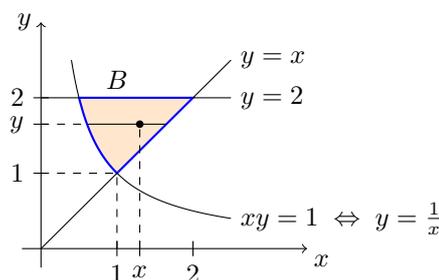
Im letzten Beispiel war die Menge  $B$ , über die integriert werden soll, schon als Normalbereich – also in der Form (9.11) oder (9.13) – gegeben. Oft ist das aber nicht der Fall, und man muss die Menge  $B$  erst auf die Form eines Normalbereichs bringen, damit man dann z.B. ein Integral berechnen kann. Das machen wir im nächsten Beispiel:

**Beispiel 9.5.2** Sei  $B$  die Kreisscheibe mit Radius  $r$  und Mittelpunkt in  $(0, 0)$ . Also

$$B = \{(x, y) \mid x^2 + y^2 \leq r^2\}.$$

Wir wollen  $B$  als Normalbereich in  $y$ -Richtung schreiben, also in der Form (9.11). Zuerst bestimmen wir den maximalen und minimalen Wert von  $x$ , der in der Menge  $B$  vorkommt. Der Punkt mit dem größten  $x$ -Wert ist der rechte Randpunkt des Kreises auf der  $x$ -Achse; da  $r$  der Radius ist, ist das der Punkt  $(r, 0)$ . Den kleinsten  $x$ -Wert haben wir am linken Rand im Punkt  $(-r, 0)$ . Für die  $x$ -Werte in  $B$  bekommen wir also das Intervall

$$-r \leq x \leq r.$$

Abbildung 9.12: Die Menge  $B$  aus Beispiel 9.5.3

Für jeden Wert von  $x$  in diesem Intervall bestimmen wir nun die  $y$ -Werte, die ein Punkt aus  $B$  für diesen festen  $x$ -Wert annehmen kann. Graphisch bedeutet das: wir betrachten die Gerade parallel zur  $y$ -Achse mit diesem konstanten  $x$ -Wert und bestimmen die  $y$ -Koordinate aller Punkte auf der Gerade, die im Kreis liegen, vergleiche Abbildung 9.11. Das sind alle Punkte zwischen den Schnittpunkten der Gerade mit dem oberen und unteren Kreisrand. Wir brauchen also die  $y$ -Koordinate dieser beiden Schnittpunkte. Der Rand des Kreises ist durch die Gleichung  $x^2 + y^2 = r^2$  gegeben. Die  $y$ -Koordinate eines Punktes auf dem Rand zu gegebenem  $x$  ist damit

$$x^2 + y^2 = r^2 \quad \Leftrightarrow \quad y = \pm \sqrt{r^2 - x^2},$$

am oberen Rand ist  $y = \sqrt{r^2 - x^2}$ , am unteren Rand  $y = -\sqrt{r^2 - x^2}$ . Alle Punkte dazwischen liegen in  $B$ , so dass für diese Punkte also gilt

$$-\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}.$$

Damit ist die gesuchte Form des Kreises  $B$  als Normalbereich in  $y$ -Richtung

$$B = \left\{ (x, y) \mid -r \leq x \leq r, -\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2} \right\}.$$

Insbesondere sind also die Funktionen  $g_1, g_2$  aus (9.11), die die Menge nach unten und oben begrenzen,

$$g_1(x) = -\sqrt{r^2 - x^2}, \quad g_2(x) = \sqrt{r^2 - x^2}.$$

Mit der Darstellung als Normalbereich können wir jetzt Funktionen  $f(x, y)$  über den Kreis  $B$  integrieren:

$$\int_B f(x, y) dF = \int_{-r}^r \int_{-\sqrt{r^2 - x^2}}^{\sqrt{r^2 - x^2}} f(x, y) dy dx.$$

Eine andere Möglichkeit zur Integration über Kreise hat man mit der Transformation auf Polarkoordinaten; wir schauen uns das in Abschnitt 9.10 an.

**Beispiel 9.5.3** Sei  $B$  die Fläche im 1. Quadranten (d.h. mit  $x \geq 0$  und  $y \geq 0$ ), die begrenzt wird von den Kurven

$$y = x, \quad xy = 1 \quad \text{und} \quad y = 2. \quad (9.15)$$

Gesucht ist  $B$  als Normalbereich. Wir skizzieren zunächst die Menge, siehe Abbildung 9.12. Der Rand von  $B$  ist durch die drei Kurven (9.15) gegeben, die wir also zuerst zeichnen. Dabei ist

$$xy = 1 \quad \Leftrightarrow \quad y = \frac{1}{x}.$$

Wir sehen, dass  $B$  in  $x$ -Richtung betrachtet eine linke und eine rechte Randfunktion hat, nämlich  $y = \frac{1}{x}$  und  $y = x$ . Bei Betrachtung in  $y$ -Richtung gibt es aber *zwei* untere Randfunktionen (wieder  $y = \frac{1}{x}$  und  $y = x$ ). Hier ist es daher sinnvoll,  $B$  als Normalbereich in  $x$ -Richtung zu betrachten.<sup>1</sup> Wir suchen also eine Darstellung von  $B$  in der Form (9.13). Aus der Skizze lesen wir als maximales  $y$ -Intervall

$$1 \leq y \leq 2$$

ab. Den Schnittpunkt  $(1, 1)$  der Kurven  $xy = 1$  und  $y = x$  erhält man dabei durch

$$xy = 1 \wedge y = x \Rightarrow x^2 = 1 \Rightarrow x = 1 \text{ (denn } x \geq 0!) \Rightarrow y = 1$$

Für ein festes  $y$  zwischen 1 und 2 liegt der  $x$ -Wert eines Punktes in  $B$  dann zwischen der linken Randkurve

$$xy = 1 \quad \Leftrightarrow \quad x = \frac{1}{y}$$

und der rechten

$$y = x \quad \Leftrightarrow \quad x = y.$$

Für  $x$  erhalten wir damit das Intervall

$$\frac{1}{y} \leq x \leq y.$$

Also hat  $B$  als Normalbereich in  $x$ -Richtung die Gestalt

$$B = \left\{ (x, y) \mid 1 \leq y \leq 2, \frac{1}{y} \leq x \leq y \right\}.$$

Aus diesem Ergebnis können wir jetzt (z.B.) den Flächeninhalt  $F$  von  $B$  berechnen (siehe (9.8)):

$$\begin{aligned} F &= \int_B 1 \, dF = \int_1^2 \int_{1/y}^y 1 \, dx \, dy = \int_1^2 [x]_{x=1/y}^{x=y} \, dy = \int_1^2 \left( y - \frac{1}{y} \right) \, dy \\ &= \left[ \frac{1}{2} y^2 - \ln y \right]_1^2 = 2 - \ln(2) - \left( \frac{1}{2} - \ln(1) \right) = \frac{3}{2} - \ln(2) \end{aligned}$$

## 9.6 Vertauschen der Integrationsreihenfolge bei Normalbereichen

Nach dem Satz von Fubini ist die Integrationsreihenfolge bei Integration über Rechtecke oder Quader vertauschbar. Auch bei der Integration über Normalbereiche kann man die Reihenfolge der Integration vertauschen, allerdings ändern

<sup>1</sup>Bei Darstellung als Normalbereich in  $y$ -Richtung würde man  $B$  zuerst in zwei Teilmengen zerlegen; ein Beispiel dafür gibt es im nächsten Abschnitt.

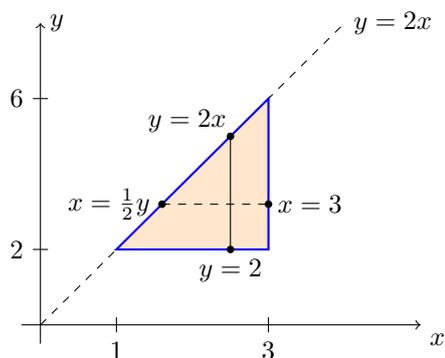


Abbildung 9.13: Vertauschen der Integrationsgrenzen, Beispiel 9.6.1(a)

sich dann die Integrationsgrenzen. Allgemein gilt

$$\int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) \, dy \, dx = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y) \, dx \, dy.$$

Ist das Integral links gegeben (Normalbereich in  $y$ -Richtung), so müssen beim Vertauschen die neuen Grenzen  $c, d, h_1(y), h_2(y)$  für das Integral rechts berechnet werden. Startet man rechts (Normalbereich in  $x$ -Richtung), so muss man die Grenzen  $a, b, g_1(x), g_2(x)$  bestimmen. Andersgesagt wird also ein Normalbereich in  $y$ -Richtung in einen Normalbereich in  $x$ -Richtung umgerechnet und andersrum. Damit kann es passieren, dass der Integrationsbereich  $B$  in mehrere Teilmengen zerlegt werden muss.

**Beispiel 9.6.1** (a) Gegeben ist das Doppelintegral

$$\int_2^6 \int_{\frac{1}{2}y}^3 f(x, y) \, dx \, dy.$$

Es soll die Reihenfolge der Integration vertauscht werden. Dazu schreiben wir das Doppelintegral zuerst als Flächenintegral über den entsprechenden Normalbereich:

$$\int_2^6 \int_{\frac{1}{2}y}^3 f(x, y) \, dx \, dy = \int_B f(x, y) \, dF$$

wobei

$$B = \left\{ (x, y) \mid 2 \leq y \leq 6, \frac{1}{2}y \leq x \leq 3 \right\}, \quad (9.16)$$

das ist also ein Normalbereich in  $x$ -Richtung. Wir müssen nun  $B$  als Normalbereich in  $y$ -Richtung schreiben, also in der Form (9.11). Um die neuen Intervallgrenzen für  $x$  und  $y$  zu bestimmen, skizzieren wir  $B$ , siehe Abbildung 9.13: Die Formel (9.16) sagt uns, dass  $y$  von 2 bis 6 läuft und  $x$ , für jeden festen Wert von  $y$ , von  $x = \frac{1}{2}y$  (linke Grenze) bis  $x = 3$  (rechte Grenze). Die linke Grenze ist äquivalent zu

$$x = \frac{1}{2}y \quad \Leftrightarrow \quad y = 2x.$$

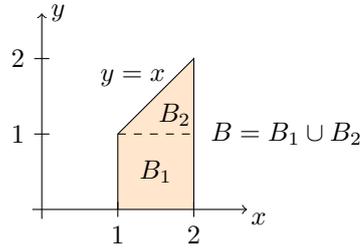


Abbildung 9.14: Vertauschen der Integrationsgrenzen, Beispiel 9.6.1(b)

Für  $B$  erhalten wir damit das skizzierte Dreieck. Um jetzt  $B$  als Normalbereich bezüglich  $y$  darzustellen, brauchen wir das maximale  $x$ -Intervall. Die  $x$ -Koordinate der linken unteren Ecke des Dreiecks ist  $x = \frac{1}{2}y = \frac{1}{2} \cdot 2 = 1$ . Damit ist das  $x$ -Intervall

$$1 \leq x \leq 3.$$

Für festen  $x$ -Wert läuft  $y$  von  $y = 2$  (untere Grenze) bis  $y = 2x$  (obere Grenze), also

$$2 \leq y \leq 2x.$$

Damit ist die Darstellung von  $B$  als Normalbereich in  $y$ -Richtung

$$B = \{(x, y) \mid 1 \leq x \leq 3, 2 \leq y \leq 2x\}.$$

Damit können wir jetzt die Reihenfolge der Integration vertauschen:

$$\int_2^6 \int_{\frac{1}{2}y}^3 f(x, y) dx dy = \int_B f(x, y) dF = \int_1^3 \int_2^{2x} f(x, y) dy dx.$$

- (b) In folgendem Beispiel muss die Menge  $B$  zum Vertauschen der Integration zerlegt werden. Gegeben ist

$$\int_1^2 \int_0^x f(x, y) dy dx$$

und die Integrationsreihenfolge soll getauscht werden. Das Integral ist

$$\int_B f(x, y) dF \quad \text{mit} \quad B = \{(x, y) \mid 1 \leq x \leq 2, 0 \leq y \leq x\}.$$

Die Menge  $B$  ist ein Normalbereich in  $y$ -Richtung, mit oberer Randfunktion  $y = x$ , skizziert in Abbildung 9.14. Wenn wir  $B$  als Normalbereich in  $x$ -Richtung betrachten, sehen wir, dass der linke Rand aus zwei Funktionen besteht,  $x = 1$  im Bereich  $0 \leq y \leq 1$  sowie  $y = x$  für  $1 \leq y \leq 2$ . Wir zerlegen daher  $B$  in zwei Teile,

$$B = B_1 \cup B_2,$$

einen Teil für jedes der  $y$ -Intervalle.  $B_1$  ist einfach ein Rechteck (sogar ein Quadrat)

$$B_1 = \{(x, y) \mid 1 \leq x \leq 2, 0 \leq y \leq 1\} = [1, 2] \times [0, 1],$$

hier können wir also die Integrationsreihenfolge mit Fubini (Satz 9.1.1) direkt vertauschen. Die Teilmenge  $B_2$  können wir als Normalbereich in  $x$ -Richtung schreiben mit linkem Rand  $y = x \Leftrightarrow x = y$  und rechtem Rand  $x = 2$ , d.h.  $y \leq x \leq 2$ , also

$$B_2 = \{(x, y) \mid 1 \leq y \leq 2, y \leq x \leq 2\}.$$

Wegen (9.7) (Additivität des Integrals) ist das Integral über  $B$  die Summe der beiden Integrale über  $B_1$  und  $B_2$  und wir erhalten damit

$$\begin{aligned} \int_1^2 \int_0^x f(x, y) dy dx &= \int_B f(x, y) dF \\ &= \int_{B_1} f(x, y) dF + \int_{B_2} f(x, y) dF \\ &= \int_0^1 \int_1^2 f(x, y) dx dy + \int_1^2 \int_y^2 f(x, y) dx dy \end{aligned}$$

## 9.7 Dreidimensionale Normalbereiche

Die Definition von Normalbereichen lässt sich leicht auf drei Dimensionen verallgemeinern. Für stetige Funktionen  $g_1(x)$ ,  $g_2(x)$ ,  $h_1(x, y)$ ,  $h_2(x, y)$  ist

$$B = \{(x, y, z) \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x), h_1(x, y) \leq z \leq h_2(x, y)\} \quad (9.17)$$

ein Normalbereich.  $B$  ist in  $z$ -Richtung durch die Funktionen  $z = h_1(x, y)$  und  $z = h_2(x, y)$  begrenzt. Die möglichen Werte von  $x$  und  $y$  liegen dabei in einer „Grundfläche“, die durch  $a \leq x \leq b$  und  $g_1(x) \leq y \leq g_2(x)$  gegeben ist (also ein zweidimensionaler Normalbereich). Integration über  $B$  erfolgt in naheliegender Weise durch ein Dreifachintegral:

$$\int_B f(x, y, z) dV = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x, y)}^{h_2(x, y)} f(x, y, z) dz dy dx.$$

Wie bei zwei Dimensionen sind Normalbereiche mit anderer Reihenfolge von  $x, y, z$  (und entsprechend anderen Funktionen) möglich, etwa

$$B = \{(x, y, z) \mid a \leq z \leq b, g_1(z) \leq y \leq g_2(z), h_1(y, z) \leq x \leq h_2(y, z)\}.$$

**Beispiel 9.7.1** Wir betrachten die Kugel mit Radius 1 um den Nullpunkt

$$B = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1\}.$$

Gesucht ist die Darstellung von  $B$  als Normalbereich. Um eine Form wie in (9.17) zu bekommen, suchen wir (in  $z$ -Richtung) obere und untere Randfunktionen  $z = h_1(x, y)$  und  $z = h_2(x, y)$  sowie eine entsprechende Grundfläche. Für die Kugel suchen wir also Gleichungen für den oberen und unteren Kugelrand (d.h. die Kugeloberfläche). Der Kugelrand sind alle Punkte mit Abstand 1 zum Mittelpunkt, also

$$x^2 + y^2 + z^2 = 1.$$

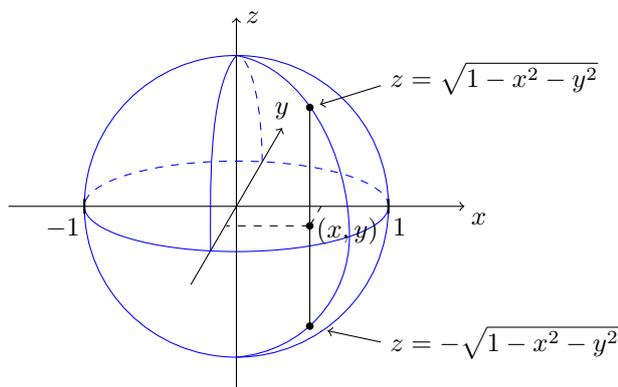


Abbildung 9.15: Kugel als Normalbereich (Beispiel 9.7.1)

Auflösen nach  $z$  gibt die gesuchten Randfunktionen:

$$\begin{aligned} x^2 + y^2 + z^2 = 1 &\Leftrightarrow z^2 = 1 - x^2 - y^2 \\ &\Leftrightarrow z = \pm\sqrt{1 - x^2 - y^2} \\ \Rightarrow h_1(x, y) = -\sqrt{1 - x^2 - y^2}, \quad h_2(x, y) = \sqrt{1 - x^2 - y^2} \end{aligned}$$

$h_1(x, y) = -\sqrt{1 - x^2 - y^2}$  beschreibt den unteren Kugelrand (weil negativ),  $h_2(x, y) = \sqrt{1 - x^2 - y^2}$  den oberen Rand, siehe Abbildung 9.15. Für die Punkte *in* der Kugel gilt damit

$$-\sqrt{1 - x^2 - y^2} \leq z \leq \sqrt{1 - x^2 - y^2}.$$

Es fehlt jetzt noch die Grundfläche, also der maximal mögliche Bereich, in dem die  $(x, y)$ -Koordinaten eines Punktes in der Kugel liegen können. Geometrisch kann man dazu entlang der  $z$ -Achse von oben auf die Kugel blicken. Man sieht dann auf die  $x, y$ -Ebene, und die Punkte der Kugel bilden von oben gesehen genau den Kreis mit Radius 1 um den Nullpunkt.  $(x, y)$  ist also ein Punkt im Kreis in der  $x, y$ -Ebene. Den Kreis haben wir schon in Beispiel 9.5.2 als Normalbereich dargestellt; hier ist der Radius  $r = 1$  und wir erhalten daher

$$-1 \leq x \leq 1, \quad -\sqrt{1 - x^2} \leq y \leq \sqrt{1 - x^2}$$

als maximalen Bereich für die  $x, y$ -Koordinaten. Insgesamt wird damit die Kugel als Normalbereich beschrieben durch

$$B = \left\{ (x, y, z) \left| \begin{array}{l} -1 \leq x \leq 1, \quad -\sqrt{1 - x^2} \leq y \leq \sqrt{1 - x^2}, \\ -\sqrt{1 - x^2 - y^2} \leq z \leq \sqrt{1 - x^2 - y^2} \end{array} \right. \right\}.$$

Für das Integral über die Kugel bekommen wir entsprechend

$$\int_B f(x, y, z) dF = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} f(x, y, z) dz dy dx.$$

Eine andere Möglichkeit zur Integration über die Kugel (oder Teile davon) werden wir später durch Transformation auf Kugelkoordinaten (Abschnitt 9.13) bekommen.

## 9.8 Eine Anwendung des Integrals: Dichten

Was ist die Bedeutung oder Interpretation des Integrals? Das hängt stark vom Anwendungszusammenhang ab. Oder anders gesagt: Mit Integralen kann man viele verschiedene Sachen ausrechnen.

Eine mögliche Interpretation des Werts von Integralen sind natürlich Flächen und Volumina, siehe die Formeln in Abschnitt 9.4. In vielen Fällen hat das Ergebnis einer Integralberechnung aber *nicht* die Bedeutung einer Fläche oder eines Volumens. Z.B. dann, wenn die Funktion die integriert wird eine Dichte ist. Wir erläutern das zunächst an der *Massendichte*:

Wir betrachten den Körper  $B \subset \mathbb{R}^3$ . Der Körper habe die Masse  $M$  und das Volumen  $V$ . Ist der Körper  $B$  *homogen* (also in jedem Punkt aus dem gleichen Material mit den gleichen Eigenschaften aufgebaut), so hat  $B$  die *Dichte* (genauer: Massendichte)

$$\varrho = \frac{M}{V}$$

(Masse pro Volumeneinheit). Die Dichte ist eine Eigenschaft des Materials, und wenn der Körper homogen ist, ist diese Dichte in jedem Punkt des Körpers gleich.

Jetzt nehmen wir an, der Körper sei nicht homogen. Dann kann die Dichte in jedem Punkt des Körpers unterschiedlich sein, d.h. die Dichte  $\varrho$  ist *ortsabhängig*,

$$\varrho = \varrho(x, y, z).$$

Wie kann man die Masse von  $B$  aus der Dichtefunktion  $\varrho(x, y, z)$  berechnen?

Dazu betrachtet man ein infinitesimales Volumen  $dV$  im Punkt  $(x, y, z)$  des Körpers, d.h. einen „unendlich kleinen“ Quader mit den Kanten  $dx, dy, dz$  und Volumen  $dV = dx dy dz$ . Das infinitesimale Volumen  $dV$  hat die (ebenfalls infinitesimale) Masse

$$dM = \varrho(x, y, z) dV,$$

denn  $dV$  hat im Punkt  $(x, y, z)$  genau die Dichte  $\varrho(x, y, z)$ . Durch Integration über den ganzen Körper  $B$  erhält man dann die Gesamtmasse  $M$  des Körpers:

$$M = \int_B dM = \int_B \varrho(x, y, z) dV. \quad (9.18)$$

Das ist analog zur Definition des Integrals, wo eine Riemannsche Summe über kleine Quader gebildet wird und sich dann im Grenzwert „unendlich kleiner“ Quader das Integral ergibt.

### Beispiel 9.8.1 Ein Würfel

$$B = [0, 1] \times [0, 1] \times [0, 1]$$

sei so beschaffen, dass seine Dichte durch

$$\varrho(x, y, z) = 1 + x + y$$

gegeben ist. Die Dichte ist also unabhängig von der Höhe  $z$ . Sie ist am kleinsten für  $x = y = 0$  ( $\Rightarrow \varrho = 1$ ), also vorne links, siehe Abbildung 9.16. Sie ist maximal mit  $\varrho = 3$  für  $x = y = 1$  (hinten rechts), und entlang der Diagonale  $x + y = 1$  (Diagonalebene von hinten links nach vorne rechts) ist sie gleich 2.

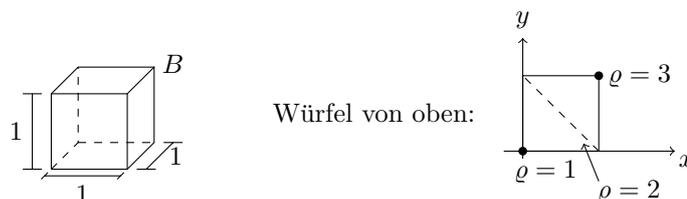


Abbildung 9.16: Beispiel zur Dichte (Beispiel 9.8.1)

Wir können jetzt die Gesamtmasse des Würfels  $B$  durch Integrieren der Dichte  $\varrho(x, y, z)$  berechnen:

$$\begin{aligned}
 M &= \int_B (1 + x + y) dV = \int_0^1 \int_0^1 \int_0^1 (1 + x + y) dz dy dx \\
 &= \int_0^1 \int_0^1 [(1 + x + y)z]_{z=0}^1 dy dx = \int_0^1 \int_0^1 (1 + x + y) dy dx \\
 &= \int_0^1 \left[ (1 + x)y + \frac{1}{2}y^2 \right]_{y=0}^1 dx = \int_0^1 \left( 1 + x + \frac{1}{2} \right) dx \\
 &= \left[ \frac{3}{2}x + \frac{1}{2}x^2 \right]_0^1 = \frac{3}{2} + \frac{1}{2} = 2
 \end{aligned}$$

Der Körper hat also die Masse 2.

In Anwendungen treten auch *andere Dichten* auf. Beispiele sind:

- Ladungsdichte  $\varrho$  (Ladung pro Volumen): Integrieren über einen Körper  $B \subset \mathbb{R}^3$  führt dann auf die Gesamtladung

$$Q = \int_B \varrho(x, y, z) dV.$$

- Flächenladungsdichte  $\sigma$ : Dies ist die Ladungsdichte auf einer (Ober-)Fläche  $B \subset \mathbb{R}^2$ , als Ladung pro Flächeneinheit. Die (Gesamt-)Ladung der Fläche ist dann

$$Q = \int_B \sigma(x, y) dF.$$

## 9.9 Der Transformationssatz

Der Transformationssatz ist die Verallgemeinerung der Substitutionsregel bei Integralen von einer Variablen auf mehrere Variablen. Die Substitutionsregel für eine Variable  $x \in \mathbb{R}$  lautet

$$\int_{\varphi(a)}^{\varphi(b)} f(x) dx = \int_a^b f(\varphi(u))\varphi'(u) du. \quad (9.19)$$

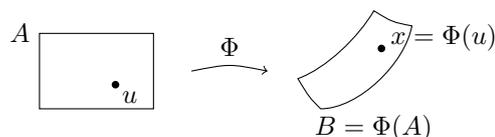


Abbildung 9.17: Transformation des Integrationsbereichs

Durch die Substitution  $x = \varphi(u)$  wird hier die neue Variable  $u$  eingeführt. In mehreren Dimensionen entspricht dies einer Koordinatentransformation

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \Phi(u_1, \dots, u_n), \quad (9.20)$$

bei der die kartesischen Koordinaten  $x = (x_1, \dots, x_n)$  durch neue Koordinaten  $u = (u_1, \dots, u_n)$  ausgedrückt werden. Die Umrechnung (Transformation) der neuen in die alten Koordinaten geschieht durch die Funktion  $x = \Phi(u)$ .

Eine Koordinatentransformation bewirkt auch eine Transformation des Integrationsbereichs  $B$  (also der Menge über die integriert wird). Im eindimensionalen Integral (9.19) entspricht das genau der Substitution der Integralgrenzen. Sei  $A \subset \mathbb{R}^n$  der Integrationsbereich in den neuen Koordinaten  $u$ . Durch die Transformation (9.20) wird daraus der entsprechende Bereich in den  $x$ -Koordinaten

$$B = \{x = \Phi(u) \mid u \in A\},$$

was wir auch kurz als  $B = \Phi(A)$  schreiben. Typischerweise wird bei der Transformation von  $A$  zu  $B$  die Menge verformt, was in Abbildung 9.17 illustriert ist.

Damit durch (9.20) eine sinnvolle Koordinatentransformation gegeben ist, muss die Zuordnung  $\Phi : u \mapsto x$  umkehrbar eindeutig sein, d.h. jeder Punkt  $x$  muss eindeutige  $u$ -Koordinaten haben. Anders gesagt dürfen zwei verschiedene Punkte  $u$  und  $\tilde{u}$  nicht auf denselben Punkt  $x$  abgebildet werden; die Funktion  $\Phi$  muss also *injektiv* sein,

$$u \neq \tilde{u} \quad \Rightarrow \quad \Phi(u) \neq \Phi(\tilde{u}).$$

Die Interpretation von  $\Phi$  als Koordinatentransformation bedeutet, dass  $u$  und  $x = \Phi(u)$  denselben Punkt in  $\mathbb{R}^n$  beschreiben, einmal in den ursprünglichen kartesischen Koordinaten  $x$ , und einmal in den neuen Koordinaten  $u$ . Entsprechend sind auch  $A$  und  $B$  „dieselbe“ Menge:  $B$  ist die Menge dargestellt in  $x$ -Koordinaten,  $A$  die Menge in  $u$ -Koordinaten.

Der folgende Transformationssatz ist nun die Verallgemeinerung der eindimensionalen Substitutionsregel:

**Satz 9.9.1 (Transformationssatz)** Sei  $\Phi : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  injektiv, stetig differenzierbar und  $\det J_\Phi(u) \neq 0$ . Sei  $A \subset G$  und  $B = \Phi(A)$ . Dann gilt

$$\int_B f(x) dx = \int_A f(\Phi(u)) |\det J_\Phi(u)| du \quad (9.21)$$

wobei  $x = (x_1, \dots, x_n)$ ,  $u = (u_1, \dots, u_n)$ .

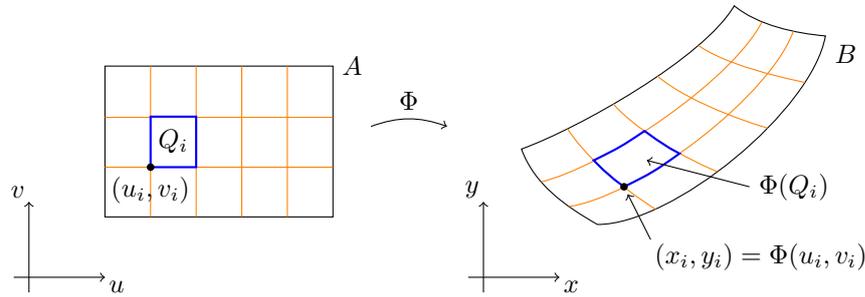


Abbildung 9.18: Zum Beweis des Transformationssatzes

**Beweis.** Der vollständige, exakte Beweis des Transformationssatzes ist sehr lang und aufwändig. Wir geben hier nur die Beweisidee für  $n = 2$  und  $n = 3$  an.

Für  $n = 2$  nennen wir die alten, kartesischen Koordinaten  $x, y$ , die neuen Koordinaten  $u, v$ . Die Transformation ist also

$$\begin{pmatrix} x \\ y \end{pmatrix} = \Phi(u, v).$$

Gemäß der Definition des Integrals betrachten wir eine Zerlegung von  $A$  in kleine Rechtecke  $Q_i$ . Über die Transformation  $\Phi$  gibt das eine Zerlegung von  $B$  in Flächenstücke  $\Phi(Q_i)$ , siehe Abbildung 9.18. Im Grenzübergang zum Integral wird aus dem Rechteck  $Q_i$  das infinitesimale Rechteck  $Q$  im Punkt  $(u, v)$  mit den infinitesimalen Seitenlängen  $du, dv$ . Um das transformierte Flächenstück  $\Phi(Q)$  zu bestimmen, betrachten wir das totale Differential von  $\Phi$ :

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = J_{\Phi}(u, v) \begin{pmatrix} du \\ dv \end{pmatrix}$$

(siehe (8.26)). Bezeichnen wir die Einträge der Jacobimatrix mit

$$J_{\Phi}(u, v) = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

so folgt

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} = du \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + dv \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (9.22)$$

Die Änderung  $(du, dv)$  des Punktes  $(u, v)$  entspricht unter der Transformation der Änderung  $(dx, dy)$  des Punktes  $(x, y)$ . Nach (9.22) wird diese Änderung von den zwei Vektoren  $du \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$  und  $dv \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$  aufgespannt, dargestellt in Abbildung 9.19. Somit ist  $\Phi(Q)$  ein Parallelogramm. Die Fläche dieses Parallelogramms können wir mit dem Vektorprodukt berechnen (vergleiche Mathematik A):

$$dF = \left| du \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times dv \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \right| = |a_1 b_2 - a_2 b_1| du dv = |\det J_{\Phi}(u, v)| du dv$$

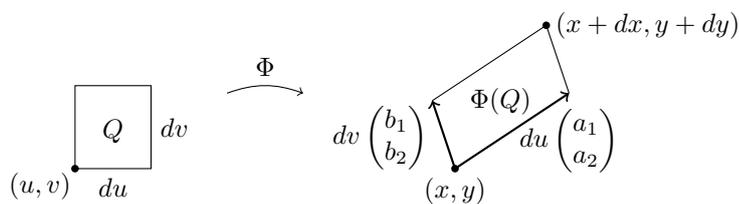


Abbildung 9.19: Transformation eines infinitesimalen Rechtecks

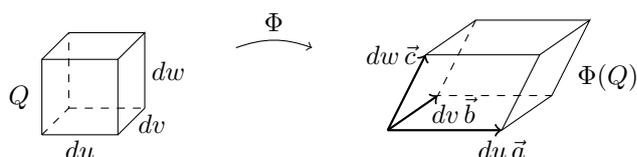


Abbildung 9.20: Transformation eines infinitesimalen Quaders

Integration über diese infinitesimale Fläche  $dF$  des Flächenstücks  $\Phi(Q)$  im Punkt  $(x, y) = \Phi(u, v)$  von  $B$  liefert damit

$$\int_B f(x, y) d(x, y) = \int_B f(x, y) dF = \int_A f(\Phi(u, v)) |\det J_\Phi(u, v)| du dv.$$

Jetzt zum Fall  $n = 3$ . Die Transformation ist dann

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \Phi(u, v, w)$$

und das totale Differential ergibt nun

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = J_\Phi(u, v, w) \begin{pmatrix} du \\ dv \\ dw \end{pmatrix} = du \vec{a} + dv \vec{b} + dw \vec{c}$$

wobei die Vektoren  $\vec{a}, \vec{b}, \vec{c}$  genau die Spalten der Jacobimatrix  $J_\Phi(u, v, w)$  sind. Diese Gleichung bedeutet jetzt, analog zum zweidimensionalen Fall, dass der infinitesimale Quader  $Q$  mit Seiten  $du, dv, dw$  auf einen Spat  $\Phi(Q)$  abgebildet wird, vergleiche Abbildung 9.20. Das Volumen dieses Spats berechnet sich mit dem Spatprodukt (aus Mathematik A...) zu

$$dV = |[du \vec{a}, dv \vec{b}, dw \vec{c}]| = |\det(\vec{a}, \vec{b}, \vec{c})| du dv dw = |\det J_\Phi(u, v, w)| du dv dw.$$

Damit folgt dann

$$\begin{aligned} \int_B f(x, y, z) d(x, y, z) &= \int_B f(x, y, z) dV \\ &= \int_A f(\Phi(u, v, w)) |\det J_\Phi(u, v, w)| du dv dw. \end{aligned}$$

□

## 9.10 Transformationssatz für Polarkoordinaten

In Polarkoordinaten ist die Transformationsfunktion

$$\begin{pmatrix} x \\ y \end{pmatrix} = \Phi(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix},$$

siehe Abschnitt 8.20; die Jacobimatrix ist

$$J_{\Phi}(r, \varphi) = \begin{pmatrix} \frac{\partial \Phi}{\partial r} & \frac{\partial \Phi}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

und somit

$$\det J_{\Phi}(r, \varphi) = r \cos^2 \varphi + r \sin^2 \varphi = r.$$

Im letzten Abschnitt war eine Forderung an die Transformationsfunktion  $\Phi$ , dass sie injektiv ist. Dafür muss man bei Polarkoordinaten den Winkel  $\varphi$  auf ein Intervall der Länge  $2\pi$  einschränken, denn der Polarkoordinatenwinkel  $\varphi$  ist allgemein nur bis auf Vielfache von  $2\pi$  bestimmt. (Oder anders gesagt:  $\varphi$  und  $\varphi \pm 2\pi$  ergeben denselben Punkt.) Typischerweise schränkt man den Winkel auf

$$0 \leq \varphi < 2\pi \quad \text{oder} \quad -\pi < \varphi \leq \pi \quad (9.23)$$

ein. Der *Transformationssatz für Polarkoordinaten* ist dann

$$\begin{aligned} \int_B f(x, y) d(x, y) &= \int_A f(\Phi(r, \varphi)) |\det J_{\Phi}(r, \varphi)| d(r, \varphi) \\ &= \int_A f(r \cos \varphi, r \sin \varphi) \cdot r d(r, \varphi) \end{aligned} \quad (9.24)$$

**Beachte:**

- Weil für den Radius in Polarkoordinaten immer  $r \geq 0$  gilt, erhalten wir in (9.24)  $|\det J_{\Phi}(r, \varphi)| = |r| = r$ .

Wichtig ist, bei Transformation auf Polarkoordinaten diesen zusätzlichen Faktor  $r$  im Integral nicht zu vergessen!

- Eine Voraussetzung im Transformationssatz 9.9.1 ist  $\det J_{\Phi} \neq 0$ , hier also  $\det J_{\Phi}(r, \varphi) = r \neq 0$ . Für den Punkt  $(x, y) = (0, 0)$  gilt  $r = 0$ , sonst ist  $r > 0$ . Im Punkt  $(0, 0)$  ist also die Voraussetzung  $\det J_{\Phi} \neq 0$  *nicht* erfüllt. Der Transformationssatz ist aber trotzdem anwendbar, da man einen einzelnen Punkt im Flächenintegral vernachlässigen kann (denn ein Punkt hat den Flächeninhalt Null).

**Beispiel 9.10.1** Wir betrachten den Integrationsbereich

$$B = \{(x, y) \mid x^2 + y^2 \leq 16, x \geq 0\}.$$

Die Ungleichung  $x^2 + y^2 \leq 16 = 4^2$  beschreibt einen Kreis mit Radius 4. Wegen  $x \geq 0$  gehören zu  $B$  aber nur die Punkte mit positiver  $x$ -Koordinate, also in der Halbebene rechts von der  $y$ -Achse.  $B$  ist daher der rechte Halbkreis mit Radius 4 um den Nullpunkt, siehe Abbildung 9.21.

Wir können  $B$  sehr schön in Polarkoordinaten beschreiben:  $B$  enthält die Punkte im Kreis mit Radius 4, in Polarkoordinaten also die Punkte mit  $r \leq 4$ .

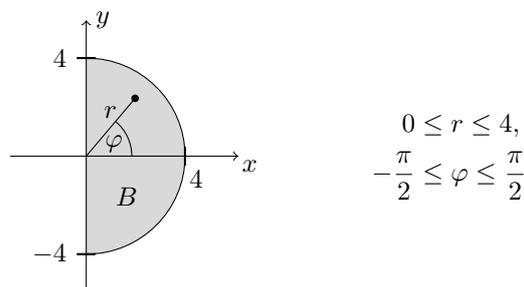


Abbildung 9.21: Halbkreis in Polarkoordinaten

Außerdem enthält  $B$  nur Punkte mit  $x \geq 0$ , in Polarkoordinaten sind das die Winkel von  $\varphi = -\frac{\pi}{2}$  (negative  $y$ -Achse) über  $\varphi = 0$  (positive  $x$ -Achse) bis  $\varphi = \frac{\pi}{2}$  (positive  $y$ -Achse), also das Intervall  $-\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2}$ . Formal ist das:

$$\begin{aligned} r^2 = x^2 + y^2 \leq 16 &\Rightarrow r \leq 4 \\ x \geq 0 &\Rightarrow -\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2} \end{aligned}$$

Damit ist dann die Menge  $B$  in Polarkoordinaten

$$B = \left\{ (r \cos \varphi, r \sin \varphi) \mid 0 \leq r \leq 4, -\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2} \right\}.$$

Alternativ können wir auch die Menge  $A$  der Paare  $(r, \varphi)$  angeben, die auch im Transformationssatz vorkommt. Damit ist

$$B = \Phi(A) \quad \text{mit} \quad A = \left\{ (r, \varphi) \mid 0 \leq r \leq 4, -\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2} \right\} = [0, 4] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Wir können nun als Beispiel noch ein Integral über  $B$  durch Transformation auf Polarkoordinaten berechnen, etwa das Integral  $\int_B x \, dF$ :

$$\begin{aligned} \int_B x \, dF &= \int_B x \, d(x, y) \stackrel{(*)}{=} \int_A r \cos \varphi \cdot r \, d(r, \varphi) \\ &= \int_0^4 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2 \cos \varphi \, d\varphi \, dr = \int_0^4 [r^2 \sin \varphi]_{\varphi=-\frac{\pi}{2}}^{\frac{\pi}{2}} \, dr \\ &= \int_0^4 2r^2 \, dr = \left[ \frac{2}{3} r^3 \right]_0^4 = \frac{128}{3} \end{aligned}$$

In  $(*)$  haben wir  $x = r \cos \varphi$  und den Transformationssatz für Polarkoordinaten (9.24) benutzt.

**Beachte:** Der Halbkreis  $B$  im Beispiel ist, wenn man ihn als Menge  $A$  im  $r, \varphi$ -Koordinatensystem darstellt,

$$A = [0, 4] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Das ist einfach ein Rechteck: das Intervall  $[0, 4]$  in  $r$ -Richtung, und das Intervall  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  in  $\varphi$ -Richtung. Für das Integral heißt das, dass man nach Transformation auf Polarkoordinaten  $r$ - und  $\varphi$ -Integrale mit *konstanten* Integralgrenzen bekommt.

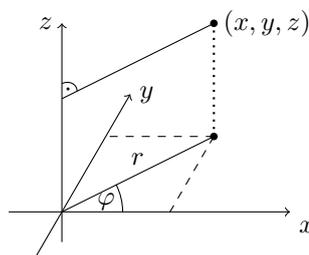


Abbildung 9.22: Zylinderkoordinaten

Im Gegensatz dazu hatten wir in Beispiel 9.5.2 das Integral über einen Kreis als Normalbereich geschrieben, was dann ein (komplizierteres) Doppelintegral mit *variablen*, von  $x$  abhängigen Grenzen liefert.

## 9.11 Zylinderkoordinaten

Wir betrachten jetzt den Transformationssatz im Dreidimensionalen. Zwei Koordinatentransformationen kommen dabei regelmäßig vor, das ersten sind *Zylinderkoordinaten*.

Zylinderkoordinaten entstehen dadurch, dass man für einen Punkt  $(x, y, z)$  im Raum für den  $x, y$ -Teil Polarkoordinaten verwendet, die  $z$ -Koordinate aber unverändert lässt, also

$$\begin{aligned}x &= r \cos \varphi \\y &= r \sin \varphi \\z &= z\end{aligned}$$

Die Transformationsfunktion  $\Phi$  ist in diesem Fall also

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \Phi(r, \varphi, z) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix}. \quad (9.25)$$

Geometrisch bedeutet das, dass der Punkt  $(x, y, z)$  senkrecht auf die  $x, y$ -Ebene projiziert wird, und dort dann die Polarkoordinaten benutzt werden. Das ist in Abbildung 9.22 dargestellt. Speziell gilt genau wie bei Polarkoordinaten

$$r^2 = x^2 + y^2.$$

In der Abbildung sieht man, dass der „Radius“  $r$  genau der senkrechte Abstand eines Punktes zur  $z$ -Achse ist.

Um den Transformationssatz in Zylinderkoordinaten hinzuschreiben, berechnen wir wieder die Determinante der Jacobimatrix von  $\Phi$ :

$$\begin{aligned}J_{\Phi}(r, \varphi, z) &= \left( \frac{\partial \Phi}{\partial r}, \frac{\partial \Phi}{\partial \varphi}, \frac{\partial \Phi}{\partial z} \right) = \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \Rightarrow \det J_{\Phi}(r, \varphi, z) &= r \cos^2 \varphi + r \sin^2 \varphi = r\end{aligned}$$

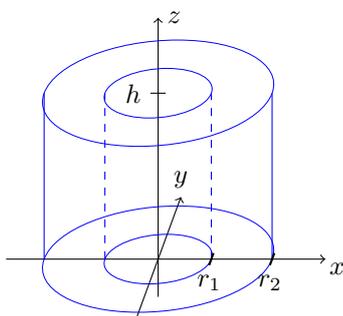


Abbildung 9.23: Kreishohlzylinder

Aus (9.21) ergibt sich damit für die Transformation auf Zylinderkoordinaten

$$\int_B f(x, y, z) dV = \int_A f(r \cos \varphi, r \sin \varphi, z) \cdot r d(r, \varphi, z) \quad (9.26)$$

mit  $B = \Phi(A)$ .

**Beispiel 9.11.1** Wir betrachten den Kreishohlzylinder mit Höhe  $h$ , innerem Radius  $r_1$  und äußerem Radius  $r_2$ , gegeben durch

$$B = \{(x, y, z) \mid r_1^2 \leq x^2 + y^2 \leq r_2^2, 0 \leq z \leq h\}.$$

Die erste Ungleichung  $r_1^2 \leq x^2 + y^2 \leq r_2^2$  bedeutet, dass der Radius  $r$  (in Zylinderkoordinaten) eines Punktes in  $B$  zwischen  $r_1$  und  $r_2$  liegen muss, denn  $x^2 + y^2 = r^2$ . Die zweite Ungleichung legt die Höhe  $z$  des Punktes zwischen 0 und  $h$  fest. Es ergibt sich damit der Hohlzylinder wie in Abbildung 9.23 skizziert. In Zylinderkoordinaten können wir den Hohlzylinder wie folgt beschreiben:

$$B = \{(r \cos \varphi, r \sin \varphi, z) \mid r_1 \leq r \leq r_2, 0 \leq \varphi \leq 2\pi, 0 \leq z \leq h\}.$$

Die Ungleichungen  $r_1 \leq r \leq r_2$  und  $0 \leq z \leq h$  haben wir schon gesehen, und für den Winkel gibt es keine Einschränkung, also  $0 \leq \varphi \leq 2\pi$ . (Genauso hätten wir auch  $-\pi \leq \varphi \leq \pi$  nehmen können.) Die entsprechende Menge  $A$  (der Punkte  $(r, \varphi, z)$ ) ist

$$A = [r_1, r_2] \times [0, 2\pi] \times [0, h] \quad (B = \Phi(A))$$

Das Integral über  $B$  in Zylinderkoordinaten ist damit

$$\int_B f(x, y, z) dV = \int_{r_1}^{r_2} \int_0^{2\pi} \int_0^h f(r \cos \varphi, r \sin \varphi, z) \cdot r dz d\varphi dr$$

(Ein konkretes Integral rechnen wir hier nicht aus.)

## 9.12 Rotationskörper

Ein Anwendung von Zylinderkoordinaten sind *Rotationskörper*. Dabei wird eine gegebene Kurve

$$r = g(z), \quad a \leq z \leq b,$$

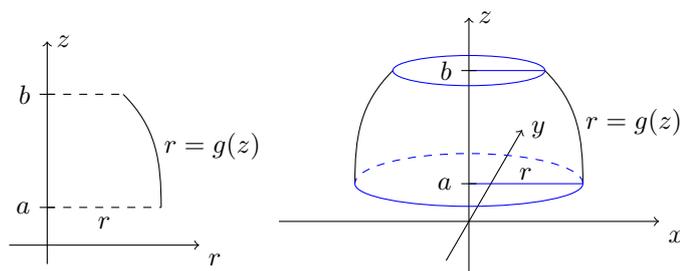


Abbildung 9.24: Konstruktion eines Rotationskörpers

um die  $z$ -Achse rotiert, siehe Abbildung 9.24. Die Kurve  $r = g(z)$  beschreibt für jede Höhe  $z$  (im Intervall  $[a, b]$ ) den Abstand  $r = \sqrt{x^2 + y^2}$  des Punktes von der  $z$ -Achse.

Für die Punkte innerhalb des Rotationskörpers gilt

$$r = \sqrt{x^2 + y^2} \leq g(z),$$

so dass der Körper als Menge durch

$$B = \left\{ (x, y, z) \mid a \leq z \leq b, \sqrt{x^2 + y^2} \leq g(z) \right\}$$

gegeben ist. In Zylinderkoordinaten ist das

$$B = \Phi(A) \quad \text{mit} \quad A = \{(r, \varphi, z) \mid a \leq z \leq b, 0 \leq \varphi \leq 2\pi, 0 \leq r \leq g(z)\},$$

womit wir jetzt das Integral über  $B$  in Zylinderkoordinaten hinschreiben können. Zu beachten ist dabei, dass die Menge  $A$  ein Normalbereich ist, denn die  $r$ -Grenze ist von  $z$  abhängig. Daher muss auch im Integral die Integration nach  $r$  vor der nach  $z$  (also innerhalb) erfolgen.

$$\int_B f(x, y, z) dV = \int_a^b \int_0^{2\pi} \int_0^{g(z)} f(r \cos \varphi, r \sin \varphi, z) \cdot r dr d\varphi dz. \quad (9.27)$$

Wir können hieraus z.B. eine Formel für das Volumen von  $B$  herleiten:

$$\begin{aligned} V &= \int_B 1 dV = \int_a^b \int_0^{2\pi} \int_0^{g(z)} r dr d\varphi dz = \int_a^b \int_0^{2\pi} \left[ \frac{1}{2} r^2 \right]_0^{g(z)} d\varphi dz \\ &= \int_a^b \int_0^{2\pi} \frac{1}{2} g(z)^2 d\varphi dz = \int_a^b 2\pi \cdot \frac{1}{2} g(z)^2 dz = \pi \int_a^b g(z)^2 dz \end{aligned}$$

Das Volumen des Rotationskörpers  $B$  ist also

$$V = \pi \int_a^b g(z)^2 dz. \quad (9.28)$$

### 9.13 Kugelkoordinaten

*Kugelkoordinaten* sind, wie auch Zylinderkoordinaten, dreidimensionale Koordinaten, also Koordinaten im Raum. Sie sind, wie der Name schon sagt, sehr

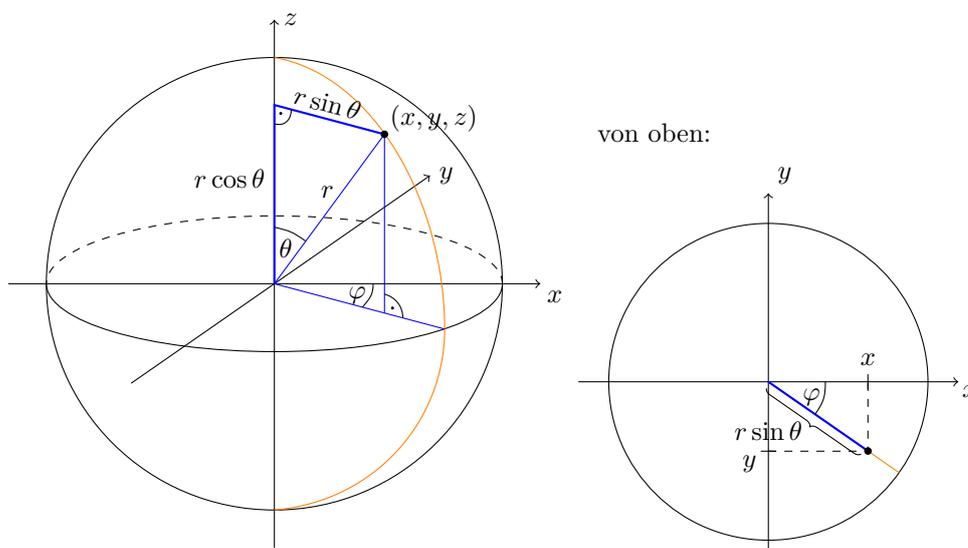


Abbildung 9.25: Kugelkoordinaten

gut zur Beschreibung von Kugeln oder Teilen von Kugeln geeignet. Man nennt sie auch *sphärische Koordinaten*.

Kugelkoordinaten sind durch

$$\begin{aligned} x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta \end{aligned} \quad (9.29)$$

definiert. Dabei ist  $r$  der Abstand des Punktes  $(x, y, z)$  zum Koordinatenursprung. Das folgt aus (9.29) durch die Rechnung

$$\begin{aligned} x^2 + y^2 + z^2 &= r^2 \sin^2 \theta \cos^2 \varphi + r^2 \sin^2 \theta \sin^2 \varphi + r^2 \cos^2 \theta \\ &= r^2 \sin^2 \theta \cdot (\cos^2 \varphi + \sin^2 \varphi) + r^2 \cos^2 \theta \\ &= r^2 \sin^2 \theta + r^2 \cos^2 \theta = r^2. \end{aligned}$$

Es gilt also

$$r = \sqrt{x^2 + y^2 + z^2} = \left| \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right|,$$

vergleiche nochmal Abschnitt 8.1, speziell Gleichung (8.1). Die Bedeutung der Winkel  $\theta$  und  $\varphi$  ist in Abbildung 9.25 illustriert.  $\theta$  ist der Winkel zwischen positiver  $z$ -Achse und dem Vektor vom Ursprung zum Punkt  $(x, y, z)$ , dem „Radiusvektor“. Mit der Verbindung von  $(x, y, z)$  senkrecht zur  $z$ -Achse ergibt sich ein rechtwinkliges Dreieck. Die Ankathete dieses Dreiecks zu  $\theta$  ist  $r \cos \theta$  und ergibt die  $z$ -Koordinate des Punkts und somit die dritte Gleichung in (9.29). Die Gegenkathete ist  $r \sin \theta$  und gleich dem senkrechten Abstand des Punkts zur  $z$ -Achse. Der Winkel  $\varphi$  ist der Winkel, mit dem das rechtwinklige Dreieck gegen die

positive  $x$ -Achse gedreht ist. Anders gesagt, ist  $\varphi$  einfach der Polarkoordinatenwinkel der Projektion des Punkts auf die  $xy$ -Ebene, oder, der Winkel zwischen positiver  $x$ -Achse und Radiusvektor „von oben“ gesehen. Da der Abstand des Punkts in der  $xy$ -Ebene zum Ursprung genau  $r \sin \theta$  ist (die Gegenkathete!), ergeben sich die ersten beiden Gleichungen in (9.29) dann durch Anwendung von Polarkoordinaten.

Ein Punkt mit den Kugelkoordinaten  $r, \theta, \varphi$  hat also den Abstand  $r$  zum Ursprung, d.h. er liegt auf der Kugeloberfläche mit Radius  $r$ . Der Winkel  $\theta$  ist der Winkelabstand des Punkts vom „Nordpol“ der Kugel. Speziell bedeutet ...

- $\theta = 0$ : der Punkt  $(x, y, z)$  ist der Nordpol der Kugel  $(0, 0, r)$
- $\theta = \frac{\pi}{2}$ : Punkt auf Äquator der Kugel (in  $xy$ -Ebene)  $(r \cos \varphi, r \sin \varphi, 0)$
- $\theta = \pi$ : der Punkt ist der Südpol  $(0, 0, -r)$
- $0 \leq \theta < \frac{\pi}{2}$ : Punkt auf der Nordhalbkugel ( $z > 0$ )
- $\frac{\pi}{2} < \theta \leq \pi$ : Punkt auf der Südhalbkugel ( $z < 0$ )

Schließlich ist  $\varphi$  der Polarkoordinatenwinkel des Punktes, gemessen in der  $xy$ -Ebene.

Die Transformationsfunktion  $\Phi$  für Kugelkoordinaten ist

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \Phi(r, \theta, \varphi) = \begin{pmatrix} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{pmatrix} \quad (9.30)$$

wobei

$$r > 0, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \varphi \leq 2\pi \quad (\text{oder } -\pi \leq \varphi \leq \pi).$$

Um den Transformationssatz für Kugelkoordinaten hinzuschreiben, berechnen wir wieder die Determinante der Jacobimatrix von  $\Phi$ . Es gilt

$$J_{\Phi}(r, \theta, \varphi) = \begin{pmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}$$

und damit

$$\begin{aligned} \det J_{\Phi}(r, \theta, \varphi) &= r^2 \cos^2 \theta \sin \theta \cos^2 \varphi + r^2 \sin^3 \theta \sin^2 \varphi \\ &\quad + r^2 \cos^2 \theta \sin \theta \sin^2 \varphi + r^2 \sin^3 \theta \cos^2 \varphi \\ &= r^2 \cos^2 \theta \sin \theta + r^2 \sin^3 \theta = r^2 \sin \theta (\cos^2 \theta + \sin^2 \theta) \\ &= r^2 \sin \theta \end{aligned}$$

Aus Satz 9.9.1 erhalten wir damit den *Transformationssatz für Kugelkoordinaten*:

$$\int_B f(x, y, z) dV = \int_A f(\Phi(r, \theta, \varphi)) \cdot r^2 \sin \theta d(r, \theta, \varphi). \quad (9.31)$$

(Man beachte, dass  $\sin \theta \geq 0$  wegen  $0 \leq \theta \leq \pi$  und somit  $|\det J_{\Phi}| = r^2 \sin \theta$ .)

**Beachte:** In *Kugelkoordinaten* ist  $r$  der Abstand des Punktes  $(x, y, z)$  vom Ursprung des Koordinatensystems, und es gilt

$$x^2 + y^2 + z^2 = r^2.$$

Bei *Zylinderkoordinaten* dagegen ist  $r$  der senkrechte Abstand vom Punkt zur  $z$ -Achse und

$$x^2 + y^2 = r^2.$$

Der senkrechte Abstand zur  $z$ -Achse in Kugelkoordinaten ist dagegen  $r \sin \theta$ .

**Beispiel 9.13.1** Wir betrachten die Kugel mit Radius  $R$

$$B = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq R^2\}.$$

Mit Kugelkoordinaten ist

$$r^2 = x^2 + y^2 + z^2 \leq R^2 \quad \Leftrightarrow \quad r \leq R.$$

Damit kann man die Kugel  $B$  in Kugelkoordinaten durch

$$\begin{aligned} B &= \Phi(A), \quad A = \{(r, \theta, \varphi) \mid 0 \leq r \leq R, 0 \leq \theta \leq \pi, 0 \leq \varphi \leq 2\pi\} \\ &= [0, R] \times [0, \pi] \times [0, 2\pi] \end{aligned}$$

beschreiben. Für das Integral über die Kugel folgt damit

$$\begin{aligned} \int_B f(x, y, z) dV &= \int_0^R \int_0^\pi \int_0^{2\pi} f(\Phi(r, \theta, \varphi)) \cdot r^2 \sin \theta d\varphi d\theta dr \\ &= \int_0^R \int_0^\pi \int_0^{2\pi} f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) \cdot r^2 \sin \theta d\varphi d\theta dr \end{aligned}$$

Als konkretes Beispiel berechnen wir damit das Volumen der Kugel:

$$\begin{aligned} V &= \int_B 1 dV = \int_0^R \int_0^\pi \int_0^{2\pi} r^2 \sin \theta d\varphi d\theta dr = 2\pi \int_0^R \int_0^\pi r^2 \sin \theta d\theta dr \\ &= 2\pi \int_0^R [-r^2 \cos \theta]_{\theta=0}^\pi dr = 2\pi \int_0^R r^2 (-\cos(\pi) + \cos(0)) dr \\ &= 2\pi \int_0^R r^2(1 + 1) dr = 4\pi \int_0^R r^2 dr = 4\pi \left[ \frac{1}{3} r^3 \right]_0^R = \frac{4}{3} \pi R^3 \end{aligned}$$

Das ist die bekannte Formel für das Volumen der Kugel mit Radius  $R$ .

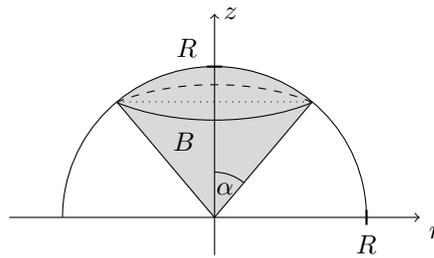
**Beispiel 9.13.2** Durch Einschränken des Winkelbereichs für  $\theta$  kann man verschiedene Teile der Kugel beschreiben.

(a) Die obere Halbkugel ist gegeben durch

$$B = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq R^2, z \geq 0\}.$$

In Kugelkoordinaten ist das

$$\begin{aligned} B &= \Phi(A), \quad A = \{(r, \theta, \varphi) \mid 0 \leq r \leq R, 0 \leq \theta \leq \frac{\pi}{2}, 0 \leq \varphi \leq 2\pi\} \\ &= [0, R] \times [0, \frac{\pi}{2}] \times [0, 2\pi] \end{aligned}$$

Abbildung 9.26: Kugelsektor  $B$ 

- (b) Jetzt betrachten wir als Menge  $B$  den Kugelsektor mit Radius  $R$  und (halbem) Öffnungswinkel  $\alpha$  um die  $z$ -Achse, siehe Abbildung 9.26. In Kugelkoordinaten können wir  $B$  durch

$$0 \leq r \leq R, \quad 0 \leq \theta \leq \alpha, \quad 0 \leq \varphi \leq 2\pi,$$

oder

$$B = \Phi(A), \quad A = [0, R] \times [0, \alpha] \times [0, 2\pi]$$

beschreiben. Damit ergibt sich z.B. für das Volumen des Kugelsektors

$$\begin{aligned} V &= \int_B 1 \, dV = \int_0^R \int_0^\alpha \int_0^{2\pi} r^2 \sin \theta \, d\varphi \, d\theta \, dr \\ &= 2\pi \int_0^R r^2 \int_0^\alpha \sin \theta \, d\theta \, dr \\ &= 2\pi \int_0^R r^2 \, dr \int_0^\alpha \sin \theta \, d\theta = 2\pi \cdot \frac{1}{3} R^3 \cdot [-\cos \theta]_0^\alpha \\ &= \frac{2}{3} \pi R^3 (1 - \cos \alpha) \end{aligned}$$