

# Analysis

**Stefan Waldmann**

Julius Maximilian University of Würzburg  
Institute of Mathematics  
Chair of Mathematics X (Mathematical Physics)  
Emil-Fischer-Straße 31  
97074 Würzburg  
Germany

`stefan.waldmann@mathematik.uni-wuerzburg.de`

2023-12-07 17:24:31 +0100

Last changes by Stefan@darkstar on 2023-12-07  
Git revision of ana: 8bb0ab1 (HEAD -> master)



# Inhaltsverzeichnis

|   |            |
|---|------------|
| <b>Einleitung</b>   | <b>1</b>   |
| <b>1 Die reellen und komplexen Zahlen</b>                       | <b>3</b>   |
| 1.1 Der Startpunkt: Von Zahlen und Ziffern . . . . .            | 3          |
| 1.2 Körper . . . . .  | 4          |
| 1.3 Angeordnete Körper . . . . .                                | 12         |
| 1.4 Archimedische Körper . . . . .                              | 15         |
| 1.5 Vollständigkeit und die reellen Zahlen . . . . .            | 17         |
| 1.6 Wurzeln und Potenzgesetze . . . . .                         | 27         |
| 1.7 Die komplexen Zahlen . . . . .                              | 33         |
| 1.8 Übungen . . . . .   | 37         |
| <b>2 Folgen und Reihen</b>                                      | <b>41</b>  |
| 2.1 Folgen und Abzählbarkeit . . . . .                          | 43         |
| 2.2 Konvergenz von Folgen . . . . .                             | 46         |
| 2.3 Rechenregeln für konvergente Folgen . . . . .               | 51         |
| 2.4 Cauchy-Folgen und Vollständigkeit . . . . .                 | 59         |
| 2.5 Häufungspunkte und Umgebungen . . . . .                     | 64         |
| 2.6 Unendliche Reihen . . . . .                                 | 73         |
| 2.7 Absolute Konvergenz von Reihen . . . . .                    | 79         |
| 2.8 Netze . . . . .   | 88         |
| 2.9 Übungen . . . . .   | 93         |
| <b>3 Stetige Funktionen</b>                                     | <b>95</b>  |
| 3.1 Funktionenlimes . . . . .                                   | 95         |
| 3.2 Stetigkeit . . . . .  | 104        |
| 3.3 Rechenregeln und Beispiele für stetige Funktionen . . . . . | 109        |
| 3.4 Kompaktheit und Zusammenhang . . . . .                      | 112        |
| 3.5 Hauptsätze für stetige Funktionen . . . . .                 | 121        |
| 3.6 Gleichmäßige Stetigkeit . . . . .                           | 132        |
| 3.7 Gleichmäßige Konvergenz . . . . .                           | 136        |
| 3.8 Konvergenz von Potenzreihen . . . . .                       | 143        |
| 3.9 Übungen . . . . .   | 147        |
| <b>4 e und <math>\pi</math></b>                                 | <b>153</b> |
| 4.1 e . . . . .   | 153        |
| 4.2 Der natürliche Logarithmus . . . . .                        | 157        |
| 4.3 Die trigonometrischen Funktionen . . . . .                  | 163        |
| 4.4 $\pi$ . . . . .   | 167        |
| 4.5 Die Arcusfunktionen und Polarkoordinaten . . . . .          | 172        |

|          |   |            |
|----------|---|------------|
| 4.6      | Übungen . . . . .   | 176        |
| <b>5</b> | <b>Differenzierbarkeit und die Ableitung</b>                      | <b>177</b> |
| 5.1      | Die Definition der Ableitung . . . . .                            | 177        |
| 5.2      | Rechenregeln für differenzierbare Funktionen . . . . .            | 184        |
| 5.3      | Hauptsätze für differenzierbare Funktionen . . . . .              | 189        |
| 5.4      | Höhere Ableitungen und $\mathcal{C}^k$ -Funktionen . . . . .      | 199        |
| 5.5      | Konvergenz von differenzierbaren Funktionen . . . . .             | 206        |
| 5.6      | Potenzreihen und Taylor-Entwicklung . . . . .                     | 213        |
| 5.7      | Asymptotische Entwicklungen . . . . .                             | 224        |
| 5.8      | Übungen . . . . .   | 229        |
| <b>6</b> | <b>Das Riemann-Integral</b>                                       | <b>231</b> |
| 6.1      | Definition des Riemann-Integrals . . . . .                        | 231        |
| 6.2      | Riemann-integrierbare Funktionen . . . . .                        | 235        |
| 6.3      | Riemann-Integral stetiger Funktionen . . . . .                    | 246        |
| 6.4      | Der Hauptsatz und Stammfunktionen . . . . .                       | 249        |
| 6.5      | Ungleichungen und Mittelwertsätze für Riemann-Integrale . . . . . | 257        |
| 6.6      | Uneigentliche Riemann-Integrale . . . . .                         | 261        |
| 6.7      | Integration und gleichmäßige Konvergenz . . . . .                 | 270        |
| 6.8      | Übungen . . . . .   | 274        |
| <b>7</b> | <b>Normierte, metrische und topologische Räume</b>                | <b>275</b> |
| 7.1      | Normierte Vektorräume und metrische Räume . . . . .               | 275        |
| 7.2      | Konvergenz in metrischen Räumen . . . . .                         | 279        |
| 7.3      | Stetigkeit und Lipschitz-Stetigkeit . . . . .                     | 283        |
| 7.4      | Vollständigkeit und der Banachsche Fixpunktsatz . . . . .         | 290        |
| 7.5      | Topologische Räume und Stetigkeit . . . . .                       | 293        |
| 7.6      | Kompaktheit und Zusammenhang in topologischen Räumen . . . . .    | 300        |
| 7.7      | Die Standardtopologie auf $\mathbb{R}^n$ . . . . .                | 304        |
| 7.8      | Übungen . . . . .   | 311        |
| <b>8</b> | <b>Differenzierbarkeit in <math>\mathbb{R}^n</math></b>           | <b>315</b> |
| 8.1      | Totale Differenzierbarkeit . . . . .                              | 316        |
| 8.2      | Richtungsableitungen und partielle Ableitungen . . . . .          | 323        |
| 8.3      | Stetig differenzierbare Funktionen . . . . .                      | 328        |
| 8.4      | $\mathcal{C}^k$ -Funktionen . . . . .                             | 335        |
| 8.5      | Funktionenfolgen und $\mathcal{C}^k$ -Konvergenz . . . . .        | 342        |
| 8.6      | Übungen . . . . .   | 350        |
| <b>9</b> | <b>Implizite Funktionen und Untermannigfaltigkeiten</b>           | <b>351</b> |
| 9.1      | Lokale Umkehrbarkeit . . . . .                                    | 351        |
| 9.2      | Implizite Funktionen . . . . .                                    | 352        |
| 9.3      | Submersionen und Immersionen . . . . .                            | 352        |
| 9.4      | Übungen . . . . .   | 352        |
|          | <b>Literaturverzeichnis</b>                                       | <b>353</b> |
|          | <b>Symbolverzeichnis</b>  | <b>355</b> |
|          | <b>Index</b>  | <b>357</b> |

# Einleitung



# Kapitel 1

## Die reellen und komplexen Zahlen

In der Analysis will man oftmals mathematische Probleme und Aufgabenstellungen zunächst näherungsweise also *approximativ* lösen. Das hat verschiedenste Gründe: So kann eine approximative Lösung, beispielsweise einer Gleichung, viel einfacher zu finden und für die Praxis aber völlig ausreichend sein. Zum anderen kann man bestimmte Probleme vielleicht nicht explizit lösen und ist daher auf Approximationen angewiesen. Dies ist zumindest dann sinnvoll, wenn man sich zuvor davon überzeugt hat, dass das Problem Lösungen besitzt und diese vielleicht aus allgemeinen Überlegungen heraus eindeutig sind.

Wichtig bei Approximationen ist dann, dass man eine gute Vorstellung davon entwickelt, wie gut die Approximation ist. Dies erfordert typischerweise den Gebrauch von *Ungleichungen*. Idealerweise findet man nicht nur eine Approximation mit einer Abschätzung ihrer Güte, sondern zugleich ein Verfahren, die Approximation zu verbessern, also neue Approximationen mit besserer Güte zu erlangen.

Kann man Approximationen sukzessive verbessern, so erhält man eine Annäherung an die tatsächliche Lösung des Problems, die dann in einem geeigneten Sinne konvergiert. Die Aufgabe der Analysis ist es, genau zu klären, was unter *Konvergenz* zu verstehen ist und zugehörige Rechenregeln und Techniken bereitzustellen, um das Konvergenzverhalten approximativer Lösungen zu kontrollieren.

In diesem einführenden Kapitel werden wir sehen, dass diese Fragestellungen bereits für die Zahlen selbst auftreten. Später werden wir nicht einzelne Zahlen sondern ganze Abbildungen von Zahlen zum Gegenstand analytischer Überlegungen machen. Damit ist aber das Ende noch lange nicht erreicht. Solche Funktionen können selbst wieder Argumente von Abbildungen sein, jetzt also Abbildungen von Funktionen zu anderen Funktionen und so weiter.

### 1.1 Der Startpunkt: Von Zahlen und Ziffern

Zunächst gilt es aber, die Natur der Zahlen selbst genauer aufzuklären. Da es sich hier offenbar um einen der grundlegendsten Begriffe der Mathematik handelt, ist zum einen besondere Sorgfalt nötig. Zum anderen ist die Frage nach der tatsächlichen *Grundlegung der Mathematik* zumindest unterschwellig gestellt. Hier gibt es verschiedene Möglichkeiten:

- i.) Wir können beispielsweise die aus der Schule bekannten *natürlichen Zahlen*  $\mathbb{N}$  als Ausgangspunkt aller Überlegungen machen und ihre Eigenschaften, beispielsweise in Form der Peano-Axiome, axiomatisch fordern. Dann müssen die *ganzen Zahlen*  $\mathbb{Z}$ , die *rationale Zahlen*  $\mathbb{Q}$  sowie die *reellen Zahlen*  $\mathbb{R}$  hieraus konstruiert werden.
- ii.) Wir können ebenso gut die reellen Zahlen selbst als Ausgangspunkt wählen und axiomatisch verwenden. Deren Existenz wird also nicht weiter hinterfragt, sondern als gegeben angenommen. Die Analysis wird dann gänzlich aus den Eigenschaften der reellen Zahlen entwickelt.

iii.) Der vermutlich aktuell am häufigsten verwendete Rahmen der Grundlegung der Mathematik ist die Mengenlehre. Hier werden Mengen und ihre Eigenschaften axiomatisch behandelt und alle weitere Mathematik daraus entwickelt. Je nach dem welches Axiomensystem der Mengenlehre man genau verwendet, ergeben sich hier verschiedene Spielarten. Das meist verwendete System ist sicherlich die Mengenlehre nach Zermelo & Fraenkel mit dem Auswahlaxiom (ZFC). Wir werden gelegentlich auf Aspekte der Mengenlehre zurückkommen aber keine systematische Diskussion anstreben, hierfür sei auf die Literatur wie etwa [1] oder [8, Anhang B] für eine kurze Übersicht verwiesen.

Wir werden den in gewisser Hinsicht einfachsten Einstieg wählen und die reellen Zahlen axiomatisch behandeln. Im Detail bedeutet dies, dass wir gewisse wünschenswerte Eigenschaften der reellen Zahlen herausarbeiten werden und dann fordern, dass es die reellen Zahlen mit genau diesen Eigenschaften gibt. Hier gilt es natürlich, ein möglichst einfaches und minimales System von Anforderungen als Ausgangspunkt zu nehmen und weitere Eigenschaften der reellen Zahlen anschließend aus diesen herzuleiten. Es bleibt dann späteren Untersuchungen vorbehalten, ihre tatsächliche Konstruktion aus den Peano-Axiomen oder der axiomatischen Mengenlehre zu gewinnen.

Aus der Schule sind die reellen Zahlen zunächst als heuristisches Konzept bekannt. Wir wollen an dieser Stelle nun auf einige Schwierigkeiten bei den üblichen Definitionen hinweisen. Eine gängige Herangehensweise ist, eine reelle Zahl als unendliche Dezimalzahl zu charakterisieren, also beispielsweise

$$\sqrt{2} = 2,414213562 \dots \quad (1.1.1)$$

oder

$$\pi = 3,141592654 \dots \quad (1.1.2)$$

Diese Schreibweise soll dann andeuten, dass die aufgelisteten Stellen nur eine Approximation an die tatsächliche reelle Zahl  $\sqrt{2}$  oder  $\pi$  darstellen. Mehr Stellen geben also eine bessere Approximation, aber den genauen Wert  $\sqrt{2}$  oder  $\pi$  wird man so nicht erhalten. Dieser erfordert ja die Angabe *aller* unendlich vielen Stellen. Während man mit dieser approximativen Natur in der Praxis gut umgehen kann, stellen sich doch einige konzeptuelle Fragen. Wie sollen beispielsweise zwei reelle Zahlen addiert oder multipliziert werden? Der übliche Algorithmus für die Addition oder Multiplikation von Zahlen mit endlich vielen nicht-trivialen Dezimalstellen sieht ja vor, *ganz rechts* zu beginnen, um die eventuell auftretenden Überträge richtig zu berücksichtigen. Diese Vorgehensweise scheidet nun bei  $\sqrt{2}$  oder  $\pi$  in der Darstellung (1.1.1) oder (1.1.2) offensichtlich aus. Die vermeintlich harmlose „Definition“ reeller Zahlen stellt sich also als tiefgründiger und deutlich problematischer heraus.

Man könnte nun tatsächlich eine Definition der reellen Zahlen geben, die auf der Dezimaldarstellung basiert *und* den Ansprüchen einer wissenschaftlichen Mathematik genügt. Ein vielleicht eher ästhetischer Gesichtspunkt, der gegen diesen Zugang spricht, ist die Auszeichnung des *dekadischen* Zahlensystems basierend auf den Ziffern 0, 1, ..., 9 vor anderen Möglichkeiten. Diese Auszeichnung hat kulturelle und historische Gründe aber sicherlich keine inner-mathematischen. Im Futurama-Universum haben die Figuren beispielsweise vier Finger an jeder Hand. Dort hätte sich vermutlich ein auf acht Ziffern 0, 1, ..., 7 basierendes Zahlssystem etabliert. Wir werden daher ein anderes Axiomensystem für die reellen Zahlen wählen und eine dekadische Darstellung später als Konsequenz daraus ableiten.

## 1.2 Körper

Um die Eigenschaften der reellen Zahlen charakterisieren zu können, benötigen wir zunächst die Rechenoperationen, die wir für Zahlen erwarten. Wir möchten zwei Verknüpfungen, *Addition* und *Multiplikation*, die dann gewissen Regeln genügen sollen. Weitere Operationen wie *Subtraktion* und



*Division* ergeben sich dann aus diesen. Die bekannten und daher gewünschten Regeln führen nun zur ersten Definition eines *Körpers*:

**Definition 1.2.1 (Körper)** Sei  $\mathbb{k}$  eine Menge mit zwei Elementen  $0, 1 \in \mathbb{k}$  und zwei Verknüpfungen

$$+, \cdot: \mathbb{k} \times \mathbb{k} \longrightarrow \mathbb{k}, \quad (1.2.1)$$

genannt *Addition und Multiplikation*. Dann heißt  $(\mathbb{k}, +, \cdot, 0, 1)$  ein *Körper*, falls folgende Eigenschaften erfüllt sind:

i.) Die Addition  $+$  ist assoziativ und kommutativ. Es gilt also

$$x + (y + z) = (x + y) + z \quad \text{und} \quad x + y = y + x \quad (1.2.2)$$

für alle  $x, y, z \in \mathbb{k}$ . Hier schreiben wir  $x + y$  für das Bild von  $(x, y)$  unter der Addition.

ii.) Das Element  $0$  ist neutral bezüglich  $+$ . Es gilt also

$$x + 0 = x \quad (1.2.3)$$

für alle  $x \in \mathbb{k}$ .

iii.) Zu jedem  $x \in \mathbb{k}$  gibt es ein Inverses bezüglich  $+$ , bezeichnet mit  $-x$ , also

$$x + (-x) = 0. \quad (1.2.4)$$

iv.) Die Multiplikation  $\cdot$  ist assoziativ und kommutativ. Es gilt also

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z \quad \text{und} \quad x \cdot y = y \cdot x \quad (1.2.5)$$

für alle  $x, y, z \in \mathbb{k}$ . Wie zuvor bei der Addition schreiben wir  $x \cdot y$  für das Bild von  $(x, y)$  unter der Multiplikation.

v.) Das Element  $1 \in \mathbb{k} \setminus \{0\}$  ist neutral bezüglich  $\cdot$ , also

$$1 \cdot x = x \quad (1.2.6)$$

für alle  $x \in \mathbb{k}$ .

vi.) Zu jedem  $x \in \mathbb{k} \setminus \{0\}$  gibt es ein Inverses bezüglich  $\cdot$ , bezeichnet mit  $x^{-1}$  oder  $\frac{1}{x}$ , also

$$x \cdot \frac{1}{x} = 1. \quad (1.2.7)$$

vii.) Es gilt das *Distributivgesetz*

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z) \quad (1.2.8)$$

zwischen  $+$  und  $\cdot$  für alle  $x, y, z \in \mathbb{k}$ .

Die obigen Eigenschaften sind aus der Schule für beispielsweise die rationalen Zahlen  $\mathbb{Q}$  bekannt. Allgemein wollen wir Elemente aus einem Körper  $\mathbb{k}$  auch *Zahlen* nennen. Im Folgenden nehmen wir an, einen Körper  $\mathbb{k}$  vorliegen zu haben. Dann sind wir daran interessiert, aus den Eigenschaften eines Körpers allgemeine Folgerungen zu ziehen: Diese gelten dann in *jedem* Körper. Später werden wir weitere Eigenschaften eines Körpers hinzunehmen, die dann weitere nützliche Konsequenzen haben werden.

Um nun allen Missverständnissen hinsichtlich der Allgemeinheit der Definition vorzubeugen, geben wir ein erstes, noch sehr obskures Beispiel eines Körpers.

**Beispiel 1.2.2 (Der Körper  $\mathbb{Z}_2$ )** Wir betrachten die Menge  $\mathbb{Z}_2 = \{0, 1\}$  bestehend aus zwei Elementen 0 und 1. Auf dieser definieren wir die Verknüpfungen  $+$  und  $\cdot$  durch

$$0 + 0 = 0, \quad 0 + 1 = 1 = 1 + 0, \quad 1 + 1 = 0 \quad (1.2.9)$$

sowie

$$0 \cdot 0 = 0, \quad 0 \cdot 1 = 0 = 1 \cdot 0, \quad 1 \cdot 1 = 1. \quad (1.2.10)$$

Durch eine einfache Fallunterscheidung zeigt man dann, dass  $\mathbb{Z}_2$  tatsächlich den Anforderungen aus Definition 1.2.1 genügt, also ein Körper ist. Dieser spielt eine große Rolle in verschiedenen anderen Bereichen der Mathematik, nicht aber so sehr für die Analysis.

Wir wollen einige neue Begriffe zu algebraischen Strukturen einführen, die der Definition eines Körpers zugrunde liegen. Da eine systematische Diskussion von Gruppen und Ringen der (linearen) Algebra vorbehalten ist, beschränken wir uns auf ein paar wenige Bemerkungen und verweisen auf die Literatur, wie beispielsweise [8, Kapitel 3] oder [3, Chap. I und Chap. II], für eine tiefergehende Ausarbeitung.

### Bemerkung 1.2.3 (Körper)

- i.) Betrachtet man nur die Addition  $+$  sowie das Nullelement  $0 \in \mathbb{k}$ , so erhält man eine *kommutative* oder *abelsche Gruppe*  $(\mathbb{k}, +, 0)$ . Verzichtet man auf die Forderung der Kommutativität  $x + y = y + x$ , so erhält man die Definition einer Gruppe im allgemeinen. Gruppen gehören zu den wichtigsten mathematischen Strukturen und treten sowohl in ihrer kommutativen Version, wie beispielsweise hier, aber auch in allgemeiner nicht-kommutativer Variante auf. In einer Gruppe hat die Gleichung  $a + x = b$  immer eine eindeutige Lösung, nämlich  $x = b + (-a)$ , was wir auch als  $x = b - a$  abkürzen. Ebenso wie  $(\mathbb{k}, +, 0)$  ist auch  $(\mathbb{k} \setminus \{0\}, \cdot, 1)$  eine abelsche Gruppe. Man schreibt auch

$$\mathbb{k}^\times = \mathbb{k} \setminus \{0\} \quad (1.2.11)$$

für die Körperelemente ungleich Null. Wieder hat die Gleichung  $a \cdot x = b$  für  $a, b \in \mathbb{k}^\times$  die eindeutige Lösung  $x = a^{-1} \cdot b$ . Man beachte, dass

$$1 \neq 0 \quad (1.2.12)$$

Teil der Definition in einem Körper ist. Damit ist also 1 multiplikativ invertierbar, und es gilt

$$1 \cdot 1 = 1. \quad (1.2.13)$$

- ii.) Weiter ist das additive Inverse  $-x$  zu  $x$  eindeutig bestimmt. Ist nämlich  $x + (-x) = 0$  sowie  $x + a = 0$ , so folgt

$$a = a + 0 = a + (x + (-x)) = (a + x) + (-x) = 0 + (-x) = -x \quad (1.2.14)$$

unter mehrfacher Verwendung der Assoziativität der Addition. Ebenso ist das multiplikative Inverse  $\frac{1}{x}$  eines Elements  $x \in \mathbb{k}^\times$  eindeutig.

- iii.) Verzichtet man auf die Existenz eines multiplikativen Inversen zu  $x \in \mathbb{k} \setminus \{0\}$  wie in vi.), so erhält man die Definition eines *kommutativen Rings* mit Eins  $1 \neq 0$ . Ein Körper ist daher der Spezialfall eines kommutativen Rings mit Eins  $1 \neq 0$ , in dem alle Elemente ungleich Null ein multiplikatives Inverses besitzen. Wie schon bei Gruppen ist es bei Ringen durchaus von Interesse, auch eine nicht-kommutative Multiplikation zuzulassen. Die Addition bleibt allerdings in der Welt der Ringe definitionsgemäß immer kommutativ. Für einen nicht-kommutativen Ring muss man zusätzlich die Distributivität von rechts ebenso fordern, also zusätzlich zu (1.2.8) auch

$$(x + y) \cdot z = (x \cdot z) + (y \cdot z) \quad (1.2.15)$$

für alle  $x, y, z$ . In einem Körper ist dies automatisch erfüllt.

- iv.) Um die Anzahl von Klammern zu reduzieren, vereinbar man die Konvention „*Punkt vor Strich*“. Wir schreiben

$$x \cdot y + z = (x \cdot y) + z \quad (1.2.16)$$

etc. und setzen nur Klammern, wenn die Addition vor der Multiplikation ausgeführt werden soll. Das Distributivgesetz lautet daher mit dieser Konvention einfach

$$x \cdot (y + z) = x \cdot y + x \cdot z. \quad (1.2.17)$$

Weiter werden wir die Multiplikation abkürzend als

$$xy = x \cdot y \quad (1.2.18)$$

schreiben, wann immer wir aus dem Zusammenhang entnehmen können, um welche Multiplikationsvorschrift es sich handelt. Da in einem Körper  $\mathbb{k}$  ja nur *eine* Multiplikation erklärt ist, führt dies typischerweise nicht zu Missverständnissen.

Nach diesen einfachen Beobachtungen wollen wir nun einige weniger offensichtliche Folgerungen aus den Körpereigenschaften aufzeigen.

**Proposition 1.2.4** *Sei  $\mathbb{k}$  ein Körper.*

- i.) *Es gilt  $x \cdot 0 = 0$  für alle  $x \in \mathbb{k}$ .*
- ii.) *Das Nullelement besitzt kein multiplikatives Inverses.*
- iii.) *Für alle  $n \in \mathbb{N}$  und  $x_1, \dots, x_n \in \mathbb{k}$  ist die Mehrfachsumme  $x_1 + \dots + x_n$  unabhängig von der Art der Klammerung und der Reihenfolge der Summanden.*
- iv.) *Für alle  $n \in \mathbb{N}$  und  $x_1, \dots, x_n \in \mathbb{k}$  ist das Mehrfachprodukt  $x_1 \cdots x_n$  unabhängig von der Art der Klammerung und der Reihenfolge der Faktoren.*
- v.) *Für  $x, y \in \mathbb{k}$  gilt*

$$(-x) \cdot y = -(xy) = x \cdot (-y). \quad (1.2.19)$$

*Insbesondere gilt  $(-1) \cdot x = -x$ .*

- vi.) *Für  $x, y_1, \dots, y_n \in \mathbb{k}$  gilt das verallgemeinerte Distributivgesetz*

$$x \cdot (y_1 + \dots + y_n) = xy_1 + \dots + xy_n. \quad (1.2.20)$$

BEWEIS: Für den ersten Teil wissen wir  $0+0=0$ , da 0 neutral bezüglich  $+$  ist. Mit der Distributivität gilt dann  $x \cdot 0 = x \cdot (0+0) = x \cdot 0 + x \cdot 0$ . Dies ist nur für  $x \cdot 0 = 0$  möglich, siehe nochmals Bemerkung 1.2.3, i.). Für ii.) nimmt man an,  $x$  sei ein multiplikatives Inverses von 0, also  $x \cdot 0 = 1$ . Nach i.) kann dies nicht sein. Das Umklammern in längeren Summen beziehungsweise Produkten macht mehrfachen Gebrauch von der Assoziativität. Der detaillierte Beweis ist etwas mühsam auszuschreiben, da die Anzahl der möglichen Klammerungen schnell mit der Anzahl der Summanden beziehungsweise der Faktoren wächst. Für  $n = 4$  hat man beispielsweise

$$\begin{aligned} x_1 + (x_2 + (x_3 + x_4)) &= x_1 + ((x_2 + x_3) + x_4) \\ &= (x_1 + (x_2 + x_3)) + x_4 \\ &= ((x_1 + x_2) + x_3) + x_4 \\ &= (x_1 + x_2) + (x_3 + x_4) \end{aligned}$$

als Möglichkeiten. Man sieht, dass man in jedem Schritt die Assoziativität von  $+$  einmal verwendet hat. Die Kommutativität von  $+$  beziehungsweise  $\cdot$  zeigt dann, dass die Reihenfolge der Summanden

beziehungsweise der Faktoren ebenfalls unerheblich ist. Auch hier gibt es eine rasch wachsende Zahl von Reihenfolgen wie beispielsweise

$$\begin{aligned} x_1 \cdot x_2 \cdot x_3 &= x_1 x_3 x_2 \\ &= x_3 x_1 x_2 \\ &= x_3 x_2 x_1 \\ &= x_2 x_3 x_1 \\ &= x_2 x_1 x_3 \end{aligned}$$

für  $n = 3$ . Die genaue Kombinatorik ist eine interessante Übung, siehe auch Übung 1.8.6. Für  $v$ ) betrachten wir

$$x \cdot y + (-x) \cdot y = (x + (-x)) \cdot y = 0 \cdot y = 0,$$

womit  $(-x) \cdot y$  das eindeutige additive Inverse von  $xy$  ist, also mit  $-(xy)$  übereinstimmt. Genauso zeigt man  $x \cdot (-y) = -(xy)$ . Den letzten Teil erhält man dann wieder durch mehrfache Anwendung des Distributivgesetzes und einer einfachen Induktion nach  $n$ .  $\square$

Wir können also in Zukunft von diesen und ähnlichen weiteren Rechenregeln für Zahlen aus einem Körper Gebrauch machen, ohne explizit auf diese Proposition zu verweisen.

Etwas ungewohnter ist die folgende Problematik. Für den Körper  $\mathbb{Z}_2$  aus Beispiel 1.2.2 haben wir gesehen, dass  $1 + 1 = 0$  gilt. Allgemein können wir in einem Körper das Einselement  $n$ -mal aufsummieren. Es stellt sich daher die Frage, ob  $1 + \dots + 1$  Null werden kann oder nicht. Dies führt zu folgender Definition:

**Definition 1.2.5 (Charakteristik)** Sei  $\mathbb{k}$  ein Körper. Die kleinste natürliche Zahl  $p \in \mathbb{N}$ , für die  $1 + \dots + 1 = 0$  mit  $p$  Summanden gilt, heißt Charakteristik von  $\mathbb{k}$ . Gibt es keine solche Zahl, so sagen wir, dass  $\mathbb{k}$  Charakteristik Null besitzt. Wir schreiben auch  $\text{char}(\mathbb{k}) = p$ , wenn  $\mathbb{k}$  Charakteristik  $p \in \mathbb{N}_0$  besitzt.

Eine kleine Überlegung zeigt, dass die Charakteristik eines Körpers entweder Null oder eine Primzahl ist. Körper mit Primzahl-Charakteristik spielen in vielen Bereichen der Mathematik eine zentrale Rolle. Für uns wird jedoch der Fall von Körpern mit Charakteristik Null die entscheidende sein. Hat  $\mathbb{k}$  Charakteristik Null, so ist

$$n = 1 + \dots + 1 \in \mathbb{k} \tag{1.2.21}$$

immer von Null verschieden. Wir können daher die natürliche Zahl  $n \in \mathbb{N}$  mit dieser Summe identifizieren und erhalten so  $\mathbb{N}$  als Teilmenge von  $\mathbb{k}$ . Damit ist auch ein additives Inverses zu  $n$  in  $\mathbb{k}$  enthalten, was eine Identifikation mit  $-n \in \mathbb{Z}$  erlaubt. Für  $n \in \mathbb{N}$  ist  $n$  in  $\mathbb{k}$  auch multiplikativ invertierbar, da nach Definition alle Zahlen in  $\mathbb{k}^\times = \mathbb{k} \setminus \{0\}$  ein multiplikatives Inverses besitzen. Dieses bezeichnen wir wie üblich als

$$\frac{1}{n} = n^{-1}. \tag{1.2.22}$$

Damit erhalten wir aber auch für  $p \in \mathbb{Z}$  und  $q \in \mathbb{N}$  ein Element

$$\frac{p}{q} = p \cdot \frac{1}{q} \in \mathbb{k}, \tag{1.2.23}$$

was dann eine Identifikation von rationalen Zahlen mit Elementen aus  $\mathbb{k}$  liefert. Insgesamt erhalten wir also folgendes Resultat:

**Proposition 1.2.6** Sei  $\mathbb{k}$  ein Körper mit Charakteristik Null.

i.) Die rationalen Zahlen  $\mathbb{Q}$  sind ein Körper mit Charakteristik Null.

ii.) Die rationalen Zahlen  $\mathbb{Q}$  können als Unterkörper von  $\mathbb{k}$  mittels (1.2.23) aufgefasst werden.

BEWEIS: Den ersten Teil können wir entweder als aus der Schule bekannt annehmen oder direkt verifizieren. Letzteres erforderte aber streng genommen zuerst eine tragfähige Konstruktion von  $\mathbb{Q}$  aus beispielsweise  $\mathbb{N}$ , wovon wir hier absehen wollen. Für den zweiten Teil bleibt noch zu zeigen, dass die Abbildung von  $\frac{p}{q} \in \mathbb{Q}$  nach  $p \cdot q^{-1} \in \mathbb{k}$  injektiv ist, so dass wir wirklich eine Identifikation erhalten. Es zeigt sich, dass derartige strukturerhaltende Abbildungen zwischen Körpern prinzipiell injektiv sind, siehe auch Übung 1.8.1.  $\square$

Wir wollen nun die allgemeine Assoziativität und Kommutativität nutzen, um eine neue Schreibweise für Mehrfachsummen und Mehrfachprodukte einzuführen. Für  $n \in \mathbb{N}$  und  $x_1, \dots, x_n \in \mathbb{k}$  setzen wir

$$\sum_{i=1}^n x_i = x_1 + \dots + x_n \quad (1.2.24)$$

sowie

$$\prod_{i=1}^n x_i = x_1 \cdots x_n, \quad (1.2.25)$$

wobei es dank Proposition 1.2.4, iii.) beziehungsweise iv.), eben weder auf die Reihenfolge noch auf die Art der Klammerung ankommt. Gelesen wird dies als *Summe über  $i$  von 1 bis  $n$*  beziehungsweise als *Produkt über  $i$  von 1 bis  $n$* . Der Laufindex  $i$  hat selbst hierbei keine Bedeutung und kann beispielsweise gefahrlos umbenannt werden. Das Rechnen mit dem Summenzeichen und dem Produktzeichen wird in Übung 1.8.3 weiter illustriert.

Wir betrachten nun  $x \in \mathbb{k}$  und  $n \in \mathbb{N}$ . Dann definiert man

$$x^n = \prod_{i=1}^n x_i \quad \text{mit} \quad x_i = x \quad (1.2.26)$$

für alle  $i = 1, \dots, n$ . Für  $n = 0$  setzt man

$$x^0 = 1. \quad (1.2.27)$$

Ist nun  $x \neq 0$ , so definiert man

$$x^{-n} = \left(\frac{1}{x}\right)^n. \quad (1.2.28)$$

Durch diese Definitionen erreicht man nun folgende Rechenregeln:

**Proposition 1.2.7** Sei  $\mathbb{k}$  ein Körper und  $x, y \in \mathbb{k}^\times = \mathbb{k} \setminus \{0\}$ .

i.) Für  $n, m \in \mathbb{Z}$  gilt

$$x^n x^m = x^{n+m}. \quad (1.2.29)$$

ii.) Für  $n \in \mathbb{Z}$  gilt

$$x^n y^n = (xy)^n. \quad (1.2.30)$$

iii.) Für alle  $n, m \in \mathbb{Z}$  gilt

$$(x^n)^m = x^{nm} \quad (1.2.31)$$

und insbesondere  $(x^n)^{-1} = x^{-n}$ .

BEWEIS: Diese Eigenschaften gelten letztlich in jeder kommutativen Gruppe wie in unserem Beispiel in  $\mathbb{k}^\times$ . Der Beweis von (1.2.29) erfolgt am einfachsten durch eine Fallunterscheidung: Sind  $n, m \in \mathbb{N}$ , so folgt (1.2.29) direkt aus der Definition

$$x^n \cdot x^m = \underbrace{(x \cdots x)}_{n\text{-mal}} \cdot \underbrace{(x \cdots x)}_{m\text{-mal}} = \underbrace{x \cdots x}_{n+m\text{-mal}} = x^{n+m}$$

und der allgemeinen Assoziativität von  $\cdot$  gemäß Proposition 1.2.4, *iv.*). Gilt  $n, m < 0$  können wir  $x$  durch  $\frac{1}{x}$  ersetzen und (1.2.29) auf den Fall  $n, m > 0$  zurückführen. Ist  $n > 0$  und  $m < 0$ , so ist beispielsweise  $n \geq -m$  und daher

$$x^n x^m = \underbrace{x \cdots x}_{n\text{-mal}} \cdot \underbrace{\frac{1}{x} \cdots \frac{1}{x}}_{(-m)\text{-mal}} = \underbrace{x \cdots x}_{(n+m)\text{-mal}} = x^{n+m},$$

weil  $-m$ -mal  $x \frac{1}{x} = 1$  verwendet werden kann. Gilt stattdessen  $n < -m$ , so ersetzen wir wieder  $x$  durch  $\frac{1}{x}$  und haben diese Situation auf den gerade diskutierten Fall zurückgeführt. Ist nun eine der beiden Zahlen Null, so gilt (1.2.29) aus trivialen Gründen ebenfalls. Der zweite Teil folgt aus der Kommutativität durch Umsortierung, siehe auch Übung 1.8.4. Für den letzten Teil überlegt man sich zunächst, dass  $x^{-n} = (\frac{1}{x})^n$  tatsächlich das multiplikative Inverse von  $x^n$  ist. Dies ist als Spezialfall in (1.2.29) enthalten. Damit kann man den dritten Teil darauf zurückführen, dass sowohl  $n$  als auch  $m$  positiv sind, da man anderenfalls geeignet zu Inversen übergeht. Der Beweis erfolgt dann durch Induktion nach  $m$  aus dem Induktionsanfang  $(x^n)^1 = x^n = x^{n+1}$  und *i.*).  $\square$

Im Hinblick auf die Definition (1.2.27) liegt es nahe, das *leere Produkt* immer als 1 zu definieren. Entsprechend definiert man die *leere Summe* als 0. Dies erleichtert manchmal Beweise und vermeidet unnötige Fallunterscheidungen.

Wir nutzen die Produktschreibweise nun für die Definition der *Fakultät* und der *Binomialkoeffizienten*:

**Definition 1.2.8 (Fakultät und Binomialkoeffizienten)** Seien  $k \leq n \in \mathbb{N}_0$ . Dann definiert man die Fakultät von  $n$  als

$$n! = \prod_{i=1}^n i \quad (1.2.32)$$

und den Binomialkoeffizienten  $\binom{n}{k}$  durch

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \quad (1.2.33)$$

Hier gilt für  $n = 0$

$$0! = 1, \quad (1.2.34)$$

da in diesem Fall das Produkt (1.2.32) leer ist und somit unsere obige Konvention zum Tragen kommt. Es gilt  $n! \in \mathbb{N}$  sowie  $\binom{n}{k} \in \mathbb{N}$ , was anhand der Definition (1.2.33) nicht unmittelbar klar ist, sondern eine kleine Zusatzüberlegung erfordert. Diese und weitere Eigenschaften der Fakultät sowie der Binomialkoeffizienten werden in Übung 1.8.5 diskutiert.

Um sich mit der Summenschreibweise besser vertraut zu machen, zeigen wir noch folgende zwei wichtige Formeln:

**Proposition 1.2.9 (Binomischer Lehrsatz)** Sei  $\mathbb{k}$  ein Körper und  $x, y \in \mathbb{k}$ . Für alle  $n \in \mathbb{N}_0$  gilt dann

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}. \quad (1.2.35)$$

BEWEIS: Der Beweis erfolgt durch Induktion nach  $n$ . Für  $n = 0$  gilt  $\binom{0}{0} = 1$ , womit beide Seiten gleich 1 sind. Für den Induktionsschritt nehmen wir die Behauptung für  $n$  als richtig an und betrachten

$$(x + y)^{n+1} = (x + y)(y + y)^n$$

$$\begin{aligned}
&= (x+y) \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \\
&= \sum_{k=0}^n \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k y^{n-k+1} \\
&\stackrel{(a)}{=} x^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=1}^{n-1} \binom{n}{k} x^k y^{n-k+1} + y^{n+1} \\
&\stackrel{(b)}{=} x^{n+1} + \sum_{k=1}^n \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=1}^n \binom{n}{k} x^k y^{n+1-k} + y^{n+1} \\
&\stackrel{(c)}{=} x^{n+1} + \sum_{k=0}^n \left( \binom{n}{k-1} + \binom{n}{k} \right) x^k y^{n+1-k} + y^{n+1} \\
&\stackrel{(d)}{=} x^{n+1} + \sum_{k=1}^n \binom{n+1}{k} x^k y^{n+1-k} + y^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k}.
\end{aligned}$$

Hier haben wir zunächst das Distributivgesetz und die Kommutativität der Multiplikation verwendet und dann in (a) den letzten beziehungsweise ersten Summanden separat geschrieben. In (b) haben wir in der ersten Summe den Index um eins verschoben, dann in (c) die beiden Summen zusammengefasst. Entscheidend ist nun die Identität

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$$

für Binomialkoeffizienten in (d), siehe erneut Übung 1.8.5. Zum Schluss nimmt man die beiden Terme  $x^{n+1}$  und  $y^{n+1}$  wieder als nullten beziehungsweise letzten Term zur Summe hinzu.  $\square$

Man beachte, dass in einem Körper mit Primzahl-Charakteristik die Binomialkoeffizienten Null werden können, ebenso wie die Fakultäten. Nichts desto trotz bleibt der Binomische Lehrsatz gültig. Die zweite wichtige Formel ist die *geometrische Summe*:

**Proposition 1.2.10 (Geometrische Summe)** Sei  $\mathbb{k}$  ein Körper und  $x \neq 1$ . Dann gilt für alle  $n \in \mathbb{N}_0$

$$\sum_{k=0}^n x^k = \frac{1-x^{n+1}}{1-x}. \quad (1.2.36)$$

BEWEIS: Man kann dies entweder durch Induktion nach  $n$  zeigen, siehe Übung ??, oder direkt. Wir entscheiden uns für den direkten Nachweis und berechnen

$$\begin{aligned}
(1-x) \sum_{k=0}^n x^k &= \sum_{k=0}^n x^k - \sum_{k=0}^n x^{k+1} \\
&= 1 + \sum_{k=1}^n x^k - \sum_{k=1}^{n+1} x^k \\
&= 1 + \sum_{k=1}^n x^k - \sum_{k=1}^n x^k - x^{n+1} \\
&= 1 - x^{n+1}.
\end{aligned}$$

Ist nun  $x \neq 1$ , können wir durch  $1-x$  teilen und erhalten (1.2.36).  $\square$

### 1.3 Angeordnete Körper

In der Analysis sind wir nicht nur an Gleichungen und deren Manipulation interessiert, sondern wir wollen Zahlen auch in ihrer Größe vergleichen können, um Ungleichungen behandeln zu können. Dieses Kerngeschäft der Analysis erfordert nun mehr als einen Körper alleine. Wir benötigen einen *angeordneten Körper*:

**Definition 1.3.1 (Angeordneter Körper)** Ein angeordneter Körper ist ein Paar  $(\mathbb{k}, \mathbb{k}^+)$  bestehend aus einem Körper  $\mathbb{k}$  und einer Teilmenge  $\mathbb{k}^+ \subseteq \mathbb{k}$  mit folgenden Eigenschaften:

i.) Der Körper  $\mathbb{k}$  ist die disjunkte Vereinigung

$$\mathbb{k} = \mathbb{k}^- \cup \{0\} \cup \mathbb{k}^+, \quad (1.3.1)$$

wobei  $\mathbb{k}^- = \{x \in \mathbb{k} \mid -x \in \mathbb{k}^+\}$ .

ii.) Es gilt

$$\mathbb{k}^+ + \mathbb{k}^+ \subseteq \mathbb{k}^+ \quad (1.3.2)$$

und

$$\mathbb{k}^+ \cdot \mathbb{k}^+ \subseteq \mathbb{k}^+. \quad (1.3.3)$$

In einem angeordneten Körper nennen wir die Elemente aus  $\mathbb{k}^+$  positiv, die Elemente aus  $\mathbb{k}^-$  entsprechend negativ.

Da wir in (1.3.1) auf einer *disjunkten* Vereinigung bestehen, ist eine Zahl in  $\mathbb{k}$  entweder positiv, negativ oder Null.

**Beispiel 1.3.2** Die rationalen Zahlen  $\mathbb{Q}$  bilden einen angeordneten Körper, wenn man für  $\mathbb{Q}^+$  die üblichen positiven rationalen Zahlen nimmt, also

$$\mathbb{Q}^+ = \left\{ \frac{p}{q} \in \mathbb{Q} \mid p, q \in \mathbb{N} \right\}. \quad (1.3.4)$$

Die negativen Zahlen  $\mathbb{Q}^-$  in  $\mathbb{Q}$  sind dann

$$\mathbb{Q}^- = \left\{ \frac{p}{q} \in \mathbb{Q} \mid -p, q \in \mathbb{N} \right\}. \quad (1.3.5)$$

Man macht sich schnell klar, dass jede rationale Zahl ungleich Null eine Darstellung als Bruch entweder mit positivem oder negativem Zähler bei positivem Nenner besitzt. Damit ist  $\mathbb{Q} = \mathbb{Q}^- \cup \{0\} \cup \mathbb{Q}^+$  tatsächlich gewährleistet. Bemerkenswerterweise ist dies die einzige Möglichkeit, den Körper  $\mathbb{Q}$  anzuordnen, siehe auch Übung 1.8.9.

Die Teilmenge  $\mathbb{k}^+ \subseteq \mathbb{k}$  in einem angeordneten Körper kann nun dazu verwendet werden, eine Ordnungsrelation auf  $\mathbb{k}$  zu definieren:

**Definition 1.3.3 (Ordnungsrelation  $<$ )** Sei  $(\mathbb{k}, \mathbb{k}^+)$  ein angeordneter Körper. Für  $x, y \in \mathbb{k}$  definiert man die Relation „ $<$ “ durch

$$x < y \iff y - x \in \mathbb{k}^+. \quad (1.3.6)$$

Entsprechend werden die Relationen „ $\leq$ “, „ $>$ “, und „ $\geq$ “ definiert.

Diese „kleiner als“-Relation verhält sich nun erwartungsgemäß so, wie man es vom Beispiel  $\mathbb{Q}$  gewohnt ist:

**Proposition 1.3.4** Sei  $(\mathbb{k}, \mathbb{k}^+)$  ein angeordneter Körper.



Abbildung 1.1: Ein angeordneter Körper als Zahlenstrahl. Die Zahlen werden der Größe nach von links nach rechts angeordnet.

- i.) Für  $x, y \in \mathbb{k}$  gilt entweder  $x < y$ , oder  $x = y$ , oder  $x > y$ .
- ii.) Für  $x, y, z \in \mathbb{k}$  mit  $x < y$  und  $y < z$  folgt  $x < z$ .
- iii.) Für  $x, y, z \in \mathbb{k}$  mit  $x < y$  folgt  $x + z < y + z$ . Gilt zudem  $z > 0$ , so folgt auch  $xz < yz$ .

BEWEIS: Für den ersten Teil betrachtet man  $z = y - x$ . Da definitionsgemäß die Zerlegung (1.3.1) disjunkt ist, gibt es drei sich wechselseitig ausschließende Möglichkeiten:  $z \in \mathbb{k}^-$ ,  $z = 0$ , oder  $z \in \mathbb{k}^+$ . Diese bedeuten entsprechend  $-z = y - x \in \mathbb{k}^+$ , also  $y > x$ , oder  $x = y$ , oder  $x > y$ . Für den zweiten Teil betrachten wir  $x < y$  und  $y < z$ , also  $y - x \in \mathbb{k}^+$  und  $z - y \in \mathbb{k}^+$ . Nach (1.3.2) folgt  $z - x = y - x + z - y \in \mathbb{k}^+$ , was aber zu  $x < z$  äquivalent ist. Sei nun  $x < y$  und  $z \in \mathbb{k}$  beliebig. Dann gilt  $y + z - (x + z) = y - x \in \mathbb{k}^+$  nach (1.3.2), also  $y + z > x + z$ . Sei zudem  $z > 0$ . Dann gilt  $zy - zx = z(y - x) \in \mathbb{k}^+$  nach (1.3.3), also  $zx < zy$ .  $\square$

Man kann umgekehrt zeigen, dass ein Körper  $\mathbb{k}$  zusammen mit einer Relation  $<$ , welche die Eigenschaften i.), ii.) und iii.) der Proposition erfüllt, zu einem angeordneten Körper wird, indem man

$$\mathbb{k}^+ = \{x \in \mathbb{k} \mid x > 0\} \subseteq \mathbb{k} \quad (1.3.7)$$

setzt, siehe Übung 1.8.10. Beide Versionen werden in der Literatur zur Definition eines angeordneten Körpers herangezogen.

Die Ordnungsrelation erlaubt es nun, einen angeordneten Körper zumindest symbolisch als *Zahlenstrahl* darzustellen, wobei wir die Zahlen von links nach rechts der Größe nach ordnen, siehe auch Abbildung 1.1. Wir werden später jedoch verschiedene konzeptionelle Schwierigkeiten mit der Darstellung als Zahlenstrahl sehen, womit diese Art der graphischen Darstellung nicht zu sehr strapaziert werden sollte.

Wir zeigen nun einige weitere Rechenregeln und Eigenschaften von angeordneten Körpern, die von den rationalen Zahlen  $\mathbb{Q}$  her vertraut sind:

**Proposition 1.3.5** Sei  $(\mathbb{k}, \mathbb{k}^+)$  ein angeordneter Körper.

- i.) Sei  $x \in \mathbb{k}$ . Es gilt genau dann  $x \neq 0$ , wenn  $x^2 > 0$ .
- ii.) Es gilt  $1 > 0$  und  $n = 1 + \dots + 1 > 0$  für alle  $n \in \mathbb{N}$ .
- iii.) Die Charakteristik von  $\mathbb{k}$  ist Null.
- iv.) Für  $x, y \in \mathbb{k}$  gilt genau dann  $xy > 0$ , wenn  $x > 0$  und  $y > 0$  oder wenn  $x < 0$  und  $y < 0$  gilt.
- v.) Sind  $x_1, \dots, x_n \in \mathbb{k}$ , so gilt genau dann  $\sum_{i=1}^n x_i^2 = 0$ , wenn  $x_1 = \dots = x_n = 0$  gilt.
- vi.) Es gilt genau dann  $x \in \mathbb{k}^+$ , wenn  $\frac{1}{x} \in \mathbb{k}^+$ .
- vii.) Für  $0 < x < y$  gilt  $0 < x^n < y^n$  für alle  $n \in \mathbb{N}$ .

BEWEIS: Sei  $x \neq 0$ , dann ist entweder  $x > 0$  oder  $x < 0$ . Im letzteren Fall ist  $-x > 0$ , also  $x^2 = (-x)(-x) > 0$ . Im Fall  $x > 0$  ist  $x^2 > 0$  klar. Da  $0^2 = 0$  gilt, folgt auch die Umkehrung, was i.) zeigt. Für das Einselement gilt  $1 = 1^2 > 0$  nach i.). Dann ist aber auch  $1 + 1 > 0$  positiv und per Induktion auch die  $n$ -fache Summe, was  $n > 0$  in  $\mathbb{k}$  zeigt. Damit ist aber insbesondere  $n \neq 0$  in  $\mathbb{k}$  für alle  $n \in \mathbb{N}$ , was iii.) zeigt. Für iv.) betrachtet man  $x, y \in \mathbb{k}^\times$  beide ungleich Null, da sonst  $xy = 0$  nicht positiv sein kann. Es bleiben vier exklusive Möglichkeiten: beide positiv, beide negativ,  $x$  positiv und  $y$  negativ oder  $x$  negativ und  $y$  positiv. Sind beide Zahlen positiv, so auch ihr Produkt nach Definition. Sind beide negativ, so gilt  $xy = (-x)(-y) > 0$ , da das Produkt zweier positiver Zahlen  $-x, -y$  definitionsgemäß positiv ist. Ist  $x > 0$  und  $y < 0$ , so ist  $-y > 0$  und daher

$-(xy) = x(-y) > 0$ , was  $xy < 0$  zeigt. Den verbleibenden Fall  $x < 0$  und  $y > 0$  erhält man genauso, was insgesamt *iv.*) zeigt. Seien nun  $x_1, \dots, x_n \in \mathbb{K}$  mit  $\sum_{i=1}^n x_i^2 = 0$  gegeben. Da  $x_i^2$  entweder Null oder positiv ist und die Summe von positiven Zahlen positiv bleibt, müssen also alle  $x_i^2$  Null gewesen sein. Damit gilt aber auch  $x_i = 0$  für alle  $i = 1, \dots, n$ . Die Umkehrung ist klar. Sei  $x \in \mathbb{K}^+$ . Da dann insbesondere  $x \neq 0$  gilt, kann  $x$  invertiert werden, und  $\frac{1}{x}$  ist ebenfalls von Null verschieden. Da  $1 = x \cdot \frac{1}{x} > 0$  gilt, liefert die Fallunterscheidung zu *iv.*) sofort die einzige verbleibende Möglichkeit  $\frac{1}{x} > 0$ . Die Umkehrung ist aus Symmetriegründen durch Vertauschen von  $x$  und  $\frac{1}{x}$  auch klar. Hierzu beachte man, dass  $\frac{1}{x} \in \mathbb{K}^+$  sowieso nur für  $x \neq 0$  möglich ist. Zum Abschluss betrachten wir  $0 < x < y$ . Wir zeigen  $x^n < y^n$  durch Induktion nach  $n$ . Der Induktionsanfang  $n = 1$  ist gerade die Voraussetzung von *vii.*) und damit trivialerweise korrekt. Es gelte also  $0 < x^n < y^n$ . Nach Proposition 1.3.4, *iii.*), gilt mit  $x > 0$  auch

$$0 = 0 \cdot x < x^n \cdot x = x^{n+1} < y^n \cdot x.$$

Wegen  $x < y$  und  $y^n > 0$  gilt schließlich auch  $y^n \cdot x < y^n \cdot y = y^{n+1}$ , womit alles gezeigt ist.  $\square$

Insbesondere gibt es also Körper wie beispielsweise  $\mathbb{Z}_2$ , die nicht angeordnet werden können. Alle Körper mit Primzahl-Charakteristik sind von dieser Art. Den vierten Punkt können wir auch als

$$\mathbb{K}^+ \cdot \mathbb{K}^+ \subseteq \mathbb{K}^+, \quad \mathbb{K}^+ \cdot \mathbb{K}^- \subseteq \mathbb{K}^- \quad \text{und} \quad \mathbb{K}^- \cdot \mathbb{K}^- \subseteq \mathbb{K}^+ \quad (1.3.8)$$

zusammenfassen.

In einem angeordneten Körper können wir nun von *Ungleichungen* sprechen. Dies ist ein wesentliches Ziel der Analysis. Wir geben ein erstes nicht-triviales Beispiel einer Ungleichung, die *Bernoulli-Ungleichung*:

**Proposition 1.3.6 (Bernoulli-Ungleichung)** *Sei  $(\mathbb{K}, \mathbb{K}^+)$  ein angeordneter Körper. Sei  $x \in \mathbb{K} \setminus \{0\}$  mit  $x > -1$  gegeben. Dann gilt für alle  $n \in \mathbb{N}$  mit  $n \geq 2$  die Bernoulli-Ungleichung*

$$(1+x)^n > 1+nx. \quad (1.3.9)$$

BEWEIS: Der Beweis erfolgt durch Induktion. Für  $n = 2$  gilt

$$(1+x)^2 = 1+2x+x^2 > 1+2x,$$

da für  $x \neq 0$  das Quadrat  $x^2$  strikt positiv ist. Gelte also (1.3.9) für  $n$ , dann betrachten wir

$$\begin{aligned} (1+x)^{n+1} &= (1+x)(1+x)^n \\ &\stackrel{(a)}{>} (1+x)(1+nx) \\ &= 1+x+nx+nx^2 \\ &> 1+(n+1)x, \end{aligned}$$

wobei wir in (a) neben der Induktionsvoraussetzung verwenden, dass  $1+x > 0$  gilt. Im letzten Schritt verwenden wir  $nx^2 > 0$  für  $x \neq 0$ .  $\square$

Für  $n = 1$  wird die Bernoulli-Ungleichung zur Gleichung, für  $x \leq -1$  ist sie im allgemeinen falsch, wie man sich anhand einfacher Gegenbeispiele überlegt.

**Bemerkung 1.3.7 (Abschätzung)** Eine Ungleichung nennt man auch eine *Abschätzung*. In einer Ungleichung der Form  $x \leq y$  wird  $x$  nach oben durch  $y$  abgeschätzt, umgekehrt wird  $y$  nach unten durch  $x$  abgeschätzt. Will man betonen, dass keine Gleichheit vorliegt, so nennt man eine Abschätzung oder Ungleichung *strikt*. In diesem Fall schreibt man zur Betonung anstelle von  $x < y$  auch  $x \lessdot y$ .

**Bemerkung 1.3.8 (Eindeutigkeit der Anordnung)** Im allgemeinen ist die Anordnung eines Körpers  $\mathbb{k}$ , also die Auswahl von  $\mathbb{k}^+$ , nicht eindeutig, sofern es überhaupt möglich ist,  $\mathbb{k}$  anzuordnen. Die Beispiele sind für uns momentan weder zugänglich noch interessant für die Analysis. Hat man einen Körper  $\mathbb{k}$  mit der Eigenschaft, dass zu jedem  $x \in \mathbb{k}$  ein  $y \in \mathbb{k}$  existiert, so dass entweder  $x = y^2$  oder  $x = -y^2$  gilt, so ist eine Anordnung auf jeden Fall eindeutig, da ja Quadrate in einem angeordneten Körper notwendigerweise positiv sind. Die Existenz einer Anordnung ist dagegen ein weitaus schwierigeres Problem und wird in der Algebra der formal reellen Körper behandelt.

Zum Abschluss dieses Abschnitts definieren wir noch den *Betrag* oder auch *Absolutbetrag* eines Elements in einem angeordneten Körper:

**Definition 1.3.9 (Absolutbetrag)** Sei  $(\mathbb{k}, \mathbb{k}^+)$  ein angeordneter Körper. Der Absolutbetrag (auch Betrag) von  $x \in \mathbb{k}$  ist dann als

$$|x| = \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases} \quad (1.3.10)$$

definiert.

**Proposition 1.3.10** Sei  $(\mathbb{k}, \mathbb{k}^+)$  ein angeordneter Körper und  $x, y \in \mathbb{k}$ .

i.) Es gilt  $|x| = 0$  genau dann, wenn  $x = 0$ .

ii.) Es gilt  $|xy| = |x| \cdot |y|$ .

iii.) Es gilt die Dreiecksungleichung

$$|x + y| \leq |x| + |y|. \quad (1.3.11)$$

BEWEIS: Der erste Teil ist klar, der zweite folgt ebenfalls direkt aus der Definition des Betrags. Für (1.3.11) benötigt man eine Fallunterscheidung gemäß der Positivität und Negativität von  $x$  und  $y$ , siehe auch Übung 1.8.12.  $\square$

Wir werden noch viele Verallgemeinerungen der Dreiecksungleichung sehen, die dann die geometrische Bedeutung in der Dreiecksgeometrie aufzeigen und so den Namen rechtfertigen.

## 1.4 Archimedische Körper

Ist  $\mathbb{k}$  ein Körper der Charakteristik Null, so enthält er  $\mathbb{Q} \subseteq \mathbb{k}$  als Unterkörper, wie wir in Proposition 1.2.6, ii.), gesehen haben. Ist  $\mathbb{k}$  sogar ein angeordneter Körper, so ist die Einbettung aus Proposition 1.2.6, ii.), sogar *ordnungserhaltend*, also

$$\mathbb{Q}^+ \subseteq \mathbb{k}^+, \quad (1.4.1)$$

siehe Übung 1.8.13. Es stellt sich daher die Frage, um wieviel größer ein angeordneter Körper im Vergleich zum kleinsten angeordneten Körper  $\mathbb{Q}$  sein kann.

Hier gibt es zwei qualitativ verschiedene Aspekte. Zum einen könnte es Zahlen in  $\mathbb{k}$  geben, die zwischen rationalen Zahlen liegen. Wir werden noch sehen, dass die rationalen Zahlen in gewisser Hinsicht hier ein großes Defizit besitzen, da gewisse erstrebenswerte Konstruktionen in  $\mathbb{Q}$  nicht möglich sind. Zum anderen besteht die Möglichkeit von Zahlen, die so groß sind, dass sie insbesondere größer als jede rationale Zahl sind. Derartige Körper gibt es tatsächlich. Wir wollen sie aber im Folgenden ausschließen. Deshalb definiert man die *archimedischen Körper*:

**Definition 1.4.1 (Archimedischer Körper)** Ein angeordneter Körper  $(\mathbb{k}, \mathbb{k}^+)$  heißt *archimedisches*, falls es zu  $x > 0$  in  $\mathbb{k}$  immer ein  $n \in \mathbb{N}$  mit

$$x < n \quad (1.4.2)$$

gibt. Andernfalls heißt  $\mathbb{k}$  *nicht-archimedisches*.

Der angeordnete Körper  $\mathbb{Q}$  der rationalen Zahlen ist archimedisch, da zu  $\frac{p}{q} \in \mathbb{Q}^+$  beispielsweise die natürliche Zahl  $n = p + 1$  das Gewünschte leistet.

**Proposition 1.4.2** Sei  $(\mathbb{k}, \mathbb{k}^+)$  ein geordneter Körper. Dann sind folgende Eigenschaften äquivalent:

- i.) Der Körper ist archimedisch.
- ii.) Zu allen  $x, y \in \mathbb{k}^+$  gibt es ein  $n \in \mathbb{N}$  mit  $y < nx$ .
- iii.) Zu allen  $x, y \in \mathbb{k}$  mit  $x < y$  gibt es ein  $q \in \mathbb{Q}$  mit  $x < q < y$ .
- iv.) Zu allen  $x \in \mathbb{k}^+$  gibt es ein  $q \in \mathbb{Q}$  mit  $0 < q < x$ .
- v.) Zu allen  $x, y \in \mathbb{k}$  mit  $x > 1$  und  $y > 0$  gibt es ein  $n \in \mathbb{N}$  mit  $x^n > y$ .

BEWEIS: Wir zeigen  $i.) \implies ii.) \implies iii.) \implies iv.) \implies v.) \implies i.)$ . Sei also  $\mathbb{k}$  archimedisch und  $x, y \in \mathbb{k}^+$ . Dann ist  $\frac{x}{y} > 0$  und es gibt ein  $n \in \mathbb{N}$  mit  $\frac{x}{y} < n$ . Dies liefert  $y < nx$ . Gelte nun  $ii.)$  und seien  $x < y$  vorgegeben. Ist  $x < 0$  und  $y > 0$ , so erfüllt  $0 \in \mathbb{Q}$  das Geforderte. Sei also zuerst  $0 < x < y$  mit  $n(y - x) > 1$  nach  $ii.)$  und daher  $ny > nx + 1$ . Sei nun  $m \in \mathbb{N}$  die kleinste natürliche Zahl, für die  $nx < m$  gilt: Es gibt solche  $m$  nach erneuter Anwendung von  $ii.)$  auf  $nx$  und 1. Dann gilt also  $nx \geq m - 1$ , da sonst  $m - 1$  ebenfalls eine solche natürliche Zahl wäre. Insgesamt gilt also

$$ny > nx + 1 \geq m > nx,$$

und damit  $y > \frac{m}{n} > x$ , was  $iii.)$  für diesen Fall zeigt. Ist nun  $x < y < 0$ , so gilt  $0 < -y < -x$  und wir können diesen Fall auf den gerade gezeigten zurückführen. Die Implikation  $iii.) \implies iv.)$  ist trivial. Sei also  $iv.)$  erfüllt. Für  $x > 1$  ist  $x = 1 + z$  mit  $z > 0$ . Wir finden daher eine rationale Zahl  $q \in \mathbb{Q}$  mit  $0 < q < \frac{z}{y}$ . Ohne Einschränkung können wir  $q = \frac{1}{n}$  annehmen. Dann gilt also  $\frac{1}{n} < \frac{z}{y}$  oder  $y < nz$ . Mit der Bernoulli-Ungleichung folgt

$$x^n = (1 + z)^n > 1 + nz > y$$

wie gewünscht. Es gelte nun also  $v.)$ . Sei weiter  $x > 0$  derart gegeben, dass es kein  $n \in \mathbb{N}$  mit  $x < n$  gibt. Wir wollen diese Annahme zum Widerspruch führen. Es gilt also für alle  $n \in \mathbb{N}$  die Ungleichung  $n \leq x$  oder eben  $\frac{n}{x} \leq 1$ . Insbesondere gilt  $\frac{1}{x} \leq 1$  und daher

$$0 < \frac{1}{x^k} \leq \frac{1}{x} \leq 1$$

für alle  $k \in \mathbb{N}$  nach Proposition 1.3.4,  $iii.)$ . Mit dem Binomischen Lehrsatz folgt dann

$$\left(1 + \frac{1}{x}\right)^n = \sum_{k=0}^n \binom{n}{k} \frac{1}{x^k} \leq 1 + \sum_{k=1}^n \binom{n}{k} \frac{1}{x} \leq 1 + \frac{2^n}{x} \leq 2,$$

da zum einen  $\sum_{k=0}^n \binom{n}{k} = 2^n$ , siehe Übung ??, und zum anderen  $2^n \in \mathbb{N}$ , womit  $\frac{2^n}{x} \leq 1$ . Damit ist aber der Widerspruch zu  $v.)$  erreicht und die Annahme, dass  $n \leq x$  für alle  $n \in \mathbb{N}$  gilt, widerlegt. Dies zeigt die letzte Implikation  $v.) \implies i.)$ .  $\square$

Insbesondere können wir bei einem archimedischen Körper zwischen zwei verschiedenen Zahlen immer noch eine weitere rationale Zahl einschieben. Man spricht daher auch davon, dass  $\mathbb{Q}$  in  $\mathbb{k}$  *dicht* liegt. Später werden wir noch eine präzisere Definition einer dichten Teilmenge kennenlernen, für den Moment ist es daher eher die Anschauung, die als Motivation für diese Sprechweise dient.



Abbildung 1.2: Auf dem Zahlenstrahl gibt es zwischen zwei  $x, y \in \mathbb{k}$  immer ein rationales  $q \in \mathbb{Q} \subseteq \mathbb{k}$ .

## 1.5 Vollständigkeit und die reellen Zahlen

Viele mathematische Probleme lassen sich als Fragestellung in  $\mathbb{Q}$  formulieren, haben aber dort keine Lösung. Als bekanntes Beispiel betrachten wir die Gleichung  $x^2 = 2$ , deren (hypothetische) Lösung die Wurzel  $\sqrt{2}$  wäre:

**Beispiel 1.5.1 (Die Wurzel  $\sqrt{2}$  ist nicht rational)** Die Gleichung

$$x^2 = 2 \quad (1.5.1)$$

benutzt nur rationale Zahlen zu ihrer Formulierung, hat aber in  $\mathbb{Q}$  keine Lösung. Wir wiederholen das bekannte Argument. Wäre  $x = \frac{r}{s} \in \mathbb{Q}$  eine (positive) Lösung von (1.5.1), so könnten wir den Bruch als vollständig gekürzt annehmen. Die Zahlen  $r, s \in \mathbb{N}$  sind also teilerfremd. Es folgt  $r^2 = 2s^2$ , womit 2 ein Teiler von  $r^2$  ist. Dann ist aber notwendigerweise auch 4 ein Teiler von  $r^2$ , da  $r^2$  ja ein Quadrat ist. Somit muss einer der beiden Faktoren  $4 = 2 \cdot 2$  ein Teiler von  $s^2$  gewesen sein. Damit ist 2 aber ein Teiler von  $r$  und von  $s$  im Widerspruch zur Annahme. Eine negative Lösung von (1.5.1) führt auf den selben Widerspruch, da dann  $-x$  eine positive Lösung wäre.

Der Körper  $\mathbb{Q}$  ist also zu klein, um gewisse, aus naheliegenden Gründen interessante, mathematische Probleme zu lösen. Die nächste Idee ist daher, zumindest eine approximative Lösung zu erreichen. Wir illustrieren dies erneut am Beispiel der Quadratwurzel  $\sqrt{2}$ :

**Beispiel 1.5.2** Wir betrachten die folgenden beiden Teilmengen

$$A = \{x \in \mathbb{Q} \mid x \leq 0 \text{ oder } x^2 \leq 2\} \quad (1.5.2)$$

und

$$B = \{x \in \mathbb{Q} \mid x > 0 \text{ und } x^2 > 2\} \quad (1.5.3)$$

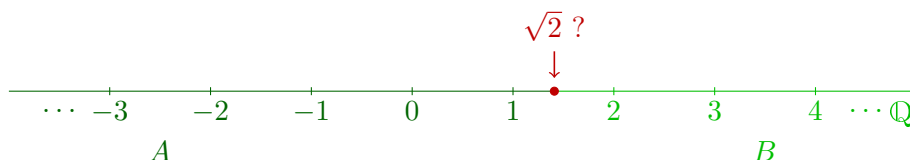
von  $\mathbb{Q}$ . Wir haben folgende elementare Eigenschaften: Es gilt für alle  $a \in A$  und  $b \in B$

$$a < b, \quad (1.5.4)$$

womit insbesondere  $A \cap B = \emptyset$ . Dies folgt sofort aus den Anordnungsaxiomen. Weiter gilt  $\mathbb{Q} = A \cup B$ , da eine Zahl  $x \in \mathbb{Q}$  entweder  $x^2 \leq 2$  oder  $x^2 > 2$  erfüllt. Symbolisiert man die angeordneten rationalen Zahlen wieder als Zahlenstrahl, so zeigt sich, dass  $\mathbb{Q}$  an der Anschlussstelle ein „Loch“ hat, siehe Abbildung 1.3. Da es keine Lösung zu (1.5.1) gibt, gilt nämlich

$$A = \{x \in \mathbb{Q} \mid x \leq 0 \text{ oder } x^2 < 2\}. \quad (1.5.5)$$

So gesehen ist die graphische Darstellung von  $\mathbb{Q}$  als Zahlenstrahl sehr irreführend, da wir an der Nahtstelle von  $A$  und  $B$  ebenso wie an vielen anderen Stellen „Löcher“ in  $\mathbb{Q}$  haben. Die in Abbildung 1.3 durchgezogene Linie entspricht also keineswegs den tatsächlichen Eigenschaften von  $\mathbb{Q}$ . Umgekehrt sehen wir, dass die Zahlen  $a \in A$  eine Approximation von unten an  $\sqrt{2}$  liefern, während die Zahlen  $b \in B$  eine Approximation von oben darstellen. Diese wird offenbar immer besser, wenn wir  $a \in A$  weiter nach rechts und  $b \in B$  weiter nach links verschieben. Als Güte der Approximation kann man dann die Differenz  $b - a$  ansehen. Je kleiner wir diese Differenz gestalten, desto bessere Approximationen  $a$  und  $b$  erhalten wir.

Abbildung 1.3: Die beiden Teilmengen  $A$  und  $B$  zerlegen  $\mathbb{Q}$  disjunkt.

Überraschenderweise enthält dieses Beispiel bereits alle relevanten Züge für das weitere Vorgehen. Die Defizite von  $\mathbb{Q}$  können auf die im Beispiel aufgezeigten Weisen genauer gefasst werden. Das erste in Beispiel 1.5.2 versteckte Konzept ist das der kleinsten oberen Schranke:

**Definition 1.5.3 (Supremum und Infimum)** Sei  $\mathbb{k}$  ein angeordneter Körper.

i.) Eine Teilmenge  $A \subseteq \mathbb{k}$  heißt nach oben beziehungsweise nach unten beschränkt, wenn es ein  $x \in \mathbb{k}$  gibt, so dass

$$a \leq x \quad \text{beziehungsweise} \quad x \leq a \quad (1.5.6)$$

für alle  $a \in A$  gilt. In diesem Fall heißt  $x$  eine obere beziehungsweise untere Schranke.

ii.) Eine Teilmenge  $A \subseteq \mathbb{k}$  heißt beschränkt, falls sie sowohl nach oben als auch nach unten beschränkt ist.

iii.) Die kleinste obere Schranke einer nach oben beschränkten Menge  $A \subseteq \mathbb{k}$  heißt Supremum von  $A$ , geschrieben als  $\sup A$ .

iv.) Die größte untere Schranke einer nach unten beschränkten Menge  $A \subseteq \mathbb{k}$  heißt Infimum von  $A$ , geschrieben als  $\inf A$ .

Kleinste obere Schranke  $x = \sup A$  zu sein, bedeutet, dass für alle anderen oberen Schranken  $y$  von  $A$  mit  $y \leq x$  bereits  $x = y$  gilt. Entsprechend ist der Begriff größte untere Schranke zu verstehen, siehe auch Übung 1.8.16 für einige einfache aber nichts desto trotz nützliche Umformulierungen.

**Beispiel 1.5.4** Sei  $\mathbb{k}$  ein angeordneter Körper.

- i.) Die Teilmenge  $\mathbb{N} \subseteq \mathbb{k}$  der natürlichen Zahlen ist genau dann nach oben unbeschränkt, wenn  $\mathbb{k}$  archimedisch ist: Zunächst ist  $1 \in \mathbb{k}$  sicherlich eine untere Schranke, die sogar zu  $\mathbb{N}$  gehört und das Infimum von  $\mathbb{N}$  ist. Die Definition eines archimedischen Körpers ist direkt zur Nicht-Existenz einer oberen Schranke äquivalent, womit die Behauptung folgt.
- ii.) Jede endliche Teilmenge  $A \subseteq \mathbb{k}$  ist beschränkt. Eine obere Schranke ist beispielsweise das Maximum von  $A$ , eine untere Schranke das Minimum.
- iii.) Die Teilmenge  $A \subseteq \mathbb{Q}$  aus Beispiel 1.5.2 ist nach oben beschränkt: Jede Zahl aus  $B$  ist eine obere Schranke. Entsprechend ist  $B$  nach unten beschränkt und jedes  $a \in A$  ist eine untere Schranke von  $B$ . Allerdings hat  $A$  kein Supremum. Wäre nämlich  $x \in \mathbb{Q}$  das Supremum von  $A$ , so gälte  $x \leq b$  für alle  $b \in B$ , da ja alle  $b \in B$  obere Schranke sind. Da es zu jedem  $b \in B$  auch ein  $b' \in B$  mit  $b' < b$  gibt, kann  $x$  kein Element von  $B$  sein. Ist nämlich  $b \in B$  und daher  $b^2 > 2$ , so erfüllt

$$b' = b - \frac{b^2 - 2}{2b} = \frac{b}{2} + \frac{1}{b} > 0$$

die Eigenschaft

$$(b')^2 = \left(b - \frac{b^2 - 2}{2b}\right)^2$$

$$\begin{aligned}
&= b^2 - (b^2 - 2) + \left(\frac{b^2 - 2}{2b}\right)^2 \\
&= 2 + \left(\frac{b^2 - 2}{2b}\right)^2 \\
&> 2,
\end{aligned}$$

und liegt damit tatsächlich in  $B$ . Es gilt aber auch  $b' < b$  wie gewünscht. Da nun  $A \cup B = \mathbb{Q}$  gilt, müsste für das Supremum  $x \in A$  gelten. Damit wäre also  $x^2 \leq 2$ . Wäre nun  $x^2 < 2$ , so gäbe es  $a \in A$  mit  $x < a$ , da man das gleiche Argument wie für  $B$  anwenden kann. Damit könnte  $x$  aber keine obere Schranke von  $A$  sein. Es bleibt also nur die Möglichkeit  $x^2 = 2$ , die wir für  $x \in \mathbb{Q}$  bereits ausgeschlossen hatten. Mit einer analogen Überlegung zeigt man, dass die nach unten beschränkte Teilmenge  $B$  kein Infimum besitzt. Dieses hypothetische Infimum von  $x$  wäre wieder eine Lösung der Gleichung  $x^2 = 2$ .

iv.) Die leere Teilmenge  $\emptyset \subseteq \mathbb{k}$  ist beschränkt, da es für alle ihre (nicht existenten) Elemente  $a$  obere und untere Abschätzungen gibt. Jedes  $x \in \mathbb{k}$  liefert sowohl eine obere als auch eine untere Schranke. Es gibt aber weder ein Supremum noch ein Infimum in  $\mathbb{k}$ .

Ist die Teilmenge  $A \subseteq \mathbb{k}$  nach oben beziehungsweise nach unten unbeschränkt, so schreiben wir auch

$$\sup A = +\infty \quad \text{beziehungsweise} \quad \inf A = -\infty. \quad (1.5.7)$$

Dies ist jedoch als symbolische Schreibweise zu verstehen:  $\pm\infty$  sind keineswegs Elemente von  $\mathbb{k}$ . Weiter setzt man  $\sup \emptyset = -\infty$  und  $\inf \emptyset = +\infty$ .

Das nächste wichtige Konzept, das wir Beispiel 1.5.2 entnehmen, ist die Aufteilung von  $\mathbb{Q}$  in die disjunkte Zerlegung  $A \cup B = \mathbb{Q}$ . Allgemein nennt man eine derartige Aufteilung einen *Dedekindschen Schnitt*:

**Definition 1.5.5 (Dedekindscher Schnitt)** Sei  $\mathbb{k}$  ein angeordneter Körper. Zwei nicht-leere Teilmengen  $A, B \subseteq \mathbb{k}$  heißen *Dedekindscher Schnitt* von  $\mathbb{k}$ , falls  $A$  nach unten unbeschränkt ist,  $B = \mathbb{k} \setminus A$  gilt und

$$a < b \quad (1.5.8)$$

für alle  $a \in A$  und  $b \in B$ .

Wir erhalten also eine disjunkte Zerlegung

$$\mathbb{k} = A \cup B \quad (1.5.9)$$

wie schon in Beispiel 1.5.2. Da für einen Dedekindschen Schnitt explizit

$$B = \mathbb{k} \setminus A \quad (1.5.10)$$

gefordert ist, bestimmt die Teilmenge  $A$  die Teilmenge  $B$  und umgekehrt.

**Beispiel 1.5.6 (Dedekindscher Schnitt mit Trennungszahl)** Sei  $\mathbb{k}$  ein angeordneter Körper und  $x \in \mathbb{k}$ . Dann definieren die beiden Teilmengen

$$A(x) = \{y \in \mathbb{k} \mid y \leq x\} \quad (1.5.11)$$

und

$$B(x) = \{y \in \mathbb{k} \mid y > x\} \quad (1.5.12)$$

einen Dedekindschen Schnitt. Genauso können wir

$$\tilde{A}(x) = \{y \in \mathbb{K} \mid y < x\} \quad (1.5.13)$$

und

$$\tilde{B}(x) = \{y \in \mathbb{K} \mid y \geq x\} \quad (1.5.14)$$

verwenden. In beiden Fällen ist die Zahl  $x \in \mathbb{K}$  nun dadurch charakterisiert, dass

$$a \leq x \leq b \quad (1.5.15)$$

für alle  $a \in A$  und  $b \in B$ . Eine solche Zahl heißt *Trennungszahl*. Das Beispiel 1.5.2 zeigt nun, dass nicht alle Dedekindschen Schnitte in  $\mathbb{Q}$  von dieser Form sind. Die benötigte Zahl  $x = \sqrt{2}$  ist nicht in  $\mathbb{Q}$  vorhanden. In diesem Fall hat der Dedekindsche Schnitt also keine Trennungszahl.

Um schließlich den Approximationsgedanken aus Beispiel 1.5.2 aufzugreifen, benötigen wir den Begriff des Intervalls:

**Definition 1.5.7 (Intervalle)** Sei  $\mathbb{K}$  ein angeordneter Körper und  $a, b \in \mathbb{K} \cup \{\pm\infty\}$ .

i.) Das offene Intervall  $(a, b)$  für  $-\infty \leq a < b \leq +\infty$  ist definiert als die Teilmenge

$$(a, b) = \{x \in \mathbb{K} \mid a < x < b\} \subseteq \mathbb{K}. \quad (1.5.16)$$

ii.) Das abgeschlossene Intervall  $[a, b]$  für  $-\infty < a \leq b < \infty$  ist definiert als die Teilmenge

$$[a, b] = \{x \in \mathbb{K} \mid a \leq x \leq b\} \subseteq \mathbb{K}. \quad (1.5.17)$$

iii.) Weiter definiert man die halboffenen Intervalle  $(a, b]$  für  $-\infty \leq a < b < \infty$  durch

$$(a, b] = \{x \in \mathbb{K} \mid a < x \leq b\} \subseteq \mathbb{K} \quad (1.5.18)$$

und  $[a, b)$  für  $-\infty < a < b \leq \infty$  durch

$$[a, b) = \{x \in \mathbb{K} \mid a \leq x < b\} \subseteq \mathbb{K}. \quad (1.5.19)$$

Hier sind die Symbole  $\pm\infty$  ausschließlich als Notation zu verstehen, um Intervalle der Form  $(-\infty, 1]$  etc. zu charakterisieren, wobei eben  $-\infty < x < +\infty$  für alle  $x \in \mathbb{K}$  gelten soll. Man beachte, dass in allen Fällen die Intervalle immer nicht-leere Teilmengen sind. Die Zahl  $a$  beziehungsweise  $b$  heißt auch die *untere* beziehungsweise *obere Grenze* des Intervalls. Mit diesen Begriffen können wir die Dedekindschen Schnitte aus Beispiel 1.5.6 als

$$(-\infty, x] \cup (x, \infty) = \mathbb{K} \quad \text{und} \quad (-\infty, x) \cup [x, \infty) = \mathbb{K} \quad (1.5.20)$$

bezeichnen. Für Intervalle lässt sich einfach entscheiden, was Supremum und Infimum sind. Es gilt

$$\sup(a, b) = \sup(a, b] = \sup[a, b) = \sup[a, b] = b \quad (1.5.21)$$

und

$$\inf(a, b) = \inf(a, b] = \inf[a, b) = \inf[a, b] = a. \quad (1.5.22)$$

Man beachte, dass dies die Fälle der unbeschränkten Intervalle mit abdeckt, siehe auch Übung 1.8.17.

Für Intervalle  $(a, b)$ ,  $[a, b)$ ,  $(a, b]$ , oder  $[a, b]$  nennen wir die Differenz  $b - a$  auch die *Länge* des Intervalls, sofern  $a, b \in \mathbb{K}$ .

Sind  $a, b \in \mathbb{K}$  mit  $a \leq b$  vorgegeben und  $x \in [a, b]$ , so können wir jede Zahl  $y \in [a, b]$  als Approximation von  $x$  verstehen, wobei der Fehler, den wir dabei machen, durch die Länge  $b - a$  des Intervalls abgeschätzt werden kann, siehe auch Abbildung 1.5. Je kleiner also diese Länge ist, desto besser ist die Approximation von  $x$  durch  $y$  geworden. Diese Idee liegt nun folgender Definition einer *Intervallschachtelung* zugrunde:



Abbildung 1.4: Bild mit typischen Intervallen

Abbildung 1.5: Ein Punkt  $x \in [a, b]$  wird durch  $y \in [a, b]$  bis auf die Länge  $b - a$  gut approximiert.

**Definition 1.5.8 (Intervallschachtelung)** Sei  $\mathbb{k}$  ein angeordneter Körper. Eine Intervallschachtelung ist eine Menge  $\{I_n\}_{n \in \mathbb{N}}$  von abgeschlossenen Intervallen  $I_n = [a_n, b_n]$  mit  $a_n, b_n \in \mathbb{k}$  derart, dass

$$I_{n+1} \subseteq I_n \quad (1.5.23)$$

für alle  $n \in \mathbb{N}$ .

Insbesondere gilt bei einer Intervallschachtelung für die Grenzen der Intervalle

$$a_1 \leq \cdots \leq a_n \leq a_{n+1} \leq \cdots \leq \cdots \leq b_{n+1} \leq b_n \leq \cdots \leq b_1, \quad (1.5.24)$$

womit man durch Induktion sofort

$$a_n \leq b_m \quad (1.5.25)$$

für alle  $n, m \in \mathbb{N}$  folgert, siehe auch Abbildung 1.6.

Wir kommen nun zur entscheidenden Formulierung der gewünschten Vollständigkeit: Die Existenz eines Supremums zu jeder nicht-leeren beschränkten Menge, die Existenz einer Trennungszahl zu jedem Dedekindschen Schnitt und ein Intervallschachtelungsprinzip sind alle gleichermaßen wünschenswert. Glücklicherweise sind alle drei Eigenschaften äquivalent, so dass wir nicht verschiedene Konzepte der Vollständigkeit erhalten. Darüber hinaus implizieren sie automatisch, dass der Körper archimedisch ist:

**Satz 1.5.9 (Vollständigkeit)** Sei  $\mathbb{k}$  ein angeordneter Körper. Dann sind folgende Aussagen äquivalent:

- i.) Der Körper  $\mathbb{k}$  ist supremumsvollständig: Jede nach oben beschränkte nicht-leere Teilmenge von  $\mathbb{k}$  besitzt ein Supremum.
- ii.) Der Körper  $\mathbb{k}$  ist schnittvollständig: Jeder Dedekindsche Schnitt von  $\mathbb{k}$  besitzt eine Trennungszahl.
- iii.) Der Körper  $\mathbb{k}$  erfüllt das Intervallschachtelungsprinzip: Zu jeder Intervallschachtelung  $\{I_n = [a_n, b_n]\}_{n \in \mathbb{N}}$  gibt es  $a, b \in \mathbb{k}$  mit  $a \leq b$  und

$$\bigcap_{n=1}^{\infty} I_n = [a, b]. \quad (1.5.26)$$

In diesem Fall ist  $\mathbb{k}$  notwendigerweise archimedisch.

BEWEIS: Wir zeigen  $i.) \implies ii.) \implies iii.) \implies i.)$ . Sei zunächst  $\mathbb{k}$  supremumsvollständig. Sei weiter  $A \subseteq \mathbb{k}$  eine nach unten unbeschränkte Teilmenge ungleich  $\mathbb{k}$  mit  $a < b$  für alle  $a \in A$  und

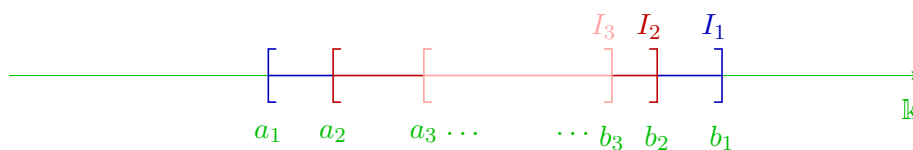


Abbildung 1.6: Eine Intervallschachtelung

$b \in B = \mathbb{k} \setminus A$ , also ein Dedekindscher Schnitt von  $\mathbb{k}$ . Da  $A$  nicht leer ist und obere Schranken besitzt, nämlich die Zahlen aus  $B$ , besitzt  $A$  ein Supremum  $x = \sup A$ . Dies ist die gewünschte Trennungszahl. Es gilt zunächst  $a \leq x$  für alle  $a \in A$ , da  $\sup A$  eine obere Schranke ist. Da  $x$  aber sogar die kleinste obere Schranke ist, gilt für alle  $b \in B$  die Ungleichung  $x \leq b$ , da  $b$  auch eine obere Schranke ist. Damit ist  $x$  aber eine (und sogar die) Trennungszahl des Dedekindschen Schnittes  $(A, B)$ .

Wir nehmen nun *ii.*) an und betrachten eine Intervallschachtelung  $\{I_n = [a_n, b_n]\}_{n \in \mathbb{N}}$  mit  $a_n \leq b_n$ . Dann definieren wir

$$A = \bigcup_{n=1}^{\infty} (-\infty, a_n).$$

Die Teilmenge  $A$  ist nicht leer, da beispielsweise  $a_1 - 1 \in (-\infty, a_1)$  gilt. Weiter ist  $\mathbb{k} \setminus A$  nicht leer, da beispielsweise  $b_1 \notin (-\infty, b_1)$  gilt, aber

$$(-\infty, a_n) \subseteq (-\infty, a_{n+1}) \subseteq (-\infty, b_n) \subseteq (-\infty, b_1)$$

zeigt, dass  $b_1 \notin A$ . Das Komplement von  $A$  lässt sich auch als

$$\mathbb{k} \setminus A = \mathbb{k} \setminus \bigcup_{n=1}^{\infty} (-\infty, a_n) = \bigcap_{n=1}^{\infty} \mathbb{k} \setminus (-\infty, a_n) = \bigcap_{n=1}^{\infty} [a_n, \infty) \quad (*)$$

schreiben. Daher gilt  $y \in \mathbb{k} \setminus A$  genau dann, wenn  $y \geq a_n$  für alle  $n$ . Für  $x \in A$  gilt dagegen, dass es ein  $n \in \mathbb{N}$  mit  $x < a_n$  gibt. Es folgt  $x < y$ , womit wir einen Dedekindschen Schnitt vorliegen haben. Nach Voraussetzung *ii.*) besitzt dieser eine Trennungszahl  $a \in \mathbb{k}$ , die entsprechend

$$x \leq a \leq y$$

für alle  $x \in A$  und  $y \in \mathbb{k} \setminus A$  erfüllt. Analog betrachten wir die Teilmenge

$$B = \bigcup_{n=1}^{\infty} (b_n, +\infty), \quad (**)$$

mit

$$\mathbb{k} \setminus B = \bigcap_{n=1}^{\infty} (-\infty, b_n]$$

von der man genauso zeigt, dass sie einen Dedekindschen Schnitt definiert, der dann eine Trennungszahl  $b \in \mathbb{k}$  besitzt. Diese ist durch

$$x \leq b \leq y$$

für alle  $y \in B$  und  $x \in \mathbb{k} \setminus B$  charakterisiert. Wir behaupten

$$I = \bigcap_{n=1}^{\infty} [a_n, b_n] = [a, b].$$

Nach (\*) gilt  $b_n \in \mathbb{k} \setminus A$  und daher  $a \leq b_n$ . Ebenso erhalten wir  $a_n \in \mathbb{k} \setminus B$  nach (\*\*) und somit  $a_n \leq b$ , jeweils für alle  $n \in \mathbb{N}$ . Etwas trickreicher ist die Eigenschaft, dass  $a_n \leq a$  für alle  $n$  gilt. Wir wissen, dass  $a_n - \varepsilon \in (-\infty, a_n) \subseteq A$  und entsprechend  $a_n - \varepsilon \leq a$  für alle  $\varepsilon > 0$ . Wäre nun  $a < a_n$ , so erreicht man für  $\varepsilon = \frac{a_n - a}{2} > 0$  den Widerspruch

$$a_n - \varepsilon = a_n - \frac{a_n - a}{2} > \frac{a_n}{2} + \frac{a}{2} > a$$

zu  $a_n - \varepsilon \leq a$  für alle  $\varepsilon > 0$ . Also folgt  $a_n \leq a$  für alle  $n$ . Insgesamt erhalten wir also für alle  $n$

$$a_n \leq a \leq b_n,$$

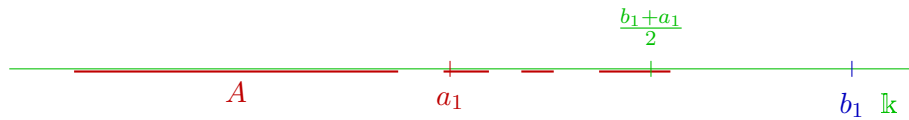


Abbildung 1.7: Halbierung des Intervalls liefert entweder eine neue obere Schranke, oder wir behalten die vorherige.

und damit  $a \in \bigcap_{n=1}^{\infty} [a_n, b_n]$ . Genauso zeigt man  $a_n \leq b \leq b_n$  und folglich  $b \in \bigcap_{n=1}^{\infty} [a_n, b_n]$ . Ist nun  $a' < a$ , so folgt aus der definierenden Eigenschaft eines Dedekindschen Schnittes, dass  $a' \in A$  gilt. Damit gibt es also ein  $n$  mit  $a' \in (-\infty, a_n)$  und folglich  $a' \notin \bigcap_{n=1}^{\infty} [a_n, b_n]$ . Wir sehen also, dass  $a$  die kleinste Zahl ist, die im Schnitt  $I$  enthalten ist. Genauso zeigt man, dass  $b$  die größte Zahl ist, die in  $I$  enthalten ist, was dann insbesondere  $a \leq b$  impliziert und  $I = [a, b]$  zur Folge hat. Damit ist auch *ii.)*  $\implies$  *iii.)* gezeigt.

Es gelte nun also *iii.)*. Wir zeigen zunächst, dass  $\mathbb{k}$  notwendigerweise archimedisch ist. Wir nehmen das Gegenteil an. Dann gibt es also eine Zahl  $b \in \mathbb{k}$  mit  $n \leq b$  für alle  $n \in \mathbb{N}$ . Wir betrachten die Intervalle  $I_n = [n, b]$ , welche eine Intervallschachtelung bilden. Es gibt also ein  $a \leq b$  mit

$$\bigcap_{n=1}^{\infty} I_n = \bigcap_{n=1}^{\infty} [n, b] = [a, b].$$

Man beachte, dass die konstante obere Grenze  $b$  auch die obere Grenze beim Durchschnitt bleibt. Für dieses  $a$  gilt also  $n \leq a$  für alle  $n \in \mathbb{N}$ , aber auch  $a - \varepsilon \notin \bigcap_{n=1}^{\infty} [n, b]$  für alle  $\varepsilon > 0$ , da  $a$  das kleinste Element des Durchschnitts ist. Es gibt also ein  $n$  mit  $a - 1 < n$  und damit  $a < n + 1$ , was einen Widerspruch liefert. Also ist  $\mathbb{k}$  notwendigerweise archimedisch.

Um nun *i.)* zu zeigen, betrachten wir eine nicht-leere nach oben beschränkte Teilmenge  $A \subseteq \mathbb{k}$ , für die wir ein Supremum konstruieren müssen. Dazu konstruieren wir induktiv eine Intervallschachtelung  $\{I_n = [a_n, b_n]\}_{n \in \mathbb{N}}$ . Wir wählen  $a_1 \in A$  und eine obere Schranke  $b_1$  von  $A$ , womit  $a_1 \leq b_1$  gilt, und wir ein Intervall  $I_1$  erhalten. Das zweite Intervall erhalten wir nun folgendermaßen. Entweder ist die Mitte zwischen  $a_1$  und  $b_1$  immer noch eine obere Schranke, dann setzen wir

$$a_2 = a_1 \quad \text{und} \quad b_2 = \frac{b_1 + a_1}{2}.$$

Ist dies dagegen nicht der Fall, behalten wir

$$b_2 = b_1$$

bei und wählen ein  $a_2 \in A$  mit der Eigenschaft

$$\frac{b_1 + a_1}{2} \leq a_2,$$

was in diesem Fall möglich ist. Auf diese Weise fahren wir induktiv fort und erhalten Intervalle  $I_n = [a_n, b_n]$  mit  $a_n \in A$  und obere Schranken  $b_n$  von  $A$  derart, dass die Intervalllänge von  $I_n$  höchstens  $\frac{1}{2^{n-1}}(b_1 - a_1)$  ist: In jedem Schritt wird die Länge mindestens halbiert, siehe auch Abbildung 1.7. Nach Voraussetzung *iii.)* gibt es also  $a \leq b$  mit

$$[a, b] = \bigcap_{n=1}^{\infty} [a_n, b_n].$$

Insbesondere gilt  $[a, b] \subseteq [a_n, b_n]$  für alle  $n$  und daher

$$b - a \leq b_n - a_n \leq \frac{1}{2^{n-1}}(b_1 - a_1).$$

Jetzt verwenden wir, dass  $\mathbb{k}$  notwendigerweise archimedisch geordnet ist. Aus

$$2^{n-1}(b-a) \leq b_1 - a_1$$

für alle  $n$  folgt  $b = a$  nach Proposition 1.4.2, *ii.*). Das Intervall  $[a, b]$  besteht also aus einem einzigen Punkt  $a = b$ . Wir behaupten, dass dies das gesuchte Supremum  $b$  ist. Zunächst ist  $b$  eine obere Schranke: Ist nämlich  $x \in A$ , so gilt für alle  $n \in \mathbb{N}$  die Ungleichung  $x \leq b_n$ , da  $b_n$  eine obere Schranke ist. Also folgt

$$b_n - b + b - x = b_n - x \geq 0,$$

und andererseits  $b_n - b \leq \frac{b_1 - a_1}{2^{n-1}}$ . Insgesamt liefert dies die Ungleichung

$$b - x \geq -\frac{b_1 - a_1}{2^{n-1}}$$

für alle  $n \in \mathbb{N}$ . Mit der bereits gezeigten Eigenschaft von  $\mathbb{k}$ , archimedisch zu sein, folgt  $b - x \geq 0$ , womit  $b$  eine obere Schranke ist. Sei nun  $b' \leq b$  eine weitere, kleinere obere Schranke. Wir müssen  $b' = b$  zeigen. Für alle  $n \in \mathbb{N}$  gilt

$$0 \leq b - b' = b - a_n + a_n - b' \leq \frac{b_1 - a_1}{2^{n-1}} + a_n - b' \leq \frac{b_1 - a_1}{2^{n-1}},$$

da  $a_n - b' \leq 0$ . Im archimedischen Körper  $\mathbb{k}$  ist dies nur für  $b - b' = 0$  möglich.  $\square$

Mit dieser gezeigten Äquivalenz können wir nun unsere Wünsche zur Vollständigkeit konkretisieren.

**Definition 1.5.10 (Vollständigkeit)** *Ein archimedisch angeordneter Körper heißt vollständig, wenn er eine der (und damit alle) drei Eigenschaften aus Satz 1.5.9 erfüllt.*

Die reellen Zahlen sind nun ein derartiger Körper. Wir wollen ihre Existenz und Eindeutigkeit an dieser Stelle nicht weiter diskutieren, sondern fordern diese axiomatisch. Dies führt zu folgender Definition:

**Definition 1.5.11 (Reelle Zahlen)** *Die reellen Zahlen sind ein archimedisch angeordneter vollständiger Körper, bezeichnet mit  $\mathbb{R}$ .*

### Bemerkung 1.5.12

- i.)* Man kann nun zeigen, dass je zwei archimedisch angeordnete vollständige Körper auf eindeutige Weise, also nur unter Benutzung der Körpereigenschaften, zueinander isomorph sind. In diesem Sinne gibt es also höchstens eine (bis auf Isomorphie eindeutige) Version der reellen Zahlen. Aus diesem Grund sprechen wir im Folgenden von *den* reellen Zahlen.
- ii.)* Die Existenz wird durch unseren Zugang nicht geklärt. Je nach Wahl des Ausgangspunktes kann man die reellen Zahlen aus beispielsweise den natürlichen Zahlen oder eben aus der axiomatischen Mengenlehre tatsächlich konstruieren.
- iii.)* Da  $\mathbb{R}$  archimedisch ist, gibt es zu jeder reellen Zahl  $x \in \mathbb{R}$  ganze Zahlen  $r, s \in \mathbb{Z}$  mit  $r < x < s$ . Verwendet man diese als Startwerte  $a_1 = r$  und  $b_1 = s$  für eine Intervallschachtelung durch Intervallhalbierung, so erhält man Intervalle  $[a_n, b_n]$  mit *rationalen* Intervallgrenzen  $a_n, b_n \in \mathbb{Q}$  und

$$\{x\} = \bigcap_{n=1}^{\infty} [a_n, b_n]. \quad (1.5.27)$$

Insbesondere gibt es zu jedem  $n \in \mathbb{N}$  rationale Zahlen  $q, p \in \mathbb{Q}$  mit

$$q < x < p \quad \text{und} \quad p - q \leq \frac{1}{n}. \quad (1.5.28)$$

Wir können reelle Zahlen also durch rationale Zahlen beliebig gut approximieren.

iv.) Mit der Äquivalenz von Supremumsvollständigkeit, Schnittvollständigkeit und Intervallschachtelungsprinzip haben wir bereits drei sehr effiziente Methoden gefunden, um über Vollständigkeit zu verfügen. Insbesondere das Intervallschachtelungsprinzip lässt sich als universellen Algorithmus verstehen, um Suprema von beschränkten Teilmengen *approximativ* zu finden: Wie im Beweis von Satz 1.5.9 erhält man zum gesuchten Supremum  $\sup A$  der Teilmenge  $A \subseteq \mathbb{R}$  Zahlen  $a_n \in A$  und  $b_n \in \mathbb{R}$  mit

$$a_n \leq \sup A \leq b_n \quad (1.5.29)$$

und  $b_n - a_n \leq \frac{b_1 - a_1}{2^{n-1}}$  für alle  $n \in \mathbb{N}$ . Daher kann man  $a_n$  oder  $b_n$  als Approximation von  $\sup A$  verstehen, wobei der Fehler eben höchstens  $\frac{b_1 - a_1}{2^{n-1}}$  ist, sich also in jedem Schritt halbiert. Sobald wir eine Fragestellung als Frage nach einem Supremum formulieren können, erhalten wir daher einen passenden Approximationsalgorithmus. Es wird in spezielleren Situationen natürlich vorteilhafter sein, andere, der tatsächlichen Fragestellung besser angepasste Approximationsverfahren zu verwenden. Die Intervallhalbierung hat jedoch den großen Vorteil, konzeptuell sehr einfach zu sein und immer zu funktionieren.

Wir schließen nun diesen Abschnitt mit einigen nützlichen Rechenregeln zu Suprema und Infima:

**Proposition 1.5.13 (Rechenregeln für  $\sup$  und  $\inf$ )** Seien  $A, B \subseteq \mathbb{R}$  nicht-leere Teilmengen von reellen Zahlen.

i.) Falls  $A \subseteq B$ , so gilt  $\sup A \leq \sup B$ .

ii.) Es gilt  $\inf A = -\sup(-A)$ , wobei

$$-A = \{-x \mid x \in A\}. \quad (1.5.30)$$

iii.) Es gilt  $\sup(A + B) = \sup A + \sup B$ , wobei

$$A + B = \{x + y \mid x \in A \text{ und } y \in B\}. \quad (1.5.31)$$

iv.) Sei zudem  $A, B \subseteq \mathbb{R}_0^+$ . Es gilt  $\sup(A \cdot B) = \sup A \cdot \sup B$ , wobei

$$A \cdot B = \{xy \mid x \in A \text{ und } y \in B\}. \quad (1.5.32)$$

v.) Gilt  $A \subseteq \mathbb{R}^+$  mit  $\inf A > 0$ , so gilt  $\sup A^{-1} = \frac{1}{\inf A}$ , wobei

$$A^{-1} = \{x^{-1} \mid x \in A\}. \quad (1.5.33)$$

BEWEIS: Die erste Aussage ist klar, da jede obere Schranke von  $B$  erst recht eine obere Schranke von  $A$  ist. Für den zweiten Teil betrachten wir eine untere Schranke  $a$  von  $A$ . Es gilt also  $a \leq x$  für alle  $x \in A$ , was  $-x \leq -a$  bedeutet. Daher ist  $-a$  eine obere Schranke von  $-A$  und  $\sup(-A) \leq -a$  folgt für alle unteren Schranken von  $A$ . Damit gilt  $-\sup(-A) \geq a \geq \inf A$ . Sei entsprechend  $a$  eine obere Schranke von  $A$ , also  $x \leq a$  für alle  $x \in A$ . Dann gilt  $-a \leq -x$ , womit  $-a$  eine untere Schranke von  $-A$  ist. Daher gilt  $-a \leq \inf(-A)$  für alle oberen Schranken  $a$  von  $A$ , also auch  $-\sup A \leq \inf(-A)$ . Ersetzen wir nun  $A$  durch  $-A$ , so erhalten wir  $-\sup(-A) \leq \inf(-(-A)) = \inf A$ , was die verbleibende Ungleichung ist, um ii.) zu folgern. Wir betrachten nun obere Schranken  $a$  von  $A$  und  $b$  von  $B$ . Dann gilt  $x + y \leq a + b$  für alle  $x \in A$  und  $y \in B$ , was zeigt, dass  $\sup A + \sup B$  eine obere Schranke von  $A + B$  ist. Es folgt  $\sup(A + B) \leq \sup A + \sup B$ , da  $\sup(A + B)$  die kleinste obere Schranke von  $A + B$  ist. Ist nun  $c < \sup A + \sup B$ , so gibt es  $a, b \in \mathbb{R}$  mit  $c = a + b$  und  $a < \sup A$  sowie  $b < \sup B$ . Damit finden wir  $x \in A$  und  $y \in B$  mit  $a < x$  und  $b < y$ , nach der Definition der Suprema. Es folgt  $c < x + y$ , womit  $c$  keine obere Schranke von  $A + B$  ist. Also gilt sogar die Gleichheit in iii.). Als nächstes betrachten wir nicht-leere Teilmengen  $A, B \in \mathbb{R}_0^+$  von nicht-negativen reellen Zahlen. Die Argumentation ist analog zu iii.). Sind  $a$  und  $b$  wieder obere Schranken von  $A$  beziehungsweise  $B$ , so

gilt  $x \leq a$  und  $y \leq b$  für alle  $x \in A$  und  $y \in B$ . Da alle Zahlen nicht-negativ sind, folgt  $xy \leq ab$ , womit  $ab$  eine obere Schranke von  $A \cdot B$  ist. Es folgt  $\sup(A \cdot B) \leq \sup A \cdot \sup B$  wie zuvor. Ist umgekehrt  $c < \sup A \cdot \sup B$ , so gibt es wieder  $a, b \in \mathbb{R}$  mit  $c = ab$  und  $a < \sup A$  sowie  $b < \sup B$ , siehe auch Übung 1.8.21. Es gibt also  $x \in A$  mit  $a < x$  und  $y \in B$  mit  $b < y$ , so dass  $c = ab < xy$ . Damit ist  $c$  keine obere Schranke von  $A \cdot B$ , und *iv.*) ist ebenfalls gezeigt. Für den letzten Teil betrachten wir  $A \subseteq \mathbb{R}^+$ . Ist  $a$  eine untere Schranke, so können wir  $a > 0$  annehmen, da  $\inf A > 0$ . Es gilt dann  $a \leq x$  für alle  $x \in A$  und damit  $\frac{1}{x} \leq \frac{1}{a}$ , womit  $\frac{1}{a}$  eine obere Schranke für  $A^{-1}$  wird. Es folgt  $\sup(A^{-1}) \leq \frac{1}{\inf A}$ . Umgekehrt gilt für eine obere Schranke  $a$  von  $A$  für jedes  $x \in A$  die Ungleichung  $x \leq a$  und damit  $\frac{1}{a} \leq \frac{1}{x}$ . Also ist  $\frac{1}{a}$  eine untere Schranke von  $A^{-1}$ , womit  $\frac{1}{\sup A} \leq \inf(A^{-1})$  folgt. Ersetzen wir wieder  $A$  durch  $A^{-1}$ , so erhalten wir  $\frac{1}{\sup(A^{-1})} \leq \inf((A^{-1})^{-1}) = \inf A$ . Dies liefert  $\sup(A^{-1}) \geq \frac{1}{\inf A}$ , was die verbleibende Ungleichung zeigt, um den Beweis von *iv.*) abzuschließen.  $\square$

Man beachte, dass die Voraussetzungen  $A, B \subseteq \mathbb{R}_0^+$  in *iv.*) und  $A \subseteq \mathbb{R}^+$  mit  $\inf A > 0$  in *v.*) im allgemeinen nicht aufgeweicht werden können, siehe auch Übung 1.8.22.

Weiter beachte man, dass die Fälle von unbeschränkten Teilmengen  $A$  und  $B$  eingeschlossen sind, sofern man die Rechenoperationen für  $\sup A = \infty$  etc. richtig interpretiert: Wir verwenden die Verabredungen

$$a + \infty = \infty \quad (1.5.34)$$

für alle  $a \in \mathbb{R} \cup \{+\infty\}$ , sowie

$$a \cdot \infty = \infty \quad \text{und} \quad b \cdot \infty = -\infty \quad (1.5.35)$$

für  $a > 0$  sowie  $b < 0$ . Schließlich ist es nützlich,

$$\frac{1}{\infty} = 0 = \frac{1}{-\infty} \quad (1.5.36)$$

als letzte Konvention zu verwenden. Man beachte aber, dass die Rechenoperationen  $+$  und  $\cdot$  eines Körpers *nicht* uneingeschränkt definiert werden: Eine Kombination wie „ $\infty - \infty$ “ oder „ $0 \cdot \infty$ “ wird ausdrücklich *nicht* erklärt. Insbesondere ist  $\mathbb{R} \cup \{\pm\infty\}$  kein Körper mehr. Trotzdem sind die obigen Verabredungen nützlich, um Fallunterscheidungen bei der Diskussion von Suprema und Infima zu vermeiden, siehe auch Übung 1.8.24.

**Bemerkung 1.5.14 (Infimumsvollständigkeit)** Der zweite Teil von Proposition 1.5.13 gestattet es, alle Aussagen über Suprema durch Aussagen über Infima zu ersetzen, siehe auch Übung 1.8.23. Insbesondere können wir eine weitere äquivalente Formulierung der Vollständigkeit in die Liste von Satz 1.5.9 mit aufnehmen: Ein geordneter Körper  $\mathbb{k}$  ist genau dann vollständig, wenn er *infimumsvollständig* ist, es also zu jeder nach unten beschränkten nicht-leeren Teilmenge ein Infimum in  $\mathbb{k}$  gibt.

Als vorerst letzte kleine Rechenregeln wollen wir das Verhalten von Suprema unter Iteration betrachten:

**Proposition 1.5.15** Sei  $A_i \subseteq \mathbb{R}$  eine nicht-leere nach oben beschränkte Teilmenge für alle  $i \in I$  mit einer Indexmenge  $I$ . Dann gilt

$$\sup\{\sup A_i \mid i \in I\} = \sup \bigcup_{i \in I} A_i \quad (1.5.37)$$

BEWEIS: Sei zunächst  $\bigcup_{i \in I} A_i \subseteq \mathbb{R}$  nach oben unbeschränkt, so dass also das Supremum auf der rechten Seite  $+\infty$  ist. In diesem Fall gibt es daher zu jedem  $n \in \mathbb{N}$  einen Index  $i_n$  und eine Zahl  $x_n \in A_{i_n}$  mit  $x_n \geq n$ . Damit ist  $\sup A_i \geq n$ , womit die Menge der Suprema  $\sup A_i$  unbeschränkt ist. Ist umgekehrt die linke Seite  $+\infty$ , so ist die Menge der Suprema  $\sup A_i$  unbeschränkt. Also finden

wir zu  $n \in \mathbb{N}$  ein  $i_n \in I$  mit  $\sup A_{i_n} \geq n$ . Damit enthält  $A_{i_n}$  aber Elemente  $x_n \geq n - 1$ , womit die Vereinigung  $\bigcup_{i \in I} A_i$  unbeschränkt ist und die rechte Seite ebenfalls  $+\infty$  liefert. Es bleibt also der Fall, dass beide Seiten endlich sind. Sei zuerst  $x \in \mathbb{R}$  eine obere Schranke von  $\bigcup_{i \in I} A_i$ . Dann gilt also  $x \geq y$  für alle  $y \in A_i$  und alle  $i \in I$ . Es folgt  $x \geq \sup A_i$  für alle  $i \in I$  und entsprechend  $x \geq \sup\{\sup A_i \mid i \in I\}$ . Da dies für alle oberen Schranken  $x$  gilt, folgt „ $\leq$ “ in (1.5.37). Sei umgekehrt  $y \in \mathbb{R}$  eine obere Schranke der Suprema  $\sup A_i$ . Dann gilt  $y \geq \sup A_i \geq x$  für alle  $x \in A_i$  und  $i \in I$  womit  $y$  eine obere Schranke der rechten Seite ist. Folglich gilt  $y \geq \sup \bigcup_{i \in I} A_i$ . Dies gilt wieder für alle oberen Schranken  $y$ , was „ $\geq$ “ in (1.5.37) zeigt.  $\square$

## 1.6 Wurzeln und Potenzgesetze

Als erste Anwendung der Vollständigkeit wollen wir nun eines der großen Defizite der rationalen Zahlen  $\mathbb{Q}$  beseitigen, indem wir zu den reellen Zahlen übergehen. In  $\mathbb{R}$  haben alle positiven Zahlen eine (sogar eindeutige) positive Wurzel. Geringfügig allgemeiner aber mit der selben Idee wie in Beispiel 1.5.2 zeigen wir, dass folgende Teilmengen Dedekindsche Schnitte bilden.

**Lemma 1.6.1** *Sei  $a \in \mathbb{R}_0^+$  eine nicht-negative reelle Zahl. Für alle  $k \in \mathbb{N}$  bildet die Teilmenge*

$$A = \left\{ x \in \mathbb{R} \mid x \geq 0 \text{ und } x^k \leq a \right\} \cup \mathbb{R}^- \quad (1.6.1)$$

*einen Dedekindschen Schnitt.*

BEWEIS: Zunächst ist  $A$  nicht leer und nach unten unbeschränkt. Für das Komplement von  $A$  gilt

$$\mathbb{R} \setminus A = \left\{ y \in \mathbb{R} \mid y \geq 0 \text{ und } y^k > a \right\}. \quad (*)$$

In der Tat, es gilt  $\mathbb{R}^- \subseteq A$ , womit  $\mathbb{R} \setminus A \subseteq \mathbb{R}_0^+$  gelten muss. Die verbleibende Bedingung  $y^k > a$  folgt dann aus  $x^k \leq a$  für alle  $x \in A \setminus \mathbb{R}^-$ , womit  $(*)$  gezeigt ist. Ist nun  $x \in A$  und  $y \in \mathbb{R} \setminus A$ , so gilt  $x < y$ , denn entweder ist  $x < 0$  sogar negativ, womit  $x < y$  klar ist, oder  $x \geq 0$  und  $x^k \leq a$  während  $y^k > a$ . Wäre nun  $x \geq y$ , so gälte auch  $x^k \geq y^k > a$ , ein Widerspruch. Damit sind aber alle Eigenschaften eines Dedekindschen Schnittes gezeigt.  $\square$

Als Dedekindscher Schnitt definiert  $A$  daher eine reelle Zahl über seine Trennungszahl. Diese stimmt allgemein mit  $\sup A$  überein, was auch mit  $\inf(\mathbb{R} \setminus A)$  übereinstimmt, siehe Übung 1.8.26.

**Definition 1.6.2 ( $k$ -te Wurzel)** *Sei  $a \in \mathbb{R}_0^+$  eine nicht-negative Zahl. Dann heißt*

$$\sqrt[k]{a} = \sup \left\{ x \mid x \geq 0 \text{ und } x^k \leq a \right\} \quad (1.6.2)$$

*die positive  $k$ -te Wurzel von  $a$ .*

Der Fall  $k = 2$  tritt besonders oft auf. Aus diesem Grunde schreibt man auch einfach  $\sqrt{a} = \sqrt[2]{a}$  für die *Quadratwurzel* von  $a$ .

**Proposition 1.6.3** *Seien  $a, b \in \mathbb{R}_0^+$  und  $k, \ell \in \mathbb{N}$ .*

- i.) Es gilt  $(\sqrt[k]{a})^k = a$ .*
- ii.) Ist  $x \in \mathbb{R}_0^+$  mit  $x^k = a$ , so gilt  $x = \sqrt[k]{a}$ .*
- iii.) Es gilt  $\sqrt[k]{a} \cdot \sqrt[k]{b} = \sqrt[k]{ab}$ .*
- iv.) Es gilt  $(\sqrt[k]{a})^\ell = \sqrt[k]{a^\ell}$ .*
- v.) Es gilt genau dann  $a < b$ , wenn  $\sqrt[k]{a} < \sqrt[k]{b}$  gilt.*

vi.) Ist  $a > 0$ , so gilt  $\sqrt[k]{\frac{1}{a}} = \frac{1}{\sqrt[k]{a}}$ .

BEWEIS: Nach Definition gilt  $(\sqrt[k]{a})^k \leq a$ . Da bei einem Dedekindschen Schnitt immer

$$\sup A = \inf(\mathbb{R} \setminus A)$$

gilt, folgt  $\sqrt[k]{a} = \inf\{x \in \mathbb{R} \mid x^k > a\}$  und damit  $(\sqrt[k]{x})^k \geq a$ , was i.) zeigt. Jede Zahl  $x \in \mathbb{R}_0^+$  mit  $x^k = a$  ist eine und damit die eindeutige Trennungszahl des Dedekindschen Schnittes  $A$ , den wir zur Definition der  $k$ -ten Wurzel verwendet haben. Dies zeigt also  $x = \sqrt[k]{a}$ . Für den dritten Teil berechnen wir  $(\sqrt[k]{a}\sqrt[k]{b})^k = (\sqrt[k]{a})^k(\sqrt[k]{b})^k = ab$ . Da  $\sqrt[k]{a}\sqrt[k]{b} \geq 0$  nicht negativ ist, ist nach der Eindeutigkeitsaussage ii.) der dritte Teil gezeigt. Für den vierten Teil berechnen wir  $((\sqrt[k]{a})^\ell)^k = ((\sqrt[k]{a})^k)^\ell = a^\ell$ . Die Eindeutigkeit aus ii.) zeigt also auch in diesem Fall die Behauptung. Wenn  $\sqrt[k]{a} < \sqrt[k]{b}$ , so folgt  $a = (\sqrt[k]{a})^k < (\sqrt[k]{b})^k = b$  wie in jedem angeordneten Körper, siehe Proposition 1.3.5, vii.). Es gilt aber auch die Umkehrung: Gilt  $a < b$  und  $\sqrt[k]{b} \leq \sqrt[k]{a}$ , so folgt  $b = (\sqrt[k]{b})^k \leq (\sqrt[k]{a})^k = a$ , ein Widerspruch. Den letzten Teil erhält man nun wieder, indem man die  $k$ -te Potenz beider Seiten berechnet und in beiden Fällen  $\frac{1}{a}$  erhält.  $\square$

Damit haben wir also die wichtigsten Rechenregeln der  $k$ -ten Wurzeln etabliert. Der entscheidende Punkt ist, dass wir überhaupt von  $k$ -ten Wurzeln sprechen können. Dies ist Dank der Vollständigkeit von  $\mathbb{R}$  möglich. Die Rechenregeln legen nun folgende Definition einer allgemeinen rationalen Potenz nahe:

**Definition 1.6.4 (Rationale Potenzen)** Sei  $a \in \mathbb{R}^+$  und  $q \in \mathbb{Q}$  mit  $q = \frac{r}{s}$  und  $r \in \mathbb{Z}$ ,  $s \in \mathbb{N}$ . Dann definiert man die  $q$ -te Potenz von  $a$  als

$$a^q = \sqrt[s]{a^r}. \quad (1.6.3)$$

Die Zahl  $q$  heißt Exponent, die Zahl  $a$  die Basis. Für  $a = 0$  setzen wir  $a^q = 0$  für alle  $q > 0$ .

Zunächst gilt es zu zeigen, dass  $a^q$  tatsächlich wohl-definiert ist. Die Darstellung  $q = \frac{r}{s}$  ist ja nicht eindeutig, da wir Brüche kürzen und erweitern können. Wir zeigen die Unabhängigkeit von der Darstellung sowie weitere Rechenregeln in der folgenden Proposition:

**Proposition 1.6.5** Seien  $a, b \in \mathbb{R}^+$ .

- i.) Für alle  $q \in \mathbb{Q}$  ist  $a^q$  gemäß (1.6.3) wohl-definiert.
- ii.) Für alle  $q, p \in \mathbb{Q}$  gilt  $a^q a^p = a^{q+p}$  und  $a^0 = 1$ .
- iii.) Für alle  $q, p \in \mathbb{Q}$  gilt  $(a^q)^p = a^{qp}$ .
- iv.) Für alle  $q \in \mathbb{Q}$  gilt  $a^q b^q = (ab)^q$ .
- v.) Es gilt  $a < b$  genau dann, wenn  $a^q < b^q$  für ein  $q \in \mathbb{Q}^+$ . In diesem Fall gilt  $a^q < b^q$  für alle  $q \in \mathbb{Q}^+$ .
- vi.) Es gilt  $\frac{1}{a^q} = a^{-q}$  für alle  $q \in \mathbb{Q}$ .
- vii.) Für  $a > 1$  und  $q, p \in \mathbb{Q}$  mit  $q < p$  gilt  $a^q < a^p$ .

BEWEIS: Sei  $q = \frac{r}{s}$  und  $k \in \mathbb{N}$ . Dann berechnen wir

$$\left(\sqrt[s]{a^{rk}}\right)^{sk} = a^{rk} = (a^r)^k$$

und daher

$$\left(\sqrt[s]{a^{rk}}\right)^s = a^r = (\sqrt[s]{a^r})^s,$$



womit  $\sqrt[s^k]{a^{rk}} = \sqrt[s]{a^r}$ . Hierbei haben wir zweimal die Eindeutigkeitsaussage von Proposition 1.6.3, *ii.*), verwendet. Da sich je zwei Darstellungen von  $q$  durch wechselseitiges Erweitern beziehungsweise Kürzen unterscheiden, folgt *i.*). Für die verbleibenden Rechenregeln müssen wir nur geschickt die Rechenregeln aus Proposition 1.6.3 verwenden. Zunächst gilt für  $q = \frac{r}{s}$  und  $p = \frac{u}{v}$

$$\begin{aligned} a^q a^p &\stackrel{i.)}{=} a^{\frac{rv}{sv}} a^{\frac{us}{vs}} \\ &\stackrel{(a)}{=} \sqrt[sv]{a^{rv} a^{us}} \\ &= \sqrt[sv]{a^{rv+us}} \\ &= a^{\frac{rv+us}{sv}} \\ &= a^{q+p}, \end{aligned}$$

wobei wir Proposition 1.6.3, *iii.*), in (a) und Proposition 1.2.7, *ii.*), für die ganzzahligen Exponenten verwendet haben. Für den dritten Teil bemerken wir zunächst, dass  $a^q > 0$  und deshalb  $(a^q)^p$  überhaupt definiert ist. Dann gilt mit  $q = \frac{r}{s}$  und  $p = \frac{u}{v}$

$$\begin{aligned} (a^q)^p &= \sqrt[v]{\left(\sqrt[s]{a^r}\right)^u} \\ &= \sqrt[v]{\left(\sqrt[s]{a}\right)^{ru}} \\ &= \sqrt[v]{\sqrt[s]{a^{ru}}}, \end{aligned}$$

wobei wir wiederholt Proposition 1.6.3, *iv.*), verwendet haben. Bilden wir nun die  $vs$ -te Potenz, so erhalten wir

$$\begin{aligned} ((a^q)^p)^{vs} &= \left( \left( \sqrt[v]{\sqrt[s]{a^{ru}}} \right)^v \right)^s \\ &= \left( \sqrt[s]{a^{ru}} \right)^s \\ &= a^{ru} \\ &= \left( a^{\frac{ru}{vs}} \right)^{vs} \\ &= (a^{pq})^{vs}, \end{aligned}$$

unter Verwendung von Proposition 1.6.3, *i.*). Die Eindeutigkeit der  $vs$ -ten Wurzel liefert dann  $(a^q)^p = a^{qp}$ . Für *iv.*) schreiben wir wieder  $q = \frac{r}{s}$  und erhalten

$$a^q b^q = \sqrt[s]{a^r} \sqrt[s]{b^r} = \sqrt[s]{a^r b^r} = \sqrt[s]{(ab)^r} = (ab)^q,$$

wie gewünscht. Da  $a^0 = 1$ , folgt *v.*) direkt aus *iii.*). Sei schließlich  $a < b$ , dann gilt  $a^r < b^r$  und somit auch  $\sqrt[s]{a^r} < \sqrt[s]{b^r}$  für alle positiven  $r, s$ , also  $a^q < b^q$  für alle  $q \in \mathbb{Q}^+$ . Hier benutzen wir Proposition 1.6.3, *v.*). Gilt umgekehrt  $a^q < b^q$  für ein  $q = \frac{r}{s} \in \mathbb{Q}^+$ , so gilt also  $\sqrt[s]{a^r} < \sqrt[s]{b^r}$  und daher  $a^r < b^r$ , wieder nach Proposition 1.6.3, *v.*). Das impliziert aber  $a < b$ , wie in jedem angeordneten Körper. Für den letzten Teil betrachten wir  $\frac{a^p}{a^q} = a^p a^{-q} = a^{p-q}$  mit einer positiven rationalen Zahl  $x = p - q \in \mathbb{Q}^+$ , da  $p > q$ . Wir schreiben daher  $x = \frac{u}{v}$  mit  $u, v \in \mathbb{N}$  und erhalten

$$a^x = \sqrt[v]{a^u} > \sqrt[v]{1} = 1$$

nach Proposition 1.6.3, *v.*), und der Eigenschaft  $a^u > 1$ , da  $a > 1$  nach Voraussetzung. Damit ist aber  $a^p > a^q$  gezeigt.  $\square$

Man erhält ebenfalls Abschätzungen für  $q \in \mathbb{Q}^-$  in Teil vi.) oder  $a < 1$  in Teil vii.), indem man geeignet zu Quotienten übergeht, siehe auch Übung 1.8.27.

Es stellt sich nun natürlich die Frage, wieso man sich auf rationale Exponenten  $q \in \mathbb{Q}$  in  $a^q$  beschränken soll, beziehungsweise wie sich die Definition auf reelle Exponenten  $x \in \mathbb{R}$  ausdehnen lässt. Dies ist tatsächlich unter Beibehaltung der Rechenregeln möglich, da wir über die Vollständigkeit von  $\mathbb{R}$  verfügen können. Zunächst zeigen wir folgendes technische Lemma:

**Lemma 1.6.6** *Sei  $a > 1$  und  $x \in \mathbb{R}$ . Dann gilt*

$$\sup\{a^q \mid q \in \mathbb{Q} \text{ mit } q \leq x\} = \inf\{a^p \mid p \in \mathbb{Q} \text{ mit } x < p\}. \quad (1.6.4)$$

BEWEIS: Sind  $p, q \in \mathbb{Q}$  mit  $q \leq x < p$  gegeben, so folgt  $a^q < a^p$  nach Proposition 1.6.5, vii.). Damit erhalten wir also wechselseitige Schranken, welche die Existenz von Supremum und Infimum zeigen, da beide Mengen offenbar nicht leer sind. Es folgt insbesondere „ $\leq$ “ in (1.6.4). Wir bezeichnen den Wert des Supremums mit  $u$  und den Wert des Infimums mit  $v$ . Wir nehmen an, dass  $u < v$  gilt und wollen dies zum Widerspruch führen. Es gilt also  $1 < \frac{v}{u} = 1 + \alpha$  mit einem  $\alpha > 0$ . Für  $q, p \in \mathbb{Q}$  mit  $q \leq x < p$  gilt daher  $v \leq a^p$  sowie  $u \geq a^q$  und folglich

$$1 < \frac{v}{u} \leq \frac{a^p}{a^q} = a^{p-q}.$$

Für ein gegebenes  $n \in \mathbb{N}$  können wir dank Bemerkung 1.5.12, iii.), die Zahlen  $q$  und  $p$  sogar so wählen, dass  $p - q \leq \frac{1}{n}$  gilt. Dann folgt aus  $1 + \alpha \leq a^{p-q} \leq a^{\frac{1}{n}}$  also

$$(1 + \alpha)^n \leq a$$

für jedes vorgegebene  $n \in \mathbb{N}$ . Das ist aber im Widerspruch zur Bernoulli-Ungleichung  $(1 + \alpha)^n > 1 + n\alpha$ , sofern  $\alpha > 0$  gilt, da  $\mathbb{R}$  archimedisch angeordnet ist. Also muss  $v \leq u$  gelten, womit insgesamt die Gleichheit in (1.6.4) gezeigt ist.  $\square$

Für  $a = 1$  ist immer  $1^q = 1 = 1^p$ , womit die Gleichheit in (1.6.4) trivialerweise immer gilt. Ist schließlich  $0 < a < 1$ , so dreht sich die Ungleichung für  $q \leq p$  zu  $a^p \leq a^q$ . Es folgt somit die Gleichheit

$$\inf\{a^q \mid q \in \mathbb{Q} \text{ mit } q \leq x\} = \sup\{a^p \mid p \in \mathbb{Q} \text{ mit } x < p\}, \quad (1.6.5)$$

indem man  $a$  durch  $\frac{1}{a}$  in Lemma 1.6.6 ersetzt. Die Gleichheit bleibt daher bestehen, nur muss man eben Infimum und Supremum vertauschen.

Um allgemeine Potenzen von  $a$  zu konstruieren, ist die Idee nun, dass wir  $a^x$  für alle  $x \in \mathbb{R}$  durch das Supremum (1.6.4) definieren. Für  $x \in \mathbb{Q}$  reproduziert dies sicherlich die vorherige Definition, da ja  $a^{q'} < a^q$  für  $q' < q$  gilt, das Supremum also für  $a^q$  angenommen wird. Das Lemma zeigt schließlich, dass die ebenso plausible Wahl des Infimums die selbe Definition von  $a^x$  liefert:

**Definition 1.6.7 (Reelle Potenzen)** *Sei  $a \geq 1$  und  $x \in \mathbb{R}$ . Dann definieren wir*

$$a^x = \sup\{a^q \mid q \in \mathbb{Q} \text{ mit } q \leq x\}. \quad (1.6.6)$$

Für  $0 < a \leq 1$  definieren wir

$$a^x = \left(\frac{1}{a}\right)^{-x} \quad (1.6.7)$$

mit  $(\frac{1}{a})^{-x}$  gemäß (1.6.6).

**Satz 1.6.8 (Potenzgesetze)** *Seien  $a, b \in \mathbb{R}^+$ .*

- i.) Die Definition von  $a^q$  für  $q \in \mathbb{Q}$  gemäß (1.6.6) beziehungsweise (1.6.7) stimmt mit der vorherigen Definition aus (1.6.3) überein.
- ii.) Es gilt  $a^0 = 1$  sowie  $1^x = 1$  und  $a^x a^y = a^{x+y}$  für alle  $x, y \in \mathbb{R}$ .
- iii.) Es gilt  $(a^x)^y = a^{xy}$  für alle  $x, y \in \mathbb{R}$ .
- iv.) Es gilt  $a^x b^x = (ab)^x$  für alle  $x \in \mathbb{R}$ .
- v.) Es gilt  $\left(\frac{1}{a}\right)^x = \frac{1}{a^x} = a^{-x}$  für alle  $x \in \mathbb{R}$ .
- vi.) Es gilt genau dann  $a < b$ , wenn  $a^x < b^x$  für ein  $x \in \mathbb{R}^+$ . In diesem Fall gilt  $a^x < b^x$  für alle  $x \in \mathbb{R}^+$ .
- vii.) Für  $a > 1$  und  $x, y \in \mathbb{R}$  mit  $x < y$  gilt  $a^x < a^y$ .

BEWEIS: Den ersten Teil hatten wir bereits erklärt. Für den zweiten Teil betrachten wir zunächst  $a > 1$ , womit (1.6.6) als Definition von  $a^x$  zur Anwendung kommt. Da  $0 \in \mathbb{Q}$ , gilt  $a^0 = 1$ . Für  $x, y \in \mathbb{R}$  gilt

$$\begin{aligned} a^x a^y &= \sup\{a^q \mid q \in \mathbb{Q} \text{ mit } q \leq x\} \cdot \sup\{a^p \mid p \in \mathbb{Q} \text{ mit } p \leq y\} \\ &\stackrel{(a)}{=} \sup\{a^q a^p \mid q, p \in \mathbb{Q} \text{ mit } q \leq x \text{ und } p \leq y\} \\ &\stackrel{(b)}{=} \sup\{a^{q+p} \mid q, p \in \mathbb{Q} \text{ mit } q \leq x \text{ und } p \leq y\} \\ &\stackrel{(c)}{=} \sup\{a^r \mid r \in \mathbb{Q} \text{ mit } r \leq x + y\} \\ &= a^{x+y}, \end{aligned}$$

wobei wir in (a) die Produktregel für Suprema aus Proposition 1.5.13, iv.), und in (b) die Potenzrechenregeln für rationale Exponenten verwenden. In (c) benutzen wir, dass die Menge der Zahlen der Form  $a^{q+p}$  mit  $q, p \in \mathbb{Q}$  und  $q \leq x$ ,  $p \leq y$  mit der Menge der Zahlen der Form  $a^r$  mit  $r \in \mathbb{Q}$  und  $r \leq x + y$  schlichtweg übereinstimmt. Dies zeigt ii.) für  $a > 1$ . Für  $a < 1$  benutzen wir (1.6.7) als Definition der Potenzen und können die Rechenregel auf  $\frac{1}{a} > 1$  anwenden. Dies liefert

$$a^x a^y = \left(\frac{1}{a}\right)^{-x} \left(\frac{1}{a}\right)^{-y} = \left(\frac{1}{a}\right)^{-x-y} = a^{x+y}$$

wie gewünscht. Der Fall  $a = 1$  ist trivialerweise richtig, da  $1^x = 1$  für alle  $x \in \mathbb{R}$ . Dann folgt v.) aber direkt, denn

$$a^x a^{-x} = a^{x-x} = a^0 = 1,$$

womit  $a^{-x}$  das multiplikative Inverse zu  $a^x$  ist. Für Teil iv.) betrachten wir wieder  $a > 1$  und  $b > 1$ , so dass (1.6.6) die relevante Definition ist. Sind  $p, q \in \mathbb{Q}$  mit  $q \leq x$  und  $p \leq x$  gegeben, so gibt es ein  $r \in \mathbb{Q}$  mit  $q \leq r \leq x$  und  $p \leq r \leq x$ , beispielsweise  $r = \max\{q, p\}$ . Entsprechend gilt  $a^q b^p \leq a^r b^r$  für ein solches  $r$ . Diese Vorüberlegung nutzen wir um die Gleichheit von

$$\sup\{a^q b^p \mid q, p \in \mathbb{Q} \text{ mit } q \leq x \text{ und } p \leq x\} = \sup\{a^r b^r \mid r \in \mathbb{Q} \text{ mit } r \leq x\}$$

zu zeigen: Die Vorüberlegung zeigt „ $\leq$ “ während „ $\geq$ “ trivialerweise nach Proposition 1.5.13, i.), gilt. Damit erhalten wir

$$\begin{aligned} a^x b^x &= \sup\{a^q \mid q \in \mathbb{Q} \text{ mit } q \leq x\} \cdot \sup\{b^p \mid p \in \mathbb{Q} \text{ mit } p \leq x\} \\ &= \sup\{a^q b^p \mid q \in \mathbb{Q} \text{ mit } q \leq x \text{ und } p \leq x\} \\ &= \sup\{a^r b^r \mid r \in \mathbb{Q} \text{ mit } r \leq x\} \\ &= \sup\{(ab)^r \mid r \in \mathbb{Q} \text{ mit } r \leq x\} \\ &= (ab)^x, \end{aligned}$$

wobei wir die gewünschte Rechenregel für rationalen Exponenten nach Proposition 1.6.5, *iv.*), ja benutzen dürfen. Als nächstes betrachten wir  $a < 1$  und  $b > 1$ . Ist dann  $ab > 1$ , so gilt

$$a^{-x}(ab)^x = \left(\frac{1}{a}\right)^x (ab)^x = \left(\frac{1}{a}ab\right)^x = b^x$$

nach den bereits gezeigten Fällen. Also folgt  $(ab)^x = a^x b^x$ , da  $a^{-x} = \frac{1}{a^x}$  gemäß *v.*). Ist dagegen  $ab < 1$ , so gilt  $\frac{1}{ab} > 1$  und daher

$$a^{-x}(ab)^x = a^{-x} \left(\frac{1}{ab}\right)^{-x} = \left(a \frac{1}{ab}\right)^{-x} = \left(\frac{1}{b}\right)^{-x} = b^x$$

nach den bereits gezeigten Situationen. Also gilt wieder  $(ab)^x = a^x b^x$ . Sind schließlich  $a < 1$  und  $b < 1$ , so verwenden wir die gezeigte Rechenregel für  $\frac{1}{a}$  und  $\frac{1}{b}$ . Die verbleibenden Fälle, wo einer der Faktoren 1 ist, oder wo  $ab = 1$  gilt, sind trivial. Mit *iv.*) erhalten wir auch

$$\left(\frac{1}{a}\right)^x \cdot a^x = \left(\frac{1}{a}a\right)^x = 1^x = 1.$$

Dies ist die verbliebene Behauptung  $\left(\frac{1}{a}\right)^x = \frac{1}{a^x}$  in *v.*). Wir zeigen nun *iii.*). Hierzu behaupten wir, dass für eine nach oben beschränkte nicht-leere Teilmenge  $A \subseteq \mathbb{R}^+$  von reellen Zahlen und  $q \in \mathbb{Q}$

$$(\sup A)^q = \sup A^q \quad (*)$$

gilt, wobei

$$A^q = \{a^q \mid a \in A\}.$$

Um dies zu zeigen, betrachten wir zunächst  $q \in \mathbb{Z}$ . In diesem Fall folgt  $(*)$  aus Proposition 1.5.13, *iv.*) und *v.*). Sei  $k \in \mathbb{N}$ , dann gilt dem gezeigten Fall  $(*)$  für  $q = k$

$$\left(\sup \sqrt[k]{A}\right)^k \stackrel{(*)}{=} \sup \left(\sqrt[k]{A}\right)^k = \sup A,$$

womit  $\sup \sqrt[k]{A}$  die  $k$ -te Wurzel aus  $\sup A$  ist, also  $(*)$  für  $q = \frac{1}{k}$  gezeigt ist. Der allgemeine Fall folgt durch Kombination dieser beiden Spezialfälle. Wir betrachten nun  $a > 1$  und  $x \in \mathbb{R}$ , so dass  $a^x > 1$  ebenfalls gilt. Dann gilt für  $y \in \mathbb{R}$

$$\begin{aligned} (a^x)^y &= \sup \{(a^x)^p \mid p \in \mathbb{Q} \text{ und } p \leq y\} \\ &= \sup \left\{ \left( \sup \{a^q \mid q \in \mathbb{Q} \text{ und } q \leq x\} \right)^p \mid p \in \mathbb{Q} \text{ und } p \leq y \right\} \\ &\stackrel{(*)}{=} \sup \left\{ \sup \{(a^q)^p \mid q \in \mathbb{Q} \text{ und } q \leq x\} \mid p \in \mathbb{Q} \text{ und } p \leq y \right\} \\ &\stackrel{(a)}{=} \sup \{a^{qp} \mid p, q \in \mathbb{Q} \text{ und } q \leq x \text{ und } p \leq y\} \\ &= \sup \{a^r \mid r \in \mathbb{Q} \text{ und } r \leq xy\} \\ &= a^{xy}, \end{aligned}$$

wobei wir in *(a)* Proposition 1.5.15 verwenden. Ist nun  $a > 1$  aber  $a^x < 1$ , so gilt  $\frac{1}{a^x} = a^{-x} > 1$  und

$$(a^x)^y \stackrel{(1.6.7)}{=} \left(\frac{1}{a^x}\right)^{-y} \stackrel{v.)}{=} (a^{-x})^{-y} = a^{xy}$$

nach dem bereits gezeigten Fall. Ist  $a < 1$  und  $a^x > 1$ , so argumentiert man mit

$$(a^x)^y \stackrel{(1.6.7)}{=} \left(\left(\frac{1}{a}\right)^{-x}\right)^y = \left(\frac{1}{a}\right)^{-xy} = a^{xy},$$

unter Verwendung des bereits Gezeigten. Schließlich betrachten wir den Fall  $a < 1$  und  $a^x < 1$ . Dann gilt

$$(a^x)^y \stackrel{(1.6.7)}{=} \left(\frac{1}{a^x}\right)^{-y} = \left(\left(\frac{1}{a}\right)^x\right)^{-y} = \left(\frac{1}{a}\right)^{-xy} = a^{xy},$$

womit der letzte interessante Fall gezeigt ist. Gilt  $a = 1$  oder  $a^x = 1$  reduziert sich die Behauptung sofort auf die triviale Fragestellung  $(1^x)^y = 1 = 1^{xy}$ . Dies zeigt nun *iii.*). Seien nun  $a < b$  vorgegeben. Wir betrachten zunächst wieder den Fall  $a > 1$ . Dann gilt  $a^q < b^q$  für alle  $q \in \mathbb{Q}^+$  und daher

$$\begin{aligned} a^x &= \sup\{a^q \mid q \in \mathbb{Q} \text{ und } q \leq x\} \\ &\leq \sup\{b^q \mid q \in \mathbb{Q} \text{ und } q \leq x\} \\ &= b^x. \end{aligned}$$

Damit ist  $a^x \leq b^x$  gezeigt. Mit *iii.*) folgt aus  $a^x = b^x$  aber  $a = (a^x)^{\frac{1}{x}} = (b^x)^{\frac{1}{x}} = b$ . Daher gilt mit  $a < b$  auch  $a^x < b^x$ . Der nächste Fall ist  $a < 1 \leq b$ . In diesem Fall gilt  $a^x \leq 1$ , da  $a^q \leq 1$  für  $q \in \mathbb{Q}^+$ . Also gilt  $a^x \leq b^x$  auch hier und mit *iii.*) sehen wir erneut  $a^x < b^x$ . Der letzte Fall  $a < b < 1$  folgt aus dem ersten mit  $1 < \frac{1}{b} < \frac{1}{a}$  und *v.*). Gilt umgekehrt  $a^x < b^x$ , so können wir mit dem Gezeigten  $a = (a^x)^{\frac{1}{x}} < (b^x)^{\frac{1}{x}} = b$  erhalten, also  $a < b$ . Sei schließlich  $a > 1$  und  $x < y$ . Dann gilt  $1 < a^{y-x}$  nach *v.*) und folglich  $a^x = a^x \cdot 1 < a^x \cdot a^{y-x} = a^y$ .  $\square$

Mit diesem Satz verfügen wir nun über eine mächtige neue Rechenvorschrift, das Potenzieren, und zugehörige Rechenregeln, die das Zusammenspiel des Potenzierens mit den bisherigen Rechenregeln eines angeordneten Körpers klären. Wir werden später eine alternative Definition von  $a^x$  sehen, wenn wir die Exponentialabbildung kennen lernen. Damit werden wir die durchaus schwierige Definition der Potenzen mit allgemeinen reellen Exponenten besser begründen können.

## 1.7 Die komplexen Zahlen

Die Existenz von allgemeinen Potenzen von positiven reellen Zahlen erlaubt es uns nun, viele Gleichungen in  $\mathbb{R}$  zu lösen, die zuvor in  $\mathbb{Q}$  keine Lösung hatten. Es bleiben jedoch unlösbare Aufgaben algebraischer Natur wie beispielsweise die Gleichung

$$x^2 = -1. \tag{1.7.1}$$

Da  $-1 \in \mathbb{R}$  negativ ist, Quadrate in einem angeordneten Körper aber immer nicht-negativ sind, kann es kein  $x \in \mathbb{R}$  mit (1.7.1) geben. Historisch war dies einer der Gründe, die hypothetische Wurzel  $\sqrt{-1}$  in der Mathematik zu verwenden, eine Zahl, der man keine reelle Natur mehr zusprechen konnte, und die daher als *imaginär* bezeichnet wurde. Die vormals rätselhafte Natur dieser *imaginären Einheit*

$$i = \sqrt{-1} \tag{1.7.2}$$

lässt sich nun jedoch einfach aufklären und verstehen: Wir erweitern erneut den zugrundeliegenden Körper und wechseln von den reellen Zahlen  $\mathbb{R}$  zu den komplexen Zahlen  $\mathbb{C}$ .

**Definition 1.7.1 (Komplexe Zahlen)** *Die komplexen Zahlen sind als die Menge*

$$\mathbb{C} = \mathbb{R} \times \mathbb{R} \tag{1.7.3}$$

*definiert und werden mit zwei Verknüpfungen*

$$+, \cdot : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}, \tag{1.7.4}$$

gegeben durch

$$(a, b) + (a', b') = (a + a', b + b') \quad (1.7.5)$$

und

$$(a, b) \cdot (a', b') = (aa' - bb', ab' + a'b), \quad (1.7.6)$$

versehen.

Mit anderen Worten sind die komplexen Zahlen einfach Paare von reellen Zahlen, für die wir die komponentenweise Addition  $+$  gemäß (1.7.5) und eine neue Multiplikation  $\cdot$  in (1.7.6) definieren. Es ist nun eine routinemäßige Verifikation, dass wir auf diese Weise einen Körper erhalten:

**Satz 1.7.2 (Der Körper  $\mathbb{C}$ )** Die komplexen Zahlen  $\mathbb{C}$  werden durch die beiden Operationen  $+$  und  $\cdot$  zu einem Körper mit Nullelement

$$0 = (0, 0) \quad (1.7.7)$$

und Einselement

$$1 = (1, 0). \quad (1.7.8)$$

Das multiplikative Inverse zu  $(a, b) \neq 0$  ist

$$(a, b)^{-1} = \left( \frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right). \quad (1.7.9)$$

Die Abbildung

$$\mathbb{R} \ni a \mapsto (a, 0) \in \mathbb{C} \quad (1.7.10)$$

ist ein (injektiver) Körpermorphismus, womit  $\mathbb{R}$  als Unterkörper von  $\mathbb{C}$  angesehen werden kann. Es gilt

$$\text{char}(\mathbb{C}) = 0. \quad (1.7.11)$$

BEWEIS: Der Beweis besteht zunächst darin, die nötigen Eigenschaften von  $+$  und  $\cdot$  nachzuprüfen: Assoziativität, Kommutativität und Distributivität. Für  $+$  ist dies einfach zu sehen und beruht auf der komponentenweisen Assoziativität und Kommutativität, die von  $\mathbb{R}$  geerbt wird. Für die neue Multiplikation ist dies eine direkte und einfache Rechnung. Dass 0 und 1 die neutralen Elemente sind, zeigt man ebenfalls direkt. Interessanter ist der Nachweis, dass (1.7.9) tatsächlich das Inverse zu  $(a, b)$  darstellt. Der entscheidende Punkt ist, dass für  $(a, b) \neq 0$  die reelle Zahl  $a^2 + b^2$  positiv und daher invertierbar ist. Hier benutzen wir entscheidend, dass  $\mathbb{R}$  ein *angeordneter* Körper ist, siehe Proposition 1.3.5, v.). Damit wird  $(a, b)^{-1} \cdot (a, b) = (1, 0)$  zu einer einfachen algebraischen Überprüfung. Dies zeigt, dass  $\mathbb{C}$  ein Körper ist. Es ist nun anhand der Formeln (1.7.5) und (1.7.6) klar, dass  $(a, 0) + (a', 0) = (a + a', 0)$  sowie  $(a, 0) \cdot (a', 0) = (aa', 0)$  die übliche Addition und Multiplikation von reellen Zahlen reproduziert. Da  $(0, 0)$  das neutrale Element von  $\mathbb{C}$  bezüglich  $+$  und  $(1, 0)$  das neutrale Element von  $\mathbb{C} \setminus \{0\}$  bezüglich  $\cdot$  ist, folgt also, dass (1.7.10) ein Körpermorphismus ist. Die Injektivität ist einerseits klar anhand der expliziten Formel (1.7.10), gilt aber sowieso für jeden Körpermorphismus nach Übung 1.8.1. Damit ist aber auch  $\text{char}(\mathbb{C}) = 0$  klar, da bereits  $\mathbb{R}$  Charakteristik Null besitzt.  $\square$

Wir betrachten nun das Element

$$i = (0, 1) \in \mathbb{C}, \quad (1.7.12)$$

für das

$$i^2 = -1 \quad (1.7.13)$$

schnell nachgerechnet ist. Dies ist also die gewünschte Wurzel aus  $-1$ , nach der wir gesucht haben. Man beachte, dass aus diesem Grunde  $\mathbb{C}$  *nicht* angeordnet werden kann, da in einem angeordneten

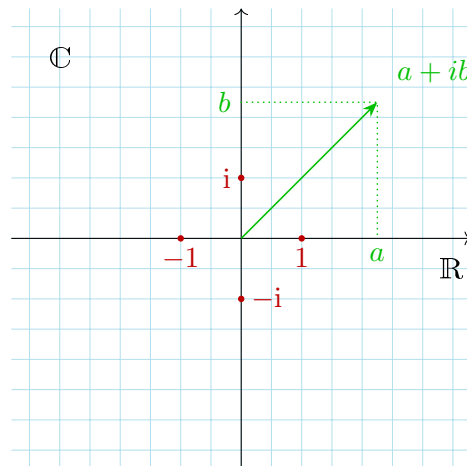


Abbildung 1.8: Die komplexe Zahlenebene

Körper Quadrate immer positiv sind,  $-1$  aber immer negativ sein muss. Mit diesem Element können wir also jede komplexe Zahl  $(a, b)$  auch als

$$(a, b) = a \cdot 1 + b \cdot i = a + ib \quad (1.7.14)$$

schreiben. In der Sprache der linearen Algebra haben wir mit  $\mathbb{C}$  einen zwei-dimensionalen reellen Vektorraum  $\mathbb{R}^2$  vorliegen, für den die beiden Elemente  $1 = (1, 0)$  und  $i = (0, 1)$  Basisvektoren bilden. Aus diesem Grunde sprechen wir auch von der *komplexen Zahlenebene* im Gegensatz zur reellen Zahlengeraden. Die Schreibweise (1.7.14) ist dann nichts anderes als die Basisentwicklung des Vektors  $(a, b) \in \mathbb{R}^2$  bezüglich der Basis  $\{1, i\}$ , siehe auch Abbildung 1.8.

Wir nennen für  $z = a + ib \in \mathbb{C}$

$$\operatorname{Re}(z) = a \quad (1.7.15)$$

den *Realteil* von  $z$  und

$$\operatorname{Im}(z) = b \quad (1.7.16)$$

den *Imaginärteil*, wobei  $a, b \in \mathbb{R}$ . Komplexe Zahlen der Form  $z = ib$  mit  $b \in \mathbb{R}$  nennt man auch *imaginär*.

Da die Gleichung (1.7.1) zwei Lösungen in  $\mathbb{C}$  besitzt, nämlich  $i$  und  $-i$ , gibt es zunächst keinen Grund,  $i$  vor  $-i$  auszuzeichnen. Es gibt daher eine Symmetrie des Körpers, die  $i$  mit  $-i$  vertauscht, die *komplexe Konjugation*:

**Definition 1.7.3 (Komplexe Konjugation)** Die Abbildung

$$\bar{\phantom{x}}: \mathbb{C} \ni z = a + ib \mapsto \bar{z} = a - ib \in \mathbb{C} \quad (1.7.17)$$

heißt *komplexe Konjugation*.

**Proposition 1.7.4** Die komplexe Konjugation ist ein involutiver Körperautomorphismus von  $\mathbb{C}$ , dessen Fixpunkte genau die reellen Zahlen sind. Es gilt also für alle  $z, w \in \mathbb{C}$

$$\overline{z + w} = \bar{z} + \bar{w}, \quad (1.7.18)$$

$$\overline{zw} = \bar{z} \bar{w}, \quad (1.7.19)$$

$$\overline{\bar{z}} = z \quad (1.7.20)$$

und

$$\bar{z} = z \iff z \in \mathbb{R} \subseteq \mathbb{C}. \quad (1.7.21)$$

BEWEIS: Der Beweis erfolgt durch Nachrechnen, siehe Übung 1.8.32.  $\square$

Die elementar-geometrische Anschauung der komplexen Zahlenebene legt nahe, die Länge des Vektors  $(a, b)$  mit Hilfe des Satzes von Pythagoras zu bestimmen. Da die Basisvektoren 1 und  $i$  senkrecht stehen, ist die Länge durch  $\sqrt{a^2 + b^2}$  gegeben. Man beachte jedoch, dass wir zum jetzigen Zeitpunkt noch nicht über mathematisch adäquate Definitionen von Winkeln und Längen verfügen. Von daher sollen diese Überlegungen lediglich als Motivation für die Definition des *Betrages* einer komplexen Zahl dienen:

**Definition 1.7.5 (Betrag)** Sei  $z \in \mathbb{C}$ . Dann ist der Betrag von  $z$  als

$$|z| = \sqrt{a^2 + b^2} \quad (1.7.22)$$

definiert, wobei  $z = a + ib$  mit  $a, b \in \mathbb{R}$ .

Der Betrag einer komplexen Zahl verhält sich tatsächlich wie die geometrische Länge. Diese und weitere nützliche Eigenschaften zeigt folgende Proposition:

**Proposition 1.7.6** Seien  $z, w \in \mathbb{C}$ .

- i.) Es gilt  $|z| \geq 0$  und  $|z| = 0$  genau dann, wenn  $z = 0$ .
- ii.) Der Betrag von  $a \in \mathbb{R}$  im Sinne von Definition 1.3.9 stimmt mit der von Definition 1.7.5 überein.
- iii.) Es gilt  $|z| = |\bar{z}|$  sowie  $|zw| = |z||w|$  und  $|1| = 1$ .
- iv.) Für  $z \neq 0$  gilt  $|\frac{1}{z}| = \frac{1}{|z|}$ .
- v.) Es gilt  $|z| = \sqrt{z\bar{z}}$ .
- vi.) Es gelten die Abschätzungen

$$|\operatorname{Re}(z)| \leq |z| \quad \text{und} \quad |\operatorname{Im}(z)| \leq |z|. \quad (1.7.23)$$

- vii.) Es gilt die Dreiecksungleichung

$$|z \pm w| \leq |z| + |w|. \quad (1.7.24)$$

- viii.) Es gilt die umgekehrte Dreiecksungleichung

$$|z - w| \geq ||z| - |w||. \quad (1.7.25)$$

BEWEIS: Mit der Definition  $|z| = \sqrt{a^2 + b^2}$  ist  $|z| \geq 0$  klar, da Wurzeln aus nicht-negativen Zahlen wieder nicht-negativ sind. Weiter gilt genau dann  $|z| = 0$ , wenn  $a^2 + b^2 = 0$  gilt, was nur für  $a = 0 = b$  möglich ist, siehe erneut Proposition 1.3.5, v.). Der zweite Teil ist klar, da  $a^2 \geq 0$  für  $a \in \mathbb{R}$  und  $\sqrt{a^2} = |a|$  gilt. Die Verträglichkeit des Betrags mit der multiplikativen Struktur von  $\mathbb{C}$  ist eine algebraische Überprüfung, die in Übung 1.8.33 erfolgt. Da  $a^2 \leq a^2 + b^2$ , folgt aus Proposition 1.6.5, vi.), die Ungleichung  $\sqrt{a} \leq \sqrt{a^2 + b^2}$ , die strikt ist, sobald  $b \neq 0$ . Dies zeigt  $|\operatorname{Re}(z)| = |a| \leq |z|$  und genauso  $|\operatorname{Im}(z)| \leq |z|$ . Die Dreiecksungleichung (1.7.24) werden wir noch in größerem Kontext sehen. Hier zeigen wir sie ad-hoc durch

$$\begin{aligned} |z \pm w|^2 &= (\overline{z \pm w})(z \pm w) \\ &= \bar{z}z \pm 2\operatorname{Re}(\bar{z}w) + \bar{w}w \\ &\leq \bar{z}z + 2|\operatorname{Re}(\bar{z}w)| + \bar{w}w \\ &\leq |z|^2 + 2|\bar{z}w| + |w|^2 \\ &= |z|^2 + 2|z||w| + |w|^2 \end{aligned}$$



$$= (|z| + |w|)^2.$$

Ziehen der Wurzel liefert dann die gewünschte Ungleichung (1.7.24), da Wurzelziehen Ungleichungen erhält. Für den letzten Teil schreiben wir

$$|z| = |z - w + w| \leq |z - w| + |w|,$$

womit  $|z| - |w| \leq |z - w|$ . Vertauschen wir die Rollen von  $w$  und  $z$ , so erhalten wir  $|w| - |z| \leq |w - z| = |z - w|$ . Insgesamt gilt also (1.7.25).  $\square$

**Bemerkung 1.7.7 (Vollständigkeit von  $\mathbb{C}$ )** Da  $\mathbb{C}$  kein angeordneter Körper sein kann, wir unsere Definition von Vollständigkeit aber auf der Existenz von Suprema aufgebaut haben, ist es zunächst nicht klar, in welchem Sinne  $\mathbb{C}$  ebenfalls vollständig sein könnte. Vollständigkeit mit Hilfe von Dedekindschen Schnitten scheitert aus dem selben Grund. Das Intervallschachtelungsprinzip wäre eine Möglichkeit, wenn man Intervalle  $[a, b] \subseteq \mathbb{R}$  beispielsweise durch ineinander geschachtelte Rechtecke in der komplexen Zahlenebene  $\mathbb{C}$  ersetzt. Im Prinzip kann man auf diese Weise einen Begriff von Vollständigkeit definieren, der dann auch tatsächlich erfüllt ist. Allerdings ist die genaue Definition wenig konzeptuell. Wir werden diese Möglichkeit daher nicht weiter verfolgen sondern in Abschnitt 2.4 einen gänzlich anderen Begriff von Vollständigkeit etablieren, der dann insgesamt deutlich allgemeiner sein wird.

Wir schließen diesen Abschnitt mit einem Satz zu den komplexen Zahlen, der den wesentlichen Unterschied zu den reellen Zahlen herausstellt:

**Satz 1.7.8 (Fundamentalsatz der Algebra)** *Der Körper  $\mathbb{C}$  ist algebraisch abgeschlossen: Jedes nicht-konstante Polynom  $p(z) = a_n z^n + \dots + a_1 z + a_0$  mit komplexen Koeffizienten  $a_0, a_1, \dots, a_n \in \mathbb{C}$  hat eine komplexe Nullstelle  $\lambda \in \mathbb{C}$ , also  $p(\lambda) = 0$ .*

Der Beweis übersteigt momentan unsere Fähigkeiten. Die einfachsten Beweise werden mit Techniken aus der Funktionentheorie erzielt, auf die wir an dieser Stelle verweisen müssen. Wir werden zu gegebener Zeit auf diesen Satz und seine Anwendungen zurückkommen. Die große Bedeutung der komplexen Zahlen liegt also unter anderem darin, dass wir nicht nur für die Gleichung  $x^2 + 1 = 0$  Lösungen finden, sondern für jedes Polynom mit komplexen Koeffizienten Nullstellen haben.

## 1.8 Übungen

**Übung 1.8.1 (Körpermorphismen)** Seien  $\mathbb{k}$  und  $\mathbb{K}$  Körper. Eine Abbildung  $\phi: \mathbb{k} \rightarrow \mathbb{K}$  heißt *Körpermorphismus*, wenn

$$\phi(a + b) = \phi(a) + \phi(b) \quad \text{und} \quad \phi(0_{\mathbb{k}}) = 0_{\mathbb{K}} \quad (1.8.1)$$

sowie

$$\phi(ab) = \phi(a)\phi(b) \quad \text{und} \quad \phi(1_{\mathbb{k}}) = 1_{\mathbb{K}} \quad (1.8.2)$$

gilt.

- i.) Zeigen Sie, dass die Identität  $\text{id}: \mathbb{k} \rightarrow \mathbb{k}$  ein Körpermorphismus ist und dass die Verkettung von Körpermorphismen wieder ein Körpermorphismus ist.
- ii.) Zeigen Sie, dass ein Körpermorphismus notwendigerweise injektiv ist.

Hinweis: Ein Körpermorphismus ist insbesondere ein Gruppenmorphismus von  $(\mathbb{k}, +, 0_{\mathbb{k}})$  nach  $(\mathbb{K}, +, 0_{\mathbb{K}})$ . Nutzen Sie daher das Kriterium für die Injektivität von Gruppenmorphismen, indem Sie den Kern von  $\phi$  betrachten.

- iii.) Zeigen Sie, dass  $\mathbb{k}$  und  $\mathbb{K}$  notwendigerweise die gleiche Charakteristik haben, sobald es einen Körpermorphismus  $\phi: \mathbb{k} \rightarrow \mathbb{K}$  gibt.

**Übung 1.8.2 (Binomische Formeln)** Sei  $\mathbb{k}$  ein Körper. Zeigen Sie für  $a, b \in \mathbb{k}$  die Binomischen Formeln

$$(a + b)^2 = a^2 + 2ab + b^2, \quad (1.8.3)$$

$$(a - b)^2 = a^2 - 2ab + b^2 \quad (1.8.4)$$

und

$$(a + b)(a - b) = a^2 - b^2. \quad (1.8.5)$$

Wie vereinfachen sich diese Formeln für  $\mathbb{k} = \mathbb{Z}_2$ ?

**Übung 1.8.3** Rechnen mit  $\sum$ ,  $\prod$ .  $n \cdot x$  hat zwei Bedeutungen! Index verschieben.

**Übung 1.8.4 (Produkt von Potenzen)** Zeigen Sie  $(xy)^n = x^n y^n$  analog zum Beweis von Proposition 1.2.7, i.).

Hinweis: Betrachten Sie zunächst  $n \in \mathbb{N}_0$  und führen Sie einen Induktionsbeweis. Wie erhält man dann daraus den Fall  $n \in \mathbb{Z} \setminus \mathbb{N}_0$ ?

**Übung 1.8.5 (Eigenschaften von Fakultät und Binomialkoeffizient)** Seien  $n, k \in \mathbb{N}_0$  natürliche Zahlen mit  $k \leq n$ .

- i.) Beweisen Sie, dass  $n! \leq n^n$  für alle  $n \in \mathbb{N}$  ist. Wann gilt die strikte Ungleichung?  
 ii.) Zeigen Sie, dass es genau  $\binom{n}{k}$  verschiedene Teilmengen mit genau  $k$  Elementen der Menge  $\mathbf{n} = \{1, \dots, n\}$  gibt.  
 iii.) Folgern Sie, dass  $\binom{n}{k} \in \mathbb{N}$  und dass

$$\sum_{k=0}^n \binom{n}{k} = 2^n \quad (1.8.6)$$

gilt. Folgern Sie insbesondere  $\binom{n}{k} \leq 2^n$  für alle  $0 \leq k \leq n$ .

- iv.) Zeigen Sie

$$\binom{n}{0} = 1 = \binom{n}{n} \quad (1.8.7)$$

und

$$\binom{n}{1} = n = \binom{n}{n-1}. \quad (1.8.8)$$

- v.) Zeigen Sie

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k} \quad (1.8.9)$$

durch eine geeignete Induktion. Folgern Sie so erneut, dass  $\binom{n}{k} \in \mathbb{N}$ . Interpretieren Sie diese Gleichung als Pascalsches Dreieck.

**Übung 1.8.6 (Klammerungen)** Sei  $\mathbb{k}$  ein Körper und  $x_1, \dots, x_n \in \mathbb{k}$  für  $n \in \mathbb{N}$ .

- i.) Betrachten Sie die Menge aller möglichen Klammerungen der  $n$  Elemente und zeigen Sie, dass deren Anzahl  $K_n$  die Rekursionsformel

$$K_n = \sum_{r=1}^{n-1} K_r K_{n-r} \quad (1.8.10)$$

mit Startwert  $K_2 = 1$  und  $K_3 = 2$  erfüllt. Bestimmen Sie  $K_4, K_5, K_6$  explizit.

- ii.) Zeigen Sie induktiv, dass in einem  $n$ -fachen Produkt die Assoziativität (1.2.5) genutzt werden kann, um die allgemeine Assoziativität zu zeigen.

Hinweis: Sei  $1 \leq m \leq n-1$  und setzen Sie  $X_m = (x_1 \cdots x_m)(x_{m+1} \cdots x_n)$ , wobei nach Induktion die Art der Klammerung bei  $m$  beziehungsweise  $n-m$  Faktoren bereits als unerheblich angenommen werden kann. Zeigen Sie nun durch wiederholtes Anwenden der Assoziativität (1.2.5), dass  $X_1 = \cdots = X_{n-1}$  gilt. Hierzu müssen Sie in den einzelnen Faktoren geeignete Klammerungen wählen.

**Übung 1.8.7 (Permutationen)** Betrachten Sie  $n \in \mathbb{N}$  sowie die Menge  $\mathbf{n} = \{1, \dots, n\}$  der ersten  $n$  natürlichen Zahlen. Eine Abbildung  $\sigma: \mathbf{n} \rightarrow \mathbf{n}$  heißt *Permutation*, wenn  $\sigma$  bijektiv ist. Die Menge der Permutation wird mit  $S_n = \text{Bij}(\mathbf{n})$  bezeichnet.

- i.) Zeigen Sie, dass die Permutationen eine Gruppe bezüglich der Verknüpfung bilden.  
 ii.) Zeigen Sie, dass  $S_n$  genau  $n!$  Elemente besitzt.

**Übung 1.8.8 (Geometrische Summe)** Sei  $\mathbb{k}$  ein Körper und  $n \in \mathbb{N}_0$ .

- i.) Zeigen Sie die geometrische Summenformel (1.2.36) durch Induktion nach  $n$ .  
 ii.) Zeigen Sie

$$\frac{x^n - y^n}{x - y} = \sum_{k=0}^{n-1} x^k y^{n-k-1} \quad (1.8.11)$$

für alle  $x, y \in \mathbb{k}$  mit  $x \neq y$ . Wie ist dies für  $n = 0$  zu verstehen?

**Übung 1.8.9 (Anordnung von  $\mathbb{Q}$ )**

**Übung 1.8.10** Umkehrung Proposition 1.3.4

**Übung 1.8.11 (Minimum und Maximum)**

**Übung 1.8.12** Zeigen der Dreiecksungleichung normal

$$|x + y| \leq |x| + |y|$$

und verallgemeinert für  $|\sum x_i|$ , sowie der umgekehrten Dreiecksungleichung.

**Übung 1.8.13**  $(\mathbb{k}, \mathbb{k}^+) \supseteq \mathbb{Q}$  ordnungserhaltend

**Übung 1.8.14**  $x^2 = p$  für  $p$  Primzahl  $x^k = p$ ,  $k \in \mathbb{N}$

**Übung 1.8.15**  $\mathbb{k}$  archimedisch  $\iff \{\frac{1}{n}, n \in \mathbb{N}\}$  hat  $\inf = 0$ .

**Übung 1.8.16 (Supremum und Infimum)**

**Übung 1.8.17** Sup/Inf von Intervallen

**Übung 1.8.18**  $\sqrt{2}$  mit Intervallschachtelung.

**Übung 1.8.19 (Intervallschachtelung mit offenen Intervallen)**

**Übung 1.8.20 (Viele rationale Zahlen)** Reelle Zahlen können wir auf verschiedene Weisen durch rationale approximieren. Hier ist eine besonders einfache Beobachtung: Sei  $\epsilon > 0$ , ein  $N \in \mathbb{N}$  und  $x_0 \in \mathbb{R}$  vorgegeben. Zeigen Sie, dass es dann teilerfremde  $p \in \mathbb{Z}$  und  $q \in \mathbb{N}$  mit  $q \geq N$  gibt, so dass  $|x - \frac{p}{q}| < \epsilon$  gilt.

Hinweis: Wählen Sie  $q$  als Primzahl mit  $q \geq \max\{N, \frac{1}{\epsilon}\}$ .

**Übung 1.8.21**  $c < ab \implies c = a'b'$  mit  $a' < a$ ,  $b' < b$  Explizite Beispiele  $a' b' c < a + b \implies c = a' + b' \dots$

**Übung 1.8.22** Gegenbeispiele zu sup-Rechenregeln.

**Übung 1.8.23** Rechenregeln aus Prop 1.5.13 für infima.

**Übung 1.8.24 (Suprema und Infima von unbeschränkten Teilmengen)** Zeigen Sie, dass die Aussagen von Proposition 1.5.13 für unbeschränkte Teilmengen ebenfalls richtig sind, wenn man die Rechenregeln gemäß den Konventionen (1.5.34), (1.5.35) und (1.5.36) interpretiert. Bestimmen Sie dazu die Eigenschaften der Teilmengen  $-A$ ,  $A+B$ , und  $A \cdot B$ , wenn  $A$  oder  $B$  nach oben (nach unten) unbeschränkt sind.

**Übung 1.8.25** Weitere Sup/Inf Rechenregeln.

**Übung 1.8.26**  $A \subseteq \mathbb{R}$  Dedek. Schnitt  $\iff \mathbb{R} \setminus A$  auch. Dann Trennungszahl gleich mit  $\sup A = \inf \mathbb{R} \setminus A$ , Eindeutigkeit der Trennungszahl, wenn existent.

**Übung 1.8.27**  $a < b \iff a^q > b^q$ ,  $q \in \mathbb{Q}^-$   $0 < a < 1 \implies a^q > a^p$ ,  $q < p$  wenn es  $a > 1$  mit  $a^x < a^y$  gibt, ist  $x < y$ .

**Übung 1.8.28**  $a^x$  für  $a < 1$  als Supremum.

**Übung 1.8.29 (Potenzen als Schnitzzahlen)** Sei  $a \in \mathbb{R}^+$  mit  $a > 1$  und  $x \in \mathbb{R}$ . Zeigen Sie, dass

$$A = \bigcup_{q \in \mathbb{Q}, q \leq x} \{y \mid y \leq a^q\} \quad (1.8.12)$$

den Dedekindschen Schnitt mit Schnitzzahl  $a^x$  definiert. Dies wäre eine weitere Variante, wie man  $a^x$  für beliebiges reelles  $x$  definieren kann.

**Übung 1.8.30**  $\sup A^x = (\sup A)^x$ ,  $x \in \mathbb{R}$ . "Stetigkeiten" von  $a^x$

**Übung 1.8.31** Details zu  $\mathbb{C}$ .

**Übung 1.8.32** Beweis von Proposition 1.7.4

**Übung 1.8.33** Verträglichkeit Betrag mit Multiplikation, *iii.) – v.)* aus Prop 1.7.6

**Übung 1.8.34** Nenner reell machen.

**Übung 1.8.35 (Real- und Imaginärteil)**

**Übung 1.8.36 (Einheitskreis  $\mathbb{S}^1$  als Gruppe)** Die komplexen Zahlen mit Betrag  $|z| = 1$  bezeichnet man als

$$\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}. \quad (1.8.13)$$

*i.)* Visualisieren Sie die Teilmenge  $\mathbb{S}^1 \subseteq \mathbb{C}$  und rechtfertigen Sie so den Namen *1-Sphäre* beziehungsweise *Einheitskreis*.

*ii.)* Zeigen Sie, dass  $\mathbb{S}^1$  eine multiplikative Untergruppe von  $\mathbb{C}^\times$ .

**Übung 1.8.37 (Polynomdivision)**

## Kapitel 2

# Folgen und Reihen

In diesem Kapitel werden wir nun die entscheidenden Begriffe der Konvergenz von Folgen und Reihen vorstellen. Folgen sind spezielle Abbildungen, bei denen wir uns für die “langfristige” Entwicklung interessieren. Dies führt auf den Begriff der Grenzwertes, den wir hier in seiner ersten Version kennenlernen. Später werden wir noch viele andere Varianten von Konvergenz und Grenzwerten sehen, aber dieses Kapitel ist grundlegend für alle diese anderen Formulierungen.

Da Folgen spezielle Abbildungen sind, lohnt es sich allgemein die algebraischen Eigenschaften von Abbildungen zu untersuchen. Wir stellen diese Überlegungen voran und spezialisieren uns anschließend auf Folgen. Unser Interesse gilt dabei Abbildungen mit Werten in entweder  $\mathbb{R}$  oder  $\mathbb{C}$ :

**Konvention 2.1 ( $\mathbb{K}$  und  $\mathbb{k}$ )** *Um sowohl reelle Zahlen als auch komplexe Zahlen anzusprechen, verabreden wir, dass das Symbol  $\mathbb{K}$  entweder für  $\mathbb{R}$  oder für  $\mathbb{C}$  steht. Tritt in einer Definition oder Aussage also  $\mathbb{K}$  auf, so ist diese sowohl für  $\mathbb{R}$  als auch für  $\mathbb{C}$  gültig. Das Symbol  $\mathbb{k}$  verwenden wir weiterhin für einen beliebigen Körper, auch wenn wir diese nicht mehr oft benötigen werden.*

Sei nun  $X$  eine beliebige Menge, die wir als nicht leer annehmen wollen. Mit

$$\text{Abb}(X, \mathbb{K}) = \{f: X \longrightarrow \mathbb{K} \text{ ist Abbildung von } X \text{ nach } \mathbb{K}\} \quad (2.1)$$

bezeichnen wir dann die *Menge der Abbildungen* von  $X$  nach  $\mathbb{K}$ . Alternativ sprechen wir auch von  *$\mathbb{K}$ -wertigen Funktionen* auf  $X$ . Da wir für  $\mathbb{K}$  bereits viele Rechenoperationen kennen, übertragen wir diese nun auf  $\text{Abb}(X, \mathbb{K})$ , indem wir *punktweise* Definitionen verwenden.

**Definition 2.2 (Rechenoperationen für Funktionen)** *Seien  $X$  eine nicht-leere Menge und  $f, g \in \text{Abb}(X, \mathbb{K})$ .*

*i.) Die Summe  $f + g \in \text{Abb}(X, \mathbb{K})$  ist punktweise für  $x \in X$  erklärt durch*

$$(f + g)(x) = f(x) + g(x). \quad (2.2)$$

*ii.) Das Vielfache  $\alpha f \in \text{Abb}(X, \mathbb{K})$  für  $\alpha \in \mathbb{K}$  ist punktweise für  $x \in X$  erklärt durch*

$$(\alpha f)(x) = \alpha f(x). \quad (2.3)$$

*iii.) Das Produkt  $fg \in \text{Abb}(X, \mathbb{K})$  ist punktweise für  $x \in X$  erklärt durch*

$$(fg)(x) = f(x)g(x). \quad (2.4)$$

*iv.) Ist  $f(x) \neq 0$  für alle  $x \in X$ , so ist der Kehrwert  $\frac{1}{f} \in \text{Abb}(X, \mathbb{K})$  punktweise für  $x \in X$  erklärt durch*

$$\left(\frac{1}{f}\right)(x) = \frac{1}{f(x)}. \quad (2.5)$$

v.) Ist  $f(x) > 0$  für alle  $x \in X$  positiv und  $\alpha \in \mathbb{R}$ , so definieren wir  $f^\alpha \in \text{Abb}(X, \mathbb{R})$  punktweise für  $x \in X$  durch

$$(f^\alpha)(x) = (f(x))^\alpha. \quad (2.6)$$

Allgemeiner definieren wir in diesem Fall für eine reellwertige Funktion  $g \in \text{Abb}(X, \mathbb{R})$  die Funktion  $f^g \in \text{Abb}(X, \mathbb{R})$  punktweise für  $x \in X$  durch

$$(f^g)(x) = (f(x))^{g(x)}. \quad (2.7)$$

vi.) Ist  $\mathbb{K} = \mathbb{C}$ , so definiert man  $\bar{f} \in \text{Abb}(X, \mathbb{C})$  punktweise für  $x \in X$  durch

$$(\bar{f})(x) = \overline{f(x)}. \quad (2.8)$$

Es sollte nun klar sein, dass wir hieraus auch weitere punktweise Operationen erhalten, wie beispielsweise  $\text{Re}(f)$ ,  $\text{Im}(f)$  oder  $|f|$ , um nur einige zu nennen. Sobald wir neue Rechenoperationen für  $\mathbb{R}$  oder  $\mathbb{C}$  erhalten, übertragen sich diese entsprechend automatisch auf  $\mathbb{K}$ -wertige Funktionen. Die Eigenschaften der punktweise definierten Operationen erhält man dann direkt aus den entsprechenden Eigenschaften für  $\mathbb{K}$ , siehe auch Übung 2.9.1.

Im Falle reellwertiger Funktionen können wir auch noch die Werte der Größe nach vergleichen. Dies liefert eine (partielle) Ordnung für  $\text{Abb}(X, \mathbb{R})$ :

**Definition 2.3 (Partielle Ordnung für  $\text{Abb}(X, \mathbb{R})$ )** Sei  $X$  eine nicht-leere Menge. Für zwei reellwertige Funktionen  $f, g \in \text{Abb}(X, \mathbb{R})$  schreiben wir  $f < g$ , falls

$$f(x) < g(x) \quad (2.9)$$

für alle  $x \in X$  gilt. Entsprechend verwenden wir die Symbole  $\leq$ ,  $>$  und  $\geq$ .

Es sei hier gewarnt, dass zwei Funktionen  $f, g \in \text{Abb}(X, \mathbb{R})$  im Allgemeinen *nicht* in ihrer Größe vergleichbar sind: Es kann durchaus passieren, dass *keine* der drei Aussagen

$$f < g \quad \text{oder} \quad f = g \quad \text{oder} \quad f > g \quad (2.10)$$

richtig ist. Der Grund ist einfach einzusehen, da es ja Punkte  $x \in X$  mit  $f(x) < g(x)$  und gleichzeitig andere Punkte  $x' \in X$  mit  $f(x') > g(x')$  geben kann. Die Relation  $<$ , und ebenso  $\leq, >, \geq$ , ist also nicht für alle Paare von Funktionen sondern nur partiell definiert. Aus diesem Grund nennt man dies eine *partielle Ordnung*, siehe auch Übung 2.9.2.

Bis jetzt lag der Fokus auf den Werten und die Menge  $X$  war bislang ohne weitere Struktur. Wir werden nun den Wechsel des Definitionsbereichs betrachten, und daher Abbildungen zwischen verschiedenen Definitionsbereichen einbeziehen. Dies führt zur Definition des *pull-backs*:

**Definition 2.4 (Pull-back)** Sei  $\Phi: X \rightarrow Y$  eine Abbildung zwischen zwei nicht-leeren Mengen. Dann ist der pull-back

$$\Phi^*: \text{Abb}(Y, \mathbb{K}) \rightarrow \text{Abb}(X, \mathbb{K}) \quad (2.11)$$

von Funktionen durch

$$(\Phi^* f)(x) = f(\Phi(x)) \quad (2.12)$$

punktweise für  $x \in X$  erklärt.

Da wir mit dem pull-back auf den Argumenten der Funktionen agieren, die bisherigen Rechenoperationen aber nur die Werte involvieren, sind pull-backs generell mit allen punktweise erklärten Rechenoperationen verträglich. Die Details finden sich in Übung 2.9.3. Wir formulieren dies hier etwas pauschal in folgender Proposition:

**Proposition 2.5** Seien  $X, Y, Z$  nicht-leere Mengen und  $\Phi: X \rightarrow Y$  sowie  $\Psi: Y \rightarrow Z$  Abbildungen.

i.) Der pull-back  $\Phi^*: \text{Abb}(Y, \mathbb{K}) \rightarrow \text{Abb}(X, \mathbb{K})$  ist mit allen punktweise erklärten Rechenoperationen von Funktionen verträglich.

ii.) Es gilt

$$(\Psi \circ \Phi)^* = \Phi^* \circ \Psi^* \quad (2.13)$$

sowie

$$\text{id}_X^* = \text{id}_{\text{Abb}(X, \mathbb{K})}. \quad (2.14)$$

BEWEIS: Der erste Teil wird in Übung 2.9.3 genauer formuliert und bewiesen. Für den zweiten Teil betrachten wir  $f \in \text{Abb}(Z, \mathbb{K})$  und  $x \in X$ . Dann gilt

$$((\Psi \circ \Phi)^* f)(x) = f((\Psi \circ \Phi)(x)) = f(\Psi(\Phi(x))) = (\Psi^* f)(\Phi(x)) = (\Phi^*(\Psi^* f))(x),$$

womit (2.13) gezeigt ist. Die Behauptung (2.14) ist klar.  $\square$

## 2.1 Folgen und Abzählbarkeit

Wir kommen nun zum ersten wichtigen Spezialfall von Abbildungen, den reellen oder komplexen Folgen.

**Definition 2.1.1 (Folge)** Eine  $\mathbb{K}$ -wertige Folge  $a$  ist eine Abbildung  $a \in \text{Abb}(\mathbb{N}, \mathbb{K})$ , also

$$a: \mathbb{N} \rightarrow \mathbb{K}. \quad (2.1.1)$$

Im Prinzip können wir auch Folgen mit Werten in anderen Mengen als  $\mathbb{K}$  betrachten. Dies wird jedoch erst zu einem späteren Zeitpunkt geschehen.

Eine Folge  $a$  ist dadurch gegeben, welchen Wert  $a_n = a(n)$  sie für  $n \in \mathbb{N}$  besitzt. Wir schreiben daher auch

$$a = (a(n))_{n \in \mathbb{N}} = (a_n)_{n \in \mathbb{N}}. \quad (2.1.2)$$

Der Index  $n \in \mathbb{N}$  kann auch bei anderen natürlichen Zahlen oder auch bei  $n = 0$  beginnen. Wir werden also insbesondere Abbildungen  $a: \mathbb{N}_0 \rightarrow \mathbb{K}$  ebenfalls als Folgen bezeichnen.

**Beispiel 2.1.2 (Folgen)**

i.) Sei  $c \in \mathbb{K}$  fest gewählt. Dann ist  $a \in \text{Abb}(\mathbb{N}, \mathbb{K})$  mit

$$a_n = c \quad (2.1.3)$$

die konstante Folge mit Wert  $c$ .

ii.) Interessantere Folgen erhalten wir, indem wir die Rechenmöglichkeiten in  $\mathbb{K}$  zum Einsatz bringen. Die Folgen

$$a_n = cn \quad \text{für festes } c \in \mathbb{K}, \quad (2.1.4)$$

$$a_n = n^k \quad \text{für festes } k \in \mathbb{Z}, \quad (2.1.5)$$

$$a_n = \alpha^n \quad \text{für festes } \alpha \in \mathbb{K}, \quad (2.1.6)$$

$$a_n = \frac{1}{n}, \quad (2.1.7)$$

$$a_n = (-1)^n, \quad (2.1.8)$$

$$a_n = n! \quad (2.1.9)$$

sind von wiederkehrendem Interesse.

iii.) Wir können Folgen auch induktiv definieren. Hierzu gibt man sich die ersten Folgenglieder  $a_1, \dots, a_N \in \mathbb{K}$  vor und spezifiziert dann eine Vorschrift, wie  $a_n$  aus  $a_{n-1}, \dots, a_{n-N}$  bestimmt werden kann. Auf diese Weise erhält man beispielsweise die *Fibonacci-Folge* mit

$$a_1 = 1, a_2 = 1 \quad \text{und} \quad a_n = a_{n-1} + a_{n-2} \quad (2.1.10)$$

für  $n \geq 3$ . Die ersten Folgenglieder lauten dann

$$a_3 = 2, \quad a_4 = 3, \quad a_5 = 5, \quad a_6 = 8, \quad \dots \quad (2.1.11)$$

Weitere interessante Eigenschaften dieser Folge finden sich in Übung 2.9.4 sowie in Übung 2.9.5.

Wir können Folgen auch so verstehen, dass Elemente des Wertebereichs durchnummeriert werden sollen. Diese Vorstellung führt auf folgende Definition:

**Definition 2.1.3 (Abzählbarkeit)** Sei  $M$  eine Menge.

i.) Die Menge  $M$  heißt *abzählbar*, wenn  $M$  entweder endlich viele Elemente besitzt, oder wenn es eine bijektive Abbildung

$$\phi: \mathbb{N} \longrightarrow M \quad (2.1.12)$$

gibt. In diesem Fall heißt  $\phi$  eine *Abzählung* von  $M$ .

ii.) Ist  $M$  nicht abzählbar, so heißt  $M$  *überabzählbar*.

Eine abzählbare unendliche Menge ist also im mengentheoretischen Sinne gleich groß wie  $\mathbb{N}$ . Eine Menge  $M$  ist genau dann unendlich, wenn es eine injektive Abbildung  $\phi: \mathbb{N} \longrightarrow M$  gibt. Die interessante Fragestellung ist also, ob es auch eine surjektive Abbildung (2.1.12) gibt. Überraschenderweise können beide Situationen auftreten. Mengen können also sehr unterschiedlich groß sein:

**Satz 2.1.4 (Überabzählbarkeit von  $\mathbb{R}$ )**

i.) Die rationalen Zahlen  $\mathbb{Q}$  sind abzählbar.

ii.) Die reellen Zahlen  $\mathbb{R}$  sind überabzählbar. Es gibt keine surjektive Abbildung

$$\phi: \mathbb{N} \longrightarrow \mathbb{R}. \quad (2.1.13)$$

BEWEIS: Der erste Teil mit vielen weiteren Beispielen von abzählbar unendlichen Mengen wird in Übung 2.9.6 besprochen. Überraschender ist daher der zweite Teil. Wir zeigen, dass es keine Surjektion (2.1.13) geben kann. Sei  $\phi: \mathbb{N} \longrightarrow \mathbb{R}$  eine Abbildung. Dann gibt es  $a_1, b_1 \in \mathbb{R}$  mit  $\phi(1) \notin [a_1, b_1]$  und  $a_1 < b_1$ . Dies ist klar, wir können beispielsweise  $a_1 = \phi(1) + 1$  und  $b_1 = \phi(1) + 2$  verwenden. Dann gibt es auch  $\alpha, \beta \in \mathbb{R}$  mit

$$a_1 < \alpha < \beta < b_1$$

da dies in jedem archimedischen Körper nach Proposition 1.4.2, iii.), möglich ist. Wir erhalten also eine Vereinigung

$$[a_1, b_1] = [a_1, \alpha] \cup [\alpha, \beta] \cup [\beta, b_1]$$

mit drei Intervallen, von denen keines leer ist. Die nächste Zahl  $\phi(2)$  kann daher nur in höchstens zwei dieser drei Intervalle sein, aber nicht in allen dreien. Es gibt daher ein abgeschlossenes Intervall, das wir mit  $[a_2, b_2]$  bezeichnen, für das  $a_2 < b_2$  sowie

$$\phi(2) \notin [a_2, b_2] \quad \text{und} \quad [a_2, b_2] \subseteq [a_1, b_1]$$

gilt. Insbesondere gilt auch  $\phi(1) \notin [a_2, b_2]$ . Induktiv können wir so zwei Folgen  $a = (a_n)_n$  und  $b = (b_n)_n$  konstruieren, so dass

$$\phi(1), \dots, \phi(n) \notin [a_n, b_n] \subseteq [a_{n-1}, b_{n-1}] \subseteq \dots \subseteq [a_1, b_1]$$



für alle  $n \in \mathbb{N}$  gilt. Die Vollständigkeit von  $\mathbb{R}$  in Form des Intervallschachtelungsprinzips nach Satz 1.5.9 zeigt dann, dass der Durchschnitt

$$I = \bigcap_{n=1}^{\infty} [a_n, b_n]$$

dieser Intervalle selbst ein nicht-leeres (abgeschlossenes) Intervall  $I$  ist. Da aber  $\phi(n) \notin [a_n, b_n]$  für alle  $n \in \mathbb{N}$  gilt, folgt  $\phi(n) \notin I$  für alle  $n \in \mathbb{N}$ . Insbesondere ist  $\phi$  nicht surjektiv.  $\square$

Es gibt also viel mehr reelle Zahlen als rationale Zahlen. Dies ist der Preis der Vollständigkeit von  $\mathbb{R}$ : Wir müssen sehr viele Zahlen hinzunehmen, um von  $\mathbb{Q}$  nach  $\mathbb{R}$  zu kommen.

**Korollar 2.1.5** *Die komplexen Zahlen sind überabzählbar.*

BEWEIS: Gäbe es eine Surjektion  $\phi: \mathbb{N} \rightarrow \mathbb{C}$ , so auch eine auf  $\mathbb{R} \subseteq \mathbb{C}$ .  $\square$

**Bemerkung 2.1.6 (Algebraische, irrationale und transzendente Zahlen)** Es gibt also Zahlen  $x \in \mathbb{R} \setminus \mathbb{Q}$ , wie wir ja auch bereits zuvor gesehen haben. Die Wurzel  $\sqrt{2}$  ist ein solches Beispiel. Reelle Zahlen, die nicht rational sind, nennt man auch *irrational*. Es folgt, dass es überabzählbar viele irrationale Zahlen gibt, siehe nochmals Übung 2.9.7. Man kann die irrationalen Zahlen nun weiter unterteilen. Man nennt  $\lambda \in \mathbb{R}$  *algebraisch*, falls es ein Polynom  $p(x) = a_n x^n + \dots + a_1 x + a_0$  mit rationalen Koeffizienten  $a_0, a_1, \dots, a_n \in \mathbb{Q}$  gibt, welches  $\lambda$  als Nullstelle  $p(\lambda) = 0$  besitzt. Insbesondere ist jede rationale Zahl  $q \in \mathbb{Q}$  algebraisch mit  $p(x) = x - q$ , aber auch  $\sqrt{2}$  ist algebraisch mit  $p(x) = x^2 - 2$ . Man kann sich nun überlegen, dass die Menge der algebraischen Zahlen in  $\mathbb{R}$  immer noch vergleichsweise klein ist. Die algebraischen Zahlen können nämlich immer noch abgezählt werden, siehe Übung 2.9.8. Die verbleibenden reellen Zahlen nennt man *transzendent*. Es gibt daher überabzählbar viele solche transzendente reelle Zahlen. Allerdings ist es trotzdem alles andere als einfach, konkrete Beispiele für transzendente reelle Zahlen zu finden, oder auch für eine gegebene reelle Zahl zu entscheiden, ob sie transzendent ist.

Wir diskutieren noch zwei wichtige Konzepte für Folgen, die nicht unmittelbar auf den Wertebereich Bezug nehmen, sondern nur auf den Definitionsbereich  $\mathbb{N}$ :

**Definition 2.1.7 (Teilfolge)** Sei  $a \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge. Ist  $\phi: \mathbb{N} \rightarrow \mathbb{N}$  streng monoton wachsend, also  $\phi(n) > \phi(m)$  für  $n > m$ , so nennt man die neue Folge  $\phi^*a \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Teilfolge von  $a$ .

Ausgeschrieben bedeutet dies, dass

$$(\phi^*a)(n) = a_{\phi(n)} \quad (2.1.14)$$

für alle  $n \in \mathbb{N}$  gilt. Anschaulich gesprochen erhält man die Elemente der Teilfolge  $\phi^*a$  dadurch, dass man unendlich viele Folgenglieder der ursprünglichen Folge auswählt, ihre Reihenfolge beibehält, und anschließend wieder durch natürliche Zahlen durchnummeriert, also beispielsweise

$$\begin{array}{ccccccc} a & = & (a_1, & a_2, & a_3, & a_4, & a_5, \dots) \\ & & & \swarrow & \searrow & \swarrow & \searrow \\ \phi^*a & = & (a_2, & a_4, & a_5, & \dots & \dots \dots) \end{array} \quad (2.1.15)$$

wenn  $\phi(1) = 2, \phi(2) = 4, \phi(3) = 5, \dots$

Man beachte, dass eine streng monoton wachsende Abbildung  $\phi: \mathbb{N} \rightarrow \mathbb{N}$  sicherlich injektiv ist. Die Umkehrung braucht aber nicht zu gelten.

Verschiedene Eigenschaften einer Folge werden nicht davon abhängen, in welcher Reihenfolge ihre Elemente angeordnet sind, wie etwa die Frage, ob alle Elemente positiv sind. Andere, interessantere Beispiele für derartige Eigenschaften werden wir noch kennenlernen. Um diese Problematik angemessen formalisieren zu können, betrachten wir eine weitere Sorte von pull-back: die *Umordnung*. Hierzu erinnern wir zunächst daran, dass die Menge der *Bijektionen*

$$\text{Bij}(\mathbb{N}) = \{\sigma \in \text{Abb}(\mathbb{N}, \mathbb{N}) \mid \sigma \text{ ist bijektiv}\} \quad (2.1.16)$$

von  $\mathbb{N}$  bezüglich der Komposition eine Gruppe bilden. Mit solchen Bijektionen können wir dann Umordnungen definieren:

**Definition 2.1.8 (Umordnung einer Folge)** Sei  $a \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge. Ist  $\sigma \in \text{Bij}(\mathbb{N})$ , so heißt  $\sigma^*a \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine *Umordnung* der Folge  $a$ .

Da  $\sigma$  eine Bijektion ist, tritt in der umgeordneten Folge  $\sigma^*a$  jedes Folgenglied von  $a$  genau einmal auf, nur eben an anderer, durch  $\sigma$  spezifizierter, Stelle.

## 2.2 Konvergenz von Folgen

In diesem Abschnitt kommen wir nun zur zentralen Definition der Theorie der Folgen. Als Motivation sei vorangestellt, dass wir das Verhalten einer Folge  $(a_n)_{n \in \mathbb{N}}$  für immer größer werdende Indizes  $n \in \mathbb{N}$  verstehen wollen. Um dieses approximative Verhalten zu spezifizieren, bedient man sich nun folgender Definition:

**Definition 2.2.1 (Konvergenz von Folgen)** Sei  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge. Eine Zahl  $a \in \mathbb{K}$  heißt *Grenzwert* (oder *Limes*) der Folge  $(a_n)_{n \in \mathbb{N}}$ , falls es für alle  $\epsilon > 0$  einen Index  $N \in \mathbb{N}$  gibt, sodass für alle Indizes  $n \geq N$

$$|a_n - a| < \epsilon \quad (2.2.1)$$

gilt. In diesem Fall schreiben wir

$$a = \lim_{n \rightarrow \infty} a_n \quad (2.2.2)$$

und sagen, dass die Folge  $(a_n)_{n \in \mathbb{N}}$  gegen  $a$  konvergiert. Andernfalls heißt die Folge *divergent*. Wir schreiben auch  $a \xrightarrow[n \rightarrow \infty]{} a$  oder einfach  $a_n \rightarrow a$  im Falle der Konvergenz.

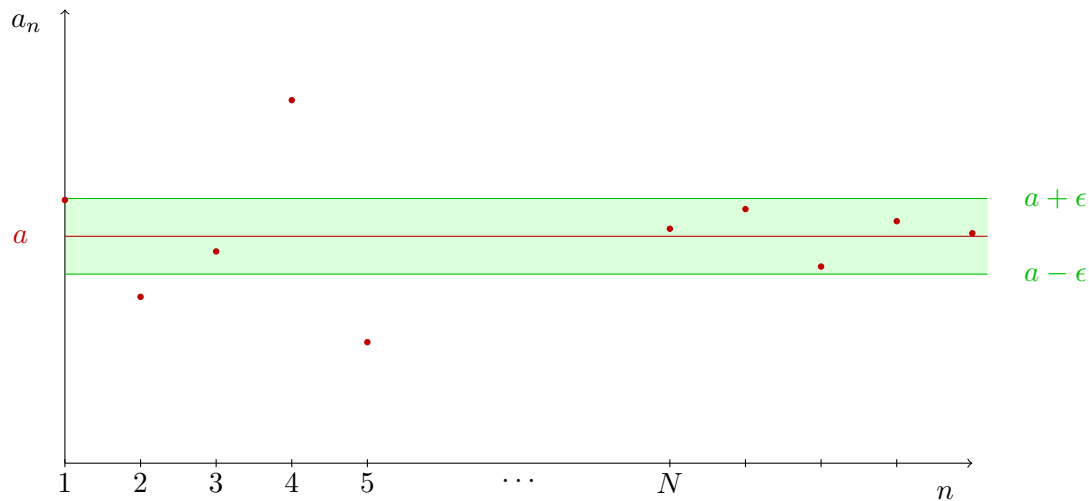
### Bemerkung 2.2.2 (Konvergenz von Folgen)

- i.) Es ist eine gute Tradition in der Mathematik, für die Definition der Konvergenz an die Quantorenschreibweise mit dem *Allquantor*  $\forall$  und dem *Existenzquantor*  $\exists$  zu erinnern. Das Symbol  $\forall$  steht einfach für *für alle*, während  $\exists$  für *es existiert ein* steht. Die Konvergenzbedingung wird dann zu

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \forall n \geq N: |a_n - a| < \epsilon. \quad (2.2.3)$$

Die Quantorenschreibweise stellt oft eine sehr effiziente Weise dar, Dinge klar zu sagen, und sollte daher gut verinnerlicht werden. Insbesondere lassen sich Aussagen leicht negieren, siehe auch Übung 2.9.11.

- ii.) Die Idee der Definition ist also, dass für eine gegebene Approximationsgüte  $\epsilon > 0$  nur endlich viele Folgenglieder  $a_n$  den Grenzwert  $a$  schlechter als bis auf  $\epsilon$  approximieren: Diese sind unter den Folgengliedern mit  $n < N$  zu finden, alle größeren Indizes  $n \geq N$  liefern bereits gute Approximationen, siehe auch Abbildung 2.1. Es ist also wichtig, hier festzuhalten, dass der Index  $N$  sehr wohl von der zuvor festgelegten Approximationsgüte  $\epsilon > 0$  abhängen darf. Aus diesem Grund schreibt man auch  $N(\epsilon)$ , wenn man diese Abhängigkeit betonen möchte.

Abbildung 2.1: Die Folge  $(a_n)_{n \in \mathbb{N}}$  konvergiert gegen  $a$ .

- iii.) Die Definition ist gleichermaßen für reelle Zahlen und komplexe Zahlen sinnvoll, da wir in beiden Fällen einen Betrag  $|\cdot|$  erklärt haben. Man beachte, dass die Identifikation von reellen Zahlen  $\mathbb{R}$  mit den besonderen komplexen Zahlen mit Imaginärteil 0 eine konsistente Sichtweise ergibt, da für  $a \in \mathbb{R}$  eben  $|a| = \sqrt{a^2}$  gilt.

Bevor wir zu den ersten Beispielen kommen, gilt es noch zu rechtfertigen, den Grenzwert einer Folge wirklich *den* Grenzwert zu nennen:

**Proposition 2.2.3** *Sei  $(a_n)_{n \in \mathbb{N}}$  eine reelle oder komplexe Folge. Konvergiert  $(a_n)_{n \in \mathbb{N}}$ , so ist der Grenzwert eindeutig bestimmt.*

BEWEIS: Seien  $a, a' \in \mathbb{K}$  zwei Grenzwerte und  $\epsilon > 0$  beliebig gewählt. Dann gibt es  $N, N' \in \mathbb{N}$  mit den Eigenschaften

$$|a_n - a| < \frac{\epsilon}{2} \quad (*)$$

für alle  $n \geq N$  und

$$|a_n - a'| < \frac{\epsilon}{2} \quad (**)$$

für alle  $n \geq N'$ . Ist nun  $n \geq \max\{N, N'\}$ , so gilt

$$|a - a'| = |a - a_n + a_n - a'| \leq |a - a_n| + |a_n - a'| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

da wir sowohl (\*) als auch (\*\*) benutzen dürfen. Es folgt also, dass für beliebiges  $\epsilon > 0$  immer  $|a - a'| < \epsilon$  gilt. Dies kann aber wegen  $0 \leq |a - a'|$  nur für  $|a - a'| = 0$  der Fall sein, womit  $a = a'$  folgt.  $\square$

### Beispiel 2.2.4 (Konvergente Folgen)

- i.) Sei  $c \in \mathbb{K}$  fest gewählt. Dann konvergiert die konstante Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n = c$  gegen  $c$ . Dies ist unmittelbar klar, da  $|a_n - c| = 0$  für alle  $n \in \mathbb{N}$  gilt.
- ii.) Wir betrachten die Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n = \frac{1}{n}$ . Diese konvergiert gegen Null. Um dies zu zeigen, geben wir ein  $\epsilon > 0$  vor. Da  $\mathbb{R}$  archimedisch ist, finden wir eine natürliche Zahl  $N \in \mathbb{N}$  mit  $N > \frac{1}{\epsilon}$ . Für  $n \geq N$  gilt dann

$$|a_n - 0| = \left| \frac{1}{n} \right| \leq \frac{1}{N} < \epsilon, \quad (2.2.4)$$

was die Konvergenz  $\frac{1}{n} \rightarrow 0$  zeigt. Eine Folge wie  $(\frac{1}{n})_{n \in \mathbb{N}}$ , die gegen Null konvergiert, heißt auch *Nullfolge*.

iii.) Wir betrachten die Folge  $(n^k)_{n \in \mathbb{N}}$  für ein festes  $k \in \mathbb{N}$ . Diese Folge divergiert. Um dies zu zeigen, nehmen wir an, sie konvergiere gegen ein  $a \in \mathbb{R}$ . Dann gibt es also zu  $\epsilon = 1$  ein  $N \in \mathbb{N}$  mit

$$|n^k - a| < 1 \quad (2.2.5)$$

für alle  $n \geq N$ . Wir können  $N$  gegebenenfalls weiter vergrößern, um zudem  $N \geq 2|a|$  zu erreichen. Dann gilt aber  $\frac{N}{2} \geq |a|$ , und somit  $N \geq |a| + \frac{N}{2}$ . Für alle  $n \geq N$  ist daher ebenfalls  $n \geq |a| + \frac{n}{2}$  gültig, womit auch  $n^k \geq |a| + \frac{n}{2}$  gilt. Es folgt nun aus (2.2.5)

$$1 > |n^k - a| \geq ||n^k| - |a|| \geq |a| + \frac{n}{2} - |a| = \frac{n}{2} \quad (2.2.6)$$

für alle  $n \geq N$ . Dies ist der gewünschte Widerspruch, womit  $(n^k)_{n \in \mathbb{N}}$  keinen Grenzwert besitzt.

iv.) Schließlich betrachten wir die Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n = (-1)^n$ . Auch diese Folge divergiert. Sei nämlich  $a$  der Grenzwert von  $(a_n)_{n \in \mathbb{N}}$ . Dann verwenden wir wieder das Konvergenzkriterium für  $\epsilon = 1$ , um ein  $N \in \mathbb{N}$  zu finden, sodass  $|a_n - a| < 1$  für alle  $n \geq N$  gilt. Nun gilt aber  $|a_{n+1} - a_n| = 2$  für alle  $n \in \mathbb{N}$ , also

$$2 = |a_{n+1} - a_n| = |a_{n+1} - a + a - a_n| \leq |a_{n+1} - a| + |a - a_n| < 1 + 1 = 2, \quad (2.2.7)$$

was den Widerspruch darstellt, den wir suchen.

Um interessantere Beispiele diskutieren zu können, fehlt uns momentan die richtige Technologie. Daher starten wir mit einigen allgemeinen Überlegungen zur Konvergenz.

**Proposition 2.2.5** *Eine Folge  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  konvergiert genau dann gegen  $a \in \mathbb{K}$ , wenn jede Teilfolge von  $(a_n)_{n \in \mathbb{N}}$  gegen  $a$  konvergiert.*

BEWEIS: Sei  $\phi: \mathbb{N} \rightarrow \mathbb{N}$  streng monoton wachsend, und  $(b_n)_{n \in \mathbb{N}} = (a_{\phi(n)})_{n \in \mathbb{N}}$  die zugehörige Teilfolge von  $(a_n)_{n \in \mathbb{N}}$ . Dann gilt  $\phi(n) \geq n$  für alle  $n \in \mathbb{N}$ . Sei nun  $\epsilon > 0$  gegeben und  $N \in \mathbb{N}$  mit  $|a_n - a| < \epsilon$  für  $n \geq N$ . Dann gilt auch  $|b_n - a| = |a_{\phi(n)} - a| < \epsilon$  da  $\phi(n) \geq n$  später in  $\mathbb{N}$  auftritt. Damit konvergiert also auch  $(b_n)_{n \in \mathbb{N}}$  gegen  $a$ . Die Umkehrung ist trivial, da  $\phi = \text{id}_{\mathbb{N}}$  eine spezielle Teilfolge liefert.  $\square$

Allerdings kann es sehr wohl passieren, dass eine Teilfolge konvergiert, die ursprüngliche Folge aber nicht. Verwendet man beispielsweise

$$\phi(n) = 2n \quad (2.2.8)$$

und die Folge  $(a_n)_{n \in \mathbb{N}} = ((-1)^n)_{n \in \mathbb{N}}$  aus (2.1.8), so ist  $\phi^*(a_n)_{n \in \mathbb{N}} = ((-1)^{\phi(n)})_{n \in \mathbb{N}} = (1)_{n \in \mathbb{N}}$  die konstante Folge und damit konvergent gegen 1. Entsprechend liefert  $\tilde{\phi}(n) = 2n + 1$  eine Teilfolge, die gegen  $-1$  konvergiert. Mit dem obigen Resultat bedeutet dies aber, dass die ursprüngliche Folge nicht konvergent sein kann, da die Teilfolgen unterschiedliche Grenzwerte besitzen. Dies ist oft ein einfacher Test auf Divergenz.

**Definition 2.2.6 (Beschränkte Folgen)** *Eine Folge  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  heißt beschränkt, falls es ein  $R > 0$  mit*

$$|a_n| \leq R \quad (2.2.9)$$

*für alle  $n \in \mathbb{N}$  gibt. Eine reelle Folge  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{R})$  heißt nach oben (nach unten) beschränkt, falls es ein  $R \in \mathbb{R}$  mit*

$$a_n \leq R \quad \text{beziehungsweise} \quad a_n \geq R \quad (2.2.10)$$

für alle  $n \in \mathbb{N}$  gibt. Ist eine Folge nicht (nach oben oder nach unten) beschränkt, so heißt sie auch (nach oben oder nach unten) unbeschränkt. Für eine beschränkte Folge definiert man die Supremumsnorm  $\|(a_n)_{n \in \mathbb{N}}\|_\infty$  durch

$$\|(a_n)_{n \in \mathbb{N}}\|_\infty = \sup\{|a_n| \mid n \in \mathbb{N}\}. \quad (2.2.11)$$

**Proposition 2.2.7** Eine konvergente Folge  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  ist beschränkt, und es gilt

$$\left| \lim_{n \rightarrow \infty} a_n \right| \leq \|(a_n)_{n \in \mathbb{N}}\|_\infty. \quad (2.2.12)$$

BEWEIS: Sei  $(a_n)_{n \in \mathbb{N}}$  konvergent gegen  $a \in \mathbb{K}$ . Sei weiter  $\epsilon > 0$  vorgegeben mit zugehörigem  $N \in \mathbb{N}$ , sodass  $|a_n - a| < \epsilon$  für alle  $n \geq N$ . Dann gilt

$$|a| = |a - a_n + a_n| \leq |a_n - a| + |a_n| = \epsilon + |a_n|$$

für diese  $n \geq N$ . Genauso erhalten wir

$$|a_n| \leq \epsilon + |a|$$

für alle  $n \geq N$ , was insbesondere

$$|a_n| \leq \max\{|a_1|, \dots, |a_{N-1}|, \epsilon + |a|\}$$

zeigt. Daher ist die Folge beschränkt. Es folgt weiterhin, dass

$$|a| \leq \epsilon + |a_n| \leq \epsilon + \sup\{|a_n| \mid n \geq N\} \leq \epsilon + \sup\{|a_n| \mid n \in \mathbb{N}\} = \epsilon + \|(a_n)_{n \in \mathbb{N}}\|_\infty$$

für alle  $\epsilon > 0$ . Damit muss aber auch (2.2.12) gelten.  $\square$

Dies Resultat erlaubt es, das Beispiel 2.2.4, iii.), auf einfachere Weise zu entscheiden: die Folgen  $(n^k)_{n \in \mathbb{N}}$  sind für alle  $k \in \mathbb{N}$  unbeschränkt. Daher konvergiert keine von ihnen.

### Beispiel 2.2.8 (Unbeschränkte Folgen)

- i.) Sei  $c \in \mathbb{K}$  mit  $c \neq 0$  gegeben. Dann ist  $(cn)_{n \in \mathbb{N}}$  unbeschränkt, da  $|cn| = |c|n$  und  $\mathbb{R}$  archimedisch ist.
- ii.) Sei  $\alpha \in \mathbb{K}$  mit  $|\alpha| > 1$  gegeben. Dann ist die Folge  $(\alpha^n)_{n \in \mathbb{N}}$  unbeschränkt, da  $|\alpha| = 1 + c$  mit  $c > 0$  und daher  $|\alpha^n| = |\alpha|^n = (1 + c)^n \geq 1 + nc > nc$  nach der Bernoulli-Ungleichung.
- iii.) Die Folge  $(n!)_{n \in \mathbb{N}}$  ist unbeschränkt, da  $n! \geq n$ .

Das nächste Resultat ist ein sehr wichtiges Kriterium zur Folgenkonvergenz und ein Paradebeispiel zur Argumentation in der Analysis. Durch Abschätzungen alleine lassen sich Grenzwerte unbekannter Folgen aus denen bekannter bestimmen:

**Proposition 2.2.9** Seien  $(a_n)_{n \in \mathbb{N}}$ ,  $(b_n)_{n \in \mathbb{N}}$  und  $(c_n)_{n \in \mathbb{N}}$  drei reelle Folgen mit

$$a_n \leq b_n \leq c_n \quad (2.2.13)$$

für alle  $n \in \mathbb{N}$ . Konvergieren dann  $(a_n)_{n \in \mathbb{N}}$  und  $(c_n)_{n \in \mathbb{N}}$  gegen den selben Grenzwert  $a \in \mathbb{R}$ , so konvergiert auch  $(b_n)_{n \in \mathbb{N}}$  gegen  $a$ .

BEWEIS: Sei  $\epsilon > 0$  und  $N \in \mathbb{N}$  derart, dass  $|a_n - a| < \epsilon$  und  $|c_n - a| < \epsilon$ , falls  $n \geq N$ . Man beachte, dass wir beide Bedingungen simultan mit gleichem  $N$  erfüllen können, indem wir gegebenenfalls zum Maximum der individuellen Wahlen übergehen. Dann gilt also

$$a_n - a \leq b_n - a \leq c_n - a$$

und damit

$$|b_n - a| \leq \max\{|a_n - a|, |c_n - a|\} < \epsilon$$

für  $n \geq N$ . Dies zeigt aber die gewünschte Konvergenz  $b_n \rightarrow a$ .  $\square$

Aus naheliegenden Gründen wird dieses Resultat auch als *Sandwich-Satz* bezeichnet.

**Beispiel 2.2.10 (Sandwich-Prinzip)** Als Anwendungen des Sandwich-Prinzips betrachten wir die beiden Nullfolgen  $(0)_{n \in \mathbb{N}}$  und  $(\frac{1}{n})_{n \in \mathbb{N}}$  und versuchen, Folgen zwischen diesen beiden Folgen einzusperren.

i.) Für alle  $k \in \mathbb{R}$  mit  $k > 1$  ist  $0 < \frac{1}{n^k} < \frac{1}{n}$  nach Satz 1.6.8, vii.). Damit folgt

$$\lim_{n \rightarrow \infty} \frac{1}{n^k} = 0. \quad (2.2.14)$$

ii.) Sei nun  $\alpha \in \mathbb{K}$  mit  $|\alpha| < 1$ . Dann gilt  $\frac{1}{|\alpha|} = 1 + c$  mit  $c > 0$ , und daher  $\frac{1}{|\alpha|^n} = (1 + c)^n \geq 1 + nc$ , womit

$$|\alpha|^n \leq \frac{1}{1 + nc} < \frac{1}{cn}. \quad (2.2.15)$$

Nun gilt  $\frac{1}{cn} \rightarrow 0$ , was daher  $|\alpha|^n \rightarrow 0$  und auch

$$\lim_{n \rightarrow \infty} \alpha^n = 0 \quad (2.2.16)$$

impliziert, da  $|\alpha^n| = |\alpha|^n$ .

Zum Abschluss dieses Abschnitts zeigen wir, dass das Konvergenzverhalten nur von den Folgengliedern aber nicht von deren Reihenfolge abhängt. Die Konvergenz ist daher eine Eigenschaft, die unter Umordnung der Folge invariant ist:

**Proposition 2.2.11** Sei  $(a_n)_{n \in \mathbb{N}}$  eine Folge und  $\sigma \in \text{Bij}(\mathbb{N})$  eine Umordnung. Die Folge  $(a_n)_{n \in \mathbb{N}}$  konvergiert genau dann gegen  $a$ , wenn die umgeordnete Folge  $\sigma^*(a_n)_{n \in \mathbb{N}} = (a_{\sigma(n)})_{n \in \mathbb{N}}$  gegen  $a$  konvergiert.

BEWEIS: Sei  $\sigma \in \text{Bij}(\mathbb{N})$  und  $\epsilon > 0$  vorgegeben. Wir nehmen an, dass  $(a_n)_{n \in \mathbb{N}}$  gegen  $a \in \mathbb{K}$  konvergiert. Es gibt also ein  $N \in \mathbb{N}$  mit  $|a_n - a| < \epsilon$ , sofern  $n \geq N$  gilt. Da  $\phi$  bijektiv ist, haben wir eindeutige Urbilder  $\sigma^{-1}(1), \dots, \sigma^{-1}(N)$  der ersten  $N$  Zahlen. Wir setzen

$$M = \max\{\sigma^{-1}(1), \dots, \sigma^{-1}(N)\}.$$

Sei nun  $m > M$ . Dann gilt  $\sigma(m) \notin \{1, \dots, N\}$  und somit  $\sigma(m) > N$ , siehe auch Abbildung 2.2. Also folgt  $|a_{\sigma(m)} - a| < \epsilon$ . Dies zeigt, dass wir zu jedem  $\epsilon > 0$  ein  $M \in \mathbb{N}$  mit  $|a_{\sigma(m)} - a| < \epsilon$  für  $m > M$  finden, womit  $(a_{\sigma(m)})_{m \in \mathbb{N}} \rightarrow a$  gezeigt ist. Gilt umgekehrt für eine Umordnung  $\sigma$  die Konvergenz  $\sigma^*(a_n)_{n \in \mathbb{N}} \rightarrow a$ , so gilt nach dem gerade Gezeigten auch  $(\sigma^{-1})^*(\sigma^*(a_n)_{n \in \mathbb{N}}) \rightarrow a$ , da  $\sigma^{-1}$  ebenfalls eine Bijektion ist. Nach Proposition 2.5, ii.), gilt aber  $(\sigma^{-1})^*(\sigma^*(a_n)_{n \in \mathbb{N}}) = (\sigma \circ \sigma^{-1})^*(a_n)_{n \in \mathbb{N}} = (a_n)_{n \in \mathbb{N}}$ , was die Umkehrung zeigt.  $\square$

Wenn bei einer unbeschränkten Folge alle Folgenglieder groß werden, wollen wir von Divergenz gegen  $+\infty$  sprechen. Dies lässt sich folgendermaßen präzisieren:

**Definition 2.2.12 (Divergenz gegen  $\pm\infty$ )** Sei  $(a_n)_{n \in \mathbb{N}}$  eine reelle Folge. Dann heißt  $(a_n)_{n \in \mathbb{N}}$  divergent gegen  $+\infty$ , geschrieben als

$$\lim_{n \rightarrow \infty} a_n = +\infty, \quad (2.2.17)$$

falls es zu jedem  $R \in \mathbb{R}$  ein  $N \in \mathbb{N}$  mit

$$a_n \geq R \quad (2.2.18)$$

für alle  $n \geq N$  gibt. Entsprechend definiert man Divergenz gegen  $-\infty$ , falls  $a_n \leq R$  in (2.2.18). Alternativ spricht man auch von Konvergenz gegen  $\pm\infty$ .

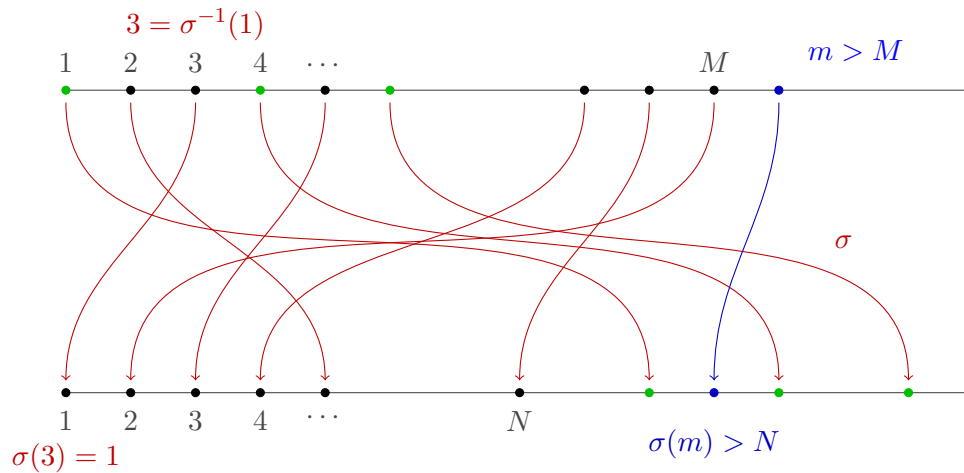


Abbildung 2.2: Für eine Bijektion  $\sigma \in \text{Bij}(\mathbb{N})$  finden sich die Urbilder der ersten  $N$  Zahlen unter den ersten  $M$  Zahlen, für ein geeignetes, von  $N$  abhängiges  $M$ . Zahlen größer als  $M$  werden auf Zahlen größer  $N$  abgebildet.

Man beachte, dass wir hier wirklich nur reelle Folgen betrachten können, da es für komplexe Folgenglieder zu viele Richtungen in  $\mathbb{C}$  gibt, um nach  $\infty$  zu entkommen. Im reellen Fall sind es eben nur zwei.

### Beispiel 2.2.13

- i.) Die Folge  $(n)_{n \in \mathbb{N}}$  divergiert gegen  $+\infty$ . Dies ist gerade die Aussage, dass  $\mathbb{R}$  archimedisch ist.
- ii.) Die Folge  $(2^n)_{n \in \mathbb{N}}$  divergiert ebenfalls gegen  $+\infty$ , als Teilfolge der Folge aus i.).
- iii.) Die Folge  $((-2)^n)_{n \in \mathbb{N}}$  ist unbeschränkt, divergiert aber weder gegen  $+\infty$  noch gegen  $-\infty$ , da wir wechselnde Vorzeichen haben.
- iv.) Die Folge  $(n!)_{n \in \mathbb{N}}$  divergiert ebenfalls gegen  $+\infty$ , da  $n! \geq n$  für alle  $n \in \mathbb{N}$ .
- v.) Die Folge  $(a_n)_{n \in \mathbb{N}}$  mit

$$a_n = \begin{cases} \frac{1}{n} & n \text{ gerade} \\ 2^n & n \text{ ungerade} \end{cases} \quad (2.2.19)$$

ist nach oben unbeschränkt, aber nicht divergent gegen  $+\infty$ .

## 2.3 Rechenregeln für konvergente Folgen

In diesem Abschnitt wollen wir untersuchen, wie die Rechenoperationen und die Ordnungsrelation von  $\mathbb{R}$  beziehungsweise  $\mathbb{C}$  sich mit der Konvergenz von Folgen vertragen, wenn wir sie gemäß Definition 2.2 und Definition 2.3 auf Folgen übertragen. Zuvor betrachten wir beschränkte Folgen:

**Definition 2.3.1 (Folgenraum  $\ell^\infty$ )** Der Folgenraum  $\ell^\infty$  ist die Menge aller komplexen beschränkten Folgen in  $\text{Abb}(\mathbb{N}, \mathbb{C})$ , also

$$\ell^\infty = \{(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{C}) \mid (a_n)_{n \in \mathbb{N}} \text{ ist beschränkt}\}. \quad (2.3.1)$$

Streng genommen haben wir zwei Varianten, den Raum aller reellen beschränkten Folgen und den der komplexen beschränkten Folgen. Da reelle Folgen als spezielle komplexe Folgen aufgefasst werden können, genügt es bei  $\ell^\infty$ , die komplexe Variante zu studieren.

**Proposition 2.3.2** Seien  $a = (a_n)_{n \in \mathbb{N}}$ ,  $b = (b_n)_{n \in \mathbb{N}} \in \ell^\infty$  zwei beschränkte Folgen.

i.) Für alle  $z, w \in \mathbb{C}$  gilt  $za + wb \in \ell^\infty$  und

$$\|za + wb\|_\infty \leq |z|\|a\|_\infty + |w|\|b\|_\infty. \quad (2.3.2)$$

ii.) Es gilt  $ab \in \ell^\infty$  und

$$\|ab\|_\infty \leq \|a\|_\infty \cdot \|b\|_\infty. \quad (2.3.3)$$

iii.) Es gilt  $\bar{a} \in \ell^\infty$  und

$$\|\bar{a}a\|_\infty = \|a\|_\infty^2 \quad (2.3.4)$$

sowie  $\|\bar{a}\|_\infty = \|a\|_\infty$ .

BEWEIS: Nach den Rechenregeln für Suprema aus Proposition 1.5.13, iii.), gilt

$$\begin{aligned} \sup\{|za_n + wb_n| \mid n \in \mathbb{N}\} &\leq \sup\{|z||a_n| + |w||b_n| \mid n \in \mathbb{N}\} \\ &\leq \sup\{|z||a_n| + |w||b_m| \mid n, m \in \mathbb{N}\} \\ &= |z| \sup\{|a_n| \mid n \in \mathbb{N}\} + |w| \sup\{|b_m| \mid m \in \mathbb{N}\} \\ &= |z|\|a\|_\infty + |w|\|b\|_\infty. \end{aligned}$$

Die zweite Abschätzung gilt, da wir zu einer größeren Teilmenge übergehen. Wegen  $a, b \in \ell^\infty$  ist die rechte Seite endlich, womit auch das Supremum auf der linken Seite endlich ist. Daher folgt  $za + wb \in \ell^\infty$ , und (2.3.2) ist gezeigt. Den zweiten Teil zeigt man genauso mit Proposition 1.5.13, iv.), durch

$$\begin{aligned} \sup\{|a_nb_n| \mid n \in \mathbb{N}\} &\leq \sup\{|a_n||b_m| \mid n, m \in \mathbb{N}\} \\ &= \sup\{|a_n| \mid n \in \mathbb{N}\} \cdot \sup\{|b_m| \mid m \in \mathbb{N}\} \\ &= \|a\|_\infty \cdot \|b\|_\infty. \end{aligned}$$

Auch der dritte Teil ist klar, da

$$\sup\{|\bar{a}_na_n| \mid n \in \mathbb{N}\} = \sup\{|a_n|^2 \mid n \in \mathbb{N}\} = \sup\{|a_n| \mid n \in \mathbb{N}\}^2 = \|a\|_\infty^2.$$

Hier verwenden wir die allgemeine Rechenregel

$$\sup\{r^2 \mid r \in A \subseteq \mathbb{R}_0^+\} = \sup\{r \mid r \in A \subseteq \mathbb{R}_0^+\}^2, \quad (*)$$

die für jede Teilmenge von  $\mathbb{R}_0^+$  gültig ist. Die eine Ungleichung „ $\leq$ “ in (\*) folgt direkt aus Proposition 1.5.13, iv.). Die andere Richtung erfordert eine zusätzliche Überlegung: Ist  $\alpha$  eine obere Schranke aller Quadrate  $a^2$  mit  $a \in A$ , so gilt also  $a^2 \leq \alpha$ . Daher ist  $\sqrt{\alpha}$  eine obere Schranke aller  $a \in A$ , und entsprechend gilt  $\sup A \leq \sqrt{\alpha}$ , was  $(\sup A)^2 \leq \sqrt{\alpha}^2 = \alpha$  zeigt. Dies ist die Ungleichung „ $\geq$ “ in (\*). Man beachte, dass das Supremum auf der linken Seite von (\*) nur über eine Teilmenge von  $A \cdot A$  gebildet wird, so dass also Proposition 1.5.13, iv.), nicht direkt zum Einsatz kommen kann. Die letzte Behauptung  $\|\bar{a}\|_\infty = \|a\|_\infty$  ist trivial, da  $|\bar{a}_n| = |a_n|$ .  $\square$

**Bemerkung 2.3.3** In der Sprache der linearen Algebra zeigt Proposition 2.3.2, i.), dass  $\ell^\infty \subseteq \text{Abb}(\mathbb{N}, \mathbb{C})$  ein komplexer Untervektorraum ist. Der zweite Teil zeigt, dass  $\ell^\infty$  eine Unteralgebra ist, welche nach dem dritten Teil unter der komplexen Konjugation stabil ist, also eine \*-Unteralgebra von  $\text{Abb}(\mathbb{N}, \mathbb{C})$  darstellt. Wir werden die Eigenschaften von  $\ell^\infty$  gemäß Proposition 2.3.2 ausgiebig verwenden, ohne jedoch das angemessene Vokabular der \*-Algebren zu vertiefen. Dies bleibt späteren Unternehmungen in der Mathematik vorbehalten.



Wir kommen nun zu den konvergenten Folgen:

**Definition 2.3.4 (Die Folgenräume  $c$  und  $c_0$ )** Der Folgenraum  $c$  ist die Menge aller komplexen konvergenten Folgen in  $\text{Abb}(\mathbb{N}, \mathbb{C})$ , also

$$c = \{(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{C}) \mid (a_n)_{n \in \mathbb{N}} \text{ konvergiert}\}. \quad (2.3.5)$$

Der Folgenraum  $c_0$  ist die Menge der komplexen Nullfolgen

$$c_0 = \{(a_n)_{n \in \mathbb{N}} \in c \mid \lim_{n \rightarrow \infty} a_n = 0\}. \quad (2.3.6)$$

Auch hier gibt es wieder reelle Varianten, die wir aber als Teilmengen von  $c$  beziehungsweise  $c_0$  auffassen können. Wir werden daher kein eigenes Symbol für die reellen Varianten verwenden.

Nach Proposition 2.2.7 wissen wir, dass konvergente Folgen beschränkt sind, womit also

$$c_0 \subseteq c \subseteq \ell^\infty \quad (2.3.7)$$

gilt. Wie schon für  $\ell^\infty$  erhalten wir einfache Rechenregeln auch für konvergente Folgen und ihre Limits:

**Proposition 2.3.5** Seien  $a = (a_n)_{n \in \mathbb{N}}$  und  $b = (b_n)_{n \in \mathbb{N}}$  komplexe konvergente Folgen mit Grenzwerten  $\alpha = \lim_{n \rightarrow \infty} a_n$  und  $\beta = \lim_{n \rightarrow \infty} b_n$ .

i.) Für alle  $z, w \in \mathbb{C}$  gilt  $za + wb \in c$  mit

$$\lim_{n \rightarrow \infty} (za_n + wb_n) = z \lim_{n \rightarrow \infty} a_n + w \lim_{n \rightarrow \infty} b_n. \quad (2.3.8)$$

ii.) Es gilt  $ab \in c$  mit

$$\lim_{n \rightarrow \infty} a_n b_n = \lim_{n \rightarrow \infty} a_n \cdot \lim_{n \rightarrow \infty} b_n. \quad (2.3.9)$$

iii.) Es gilt  $\bar{a} \in c$  mit

$$\lim_{n \rightarrow \infty} \bar{a}_n = \overline{\lim_{n \rightarrow \infty} a_n}. \quad (2.3.10)$$

iv.) Ist  $b_n \neq 0$  für alle  $n \in \mathbb{N}$  und  $\beta \neq 0$ , so gilt  $\frac{a}{b} \in c$  mit

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}. \quad (2.3.11)$$

BEWEIS: Für den ersten Teil wählen wir  $\epsilon > 0$ . Dann finden wir zu  $\epsilon' = \frac{\epsilon}{2(|z|+1)} > 0$  ein  $N$  mit

$$|a_n - \alpha| < \epsilon' = \frac{\epsilon}{2(|z|+1)}, \quad (*)$$

sofern  $n \geq N$ . Genauso finden wir zu  $\epsilon'' = \frac{\epsilon}{2(|w|+1)} > 0$  ein  $N'$  mit

$$|b_n - \beta| < \epsilon'' = \frac{\epsilon}{2(|w|+1)}, \quad (**)$$

sofern  $n \geq N'$ . Wir setzen  $M = \max\{N, N'\}$ , womit beide Ungleichungen  $(*)$  und  $(**)$  gelten, sobald  $n \geq M$ . Diese Argumentation werden wir im Folgenden öfters sehen: Endlich viele Bedingungen können wir durch eventuelles Vergrößern der Schranken immer auch simultan erfüllen. Es gilt nun für  $n \geq M$

$$\begin{aligned} |za_n + wb_n - (z\alpha + w\beta)| &= |z(a_n - \alpha) + w(b_n - \beta)| \\ &\leq |z||a_n - \alpha| + |w||b_n - \beta| \end{aligned}$$

$$\begin{aligned}
&< |z| \frac{\epsilon}{2(|z|+1)} + |w| \frac{\epsilon}{2(|w|+1)} \\
&< \epsilon,
\end{aligned}$$

da ja  $\frac{|z|}{2(|z|+1)} < \frac{1}{2}$  ebenso wie  $\frac{|w|}{2(|w|+1)} < \frac{1}{2}$ . Dies zeigt *i.*). Für den zweiten Teil geben wir wieder  $\epsilon > 0$  beliebig vor. Dann finden wir zu  $\epsilon' = \frac{\epsilon}{2(|\beta|+1)} > 0$  ein  $N$  mit

$$|a_n - \alpha| < \epsilon' = \frac{\epsilon}{2(|\beta|+1)}$$

für alle  $n \geq N$ . Eventuelles Vergrößern von  $N$  wie schon in *i.*) erlaubt es auch, zu  $\epsilon'' = \frac{\epsilon}{2(\|a\|_\infty+1)} > 0$

$$|b_n - \beta| < \epsilon'' = \frac{\epsilon}{2(\|a\|_\infty+1)}$$

für  $n \geq N$  zu schließen. Hier verwenden wir, dass eine konvergente Folge  $(a_n)_{n \in \mathbb{N}}$  beschränkt ist, um  $\|a\|_\infty < \infty$  zum Einsatz zu bringen. Zusammen liefert dies

$$\begin{aligned}
|a_n b_n - \alpha \beta| &= |a_n b_n - a_n \beta + a_n \beta - \alpha \beta| \\
&\leq |a_n| |b_n - \beta| + |\beta| |a_n - \alpha| \\
&\leq \|a\|_\infty |b_n - \beta| + |\beta| |a_n - \alpha| \\
&< \|a\|_\infty \frac{\epsilon}{2(\|a\|_\infty+1)} + |\beta| \frac{\epsilon}{2(|\beta|+1)} \\
&< \epsilon
\end{aligned}$$

für alle  $n \geq N$ , womit auch *ii.*) gezeigt ist. Da  $|\bar{a}_n - \bar{\alpha}| = |a_n - \alpha|$  für alle  $n \in \mathbb{N}$  gilt, ist der dritte Teil klar. Sei also auch für den vierten Teil  $\epsilon > 0$  vorgegeben. Zunächst finden wir ein  $N_1 \in \mathbb{N}$  mit

$$|b_n - \beta| < \frac{1}{2} |\beta|$$

für alle  $n \geq N_1$ , da  $\frac{1}{2} |\beta| > 0$  nach Voraussetzung. Damit gilt

$$|b_n| = |b_n - \beta + \beta| \geq |b_n - \beta| - |\beta| > \frac{1}{2} |\beta|$$

für diese  $n \geq N_1$ . Als nächstes finden wir zu  $\epsilon' = \frac{1}{2} |\beta|^2 \epsilon > 0$  ein  $N_2 \in \mathbb{N}$  mit

$$|b_n - \beta| < \epsilon' = \frac{1}{2} \epsilon \beta^2$$

für alle  $n \geq N_2$ . Sei schließlich  $n \geq N = \max\{N_1, N_2\}$ . Dann gilt

$$\left| \frac{1}{b_n} - \frac{1}{\beta} \right| = \left| \frac{b_n - \beta}{b_n \beta} \right| = \frac{|b_n - \beta|}{|b_n| |\beta|} \leq \frac{2}{|\beta|^2} |b_n - \beta| < \frac{2}{|\beta|^2} 2\epsilon |\beta|^2 = \epsilon$$

für  $n \geq N$ . Dies zeigt, dass die Folge  $\frac{1}{b} = \left(\frac{1}{b_n}\right)_{n \in \mathbb{N}}$  konvergiert mit Grenzwert  $\frac{1}{\beta}$ . Aus diesem Spezialfall von (2.3.11) für  $a = (a_n)_{n \in \mathbb{N}} = (1)_{n \in \mathbb{N}}$  erhalten wir dann den allgemeinen Fall nach *ii.*).  $\square$

**Proposition 2.3.6** Sei  $a = (a_n)_{n \in \mathbb{N}} \in \ell^\infty$  eine beschränkte komplexe Folge und  $b = (b_n)_{n \in \mathbb{N}} \in c_0$  eine komplexe Nullfolge. Dann gilt  $ab \in c_0$ .

BEWEIS: Dies ist vergleichsweise einfach. Da  $b$  eine Nullfolge ist, finden wir zu  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit  $|b_n| < \frac{\epsilon}{\|a\|_\infty+1}$  für alle  $n \geq N$ . Damit erhalten wir

$$|a_n b_n| = |a_n| |b_n| \leq \|a\|_\infty |b_n| < \|a\|_\infty \frac{\epsilon}{\|a\|_\infty+1} < \epsilon$$

für  $n \geq N$ , was  $ab \in c_0$  zeigt.  $\square$

**Bemerkung 2.3.7** Mit Proposition 2.3.5, i.), sehen wir, dass  $c \subseteq \ell^\infty$  ein Untervektorraum ist. Zudem ist

$$\lim_{n \rightarrow \infty} : c \longrightarrow \mathbb{C} \quad (2.3.12)$$

ein lineares Funktional. Dies ist gerade die Aussage von (2.3.8). Teil ii.) und iii.) von Proposition 2.3.5 zeigen nun, dass  $c \subseteq \ell^\infty$  sogar eine \*-Unteralgebra ist. Die Nullfolgen können wir auch als Kern von (2.3.12) auffassen,

$$c_0 = \ker \left( \lim_{n \rightarrow \infty} \right). \quad (2.3.13)$$

Mit (2.3.9) sehen wir, dass das Limes-Funktional (2.3.12) sogar multiplikativ ist. Solche Funktionale nennt man auch *Charaktere*. Schließlich zeigt Proposition 2.3.6, dass  $c_0 \subseteq \ell^\infty$  nicht nur eine Unteralgebra sondern sogar ein Ideal ist. Wir werden diese konzeptuellen Gesichtspunkte vorerst nicht weiter vertiefen, aber ausgiebigen Gebrauch der zugrundeliegenden Rechenregeln machen.

Zusätzlich zu den algebraischen Eigenschaften von  $c$  haben wir auch eine gute Verträglichkeit mit den Ordnungsstrukturen:

**Proposition 2.3.8** Seien  $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \in c$  reelle konvergente Folgen. Gilt dann  $a_n \leq b_n$  für alle  $n \in \mathbb{N}$ , so gilt auch

$$\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n. \quad (2.3.14)$$

BEWEIS: Wir bezeichnen die Grenzwerte mit  $\alpha = \lim_{n \rightarrow \infty} a_n$  und  $\beta = \lim_{n \rightarrow \infty} b_n$ . Wäre  $\alpha > \beta$ , so wäre  $\epsilon = \frac{\alpha - \beta}{2} > 0$ . Dann fänden wir ein  $N$  mit

$$|a_n - \alpha| < \epsilon \quad \text{ebenso wie} \quad |b_n - \beta| < \epsilon$$

für  $n \geq N$ . Damit folgt aber einerseits

$$\alpha - \beta > a_n - \alpha + \beta - b_n$$

und somit andererseits

$$a_n = a_n - \alpha + \alpha - \beta + \beta - b_n + b_n > b_n,$$

ein Widerspruch. □

**Bemerkung 2.3.9** Durch Kombination dieser allgemeinen Eigenschaften lassen sich nun viele weitere nützliche Resultate gewinnen, wie beispielsweise Folgendes: Eine komplexe Folge  $(z_n)_{n \in \mathbb{N}} = (a_n + ib_n)_{n \in \mathbb{N}}$  konvergiert genau dann, wenn die reellen Folgen der Real- und Imaginärteile konvergieren. In diesem Fall gilt

$$\lim_{n \rightarrow \infty} z_n = \lim_{n \rightarrow \infty} a_n + i \lim_{n \rightarrow \infty} b_n \quad (2.3.15)$$

Dies erhält man aus  $\operatorname{Re}(z) = \frac{1}{2}(z + \bar{z})$  sowie  $\operatorname{Im}(z) = \frac{1}{2i}(z - \bar{z})$  und Proposition 2.3.5.

Als letztes Resultat zu Grenzwerten wollen wir Wurzelziehen und allgemeiner Potenzieren betrachten. Wir starten mit folgender Grenzwertbetrachtung:

**Proposition 2.3.10** Sei  $(x_n)_{n \in \mathbb{N}} \in c$  eine konvergente reelle Folge und  $a > 0$ . Dann ist auch  $(a^{x_n})_{n \in \mathbb{N}}$  konvergent und es gilt

$$\lim_{n \rightarrow \infty} a^{x_n} = a^{\lim_{n \rightarrow \infty} x_n}. \quad (2.3.16)$$

BEWEIS: Gilt  $a = 1$ , so ist (2.3.16) trivialerweise richtig, da  $1^x = 1$  für alle  $x \in \mathbb{R}$ . Gilt  $a < 1$ , so ist  $\frac{1}{a} > 1$  und  $a^{x_n} = (\frac{1}{a})^{-x_n}$ . Die Folge  $(-x_n)_{n \in \mathbb{N}}$  konvergiert gegen  $-\lim_{n \rightarrow \infty} x_n$ , womit wir diesen Fall auf den Fall  $a > 1$  zurückgeführt haben. Es genügt daher  $a > 1$  zu betrachten. Wir beginnen mit folgender Vorüberlegung und betrachten die spezielle Nullfolge  $(\frac{1}{n})_{n \in \mathbb{N}}$ . Da  $a^{\frac{1}{n}} > 1$ , finden wir Zahlen  $a_n > 0$  mit  $a^{\frac{1}{n}} = 1 + a_n$ . Daher gilt

$$a = (1 + a_n)^n \geq 1 + na_n$$

nach der Bernoulli-Ungleichung. Es folgt

$$0 \leq a_n \leq \frac{a - 1}{n}$$

für alle  $n \in \mathbb{N}$ . Nach dem Sandwich-Prinzip ist  $a_n$  also eine Nullfolge, was  $a^{\frac{1}{n}} \rightarrow 1$  zeigt, als Spezialfall von (2.3.16). Da  $a^{\frac{1}{n}} > 0$  gilt, können wir auch den Kehrwert betrachten und erhalten damit die Konvergenz  $(\frac{1}{a})^{\frac{1}{n}} = a^{-\frac{1}{n}} \rightarrow 1$ . Den allgemeinen Fall erhalten wir daraus so: Sei  $x = \lim_{n \rightarrow \infty} x_n$  der Grenzwert der Folge  $(x_n)_{n \in \mathbb{N}}$ . Dann ist die neue Folge  $(x_n - x)_{n \in \mathbb{N}}$  eine Nullfolge. Zu jedem  $k \in \mathbb{N}$  gibt es daher ein  $N \in \mathbb{N}$  mit  $|x_n - x| < \frac{1}{k}$ , sofern  $n \geq N$ . Dies bedeutet aber

$$-\frac{1}{k} < x_n - x < \frac{1}{k}.$$

Daraus erhalten wir

$$a^{-\frac{1}{k}} < a^{x_n - x} < a^{\frac{1}{k}} \quad (*)$$

für alle  $n \geq N$ , da Exponenzieren nach Satz 1.6.8, vii.), Ungleichungen erhält. Sei nun  $\epsilon > 0$  vorgegeben. Da  $a^{\frac{1}{k}} \rightarrow 1$  ebenso wie  $a^{-\frac{1}{k}} \rightarrow 1$  gilt, finden wir zunächst ein  $k$  mit

$$|a^{\frac{1}{k}} - 1| < \frac{\epsilon}{2} \quad \text{und} \quad |a^{-\frac{1}{k}} - 1| < \frac{\epsilon}{2},$$

so dass insgesamt  $|a^{-\frac{1}{k}} - a^{\frac{1}{k}}| < \epsilon$  folgt. Für  $N$  gemäß (\*) gilt dann für alle  $n \geq N$

$$|a^{x_n - x} - 1| < \epsilon$$

und daher

$$\lim_{n \rightarrow \infty} a^{x_n - x} = 1.$$

Multiplikation mit  $a^x$  liefert dann (2.3.16) □

**Proposition 2.3.11** Sei  $(a_n)_{n \in \mathbb{N}} \in c$  eine konvergente Folge mit  $a_n > 0$  und  $\alpha = \lim_{n \rightarrow \infty} a_n > 0$ . Dann ist die Folge  $(a_n^x)_{n \in \mathbb{N}}$  für  $x \in \mathbb{R}$  ebenfalls konvergent und

$$\lim_{n \rightarrow \infty} a_n^x = \left( \lim_{n \rightarrow \infty} a_n \right)^x. \quad (2.3.17)$$

Gilt zudem  $x > 0$ , so ist (2.3.17) auch für  $\alpha = 0$  richtig.

BEWEIS: Sei  $\epsilon > 0$  vorgegeben. Dann wählen wir zunächst ein  $M \in \mathbb{N}$  derart, dass  $|2^{\frac{x}{M}} - 1| < \frac{\epsilon}{2}$  ebenso wie  $|1 - 2^{-\frac{x}{M}}| < \frac{\epsilon}{2}$  gilt. Nach der Konvergenz von  $\frac{x}{M} \xrightarrow{M \rightarrow \infty} 0$  und Proposition 2.3.10 ist dies möglich, da ja  $2^{\frac{x}{M}} \rightarrow 1$  ebenso wie  $2^{-\frac{x}{M}} \rightarrow 1$  konvergiert, siehe auch Abbildung 2.3. Wir betrachten dann die Folge  $b_n = \frac{a_n}{\alpha}$ , welche nun gegen 1 konvergiert. An dieser Stelle benutzen wir  $\alpha > 0$ . Da  $b_n \rightarrow 1$  gibt es ein  $N \in \mathbb{N}$  derart, dass

$$2^{-\frac{1}{M}} < b_n < 2^{\frac{1}{M}}$$

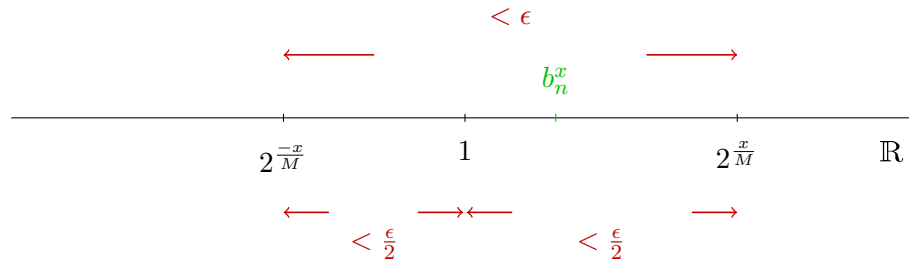


Abbildung 2.3: Die Folge  $b_n$  konvergiert gegen 1. Ist  $x < 0$  so vertauschen die beiden Grenzen ihre Rolle.

für alle  $n \geq N$  gilt. Ist nun  $x > 0$  bleibt diese Ungleichung unter Potenzieren erhalten und liefert

$$2^{-\frac{x}{M}} < b_n^x < 2^{\frac{x}{M}}.$$

Ist  $x < 0$  gilt entsprechend

$$2^{\frac{x}{M}} < b_n^x < 2^{-\frac{x}{M}}.$$

In beiden Fällen folgt aber

$$|b_n^x - 1| < |2^{\frac{x}{M}} - 2^{-\frac{x}{M}}| < \epsilon,$$

da sowohl  $b_n^x$  als auch 1 im Intervall zwischen diesen beiden Grenzwerten liegen und das Intervall nach Wahl von  $M$  eine Länge kleiner  $\epsilon$  besitzt. Dies zeigt (2.3.17), da der verbleibende Fall  $x = 0$  trivial ist. Sei zum Abschluss  $\alpha = 0$  und  $x > 0$ . Es gilt dann also zu zeigen, dass  $(a_n^x)_{n \in \mathbb{N}}$  immer noch eine Nullfolge ist. Zu  $\epsilon > 0$  ist auch  $\epsilon^{\frac{1}{x}} > 0$ . Daher finden wir ein  $N$  mit  $|a_n| = a_n < \epsilon^{\frac{1}{x}}$ , sofern  $n \geq N$ . Für positive Exponenten werden Ungleichungen erhalten, womit  $a_n^x < (\epsilon^{\frac{1}{x}})^x = \epsilon$  folgt. Da sowieso  $a_n^x > 0$  gilt, folgt also  $a_n^x \rightarrow 0$  wie gewünscht.  $\square$

Man beachte, dass für negative  $x \in \mathbb{R}^-$  und eine Nullfolge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n > 0$  die Folge  $(a_n^x)_{n \in \mathbb{N}}$  nicht konvergiert, ja nicht einmal beschränkt ist. Dieser Fall muss also ausgeschlossen werden.

Durch Kombination der vorherigen Rechenregeln können wir nun bereits viele Grenzwerte bestimmen, indem wir unser Wissen über einige wenige Folgen geschickt kombinieren. Beispiele finden sich hierzu in den Übungen. Es gibt allerdings Folgen, für die unsere bisherigen Künste nicht ausreichend sind und daher neue Ideen erfordern. Wir illustrieren dies anhand einiger wichtiger Beispiele.

### Beispiel 2.3.12 (Weitere Grenzwerte)

- i.) Wir betrachten die Folge  $(\sqrt{n+1} - \sqrt{n})_{n \in \mathbb{N}}$ . Da beide Folgen  $(\sqrt{n+1})_{n \in \mathbb{N}}$  ebenso wie  $(\sqrt{n})_{n \in \mathbb{N}}$  gegen  $+\infty$  divergieren, können wir die Rechenregel aus Proposition 2.3.5, i.), nicht anwenden. Diese würde zum unbestimmten „Ergebnis“  $+\infty - \infty = ?$  führen. Die Voraussetzungen zur Proposition 2.3.5 sind eben schlichtweg nicht erfüllt. Wir argumentieren daher anders und formen die Terme zuerst geeignet um. Es gilt

$$\sqrt{n+1} - \sqrt{n} = \frac{n+1-n}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}} \leq \frac{1}{\sqrt{n}}. \quad (2.3.18)$$

Nun können wir unsere gesammelten Rechenregeln verwenden, denn  $(\frac{1}{\sqrt{n}})_{n \in \mathbb{N}}$  ist eine Nullfolge und  $\sqrt{n+1} - \sqrt{n}$  daher mit dem Sandwich-Prinzip ebenfalls. Es gilt also

$$\lim_{n \rightarrow \infty} \sqrt{n+1} - \sqrt{n} = 0. \quad (2.3.19)$$

ii.) Sei  $p \geq 1$  fest gewählt. Dann betrachten wir die Folge  $(\sqrt[p]{n^p})_{n \in \mathbb{N}}$ . Für die Folge  $(\sqrt[p]{a})_{n \in \mathbb{N}}$  wissen wir nach Proposition 2.3.10

$$\lim_{n \rightarrow \infty} \sqrt[p]{a} = a^{\lim_{n \rightarrow \infty} \frac{1}{n}} = a^0 = 1. \quad (2.3.20)$$

Im vorliegenden Falle ist aber der Radikand  $n^p$  selbst von  $n$  abhängig und divergiert gegen  $+\infty$ . Wir können somit weder Proposition 2.3.10 noch Proposition 2.3.11 anwenden. Wir argumentieren daher wie folgt. Zuerst betrachten wir den Fall  $p = 1$  und schreiben  $\sqrt[n]{n} = 1 + a_n$  mit eindeutig bestimmten Zahlen  $a_n \geq 0$ . Für  $n \geq 2$  gilt  $a_n > 0$ . Dann gilt nach dem binomischen Lehrsatz

$$n = (1 + a_n)^n = \sum_{k=0}^n \binom{n}{k} a_n^k \geq \binom{n}{2} a_n^2 = \frac{n(n+1)}{2} a_n^2,$$

womit

$$a_n \leq \sqrt{\frac{2n}{n(n+1)}} = \sqrt{\frac{2}{n+1}} \quad (2.3.21)$$

folgt. Da die rechte Seite eine Nullfolge ist, können wir  $\lim_{n \rightarrow \infty} a_n = 0$  nach dem Sandwich-Prinzip schließen. Es folgt

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1. \quad (2.3.22)$$

Dann erhalten wir aber sofort auch

$$\lim_{n \rightarrow \infty} \sqrt[p]{n^p} = \lim_{n \rightarrow \infty} (\sqrt[n]{n})^p = \left( \lim_{n \rightarrow \infty} \sqrt[n]{n} \right)^p = 1 \quad (2.3.23)$$

für alle  $p \geq 1$  nach Proposition 2.3.11.

iii.) Sei  $a > 1$  und  $p > 0$ . Dann ist  $(n^p)_{n \in \mathbb{N}}$  divergent gegen  $+\infty$  ebenso wie  $(a^n)_{n \in \mathbb{N}}$ . Um zu entscheiden, welche Folge „schneller“ divergiert, bilden wir den Quotienten  $(\frac{n^p}{a^n})_{n \in \mathbb{N}}$ . Eine naive Anwendung von Proposition 2.3.5, iv.), liefert das unbestimmte „Ergebnis“  $\frac{\infty}{\infty} = ?$ . Die Voraussetzungen sind ja aber auch nicht erfüllt, da sowohl die Folge im Zähler als auch die Folge im Nenner divergent sind. Wir müssen also anders argumentieren. Dazu schreiben wir zunächst

$$\frac{n^p}{a^n} = \frac{n^p}{(a^{\frac{n}{p}})^p} = \left( \frac{n}{(a^{\frac{1}{p}})^n} \right)^p. \quad (2.3.24)$$

Da wegen  $p > 0$  immer noch  $a^{\frac{1}{p}} > 1$  gilt, können wir zunächst den Fall  $p = 1$  betrachten. Wir schreiben  $a = 1 + b$  mit  $b > 0$  und erhalten wie in ii.)

$$a^n = (1 + b)^n \geq \frac{1}{2} n(n-1) b^2$$

mit Hilfe des binomischen Lehrsatzes. Daher folgt

$$\frac{n}{a^n} = \frac{n}{\frac{1}{2} n(n-1) b^2} = \frac{2}{(n-1) b^2},$$

womit wir  $\frac{n}{a^n}$  durch eine Nullfolge nach oben abgeschätzt haben. Da auch nach unten  $0 \leq \frac{n}{a^n}$  gilt, folgt nach dem Sandwich-Prinzip

$$\lim_{n \rightarrow \infty} \frac{n}{a^n} = 0. \quad (2.3.25)$$

Aus (2.3.24) erhalten wir dann mit Proposition 2.3.11 für alle  $p > 1$

$$\lim_{n \rightarrow \infty} \frac{n^p}{a^n} = 0. \quad (2.3.26)$$

iv.) Sei nun  $a > 1$  beliebig. Dann betrachten wir die Folge  $(\frac{a^n}{n!})_{n \in \mathbb{N}}$ . In iii.) haben wir gesehen, dass  $a^n$  schneller wächst als jede feste Potenz  $n^p$ . Es stellt sich nun also die Frage, welche Folge hier die Oberhand behält und schneller wächst: die Exponentialfolge  $(a^n)_{n \in \mathbb{N}}$  oder die Fakultät  $(n!)_{n \in \mathbb{N}}$ . Es gibt nun ein  $N \in \mathbb{N}$  mit  $a < N$ . Dieses wählen wir fest, beispielsweise das kleinste solche  $N$ . Dann gilt

$$\frac{a^n}{n!} = \frac{a}{n} \cdot \underbrace{\frac{a}{(n-1)}}_{<1} \cdots \underbrace{\frac{a}{(N+1)}}_{<1} \cdot \underbrace{\frac{a}{N}}_{<1} \cdot \frac{a}{(N-1)} \cdots \frac{a}{1} \leq \frac{a}{n} \cdot \frac{a}{(N-1)} \cdots \frac{a}{1} = \frac{1}{n} \frac{a^N}{(N-1)!}$$

für alle  $n > N$ . Hier erhalten wir die Abschätzung, indem wir Faktoren kleiner 1 weglassen. Da das Produkt  $\frac{a^N}{(N-1)!}$  einen festen Wert besitzt, ist die rechte Seite eine Nullfolge. Damit haben wir die Folge  $(\frac{a^n}{n!})_{n \in \mathbb{N}}$  für alle Folgenglieder mit  $n > N$  durch eine Nullfolge nach oben abgeschätzt. Da sowieso  $0 \leq \frac{a^n}{n!}$  gilt, können wir mit dem Sandwich-Prinzip

$$\lim_{n \rightarrow \infty} \frac{a^n}{n!} = 0 \quad (2.3.27)$$

folgern. Mit anderen Worten wächst die Fakultät  $n!$  schneller gegen  $+\infty$  als jede Exponentialfolge  $a^n$ .

v.) Jetzt erhalten wir durch Kombination einige weitere interessante Grenzwerte, wie etwa

$$\lim_{n \rightarrow \infty} \frac{n^p}{n!} = 0 \quad (2.3.28)$$

für alle  $p > 0$ . Dies ist nun einfach, da wir  $\frac{n^p}{n!} = \frac{n^p}{2^n} \frac{2^n}{n!}$  schreiben können, und beide Faktoren nach iii.) beziehungsweise iv.) Nullfolgen sind.

Bei der Frage nach Konvergenz oder Divergenz einer Folge spielen die ersten Folgenglieder keine Rolle, da die entscheidende Ungleichung (2.2.1) nur für alle  $n$  größer einer gewissen Schranke  $N$  gelten muss. Dies legt folgende Sprechweise nahe:

**Definition 2.3.13 (Fast alle)** Seien  $A(n)$  mathematische Aussagen für alle  $n \in \mathbb{N}$ . Wir sagen, dass fast alle  $A(n)$  gelten, wenn alle bis auf endlich viele der Aussagen  $A(n)$  gelten.

**Bemerkung 2.3.14** Die Definition der Konvergenz der Folge  $(a_n)_{n \in \mathbb{N}}$  gegen  $\alpha \in \mathbb{K}$  können wir also folgendermaßen umschreiben. Für jedes  $\epsilon > 0$  liegen fast alle Folgenglieder näher als  $\epsilon$  bei  $\alpha$ . Es zeigt sich, dass wir bei den Rechenregeln, dem Sandwichprinzip und den übrigen Konvergenzsätzen immer endlich viele Ausnahmen bei den Voraussetzungen an die Folgenglieder machen können, ohne dass die Schlussfolgerungen der Sätze falsch werden. Dies werden wir von nun an stillschweigend tun, siehe auch Übung ??.

## 2.4 Cauchy-Folgen und Vollständigkeit

Unsere Definition einer konvergenten Folge setzt voraus, dass wir bereits einen Kandidaten für den Grenzwert haben. Dies ist in vielen Fällen leider nicht so einfach zu bewerkstelligen. Wir werden sehen, dass viele wichtige Zahlen in der Mathematik, wie etwa die Kreiszahl  $\pi$ , sich als Grenzwerte von bestimmten Folgen definieren lassen. Doch wie soll man nachweisen, dass diese Folgen überhaupt konvergieren, ohne den Grenzwert explizit zu kennen? Diese Frage wird durch den Begriff der Cauchy-Folge beantwortet.

**Definition 2.4.1 (Cauchy-Folge)** Eine Folge  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  heißt *Cauchy-Folge*, falls es zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  gibt, so dass

$$|a_n - a_m| < \epsilon \quad (2.4.1)$$

für alle  $n, m \geq N$  gilt.

Die Idee bei dieser Definition ist, dass spätere Folgenglieder alle nahe beieinander liegen. Der entscheidende Vorteil dieses Begriffs ist, dass wir die Eigenschaft, Cauchy-Folge zu sein, allein mit der Kenntnis der Folgenglieder entscheiden können. Weiterhin beachte man, dass die Definition für jeden angeordneten Körper  $\mathbb{k}$  anstelle von  $\mathbb{R}$  sinnvoll ist und auch für die komplexen Zahlen  $\mathbb{C}$  gestellt werden kann, da wir ja auch hier über einen Betrag  $|\cdot|$  verfügen. Insbesondere zeigt man schnell, dass es für die Cauchy-Eigenschaft genügt,  $\epsilon_n > 0$  mit  $\epsilon_n \rightarrow 0$  zu verwenden, siehe auch Übung 2.9.10 und Übung 2.9.13. Damit kann man sich insbesondere auf  $\epsilon \in \mathbb{Q} \subseteq \mathbb{k}$  beschränken.

Die erste Beobachtung ist, dass konvergente Folgen immer Cauchy-Folgen sind:

**Proposition 2.4.2** Sei  $a = (a_n)_{n \in \mathbb{N}} \in c$  eine konvergente Folge. Dann ist  $a$  auch eine Cauchy-Folge.

BEWEIS: Sei  $\epsilon > 0$  vorgegeben. Wir bezeichnen den Grenzwert von  $a$  mit  $\alpha = \lim_{n \rightarrow \infty} a_n \in \mathbb{K}$ . Dann gibt es also ein  $N \in \mathbb{N}$  mit

$$|a_n - \alpha| < \frac{\epsilon}{2}$$

für alle  $n \geq N$ . Für  $n, m \geq N$  gilt entsprechend

$$|a_n - a_m| = |a_n - \alpha + \alpha - a_m| \leq |a_n - \alpha| + |a_m - \alpha| < \epsilon,$$

womit  $a$  eine Cauchy-Folge ist. □

Umgekehrt können wir immer zeigen, dass Cauchy-Folgen zumindest beschränkt sind:

**Proposition 2.4.3** Sei  $a = (a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Cauchy-Folge. Dann ist  $a$  beschränkt, also  $a \in \ell^\infty$ .

BEWEIS: Wir betrachten beispielsweise  $\epsilon = 1$  in Definition 2.4.1. Dann gibt es also ein  $N \in \mathbb{N}$  mit

$$|a_n - a_m| < 1$$

für alle  $n, m \geq N$ . Insbesondere ist  $|a_n - a_N| < 1$  für alle  $n \geq N$  und daher

$$|a_n| = |a_n - a_N + a_N| \leq |a_n - a_N| + |a_N| < 1 + |a_N|.$$

Damit sind aber fast alle Folgenglieder betragsmäßig durch  $1 + |a_N|$  beschränkt. Die verbleibenden endlich vielen Folgenglieder  $a_1, \dots, a_N$  haben sicherlich ein betragsmäßiges Maximum, so dass  $(a_n)_{n \in \mathbb{N}}$  insgesamt beschränkt ist. □

Wir wollen nun zeigen, dass eine Cauchy-Folge notwendigerweise konvergiert. Dies kann man beispielsweise mit Hilfe einer Intervallschachtelung zeigen:

**Satz 2.4.4 (Cauchy-Folgen sind konvergent)** Sei  $a = (a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Cauchy-Folge. Dann ist  $a$  konvergent in  $\mathbb{K}$ .



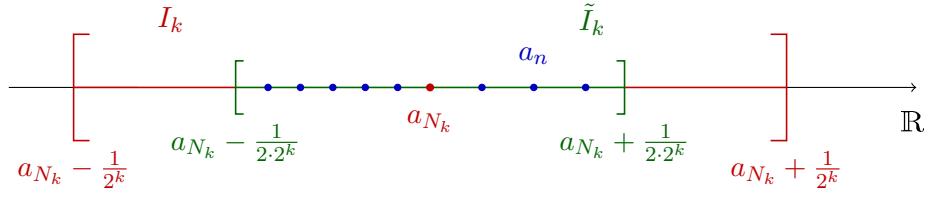


Abbildung 2.4: Die Folgenglieder  $a_n$  liegen im Intervall  $\tilde{I}_k$ , falls  $n \geq N_k$ .

BEWEIS: Wir betrachten zuerst den Fall  $\mathbb{K} = \mathbb{R}$ . Zu jedem  $k \in \mathbb{N}$  finden wir ein  $N_k \in \mathbb{N}$  derart, dass

$$|a_n - a_m| < \frac{1}{2 \cdot 2^k} \quad (*)$$

für alle  $n, m \geq N_k$  gilt. Ohne Einschränkung können wir  $N_{k+1} \geq N_k$  für alle  $k$  wählen. Dies ist die Cauchy-Bedingung für  $\epsilon = \frac{1}{2 \cdot 2^k} > 0$ . Wir betrachten das Intervall

$$I_k = \left[ a_{N_k} - \frac{1}{2^k}, a_{N_k} + \frac{1}{2^k} \right],$$

siehe auch Abbildung 2.4. Mit (\*) folgt insbesondere  $|a_n - a_{N_k}| < \frac{1}{2 \cdot 2^k}$  für alle  $n \geq N_k$ , womit

$$a_n \in \tilde{I}_k = \left[ a_{N_k} - \frac{1}{2 \cdot 2^k}, a_{N_k} + \frac{1}{2 \cdot 2^k} \right] \subseteq I_k.$$

Da der nächste Index  $N_{k+1}$  ein Folgenglied  $a_{N_{k+1}}$  in  $\tilde{I}_k$  liefert, gilt

$$a_{N_{k+1}} \pm \frac{1}{2^{k+1}} \in I_k$$

und somit

$$I_{k+1} \subseteq I_k. \quad (**)$$

Man beachte, dass  $\tilde{I}_{k+1}$  nicht notwendigerweise ganz in  $\tilde{I}_k$  liegen muss, obwohl alle Folgenglieder  $a_n$  mit  $n \geq N_{k+1}$  natürlich auch in  $\tilde{I}_k$  liegen. Mit (\*\*) haben wir also eine Intervallschachtelung konstruiert, wobei die Länge der Intervalle  $I_k$  gerade  $\frac{2}{2^k}$  und damit eine Nullfolge ist. Der Schnitt

$$I = \bigcap_{k=1}^{\infty} I_k$$

ist daher ein abgeschlossenes Intervall nach Satz 1.5.9, *iii.*), beziehungsweise aufgrund der Vollständigkeit von  $\mathbb{R}$ . Die Länge von  $I$  ist aber Null, daher folgt  $I = \{\alpha\}$  für eine reelle Zahl  $\alpha$ . Da  $\alpha \in I_k$  für alle  $k$ , folgt

$$|a_n - \alpha| < \frac{2}{2^k}$$

für alle  $n \geq N_k$ , da ja auch  $a_n \in I_k$  und  $I_k$  die Länge  $\frac{2}{2^k}$  besitzt. Ein gegebenes  $\epsilon > 0$  können wir durch die Nullfolge  $(\frac{2}{2^n})_{n \in \mathbb{N}}$  unterbieten, was dann die Konvergenz  $a_n \rightarrow \alpha$  zeigt. Der komplexe Fall ist nun einfach: aus  $|\operatorname{Re}(z)| \leq |z|$  und  $|\operatorname{Im}(z)| \leq |z|$  folgt sofort, dass zu einer komplexen Cauchy-Folge  $(a_n)_{n \in \mathbb{N}}$  die Folgen  $(\operatorname{Re}(a_n))_{n \in \mathbb{N}}$  und  $(\operatorname{Im}(a_n))_{n \in \mathbb{N}}$  beides reelle Cauchy-Folgen sind. Also konvergieren die Realteil- und Imaginärteilstfolgen, was unmittelbar die Konvergenz von  $(a_n)_{n \in \mathbb{N}}$  zur Folge hat.  $\square$

Überraschenderweise ist die Konvergenz von Cauchy-Folgen für einen archimedischen Körper äquivalent zur Vollständigkeit. Dass Vollständigkeit im Sinne von Definition 1.5.10 die *Cauchy-Vollständigkeit* impliziert, ist gerade die Aussage von Satz 2.4.4. Die Umkehrung ist Gegenstand folgender Proposition, deren Idee bereits im Beweis von Satz 1.5.9, *iii.*)  $\implies$  *i.*) enthalten ist.

**Proposition 2.4.5** *Sei  $\mathbb{k}$  ein archimedisch angeordneter Körper, in dem jede Cauchy-Folge konvergiert. Dann ist  $\mathbb{k}$  vollständig.*

BEWEIS: Wir können dem Beweis von Satz 1.5.9, iii.)  $\implies$  i.) in weiten Teilen folgen. Der Körper ist als archimedisch angenommen. Wir betrachten dann eine nach oben beschränkte, nicht-leere Teilmenge  $A \subseteq \mathbb{k}$ , für die wir ein Supremum konstruieren wollen. Wie im Beweis von Satz 1.5.9 erhalten wir eine Intervallschachtelung  $\{I_n\}_{n \in \mathbb{N}}$  aus abgeschlossenen Intervallen  $I_n = [a_n, b_n]$ , wobei  $a_n \in A$  für alle  $n \in \mathbb{N}$  gilt und  $b_n$  eine obere Schranke von  $A$  ist. Zudem ist die Intervalllänge von  $I_n$  höchstens  $\frac{b_1 - a_1}{2^{n-1}}$  und damit eine Nullfolge. Da  $I_{n+1} \subseteq I_n$  für alle  $n \in \mathbb{N}$  gilt, folgt auch  $a_m \in I_n$  für alle  $m \geq n$ . Da  $|a_n - a_m|$  höchstens die Intervalllänge von  $I_n$  sein kann, bilden die  $(a_n)_{n \in \mathbb{N}}$  eine Cauchy-Folge. Nach Voraussetzung konvergiert diese gegen ein  $\alpha \in \mathbb{k}$ . Wie im Beweis von Satz 1.5.9 zeigt man, dass  $\alpha = \sup A$  das Supremum von  $A$  ist.  $\square$

**Bemerkung 2.4.6 (Cauchy-Vollständigkeit)** Für die Definition von  $\mathbb{R}$  hätten wir alternativ zu Definition 1.5.10 auch die *Cauchy-Vollständigkeit* als Grundlage verwenden können. Man beachte, dass sowohl für die Cauchy-Eigenschaft als auch für Konvergenz nur Zahlen aus  $\mathbb{Q}$  benötigt werden, siehe Übung 2.9.10 und Übung 2.9.13. Es zeigt sich nun, dass Cauchy-Vollständigkeit der tatsächlich viel allgemeinere Begriff von Vollständigkeit ist. Wir haben bereits gesehen, dass  $\mathbb{C}$  ebenfalls Cauchy-vollständig ist, ohne dass wir auch nur ein Analogon zur Definition 1.5.10 für  $\mathbb{C}$  zur Verfügung haben, da  $\mathbb{C}$  kein angeordneter Körper ist, siehe erneut die Kommentare in Bemerkung 1.7.7. Damit ist noch lange nicht die Grenze der Leistungsfähigkeit der Cauchy-Vollständigkeit erreicht. Wir werden noch viele andere Szenarien und mathematische Objekte kennenlernen, die einen sinnvollen Begriff von Cauchy-Folgen erlauben. In diesem Fall ist dann Cauchy-Vollständigkeit ebenso definiert. Schließlich wird es einen sehr allgemeinen Mechanismus geben, wie man Cauchy-Vollständigkeit erzwingen kann, indem man die eventuell noch fehlenden Grenzwerte hinzunimmt. Diese *Cauchy-Vervollständigung* werden wir später noch im Detail kennenlernen.

Aus dem Beweis entnehmen wir noch folgende nützliche Eigenschaft von Suprema:

**Korollar 2.4.7** *Sei  $A \subseteq \mathbb{R}$  eine nicht-leere nach oben beschränkte Teilmenge. Dann gibt es eine Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n \in A$ , sowie*

$$a_n \leq a_{n+1} \quad (2.4.2)$$

für alle  $n \in \mathbb{N}$  und

$$\lim_{n \rightarrow \infty} a_n = \sup A. \quad (2.4.3)$$

BEWEIS: Die unteren Intervallgrenzen der Intervallschachtelung im Beweis von Proposition 2.4.5 haben genau diese Eigenschaften.  $\square$

Folgen mit der Eigenschaft 2.4.2 nennt man monoton wachsend:

**Definition 2.4.8 (Monotone Folgen)** *Sei  $a = (a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{R})$  eine reelle Folge.*

i.) *Die Folge  $a$  heißt monoton wachsend (fallend), wenn*

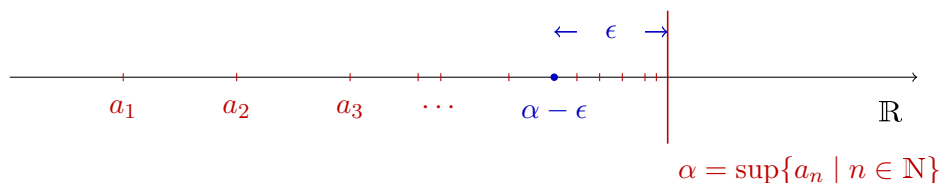
$$a_n \leq a_{n+1} \quad \text{beziehungsweise} \quad a_n \geq a_{n+1} \quad (2.4.4)$$

für alle  $n \in \mathbb{N}$ .

ii.) *Die Folge  $a$  heißt streng monoton wachsend (fallend), wenn*

$$a_n < a_{n+1} \quad \text{beziehungsweise} \quad a_n > a_{n+1} \quad (2.4.5)$$

für alle  $n \in \mathbb{N}$ .

Abbildung 2.5: Eine monoton wachsende Folge konvergiert gegen  $\sup\{a_n \mid n \in \mathbb{N}\}$ .

Die Beobachtung aus Korollar 2.4.7 hat eine Umkehrung in folgendem Sinne. Für monotone Folgen können wir leicht entscheiden, wann sie konvergieren:

**Proposition 2.4.9** *Eine monotone Folge  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{R})$  konvergiert genau dann, wenn sie beschränkt ist.*

BEWEIS: Eine konvergente Folge ist immer beschränkt nach Proposition 2.2.7. Sei also umgekehrt  $(a_n)_{n \in \mathbb{N}}$  monoton und beschränkt. Wir betrachten den Fall einer monoton wachsenden Folge, den monoton fallenden Fall zeigt man analog. Da  $(a_n)_{n \in \mathbb{N}}$  beschränkt ist, gibt es das Supremum

$$\alpha = \sup\{a_n \mid n \in \mathbb{N}\}.$$

Ist nun  $\epsilon > 0$ , so gibt es ein  $N \in \mathbb{N}$  mit

$$a_N \geq \alpha - \epsilon,$$

da  $\alpha$  die kleinste obere Schranke ist. Da  $(a_n)_{n \in \mathbb{N}}$  aber monoton wachsend ist, gilt damit auch

$$a_n \geq a_N \geq \alpha - \epsilon$$

für alle  $n \geq N$ , dies zeigt man induktiv mit  $a_{n+1} \geq a_n$ . Da andererseits aber  $a_n \leq \alpha$  für alle  $n \in \mathbb{N}$  gilt, folgt

$$|a_n - \alpha| < \epsilon$$

für  $n \geq N$ , siehe auch Abbildung 2.5. Dies ist die behauptete Konvergenz.  $\square$

Mit der bereits getroffenen Vereinbarung zur Divergenz gegen  $\pm\infty$  gilt das Resultat auch für divergente monotone Folgen in dem Sinne, dass eine divergente monoton wachsende Folge  $(a_n)_{n \in \mathbb{N}}$  gegen  $+\infty$  divergiert. Wir haben also sowohl im konvergenten als auch im divergenten Fall

$$\lim_{n \rightarrow \infty} a_n = \sup\{a_n \mid n \in \mathbb{N}\} \quad (2.4.6)$$

für jede monoton wachsende Folge. Ist  $(a_n)_{n \in \mathbb{N}}$  entsprechend monoton fallend, gilt

$$\lim_{n \rightarrow \infty} a_n = \inf\{a_n \mid n \in \mathbb{N}\}, \quad (2.4.7)$$

wobei wieder  $\inf\{a_n \mid n \in \mathbb{N}\} = -\infty$  zugelassen ist, was genau dann auftritt, wenn die monoton fallende Folge nach unten unbeschränkt ist und damit gegen  $-\infty$  divergiert.

Zum Abschluss dieses Abschnitts zu Cauchy-Folgen zeigen wir folgenden fundamentalen Satz von Bolzano und Weierstraß:

**Satz 2.4.10 (Bolzano-Weierstraß)** *Sei  $a = (a_n)_{n \in \mathbb{N}} \in \ell^\infty$  eine beschränkte Folge. Dann besitzt  $(a_n)_{n \in \mathbb{N}}$  eine konvergente Teilfolge.*

BEWEIS: Wir betrachten zunächst den Fall  $\mathbb{K} = \mathbb{R}$ . Dann gilt also  $-\|a\|_\infty \leq a_n \leq \|a\|_\infty$  für alle  $n \in \mathbb{N}$ . Wir halbieren das Intervall  $[-\|a\|_\infty, \|a\|_\infty]$ . Da die Folge insgesamt unendlich viele Folgenglieder besitzt, enthält mindestens eines der beiden Intervalle  $[-\|a\|_\infty, 0]$  oder  $[0, \|a\|_\infty]$  ebenfalls unendlich viele Folgenglieder. Wir bezeichnen dieses (oder die Wahl eines, falls beide unendlich viele Folgenglieder enthalten) mit  $I_1$ . Dann halbieren wir  $I_1$  und erhalten wieder ein abgeschlossenes Intervall  $I_2 \subseteq I_1$  der halben Länge wie  $I_1$ , welches unendlich viele Folgenglieder enthält. Induktiv erhalten wir also abgeschlossene Intervalle  $I_1 \supseteq I_2 \supseteq \cdots \supseteq I_k \supseteq I_{k+1} \supseteq \cdots$  mit jeweils halbiertem Länge, so dass jedes  $I_k$  unendlich viele Folgenglieder enthält. Nach dem Intervallschachtelungsprinzip folgt also, dass der Durchschnitt

$$\bigcap_{k=1}^{\infty} I_k = \{\alpha\}$$

genau eine reelle Zahl enthält. Wir wählen nun induktiv in jedem Intervall  $I_k$  dasjenige Folgenglied  $a_{n_k}$  aus, welches den kleinsten Index  $n_k$  mit  $n_k > n_{k-1}$  besitzt. Da wir in  $I_{k+1}$  ja wieder unendlich viele Folgenglieder zur Auswahl haben, gibt es solche Folgenglieder und eines hat den kleinsten Index. Dies definiert eine Teilfolge  $(a_{n_k})_{k \in \mathbb{N}}$ , welche wegen  $a_{n_k} \in I_k$  die Eigenschaft

$$|a_{n_k} - \alpha| \leq \frac{2\|a\|_\infty}{2^k}$$

besitzt, da die Länge von  $I_k$  gerade  $\frac{2\|a\|_\infty}{2^k}$  ist. Da aber alle nachfolgenden Elemente ebenfalls in  $I_k$  liegen, folgt sogar

$$|a_{n_\ell} - \alpha| \leq \frac{2\|a\|_\infty}{2^k},$$

sofern nur  $\ell \geq k$ . Dies liefert sofort die Konvergenz  $a_{n_k} \rightarrow \alpha$  für  $k \rightarrow \infty$ .

Im Fall  $\mathbb{K} = \mathbb{C}$  ist die Realteilfolge  $(\operatorname{Re}(a_n))_{n \in \mathbb{N}}$  ebenso wie die Imaginärteilfolge  $(\operatorname{Im}(a_n))_{n \in \mathbb{N}}$  beschränkt. Wir finden also eine Teilfolge  $(a_{n_k})_{k \in \mathbb{N}}$ , so dass  $(\operatorname{Re}(a_{n_k}))_{k \in \mathbb{N}}$  konvergiert. Da  $(\operatorname{Im}(a_{n_k}))_{k \in \mathbb{N}}$  immer noch beschränkt ist, finden wir eine weitere Teilfolge  $(a_{n_{k_j}})_{j \in \mathbb{N}}$ , so dass  $(\operatorname{Im}(a_{n_{k_j}}))_{j \in \mathbb{N}}$  konvergiert. Dann bleibt aber auch  $(\operatorname{Re}(a_{n_{k_j}}))_{j \in \mathbb{N}}$  als Teilfolge einer konvergenten Folge konvergent, so dass insgesamt  $(a_{n_{k_j}})_{j \in \mathbb{N}}$  konvergent ist.  $\square$

Es ist also kein Zufall, dass die divergente aber beschränkte Folge  $((-1)^n)_{n \in \mathbb{N}}$  konvergente Teilfolgen besitzt. In diesem Beispiel erhalten wir Teilfolgen die gegen 1 konvergieren oder gegen -1 konvergieren.

## 2.5 Häufungspunkte und Umgebungen

Bei einer beschränkten Folge gibt es nach dem Satz von Bolzano-Weierstraß konvergente Teilfolgen. Es stellt sich also die Frage, welche Zahlen als Grenzwerte von konvergenten Teilfolgen auftreten können. Diese Frage ist immer noch sinnvoll, selbst wenn die ursprüngliche Folge nicht beschränkt ist. Dann kann es natürlich sein, dass sie keine konvergenten Teilfolgen besitzt.

Um dieser Frage nachzugehen, benötigen wir den Begriff der *Umgebung*.

**Definition 2.5.1 (Umgebung, offene und abgeschlossene Teilmengen)** Sei  $z \in \mathbb{K}$ .

i.) Sei  $r > 0$ . Dann heißt die Teilmenge

$$B_r(z) = \{w \in \mathbb{K} \mid |z - w| < r\} \quad (2.5.1)$$

die offene Kugel (englisch ball) vom Radius  $r$  um  $z$ . Die Teilmenge

$$B_r(z)^{\text{cl}} = \{w \in \mathbb{K} \mid |z - w| \leq r\} \quad (2.5.2)$$

heißt abgeschlossene Kugel vom Radius  $r$  um  $z$ .

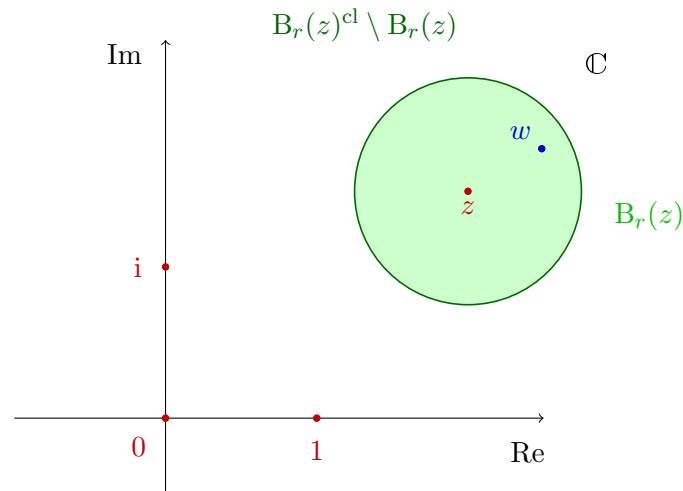


Abbildung 2.6: In der komplexen Ebene ist die „Kugel“ vom Radius  $r$  eine Kreisscheibe. Bei der offenen Kreisscheibe ist der Kreis selbst nicht Teil von  $B_r(z)$ . Bei  $B_r(z)^{cl}$  gehört die Kreislinie mit dazu.

ii.) Eine Teilmenge  $U \subseteq \mathbb{K}$  heißt Umgebung von  $z$ , falls es ein  $r > 0$  gibt, sodass

$$B_r(z) \subseteq U \quad (2.5.3)$$

gilt.

iii.) Eine Teilmenge  $O \subseteq \mathbb{K}$  heißt offen, wenn sie Umgebung aller ihrer Punkte  $z \in O$  ist.

iv.) Eine Teilmenge  $A \subseteq \mathbb{K}$  heißt abgeschlossen, falls  $\mathbb{K} \setminus A$  offen ist.

Für  $\mathbb{K} = \mathbb{R}$  sind die offenen Kugeln gerade die offenen Intervalle

$$B_r(z) = (z - r, z + r). \quad (2.5.4)$$

Im komplexen Fall  $\mathbb{K} = \mathbb{C}$  sind es offene Kreisscheiben um  $z$  vom Radius  $r$ , siehe auch Abbildung 2.6.

**Proposition 2.5.2** Sei  $z \in \mathbb{K}$ .

- i.) Ist  $U \subseteq \mathbb{K}$  eine Umgebung von  $z$  und  $U \subseteq U'$ , so ist auch  $U'$  eine Umgebung von  $z$ .
- ii.) Sind  $U_1, \dots, U_n \subseteq \mathbb{K}$  endlich viele Umgebungen von  $z$ , so ist auch  $U_1 \cap \dots \cap U_n$  eine Umgebung von  $z$ .
- iii.) Eine offene Kugel  $B_r(z)$  ist offen.
- iv.) Die leere Menge  $\emptyset \subseteq \mathbb{K}$  und  $\mathbb{K}$  sind offen.
- v.) Sind  $\{O_i\}_{i \in I}$  offene Teilmengen von  $\mathbb{K}$ , so ist auch  $\bigcup_{i \in I} O_i \subseteq \mathbb{K}$  offen.
- vi.) Sind  $O_1, \dots, O_n \subseteq \mathbb{K}$  endlich viele offene Teilmengen, so ist auch  $O_1 \cap \dots \cap O_n \subseteq \mathbb{K}$  offen.
- vii.) Abgeschlossene Kugeln  $B_r(z)^{cl}$  sind abgeschlossen. Zudem ist  $\{z\} \subseteq \mathbb{K}$  eine abgeschlossene Teilmenge.
- viii.) Ist  $U \subseteq \mathbb{K}$  eine Umgebung von  $z$ , so gibt es eine Umgebung  $V \subseteq U$  von  $z$ , die offen ist.

BEWEIS: Der erste Teil ist einfach: Ist  $U$  eine Umgebung von  $z$ , so gibt es ein  $r > 0$  mit  $B_r(z) \subseteq U$ . Ist nun  $U \subseteq U'$ , so gilt erst recht  $B_r(z) \subseteq U'$ . Sind  $U_1, \dots, U_n$  endlich viele Umgebungen von  $z$ , so gibt es  $r_1, \dots, r_n > 0$  mit  $B_{r_i}(z) \subseteq U_i$  für  $i = 1, \dots, n$ . Sei

$$r = \min\{r_1, \dots, r_n\},$$

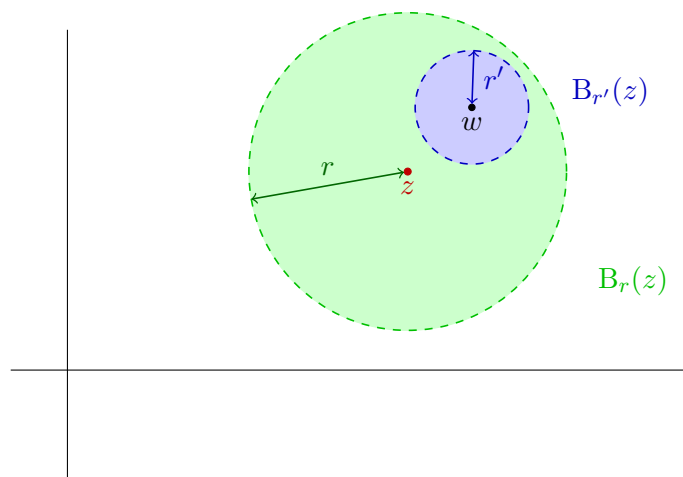


Abbildung 2.7: Die offene Kugel um  $z$  enthält kleinere offene Kugeln um jeden Punkt  $w \in B_r(z)$ , wobei der Radius  $r'$  entsprechend klein genug gewählt werden muss.

dann gilt  $B_r(z) \subseteq B_{r_i}(z) \subseteq U_i$  für alle  $i = 1, \dots, n$ , also  $B_r(z) \subseteq U_1 \cap \dots \cap U_n$ , was *ii.*) zeigt. Der dritte Teil ist interessanter: Sei  $r > 0$  und  $w \in B_r(z)$ . Dann ist  $r - |z - w| > 0$  und wir finden ein  $r' > 0$  mit  $r' < r - |z - w|$ . Für  $u \in B_{r'}(z)$  gilt dann

$$|z - u| \leq |z - w| + |w - u| < |z - w| + r' < r,$$

womit  $B_{r'}(z) \subseteq B_r(z)$ , siehe auch Abbildung 2.7. Damit ist  $B_r(z)$  aber eine Umgebung von  $w$  für alle  $w \in B_r(z)$ , also offen im Sinne der Definition 2.5.1, *iii.*). Da die leere Menge keine Punkte enthält, ist sie Umgebung aller ihrer Punkte. Die Offenheit von  $\mathbb{K}$  ist klar. Sei nun  $\{O_i\}_{i \in I}$  eine beliebige Menge von offenen Teilmengen und  $z \in \bigcup_{i \in I} O_i$ . Dann gibt es (mindestens) ein  $i_0 \in I$  mit  $z \in O_{i_0}$ . Da  $O_{i_0}$  offen ist, gibt es ein  $r > 0$  mit  $B_r(z) \subseteq O_{i_0}$ . Es folgt  $B_r(z) \subseteq \bigcup_{i \in I} O_i$ , was die Offenheit der Vereinigung zeigt, da  $z$  beliebig war. Der sechste Teil folgt aus *ii.*): Ist  $z \in O_1 \cap \dots \cap O_n$ , so gilt  $z \in O_i$  für alle  $i = 1, \dots, n$ . Da  $O_i$  offen ist, ist  $O_i$  eine Umgebung von  $z$ . Nach *ii.*) ist  $O_1 \cap \dots \cap O_n$  dann auch eine Umgebung von  $z$ . Da dies für alle  $z \in O_1 \cap \dots \cap O_n$  gilt, ist  $O_1 \cap \dots \cap O_n$  offen. Für *vii.*) betrachten wir  $w \in \mathbb{K} \setminus B_r(z)^{\text{cl}}$ . Dies bedeutet  $|z - w| > r$ . Damit gibt es aber ein  $r' > 0$  mit  $r' < |z - w| - r$ . Für  $u \in B_{r'}(w)$  gilt dann

$$|z - u| \leq |z - w| + |w - u| < |z - w| + r' < |z - w| - r,$$

womit  $|z - u| > r$  folgt, also  $u \notin B_r(z)^{\text{cl}}$ . Damit ist aber  $u \in \mathbb{K} \setminus B_r(z)^{\text{cl}}$  und somit  $B_{r'}(w) \subseteq \mathbb{K} \setminus B_r(z)^{\text{cl}}$ , was die Offenheit von  $\mathbb{K} \setminus B_r(z)^{\text{cl}}$  zeigt, siehe auch Abbildung ???. Dies bedeutet, dass  $B_r(z)^{\text{cl}}$  abgeschlossen ist. Die Argumentation ist offenbar immer noch richtig, falls  $r = 0$  gilt. Dies entspricht dann der einpunktigen Teilmenge  $\{z\}$ . Der letzte Teil ist klar, da nach Definition einer Umgebung  $U$  von  $z$  ein  $r > 0$  existiert, so dass  $B_r(z) \subseteq U$  gilt. Nach *iii.*) ist  $B_r(z)$  offen, womit  $V = B_r(z)$  allen Anforderungen genügt.  $\square$

**Bemerkung 2.5.3** Wir werden sowohl die Begriffsbildungen von Definition 2.5.1 als auch die Resultate von Proposition 2.5.2 in einen größeren Kontext stellen, wenn wir uns in Kapitel 7 metrischen und topologischen Räumen zuwenden. Weitere Eigenschaften von offenen und abgeschlossenen Teilmengen von  $\mathbb{K}$  werden in Übung 2.9.14 und Übung 2.9.15 diskutiert.

Mit dem neuen Begriff der Umgebung können wir die Konvergenz von Folgen umformulieren:

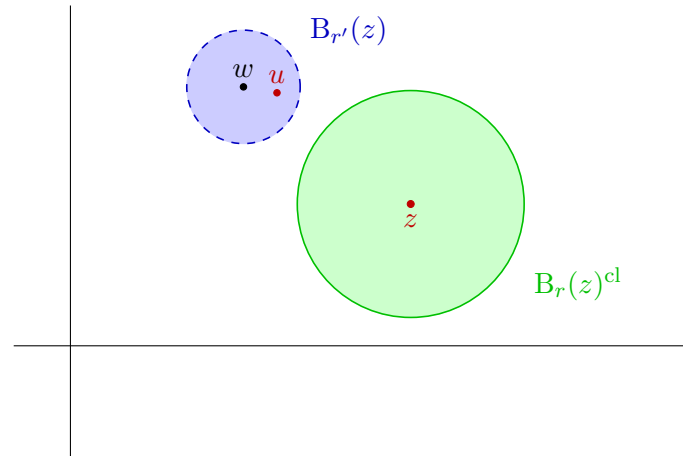


Abbildung 2.8: Im Komplement von  $B_r(z)^{cl}$  finden wir eine offene Kugel um  $w \in \mathbb{K} \setminus B_r(z)^{cl}$ .

**Proposition 2.5.4** Sei  $a = (a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge. Die Folge  $a$  konvergiert genau dann gegen  $\alpha \in \mathbb{K}$ , wenn jede Umgebung von  $\alpha$  fast alle Folgenglieder von  $a$  enthält.

BEWEIS: Sei zunächst  $(a_n)_{n \in \mathbb{N}}$  konvergent gegen  $\alpha$ . Für eine Umgebung  $U \subseteq \mathbb{K}$  von  $\alpha$  finden wir dann ein  $r > 0$  mit  $B_r(\alpha) \subseteq U$ . Verwenden wir  $\epsilon = r$  in der Definition von Konvergenz, so gibt es also ein  $N \in \mathbb{N}$  mit  $|a_n - \alpha| < \epsilon$  sofern  $n \geq N$ . Dies bedeutet aber  $a_n \in B_r(\alpha) \subseteq U$  für alle  $n \geq N$ . Damit sind fast alle Folgenglieder von  $a$  in  $U$  enthalten. Für die umgekehrte Richtung geben wir  $\epsilon > 0$  vor. Nach Voraussetzung sind dann fast alle Folgenglieder von  $a$  in der (offenen) Umgebung  $B_\epsilon(\alpha)$  enthalten, also alle bis auf endlich viele. Diese endlich vielen seien  $a_{n_1}, \dots, a_{n_k}$  für geeignete Indizes  $n_1, \dots, n_k \in \mathbb{N}$ . Setzen wir  $N = \max\{n_1, \dots, n_k\} + 1$ , so sind alle  $a_n$  mit  $n \geq N$  in  $B_\epsilon(\alpha)$  enthalten, was gerade  $|a_n - \alpha| < \epsilon$  bedeutet. Dies ist die Konvergenz von  $a$  gegen  $\alpha$ .  $\square$

Abgeschlossene Teilmengen können wir mit Hilfe von konvergenten Folgen charakterisieren.

**Proposition 2.5.5** Eine Teilmenge  $A \subseteq \mathbb{K}$  ist genau dann abgeschlossen, wenn für jede konvergente Folge  $a = (a_n)_{n \in \mathbb{N}}$  in  $A$ , also  $a_n \in A$  für alle  $n \in \mathbb{N}$ , auch der Grenzwert  $\alpha = \lim_{n \rightarrow \infty} a_n \in A$  in  $A$  liegt.

BEWEIS: Sei zunächst  $A \subseteq \mathbb{K}$  abgeschlossen. Um Trivialitäten vorzubeugen, können wir  $A \neq \emptyset$  annehmen. Sei dann also eine Folge  $(a_n)_{n \in \mathbb{N}}$  mit Folgengliedern in  $A$  vorgegeben, die gegen  $\alpha = \lim_{n \rightarrow \infty} a_n \in \mathbb{K}$  konvergiert. Wir nehmen an, dass  $\alpha \notin A$ , also  $\alpha \in \mathbb{K} \setminus A$ . Dies ist nach Definition eine offene Teilmenge, womit es ein  $r > 0$  mit  $B_r(\alpha) \subseteq \mathbb{K} \setminus A$  gibt. Also ist  $B_r(\alpha) \cap A = \emptyset$ . Die Konvergenz von  $(a_n)_{n \in \mathbb{N}}$  gegen  $\alpha$  bedeutet aber, dass fast alle Folgenglieder von  $(a_n)_{n \in \mathbb{N}}$  in der Umgebung  $B_r(\alpha)$  liegen müssen, in klarem Widerspruch zu  $a_n \in A$ . Für die Umkehrung müssen wir zeigen, dass  $\mathbb{K} \setminus A$  offen ist. Sei  $z \in \mathbb{K} \setminus A$ . Dann ist  $z$  nicht Grenzwert einer Folge  $(a_n)_{n \in \mathbb{N}}$  von Punkten  $a_n \in A$ . Dann gibt es ein  $n_0 \in \mathbb{N}$  derart, dass  $B_{\frac{1}{n_0}}(z) \cap A = \emptyset$ . Andernfalls gäbe es zu jedem  $n \in \mathbb{N}$  ein  $a_n \in B_{\frac{1}{n}}(z) \cap A$ . Dies definiert nun eine Folge, welche zum einen in  $A$  liegt, zum anderen aber  $a_n \in B_{\frac{1}{n}}(z)$  für alle  $n \in \mathbb{N}$  erfüllt. Da  $B_{\frac{1}{n+1}}(z) \subseteq B_{\frac{1}{n}}(z)$  gilt, folgt dass  $(a_n)_{n \in \mathbb{N}}$  gegen  $z$  konvergiert, ein Widerspruch. Also gibt es ein  $n_0 \in \mathbb{N}$  mit  $B_{\frac{1}{n_0}}(z) \cap A = \emptyset$  und damit  $B_{\frac{1}{n_0}}(z) \subseteq \mathbb{K} \setminus A$ . Dies zeigt die Offenheit von  $\mathbb{K} \setminus A$  und damit die Abgeschlossenheit von  $A$ .  $\square$

Wir kommen nun zum Begriff des Häufungspunkts einer Folge. Als Motivation dient Proposition 2.5.4, die Grenzwerte charakterisiert. Diesen Begriff wollen wir abschwächen, so dass beispielsweise die beiden Punkte  $\pm 1$  nun Häufungspunkte von  $((-1)^n)_{n \in \mathbb{N}}$  werden.

**Definition 2.5.6 (Häufungspunkt einer Folge)** Sei  $a = (a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge. Dann heißt  $\alpha \in \mathbb{K}$  Häufungspunkt von  $a$ , wenn jede Umgebung von  $\alpha$  unendlich viele Folgenglieder von  $a$  enthält. Die Menge der Häufungspunkte von  $a$  wird mit  $\text{HP}(a)$  bezeichnet.

Im Hinblick auf Proposition 2.5.4 ist dies tatsächlich eine Abschwächung des Grenzwertbegriffs. Insbesondere ist ein (der!) Grenzwert immer ein Häufungspunkt und sogar der einzige, sofern die Folge überhaupt konvergiert:

**Proposition 2.5.7** Eine konvergente Folge  $(a_n)_{n \in \mathbb{N}}$  besitzt genau einen Häufungspunkt, den Grenzwert.

BEWEIS: Nach Proposition 2.5.4 ist der Grenzwert sicherlich ein Häufungspunkt. Ist nun  $x \in \mathbb{K}$  ein Häufungspunkt von  $(a_n)_{n \in \mathbb{N}}$ , der nicht gleich dem Grenzwert  $\alpha = \lim_{n \rightarrow \infty} a_n$  ist, so gibt es unendlich viele Folgenglieder in der offenen Kugel  $B_\epsilon(x)$  mit  $\epsilon = \frac{|x-\alpha|}{2} > 0$ . Diese liegen dann aber nicht in der offenen Kugel  $B_\epsilon(\alpha)$ , womit diese nicht fast alle Folgenglieder enthält, ein Widerspruch.  $\square$

**Beispiel 2.5.8** Es gibt aber auch durchaus divergente Folgen mit nur einem Häufungspunkt. Die Folge  $(a_n)_{n \in \mathbb{N}}$  mit

$$a_n = \begin{cases} n & \text{für } n \text{ gerade} \\ \frac{1}{n} & \text{für } n \text{ ungerade} \end{cases} \quad (2.5.5)$$

ist sicherlich divergent, da sie unbeschränkt ist. Trotzdem ist  $0 \in \mathbb{K}$  ein Häufungspunkt. Es ist nun einfach zu sehen, dass es keinen weiteren Häufungspunkt gibt. Die Aussage der Proposition lässt sich also ohne weitere Voraussetzungen nicht verschärfen.

Eine äquivalente Formulierung mit einer einfachen Möglichkeit, auf Häufungspunkte zu testen, erhält man durch Teilfolgen:

**Proposition 2.5.9** Sei  $a = (a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge. Eine Zahl  $\alpha \in \mathbb{K}$  ist genau dann ein Häufungspunkt von  $a$ , falls es eine Teilfolge von  $a$  gibt, die  $\alpha$  als Grenzwert besitzt.

BEWEIS: Sei zunächst  $\alpha$  ein Häufungspunkt. Zu  $k \in \mathbb{N}$  gibt es dann unendlich viele Folgenglieder  $a_n \in B_{\frac{1}{k}}(\alpha)$ . Wir wählen ein  $a_{n_1} \in B_1(\alpha)$  und induktiv

$$a_{n_k} \in B_{\frac{1}{k}}(\alpha)$$

derart, dass  $n_k$  der kleinste Index mit  $n_k > n_{k-1}$  ist. Da es unendlich viele Folgenglieder in  $B_{\frac{1}{n}}(\alpha)$  gibt, gibt es auch eines mit kleinstem Index. Dies definiert dann eine Teilfolge  $(a_{n_k})_{k \in \mathbb{N}}$  der ursprünglichen Folge  $a$ , welche nun gegen  $\alpha$  konvergiert: In jedem  $\epsilon$ -Ball  $B_\epsilon(\alpha)$  liegt ein offener Ball der Form  $B_{\frac{1}{k_0}}(\alpha)$  mit  $k_0$  groß genug. Dann liegen fast alle Folgenglieder von  $(a_{n_k})_{k \in \mathbb{N}}$  in  $B_{\frac{1}{k_0}} \subseteq B_\epsilon(\alpha)$ , was die Konvergenz zeigt. Ist umgekehrt  $(a_{n_k})_{k \in \mathbb{N}}$  eine Teilfolge von  $a$ , welche gegen  $\alpha$  konvergiert, so liegen zu jeder offenen Umgebung von  $\alpha$  fast alle Folgenglieder von  $(a_{n_k})_{k \in \mathbb{N}}$  in dieser Umgebung. Dies sind aber unendlich viele, womit  $\alpha$  ein Häufungspunkt ist.  $\square$

Aus dem Satz von Bolzano-Weierstraß erhält man damit sofort folgende Existenz von Häufungspunkten:

**Korollar 2.5.10** Sei  $a = (a_n)_{n \in \mathbb{N}} \in \ell^\infty$  eine beschränkte Folge. Dann hat  $a$  mindestens einen Häufungspunkt.

Es stellt sich also nun die Frage, wieviele Häufungspunkte eine Folge besitzen kann. Das zuvor betrachtete Beispiel der Folge  $a = ((-1)^n)_{n \in \mathbb{N}}$  hat offenbar genau zwei Häufungspunkte, nämlich  $+1$  und  $-1$ , womit also durchaus mehr als nur ein Häufungspunkt auftreten kann. Der folgenden Satz zeigt, dass die Menge der Häufungspunkte sehr allgemein sein kann:



**Satz 2.5.11 (Häufungspunkte)**

- i.) Sei  $a = (a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge. Dann ist die Menge  $\text{HP}(a)$  der Häufungspunkte von  $a$  abgeschlossen.
- ii.) Ist  $A \subseteq \mathbb{K}$  eine abgeschlossene Teilmenge von  $\mathbb{K}$ , so gibt es eine Folge  $a = (a_n)_{n \in \mathbb{N}}$  in  $\mathbb{K}$ , die genau diese Menge als Häufungspunkte besitzt.

BEWEIS: Um zu zeigen, dass  $\text{HP}(a)$  abgeschlossen ist, zeigen wir, dass das Komplement  $\mathbb{K} \setminus \text{HP}(a)$  offen ist. Sei also  $z \in \mathbb{K} \setminus \text{HP}(a)$ . Da  $z$  kein Häufungspunkt ist, gibt es also eine Umgebung  $U$  von  $z$ , die nur endlich viele Folgenglieder, etwa  $a_{i_1}, \dots, a_{i_k}$  enthält. Durch verkleinern von  $U$  können wir erreichen, dass  $U$  bereits offen ist. Ist dann  $w \in U$  ein weiterer Punkt, so ist  $U$  auch eine Umgebung von  $w$ , die immer noch nur endlich viele Folgenglieder enthält. Also ist  $w$  ebenfalls kein Häufungspunkt. Damit folgt aber  $U \subseteq \mathbb{K} \setminus \text{HP}(a)$ . Da wir zu jedem Punkt  $z \in \mathbb{K} \setminus \text{HP}(a)$  eine solche Umgebung finden können, ist  $\mathbb{K} \setminus \text{HP}(a)$  offen. Der zweite Teil stellt gewissermaßen die Umkehrung dar. Sei also  $A \subseteq \mathbb{K}$  abgeschlossen. Ist  $A = \emptyset$ , so leistet die Folge  $a = (n)_{n \in \mathbb{N}}$  das Gewünschte, da diese keine Häufungspunkte besitzt. Wir können also  $A \neq \emptyset$  annehmen. Als kleine Vorüberlegung zeigen wir, dass es zu  $a \in A$  eine Folge  $(z_n)_{n \in \mathbb{N}}$  in  $\mathbb{K}$  gibt, die zum einen gegen  $a$  konvergiert, und zum anderen aus rationalen Zahlen  $z_n \in \mathbb{Q}$  beziehungsweise aus komplexen Zahlen von der Form  $z_n = q_n + iq'_n \in \mathbb{Q} + i\mathbb{Q}$  besteht. Für den reellen Fall ist dies klar, da wir beispielsweise ein  $N \in \mathbb{N}$  mit  $a \in [-N, N]$  finden und durch Intervallhalbierung eine solche Folge konstruieren können. Den allgemeineren komplexen Fall erhalten wir dann, indem wir Real- und Imaginärteil von  $a$  getrennt betrachten und das gerade gezeigte Resultat für reelles  $a$  verwenden. Wir definieren nun

$$Q = \left\{ \begin{array}{ll} \mathbb{Q} & \text{falls } \mathbb{K} = \mathbb{R} \\ \mathbb{Q} + i\mathbb{Q} & \text{falls } \mathbb{K} = \mathbb{C} \end{array} \right\} \subseteq \mathbb{K}$$

Diese Teilmenge von  $\mathbb{K}$  ist abzählbar, da  $\mathbb{Q}$  abzählbar ist und ein kartesisches Produkt zweier abzählbarer Mengen wieder abzählbar ist, siehe auch Übung 2.9.7. Weiter benötigen wir die Menge

$$M = \left\{ \left( \frac{1}{k}, q \right) \mid k \in \mathbb{N}, q \in Q \right\} \subseteq \mathbb{R}^+ \times \mathbb{K},$$

welche mit dem gleichen Argument wieder abzählbar ist. Zum Schluss definieren wir die Menge

$$B = \left\{ \left( \frac{1}{k}, q \right) \in M \mid A \cap B_{\frac{1}{k}}(q) \neq \emptyset \right\} \subseteq M,$$

die als Teilmenge einer abzählbaren Menge selbst wieder abzählbar ist. Ist nun  $a \in A$ , so gibt es eine Folge  $(z_n)_{n \in \mathbb{N}}$  mit  $z_n \in Q$ , mit  $\lim_{n \rightarrow \infty} z_n = a$ . Dies zeigt aber, dass es unendlich viele  $(\frac{1}{k}, q) \in B$  mit  $a \in B_{\frac{1}{k}}(q)$  gibt: Man kann beispielsweise die  $z_n$  aus obiger Folge als  $q$  verwenden und die Radien  $\frac{1}{k}$  entsprechend der Konvergenz  $z_n \rightarrow a$  einrichten, siehe auch Abbildung 2.9. Insbesondere gibt es also unendlich viele Elemente in  $B$ , womit  $B$  abzählbar unendlich ist. Wir wählen daher eine Bijektion

$$\mathbb{N} \ni n \mapsto \left( \frac{1}{k_n}, a_n \right) \in B$$

und definieren so eine Folge  $a = (a_n)_{n \in \mathbb{N}}$ . Wir behaupten, dass für diese Folge  $\text{HP}(a) = A$  gilt. Sei zunächst  $\alpha \in A$ . Da es  $z_m \in Q$  mit  $z_m \rightarrow a$  gibt, gibt es auch zugehörige Radien  $\frac{1}{k_m}$  mit  $|\alpha - z_m| < \frac{1}{k_m}$  für alle  $m \in \mathbb{N}$ . Diese Folge  $(z_m)_{m \in \mathbb{N}}$  ist dann aber eine Teilfolge von  $(a_n)_{n \in \mathbb{N}}$ , womit  $\alpha$  ein Häufungspunkt von  $a$  ist. Es folgt  $A \subseteq \text{HP}(a)$ . Sei umgekehrt  $\beta \notin A$ . Da  $A$  abgeschlossen ist, gibt es ein  $\epsilon > 0$  mit  $B_\epsilon(\beta) \subseteq \mathbb{K} \setminus A$ , da  $\mathbb{K} \setminus A$  ja selbst wieder offen ist. Sei nun  $k$  so groß, dass  $\frac{1}{k} < \epsilon$ . Dann gibt es keinen Punkt  $q \in Q \cap B_{\frac{1}{2k}}(\beta)$ . Wäre nämlich  $q$  ein solcher Punkt, so gälte  $(\frac{1}{2k}, q) \in B$ , womit es ein  $\alpha \in A$  gäbe, so dass  $\alpha \in B_{\frac{1}{2k}}(q)$ , also  $|\alpha - q| < \frac{1}{2k}$ . Dann gälte aber auch

$$|\beta - \alpha| \leq |\beta - q| + |q - \alpha| < \frac{1}{2k} + \frac{1}{2k} = \frac{1}{k},$$

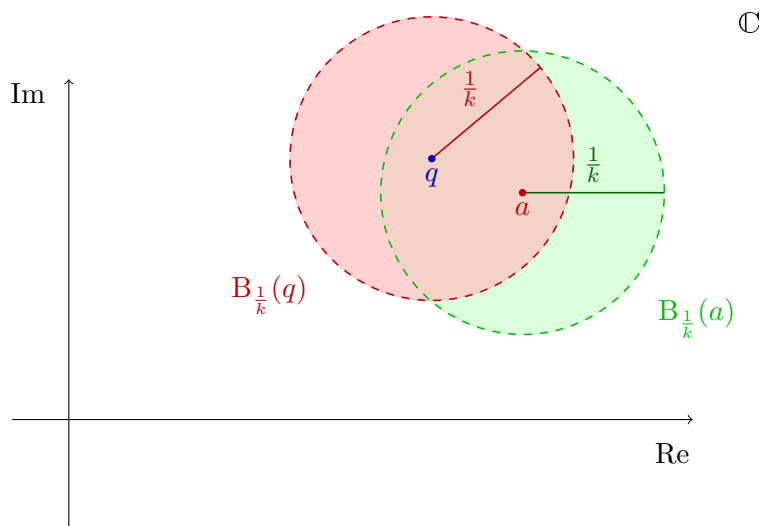


Abbildung 2.9: Es gilt  $a \in B_{\frac{1}{k}}(q)$  genau dann, wenn  $q \in B_{\frac{1}{k}}(a)$ .

ein Widerspruch zu  $A \cap B_\epsilon(\beta) = \emptyset$ . Daher gibt es kein Folgenglied der Folge  $(a_n)_{n \in \mathbb{N}}$  in der Umgebung  $B_\epsilon(\beta)$ . Der Punkt  $\beta$  kann daher kein Häufungspunkt von  $a$  sein,  $\beta \notin \text{HP}(a)$ . Dies zeigt  $\text{HP}(a) \subseteq A$  und somit insgesamt die Gleichheit.  $\square$

**Bemerkung 2.5.12** Dieser Satz ist insofern bemerkenswert, als dass es so viele und komplizierte abgeschlossene Teilmengen von  $\mathbb{K}$  gibt. Wir werden hier noch etliche Beispiele sehen wie etwa in Übung ???. Insbesondere können abgeschlossene Teilmengen  $A \subseteq \mathbb{K}$  sehr groß sein: Die Intervalle  $[a, b]$  mit  $a < b$  sind beispielsweise überabzählbar unendlich. Im Umkehrschluss zeigt dies, dass eine Folge sehr viele Teilfolgen hat, nämlich überabzählbar viele. Insgesamt also ein eher kontraintuitives Resultat.

Eine Folge muss keine Häufungspunkte besitzen. Im reellen wie im komplexen Fall bedeutet dies, dass die Beträge  $|a_n|$  gegen  $+\infty$  divergieren, da sonst eine beschränkte Teilfolge existiert, die dann Häufungspunkte besitzt. Aus diesem Grunde werden wir im reellen Fall folgende Verabredung treffen:

**Definition 2.5.13 (Uneigentlicher Häufungspunkt)** Sei  $a = (a_n)_{n \in \mathbb{N}}$  eine reelle Folge. Ist  $a$  nach oben unbeschränkt, so heißt  $+\infty$  uneigentlicher Häufungspunkt. Ist  $a$  nach unten unbeschränkt, so heißt  $-\infty$  uneigentlicher Häufungspunkt.

Nehmen wir die möglicherweise uneigentlichen Häufungspunkte von  $a$  zur Menge  $\text{HP}(a)$  hinzu, so hat jede reelle Folge  $a$  (uneigentliche) Häufungspunkte,

$$\text{HP}(a) \neq \emptyset. \quad (2.5.6)$$

Diese Konvention erleichtert einige Fallunterscheidungen, wie beispielsweise in folgender Situation:

**Definition 2.5.14 (Limes superior und Limes inferior)** Sei  $a = (a_n)_{n \in \mathbb{N}}$  eine reelle Folge.

i.) Der Limes superior von  $a$  ist als

$$\limsup_{n \rightarrow \infty} a_n = \sup \text{HP}(a) \quad (2.5.7)$$

definiert.

ii.) Der Limes inferior von  $a$  ist als

$$\liminf_{n \rightarrow \infty} a_n = \inf \text{HP}(a) \quad (2.5.8)$$

definiert.

**Bemerkung 2.5.15** Das Supremum  $\sup \text{HP}(a)$  kann dabei natürlich  $+\infty$  sein. Ist  $a$  eine Folge, die nach  $+\infty$  divergiert, so ist zudem auch  $\inf \text{HP}(a) = +\infty$ , da  $\text{HP}(a) = \{+\infty\}$  nur aus einem, in diesem Falle uneigentlichen Häufungspunkt besteht. Allgemein gilt immer

$$\liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n. \quad (2.5.9)$$

Es gilt weiter genau dann Gleichheit in (2.5.9), wenn  $a$  konvergiert. In diesem Fall gilt

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} a_n = \liminf_{n \rightarrow \infty} a_n, \quad (2.5.10)$$

wobei wir hier den Fall der Divergenz gegen  $+\infty$  oder  $-\infty$  mit einschließen müssen. In diesem Sinne ist der Limes superior also der größte Häufungspunkt, während der Limes inferior der kleinste ist: Man beachte, dass

$$\limsup_{n \rightarrow \infty} a_n \in \text{HP}(a) \quad (2.5.11)$$

beziehungsweise

$$\liminf_{n \rightarrow \infty} a_n \in \text{HP}(a) \quad (2.5.12)$$

gilt, da  $\text{HP}(a)$  eine abgeschlossene Teilmenge ist, beziehungsweise  $\pm\infty$  definitionsgemäß enthält.

Eine alternative Charakterisierung von Limes superior und Limes inferior erhält man folgendermaßen:

**Proposition 2.5.16** Sei  $a = (a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{R})$  eine reelle Folge. Dann gilt

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \sup \{a_k \mid k \geq n\} \quad (2.5.13)$$

sowie

$$\liminf_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \inf \{a_k \mid k \geq n\}. \quad (2.5.14)$$

BEWEIS: Wir betrachten die Folge  $(b_n)_{n \in \mathbb{N}}$  mit

$$b_n = \sup \{a_k \mid k \geq n\} \in \mathbb{R} \cup \{+\infty\}.$$

Da die Suprema über immer kleiner werdende Mengen gebildet werden, gilt

$$b_n \geq b_{n+1}$$

für alle  $n \in \mathbb{N}$ . Wir unterscheiden nun einige Fälle: Ist  $(a_n)_{n \in \mathbb{N}}$  nach oben unbeschränkt, so gilt  $b_n = +\infty$  für alle  $n$ , und damit  $\lim_{n \rightarrow \infty} b_n = +\infty$ . Damit gilt (2.5.13). Sei nun  $(a_n)_{n \in \mathbb{N}}$  nach oben beschränkt, so dass  $b_n \in \mathbb{R}$  für alle  $n \in \mathbb{N}$ . Der nächste Fall von Interesse ist, dass die Folge  $b_n$  nach unten beschränkt ist. Nach Proposition 2.4.9 konvergiert  $(b_n)_{n \in \mathbb{N}}$  gegen das Infimum der Folgenglieder, also

$$\beta = \lim_{n \rightarrow \infty} b_n = \inf \{b_n \mid n \in \mathbb{N}\}.$$

Insbesondere gilt  $b_n \geq \beta$  für alle  $n \in \mathbb{N}$ . Sei  $\epsilon > 0$  vorgegeben. Dann liegen nur endlich viele  $b_n$  oberhalb von  $\beta + \epsilon$ , und entsprechend gilt für fast alle  $n \in \mathbb{N}$

$$b_n \leq \beta + \epsilon.$$

Da aber  $a_k \leq b_n$  für alle  $k \geq n$ , folgt

$$a_n \leq \beta + \epsilon \quad (*)$$

für fast alle  $n \in \mathbb{N}$ . Insbesondere hat die ursprüngliche Folge keinen Häufungspunkt größer als  $\beta + \epsilon$  für beliebiges  $\epsilon > 0$ . Mit anderen Worten, der größte Häufungspunkt von  $a$  ist höchstens  $\beta$ , also

$$\sup \text{HP}(a) \leq \beta = \lim_{n \rightarrow \infty} b_n.$$

Da  $b_n \geq \beta$  gilt, folgt  $b_n > \beta - \frac{\epsilon}{2}$  für alle  $n \in \mathbb{N}$ . Da  $b_n = \sup\{a_k \mid k \geq n\}$  gibt es zu jedem  $n$  ein  $k \geq n$  mit  $a_k > b_n - \frac{\epsilon}{2}$ , womit insgesamt

$$a_k > \beta - \epsilon \quad (**)$$

folgt. Da wir dies für jedes  $n \in \mathbb{N}$  erreichen, gibt es also unendlich viele  $a_k > \beta - \epsilon$ . Beide Aussagen (\*) und (\*\*) zusammen liefern also insgesamt noch unendlich viele  $a_n \in (\beta - \epsilon, \beta + \epsilon)$ . Da  $\epsilon > 0$  beliebig war, ist  $\beta \in \text{HP}(a)$  ein Häufungspunkt von  $a$ , womit (2.5.13) in diesem Fall auch gilt. Sei schließlich  $b_n$  nach unten unbeschränkt. Ist  $N \in \mathbb{N}$  vorgegeben, so gilt

$$b_n < -N$$

für fast alle  $n$ , da  $b_n$  ja auch monoton fallend ist. Da allgemein  $a_k \leq b_n$  für  $k \geq n$ , gilt auch

$$a_n < -N$$

für fast alle  $n$ . Die Folge  $(a_n)_{n \in \mathbb{N}}$  hat damit nur den uneigentlichen Grenzwert  $-\infty$  als einzigen uneigentlichen Häufungspunkt, so dass auch in diesem letzten Fall (2.5.13) gezeigt ist. Den zweiten Teil (2.5.14) erhält man analog, indem man die Folge  $a$  durch  $-a$  ersetzt.  $\square$

Zum Abschluss zeigen wir noch eine Ungleichung, die bei der Diskussion von Reihen sehr nützlich sein wird.

**Proposition 2.5.17** Sei  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge mit  $a_n \neq 0$ . Dann gilt

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \leq \liminf_{n \rightarrow \infty} \sqrt[n]{|a_n|} \leq \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \leq \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|. \quad (2.5.15)$$

BEWEIS: Die mittlere Ungleichung gilt trivialerweise nach (2.5.9). Wir zeigen die rechte Ungleichung. Sei

$$\alpha = \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

der Limes superior der Folge  $(\left| \frac{a_{n+1}}{a_n} \right|)_{n \in \mathbb{N}}$ . Gilt  $\alpha = +\infty$ , so ist die rechte Ungleichung in (2.5.15) trivialerweise richtig, unabhängig davon,  $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$  ist. Wir können daher  $\alpha < \infty$  annehmen. Wir wählen ein  $\epsilon > 0$  und eine Zahl  $\beta \in (\alpha, \alpha + \epsilon)$ . Nach Definition des Limes superior sind dann fast alle Folgenglieder  $\left| \frac{a_{n+1}}{a_n} \right|$  kleiner als  $\beta$ . Es gibt also ein  $N \in \mathbb{N}$  mit  $\left| \frac{a_{n+1}}{a_n} \right| < \beta$  für alle  $n \geq N$ . Ist  $k > 0$ , so folgt durch Induktion

$$|a_{N+k}| < \beta |a_{N+k-1}| < \cdots < \beta^k |a_N|.$$

Schreiben wir wieder  $n = N + k$ , so gilt also

$$|a_n| < \beta^{n-N} |a_N|$$

und damit

$$|a_n|^{\frac{1}{n}} < \beta^{\frac{n-N}{n}} \sqrt[n]{\frac{|a_N|}{\beta^N}}$$

für alle  $n > N$ . Für die rechte Seite wissen wir bereits

$$\lim_{n \rightarrow \infty} \beta \sqrt[n]{\frac{|a_N|}{\beta^N}} = \beta,$$

siehe Beispiel 2.3.12, *ii.*). Damit ist aber

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \leq \beta < \alpha + \epsilon$$

für alle  $\epsilon > 0$ . Dies zeigt schließlich, dass auch  $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \leq \alpha$  gelten muss, was gerade die rechte Ungleichung in (2.5.15) ist. Die verbleibende linke Ungleichung zeigt man durch analoge Überlegungen, siehe Übung ??.

## 2.6 Unendliche Reihen

Unendliche Reihen sind spezielle Folgen, deren Konvergenzverhalten nun studiert werden soll. Alle Resultate zu Folgen sind daher zumindest prinzipiell anwendbar. Darüber hinaus gibt es jedoch neue Phänomene, die eine weiterführende Diskussion erforderlich machen.

Ist  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  eine Folge, so können wir daraus eine neue Folge  $(s_n)_{n \in \mathbb{N}}$  konstruieren, indem wir die ersten  $n$  Folgenglieder addieren. Wir setzen also

$$s_n = \sum_{k=1}^n a_k. \quad (2.6.1)$$

Wir schreiben für die zugehörige Folge  $(s_n)_{n \in \mathbb{N}}$  auch

$$(s_n)_{n \in \mathbb{N}} = \sum_{n=1}^{\infty} a_n \quad (2.6.2)$$

und nennen dies die *unendliche Reihe* der Folgenglieder  $a_n$ . Da

$$s_1 = a_1 \quad \text{und} \quad s_{n+1} - s_n = a_{n+1}, \quad (2.6.3)$$

enthält die Folge  $(s_n)_{n \in \mathbb{N}}$  also die selbe Information wie die ursprüngliche Folge  $(a_n)_{n \in \mathbb{N}}$ . Wir nennen  $(s_n)_{n \in \mathbb{N}}$  auch die Folge der *Partialsummen*.

Konvergiert nun die Folge der Partialsummen, so heißt der Grenzwert

$$\sum_{n=1}^{\infty} a_n = \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{k=1}^n a_k \quad (2.6.4)$$

der *Wert* der unendlichen Reihe. Man beachte den hier stattfindenden aber sehr üblichen Notationsmissbrauch: Das Symbol  $\sum_{k=1}^{\infty} a_n$  ist doppelt belegt, da es einerseits die unendliche Reihe (2.6.2) und andererseits, im Falle der Konvergenz, auch deren Wert, also den Grenzwert (2.6.4), kennzeichnet.

Schließlich sei angemerkt, dass wir den Summationsindex auch bei anderen Werten in  $\mathbb{Z}$  anstelle von 1 starten lassen können. Die Folge der Partialsummen startet dann entsprechend auch bei einem anderen Index als 1. Dies liefert konzeptuell nichts neues, da wir durch eine Indexverschiebung immer den Startwert 1 erreichen können. Allerdings werden verschiedene Formeln in Beispielen einfacher zu interpretieren sein, so dass wir von dieser Freiheit Gebrauch machen werden.

Wir beginnen nun mit einigen Beispielen für konvergente und divergente Reihen:

**Beispiel 2.6.1 (Geometrische Reihe)** Sei  $z \in \mathbb{K}$  mit  $|z| < 1$ . Dann konvergiert die *geometrische Reihe*

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}. \quad (2.6.5)$$

Für  $|z| \geq 1$  ist die geometrische Reihe divergent. Um dies zu sehen erinnern wir an die geometrische Summenformel

$$\sum_{n=0}^N z^n = \frac{1-z^{N+1}}{1-z} \quad (2.6.6)$$

aus Proposition 1.2.10 für  $z \neq 1$ . Im Grenzwert  $N \rightarrow \infty$  konvergiert  $z^{N+1} \rightarrow 0$  genau dann, wenn  $|z| < 1$  nach Beispiel 2.2.10, *ii.*). Ist dagegen  $|z| \geq 1$ , so ist  $(z^{N+1})_{N \in \mathbb{N}}$  divergent, siehe auch Übung 2.9.17. Der Fall  $z = 1$  kann direkt gesehen werden, da dann  $\sum_{n=0}^N 1^n = N+1$  gegen  $+\infty$  divergiert. Überraschenderweise ist diese einfache Beobachtung von fundamentaler Bedeutung in der Mathematik.

**Beispiel 2.6.2 (Harmonische Reihe)** Die harmonische Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty \quad (2.6.7)$$

ist divergent gegen  $+\infty$ . Dies ist in dem Moment überraschend, wenn man auf dem Computer die Folge der Partialsummen

$$s_N = \sum_{n=1}^N \frac{1}{n} \quad (2.6.8)$$

numerisch berechnet. Die Divergenz gegen  $+\infty$  ist unglaublich langsam, und je nach verwendetem Computerprogramm aufgrund von Rundungsfehlern nicht zu sehen. Um die Divergenz zu zeigen, betrachten wir

$$\begin{aligned} s_{2^{k+1}} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \cdots + \left(\frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}}\right) \\ &> 1 + \frac{1}{2} + \frac{2}{4} + \cdots + \frac{2^k}{2^{k+1}} \\ &> \frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2} \\ &= \frac{k+1}{2}. \end{aligned}$$

Damit gilt also

$$s_{2^{k+1}} > \frac{k+1}{2}, \quad (2.6.9)$$

womit die Divergenz gegen  $+\infty$  gezeigt ist. Wir sehen aber auch, wieso die Divergenz so langsam ist: Um einen linearen Term in  $k$  zu erreichen, brauchen wir exponentiell viele Summanden  $N = 2^{k+1}$ .

Wir wollen nun einige Konvergenzkriterien zu unendlichen Reihen finden, die sich zum einen aus den Konvergenzkriterien von Folgen ergeben, zum anderen aber auch neue Phänomene bereithalten.

**Proposition 2.6.3** Eine unendliche Reihe  $\sum_{n=1}^{\infty} a_n$  mit  $a_n \in \mathbb{K}$  konvergiert genau dann, wenn sie eine Cauchy-Reihe ist, also wenn es zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  gibt, so dass für alle  $n \geq m \geq N$

$$\left| \sum_{k=m}^n a_k \right| < \epsilon \quad (2.6.10)$$

*gilt.*

BEWEIS: Eine Reihe konvergiert genau dann, wenn die Folge der Partialsummen

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

konvergiert. Dies ist genau dann der Fall, wenn  $(s_n)_{n \in \mathbb{N}}$  eine Cauchy-Folge ist, siehe Satz 2.4.4 und Proposition 2.4.2. Die Cauchy-Bedingung für die Folge  $(s_n)_{n \in \mathbb{N}}$  bedeutet, dass es zu  $\epsilon > 0$  ein  $N \in \mathbb{N}$  gibt, so dass für alle  $n, m \geq N$

$$|s_n - s_m| < \epsilon$$

gilt. Ohne Einschränkung können wir  $n \geq m$  annehmen. Dann gilt aber

$$|s_n - s_m| = \left| \sum_{k=0}^n a_k - \sum_{k=0}^m a_k \right| = \left| \sum_{k=m+1}^n a_k \right|,$$

womit der Beweis vollbracht ist. □

**Korollar 2.6.4** Sei  $\sum_{k=1}^{\infty} a_k$  eine konvergente Reihe.

i.) Die Folge  $(a_n)_{n \in \mathbb{N}}$  ist notwendigerweise eine Nullfolge.

ii.) Die Restreihen  $(\sum_{k=n}^{\infty} a_k)_{n \in \mathbb{N}}$  bilden eine Nullfolge.

BEWEIS: Wendet man das Cauchy-Kriterium (2.6.10) auf  $n > n-1 \geq N$  an, erhält man  $|a_n| < \epsilon$ , also eine Nullfolge. Die Restreihen

$$\sum_{k=n}^{\infty} a_k = \sum_{k=1}^{\infty} a_k - (a_1 + \cdots + a_{n-1})$$

bilden ebenfalls nach dem Cauchy-Kriterium eine Nullfolge, da

$$\left| \sum_{k=m}^n a_k \right| < \epsilon$$

für  $n \geq m \geq N$  und daher auch

$$\left| \sum_{k=m}^{\infty} a_k \right| \leq \limsup_{n \rightarrow \infty} \left| \sum_{k=m}^n a_k \right| \leq \epsilon,$$

sofern  $m \geq N$ . □

Die harmonische Reihe aus Beispiel 2.6.2 zeigt, dass die Bedingung i.) in Korollar 2.6.4 notwendig, aber im Allgemeinen nicht hinreichend ist: Es gibt Nullfolgen  $(a_n)_{n \in \mathbb{N}}$  wie beispielsweise  $(\frac{1}{n})_{n \in \mathbb{N}}$ , so dass  $\sum_{n=1}^{\infty} a_n$  trotzdem divergiert.

Die Rechenregeln für Grenzwerte von Folgen liefern direkt entsprechende Rechenregeln für Reihen, wie beispielsweise folgende Linearitätseigenschaften:

**Proposition 2.6.5** Seien  $\sum_{n=1}^{\infty} a_n$  und  $\sum_{n=1}^{\infty} b_n$  konvergente Reihen.

i.) Für alle  $z, w \in \mathbb{K}$  ist auch  $\sum_{n=1}^{\infty} (za_n + wb_n)$  konvergent, und es gilt

$$\sum_{n=1}^{\infty} (za_n + wb_n) = z \sum_{n=1}^{\infty} a_n + w \sum_{n=1}^{\infty} b_n. \quad (2.6.11)$$

ii.) Die Reihe  $\sum_{n=1}^{\infty} \bar{a}_n$  ist konvergent, und es gilt

$$\sum_{n=1}^{\infty} \bar{a}_n = \overline{\sum_{n=1}^{\infty} a_n}. \quad (2.6.12)$$

BEWEIS: Seien  $s_n = \sum_{k=1}^n a_k$  und  $t_n = \sum_{k=1}^n b_k$  die jeweiligen Partialsummen. Dann gilt

$$\sum_{k=1}^n (za_k + wb_k) = zs_n + wt_n,$$

womit i.) aus Proposition 2.3.5, i.), angewandt auf die Folgen der Partialsummen, folgt. Da  $\bar{s}_n = \sum_{k=1}^n \bar{a}_k$  gilt, erhält man ii.) aus Proposition 2.3.5, iii.).  $\square$

Eine Verträglichkeit mit Produktbildung ist komplizierter. Wir werden darauf später zurückkommen. Bereits die obige Linearität erlaubt es in vielen Beispielen, nicht nur die Konvergenz zu entscheiden, sondern auch die Werte zu bestimmen.

Eine besonders einfache Situation tritt ein, wenn alle Folgenglieder positiv sind, da in diesem Fall die Folge der Partialsummen monoton wachsend ist:

**Proposition 2.6.6** Sei  $(a_n)_{n \in \mathbb{N}}$  eine reelle Folge mit  $a_n \geq 0$  für fast alle  $n \in \mathbb{N}$ . Die unendliche Reihe  $\sum_{n=1}^{\infty} a_n$  konvergiert genau dann, wenn ihre Folge von Partialsummen nach oben beschränkt ist.

BEWEIS: Da  $a_n \geq 0$  für fast alle  $n \in \mathbb{N}$  gilt, gibt es ein  $N \in \mathbb{N}$  mit  $a_n \geq 0$  für  $n \geq N$ . Von da an sind die Partialsummen  $s_n = a_1 + \dots + a_n$  monoton wachsend, also

$$s_{n+1} = s_n + a_{n+1} \geq s_n$$

für alle  $n \geq N$ . Nach Proposition 2.4.9 ist  $(s_n)_{n \in \mathbb{N}}$  also genau dann konvergent, wenn diese Folge (nach oben) beschränkt ist.  $\square$

Wir können dieses einfache Kriterium nun nutzen, um folgendes *Verdichtungskriterium* von Cauchy zu zeigen:

**Proposition 2.6.7** Sei  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{R})$  eine monoton fallende Nullfolge, also insbesondere  $a_n \geq 0$  für alle  $n \in \mathbb{N}$ . Dann konvergiert  $\sum_{n=1}^{\infty} a_n$  genau dann, wenn die verdichtete Reihe  $\sum_{n=1}^{\infty} 2^n a_{2^n}$  konvergiert. In diesem Fall ist die Folge  $(2^n a_{2^n})_{n \in \mathbb{N}}$  also immer noch eine Nullfolge.

BEWEIS: Sei zunächst  $\sum_{n=1}^{\infty} a_n$  konvergent mit Wert  $s$ . Dann gilt für alle  $n \in \mathbb{N}$

$$\begin{aligned} s &= \sum_{k=1}^{\infty} a_k \\ &\geq \sum_{k=1}^{2^n} a_k \\ &= a_1 + a_2 + (a_3 + a_4) + \dots + (a_{2^{n-1}-1} + \dots + a_{2^n}) \\ &\geq \frac{a_1}{2} + a_2 + 2a_4 + \dots + 2^{n-1}a_{2^n}, \end{aligned}$$

da  $(a_n)_{n \in \mathbb{N}}$  monoton fallend ist. Dies zeigt, dass

$$a_1 + 2a_2 + 4a_4 + \dots + 2^n a_{2^n} \geq 2s$$



für alle  $n \in \mathbb{N}$ , womit die Partialsummen der Reihe  $\sum_{n=1}^{\infty} 2^n a_{2^n}$  durch  $2 \sum_{n=1}^{\infty} a_n = 2s$  beschränkt sind. Nach Proposition 2.6.6 konvergiert die verdichtete Reihe. Sei umgekehrt  $\sum_{n=1}^{\infty} 2^n a_{2^n}$  konvergent. Dann gilt für alle  $n \in \mathbb{N}$

$$\begin{aligned} \sum_{k=1}^n a_k &= a_1 + \cdots + a_n \\ &\leq a_1 + \cdots + a_{2^{n+1}-1} \\ &= a_1 + (a_2 + a_3) + \cdots + (a_{2^n} + \cdots + a_{2^{n+1}-1}) \\ &\leq a_1 + 2a_2 + \cdots + 2^n a_{2^n} \\ &\leq \sum_{k=1}^{\infty} 2^k a_{2^k}, \end{aligned}$$

da  $(a_n)_{n \in \mathbb{N}}$  monoton fallend ist. Damit sind die Partialsummen von  $\sum_{k=1}^{\infty} a_k$  durch  $\sum_{k=1}^{\infty} 2^k a_{2^k}$  beschränkt, womit die Konvergenz von  $\sum_{k=1}^{\infty} a_k$  wieder aus Proposition 2.6.6 folgt. Dass dann die verdichtete Folge  $(2^n a_{2^n})_{n \in \mathbb{N}}$  eine Nullfolge sein muss, folgt wieder aus Korollar 2.6.4, i.).  $\square$

In gewisser Hinsicht ist dieses Verdichtungskriterium die Abstraktion des Vorgehens bei der Divergenz der harmonischen Reihe in Beispiel 2.6.1. Wir können es nun nutzen, um die Konvergenz folgender Varianten der harmonischen Reihe zu zeigen:

**Beispiel 2.6.8 (Riemannsche  $\zeta$ -Funktion)** Sei  $\alpha \in \mathbb{R}$ . Die Reihe

$$\zeta(\alpha) = \sum_{n=1}^{\infty} \frac{1}{n^\alpha} \quad (2.6.13)$$

konvergiert für  $\alpha > 1$  und divergiert für  $\alpha \leq 1$ . Nach unseren Rechenregeln für Potenzen ist zunächst klar, dass  $(\frac{1}{n^\alpha})_{n \in \mathbb{N}}$  für  $\alpha > 0$  eine (streng) monoton fallende Folge ist, die gegen 0 konvergiert. Für  $\alpha \leq 0$  ist sie keine Nullfolge, womit die Divergenz der Reihe klar ist. Wir wenden nun Proposition 2.6.7 an, und wissen daher, dass (2.6.13) genau dann konvergiert, wenn die verdichtete Reihe

$$\sum_{n=1}^{\infty} 2^n \frac{1}{(2^n)^\alpha} = \sum_{n=1}^{\infty} \left( \frac{1}{2^{\alpha-1}} \right)^n \quad (2.6.14)$$

konvergiert. Dies ist aber, bis auf den Summanden 1 für  $n = 0$ , gerade die geometrische Reihe. Daher konvergiert (2.6.14) genau dann, wenn  $\frac{1}{2^{\alpha-1}} < 1$  gilt, was gleichbedeutend mit  $\alpha > 1$  ist, siehe Beispiel 2.6.1. Der Wert der Reihe (2.6.13) ist die berühmte *Riemannsche  $\zeta$ -Funktion*. Die tatsächliche Berechnung von einzelnen, speziellen Werten  $\zeta(\alpha)$  werden wir später nachholen.

Eine reelle unendliche Reihe  $\sum_{n=1}^{\infty} a_n$  hat eine bessere Chance zu konvergieren, wenn die Vorzeichen der Koeffizienten wechseln, da sich in diesem Fall Terme wegheben können und  $|\sum_{k=n}^m a_k|$  schneller klein wird. Diese heuristische Vorstellung lässt sich in folgendem Konvergenzkriterium von Leibniz für *alternierende Reihen* präzisieren:

**Proposition 2.6.9** Sei  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{R})$  eine monoton fallende Nullfolge, also insbesondere  $a_n \geq 0$  für alle  $n \in \mathbb{N}$ . Dann konvergiert die unendliche Reihe  $\sum_{n=1}^{\infty} (-1)^n a_n$ . Es gilt die Fehlerabschätzung

$$\left| \sum_{k=1}^{\infty} (-1)^k a_k - \sum_{k=1}^n (-1)^k a_k \right| \leq a_{n+1}. \quad (2.6.15)$$

BEWEIS: Wir betrachten die Folge der Partialsummen  $s_n = \sum_{k=1}^n (-1)^k a_k$  der Reihe. Dann gilt

$$s_n - s_{n-2} = (-1)^n a_n + (-1)^{n-1} a_{n-1} = (-1)^n (a_n - a_{n-1}).$$

Da  $(a_n)_{n \in \mathbb{N}}$  monoton fallend ist, gilt  $a_n - a_{n-1} \leq 0$ , womit

$$s_{2n} - s_{2n-2} = a_{2n} - a_{2n-1} \leq 0$$

und

$$s_{2n+1} - s_{2n-1} = -(a_{2n+1} - a_{2n}) \geq 0$$

gilt. Die Teilfolge  $(s_{2n})_{n \in \mathbb{N}}$  der Partialsummen ist also monoton fallend, während die Teilfolge  $(s_{2n+1})_{n \in \mathbb{N}}$  monoton wachsend ist. Weiter gilt

$$s_{2n} = s_{2n+1} + a_{2n+1} \geq s_{2n+1}$$

und

$$s_{2n+1} = s_{2n} - a_{2n+1} \leq s_{2n}.$$

Da  $(s_{2n+1})_{n \in \mathbb{N}}$  monoton wachsend ist, ist insbesondere  $s_{2n+1} \geq s_1$  für alle  $n \in \mathbb{N}$ . Genauso erhalten wir  $s_{2n} \leq s_2$  für alle  $n \in \mathbb{N}$ . Dies zeigt

$$s_{2n} \geq s_{2n+1} \geq s_1 \quad \text{und} \quad s_{2n+1} \leq s_{2n} \leq s_2,$$

womit  $(s_{2n})_{n \in \mathbb{N}}$  nach oben beschränkt ist, während  $(s_{2n+1})_{n \in \mathbb{N}}$  nach unten beschränkt ist. Daher sind beide Folgen nach Proposition 2.4.9 konvergent. Weiter gilt

$$s_{2n} - s_{2n-1} = a_{2n}.$$

Da  $(a_{2n})_{n \in \mathbb{N}}$  eine Nullfolge ist, konvergieren die beiden konvergenten Folgen  $(s_{2n})_{n \in \mathbb{N}}$  und  $(s_{2n+1})_{n \in \mathbb{N}}$  gegen den selben Grenzwert  $s \in \mathbb{R}$ . Dies ist die Konvergenz der Reihe, da für fast alle  $n$  sowohl

$$|s_{2n} - s| < \epsilon \quad \text{als auch} \quad |s_{2n+1} - s| < \epsilon$$

gilt, also insgesamt für fast alle  $n$  die Abschätzung  $|s_n - s| < \epsilon$  gültig ist. Da die geraden Partialsummen  $s_{2n}$  alle

$$s_{2n} \geq \sum_{k=1}^{\infty} (-1)^k a_k$$

erfüllen, während die ungeraden Partialsummen

$$s_{2n+1} \leq \sum_{k=1}^{\infty} (-1)^k a_k$$

erfüllen, gilt die Abschätzung in (2.6.15), da ja  $|s_{2n+1} - s_{2n}| = a_{2n+1}$  und  $|s_{2n} - s_{2n-1}| = a_{2n}$ , siehe auch Abbildung 2.10.  $\square$

**Beispiel 2.6.10** Ein erstes Beispiel erhalten wir aus der Nullfolge  $(\frac{1}{n})_{n \in \mathbb{N}}$ , dass die Reihe

$$s = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \tag{2.6.16}$$

nach dem Leibniz-Kriterium für alternierende Reihen konvergiert. Den Wert  $s$  dieser Reihe werden wir später noch bestimmen. An diesem Beispiel sieht man gut, dass die zusätzlichen alternierenden Vorzeichen für die Konvergenz entscheidend sind. Ohne diese Vorzeichen ist die Reihe  $\sum_{n=1}^{\infty} \frac{1}{n} = +\infty$  die bekannte divergente harmonische Reihe.

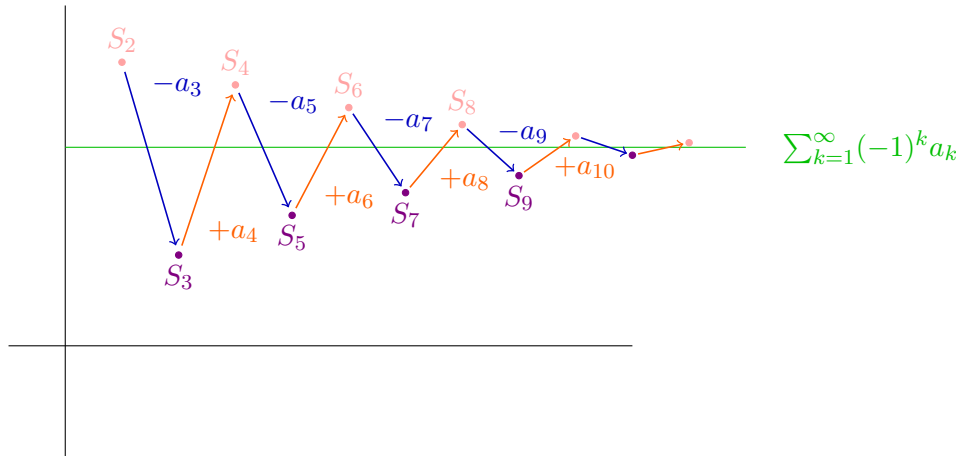


Abbildung 2.10: Die Partialsummen pendeln um den Grenzwert  $\sum_{k=1}^{\infty} (-1)^k a_k$ , wobei der Abstand zweier aufeinander folgender Partialsummen immer durch das entsprechende Folgenglied gegeben ist.

## 2.7 Absolute Konvergenz von Reihen

Gerade das Leibniz-Kriterium für alternierende Reihen zeigt, dass die Vorzeichen bei der Konvergenz von unendlichen Reihen eine entscheidende Rolle spielen können. Man erhält also einen strikteren Konvergenzbegriff, wenn man die Konvergenz der Absolutbeträge fordert:

**Definition 2.7.1 (Absolute Konvergenz)** Eine unendliche Reihe  $\sum_{n=1}^{\infty} a_n$  mit Koeffizientenfolge  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  heißt absolut konvergent, falls die Reihe  $\sum_{n=1}^{\infty} |a_n|$  konvergiert.

Dies ist tatsächlich eine Verschärfung des bisherigen Konvergenzbegriffs:

**Proposition 2.7.2** Sei  $\sum_{n=1}^{\infty} a_n$  eine absolut konvergente Reihe mit  $a_n \in \mathbb{K}$  für alle  $n \in \mathbb{N}$ . Dann konvergiert die Reihe  $\sum_{n=1}^{\infty} a_n$ , und es gilt

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n|. \quad (2.7.1)$$

BEWEIS: Die übliche Dreiecksungleichung für endliche Summen zeigt, dass

$$\left| \sum_{k=n}^m a_k \right| \leq \sum_{k=n}^m |a_k| \quad (*)$$

für alle  $m \geq n \in \mathbb{N}$  gilt. Da die Reihe der Absolutbeträge konvergiert, ist  $\sum_{n=1}^{\infty} |a_n|$  eine Cauchy-Reihe. Die rechte Seite in (\*) wird also klein sofern  $n$  und  $m$  groß genug sind, womit die Cauchy-Eigenschaft für die ursprüngliche Reihe  $\sum_{n=1}^{\infty} a_n$  erfüllt ist. Nach Proposition 2.6.3 konvergiert also die ursprüngliche Reihe. Es gilt nun für alle  $N \in \mathbb{N}$

$$\left| \sum_{n=1}^N a_n \right| \stackrel{(*)}{\leq} \sum_{n=1}^N |a_n| \leq \sum_{n=1}^{\infty} |a_n|,$$

womit die Folge der Partialsummen von  $\sum_{n=1}^{\infty} a_n$  betragsmäßig durch den Wert der Reihe  $\sum_{n=1}^{\infty} |a_n|$  beschränkt ist. Dies zeigt (2.7.1).  $\square$

Wir nennen die Ungleichung (2.7.1) aus naheliegenden Gründen auch *Dreiecksungleichung* für die absolut konvergente Reihe.

**Beispiel 2.7.3 (Absolute Konvergenz I)**

- i.) Die alternierende Reihe  $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$  ist nach Beispiel 2.6.10 konvergent, aber nicht absolut konvergent. Die Umkehrung der Aussage 2.7.2 gilt also im Allgemeinen nicht.
- ii.) Ist  $a_n \geq 0$  für alle  $n \in \mathbb{N}$ , so ist die Reihe  $\sum_{n=1}^{\infty} a_n$  genau dann konvergent, wenn sie absolut konvergent ist. Weiter gilt, dass die Reihe genau dann divergent ist, wenn ihre Partialsummenfolge gegen  $+\infty$  divergiert.
- iii.) Die geometrische Reihe  $\sum_{n=1}^{\infty} z^n$  ist absolut konvergent für  $|z| < 1$  und entsprechend gilt

$$\left| \frac{1}{1-z} \right| = \left| \sum_{n=1}^{\infty} z^n \right| \leq \sum_{n=1}^{\infty} |z|^n = \frac{1}{1-|z|}, \quad (2.7.2)$$

eine Ungleichung, die man natürlich auch elementar sehen kann.

Die absolute Konvergenz lässt sich durch verschiedene Kriterien überprüfen. Das erste und einfachste ist das *Majorantenkriterium*, das aber das Vorwissen um eine *Majorante* voraussetzt.

**Proposition 2.7.4 (Majorantenkriterium)** Sei  $\sum_{n=1}^{\infty} a_n$  eine reelle oder komplexe Reihe. Gibt es dann eine absolut konvergente Reihe  $\sum_{n=1}^{\infty} b_n$  mit  $b_n \geq 0$  und

$$|a_n| \leq b_n \quad (2.7.3)$$

für fast alle  $n \in \mathbb{N}$ , so ist  $\sum_{n=1}^{\infty} a_n$  absolut konvergent.

BEWEIS: Dies ist einfach, da mit (2.7.3) ein  $N \in \mathbb{N}$  existiert, so dass  $|a_n| \leq b_n$  für  $n \geq N$  und somit

$$\sum_{n=N}^m |a_n| \leq \sum_{n=N}^m b_n \leq \sum_{n=0}^{\infty} b_n < \infty.$$

Dies zeigt, dass  $\sum_{n=1}^{\infty} |a_n|$  eine Cauchy-Reihe ist, da ja die Restreihen  $\sum_{n=N}^m b_n$  der konvergenten Reihe  $\sum_{n=1}^{\infty} b_n$  klein werden, wenn  $N$  groß genug wird.  $\square$

In diesem Fall heißt die Reihe  $\sum_{n=1}^{\infty} b_n$  auch Majorante der Reihe  $\sum_{n=1}^{\infty} a_n$ .

Eine Trennung zwischen absoluter Konvergenz und Divergenz erlaubt das *Wurzelkriterium*:

**Proposition 2.7.5 (Wurzelkriterium)** Sei  $\sum_{n=1}^{\infty} a_n$  eine reelle oder komplexe Reihe.

i.) Gilt

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1, \quad (2.7.4)$$

so konvergiert  $\sum_{n=1}^{\infty} a_n$  absolut.

ii.) Gilt

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1, \quad (2.7.5)$$

so divergiert die Reihe  $\sum_{n=1}^{\infty} a_n$ .

BEWEIS: Wir betrachten zunächst den Fall i.). Nach Definition der Limes superior als größter Häufungspunkt gibt es ein  $N \in \mathbb{N}$  und eine Zahl  $0 \leq q < 1$  derart, dass

$$\sqrt[n]{|a_n|} \leq q$$

für alle  $n \geq N$ . Dies bedeutet, dass für fast alle  $n$  die Abschätzung  $|a_n| \leq q^n$  gilt. Da die geometrische Reihe  $\sum_{n=1}^{\infty} q^n$  dank  $0 \leq q < 1$  absolut konvergiert, haben wir nach Proposition 2.7.4 eine Majorante gefunden, was i.) zeigt. Gilt umgekehrt (2.7.5), so gibt es unendlich viele Indizes  $n \in \mathbb{N}$  mit  $\sqrt[n]{|a_n|} > 1$  und daher  $|a_n| > 1$ . Die Folge  $(a_n)_{n \in \mathbb{N}}$  kann also nicht einmal eine Nullfolge sein.  $\square$

Das letzte allgemeine Kriterium ist das *Quotientenkriterium*, das meist etwas einfacher zu handhaben ist als das Wurzelkriterium.

**Proposition 2.7.6 (Quotientenkriterium)** Sei  $\sum_{n=1}^{\infty} a_n$  eine reelle oder komplexe Reihe mit  $a_n \neq 0$  für fast alle  $n \in \mathbb{N}$ .

i.) Gilt

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1, \quad (2.7.6)$$

so konvergiert  $\sum_{n=1}^{\infty} a_n$  absolut.

ii.) Gilt für fast alle  $n \in \mathbb{N}$

$$\left| \frac{a_{n+1}}{a_n} \right| > 1, \quad (2.7.7)$$

so divergiert die Reihe  $\sum_{n=1}^{\infty} a_n$ .

BEWEIS: Nach Proposition 2.5.17 gilt

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \leq \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|,$$

womit wir im Fall i.) das Wurzelkriterium aus Proposition 2.7.5, i.), zum Einsatz bringen können. Gilt dagegen die Voraussetzung in ii.), so folgt  $|a_{n+1}| \geq |a_n|$  für fast alle  $n \in \mathbb{N}$ . Da fast alle  $a_n \geq 0$  sein sollen, gibt es also ein  $N \in \mathbb{N}$  mit  $a_N \neq 0$  und  $|a_{n+1}| \geq |a_n|$  für alle  $n \geq N$ . Es folgt

$$|a_{n+1}| \geq |a_n| \geq \cdots \geq |a_N| > 0,$$

so dass auch in diesem Fall  $(a_n)_{n \in \mathbb{N}}$  keine Nullfolge ist. □

### Beispiel 2.7.7 (Absolute Konvergenz II)

i.) Sei  $\delta > 0$  und  $p \geq 1$  fest gewählt. Dann konvergiert die Reihe

$$\sum_{n=1}^{\infty} \frac{n^p}{(1+\delta)^n} < \infty \quad (2.7.8)$$

absolut. Es gilt nämlich

$$\lim_{n \rightarrow \infty} \sqrt[n]{\frac{n^p}{(1+\delta)^n}} = \lim_{n \rightarrow \infty} \frac{\sqrt[n]{n^p}}{(1+\delta)} = \frac{1}{1+\delta} < 1 \quad (2.7.9)$$

nach Beispiel 2.3.12, ii.). Da die Folge sogar konvergent ist, stimmt der Grenzwert mit dem Limes superior überein, womit

$$\limsup_{n \rightarrow \infty} \sqrt[n]{\frac{n^p}{(1+\delta)^n}} = \frac{1}{1+\delta} < 1 \quad (2.7.10)$$

folgt. Nach dem Wurzelkriterium konvergiert (2.7.8) daher absolut.

ii.) Für die harmonische Reihe erhalten wir aus Beispiel 2.3.12, ii.),

$$\limsup_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n}} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n}} = 1, \quad (2.7.11)$$

so dass wir mit Hilfe des Wurzelkriteriums nicht entscheiden können, dass die harmonische Reihe divergiert. Es gilt aber auch für  $\alpha > 1$

$$\limsup_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n^\alpha}} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n^\alpha}} = 1. \quad (2.7.12)$$

Da wir die Konvergenz von  $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$  für  $\alpha > 1$  nach Beispiel 2.6.8 aber wissen, zeigt dies, dass beim Wurzelkriterium der verbleibende Fall  $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 1$  wirklich unentschieden bleibt: Je nach Reihe kann Konvergenz vorliegen oder nicht.

iii.) Als Beispiel für das Majorantenkriterium betrachten wir  $a_n = \frac{1}{n(n+1)}$  und die zugehörige Reihe  $\sum_{n=1}^{\infty} a_n$ . Da

$$\frac{1}{n(n+1)} < \frac{1}{n^2} \quad (2.7.13)$$

für alle  $n \in \mathbb{N}$  gilt, konvergiert die Reihe absolut, und es gilt

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} = \zeta(2). \quad (2.7.14)$$

Den tatsächlichen Wert von  $\zeta(2)$  haben wir noch nicht bestimmt.

iv.) Sei  $z \in \mathbb{C}$  und  $\epsilon > 0$ . Wir betrachten die Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n = \frac{z^n}{n!^\epsilon}$ . Hier können wir das Quotientenkriterium für  $z \neq 0$  benutzen. Es gilt

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{z^{n+1}}{(n+1)!^\epsilon} \cdot \frac{n!^\epsilon}{z^n} \right| = \frac{|z|}{(n+1)^\epsilon}. \quad (2.7.15)$$

Da  $\frac{1}{n^\epsilon}$  eine Nullfolge ist, sobald  $\epsilon > 0$ , folgt

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 0. \quad (2.7.16)$$

Damit können wir auf die absolute Konvergenz der Reihe  $\sum_{n=1}^{\infty} \frac{z^n}{n!^\epsilon}$  schließen, indem wir das Quotientenkriterium aus Proposition 2.7.6, i.), anwenden.

Bei einer Folge  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  hatten wir in Proposition 2.2.11 gesehen, dass die Konvergenz nicht davon abhängt, in welcher Reihenfolge man die Folgenglieder durchläuft. Jede Umordnung von  $(a_n)_{n \in \mathbb{N}}$  liefert Konvergenz gegen den selben Grenzwert. Es zeigt sich nun, dass die Situation bei Reihen deutlich komplizierter ist. Während wir bei einer endlichen Summe die Summanden in einer beliebigen Reihenfolge aufsummieren können, da  $+$  ja kommutativ ist, hängt das Verhalten bei einer unendlichen Reihe eventuell sensitiv von der Reihenfolge ab. Dies motiviert folgende Definition:

**Definition 2.7.8 (Unbedingte Konvergenz)** Sei  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$ . Dann heißt eine Reihe  $\sum_{n=1}^{\infty} a_n$  unbedingt konvergent gegen  $s \in \mathbb{K}$ , falls für jede Bijektion  $\sigma \in \text{Bij}(\mathbb{N})$  die Umordnung  $\sum_{n=1}^{\infty} a_{\sigma(n)}$  ebenfalls gegen  $s$  konvergiert, also

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = \sum_{n=1}^{\infty} a_n \quad (2.7.17)$$

gilt. Eine Reihe heißt bedingt konvergent, wenn sie konvergent aber nicht absolut konvergent ist.

Die erste Beobachtung ist, dass absolute Konvergenz unbedingte Konvergenz nach sich zieht. Wir können in einer absolut konvergenten Reihe beliebig umordnen, ohne die Konvergenz zu ändern:

**Proposition 2.7.9** Eine absolut konvergente Reihe ist unbedingt konvergent.

BEWEIS: Sei  $\sum_{n=1}^{\infty} a_n = s$  absolut konvergent und  $\sigma \in \text{Bij}(\mathbb{N})$  eine Umordnung. Die absolute Konvergenz zeigt, dass es zu  $\epsilon > 0$  einen Index  $N \in \mathbb{N}$  mit

$$\sum_{n=N}^{\infty} |a_n| < \frac{\epsilon}{2}$$

gibt. Daher folgt also

$$\left| \sum_{n=1}^{N-1} a_n - s \right| = \left| \sum_{n=N}^{\infty} a_n \right| \leq \sum_{n=N}^{\infty} |a_n| < \frac{\epsilon}{2}.$$

Wie schon im Beweis von Proposition 2.2.11 setzen wir

$$M = \max\{\sigma^{-1}(1), \dots, \sigma^{-1}(N-1)\},$$

womit also

$$\{\sigma(1), \dots, \sigma(M)\} \supseteq \{1, \dots, N-1\} \quad (*)$$

gilt. Dann betrachten wir

$$\begin{aligned} \left| \sum_{m=1}^M a_{\sigma(m)} - s \right| &\leq \left| \sum_{m=1}^M a_{\sigma(m)} - \sum_{n=1}^{N-1} a_n \right| + \left| \sum_{n=1}^{N-1} a_n - s \right| \\ &< \sum_{n=N}^{\infty} |a_n| + \frac{\epsilon}{2} \\ &< \epsilon. \end{aligned}$$

Hier verwenden wir, dass in der Differenz  $\sum_{m=1}^M a_{\sigma(m)} - \sum_{n=1}^{N-1} a_n$  nur noch solche Folgenglieder mit Index echt größer  $N-1$  auftreten: Die ersten  $n = 1, \dots, N-1$  Terme treten alle auch in der Summe  $\sum_{m=1}^M a_{\sigma(m)}$  wegen  $(*)$  auf und heben sich daher weg. Die verbleibenden Summanden können wir dann betragsmäßig durch das Endstück  $\sum_{n=N}^{\infty} |a_n|$  abschätzen. Damit ist aber die Konvergenz gegen  $s$  gezeigt.  $\square$

**Korollar 2.7.10** *Ist  $\sum_{n=1}^{\infty} a_n$  eine absolut konvergente Reihe, so ist auch jede Umordnung davon wieder absolut konvergent.*

BEWEIS: Da  $\sum_{n=1}^{\infty} a_n$  absolut konvergent ist, konvergiert  $\sum_{n=1}^{\infty} |a_n|$ . Da  $|a_n| \geq 0$ , ist die Reihe  $\sum_{n=1}^{\infty} |a_n|$  automatisch absolut konvergent. Nach Proposition 2.7.9 konvergiert auch jede Umordnung  $\sum_{n=1}^{\infty} |a_{\sigma(n)}|$  für  $\sigma \in \text{Bij}(\mathbb{N})$ . Dies ist dann die absolute Konvergenz von  $\sum_{n=1}^{\infty} a_{\sigma(n)}$ .  $\square$

Im reellen Fall können wir bei einer bedingt konvergenten Reihe durch Umordnung nicht nur den Grenzwert verändern, sondern sogar jede reelle Zahl als Grenzwert erreichen: Dies ist gerade die Aussage des *Riemannschen Umordnungssatzes*:

**Satz 2.7.11 (Riemannscher Umordnungssatz)** *Sei  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{R})$  eine reelle Folge, so dass  $\sum_{n=1}^{\infty} a_n$  bedingt konvergiert. Dann gibt es zu jeder reellen Zahl  $s \in \mathbb{R}$  eine Umordnung  $\sigma \in \text{Bij}(\mathbb{N})$  mit*

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = s. \quad (2.7.18)$$

BEWEIS: Für eine reelle Zahl  $x \in \mathbb{R}$  schreiben wir

$$x^{\pm} = \frac{1}{2}(|x| \pm x) \geq 0.$$

Es gilt also

$$x^+ = \begin{cases} x & \text{falls } x \geq 0 \\ 0 & \text{falls } x < 0 \end{cases} \quad \text{und} \quad x^- = \begin{cases} 0 & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases}$$

sowie

$$x = x^+ - x^- \quad \text{und} \quad |x| = x^+ + x^-.$$

Wir betrachten nun die beiden Reihen  $\sum_{n=1}^{\infty} a_n^{\pm}$  und behaupten, dass beide notwendigerweise divergieren. Wäre beispielsweise  $\sum_{n=1}^{\infty} a_n^+$  konvergent, so auch

$$\sum_{n=1}^{\infty} a_n^- = - \sum_{n=1}^{\infty} a_n^+ - \sum_{n=1}^{\infty} a_n$$

und somit auch

$$\sum_{n=1}^{\infty} |a_n| = \sum_{n=1}^{\infty} a_n^+ + \sum_{n=1}^{\infty} a_n^-,$$

was aber für die bedingt konvergente Reihe  $\sum_{n=1}^{\infty} a_n$  nicht der Fall ist. Denn dann wäre sie selbst absolut konvergent und nach Proposition 2.7.9 auch unbedingt konvergent. Es müssen also beide Reihen

$$\sum_{n=1}^{\infty} a_n^{\pm} = +\infty.$$

divergieren. Man beachte, dass die Divergenz automatisch eine Divergenz gegen  $+\infty$  bedeutet, da alle Folgenglieder nicht-negativ sind. Für die ursprüngliche Reihe  $\sum_{n=1}^{\infty} a_n$  wollen wir nun zusätzlich annehmen, dass alle Summanden  $a_n \neq 0$  erfüllen. Gibt es  $a_n = 0$ , so können wir diese getrost weglassen, da sie das Konvergenzverhalten weder von  $\sum_{n=1}^{\infty} a_n$  noch von jeder Umordnung  $\sum_{n=1}^{\infty} a_{\sigma(n)}$  beeinflussen. Es folgt, dass für jedes  $n \in \mathbb{N}$  genau eine der beiden Zahlen  $a_n^+$  oder  $a_n^-$  von Null verschieden (und damit positiv) ist, während die andere verschwindet, siehe Abbildung 2.11. Wir definieren nun zwei neue Folgen, indem wir die Null-Einträge in  $(a_n^{\pm})_{n \in \mathbb{N}}$  weglassen, die wir mit  $(p_n)_{n \in \mathbb{N}}$  und  $(q_n)_{n \in \mathbb{N}}$  bezeichnen. Entsprechend gilt nun  $p_n, q_n > 0$  für alle  $n \in \mathbb{N}$  und sowohl  $\sum_{n=1}^{\infty} p_n$  als auch  $\sum_{n=1}^{\infty} q_n$  sind divergent gegen  $+\infty$ . Da  $\sum_{n=1}^{\infty} p_n = +\infty$  divergiert, gibt es einen minimalen (und damit eindeutigen) Index  $n_1 \in \mathbb{N}$  mit der Eigenschaft

$$\sum_{n=1}^{n_1} p_n > s.$$

Da auch  $\sum_{n=1}^{\infty} q_n = +\infty$  divergiert, gibt es einen minimalen (und damit eindeutigen) Index  $m_1 \in \mathbb{N}$  mit

$$\sum_{n=1}^{n_1} p_n - \sum_{n=1}^{m_1} q_n < s.$$

Induktiv konstruieren wir so Indizes  $n_1 < n_2 < \dots$  und  $m_1 < m_2 < \dots$  derart, dass

$$\sum_{n=1}^{n_1} p_n - \sum_{n=1}^{m_1} q_n + \sum_{n=n_1+1}^{n_2} p_n - \dots + \sum_{n=n_k+1}^{n_{k+1}} p_n > s$$

sowie

$$\sum_{n=1}^{n_1} p_n - \sum_{n=1}^{m_1} q_n + \dots + \sum_{n=n_k+1}^{n_{k+1}} p_n - \sum_{n=m_k+1}^{m_{k+1}} q_n < s$$

für alle  $k \in \mathbb{N}$ . Da alle Folgenglieder  $(p_n)_{n \in \mathbb{N}}$  gerade die strikt positiven Folgenglieder von  $(a_n)_{n \in \mathbb{N}}$  sind während die  $(-q_n)_{n \in \mathbb{N}}$  gerade die negativen durchlaufen, tritt in der resultierenden Reihe

$$p_1 + \dots + p_{n_1} - q_1 - \dots - q_{m_1} + p_{n_1+1} + \dots + p_{n_2} - q_{m_1+1} - \dots - q_{m_2} + \dots \quad (*)$$

jedes  $a_n$  genau einmal auf. Die so erhaltene Reihe ist also eine Umordnung der ursprünglichen Reihe, siehe auch Abbildung 2.12. Betrachten wir nun eine Teilsumme der Reihe  $(*)$  die mit  $q$ 's endet. Dann ist

$$p_1 + \dots + p_{n_1} - q_1 - \dots - q_{m_1} + \dots - q_{m_k+1} - \dots - q_i - s$$



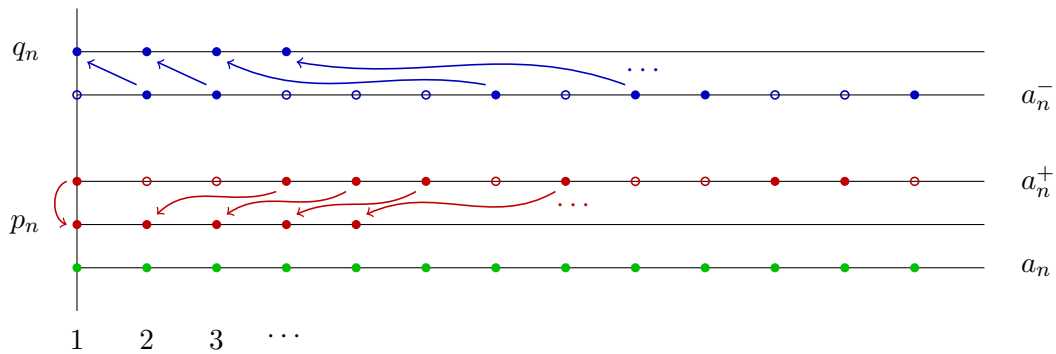


Abbildung 2.11: Die Folgen  $(a_n^+)_{n \in \mathbb{N}}$  und  $(a_n^-)_{n \in \mathbb{N}}$  enthalten viele Nulleinträge: Immer wenn  $a_n^+ \neq 0$  gilt, ist  $a_n^- = 0$  und umgekehrt.

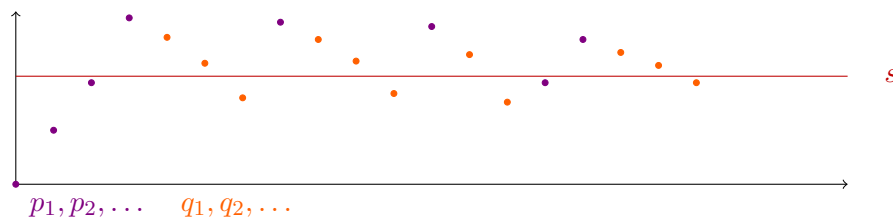


Abbildung 2.12: Die Reihe  $(*)$  steigt erst strikt monoton, bevor sie das erste Mal den Wert  $s$  übertrifft. Dann fällt sie monoton bis sie das erste mal den Wert  $s$  unterschreitet. Dies wiederholt sich nun.

positiv, solange  $i < m_{k+1}$ . Da  $p_{n_{k+1}}$  minimal gewählt wurde, ist diese Differenz aber immer kleiner  $p_{m_{k+1}}$ : Die „schlechteste“ Approximation an  $s$  erhalten wir, wenn wir noch keine  $q$ 's abgezogen haben, siehe auch Abbildung 2.13. Ist dagegen  $i = m_{k+1}$ , so wird die Differenz negativ aber betragsmäßig höchstens  $q_{m_{k+1}}$ . Endet die Partialsumme mit  $p$ 's, so können wir analog argumentieren. Da nun beide Folgen  $(p_n)_{n \in \mathbb{N}}$  und  $(q_n)_{n \in \mathbb{N}}$  Nullfolgen sind, haben wir die Konvergenz gegen  $s$  gezeigt.  $\square$

**Bemerkung 2.7.12** Mit einer leichten Modifikation des Beweises kann man für eine bedingt konvergente Reihe  $\sum_{n=1}^{\infty} a_n$  Umordnungen  $\sigma, \sigma' \in \text{Bij}(\mathbb{N})$  finden, so dass  $\sum_{n=1}^{\infty} a_{\sigma(n)}$  gegen  $+\infty$  divergiert, während  $\sum_{n=1}^{\infty} a_{\sigma'(n)}$  gegen  $-\infty$  divergiert, siehe auch Übung ??.

Die Bedeutung des Umordnungssatzes liegt nicht so sehr in seiner direkten Aussage, sondern in den konzeptuellen Konsequenzen, wie beispielsweise folgender:

**Korollar 2.7.13** Eine Reihe  $\sum_{n=1}^{\infty} a_n$  mit  $a_n \in \mathbb{K}$  konvergiert genau dann absolut, wenn sie unbedingt konvergiert.

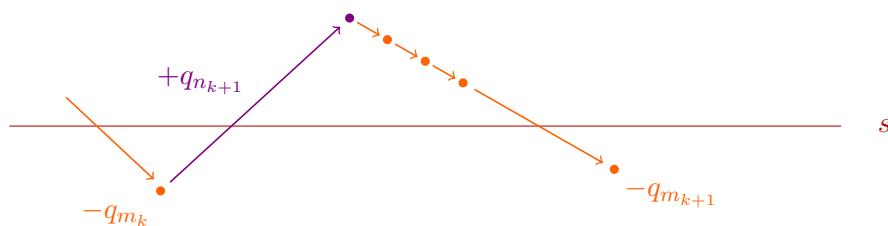


Abbildung 2.13: Die Reihe konvergiert gegen  $s$ .

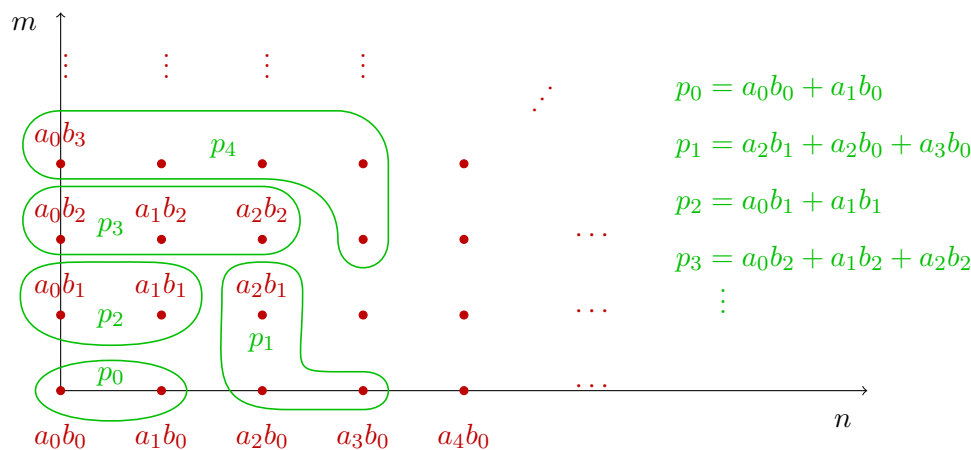


Abbildung 2.14: Die einzelnen Beiträge beim unendlichen Ausmultiplizieren geben viele Möglichkeiten zur Summation.

BEWEIS: Die eine Implikation ist gerade Proposition 2.7.2. Sei  $\sum_{n=1}^{\infty} a_n$  nicht absolut konvergent. Dann divergiert also  $\sum_{n=1}^{\infty} |a_n|$  gegen  $+\infty$ . Ist  $(a_n)_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{C})$  eine komplexe Folge, folgt aus

$$|\text{Re}(z)|, |\text{Im}(z)| \leq |z| \leq |\text{Re}(z)| + |\text{Im}(z)|$$

dass mindestens eine der beiden Reihen  $\sum_{n=1}^{\infty} \text{Re}(a_n)$  oder  $\sum_{n=1}^{\infty} \text{Im}(a_n)$  nicht absolut konvergiert. Nach dem Umordnungssatz angewandt auf diese *reelle* Reihe finden wir daher eine Umordnung  $\sigma \in \text{Bij}(\mathbb{N})$ , so dass die umgeordnete komplexe Reihe  $\sum_{n=1}^{\infty} a_{\sigma(n)}$  einen anderen Real- beziehungsweise Imaginärteil besitzt. Damit ist die Reihe aber nicht unbedingt konvergent gewesen. Der reelle Fall  $a_n \in \mathbb{R}$  ist eine direkte Folge aus Satz 2.7.11.  $\square$

**Bemerkung 2.7.14** Als Ausblick auf die Funktionalanalysis sei hier erwähnt, dass Reihen auch für andere Koeffizienten als  $a_n \in \mathbb{K}$  in der Mathematik von Interesse sind. Es zeigt sich, dass dann die beiden Begriffe der absoluten und unbedingten Konvergenz im Allgemeinen durchaus verschieden sein können. Absolute Konvergenz impliziert unbedingte Konvergenz, aber nicht notwendigerweise umgekehrt.

Als eine vorerst letzte wichtige Anwendung von absoluter Konvergenz wollen wir die Multiplikation von Reihen betrachten. Für endliche Summen gilt

$$(a_1 + \cdots + a_N)(b_1 + \cdots + b_M) = \sum_{i=1}^N \sum_{j=1}^M a_i b_j, \quad (2.7.19)$$

wobei wir in der Doppelsumme auf der rechten Seite beliebig umordnen können, da die Addition kommutativ ist. Für unendliche Reihen stellt sich nun die Frage, wie das „Ausmultiplizieren“ und Aufsummieren zu geschehen hat. Sind zwei Reihen  $\sum_{n=0}^{\infty} a_n$  und  $\sum_{m=0}^{\infty} b_m$  vorgegeben, so können wir die Produkte  $a_n b_m$  in einem unendlichen quadratischen Schema anordnen, siehe Abbildung 2.14.

Wir können nun immer endlich viele Produkte  $a_i b_j$  zu einer Summe zusammenfassen, und erhalten dann eine Folge von Summen  $p_0, p_1, p_2, \dots$  derart, dass jedes Produkt als Summand in genau einer solchen Summe  $p_k$  auftritt. Insbesondere können wir für die  $(p_k)_{k \in \mathbb{N}_0}$  einfach eine Abzählung von dem quadratischen Schema aus Abbildung 2.14 verwenden. In diesem Fall wäre also jedes  $p_k$  eine Summe mit nur einem Summanden. Ebenfalls naheliegend wäre die Wahl

$$p_n = \sum_{i+j=n} a_i b_j = \sum_{i=0}^n a_i b_{n-i}, \quad (2.7.20)$$

die in Abbildung 2.14 längs der Diagonalen summiert. Egal wie man vorgeht, am Ende stellt sich die Frage, ob die Reihe der  $p_n$  konvergiert und

$$\sum_{n=0}^{\infty} p_n = \left( \sum_{k=0}^{\infty} a_k \right) \left( \sum_{\ell=0}^{\infty} b_{\ell} \right). \quad (2.7.21)$$

erfüllt. Im Allgemeinen ist dies falsch, aber absolute Konvergenz ist die einfache hinreichende Bedingung:

**Satz 2.7.15 (Multiplikation von Reihen)** *Seien  $\sum_{n=0}^{\infty} a_n$  und  $\sum_{m=0}^{\infty} b_m$  zwei absolut konvergente Reihen mit  $a_n, b_m \in \mathbb{K}$ . Ist dann  $(p_n)_{n \in \mathbb{N}_0}$  eine Folge von endlichen Summen der  $a_i b_j$ , so dass jedes Produkt  $a_i b_j$  in genau einem Folgenglied  $p_n$  genau einmal auftritt, dann konvergiert die Reihe  $\sum_{n=0}^{\infty} p_n$  ebenfalls absolut, und es gilt*

$$\sum_{n=0}^{\infty} p_n = \left( \sum_{k=0}^{\infty} a_k \right) \left( \sum_{\ell=0}^{\infty} b_{\ell} \right). \quad (2.7.22)$$

BEWEIS: Besteht ein  $p_n$  aus mehr als einem Summanden, so können wir diese einzeln hinzunehmen und erhalten eine neue Abzählung auf der linken Seite mit jeweils einem Summanden in einem  $p_n$ , ohne das Konvergenzverhalten zu ändern. Wir werden daher ohne Einschränkung annehmen, dass die Folge  $(p_n)_{n \in \mathbb{N}_0}$  eine Abzählung der Produkte  $(a_i b_j)_{i,j \in \mathbb{N}_0}$  ist. Es gibt also zu jedem  $N \in \mathbb{N}$  ein  $M \in \mathbb{N}$  derart, dass

$$p_0, \dots, p_N \in \{a_i b_j\}_{i,j=0,\dots,M}$$

gilt. Dann gilt

$$|p_0| + \dots + |p_N| \leq (|a_0| + \dots + |a_M|)(|b_0| + \dots + |b_M|),$$

da die rechte Seite im Allgemeinen mehr positive Summanden besitzt, aber alle Terme der linken Seite auftreten. Da die beiden Reihen  $\sum_{n=0}^{\infty} a_n$  und  $\sum_{n=0}^{\infty} b_n$  absolut konvergieren, folgt

$$|p_0| + \dots + |p_N| \leq \left( \sum_{n=0}^{\infty} |a_n| \right) \left( \sum_{m=0}^{\infty} |b_m| \right)$$

für alle  $N \in \mathbb{N}$ . Dies zeigt

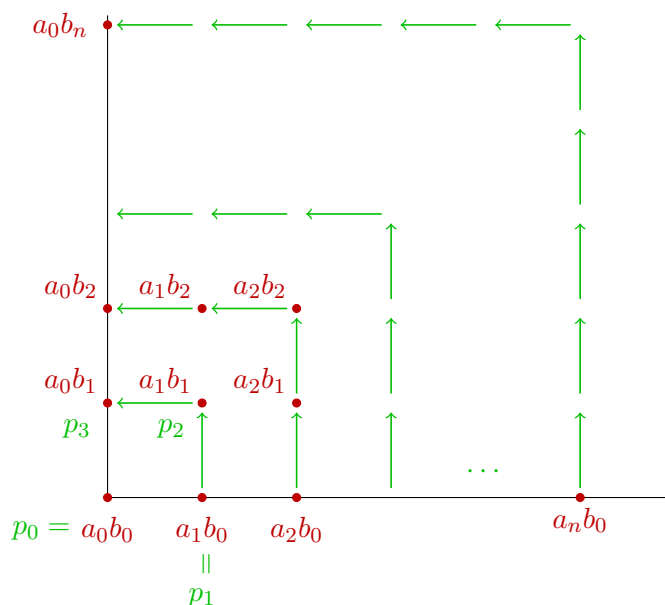
$$\sum_{n=0}^{\infty} |p_n| \leq \left( \sum_{k=0}^{\infty} |a_k| \right) \left( \sum_{\ell=0}^{\infty} |b_{\ell}| \right)$$

und somit die absolute Konvergenz der Reihe  $\sum_{n=0}^{\infty} p_n$ . Nach Proposition 2.7.9 folgt, dass alle Wahlen einer Abzählung der Produkte  $a_i b_j$  zu einer konvergenten Reihe mit gleichem Resultat führen. Es bleibt also zu zeigen, dass der Wert der Reihe  $\sum_{n=0}^{\infty} p_n$  gerade (2.7.22) erfüllt. Um dies zu sehen, betrachten wir eine spezielle Abzählung, nämlich wie in Abbildung 2.15. Dann gilt

$$p_0 + p_1 + \dots + p_{(n+1)^2-1} = (a_0 + \dots + a_n)(b_0 + \dots + b_n),$$

da wir mit den ersten  $(n+1)^2 - 1$  Termen von  $(p_n)_{n \in \mathbb{N}_0}$  gerade alle Produkte der ausmultiplizierten rechten Seite erreichen. Diese Teilfolge der Partialsummen von  $\sum_{n=0}^{\infty} p_n$  konvergiert aber gegen

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{n=0}^{(N+1)^2-1} &= \lim_{N \rightarrow \infty} \left( \sum_{k=0}^N a_k \right) \left( \sum_{\ell=0}^N b_{\ell} \right) \\ &= \left( \lim_{N \rightarrow \infty} \sum_{k=0}^N a_k \right) \left( \lim_{N \rightarrow \infty} \sum_{\ell=0}^N b_{\ell} \right) \end{aligned}$$



Abbildungung 2.15: Die spezielle Abzählung folgt parallel zu den Achsen längs der grünen Pfeile.

$$= \left( \sum_{k=0}^{\infty} a_k \right) \left( \sum_{\ell=0}^{\infty} b_{\ell} \right).$$

Da wir bereits wissen, dass  $\sum_{k=0}^{\infty} p_n$  konvergiert, genügt es, die Konvergenz einer Teilfolge der Partialsummen auszuwerten, um den Wert der Reihe zu bestimmen. Damit ist also (2.7.22) gezeigt.  $\square$

**Korollar 2.7.16 (Cauchy-Produkt)** Sind  $\sum_{n=0}^{\infty} a_n$  und  $\sum_{n=0}^{\infty} b_n$  absolut konvergente Reihen mit  $a_n, b_n \in \mathbb{K}$ , so konvergiert ihr Cauchy-Produkt  $\sum_{n=0}^{\infty} c_n$  mit

$$c_n = \sum_{k=0}^n a_k b_{n-k} \quad (2.7.23)$$

ebenfalls absolut mit Wert

$$\sum_{n=0}^{\infty} c_n = \left( \sum_{k=0}^{\infty} a_k \right) \left( \sum_{\ell=0}^{\infty} b_{\ell} \right). \quad (2.7.24)$$

## 2.8 Netze

Folgen indizieren mathematisch interessante Größen, wie beispielsweise reelle oder komplexe Zahlen, mit natürlichen Zahlen. Nimmt man den Standpunkt der Abbildungen  $\text{Abb}(\mathbb{N}, \mathbb{K})$  ein, so wird schnell klar, dass auch andere Parameterbereiche  $I$  als  $\mathbb{N}$  von Interesse sein können. Um nun von Konvergenz sprechen zu können, brauchen wir auf der Indexmenge einen Begriff von „größer“ oder „später“, da unsere Konvergenzbegriffe im Falle von  $I = \mathbb{N}$  ja explizit von der Ordnung „ $\leq$ “ auf  $\mathbb{N}$  Gebrauch machen. Um dies gut verallgemeinern zu können, bietet sich der Begriff der *gerichteten Menge* an:

**Definition 2.8.1 (Gerichtete Menge)** Sei  $I$  eine nicht-leere Menge mit einer partiell definierten Relation  $\preceq$ . Dann heißt  $(I, \preceq)$  gerichtete Menge, falls  $\preceq$  folgende Eigenschaften erfüllt:

i.) Es gilt  $i \preceq i$  für alle  $i \in I$ .

Abbildung 2.16: Schematische Darstellung einer gerichteten Menge. Nicht alle Punkte sind vergleichbar, aber es gibt immer gemeinsame Nachfolger.

- ii.) Gelten für  $i, j, k \in I$  die Relationen  $i \preccurlyeq j$  und  $j \preccurlyeq k$ , so gilt auch  $i \preccurlyeq k$ .
- iii.) Für alle  $i, j \in I$  gibt es ein  $k \in I$  mit  $i \preccurlyeq k$  und  $j \preccurlyeq k$ .

Die Relation  $\preccurlyeq$  heißt auch *Richtung*. Gilt  $i \preccurlyeq j$ , so heißt  $j$  *später* als  $i$  beziehungsweise  $i$  *früher* als  $j$ .

**Bemerkung 2.8.2 (Gerichtete Mengen)** Sei  $(I, \preccurlyeq)$  eine gerichtete Menge.

- i.) Wir nennen  $i \in I$  und  $j \in I$  *vergleichbar*, wenn  $i \preccurlyeq j$  oder  $j \preccurlyeq i$  gilt. Da die Relation  $\preccurlyeq$  nur partiell definiert ist, kann es durchaus vorkommen, dass zwei verschiedene Elemente  $i$  und  $j$  nicht vergleichbar sind. Abbildung 2.16 liefert hier eine graphische Intuition.
- ii.) Erfüllt  $\preccurlyeq$  nur die ersten beiden Eigenschaften von Definition 2.8.1, so nennt man  $(I, \preccurlyeq)$  auch *prä-geordnete Menge*.
- iii.) Im Allgemeinen fordert man *nicht*, dass  $i \preccurlyeq j$  und  $j \preccurlyeq i$  nur für  $i = j$  gültig ist. Gilt diese Eigenschaft für eine prä-geordnete Menge, so nennt man  $(I, \preccurlyeq)$  eine *partiell geordnete Menge*. Eine partiell geordnete Menge  $(I, \preccurlyeq)$  heißt *geordnete Menge*, falls je zwei Elemente  $i, j \in I$  auch vergleichbar sind. Wir benötigen diese Eigenschaften jedoch nicht. Für viele Beispiele sind sie zu restriktiv. Wichtig ist dagegen die dritte Eigenschaft. Auch wenn wir  $i$  und  $j$  eventuell nicht vergleichen können, so gibt es immer ein weiteres Element  $h$ , welches später als die beiden Elemente  $i$  und  $j$  ist.

Wir kommen nun zu einigen wichtigen Beispielen für gerichtete Mengen:

**Beispiel 2.8.3 (Gerichtete Mengen)**

- i.) Die natürlichen Zahlen  $(\mathbb{N}, \preccurlyeq)$  mit  $\preccurlyeq = \leq$  sind eine gerichtete Menge. Zudem ist die *Kleiner-gleich*-Relation eine Ordnung. Dieses Beispiel liegt dem Folgenbegriff zugrunde.
- ii.) Ebenfalls bezüglich  $\leq$  sind auch  $\mathbb{Z}, \mathbb{Q}$  und  $\mathbb{R}$  gerichtete Mengen. Auch hier ist  $\leq$  sogar eine Ordnung, was wir aber im Folgenden nicht verwenden müssen.
- iii.) Sei  $M$  eine Menge und  $2^M$  ihre Potenzmenge, also die Menge aller Teilmengen von  $M$ . Die Relation  $U \subseteq V$  für Teilmengen  $U, V \in 2^M$  liefert dann eine gerichtete Menge  $(2^M, \subseteq)$ , da zu je zwei Teilmengen  $U, V \subseteq M$  eine größere Teilmenge  $X \subseteq M$  mit  $U \subseteq X$  und  $V \subseteq X$  existiert, beispielsweise  $U \cup V$  oder einfach  $M$ . Die *Enthaltensein*-Relation ist zudem eine partielle Ordnung, aber im Allgemeinen keine Ordnung.
- iv.) Ist  $I$  eine nicht-leere Menge, so definiert man  $\preccurlyeq$  dadurch, dass  $i \preccurlyeq j$  für alle Paare  $(i, j)$  gilt. Diese sehr langweilige Richtung heißt *triviale Richtung*. Allgemein nennen wir eine Richtung  $\preccurlyeq$  schwächer als eine andere  $\leq$ , falls  $i \leq j$  impliziert, dass auch  $i \preccurlyeq j$  gilt. In diesem Sinne ist die triviale Richtung die *schwächste Richtung* auf der Menge  $I$ .
- v.) Sind  $(I, \preccurlyeq_I)$  und  $(J, \preccurlyeq_J)$  zwei gerichtete Mengen, so wird das kartesische Produkt  $I \times J$  wieder eine gerichtete Menge, indem man

$$(i, j) \preccurlyeq_{I \times J} (i', j') \quad \text{falls} \quad i \preccurlyeq_I i' \text{ und } j \preccurlyeq_J j' \quad (2.8.1)$$

definiert. Eine schnelle Verifikation zeigt, dass dies tatsächlich eine gerichtete Menge definiert. Diese Richtung nennen wir die *kanonische Richtung* auf dem Produkt  $I \times J$ .

- vi.) Einen Spezialfall von v.) erhält man für eine gerichtete Menge  $(I, \preccurlyeq)$  und eine beliebige nichtleere Menge  $X$ . Wir versehen  $X$  mit der trivialen Richtung, dann ist  $I \times X$  gerichtet, wobei  $(i, x) \preccurlyeq$

$(i', x')$  genau dann gilt, wenn  $i \preccurlyeq i'$  gilt. Die zusätzlichen Komponenten  $x, x'$  spielen also in diesem Fall eine Zuschauerrolle. Überraschenderweise wird dieses Beispiel bei der Konstruktion von Riemann-Integralen eine entscheidende Rolle spielen.

Wir können nun den Begriff einer Folge zu dem eines Netzes verallgemeinern und erhalten direkt eine passende Definition von Konvergenz:

**Definition 2.8.4 (Netze und Konvergenz)** Sei  $(I, \preccurlyeq)$  eine gerichtete Menge.

- i.) Ein Netz von reellen oder komplexen Zahlen ist eine Abbildung  $a \in \text{Abb}(I, \mathbb{K})$ . Wir schreiben  $a = (a_i)_{i \in I}$  mit  $a_i = a(i)$  für  $i \in I$  wie zuvor bei Folgen.
- ii.) Ein Netz  $a = (a_i)_{i \in I} \in \text{Abb}(I, \mathbb{K})$  konvergiert gegen  $\alpha \in \mathbb{K}$ , falls es zu jedem  $\epsilon > 0$  einen Index  $i \in I$  gibt, so dass

$$|a_j - \alpha| < \epsilon \quad (2.8.2)$$

für alle  $j \in I$  mit  $i \preccurlyeq j$ . In diesem Fall schreiben wir

$$\alpha = \lim_{i \in I} a_i \quad (2.8.3)$$

und nennen  $\alpha$  den Grenzwert (Limes) von  $a$ .

Mit dieser Definition ist klar, dass wir den bisherigen Begriff der Konvergenz im Beispiel  $I = \mathbb{N}$  reproduzieren. Somit haben wir eine drastische Verallgemeinerung gefunden.

Es stellt sich aber die Frage, wieso wir eine gerichtete Menge fordern und nicht etwa nur eine prä-geordnete. Die folgende Proposition illustriert dies eindrücklich:

**Proposition 2.8.5** Sei  $(I, \preccurlyeq)$  eine gerichtete Menge und  $a = (a_i)_{i \in I} \in \text{Abb}(I, \mathbb{K})$  ein Netz. Besitzt  $a$  einen Grenzwert, so ist dieser eindeutig.

BEWEIS: Wir nehmen an, dass  $\alpha \neq \alpha'$  zwei verschiedene Grenzwerte von  $a$  sind. Dann betrachten wir  $\epsilon = \frac{1}{2}|\alpha - \alpha'| > 0$  und die zugehörigen Konvergenzbedingungen. Wir finden demnach  $i, i' \in I$  mit

$$|a_j - \alpha| < \epsilon \quad \text{und} \quad |a_{j'} - \alpha'| < \epsilon$$

für  $i \preccurlyeq j$  und  $i' \preccurlyeq j'$ . Da  $I$  gerichtet ist, finden wir sogar ein  $k \in I$  mit  $i \preccurlyeq k$  und  $i' \preccurlyeq k$ . Für  $j \succcurlyeq k$  gilt dann  $j \succcurlyeq i$  und  $j \succcurlyeq i'$ , also

$$|a_j - \alpha| < \epsilon \quad \text{und} \quad |a_j - \alpha'| < \epsilon.$$

Insbesondere können wir  $j = k$  wählen. Dies ist aber ein Widerspruch, da

$$|\alpha - \alpha'| \leq |\alpha - a_j| + |a_j - \alpha'| < \epsilon + \epsilon = |\alpha - \alpha'|.$$

Es kann also höchstens einen Grenzwert geben. □

Wir können also auch bei allgemeinen Netzen von dem Grenzwert sprechen, sobald überhaupt Konvergenz vorliegt.

Als nächstes wollen wir verstehen, wie die Rechenregeln für konvergente Folgen auf konvergente Netze übertragen werden können. Als Vorüberlegung betrachten wir hierzu besondere „Teilfolgen“ eines Netzes:

**Proposition 2.8.6** Sei  $(I, \preccurlyeq)$  eine gerichtete Menge und  $a = (a_i)_{i \in I} \in \text{Abb}(I, \mathbb{K})$  ein Netz. Konvergiert  $a$  gegen  $\alpha \in \mathbb{K}$ , so gibt es eine Abbildung

$$\sigma: \mathbb{N} \longrightarrow I \quad (2.8.4)$$

derart, dass die Folge  $a \circ \sigma = (a_{\sigma(n)})_{n \in \mathbb{N}} \in \text{Abb}(\mathbb{N}, \mathbb{K})$  gegen  $\alpha$  konvergiert.

BEWEIS: Sei  $n \in \mathbb{N}$  gegeben. Dann betrachten wir für  $\epsilon = \frac{1}{n} > 0$  die Konvergenzbedingung und erhalten einen Index  $i_n \in I$  derart, dass

$$|a_j - \alpha| < \epsilon = \frac{1}{n}$$

für alle  $j \succ i_n$  gilt. Setzen wir nun  $\sigma(n) = i_n$ , so liefert dies eine Folge  $(a_{\sigma(n)})_{n \in \mathbb{N}}$  welche gegen  $\alpha$  konvergiert. Ist nämlich  $\epsilon > 0$  beliebig vorgegeben, so wählen wir  $N \in \mathbb{N}$  mit  $\frac{1}{N} < \epsilon$ . Dann gilt für alle  $n \geq N$  die Abschätzung

$$|a_{\sigma(n)} - \alpha| < \frac{1}{n} \leq \frac{1}{N} < \epsilon,$$

was die gewünschte Folgenkonvergenz bedeutet.  $\square$

Wir können uns also bei einem konvergenten Netz auch immer mit einer Folge von Elementen des Netzes an den Grenzwert heranpirschen. Es sei hier jedoch angemerkt, dass eine derartige Folge nicht der üblichen Definition einer Teilfolge des Netzes entspricht. Wir wollen die richtige Definition an dieser Stelle nicht diskutieren, sondern verweisen beispielsweise auf [7, Sect. 4.2].

Mit den selben Ideen wie in Abschnitt 2.3 erhalten wir nun die folgenden Rechenregeln für Netze:

**Proposition 2.8.7** *Sei  $(I, \preccurlyeq)$  eine gerichtete Menge, und seien  $a = (a_i)_{i \in I}, b = (b_i)_{i \in I} \in \text{Abb}(I, \mathbb{K})$  konvergente Netze mit Grenzwerten  $\alpha$  und  $\beta$ .*

i.) *Für alle  $z, w \in \mathbb{K}$  ist das Netz  $za + wb = (za_i + wb_i)_{i \in I}$  wieder konvergent und erfüllt*

$$\lim_{i \in I} (za_i + wb_i) = z \lim_{i \in I} a_i + w \lim_{i \in I} b_i. \quad (2.8.5)$$

ii.) *Das Netz  $ab = (a_i b_i)_{i \in I}$  ist konvergent und erfüllt*

$$\lim_{i \in I} a_i b_i = (\lim_{i \in I} a_i) (\lim_{i \in I} b_i). \quad (2.8.6)$$

iii.) *Das Netz  $\bar{a} = (\bar{a}_i)_{i \in I}$  ist konvergent mit*

$$\lim_{i \in I} \bar{a}_i = \overline{\lim_{i \in I} a_i}. \quad (2.8.7)$$

iv.) *Gilt  $b_i \neq 0$  für alle  $i \in I$  und  $\beta \neq 0$ , so ist das Netz  $\frac{a}{b} = (\frac{a_i}{b_i})_{i \in I}$  konvergent mit*

$$\lim_{i \in I} \frac{a_i}{b_i} = \frac{\lim_{i \in I} a_i}{\lim_{i \in I} b_i}. \quad (2.8.8)$$

v.) *Gilt  $a_i \leq b_i$  für alle  $i \in I$ , so gilt auch*

$$\lim_{i \in I} a_i \leq \lim_{i \in I} b_i. \quad (2.8.9)$$

vi.) *Ist  $a \in \text{Abb}(I, \mathbb{R})$  reell und  $x > 0$ , so ist das Netz  $(x^{a_i})_{i \in I}$  konvergent mit*

$$\lim_{i \in I} x^{a_i} = x^{\lim_{i \in I} a_i}. \quad (2.8.10)$$

vii.) *Gilt  $a_i > 0$  für alle  $i \in I$  und  $\alpha > 0$ , so ist das Netz  $(a_i^x)_{i \in I}$  für alle  $x \in \mathbb{R}$  konvergent mit*

$$\lim_{i \in I} a_i^x = (\lim_{i \in I} a_i)^x. \quad (2.8.11)$$

BEWEIS: Der Beweis folgt den Argumenten in Proposition 2.3.5, Proposition 2.3.8, Proposition 2.3.10 und Proposition 2.3.11 mit geringfügigen Modifikationen. Wir werden die Details hier nicht ausführen sondern auf Übung 2.9.18 verweisen.  $\square$

Im Hinblick auf die Rechenregeln stellt sich nun berechtigterweise die Frage, worin sich Netze von Folgen unterscheiden und für welche Situationen ihr Gebrauch unausweichlich wird. Einen ersten Eindruck gibt folgender Satz zur Vertauschung von Grenzwerten.

**Satz 2.8.8 (Vertauschen von Netz-Limites)** *Seien  $(I, \preccurlyeq_I)$  und  $(J, \preccurlyeq_J)$  zwei gerichtete Mengen und sei  $I \times J$  mit einer Richtung versehen, die schwächer als die kanonische Richtung  $\preccurlyeq_{I \times J}$  ist. Sei weiter  $(a_{ij})_{(i,j) \in I \times J} \in \text{Abb}(I \times J, \mathbb{K})$  ein konvergentes Netz.*

*i.) Sind die Netze  $(a_{ij})_{i \in I} \in \text{Abb}(I, \mathbb{K})$  für alle  $j \in J$  konvergent mit Grenzwerten  $a_j = \lim_{i \in I} a_{ij}$ , so ist auch das resultierende Netz  $(a_j)_{j \in J} \in \text{Abb}(J, \mathbb{K})$  der Grenzwerte konvergent, und es gilt*

$$\lim_{j \in J} a_j = \lim_{j \in J} \left( \lim_{i \in I} a_{ij} \right) = \lim_{(i,j) \in I \times J} a_{ij}. \quad (2.8.12)$$

*ii.) Sind die Netze  $(a_{ij})_{j \in J} \in \text{Abb}(J, \mathbb{K})$  für alle  $i \in I$  konvergent mit Grenzwerten  $a_i = \lim_{j \in J} a_{ij}$ , so ist auch das resultierende Netz  $(a_i)_{i \in I} \in \text{Abb}(I, \mathbb{K})$  der Grenzwerte konvergent, und es gilt*

$$\lim_{i \in I} a_i = \lim_{i \in I} \left( \lim_{j \in J} a_{ij} \right) = \lim_{(i,j) \in I \times J} a_{ij}. \quad (2.8.13)$$

*iii.) Gelten die Voraussetzungen von i.) und ii.), so gilt*

$$\lim_{i \in I} \left( \lim_{j \in J} a_{ij} \right) = \lim_{j \in J} \left( \lim_{i \in I} a_{ij} \right). \quad (2.8.14)$$

BEWEIS: Wir zeigen i.). Der zweite Teil erfolgt dann genauso, und der dritte ist damit eine einfache Konsequenz. Sei also  $\epsilon > 0$  vorgegeben. Die Konvergenz von  $(a_{ij})_{(i,j) \in I \times J}$  gegen  $\alpha \in \mathbb{K}$  bedeutet dann, dass es ein  $(i_0, j_0) \in I \times J$  gibt, sodass

$$|a_{ij} - \alpha| < \epsilon \quad (*)$$

für alle  $(i, j) \succcurlyeq (i_0, j_0)$ . Da die Richtung auf  $I \times J$  schwächer als die kanonische Richtung sein soll, liefert (\*), dass für festes  $i \succcurlyeq i_0$  schon

$$|a_{ij} - \alpha| < \epsilon \quad (**)$$

für alle  $j \succcurlyeq j_0$  gilt. Nach Annahme existieren die Netzlimits  $a_i = \lim_{j \in J} a_{ij}$  für alle festen  $i \in I$ , womit aus (\*\*) die Ungleichung

$$|a_i - \alpha| = \left| \lim_{j \in J} a_{ij} - \alpha \right| = \lim_{j \in J} |a_{ij} - \alpha| \leq \epsilon \quad (\odot)$$

folgt, da die Netzlimits mit algebraischen Operationen und Ungleichungen verträglich sind. Man beachte, dass in ( $\odot$ ) jetzt nur ein  $\leq$  steht. Für die Konvergenz von  $(a_i)_{i \in I}$  ist dies jedoch unerheblich, da wir sowieso über alle  $\epsilon > 0$  verfügen. Daher zeigt ( $\odot$ ), dass  $(a_i)_{i \in I}$  gegen  $\alpha$  konvergiert, womit i.) und somit auch ii.) und iii.) gezeigt sind.  $\square$

Wollte man einen solchen Satz allein für Folgen formulieren, bräuchte man also Doppelfolgen  $(a_{nm})_{n,m \in \mathbb{N}}$ , für die man zunächst einen Konvergenzbegriff im Sinne einer Doppelfolge formulieren müsste. Dies wäre gerade die Konvergenz des Netzes  $(a_{ij})_{(i,j) \in I \times J}$ . Auch wenn dies möglich ist, liegt der Vorteil der Sprache der Netze auf der Hand: das Produkt  $I \times J$  kann direkt wieder als gerichtete Menge aufgefasst werden, womit man keinen neuen Konvergenzbegriff benötigt. Mit Folgen dagegen ist  $\mathbb{N} \times \mathbb{N}$  nicht mehr gleich zur Indexmenge  $\mathbb{N}$  mit ihrer üblichen Ordnung, sondern eben nur noch eine gerichtete Menge.

Wir werden einige interessante Anwendungen von Satz 2.8.8 sehen, wenn wir verschiedenste Grenzprozesse vertauschen wollen. Zum Abschluss sei noch angemerkt, dass wir in Satz 2.8.8 insbesondere die kanonische Richtung  $\preccurlyeq_{I \times J}$  verwenden können.



## 2.9 Übungen

Übung 2.9.1 (Punktweise Operationen für Funktionen)

Übung 2.9.2 (Partielle Ordnung)

Übung 2.9.3 (Eigenschaften des pull-backs)

Übung 2.9.4 (Fibonacci-Folge)

Übung 2.9.5 (Lineare Rekursionen)

Übung 2.9.6 (Die rationalen Zahlen sind abzählbar)

Übung 2.9.7 (Hilberts Hotel)

Übung 2.9.8 (Die algebraische Zahlen sind abzählbar)

Übung 2.9.9 (Kleine Bijektionen)

**Übung 2.9.10 (Konvergenz mit Nullfolgen)** Sei  $(\epsilon_n)_{n \in \mathbb{N}}$  eine Nullfolge mit positiven Folgengliedern  $\epsilon_n > 0$  für alle  $n \in \mathbb{N}$ . Sei weiter  $(a_n)_{n \in \mathbb{N}}$  eine beliebige Folge in  $\mathbb{K}$ .

i.) Zeigen Sie, dass  $(a_n)_{n \in \mathbb{N}}$  genau dann gegen  $\alpha \in \mathbb{K}$  konvergiert, wenn es zu jedem  $n \in \mathbb{N}$  ein  $N \in \mathbb{N}$  gibt, so dass

$$|a_m - \alpha| < \epsilon_n \quad (2.9.1)$$

für alle  $m \geq N$  gilt.

ii.) Folgern Sie, dass  $(a_n)_{n \in \mathbb{N}}$  genau dann gegen  $\alpha \in \mathbb{K}$  konvergiert, wenn zu jedem  $n \in \mathbb{N}$  ein  $N \in \mathbb{N}$  existiert, so dass

$$|a_m - \alpha| < \frac{1}{n} \quad (2.9.2)$$

für alle  $m \geq N$  gilt.

Insbesondere muss die Konvergenz von Folgen also nicht für alle  $\epsilon > 0$  geprüft werden. Es reicht vielmehr, Folgenglieder einer Nullfolge zu verwenden. Damit genügen insbesondere auch *rationale*  $\epsilon > 0$ , um Konvergenz zu entscheiden.

**Übung 2.9.11 (Divergenz von Folgen)** Verwenden Sie die Quantorenschreibweise, um ein Kriterium für die Divergenz einer Folge zu erhalten.

**Übung 2.9.12 (Besondere Folgen)** Sei  $x \in \mathbb{R}$  eine vorgegebene reelle Zahl. Zeigen Sie, dass es dann Folgen  $(a_n)_{n \in \mathbb{N}}$  und  $(b_n)_{n \in \mathbb{N}}$  von reellen Zahlen gibt, die gegen  $x$  konvergieren und

$$a_n \in \mathbb{Q} \quad \text{und} \quad a_n \neq 0 \quad (2.9.3)$$

sowie

$$b_n \in \mathbb{R} \setminus \mathbb{Q} \quad \text{und} \quad b_n \neq 0 \quad (2.9.4)$$

für alle  $n \in \mathbb{N}$  erfüllen. Dies zeigt erneut die Reichhaltigkeit von sowohl rationalen als auch irrationalen Zahlen.

Hinweis: Sei  $n \in \mathbb{N}$ . Finden Sie in  $B_{\frac{1}{n}}(x)$  sowohl irrationale als auch rationale Zahlen.

**Übung 2.9.13 (Cauchy-Folgen mit Nullfolgen)** Sei  $(\epsilon_n)_{n \in \mathbb{N}}$  eine Nullfolge mit positiven Folgengliedern  $\epsilon_n > 0$  für alle  $n \in \mathbb{N}$ . Sei weiter  $(a_n)_{n \in \mathbb{N}}$  eine beliebige Folge in  $\mathbb{K}$ .

- i.) Zeigen Sie, dass  $(a_n)_{n \in \mathbb{N}}$  genau dann eine Cauchy-Folge ist, wenn es zu jedem  $n \in \mathbb{N}$  ein  $N \in \mathbb{N}$  gibt, so dass

$$|a_m - a_{m'}| < \epsilon_n \quad (2.9.5)$$

für alle  $m, m' \geq N$  gilt.

- ii.) Folgern Sie, dass  $(a_n)_{n \in \mathbb{N}}$  genau dann eine Cauchy-Folge ist, wenn zu jedem  $n \in \mathbb{N}$  ein  $N \in \mathbb{N}$  existiert, so dass

$$|a_m - a_{m'}| < \frac{1}{n} \quad (2.9.6)$$

für alle  $m, m' \geq N$  gilt.

Wie auch in Übung 2.9.10 zur Konvergenz von Folgen muss also auch die Cauchy-Eigenschaft von Folgen nicht für alle  $\epsilon > 0$  geprüft werden. Es genügt vielmehr, Folgenglieder einer Nullfolge zu verwenden. Damit genügen insbesondere auch hier *rationale*  $\epsilon > 0$ , um die Cauchy-Eigenschaft zu entscheiden.

**Übung 2.9.14 (Offene Mengen und offene Kugeln)** Sei  $O \subseteq \mathbb{K}$  eine Teilmenge.

- i.) Zeigen Sie, dass  $O$  genau dann offen ist, wenn  $O$  eine Vereinigung von offenen Kugeln ist.  
 ii.) Zeigen Sie, dass eine offene Menge  $O$  sogar eine Vereinigung von offenen Kugeln der Form  $B_r(q)$  mit  $r \in \mathbb{Q}^+ \subseteq \mathbb{R}^+$  und  $q \in \mathbb{Q} \subseteq \mathbb{R}$  beziehungsweise  $q \in \mathbb{Q} + i\mathbb{Q} \subseteq \mathbb{C}$  ist.

Hinweis: Hier gibt der Beweis von Satz 2.5.11 vielleicht etwas Inspiration.

**Übung 2.9.15 (Eigenschaften von abgeschlossenen Mengen)**

**Übung 2.9.16 (Beschränkte Folgen mit einem Häufungspunkt)**

**Übung 2.9.17 (Geometrische Reihe auf dem Einheitskreis)** Sei  $z \in \mathbb{C}$  mit  $|z| = 1$ . Zeigen Sie, dass die geometrische Reihe  $\sum_{n=0}^{\infty} z^n$  divergiert.

**Übung 2.9.18 (Rechenregeln für konvergente Netze)** Übertragen Sie die Beweise von Proposition 2.3.5, Proposition 2.3.8, Proposition 2.3.10 und Proposition 2.3.11, um die entsprechenden Aussagen für konvergente Netze aus Proposition 2.8.7 zu erhalten. Die Schwierigkeit besteht im wesentlichen darin, sich die jeweiligen Konvergenzbegriffe gut klar zu machen, dann geschieht die Übertragung ohne weitere Ideen und Argumente.

**Übung 2.9.19 (Monotone Netze)** Sei  $(a_i)_{i \in I}$  ein Netz von reellen Zahlen. Dann heißt  $(a_i)_{i \in I}$  beschränkt, wenn es eine reelle Zahl  $\alpha$  mit  $a_i \leq \alpha$  für alle  $i \in I$  gibt. Weiter heißt  $(a_i)_{i \in I}$  monoton wachsend, falls für alle  $i, j \in I$  mit  $i \preccurlyeq j$

$$a_i \leq a_j \quad (2.9.7)$$

gilt. Zeigen Sie, dass ein monoton wachsendes Netz genau dann konvergiert, wenn es beschränkt ist.

Hinweis: Verfahren Sie wie im Beweis von Proposition 2.4.9.

# Kapitel 3

## Stetige Funktionen

Nach Folgen wollen wir nun Abbildungen mit Werten in  $\mathbb{K}$  betrachten, deren Definitionsbereich nicht länger  $\mathbb{N}$  sondern eine größere Teilmenge von  $\mathbb{K}$  ist. Wir sind dann insbesondere daran interessiert, das Verhalten der Funktionswerte  $f(x)$  für Punkte  $x$  in der Nähe eines zuvor festgelegten Punktes  $x_0$  zu verstehen. Um sinnvoll von „in der Nähe“ sprechen zu können, werden wir typischerweise offene Definitionsbereiche oder zumindest Definitionsbereiche betrachten, die große offene Teilmengen enthalten. Die Frage nach guten und interessanten Definitionsbereichen wird uns auf weitere Sorten von Teilmengen von  $\mathbb{K}$  neben den bereits bekannten offenen und abgeschlossenen Teilmengen führen. Hier lernen wir das erste Mal kompakte Teilmengen kennen, die sich in vielerlei Hinsicht fast wie endlich Teilmengen verhalten, aber diese trotzdem drastisch verallgemeinern. Die Funktionen, welche wir in diesem Kapitel betrachten, sollen stetig sein. Anschaulich gesprochen sollen sich die Funktionswerte  $f(x)$  nur wenig ändern, wenn sich die Argumente  $x$  wenig ändern. Hierfür müssen wir eine tragfähige mathematische Definition finden, wobei wir verschiedene äquivalente Herangehensweisen kennenlernen werden. Es folgen dann wichtige Rechenregeln für stetige Funktionen und einige ihrer fundamentalen Eigenschaften. Schließlich werden wir nun zum ersten Mal Folgen von Funktionen, hier eben stetigen Funktionen, betrachten und deren Konvergenzeigenschaften studieren. Es zeigt sich, dass die naive punktweise Konvergenz die Stetigkeit typischerweise zerstört, während die gleichmäßige Konvergenz sie erhält. Als erste Anwendung betrachten wir Potenzreihen und zeigen deren Stetigkeit.

### 3.1 Funktionenlimites

Sei  $X \subseteq \mathbb{K}$  eine nicht-leere Teilmenge von reellen oder komplexen Zahlen. Wir wollen dann Funktionen

$$f: X \longrightarrow \mathbb{K} \tag{3.1.1}$$

betrachten. Um das Verhalten der Funktionswerte  $f(x)$  für  $x$  nahe bei  $x_0$  zu studieren, können wir auf folgende Konzepte zurückgreifen:

**Definition 3.1.1 (Häufungspunkt einer Teilmenge)** Sei  $X \subseteq \mathbb{K}$  eine nicht-leere Teilmenge. Ein Punkt  $x_0 \in \mathbb{K}$  heißt Häufungspunkt der Teilmenge  $X$ , wenn es in jeder Umgebung von  $x_0$  unendlich viele Punkte aus  $X$  gibt.

**Definition 3.1.2 (Berührungspunkt einer Teilmenge)** Sei  $X \subseteq \mathbb{K}$  eine nicht-leere Teilmenge. Ein Punkt  $x_0 \in \mathbb{K}$  heißt Berührungspunkt von  $X$ , falls es in jeder Umgebung von  $x_0$  einen Punkt aus  $X$  gibt.

**Proposition 3.1.3** Sei  $X \subseteq \mathbb{K}$  eine nicht-leere Teilmenge.

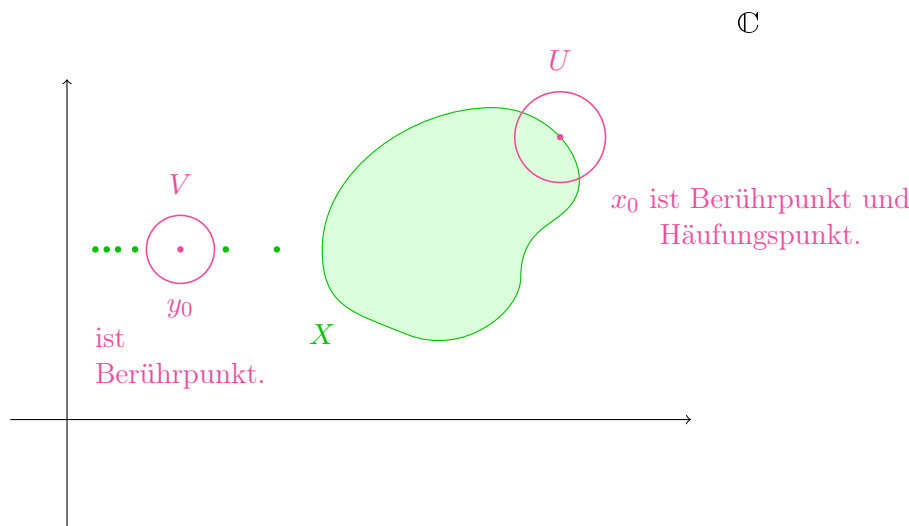


Abbildung 3.1: Häufungspunkte und Berührungspunkte: in jeder Umgebung  $U$  von  $x_0$  sind unendlich viele Punkte aus  $X$ . In jeder Umgebung von  $y_0$  sind Punkte aus  $X$ , nämlich  $y_0$ , aber im allgemeinen nur dieser eine.

- i.) Ein Punkt  $x_0 \in \mathbb{K}$  ist genau dann Häufungspunkt von  $X$ , wenn zu jedem  $\epsilon > 0$  die Teilmenge  $B_\epsilon(x_0) \cap X$  unendlich viele Punkte enthält.
- ii.) Ein Punkt  $x_0 \in \mathbb{K}$  ist genau dann ein Berührungspunkt von  $X$ , wenn für jedes  $\epsilon > 0$

$$B_\epsilon(x_0) \cap X \neq \emptyset. \quad (3.1.2)$$

Insbesondere ist jeder Punkt  $x_0 \in X$  ein Berührungspunkt von  $X$ .

- iii.) Jeder Häufungspunkt von  $X$  ist ein Berührungspunkt.
- iv.) Die Menge der Berührungspunkte von  $X$  ist die kleinste abgeschlossene Teilmenge von  $\mathbb{K}$ , die  $X$  enthält.
- v.) Ist  $X$  offen, so ist jeder Punkt  $x_0 \in X$  ein Häufungspunkt von  $X$ .

BEWEIS: Jede Umgebung  $U \subseteq \mathbb{K}$  von  $x_0$  enthält eine kleine offene Kugel  $B_\epsilon(x_0) \subseteq U$ . Enthält also  $B_\epsilon(x_0)$  unendlich viele Punkte aus  $X$ , so auch  $U$ . Da umgekehrt  $B_\epsilon(x_0)$  selbst eine Umgebung von  $x_0$  ist, folgt der erste Teil. Den zweiten Teil zeigt man analog. Da  $x_0 \in B_\epsilon(x_0)$  für alle offenen Kugeln  $B_\epsilon(x_0)$  um  $x_0$  gilt, ist  $x_0 \in X$  ein Berührungspunkt von  $X$ . Der dritte Teil ist klar. Der vierte Teil ist interessanter: zunächst ist klar, dass es eine kleinste abgeschlossene Teilmenge von  $\mathbb{K}$  gibt, die  $X$  enthält. Zunächst ist nämlich  $X \subseteq \mathbb{K}$  und  $\mathbb{K}$  ist abgeschlossen. Es gibt also abgeschlossene Teilmengen, die  $X$  enthalten. Dann setzen wir

$$X^{\text{cl}} = \bigcap_{\substack{A \subseteq \mathbb{K} \text{ abgeschlossen} \\ X \subseteq A}} A.$$

Da der beliebige Schnitt von abgeschlossenen Teilmengen nach Übung 2.9.15 wieder abgeschlossen ist, ist  $X^{\text{cl}}$  abgeschlossen. Es gilt nach Konstruktion  $X \subseteq X^{\text{cl}}$  und  $X^{\text{cl}}$  ist die kleinste abgeschlossene solche Teilmenge. Die Behauptung ist also, dass  $X^{\text{cl}}$  gerade die Berührungspunkte von  $X$  sind. Ist  $p \in \mathbb{K} \setminus X^{\text{cl}}$ , so gibt es ein  $\epsilon > 0$  mit  $p \in B_\epsilon(p) \subseteq \mathbb{K} \setminus X^{\text{cl}}$ , da  $X^{\text{cl}}$  abgeschlossen und entsprechend  $\mathbb{K} \setminus X^{\text{cl}}$  offen ist. Das bedeutet aber  $X^{\text{cl}} \cap B_\epsilon(p) = \emptyset$  und mit  $X \subseteq X^{\text{cl}}$  erst recht  $X \cap B_\epsilon(p) = \emptyset$ . Also ist  $p$  kein Berührungspunkt. Ist umgekehrt  $p$  kein Berührungspunkt, so gibt es ein  $\epsilon > 0$  mit  $B_\epsilon(p) \cap X = \emptyset$ . Da  $B_\epsilon(p)$  offen

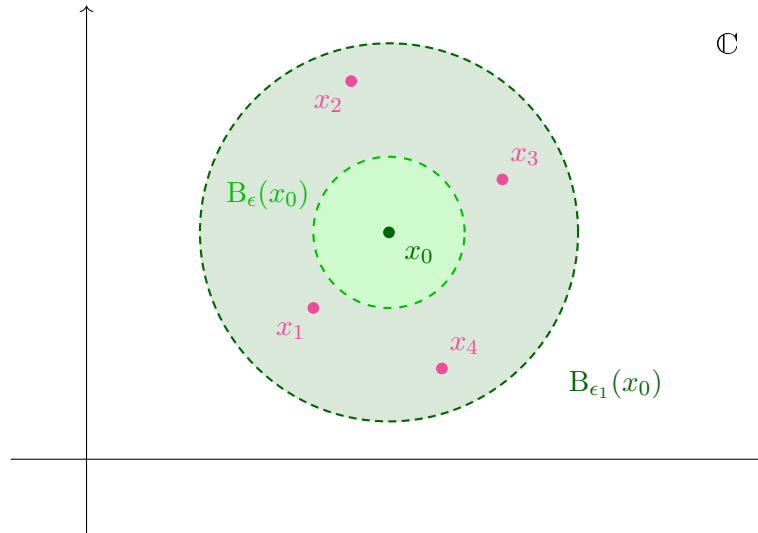


Abbildung 3.2: Die Kugel  $B_{\epsilon_1}(x_0)$  enthält endlich viele Punkte aus  $X$ .

ist, ist  $B_\epsilon(p)$  eine Umgebung jedes Punktes  $q \in B_\epsilon(p)$ , so dass auch  $q$  kein Berührungspunkt ist. Dies zeigt dann, dass die Menge der Punkte, die keine Berührungspunkte von  $X$  sind, offen ist. Ihr Komplement, die Menge der Berührungspunkte, ist also abgeschlossen. Da wir bereits gezeigt haben, dass die Berührungspunkte in  $X^{\text{cl}}$  liegen, gilt die Gleichheit, da  $X^{\text{cl}}$  die kleinste abgeschlossene Teilmenge ist, die  $X$  enthält. Der letzte Teil ist wieder einfach, da zu  $x_0 \in X$  ein  $\epsilon_0 > 0$  mit  $B_{\epsilon_0}(x_0) \subseteq X$  existiert. Diese offene Kugel enthält dann aber unendlich viele Punkte, die zu  $X$  gehören. Ist nun  $\epsilon > 0$  beliebig, so ist entweder  $\epsilon > \epsilon_0$ , also  $B_{\epsilon_0}(x_0) \subseteq B_\epsilon(x_0)$ , was unendlich viele Punkte in  $B_\epsilon(x_0)$  liefert, oder  $\epsilon_0 \leq \epsilon$  und damit  $B_\epsilon(x_0) \subseteq X$ , was auch wieder unendlich viele Punkte aus  $X$  in  $B_\epsilon(x_0)$  zeigt.  $\square$

**Definition 3.1.4 (Abschluss)** Ist  $X \subseteq \mathbb{K}$  eine Teilmenge, so heißt die kleinste abgeschlossene Teilmenge von  $\mathbb{K}$ , die  $X$  enthält, der Abschluss von  $X$ , bezeichnet mit  $X^{\text{cl}}$ .

Die nächste Proposition zeigt dann, dass die Intuition in Abbildung 3.1 zu Berührungspunkten im Wesentlichen richtig ist.

**Proposition 3.1.5** Sei  $X \subseteq \mathbb{K}$  eine nicht-leere Teilmenge und  $x_0 \in \mathbb{K}$  ein Berührungspunkt von  $X$ , der kein Häufungspunkt von  $X$  ist. Dann gilt  $x_0 \in X$  und es gibt ein  $\epsilon > 0$  mit

$$B_\epsilon(x_0) \cap X = \{x_0\}. \quad (3.1.3)$$

BEWEIS: Wenn  $x_0$  kein Häufungspunkt ist, gibt es also ein  $\epsilon_1 > 0$  derart, dass  $B_{\epsilon_1}(x_0) \cap X$  nur endlich viele Punkte enthält, die wir mit  $x_1, \dots, x_n \in X$  bezeichnen, siehe Abbildung 3.2. Ist nun  $x_0 \notin \{x_1, \dots, x_n\}$ , so gibt es ein  $\epsilon > 0$  mit  $\epsilon < \min\{|x_1 - x_0|, \dots, |x_n - x_0|\}$ . Die Kugel  $B_\epsilon(x_0)$  enthält dann keinen der Punkte  $x_1, \dots, x_n$  womit  $B_\epsilon(x_0) \cap X = \emptyset$ . Da  $x_0$  aber ein Berührungspunkt von  $X$  sein soll, kann dies nicht der Fall sein. Also ist  $x_0 \in \{x_1, \dots, x_n\}$ , etwa  $x_0 = x_1$ . Dann wählen wir  $\epsilon > 0$  mit  $\epsilon < \min\{|x_2 - x_0|, \dots, |x_n - x_0|\}$  und erhalten  $B_\epsilon(x_0) \cap X = \{x_0\}$  wie gewünscht.  $\square$

Einen solchen Punkt von  $X$  nennen wir auch *isolierten Punkt*. Die Menge der Berührungspunkte von  $X$ , also der Abschluss  $X^{\text{cl}}$ , besteht daher aus der disjunkten Vereinigung von isolierten Punkten und Häufungspunkten

$$X^{\text{cl}} = \{\text{Häufungspunkte von } X\} \cup \{\text{isolierte Punkte von } X\}, \quad (3.1.4)$$

wie es in Abbildung 3.1 skizziert ist. Wollen wir nun Folgen  $(x_n)_{n \in \mathbb{N}}$  in  $X$  betrachten, die gegen einen Punkt  $x_0 \in X^{\text{cl}}$  konvergieren, so sind die isolierten Punkte recht langweilig: es gilt  $x_n \rightarrow x_0$  genau dann, wenn

$$x_n = x_0 \quad (3.1.5)$$

für fast alle  $n$  gilt. Die Folge ist bis auf endlich viele Folgenglieder die konstante Folge. Interessanter ist es daher, Häufungspunkte  $x_0$  zu betrachten. Für diese gibt es viele nicht-triviale Folgen in  $X$ , die gegen  $x_0$  konvergieren.

**Beispiel 3.1.6** Wir betrachten die Teilmenge

$$X = \left\{ \frac{1}{n} \mid n \in \mathbb{N} \right\} \subseteq \mathbb{R}. \quad (3.1.6)$$

Dann gilt

$$X^{\text{cl}} = \left\{ \frac{1}{n} \mid n \in \mathbb{N} \right\} \cup \{0\}, \quad (3.1.7)$$

und 0 ist der einzige Häufungspunkt von  $X$ . Zur Vorsicht sei hier angemerkt, dass wir den Begriff Häufungspunkt auch für Folgen  $(a_n)_{n \in \mathbb{N}}$  definiert hatten. Hier ist zu beachten, dass ein Häufungspunkt der Wertemenge

$$A = \{a_n \mid n \in \mathbb{N}\} \subseteq \mathbb{K} \quad (3.1.8)$$

einer Folge auch ein Häufungspunkt der Folge im Sinne von Definition 2.5.6 ist. Die Umkehrung ist im allgemeinen falsch, wie die konstante Folge mit  $a_n = \alpha \in \mathbb{K}$  zeigt. Der Grenzwert  $\alpha$  ist natürlich ein Häufungspunkt der Folge, die Menge  $\{\alpha\} \subseteq \mathbb{K}$  hat dagegen keinen Häufungspunkt im Sinne von Definition 3.1.1. Sie besteht vielmehr aus einem isolierten Punkt, nämlich  $\alpha$ , siehe auch Übung 3.9.1.

Wir suchen nun eine tragfähige Definition für die Konvergenz von Funktionswerten  $f(x)$ , falls die Argumente  $x$  der Funktion gegen einen Punkt  $x_0$  konvergieren. Hier bietet sich der Standpunkt der Netze an:

**Beispiel 3.1.7 (Netze von Punkten)** Sei  $X \subseteq \mathbb{K}$  eine nicht-leere Teilmenge.

i.) Ist  $x_0 \in X$ , so definiert man eine Richtung  $\succsim$  für Punkte aus  $X$  durch  $x \succsim y$ , falls

$$|x - x_0| \leq |y - x_0|. \quad (3.1.9)$$

Man zeigt schnell, dass dies wirklich eine gerichtete Menge liefert. Die Richtung bedeutet geometrisch *näher bei*  $x_0$ . Da  $x_0 \in X$ , gilt für diese Richtung

$$x \preccurlyeq x_0 \quad (3.1.10)$$

für alle  $x \in X$ , da  $x_0$  ja näher an  $x_0$  als alle anderen Punkte aus  $X$  ist. Die gerichtete Menge  $(X, \preccurlyeq)$  besitzt also ein *Maximum*.

ii.) Ist  $x_0 \in \mathbb{K}$  ein Häufungspunkt, so definiert man auf  $X \setminus \{x_0\}$  eine Richtung  $\succsim$  wie in (3.1.9). Wieder bedeutet diese Richtung, näher bei  $x_0$  zu sein. Jetzt ist aber  $x_0$  nicht länger Element der Indexmenge, womit (3.1.10) bedeutungslos geworden ist. Man beachte, dass es zu jedem  $\epsilon > 0$  Punkte  $x \in X \setminus \{x_0\}$  gibt, die  $|x - x_0| < \epsilon$  erfüllen. Insbesondere gibt es zu  $x, y \in X \setminus \{x_0\}$  einen Punkt  $z \in X \setminus \{x_0\}$ , welcher strikt näher als beide Punkte  $x, y$  bei  $x_0$  liegt. Hier benötigen wir, dass  $x_0$  ein Häufungspunkt ist, siehe auch Übung 3.9.2.

iii.) Ist  $X \subseteq \mathbb{R}$  nach oben unbeschränkt, so erhalten wir eine Richtung  $\succsim$  durch  $x \succsim y$ , falls  $x \geq y$ . Diese Richtung bedeutet also *näher bei*  $+\infty$ . Mit ihr werden wir Divergenz gegen  $+\infty$  kodieren können. Analog erhalten wir für eine nach unten beschränkte Teilmenge  $X \subseteq \mathbb{R}$  eine Richtung  $\succsim$  durch  $x \succsim y$ , falls  $x \leq y$ , die entsprechend *näher bei*  $-\infty$  bedeutet.

iv.) Ist  $X \subseteq \mathbb{R}$  und  $x_0 \in \mathbb{K}$  ein Häufungspunkt von  $X$ , so erhalten wir zwei gerichtete Mengen, die *näher bei  $x_0$  von links* und *näher bei  $x_0$  von rechts* kodieren. Man betrachte

$$X_{<x_0} = \{x \in X \mid x < x_0\} \quad (3.1.11)$$

beziehungsweise

$$X_{>x_0} = \{x \in X \mid x > x_0\} \quad (3.1.12)$$

und versieht diese Teilmengen mit der Richtung (3.1.9), falls sie nicht leer sind. Es kann natürlich passieren, dass  $X_{<x_0}$  oder  $X_{>x_0}$  leer sind. Da aber  $x_0$  ein Häufungspunkt von  $X$  ist, ist mindestens eine der beiden Teilmengen  $X_{<x_0}$  oder  $X_{>x_0}$  sogar unendlich. In diesem Fall erhält man tatsächlich eine gerichtete Menge, siehe erneut Übung 3.9.2.

**Definition 3.1.8 (Konvergenz von Funktionswerten)** Sei  $X \subseteq \mathbb{K}$  eine nicht-leere Teilmenge, und seien  $f \in \text{Abb}(X, \mathbb{K})$  eine Funktion sowie  $x_0 \in X^{\text{cl}}$  ein Berührungspunkt von  $X$ .

i.) Die Funktionswerte von  $f(x)$  konvergieren gegen  $\alpha \in \mathbb{K}$  für  $x \rightarrow x_0$ , falls für jedes konvergente Netz  $(x_i)_{i \in I}$  in  $X$  mit  $\lim_{i \in I} x_i = x_0$  gilt, dass  $(f(x_i))_{i \in I}$  konvergiert mit

$$\lim_{i \in I} f(x_i) = \alpha. \quad (3.1.13)$$

Wir schreiben in diesem Fall

$$\lim_{x \rightarrow x_0} f(x) = \alpha. \quad (3.1.14)$$

ii.) Sind die Funktionswerte von  $f$  reell, so heißt  $f$  divergent gegen  $+\infty$  beziehungsweise gegen  $-\infty$  für  $x \rightarrow x_0$ , falls für jedes konvergente Netz  $(x_i)_{i \in I}$  in  $X$  mit  $\lim_{i \in I} x_i = x_0$  gilt, dass das Netz  $(f(x_i))_{i \in I}$  gegen  $+\infty$  beziehungsweise gegen  $-\infty$  divergiert. Wir schreiben in diesem Fall

$$\lim_{x \rightarrow x_0} f(x) = +\infty \quad \text{beziehungsweise} \quad \lim_{x \rightarrow x_0} f(x) = -\infty. \quad (3.1.15)$$

iii.) Ist  $X \subseteq \mathbb{R}$  nach oben beziehungsweise nach unten unbeschränkt, so heißt  $f$  konvergent gegen  $\alpha \in \mathbb{K}$  für  $x \rightarrow +\infty$  beziehungsweise für  $x \rightarrow -\infty$ , wenn für jedes Netz  $(x_i)_{i \in I}$  in  $X$  mit  $x_i \rightarrow +\infty$  beziehungsweise  $x_i \rightarrow -\infty$  gilt, dass das Netz  $(f(x_i))_{i \in I}$  konvergiert mit

$$\lim_{i \in I} f(x_i) = \alpha. \quad (3.1.16)$$

Wir schreiben in diesem Fall

$$\lim_{x \rightarrow +\infty} f(x) = \alpha \quad \text{beziehungsweise} \quad \lim_{x \rightarrow -\infty} f(x) = \alpha. \quad (3.1.17)$$

iv.) Ist  $X \subseteq \mathbb{R}$ , so heißt  $f$  konvergent gegen  $\alpha$  für  $x \rightarrow x_0$  von links (von rechts), wenn für jedes Netz  $(x_i)_{i \in I}$  in  $X$  mit  $x_i \leq x_0$  (oder  $x_i \geq x_0$ ) und  $x_i \rightarrow x_0$  das Netz  $(f(x_i))_{i \in I}$  gegen  $\alpha$  konvergiert. In diesem Fall schreiben wir

$$\lim_{\substack{x \rightarrow x_0 \\ x < x_0}} f(x) = \lim_{x \nearrow x_0} f(x) = \alpha \quad \text{beziehungsweise} \quad \lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f(x) = \lim_{x \searrow x_0} f(x) = \alpha. \quad (3.1.18)$$

Die Voraussetzung  $x_0 \in X^{\text{cl}}$  ist deshalb sinnvoll, damit es überhaupt Netze in  $X$  gibt, die gegen  $x_0$  konvergieren. Diese Definition sollte die wichtigsten Fälle abdecken, es gibt jedoch etliche weitere Situationen, in denen man Konvergenz von Funktionswerten betrachten möchte. Wichtig an der Definition ist, dass man die Konvergenzeigenschaften nicht mehr nur für *ein* Netz  $(x_i)_{i \in I}$  prüft, sondern für *alle* Netze, die den obigen Anforderungen genügen. Dies sieht zunächst wie ein sehr hoffnungsloses Unterfangen aus. Es zeigt sich aber, dass wir die Konvergenz geeignet umformulieren können, was dann handhabbare Kriterien liefert:

**Satz 3.1.9 (Konvergenz von Funktionswerten)** Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen. Seien weiter  $f \in \text{Abb}(X, \mathbb{K})$  eine Funktion und  $x_0 \in X^{\text{cl}}$  ein Berührungspunkt.

i.) Ist  $x_0$  ein isolierter Punkt, so gilt immer

$$\lim_{x \rightarrow x_0} f(x) = f(x_0). \quad (3.1.19)$$

ii.) Sind  $x_0$  ein Häufungspunkt und  $\alpha \in \mathbb{K}$ , so sind folgende Aussagen äquivalent:

- a) Es gilt  $\lim_{x \rightarrow x_0} f(x) = \alpha$ .
- b) Es gilt  $\lim_{x \in X \setminus \{x_0\}} f(x) = \alpha$  für die gerichtete Menge  $X \setminus \{x_0\}$  aus Beispiel 3.1.7, ii.).
- c) Für jede konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  in  $X$  mit  $x_n \rightarrow x_0$  gilt

$$\lim_{n \rightarrow \infty} f(x_n) = \alpha. \quad (3.1.20)$$

d) Für jedes  $\epsilon > 0$  gibt es ein  $\delta > 0$  derart, dass für alle  $x \in X$  mit  $|x - x_0| < \delta$  gilt, dass

$$|f(x) - \alpha| < \epsilon. \quad (3.1.21)$$

iii.) Ist  $x_0$  ein isolierter Punkt, so sind a), c) und d) trivialerweise für  $\alpha = f(x_0)$  erfüllt.

BEWEIS: Ist  $x_0$  ein isolierter Punkt von  $X$  und  $(x_i)_{i \in I}$  ein Netz in  $X$  mit  $x_i \rightarrow x_0$ , so gibt es einen Index  $i_0 \in I$  mit  $x_i = x_0$  für  $i \succ i_0$ , da in einer geeigneten kleinen Umgebung von  $x_0$  keine weiteren Punkte aus  $X$  sind. Damit ist aber die Konvergenz von  $(f(x_i))_{i \in I}$  gegen  $f(x_0)$  klar. Wir betrachten daher den interessanten Fall, dass  $x_0$  ein Häufungspunkt ist. Wir zeigen  $a) \implies b) \implies c) \implies d) \implies a)$ . Die Implikation  $a) \implies b)$  ist trivial. Sei also  $x_n \in X$  für  $n \in \mathbb{N}$  mit  $x_n \rightarrow x_0$  gegeben. Wir geben  $\epsilon > 0$  vor, dann finden wir ein  $x \in X \setminus \{x_0\}$  in der gerichteten Menge  $X \setminus \{x_0\}$  mit der Eigenschaft  $|f(y) - \alpha| < \epsilon$  für alle  $y \succ x$ , also für alle  $y \in X \setminus \{x_0\}$  mit  $|y - x_0| \leq |x - x_0|$ . Wir setzen  $\delta = |x - x_0| > 0$ , dann gilt für alle  $y \in X \setminus \{x_0\}$  mit  $|y - x_0| < \delta$  die Ungleichung

$$|f(y) - \alpha| < \epsilon. \quad (*)$$

Die Konvergenz  $x_n \rightarrow x_0$  liefert nun ein  $N \in \mathbb{N}$  mit

$$|x_n - x_0| < \delta \quad (**)$$

für alle  $n \geq N$ . Dies zeigt

$$|f(x_n) - \alpha| < \epsilon$$

für alle  $n \geq N$ , da wir  $(*)$  für  $y = x_n$  nach  $(**)$  verwenden dürfen. Also folgt  $b) \implies c)$ . Es gelte nun  $c)$ . Wir nehmen an,  $d)$  sei falsch. Dann gibt es also ein  $\epsilon > 0$ , so dass es zu jedem  $\delta > 0$  ein  $x \in B_\delta(x_0) \cap X$  mit

$$|f(x) - \alpha| \geq \epsilon \quad (\odot)$$

gibt. Wählen wir  $\delta = \frac{1}{n}$ , so finden wir also  $x_n \in B_{\frac{1}{n}}(x_0) \cap X$  derart, dass  $|f(x_n) - \alpha| \geq \epsilon$  für alle  $n \in \mathbb{N}$ . Diese Folge  $(x_n)_{n \in \mathbb{N}}$  erfüllt nun die Anforderungen von  $c)$ , womit  $f(x_n) \rightarrow \alpha$ , ein Widerspruch zu  $(\odot)$ . Also war  $d)$  richtig und wir haben  $c) \implies d)$  gezeigt. Sei schließlich  $(x_i)_{i \in I}$  ein beliebiges Netz in  $X$  mit  $x_i \rightarrow x_0$ . Sei  $\epsilon > 0$  vorgegeben und  $\delta > 0$  so gewählt, dass (3.1.21) gilt. Die Konvergenz  $x_i \rightarrow x_0$  zeigt dann, dass es einen Index  $i_0 \in I$  gibt, so dass  $|x_i - x_0| < \delta$  für alle  $i \succ i_0$  gilt. Also können wir (3.1.21) für diese  $x_i$  anwenden und erhalten  $|f(x_i) - \alpha| < \epsilon$  für alle  $i \succ i_0$ . Dies zeigt  $a)$ , womit wir den Beweis für ii.) erbracht haben. Der dritte Teil ist trivial.  $\square$



Die Konvergenz von Funktionswerten können wir also entweder mit *einem besonderen* Netz wie in b) zeigen oder mit Hilfe von Folgen oder mit einer direkten Abschätzung, siehe Übung 3.9.7 für eine weitere, auf Umgebungen basierende Formulierung. Entsprechend nennt man c) das *Folgenkriterium* und d) das  $\epsilon\delta$ -*Kriterium*. In der Literatur werden alle diese (äquivalenten) Kriterien zur Definition herangezogen.

**Bemerkung 3.1.10 (Konvergenz von Funktionswerten)** Wir haben noch weitere Szenarien für Konvergenz von Funktionswerten in Definition 3.1.8 kennengelernt. Auch für diese können wir Netze durch Folgen ersetzen oder ein sehr speziell gewähltes Netz benutzen, nämlich die passenden Netze aus Beispiel 3.1.7. Zudem gibt es auch wieder ein  $\epsilon\delta$ -Kriterium, was geringfügig angepasst werden muss. Details zur Äquivalenz der jeweiligen Kriterien finden sich in Übung 3.9.6. Schließlich können wir alle Rechenregeln für konvergente Folgen auf Konvergenz von Funktionswerten übertragen, siehe auch Übung 3.9.8. Wir werden dies im Folgenden zunehmend stillschweigender tun.

Wir betrachten nun einige Beispiele, die von wiederkehrendem Interesse sind:

**Beispiel 3.1.11 (Polynome)** Sei  $f(x) = a_N x^N + a_{N-1} x^{N-1} + \dots + a_1 x + a_0$  ein Polynom  $N$ -ten Grades, aufgefasst als Funktion  $f: \mathbb{K} \rightarrow \mathbb{K}$ , wobei  $a_N, \dots, a_0 \in \mathbb{K}$ . Ist  $x_0 \in \mathbb{K}$  ein beliebiger Punkt, so gilt für jede Folge  $(x_n)_{n \in \mathbb{N}}$  mit  $x_n \rightarrow x_0$  auch

$$f(x_n) \rightarrow f(x_0), \quad (3.1.22)$$

da Folgenkonvergenz mit den algebraischen Operationen des Körpers  $\mathbb{K}$  verträglich ist, siehe Proposition 2.3.5. Es gilt also

$$\lim_{x \rightarrow x_0} f(x) = f(x_0) \quad (3.1.23)$$

nach Satz 3.1.9.

**Beispiel 3.1.12 (Rationale Funktionen)** Seien  $p(x) = p_N x^N + \dots + p_1 x + p_0$  und  $q(x) = q_M x^M + \dots + q_1 x + q_0$  Polynome, wobei  $q_M \neq 0$  gelte und  $N, M \in \mathbb{N}_0$ . Dann heißt die Funktion

$$f(x) = \frac{p(x)}{q(x)} \quad (3.1.24)$$

eine *rationale Funktion*, wobei der *natürliche (maximale) Definitionsbereich*  $X \subseteq \mathbb{K}$  alle die Punkte  $x$  sind, für die  $q(x) \neq 0$ . Da nach dem Fundamentalsatz der Algebra jedes nicht konstante Polynom vom Grad  $M$  genau  $M$  komplexe Nullstellen (mit Vielfachheit gezählt) besitzt, ist  $X$  ganz  $\mathbb{K}$  bis auf endlich viele Punkte. Für  $x_0 \in X$  ist also  $q(x_0) \neq 0$ . Ist dann  $(x_i)_{i \in I}$  ein Netz in  $X$ , so gilt zum einen  $q(x_i) \neq 0$ , weil  $x_i \in X$ . Zum anderen impliziert die Konvergenz  $x_i \rightarrow x_0$  dann  $q(x_i) \rightarrow q(x_0)$  sowie  $p(x_i) \rightarrow p(x_0)$  nach Beispiel 3.1.11. Die Rechenregeln für konvergente Netze aus Proposition 2.8.7 oder für Folgen wie in Proposition 2.3.5 zeigen dann  $\frac{p(x_i)}{q(x_i)} \rightarrow \frac{p(x_0)}{q(x_0)}$  und somit

$$\lim_{x \rightarrow x_0} \frac{p(x)}{q(x)} = \frac{p(x_0)}{q(x_0)} \quad (3.1.25)$$

für alle  $x_0 \in X$ . Ist dagegen  $x_0$  eine Nullstelle des Nenners, können verschiedene Effekte auftreten:

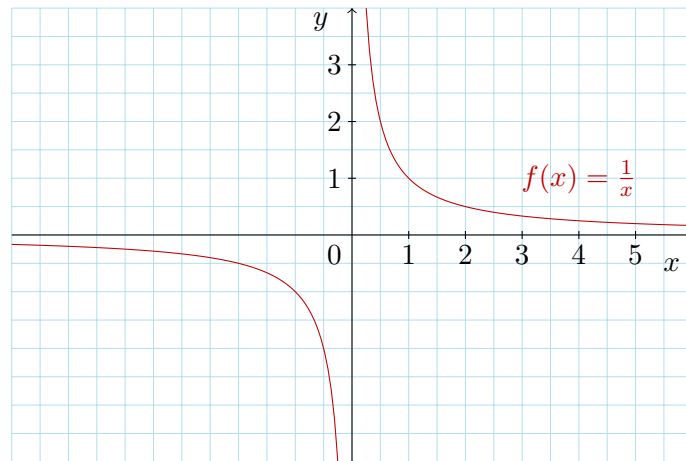
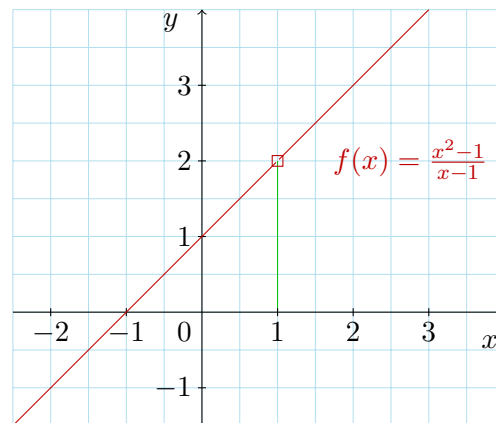
i.) Für die rationale Funktion

$$f(x) = \frac{1}{x} \quad (3.1.26)$$

auf dem reellen Definitionsbereich  $X = \mathbb{R} \setminus \{0\}$  gilt beispielsweise

$$\lim_{x \searrow 0} f(x) = +\infty \quad \text{und} \quad \lim_{x \nearrow 0} f(x) = -\infty, \quad (3.1.27)$$

was man sich gut anhand von Folgen überlegen kann, siehe auch Abbildung 3.3.

Abbildung 3.3: Der Graph der Funktion  $f(x) = \frac{1}{x}$ .Abbildung 3.4: Der Graph der Funktion  $f(x) = \frac{x^2 - 1}{x - 1}$  mit dem Verhalten für  $x \rightarrow 1$ .

ii.) Für

$$f(x) = \frac{x^2 - 1}{x - 1} \quad (3.1.28)$$

ist der natürliche komplexe Definitionsbereich zunächst einmal  $X = \mathbb{C} \setminus \{1\}$ . Um für eine Folge  $(x_n)_{n \in \mathbb{N}}$  in  $X$  mit  $x_n \rightarrow 1$  das Verhalten von  $f(x)$  zu bestimmen, schreiben wir  $f(x_n)$  zunächst als

$$f(x_n) = \frac{x_n^2 - 1}{x_n - 1} = \frac{(x_n + 1)(x_n - 1)}{x_n - 1} = x_n + 1.$$

Dann gilt offenbar  $\lim_{x_n \rightarrow 1} f(x_n) = 2$  für jede solche Folge, womit

$$\lim_{x \rightarrow 1} \frac{x^2 - 1}{x - 1} = 2 \quad (3.1.29)$$

gezeigt ist. In diesem Beispiel kann die Nullstelle des Nenners also zunächst gegen die passende Nullstelle des Zählers gekürzt werden, was dann zu einem endlichen Grenzwert in (3.1.29) führt, siehe Abbildung 3.4 für das Verhalten auf der reellen Achse.

iii.) Wir wollen schließlich das Verhalten von rationalen Funktionen für reelle  $x \rightarrow \pm\infty$  verstehen. Zu diesem Zweck schreiben wir  $f(x)$  aus (3.1.24) zunächst für  $x \neq 0$  folgendermaßen um. Für

$x \neq 0$  gilt

$$f(x) = x^{N-M} \frac{p_N + p_{N-1} \frac{1}{x} + \cdots + p_1 \frac{1}{x^{N-1}} + p_0 \frac{1}{x^N}}{q_M + q_{M-1} \frac{1}{x} + \cdots + q_1 \frac{1}{x^{M-1}} + q_0 \frac{1}{x^M}}. \quad (3.1.30)$$

Für die rationale Funktion

$$g(x) = \frac{1}{x} \quad (3.1.31)$$

ist der natürliche (reelle) Definitionsbereich  $\mathbb{R} \setminus \{0\}$ , siehe *i.*). Ist nun  $x_n \neq 0$  für  $n \in \mathbb{N}$  derart, dass  $x_n \rightarrow +\infty$ , so gilt nach Proposition 2.3.5, *iv.*), dass  $\frac{1}{x_n} \rightarrow 0$  konvergiert. Da dies für alle solchen Folgen der Fall ist, folgt

$$\lim_{x \rightarrow +\infty} \frac{1}{x} = 0 \quad \text{ebenso wie} \quad \lim_{x \rightarrow -\infty} \frac{1}{x} = 0. \quad (3.1.32)$$

Verwendet man erneut die Rechenregeln aus Proposition 2.3.5 für Folgen  $x_n \in X$  mit  $x_n \rightarrow +\infty$ , so ist zunächst  $x_n \neq 0$  für fast alle  $n \in \mathbb{N}$ , eventuell sogar für alle  $n \in \mathbb{N}$ , falls  $0 \notin X$ . Daher schließen wir

$$\lim_{x \rightarrow +\infty} \frac{p_N + p_{N-1} \frac{1}{x} + \cdots + p_1 \frac{1}{x^{N-1}} + p_0 \frac{1}{x^N}}{q_M + q_{M-1} \frac{1}{x} + \cdots + q_1 \frac{1}{x^{M-1}} + q_0 \frac{1}{x^M}} = \frac{p_N}{q_M} \quad (3.1.33)$$

und genauso für  $x \rightarrow -\infty$ . Das Verhalten der ursprünglichen Funktion  $f(x)$  hängt daher nur noch vom ersten Faktor  $x^{N-M}$  in (3.1.30) ab. Hier gibt es folgende Möglichkeiten:

a) Ist  $N < M$ , so gilt  $\lim_{x \rightarrow \pm\infty} x^{N-M} = 0$  und daher

$$\lim_{x \rightarrow \pm\infty} \frac{p(x)}{q(x)} = 0. \quad (3.1.34)$$

b) Ist  $N = M$ , so ist  $x^{N-M} = 1$  konstant. Daher folgt

$$\lim_{x \rightarrow \pm\infty} \frac{p(x)}{q(x)} = \frac{p_N}{q_N}. \quad (3.1.35)$$

c) Ist  $N > M$ , so gilt

$$\lim_{x \rightarrow +\infty} x^{N-M} = +\infty. \quad (3.1.36)$$

Ist  $\frac{p_N}{q_M} > 0$ , so folgt

$$\lim_{x \rightarrow +\infty} \frac{p(x)}{q(x)} = +\infty, \quad (3.1.37)$$

während für  $\frac{p_N}{q_M} < 0$

$$\lim_{x \rightarrow -\infty} \frac{p(x)}{q(x)} = -\infty, \quad (3.1.38)$$

gilt. Für  $x \rightarrow -\infty$  erhalten wir je nach dem, ob  $N - M$  gerade oder ungerade ist,

$$\lim_{x \rightarrow -\infty} x^{N-M} = \begin{cases} +\infty & N - M \text{ gerade,} \\ -\infty & N - M \text{ ungerade.} \end{cases} \quad (3.1.39)$$

Entsprechend erhalten wir für  $\lim_{x \rightarrow -\infty} \frac{p(x)}{q(x)}$  entweder  $+\infty$  oder  $-\infty$ , je nach Vorzeichen von  $\frac{p_N}{q_M}$ .

In den Übungen findet man noch viele andere Beispiele für Grenzwerte von Funktionswerten.

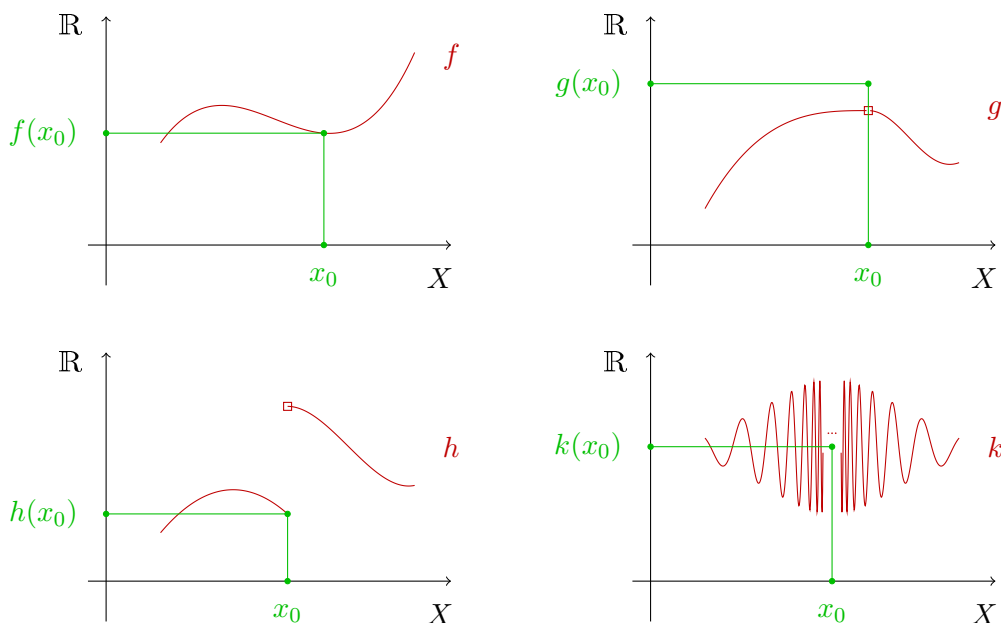


Abbildung 3.5: Die Funktion  $f$  entspricht der heuristischen Vorstellung von Stetigkeit bei  $x_0$ . Dieses Verhalten wird für  $g$  sabotiert, indem man den „falschen“ Funktionswert  $g(x_0)$  bei  $x_0$  vorschreibt. Bei der Funktion  $h$  ist schließlich Stetigkeit von der linken Seite gegeben, aber nicht von der rechten. Gänzlich unstetiges Verhalten legt das immer schnellere Oszillieren der Funktion  $k$  bei  $x_0$  an den Tag, egal welchen Funktionswert  $k$  bei  $x_0$  hat.

## 3.2 Stetigkeit

Wir kommen nun zu einem der wichtigsten Begriffe in der Analysis überhaupt, der Stetigkeit von Funktionen. Wie zuvor betrachten wir einen Definitionsbereich  $X$  von reellen oder komplexen Zahlen sowie eine reellwertige oder komplexwertige Funktion  $f \in \text{Abb}(X, \mathbb{K})$ . Hier sind alle vier Kombinationen von Interesse. Später werden wir den Stetigkeitsbegriff noch auf deutlich allgemeinere Situationen ausdehnen.

Ist nun  $X = (a, b) \subseteq \mathbb{R}$  ein reelles Intervall und  $f: X \rightarrow \mathbb{R}$  reellwertig, so ist die Wunschvorstellung, dass eine stetige Funktion  $f$  aus dem zusammenhängenden Intervall wieder etwas Zusammenhängendes im Bild liefert. Auch wenn dies die heuristische Idee bei Stetigkeit ist, sei hier gewarnt, dass wir momentan über keine tragfähige Definition des Begriffs zusammenhängend verfügen. Bildlich kann man diese Vorstellung am besten anhand des Graphen

$$\text{graph}(f) = \{(x, f(x)) \mid x \in X\} \subseteq X \times \mathbb{K} \quad (3.2.1)$$

der Funktion zeigen. Einige typische Situationen finden sich in Abbildung 3.5. Auch wenn diese Skizzen noch keinen mathematischen Sachverhalt liefern, geben sie doch eine ganz gute Vorstellung von den Möglichkeiten, wie Stetigkeit oder Unstetigkeit auftreten kann. Leider werden wir noch sehen, dass die Welt der stetigen wie auch unstetigen Funktionen noch deutlich mehr Phänomene bereithält, als durch Abbildung 3.5 erfasst werden. Es gilt also zunächst die tatsächliche Definition von Stetigkeit zu formulieren:

**Definition 3.2.1 (Stetigkeit)** Sei  $f \in \text{Abb}(X, \mathbb{K})$  eine Funktion und  $x_0 \in X$ .

i.) Die Funktion heißt stetig bei  $x_0$ , wenn

$$\lim_{x \rightarrow x_0} f(x) = f(x_0). \quad (3.2.2)$$

- ii.) Die Funktion heißt stetig auf  $X$ , falls  $f$  stetig bei allen Punkten  $x \in X$  ist.  
 iii.) Die Menge der stetigen Funktionen auf  $X$  wird mit

$$\mathcal{C}(X, \mathbb{K}) = \{f \in \text{Abb}(X, \mathbb{K}) \mid f \text{ ist stetig auf } X\} \quad (3.2.3)$$

bezeichnet.

In (3.2.2) sind implizit zwei Aussagen enthalten: zum einen soll der Grenzwert  $\lim_{x \rightarrow x_0} f(x)$  der Funktionswerte existieren, was ja bereits eine nicht triviale Forderung ist. Zum anderen ist die Gleichheit mit dem Funktionswert  $f(x_0)$  gefordert.

**Proposition 3.2.2** Sei  $f \in \text{Abb}(X, \mathbb{K})$  und  $x_0 \in X$  kein Häufungspunkt. Dann ist  $f$  stetig bei  $x_0$ .

BEWEIS: Da  $x_0 \in X$ , folgt nach Proposition 3.1.5, dass  $x_0$  ein isolierter Punkt von  $X$  ist. Nach Satz 3.1.9, i.), folgt dann die Behauptung.  $\square$

Stetigkeit bei  $x_0$  wird also nur dann eine interessante Bedingung, wenn es interessante Folgen oder Netze in  $X$  gibt, die gegen  $x_0$  konvergieren. Für isolierte Punkte ist dies eben nicht der Fall. Aus der Äquivalenz für die Konvergenz von Funktionswerten gemäß Satz 3.1.9 erhalten wir sofort folgende äquivalente Formulierungen der Stetigkeit:

**Proposition 3.2.3** Sei  $f \in \text{Abb}(X, \mathbb{K})$  und  $x_0 \in X$ . Dann sind folgende Aussagen äquivalent:

- i.) Die Funktion  $f$  ist stetig bei  $x_0$ .  
 ii.) Für jede konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  in  $X$  mit  $\lim_{n \rightarrow \infty} x_n = x_0$  ist  $(f(x_n))_{n \in \mathbb{N}}$  konvergent und

$$\lim_{n \rightarrow \infty} f(x_n) = f(x_0). \quad (3.2.4)$$

- iii.) Zu jedem  $\epsilon > 0$  gibt es ein  $\delta > 0$  derart, dass

$$|f(x) - f(x_0)| < \epsilon \quad (3.2.5)$$

für alle  $x \in B_\delta(x_0) \cap X$  gilt.

Entsprechend nennt man die Eigenschaft ii.) auch die *Folgenstetigkeit* von  $f$  bei  $x_0$ , während iii.) die  $\epsilon\delta$ -*Stetigkeit* von  $f$  bei  $x_0$  ist. Da wir unsere Definition von Stetigkeit auf dem netzbasierten Grenzwertbegriff von Funktionswerten aufgebaut haben, nennt man Definition 3.2.1, i.), auch *Netzstetigkeit* von  $f$  bei  $x_0$ . Die obige Proposition besagt, dass alle diese drei Konzepte äquivalente Begriffe von Stetigkeit liefern.

Wir wollen nun zwei weitere Charakterisierungen von Stetigkeit geben, die sowohl konzeptuell interessant als auch für praktische Anwendungen nützlich sind. Dazu benötigen wir eine Verfeinerung unserer Begriffe von Umgebung und offenen Mengen:

**Definition 3.2.4 (Umgebung und offene Mengen in  $X$ )** Sei  $X \subseteq \mathbb{K}$  eine Teilmenge.

- i.) Eine Teilmenge  $U \subseteq X$  heißt *Umgebung* von  $x \in X$  in  $X$  (oder: relativ zu  $X$ ), falls es eine Umgebung  $\tilde{U} \subseteq \mathbb{K}$  von  $x$  gibt, so dass

$$U = X \cap \tilde{U} \quad (3.2.6)$$

gilt.

- ii.) Eine Teilmenge  $O \subseteq X$  heißt *offen* in  $X$  (oder: relativ zu  $X$ ), falls es eine offene Teilmenge  $\tilde{O} \subseteq \mathbb{K}$  gibt, so dass

$$O = X \cap \tilde{O}. \quad (3.2.7)$$

iii.) Eine Teilmenge  $A \subseteq X$  heißt abgeschlossen in  $X$  (oder: relativ zu  $X$ ), falls  $X \setminus A$  offen in  $X$  ist.

Die Betonung bei dieser Definition liegt auf dem Zusatz *in  $X$* . Damit ist auch klar, dass eine Umgebung (offene oder abgeschlossene Menge) in  $\mathbb{K}$  eine Umgebung (offene oder abgeschlossene Menge) im vorherigen Sinne ist. Überraschenderweise verhalten sich Umgebungen, offene und abgeschlossene Teilmengen in  $X$  formal genauso wie die bisherigen Umgebungen, offenen und abgeschlossenen Teilmengen in  $\mathbb{K}$ . Wir sammeln die entscheidenden Eigenschaften erneut in einer Proposition:

**Proposition 3.2.5** Sei  $X \subseteq \mathbb{K}$  eine Teilmenge.

- i.) Ist  $U \subseteq X$  eine Umgebung von  $x \in X$  in  $X$ , so ist jede Teilmenge  $U' \subseteq X$  mit  $U \subseteq U'$  ebenfalls eine Umgebung von  $x \in X$  in  $X$ .
- ii.) Sind  $U_1, \dots, U_n \subseteq X$  Umgebungen von  $x \in X$  in  $X$ , so ist  $U_1 \cap \dots \cap U_n$  ebenfalls eine Umgebung von  $x$  in  $X$ .
- iii.) Der Schnitt  $B_r(x) \cap X$  für  $r > 0$  und  $x \in X$  ist offen in  $X$ .
- iv.) Die leere Menge  $\emptyset \subseteq X$  und  $X$  sind offen in  $X$ .
- v.) Sind  $\{O_i\}_{i \in I}$  offene Teilmengen  $O_i \subseteq X$  in  $X$ , so ist auch die Vereinigung  $\bigcup_{i \in I} O_i \subseteq X$  offen in  $X$ .
- vi.) Sind  $O_1, \dots, O_n \subseteq X$  offene Teilmengen in  $X$ , so ist auch  $O_1 \cap \dots \cap O_n \subseteq X$  offen in  $X$ .
- vii.) Eine Teilmenge  $A \subseteq X$  ist genau dann abgeschlossen in  $X$ , wenn es eine abgeschlossene Teilmenge  $\tilde{A} \subseteq \mathbb{K}$  mit  $A = X \cap \tilde{A}$  gibt.
- viii.) Ist  $U \subseteq X$  eine Umgebung von  $x \in X$  in  $X$ , so gibt es eine Umgebung  $V \subseteq U$  von  $x \in X$  in  $X$ , die offen in  $X$  ist.

BEWEIS: Anders als der Beweis von Proposition 2.5.2 ist dieser Beweis sehr einfach und beruht lediglich auf elementaren Rechenregeln zu Vereinigungen, Schnitten und Komplementen von Mengen.

- i.) Ist  $U \subseteq X$  Umgebung von  $x \in X$  in  $X$ , so gibt es eine Umgebung  $\tilde{U} \subseteq \mathbb{K}$  von  $x$  mit  $U = \tilde{U} \cap X$ . Dann ist  $\tilde{U} \cup U' \subseteq \mathbb{K}$  immer noch eine Umgebung nach Proposition 2.5.2, i.). Es gilt  $(\tilde{U} \cap U') \cap X = (\tilde{U} \cap X) \cup (U' \cap X) = U \cup U' = U'$ , da  $U' \subseteq X$  und  $\tilde{U} \cap X = U$ . Also ist  $U'$  eine Umgebung in  $X$ , siehe auch Abbildung 3.6.
- ii.) Zu den Umgebungen  $U_1, \dots, U_n \subseteq X$  von  $x \in X$  in  $X$  finden wir Umgebungen  $\tilde{U}_1, \dots, \tilde{U}_n \subseteq \mathbb{K}$  von  $x$  mit  $U_i = \tilde{U}_i \cap X$ . Deren Schnitt  $\tilde{U}_1 \cap \dots \cap \tilde{U}_n$  ist eine Umgebung nach Proposition 2.5.2, ii.). Es gilt

$$(\tilde{U}_1 \cap \dots \cap \tilde{U}_n) \cap X = (\tilde{U}_1 \cap X) \cap \dots \cap (\tilde{U}_n \cap X) = U_1 \cap \dots \cap U_n,$$

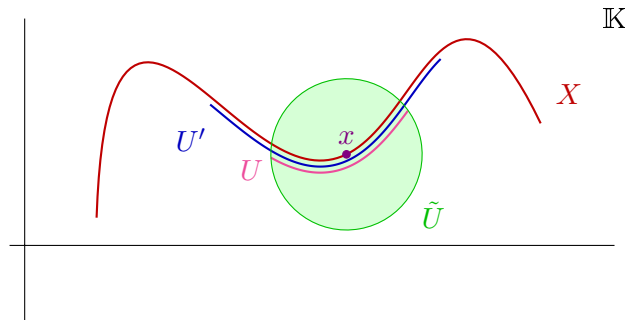
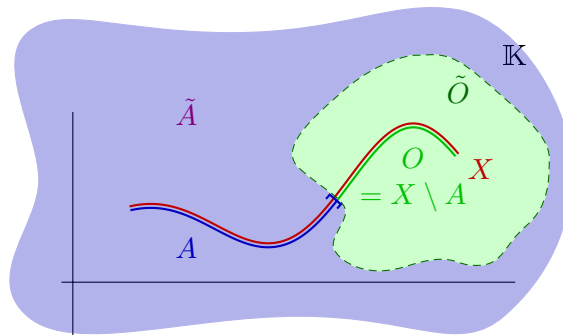
womit  $U_1 \cap \dots \cap U_n$  eine Umgebung von  $x$  in  $X$  ist.

- iii.) Da  $B_r(x)$  offen in  $\mathbb{K}$  ist, ist dies gerade die Definition offener Mengen in  $X$ .
- iv.) Es gilt  $\emptyset = X \cap \emptyset$  und  $X = X \cap \mathbb{K}$ . Da  $\emptyset$  und  $\mathbb{K}$  nach Proposition 2.5.2, iv.) offen sind, sind  $\emptyset$  und  $X$  offen in  $X$ .
- v.) Zu den offenen Teilmengen  $O_i \subseteq X$  in  $X$  finden wir offene Teilmengen  $\tilde{O}_i \subseteq \mathbb{K}$  mit  $O_i = \tilde{O}_i \cap X$  für alle  $i \in I$ . Nach Proposition 2.5.2, v.) ist  $\bigcup_{i \in I} \tilde{O}_i = \tilde{O}$  offen. Es gilt

$$\left( \bigcup_{i \in I} \tilde{O}_i \right) \cap X = \bigcup_{i \in I} (\tilde{O}_i \cap X) = \bigcup_{i \in I} O_i,$$

womit  $\bigcup_{i \in I} O_i$  offen in  $X$  ist.

- vi.) Dies zeigt man wie v.).

Abbildung 3.6: Obermengen von Umgebungen in  $X$  sind Umgebungen in  $X$ .Abbildung 3.7: Die Teilmenge  $A \subseteq X$  ist genau dann abgeschlossen in  $X$ , wenn  $A = X \cap \tilde{A}$  mit einer abgeschlossenen Teilmenge  $\tilde{A} \subseteq \mathbb{K}$ .

viii.) Eine Teilmenge  $A \subseteq X$  ist abgeschlossen in  $X$ , wenn  $O = X \setminus A$  offen in  $X$  ist, es also eine offene Teilmenge  $\tilde{O} \subseteq \mathbb{K}$  mit  $O = X \cap \tilde{O}$  gibt. Es gilt

$$A = X \setminus O = X \setminus (X \cap \tilde{O}) = X \cap (\mathbb{K} \setminus \tilde{O}),$$

siehe auch Abbildung 3.7. Dann ist  $A = X \cap \tilde{A}$  mit der abgeschlossenen Menge  $\tilde{A} = \mathbb{K} \setminus \tilde{O}$ .

viii.) Sei  $U \subseteq X$  eine Umgebung von  $x \in X$  in  $X$  mit passender Umgebung  $\tilde{U} \subseteq \mathbb{K}$  von  $x$ , so dass  $U = X \cap \tilde{U}$ . Nach Proposition 2.5.2, viii.) gibt es eine offene Umgebung  $\tilde{V} \subseteq \tilde{U}$  von  $x$ . Dann ist  $V = X \cap \tilde{V}$  offen in  $X$  und  $x \in V$ . Weiter ist  $V \subseteq U$ , womit auch der letzte Teil gezeigt ist.  $\square$

**Bemerkung 3.2.6** Die Teilmengen von  $X$  können offen in  $X$  sein, ohne dass sie offene Teilmengen von  $\mathbb{K}$  sind, ebenso für abgeschlossene Teilmengen oder Umgebungen. Beispielsweise ist  $X$  immer offen in  $X$ , aber nicht notwendigerweise in  $\mathbb{K}$ , siehe Abbildung 3.8. Als konkretes Beispiel denke man etwa an  $X = [a, b) \subseteq \mathbb{R}$  mit  $a < b$ , was weder abgeschlossen noch offen in  $\mathbb{R}$  ist, aber offen und abgeschlossen in  $X$ .

Nachdem wir nun die Begriffe der Umgebung sowie der offenen und abgeschlossenen Teilmengen relativ zu einer vorgegebenen Teilmenge  $X$  der reellen oder komplexen Zahlen zur Verfügung haben, können wir Stetigkeit auch auf folgende Weise charakterisieren:

**Satz 3.2.7 (Stetigkeit)** Seien  $X, Y$  Teilmengen der reellen oder komplexen Zahlen und  $f: X \rightarrow Y$  eine Abbildung.

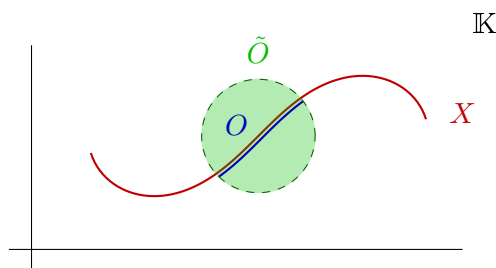


Abbildung 3.8: Die Teilmenge  $O \subseteq X$  ist offen in  $X$  aber nicht offen in  $\mathbb{K}$ .

- i.) Die Funktion  $f$  ist genau dann stetig bei  $x_0 \in X$ , wenn für jede Umgebung  $V$  von  $f(x_0) \in Y$  in  $Y$  das Urbild  $f^{-1}(V)$  eine Umgebung von  $x_0$  in  $X$  ist.
- ii.) Die Funktion  $f$  ist genau dann stetig, falls für jede offene Teilmenge  $O \subseteq Y$  in  $Y$  das Urbild  $f^{-1}(O) \subseteq X$  offen in  $X$  ist.
- iii.) Die Funktion  $f$  ist genau dann stetig, falls für jede abgeschlossene Teilmenge  $A \subseteq Y$  in  $Y$  das Urbild  $f^{-1}(A) \subseteq X$  abgeschlossen in  $X$  ist.

BEWEIS: Für den ersten Teil betrachten wir zunächst den Fall, dass  $f$  stetig bei  $x_0 \in X$  ist. Ist  $U \subseteq Y$  eine Umgebung von  $f(x_0)$ , so gibt es nach Definition 3.2.4, i.), und Definition 2.5.1, ii.), ein  $\epsilon > 0$  derart, dass  $B_\epsilon(f(x_0)) \cap Y \subseteq U$  gilt. Da  $f$  stetig bei  $x_0$  ist, gibt es ein  $\delta > 0$  derart, dass für alle  $x \in X$  mit  $|x - x_0| < \delta$  auch  $|f(x_0) - f(x)| < \epsilon$  gilt. Dies bedeutet dann, dass

$$X \cap B_\delta(x_0) \subseteq f^{-1}(B_\epsilon(f(x_0))).$$

Es gilt  $f^{-1}(B_\epsilon(f(x_0))) \subseteq f^{-1}(U)$ , da Urbilder Inklusionen generell erhalten. Damit gilt also

$$X \cap B_\delta(x_0) \subseteq f^{-1}(U),$$

womit  $f^{-1}(U)$  eine Umgebung von  $x$  in  $X$ , nämlich  $X \cap B_\delta(x_0)$ , enthält. Also ist  $f^{-1}(U)$  nach Proposition 3.2.5, i.), selbst eine Umgebung von  $x$  in  $X$ . Für die Umkehrung sei  $f^{-1}(U)$  eine Umgebung von  $x_0 \in X$  in  $X$  für jede Umgebung  $U \subseteq Y$  von  $f(x_0)$  in  $Y$ . Sei  $\epsilon > 0$  vorgegeben, dann ist  $U = B_\epsilon(f(x_0)) \cap Y$  eine (sogar offene) Umgebung von  $f(x_0)$  in  $Y$ . Also ist  $f^{-1}(B_\epsilon(f(x_0)) \cap Y)$  eine Umgebung von  $x \in X$  in  $X$ . Als solche enthält sie ein  $B_\delta(x_0) \cap X$  für ein geeignetes  $\delta > 0$ . Also gilt

$$B_\delta(x_0) \cap X \subseteq f^{-1}(B_\epsilon(f(x_0)) \cap Y).$$

Dies bedeutet aber, dass für  $x \in B_\delta(x_0) \cap X$  der Funktionswert  $f(x)$  in  $B_\epsilon(f(x_0)) \cap Y$  liegt, also  $|f(x) - f(x_0)| < \epsilon$  erfüllt. Also haben wir die  $\epsilon\delta$ -Stetigkeit von  $f$  bei  $x_0$  gezeigt. Für den zweiten Teil betrachten wir zunächst wieder eine stetige Abbildung  $f$  sowie eine offene Teilmenge  $O$  in  $Y$ . Also gibt es eine offene Teilmenge  $\tilde{O}$  mit  $O = \tilde{O} \cap Y$ . Damit sehen wir, dass  $O$  eine (offene) Umgebung in  $Y$  von jedem Punkt  $y_0 \in O$  ist. Sei also  $x_0 \in f^{-1}(O)$  ein Urbild von  $f(x_0) = y_0 \in O$ . Dann liefert die Stetigkeit von  $f$  bei  $x_0$  nach i.), dass  $f^{-1}(O)$  eine Umgebung von  $x_0$  in  $X$  ist. Da dies für alle Punkte aus  $f^{-1}(O)$  der Fall ist, ist  $f^{-1}(O) \subseteq X$  offen in  $X$ . Dies gilt ganz allgemein nach Proposition 3.2.5, viii.), und folgender Zusatzüberlegung. Zu  $x_0$  gibt es eine offene Umgebung  $V_{x_0} \subseteq f^{-1}(O)$  in  $X$ . Dann gilt

$$f^{-1}(O) \subseteq \bigcup_{x_0 \in f^{-1}(O)} V_{x_0} \subseteq f^{-1}(O),$$

da  $x_0 \in V_{x_0}$ , womit wir Gleichheit gezeigt haben. Also ist  $f^{-1}(O)$  eine Vereinigung von offenen Mengen in  $X$  und somit selbst offen nach Proposition 3.2.5, v.). Dies zeigt die eine Richtung. Für



die andere betrachten wir  $x_0 \in X$ . Ist nun  $U \subseteq Y$  eine Umgebung von  $f(x_0)$  in  $Y$ , so enthält diese eine offene Umgebung  $O \subseteq U$  von  $f(x_0)$  in  $Y$  nach Proposition 3.2.5, *viii.*). Nach Voraussetzung ist  $f^{-1}(O)$  offen in  $X$ , also eine sogar offene Umgebung von  $x_0$  in  $X$ . Nach *i.*) zeigt dies, dass  $f$  bei  $x_0$  stetig ist. Da  $x_0 \in X$  beliebig war, folgt die Stetigkeit überall. Der letzte Punkt ist einfach, da eine Teilmenge definitionsgemäß genau dann abgeschlossen ist, wenn ihr Komplement offen ist. Ist also  $A \subseteq Y$  abgeschlossen in  $Y$ , so ist  $O = Y \setminus A$  offen in  $Y$ . Nun gilt aber für Urbilder

$$f^{-1}(Y \setminus A) = X \setminus f^{-1}(A).$$

Damit ist  $f^{-1}(A)$  also genau dann abgeschlossen in  $X$ , wenn  $f^{-1}(Y \setminus A)$  offen in  $X$  ist. Die Äquivalenz in *iii.*) folgt damit aus der entsprechenden Äquivalenz in *ii.*)  $\square$

**Bemerkung 3.2.8** Mit diesem Satz haben wir die wirklich verallgemeinerungsfähige Variante von Stetigkeit gefunden. Dies werden wir an verschiedenen Stellen, insbesondere in Kapitel 7, noch sehen. Ein konzeptionelles Studium der Stetigkeit basierend auf offenen Mengen und Umgebungen findet jedoch erst in der (mengen-theoretischen) Topologie statt, siehe etwa [7]. Bemerkenswert für uns ist, dass wir den Stetigkeitsbegriff gänzlich auf Umgebungen oder offene Teilmengen *im Definitionsbereich* beziehungsweise *im Bildbereich* der Abbildung  $f$  zurückgeführt haben. Der umgebende Raum  $\mathbb{K}$  wird nur benötigt, um aus seinen offenen Teilmengen die offenen Teilmengen *in*  $X$  beziehungsweise *in*  $Y$  zu erlangen. Natürlich können wir eine Abbildung  $f: X \rightarrow Y \subseteq \mathbb{K}$  immer auch als eine Abbildung  $f: X \rightarrow \mathbb{K}$  auffassen. Die Stetigkeit ändert sich dadurch nicht. Die nun gewonnene Freiheit der Ko-Einschränkung auf eine beliebige Teilmenge  $Y \subseteq \mathbb{K}$ , welche das Bild  $f(X)$  natürlich umfassen muss, ist trotzdem manchmal eine sehr nützliche Option.

### 3.3 Rechenregeln und Beispiele für stetige Funktionen

Nachdem wir nun viele äquivalente Formulierungen für Stetigkeit gefunden haben, wollen wir diese jetzt nutzen, um konkrete Rechenregeln für stetige Funktionen zu etablieren und entsprechende Beispiele kennenzulernen:

**Beispiel 3.3.1 (Polynome)** Eine polynomiale Funktion  $f(x) = a_N x^N + a_{N-1} x^{N-1} + \cdots + a_1 x + a_0$  mit  $a_N, \dots, a_0 \in \mathbb{K}$  ist eine stetige Abbildung, also

$$f \in \mathcal{C}(\mathbb{K}, \mathbb{K}). \quad (3.3.1)$$

Dies haben wir implizit im Beispiel 3.1.11 gezeigt, wenn man das Folgenkriterium für Stetigkeit verwendet.

**Beispiel 3.3.2 (Rationale Funktionen)** Auch eine rationale Funktion

$$f(x) = \frac{p(x)}{q(x)} = \frac{p_N x^N + \cdots + p_1 x + p_0}{p_M x^M + \cdots + q_1 x + q_0} \quad (3.3.2)$$

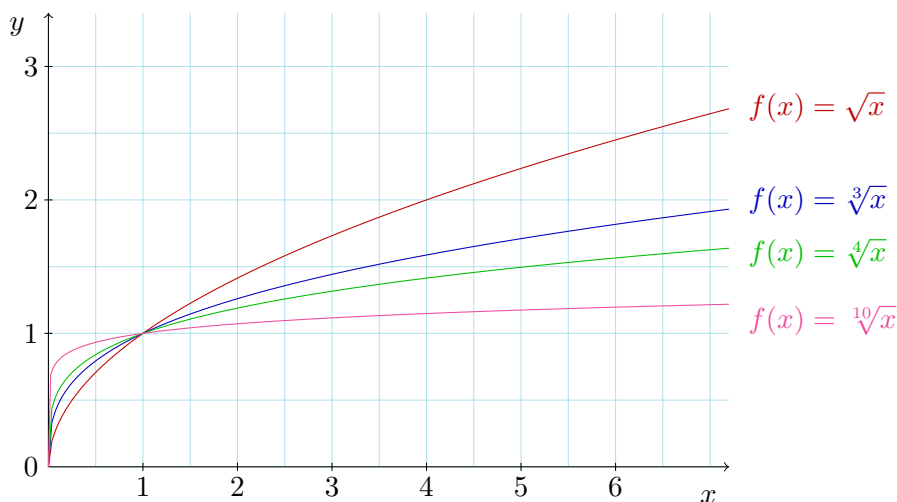
ist stetig auf ihrem natürlichen Definitionsbereich  $X$ , der eben gerade die Nennernullstellen ausschließt. Dies haben wir implizit in Beispiel 3.1.12 wieder unter Verwendung des Folgenkriteriums gezeigt.

**Beispiel 3.3.3 (Wurzelfunktionen)** Sei  $p > 0$ . Dann ist die Funktion

$$\mathbb{R}_0^+ \ni x \mapsto x^p \in \mathbb{R} \quad (3.3.3)$$

stetig. Dies folgt aus dem Folgenkriterium und Proposition 2.3.11. Ist dagegen  $p \in \mathbb{R}$  beliebig, so ist zumindest

$$\mathbb{R}^+ \ni x \mapsto x^p \in \mathbb{R} \quad (3.3.4)$$

Abbildung 3.9: Die Wurzelfunktionen sind stetig auf  $\mathbb{R}_0^+$ .

noch stetig mit Hilfe des gleichen Arguments. Man beachte, dass wir für die Stetigkeit von (3.3.4) für  $p = 0$  den Punkt  $x = 0$  explizit weglassen müssen, es sei denn, wir setzen  $0^0 = 1$ . Im Hinblick auf die Stetigkeit wäre dies also eine vernünftige *Konvention*. Allerdings zeigt Übung 3.9.14, dass es berechnete Interessen gibt, den Wert von  $0^0$  auch anders festzulegen. Auf jeden Fall sind also alle Wurzelfunktionen

$$\mathbb{R}_0^+ \ni x \mapsto \sqrt[n]{x} \in \mathbb{R} \quad (3.3.5)$$

für  $n \in \mathbb{N}$  stetig, siehe auch Abbildung 3.9.

**Beispiel 3.3.4 (Komplexe Konjugation und Betrag)** Die komplexe Konjugation

$$\bar{\cdot}: \mathbb{C} \ni z \mapsto \bar{z} \in \mathbb{C} \quad (3.3.6)$$

ist stetig. Ist nämlich  $(z_n)_{n \in \mathbb{N}}$  eine beliebige komplexe Folge mit  $z_n \rightarrow z$ , so gilt nach Proposition 2.3.5, iii.), auch  $\bar{z}_n \rightarrow \bar{z}$ . Nach dem Folgenkriterium für Stetigkeit ist (3.3.6) stetig. Damit ist auch

$$|\cdot|: \mathbb{C} \ni z \mapsto |z| \in \mathbb{R} \quad (3.3.7)$$

stetig. Ist nämlich wieder  $(z_n)_{n \in \mathbb{N}}$  eine beliebige komplexe Folge mit  $z_n \rightarrow z$ , so gilt zunächst  $\bar{z}_n z_n \rightarrow \bar{z} z$  nach den Rechenregeln aus Proposition 2.3.5. Mit der Folgenstetigkeit der Quadratwurzel aus Beispiel 3.3.3 folgt dann auch  $\sqrt{\bar{z}_n z_n} \rightarrow \sqrt{\bar{z} z}$ , was aber gerade

$$|z_n| \rightarrow |z| \quad (3.3.8)$$

zeigt. Dies ist die Folgenstetigkeit von (3.3.7).

**Beispiel 3.3.5 (Exponentialfunktionen)** Sei  $a > 0$ . Dann ist die Exponentialfunktion

$$\mathbb{R} \ni x \mapsto a^x \in \mathbb{R} \quad (3.3.9)$$

stetig, denn nach Proposition 2.3.10 ist (3.3.9) folgenstetig.

Damit haben wir also die Stetigkeit einiger besonders wichtiger Funktionen gezeigt. Es sind aber keineswegs alle Funktionen stetig:

**Beispiel 3.3.6 (Unstetige Funktionen)** Wir betrachten als Definitionsbereich  $\mathbb{R}$ .

i.) Die Funktion

$$f(x) = \begin{cases} x & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases} \quad (3.3.10)$$

ist stetig bei 0 aber unstetig bei allen anderen Punkten  $\mathbb{R} \setminus \{0\}$ . Für die Stetigkeit bei  $x_0 = 0$  geben wir  $\epsilon > 0$  vor. Mit  $\delta = \epsilon > 0$  gilt dann für  $|x| < \delta$  entweder  $f(x) = x$  also  $|f(x)| < \epsilon$  im Fall  $x \in \mathbb{Q}$  oder  $f(x) = 0$  also  $|f(x)| < \epsilon$  für  $x \in \mathbb{R} \setminus \mathbb{Q}$ . Dies ist die  $\epsilon\delta$ -Stetigkeit bei  $x_0 = 0$ . Für einen anderen Punkt  $x_0 \neq 0$  gibt es eine Folge  $(x_n)_{n \in \mathbb{N}}$  mit rationalen Zahlen  $x_n$ , die gegen  $x_0$  konvergiert. Es folgt  $f(x_n) = x_n \rightarrow x_0$  für eine solche Folge. Andererseits gibt es auch eine Folge  $(y_n)_{n \in \mathbb{N}}$  mit irrationalen Zahlen  $y_n \in \mathbb{R} \setminus \mathbb{Q}$ , die gegen  $x_0$  konvergiert, siehe Übung 2.9.12. Für diese gilt dann  $f(y_n) = 0 \rightarrow 0 \neq x_0$ . Damit kann  $f$  nicht folgenstetig bei  $x_0$  sein, da ja *alle* Folgen, die gegen  $x_0$  konvergieren, den Grenzwert  $f(x_0)$  für die Funktionswerte und damit insbesondere den *selben* Grenzwert haben müssen.

ii.) Eine gänzlich unstetige Funktion erhalten wir beispielsweise durch die charakteristische Funktion  $\chi_{\mathbb{Q}}$  von  $\mathbb{Q} \subseteq \mathbb{R}$ . Ist  $A \subseteq M$  eine Teilmenge einer Menge  $M$ , so heißt  $\chi_A: M \rightarrow \mathbb{R}$  mit

$$\chi_A(p) = \begin{cases} 1 & \text{falls } p \in A \\ 0 & \text{sonst} \end{cases} \quad (3.3.11)$$

die *charakteristische Funktion* von  $A$ . Für den Fall  $M = \mathbb{R}$  und  $A = \mathbb{Q}$  nennt man  $\chi_{\mathbb{Q}}$  auch die *Dirichlet-Funktion*. Diese ist überall unstetig, was man mit einem Folgenargument wie in i.) sehen kann.

iii.) Ein vielleicht noch extremeres Beispiel einer unstetigen Funktion ist folgende Variante von  $\chi_{\mathbb{Q}}$ . Wir definieren  $f: \mathbb{R} \rightarrow \mathbb{R}$  durch

$$f(x) = \begin{cases} q & \text{für } x = \frac{p}{q} \in \mathbb{Q} \setminus \{0\} \text{ mit } p, q \text{ teilerfremd und } q \in \mathbb{N} \\ 1 & \text{für } x = 0 \\ 0 & \text{für } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases} \quad (3.3.12)$$

Die Unstetigkeit zeigt man dann wieder mit dem Folgenkriterium wie für  $\chi_{\mathbb{Q}}$ . Neu ist nun folgende Eigenschaft: Ist  $x_0 \in \mathbb{R}$  ein beliebiger Punkt, so gibt es zu jedem  $\epsilon > 0$  und jedem  $N$  eine rationale Zahl  $\frac{p}{q}$  mit teilerfremden  $p$  und  $q$  derart, dass  $q \geq N$  und  $|x_0 - \frac{p}{q}| < \epsilon$ , siehe auch Übung 1.8.20. Damit ist der Funktionswert also  $f(\frac{p}{q}) \geq N$ , womit für jedes  $\epsilon > 0$

$$\sup\{f(x) \mid x \in B_\epsilon(x_0)\} = +\infty. \quad (3.3.13)$$

Die Funktion  $f$  ist also auf *jeder* Umgebung von  $x_0 \in \mathbb{R}$  unbeschränkt.

iv.) Anstelle von (3.3.12) definiert man

$$f(x) = \begin{cases} \frac{1}{q} & \text{für } x = \frac{p}{q} \in \mathbb{Q} \setminus \{0\} \text{ mit } p, q \text{ teilerfremd und } q \in \mathbb{N} \\ 1 & \text{für } x = 0 \\ 0 & \text{sonst.} \end{cases} \quad (3.3.14)$$

Anders als  $\chi_{\mathbb{Q}}$  ist  $f$  bei allen irrationalen Punkten stetig: Um eine irrationale Zahl  $x_0$  zu approximieren, müssen die Nenner und Zähler einer rationalen Zahl immer größer werden. Bei rationalen Zahlen ist die Funktion  $f$  dagegen unstetig.

Es sollte für diese vier Beispiele klar sein, dass eine Diskussion der Stetigkeitseigenschaften anhand des Graphen nicht möglich ist, da dieser in keinem vernünftigen Sinne „gezeichnet“ werden kann.

Nach diesen Beispielen und Gegenbeispielen wollen wir nun einige allgemeine Rechenregeln zu stetigen Funktionen etablieren. Insbesondere wollen wir die Eigenschaften der Teilmenge  $\mathcal{C}(X, \mathbb{K}) \subseteq \text{Abb}(X, \mathbb{K})$  aller stetigen Funktionen klären:

**Proposition 3.3.7** *Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen. Dann bildet  $\mathcal{C}(X, \mathbb{K})$  eine Unteralgebra von  $\text{Abb}(X, \mathbb{K})$  bezüglich der punktweise erklärten Algebrastruktur aus Definition 2.2. Ausgeschrieben bedeutet dies, dass zu  $f, g \in \mathcal{C}(X, \mathbb{K})$  und  $z, w \in \mathbb{K}$  auch  $zf + wg$  ebenso wie  $fg$  wieder stetig sind. Zudem sind die konstanten Funktionen stetig. Für  $\mathbb{K} = \mathbb{C}$  ist auch  $\overline{f} \in \mathcal{C}(X, \mathbb{C})$ .*

BEWEIS: Mit beispielsweise dem Folgenkriterium für Stetigkeit folgt dies direkt aus Proposition 2.3.5, *i.*), *ii.*), *iii.*), sowie der Tatsache, dass konstante Funktionen trivialerweise stetig sind.  $\square$

Ist zudem  $g \in \mathcal{C}(X, \mathbb{K})$  auf  $X$  ungleich 0, also  $g(x) \neq 0$  für alle  $x \in X$ , so ist für  $f \in \mathcal{C}(X, \mathbb{K})$  auch der Quotient

$$\frac{f}{g} \in \mathcal{C}(X, \mathbb{K}) \quad (3.3.15)$$

wieder stetig. Auch dies folgt aus Proposition 2.3.5, *iv.*), und der Charakterisierung von Stetigkeit mittels Folgenstetigkeit.

**Bemerkung 3.3.8 (Stetigkeit an einem Punkt)** Hat man für  $f, g: X \rightarrow \mathbb{K}$  nur Stetigkeit bei einem Punkt  $x_0 \in X$  vorliegen, so zeigt das Folgenkriterium für Stetigkeit in Verbindung mit Proposition 2.3.5, *i.*), *ii.*) und *iii.*), dass die Funktionen  $zf + wg, fg$  und  $\overline{f}$  bei  $x_0$  wieder stetig sind. Gilt  $g(x_0) \neq 0$  so ist  $\frac{f}{g}$  auch wieder bei  $x_0$  stetig, siehe Übung 3.9.15.

### 3.4 Kompaktheit und Zusammenhang

Bevor wir unser Studium stetiger Funktionen fortsetzen können, benötigen wir noch zwei fundamentale Eigenschaften von Intervallen in  $\mathbb{R}$ . Wie zu erwarten, wird die Vollständigkeit von  $\mathbb{R}$  auch für diese Eigenschaften von grundlegender Bedeutung sein. Wir beginnen mit offenen Überdeckungen:

**Definition 3.4.1 (Offene Überdeckung)** *Sei  $X \subseteq \mathbb{K}$  eine Teilmenge und  $\{O_i\}_{i \in I}$  eine Familie von offenen Teilmengen in  $\mathbb{K}$ . Dann heißt  $\{O_i\}_{i \in I}$  offene Überdeckung von  $X$ , falls*

$$X \subseteq \bigcup_{i \in I} O_i. \quad (3.4.1)$$

**Beispiel 3.4.2 (Offene Überdeckung)** Wir betrachten  $\mathbb{Q} \subseteq \mathbb{R}$  und die folgenden offenen Mengen:

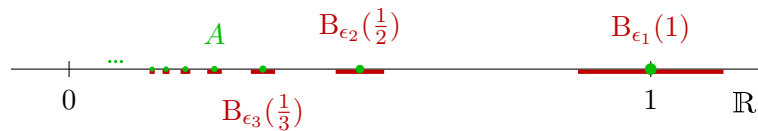
- i.)* Zuerst wählen wir eine Abzählung  $(q_n)_{n \in \mathbb{N}}$  von  $\mathbb{Q}$ . Dann betrachten wir  $\{B_{\epsilon_n}(q_n)\}_{n \in \mathbb{N}}$ , wobei  $(\epsilon_n)_{n \in \mathbb{N}}$  eine beliebige Folge von positiven Zahlen ist. Da  $q_n \in B_{\epsilon_n}(q_n)$  gilt, folgt also

$$\mathbb{Q} \subseteq \bigcup_{n=1}^{\infty} B_{\epsilon_n}(q_n), \quad (3.4.2)$$

so dass wir eine (abzählbare) offene Überdeckung von  $\mathbb{Q}$  gefunden haben.

- ii.)* Als kleine Abwandlung betrachten wir die Teilmenge

$$A = \left\{ \frac{1}{n} \mid n \in \mathbb{N} \right\}. \quad (3.4.3)$$

Abbildungung 3.10: Eine offene Überdeckung von  $\{\frac{1}{n} \mid n \in \mathbb{N}\}$ .

Die offenen  $\{B_{\epsilon_n}(\frac{1}{n})\}_{n \in \mathbb{N}}$  liefern dann wieder für jede Folge  $(\epsilon_n)_{n \in \mathbb{N}}$  eine offene Überdeckung. Durch geschickte Wahl der  $\epsilon_n$ , beispielsweise durch

$$\epsilon_n = \frac{1}{4} \left( \frac{1}{n} - \frac{1}{n+1} \right), \quad (3.4.4)$$

können wir  $B_{\epsilon_n}(\frac{1}{n}) \cap B_{\epsilon_{n'}}(\frac{1}{n'}) = \emptyset$  für  $n \neq n'$  erreichen, siehe auch Abbildung 3.10. Insbesondere gibt es keine Teilmenge  $I \subseteq \mathbb{N}$  derart, dass auch  $\{B_{\epsilon_n}(\frac{1}{n})\}_{n \in I}$  schon eine Überdeckung von  $A$  liefert.

Gerade im Hinblick auf die Nichtexistenz einer Teilüberdeckung in Beispiel 3.4.2, *ii.*), verhält sich das abgeschlossene Intervall gänzlich anders:

**Satz 3.4.3 (Kompaktheit von  $[a, b]$ )** Seien  $a < b$  reelle Zahlen. Dann besitzt jede offene Überdeckung  $\{O_i\}_{i \in I}$  des abgeschlossenen Intervalls  $[a, b]$  eine endliche Teilüberdeckung  $O_{i_1}, \dots, O_{i_n}$ , also

$$[a, b] \subseteq O_{i_1} \cup \dots \cup O_{i_n}. \quad (3.4.5)$$

BEWEIS: Wir nehmen an, es gibt eine offene Überdeckung  $\{O_i\}_{i \in I}$  von  $[a, b]$ , die keine endliche Teilüberdeckung  $O_{i_1}, \dots, O_{i_n}$  zulässt. Wir teilen das Intervall in der Mitte  $[a, b] = [a, \frac{b+a}{2}] \cup [\frac{b+a}{2}, b]$ . Die Überdeckung von  $[a, b]$  überdeckt dann sowohl  $[a, \frac{b+a}{2}]$  als auch  $[\frac{b+a}{2}, b]$ . Hätten nun beide Teilintervalle eine jeweils endliche Teilüberdeckung nämlich  $O_{i_1}, \dots, O_{i_n}$  für  $[a, \frac{b+a}{2}]$  und  $O_{j_1}, \dots, O_{j_m}$  für  $[\frac{b+a}{2}, b]$ , so wäre  $O_{i_1}, \dots, O_{i_n}, O_{j_1}, \dots, O_{j_m}$  eine endliche Teilüberdeckung des ursprünglichen Intervalls, die es aber nach der Annahme nicht geben soll. Daher hat mindestens eins der beiden Teilintervalle ebenfalls die Eigenschaft, dass  $\{O_i\}_{i \in I}$  keine endliche Teilüberdeckung besitzt. Wir wählen dieses aus und bezeichnen es mit  $[a_2, b_2]$ . Durch fortlaufende Intervallhalbierung erhalten wir induktiv Intervalle  $[a_2, b_2] \supseteq \dots \supseteq [a_n, b_n] \supseteq \dots$ , die keine endliche Teilüberdeckung besitzen. Da die Länge der Intervalle in jedem Schritt halbiert wird, haben wir eine Intervallschachtelung, die nach Satz 1.5.9, *iii.*), genau eine reelle Zahl  $x \in [a, b]$  im Schnitt hat, also

$$\{x\} = \bigcap_{n=1}^{\infty} [a_n, b_n].$$

Da  $x \in [a, b]$  gibt es mindestens ein  $O_{i_0}$  mit  $x \in O_{i_0}$ , weil wir eine Überdeckung haben. Da  $O_{i_0}$  offen ist, gibt es ein  $\epsilon > 0$  mit  $B_\epsilon(x) \subseteq O_{i_0}$ , siehe auch Abbildung 3.11. Da aber  $a_n \rightarrow x$  und  $b_n \rightarrow x$  mit  $x \in [a_n, b_n]$ , gibt es ein  $n_0$ , so dass  $[a_{n_0}, b_{n_0}] \subseteq B_\epsilon(x)$  gilt. Also gilt für dieses  $n_0$  sogar  $[a_{n_0}, b_{n_0}] \subseteq O_{i_0}$ , womit dieses Intervall eine endliche Teilüberdeckung, nämlich durch eine einzige Menge  $O_{i_0}$ , besitzt, ein Widerspruch.  $\square$

Wir nehmen diese Eigenschaft nun als Definition kompakter Teilmengen der reellen Achse und genauso für Teilmengen der komplexen Zahlen:

**Definition 3.4.4 (Kompaktheit)** Sei  $K \subseteq \mathbb{K}$  eine Teilmenge. Dann heißt  $K$  kompakt, falls jede offene Überdeckung von  $K$  eine endliche Teilüberdeckung besitzt.

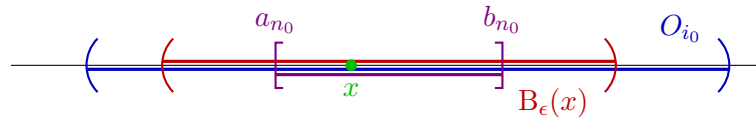


Abbildung 3.11: Zur Kompaktheit des abgeschlossenen Intervalls.

**Bemerkung 3.4.5 (Kompaktheit)**

- i.) Das abgeschlossene Intervall  $[a, b] \subseteq \mathbb{R}$  ist nach Satz 3.4.3 für alle  $a, b \in \mathbb{R}$  mit  $a < b$  kompakt.
- ii.) Endliche Teilmengen von  $\mathbb{R}$  oder  $\mathbb{C}$  sind aus trivialen Gründen kompakt. Damit sind zusammen mit dem ersten Teil also alle abgeschlossenen Intervalle  $[a, b]$  mit  $a \leq b$  kompakt.
- iii.) Der obige Kompaktheitsbegriff wird auch *Überdeckungskompaktheit* oder *topologische Kompaktheit* genannt. Die recht unanschauliche Definition stellt trotzdem einen der wichtigsten Begriffe der Mathematik dar. Wir werden noch alternative Charakterisierungen kennenlernen aber auch diese Variante direkt benutzen.

Momentan verfügen wir noch nicht über sehr viele Beispiele kompakter Teilmengen. Es wird also nützlich sein, aus bekannten kompakten Mengen neue zu konstruieren, siehe Übung 3.9.17 für eine solche allgemeine Konstruktion. Um weitere Beispiele zu konstruieren, wird folgende Umformulierung der Definition nützlich sein:

**Proposition 3.4.6** *Eine Teilmenge  $K \subseteq \mathbb{K}$  ist genau dann kompakt, wenn für jede Familie  $\{A_i\}_{i \in I}$  von abgeschlossenen Teilmengen in  $K$  mit  $\bigcap_{i \in I} A_i = \emptyset$  bereits endlich viele  $A_{i_1}, \dots, A_{i_n}$  einen leeren Durchschnitt besitzen.*

BEWEIS: Ist  $\{A_i\}_{i \in I}$  eine Familie von abgeschlossenen Teilmengen in  $K$ , so ist  $O_i = K \setminus A_i$  offen in  $K$ . Es gilt

$$\bigcap_{i \in I} A_i = \bigcap_{i \in I} (K \setminus O_i) = K \setminus \bigcup_{i \in I} O_i,$$

womit die  $\{O_i\}_{i \in I}$  genau eine offene Überdeckung von  $K$  bilden, wenn der Durchschnitt der  $A_i$  leer ist. Damit folgt die Äquivalenz der beiden Aussagen.  $\square$

**Proposition 3.4.7** *Sei  $K \subseteq \mathbb{K}$  kompakt.*

- i.) *Zu  $p \in \mathbb{K} \setminus K$  gibt es offene Teilmengen  $O_1, O_2 \subseteq \mathbb{K}$  derart, dass  $O_1 \cap O_2 = \emptyset$  und*

$$p \in O_1 \quad \text{sowie} \quad K \subseteq O_2. \quad (3.4.6)$$

- ii.) *Die Teilmenge  $K$  ist abgeschlossen.*

- iii.) *Jede abgeschlossene Teilmenge  $A \subseteq K$  in  $K$  ist ebenfalls kompakt.*

BEWEIS: Zu  $x \in K$  finden wir eine offene Kugel  $B_{\epsilon_x}(x)$  um  $x$  sowie eine offene Kugel  $B_{\delta_x}(p)$  um  $p$  mit

$$B_{\epsilon_x}(x) \cap B_{\delta_x}(p) = \emptyset,$$

siehe auch Abbildung 3.12, da ja  $x \neq p$  und wir deshalb beispielsweise  $\epsilon_x = \delta_x = \frac{1}{2}|x - p| > 0$  verwenden können. Nun gilt  $x \in B_{\epsilon_x}(x)$  und daher

$$K \subseteq \bigcup_{x \in K} B_{\epsilon_x}(x),$$

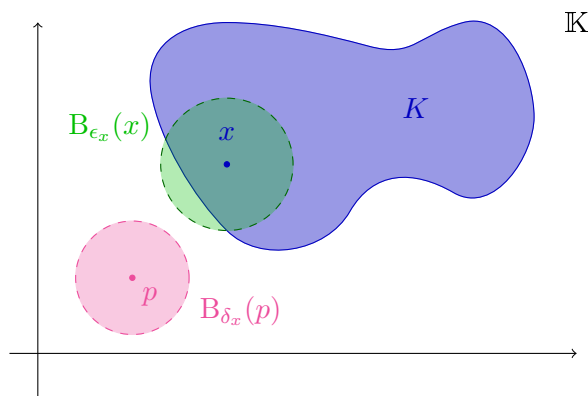


Abbildung 3.12: Punkte in  $K$  können von  $p$  durch offene Kugeln mit geeignetem Radius getrennt werden.

womit wir eine offene Überdeckung von  $K$  gefunden haben. Dank der Kompaktheit von  $K$  finden wir also endlich viele  $x_1, \dots, x_n \in K$  mit

$$K \subseteq B_{\epsilon_{x_1}}(x_1) \cup \dots \cup B_{\epsilon_{x_n}}(x_n).$$

Wir setzen

$$\delta = \min\{\delta_{x_1}, \dots, \delta_{x_n}\} > 0$$

und betrachten die offene Kugel  $B_\delta(p)$ . Es gilt nun  $B_\delta(p) \cap B_{\epsilon_{x_i}}(x_i) \subseteq B_{\delta_{x_i}}(p) \cap B_{\epsilon_{x_i}}(x_i) = \emptyset$  für alle  $i = 1, \dots, n$ , da  $\delta \leq \delta_{x_i}$ . Damit gilt auch

$$\underbrace{B_\delta(p)}_{=O_1} \cap \underbrace{(B_{\epsilon_1}(x_1) \cup \dots \cup B_{\epsilon_n}(x_n))}_{=O_2} = \emptyset,$$

womit wir die benötigten zwei offenen Mengen  $O_1$  und  $O_2$  gefunden haben. Man betrachte, dass die Endlichkeit der Überdeckung  $B_{\epsilon_1}(x_1), \dots, B_{\epsilon_n}(x_n)$  entscheidend ist, um das Minimum der Radien  $\delta$  überhaupt definieren zu können. Für den zweiten Teil wählen wir zu jedem Punkt  $p \in \mathbb{K} \setminus K$  eine offene Teilmenge  $O_p \subseteq \mathbb{K}$  mit  $p \in O_p$  und  $K \cap O_p = \emptyset$ , indem wir *i.*) benutzen. Dann gilt mit  $p \in O_p$

$$\mathbb{K} \setminus K = \bigcup_{p \in \mathbb{K} \setminus K} O_p,$$

was eine Vereinigung von offenen Mengen und damit selbst offen ist. Also ist das Komplement  $K$  abgeschlossen wie behauptet. Sei nun  $A \subseteq K$  abgeschlossen in  $K$ . Dann ist  $K \setminus A$  offen in  $K$ . Es gibt also eine offene Teilmenge  $O \subseteq \mathbb{K}$  mit

$$K \setminus A = K \cap O,$$

siehe auch Abbildung 3.13. Da zudem  $K$  selbst abgeschlossen ist, ist  $\mathbb{K} \setminus K \subseteq \mathbb{K}$  offen, womit

$$\mathbb{K} \setminus A = (\mathbb{K} \setminus K) \cup O$$

die Vereinigung zweier offener Mengen und damit selbst wieder offen ist. Damit ist  $A \subseteq \mathbb{K}$  abgeschlossen, siehe auch Übung 3.9.11 für eine allgemeine Diskussion. Seien nun  $A_i \subseteq A$  abgeschlossene Teilmengen in  $A$  für  $i \in I$ . Nach dem gerade gezeigten ist dies äquivalent zur Abgeschlossenheit  $A_i \subseteq K$  in  $K$ . Gilt nun  $\bigcap_{i \in I} A_i = \emptyset$ , so erhalten wir wegen  $A_i \subseteq K$  und der Abgeschlossenheit der  $A_i$ , dass es endlich viele  $A_{i_1}, \dots, A_{i_n}$  mit  $A_{i_1} \cap \dots \cap A_{i_n} = \emptyset$  gibt, da  $K$  kompakt ist, siehe Proposition 3.4.6. Damit ist nach Proposition 3.4.6 auch  $A$  kompakt.  $\square$

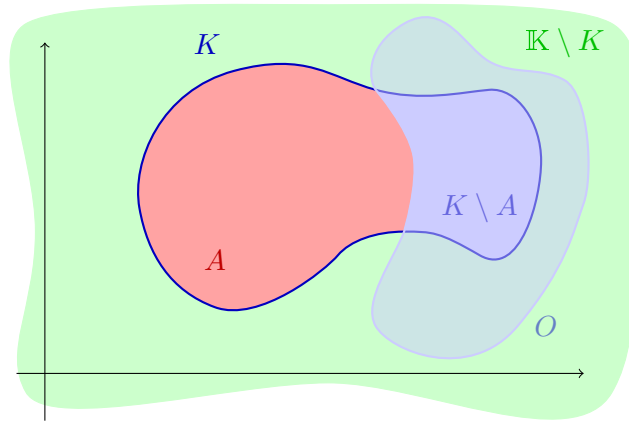


Abbildung 3.13: Eine abgeschlossene Teilmenge  $A \subseteq K$  ist abgeschlossen in  $\mathbb{K}$ .

Der nächste Satz erlaubt es, aus kompakten Teilmengen von  $\mathbb{R}$  kompakte Teilmengen von  $\mathbb{C}$  zu konstruieren. Wir benötigen dazu eine kleine Vorüberlegung:

**Lemma 3.4.8** Sei  $K \subseteq \mathbb{R}$  kompakt und  $x \in \mathbb{R}$ . Dann sind die beiden Teilmengen

$$x + iK = \{z = x + iy \mid y \in K\} \subseteq \mathbb{C} \quad (3.4.7)$$

und

$$K + ix = \{z = y + ix \mid y \in K\} \subseteq \mathbb{C} \quad (3.4.8)$$

kompakt.

BEWEIS: Die Abbildung

$$f: \mathbb{R} \ni y \mapsto x + iy \in \mathbb{C}$$

ist nach unseren bereits gezeigten Rechenregeln für stetige Abbildungen stetig. Ist also  $O \subseteq \mathbb{C}$  offen, so auch  $f^{-1}(O) \subseteq \mathbb{R}$ . Ist daher  $\{O_i\}_{i \in I}$  eine offene Überdeckung von  $x + iK = f(K) \subseteq \mathbb{C}$ , so ist  $\{f^{-1}(O_i)\}_{i \in I}$  eine offene Überdeckung von  $K$ . Als solche hat sie eine endliche Teilüberdeckung  $f^{-1}(O_{i_1}), \dots, f^{-1}(O_{i_n})$ , die dann eine endliche Teilüberdeckung  $O_{i_1}, \dots, O_{i_n}$  von  $f(K)$  liefert. Der zweite Fall ist analog.  $\square$

**Satz 3.4.9 (Tikhonov I)** Seien  $K_1, K_2 \subseteq \mathbb{R}$  kompakt. Dann ist

$$K_1 + iK_2 = \{x + iy \mid x \in K_1, y \in K_2\} \subseteq \mathbb{C} \quad (3.4.9)$$

wieder kompakt.

BEWEIS: Sei  $\{O_j\}_{j \in I}$  eine offene Überdeckung von  $K = K_1 + iK_2$ . Da eine offene Teilmenge  $O_j \subseteq \mathbb{C}$  eine geeignete Vereinigung von offenen Kugeln ist, können wir auch direkt annehmen, dass wir eine offene Überdeckung durch offene Kugeln vorliegen haben. Ist  $w \in B_r(z) \subseteq \mathbb{C}$ , so gibt es ein  $\epsilon_w > 0$  mit der Eigenschaft, dass

$$(w_1 - \epsilon_w, w_1 + \epsilon_w) + i(w_2 - \epsilon_w, w_2 + \epsilon_w) \subseteq B_r(z)$$

gilt, siehe auch Abbildung 3.14. Jeder Punkt  $w \in B_r(z)$  besitzt also eine kleine offene Umgebung



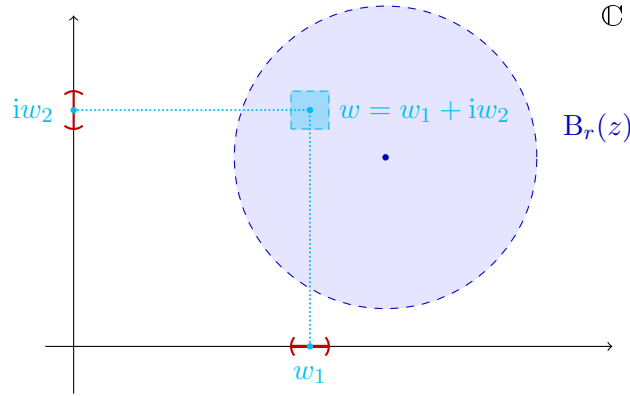
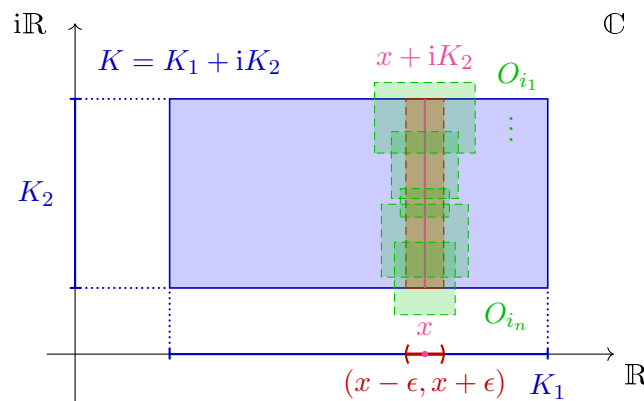


Abbildung 3.14: Offene Kugeln erhält man als Vereinigung von offenen Quadraten.

Abbildung 3.15: Endliche Überdeckung von  $x + iK_2$ .

innerhalb von  $B_r(z)$  von der Form eines Quadrats. Dies zeigt aber

$$B_r(z) = \bigcup_{w \in B_r(z)} (w_1 - \epsilon_w, w_1 + \epsilon_w) + i(w_2 - \epsilon_w, w_2 + \epsilon_w),$$

womit jede offene Kugel die Vereinigung offener Quadrate ist. Um also die Kompaktheit von  $K$  zu zeigen, können wir annehmen, dass die Überdeckung  $\{O_j\}_{j \in I}$  bereits aus offenen Quadraten besteht: Zeigen wir nämlich, dass dann, wenn  $K$  die endliche Überdeckungseigenschaft für jede Überdeckung durch offene Quadrate besitzt, von den größeren, beliebigen offenen Mengen erst recht endlich viele reichen. Seien also  $O_j = U_j + iV_j$  mit offenen Intervallen  $U_j, V_j \subseteq \mathbb{R}$  derart, dass  $K \subseteq \bigcup_{j \in I} O_j$ . Wir wählen ein  $x \in K_1$  aus und betrachten die nach Lemma 3.4.8 kompakte Teilmenge  $x + iK_2 \subseteq K_1 + iK_2$ . Diese wird durch  $\{O_j\}_{j \in I}$  erst recht überdeckt, womit es also endlich viele  $O_{j_1}, \dots, O_{j_n}$  gibt, so dass

$$x + iK_2 \subseteq O_{j_1} \cup \dots \cup O_{j_n},$$

siehe auch Abbildung 3.15. Da  $O_j = U_j + iV_j$  ein Quadrat ist, gibt es also  $\epsilon_1, \dots, \epsilon_n$  und Mittelpunkte  $x_1, \dots, x_n$  mit

$$U_{j_1} = (x_1 - \epsilon_1, x_1 + \epsilon_1), \dots, U_{j_n} = (x_n - \epsilon_n, x_n + \epsilon_n).$$

Es gilt  $x \in U_{j_1} \cap \dots \cap U_{j_n}$  und dieser endliche Schnitt von offenen Intervallen ist wieder offen. Daher gibt es ein  $\epsilon > 0$  mit

$$(x - \epsilon, x + \epsilon) \subseteq U_{j_1} \cap \dots \cap U_{j_n}.$$

Es folgt, dass

$$\begin{aligned}(x - \epsilon, x + \epsilon) + iK_2 &\subseteq (U_{j_1} \cap \cdots \cap U_{j_n}) + iK_2 \\ &\subseteq (U_{j_1} + iV_{j_1}) \cap \cdots \cap (U_{j_n} + iV_{j_n}) \\ &= O_{j_1} \cup \cdots \cup O_{j_n}.\end{aligned}$$

Geometrisch bedeutet dies, dass der „Streifen“  $(x - \epsilon, x + \epsilon) + iK_2$  immer noch von den endlich vielen  $O_{j_1} \cup \cdots \cup O_{j_n}$  überdeckt wird. Wir können nun für jedes  $x \in K_1$  ein passendes  $\epsilon_x > 0$  mit dieser Eigenschaft finden. Da  $K_1$  kompakt ist, überdecken aber endlich viele der offenen Intervalle  $(x_1 - \epsilon_1, x_1 + \epsilon_1), \dots, (x_m - \epsilon_m, x_m + \epsilon_m)$  bereits  $K_1$ , denn alle diese Intervalle bilden wegen  $x \in (x - \epsilon_x, x + \epsilon_x)$  sicherlich eine offene Überdeckung. Damit werden die endlich vielen zugehörigen Streifen aber jeweils durch endlich viele offene Mengen aus der ursprünglichen Überdeckung überdeckt, und wir haben insgesamt eine endliche Teilüberdeckung für alle Streifen gefunden. Diese Streifen zusammen enthalten aber  $K_1 + iK_2$ .  $\square$

**Bemerkung 3.4.10 (Tikhonov)** Identifiziert man  $\mathbb{C} = \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$  als kartesisches Produkt, so entspricht  $K_1 + iK_2$  gerade dem Produkt  $K_1 \times K_2$ . Damit haben wir gezeigt, dass das kartesische Produkt von zwei kompakten Teilmengen aus  $\mathbb{R}$  eine kompakte Teilmenge des kartesischen Produkts  $\mathbb{R} \times \mathbb{R}$  wird. Der wirkliche Satz von Tikhonov verallgemeinert diese Aussage drastisch und wird in der mengen-theoretischen Topologie bewiesen. Er ist eines der Glanzstücke der Topologie und besitzt unzählige Anwendungen in der Funktionalanalysis und anderen Gebieten der Mathematik. Unsere Formulierung ist daher die Baby-Version, enthält allerdings so gut wie alle wesentlichen Elemente des allgemeinen Beweises. Wir kommen auf eine Verallgemeinerung in Abschnitt 7.6.

Insbesondere sind also Rechtecke der Form

$$[a_1, b_1] \times [a_2, b_2] = [a_1, b_1] + i[a_2, b_2] \subseteq \mathbb{C} \quad (3.4.10)$$

wieder kompakte Teilmengen von  $\mathbb{C}$ , ganz in Analogie zu den kompakten Intervallen  $[a, b] \subseteq \mathbb{R}$ . Hier hätte man auch anders die Kompaktheit nachprüfen können: Wie im Beweis von Satz 3.4.3 für ein einzelnes Intervall  $[a, b]$  könnte man (3.4.10) durch „Rechteckviertelung“ als kompakt erweisen, siehe auch Übung 3.9.18. Mit dem Resultat, dass die abgeschlossenen Rechtecke der Form (3.4.10) kompakt in  $\mathbb{C}$  sind, erhalten wir dann folgende umfassende Charakterisierung kompakter Teilmengen von  $\mathbb{R}$  und  $\mathbb{C}$ :

**Satz 3.4.11 (Heine-Borel)** Sei  $K \subseteq \mathbb{K}$  eine Teilmenge. Dann sind folgende Aussagen äquivalent:

- i.) Die Teilmenge  $K$  ist kompakt.
- ii.) Die Teilmenge  $K$  ist abgeschlossen und beschränkt.
- iii.) Die Teilmenge  $K$  ist folgenkompakt in dem Sinne, dass jede Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n \in K$  eine konvergente Teilfolge mit Grenzwert in  $K$  besitzt.

BEWEIS: Mit einiger Redundanz zeigen wir alle Äquivalenzen einzeln, da hier jeweils interessante zusätzliche Beweisideen versteckt sind. Sei also zunächst  $K$  kompakt. Sei nun für jeden Punkt  $z \in \mathbb{K}$  die offene Kugel  $B_1(z)$  gegeben, was sicherlich eine offene Überdeckung  $\{B_1(z)\}_{z \in \mathbb{K}}$  von ganz  $\mathbb{K}$  und daher auch von  $K$  liefert. Endlich viele Kugeln  $B_1(z_1), \dots, B_1(z_n)$  überdecken also  $K$ , siehe Abbildung 3.16. Damit ist  $K$  aber sicherlich beschränkt, da der „Durchmesser“ von  $K$  höchstens  $2n$  ist. Nach Proposition 3.4.7, ii.), ist  $K$  auch abgeschlossen, was i.)  $\implies$  ii.) zeigt.

Sei nun  $K \subseteq \mathbb{K}$  abgeschlossen und beschränkt. Dann gibt es also ein großes Rechteck  $[a_1, b_1] \times [a_2, b_2] \subseteq \mathbb{C}$  beziehungsweise ein großes Intervall  $[a, b] \subseteq \mathbb{R}$ , mit

$$K \subseteq [a_1, b_1] \times [a_2, b_2] \quad \text{beziehungsweise} \quad K \subseteq [a, b],$$

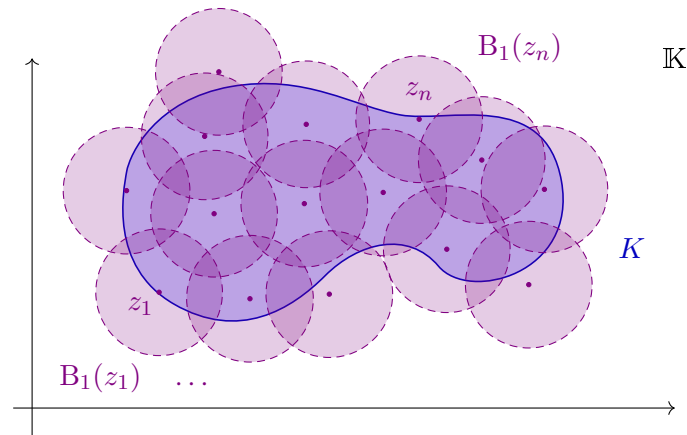


Abbildung 3.16: Eine kompakte Teilmenge  $K$  wird von endlich vielen Kugeln vom Radius 1 überdeckt.

da  $K$  ja beschränkt ist. Nach Satz 3.4.3 beziehungsweise Satz 3.4.9 sind diese Teilmengen aber kompakt, womit  $K$  als abgeschlossene Teilmenge einer kompakten Menge nach Proposition 3.4.6 wieder kompakt ist. Damit ist  $ii.) \implies i.)$  gezeigt.

Sei wieder  $K$  kompakt und  $(a_n)_{n \in \mathbb{N}}$  eine Folge in  $K$ . Nach  $ii.)$  ist  $(a_n)_{n \in \mathbb{N}} \in \ell^\infty$  eine beschränkte Folge und hat daher nach dem Satz von Bolzano-Weierstraß, siehe Satz 2.4.10, eine konvergente Teilfolge. Deren Grenzwert ist ein Berührungspunkt von  $K$  und nach Proposition 3.1.3,  $iv.)$ , in der kleinsten abgeschlossenen Teilmenge  $K^{\text{cl}}$  enthalten, die  $K$  umfasst. Da aber  $K$  selbst abgeschlossen ist, gilt  $K = K^{\text{cl}}$  und der Grenzwert der Teilfolge ist ein Element von  $K$ . Dies zeigt  $i.) \implies iii.)$ .

Es gelte schließlich  $iii.)$ . Wäre  $K$  unbeschränkt, so gäbe es eine Folge  $(a_n)_{n \in \mathbb{N}}$  in  $K$  mit  $|a_n| \rightarrow \infty$ . Insbesondere hat eine solche Folge keine konvergente Teilfolge in  $\mathbb{K}$ , also auch nicht in  $K$ . Damit ist  $K$  notwendigerweise beschränkt. Sei nun  $x \in K^{\text{cl}} \setminus K$  ein Punkt im Abschluss von  $K$ , der nicht in  $K$  enthalten ist. Da  $K^{\text{cl}}$  die disjunkte Vereinigung von isolierten Punkten aus  $K$  und Häufungspunkten von  $K$  ist, siehe erneut Proposition 3.1.5 und (3.1.4), folgt, dass  $x$  ein Häufungspunkt von  $K$  ist. Nach Definition eines Häufungspunktes gibt es eine Folge  $(a_n)_{n \in \mathbb{N}}$  von Elementen in  $K$ , die gegen  $x$  konvergiert. Nach  $iii.)$  folgt  $x \in K$ , ein Widerspruch. Also gilt bereits  $K = K^{\text{cl}}$ , was  $iii.) \implies ii.)$  zeigt.  $\square$

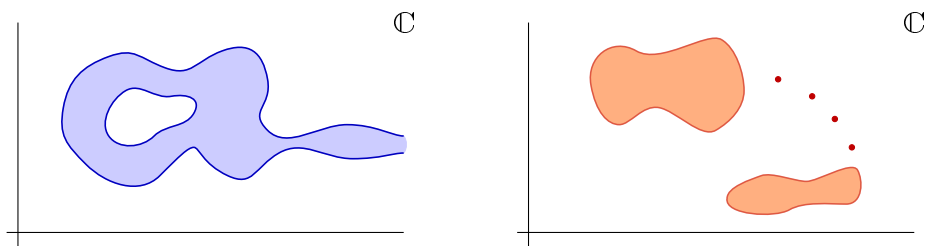
Der Satz von Heine-Borel ist die Äquivalenz von  $i.)$  und  $ii.)$ . Die Äquivalenz von  $i.)$  und  $iii.)$  wird oft genutzt, um Kompaktheit von Teilmengen aus  $\mathbb{K}$  zu *definieren*. Diese alternative Definition hat den Vorteil, dass die Argumentation mit der Überdeckungskompaktheit anfangs oftmals als unanschaulich und schwierig empfunden wird. Trotzdem erweist sich die Überdeckungskompaktheit als die viel weiterführende Variante als die Folgenkompaktheit. Letztlich ist es wichtig, dass man deren Äquivalenz für Teilmengen von  $\mathbb{K}$  kennt. Welche Variante man dann als Definition betrachtet, und welche als Konsequenz, ist eine Frage der persönlichen Vorlieben.

Die zweite wichtige Eigenschaft von Teilmengen von  $\mathbb{K}$ , die wir hier diskutieren wollen, ist der Zusammenhang. Wieder ist eine vor-mathematische Anschauung vorhanden, was eine zusammenhängende Teilmenge sein soll. Eine tatsächliche Formalisierung als mathematisches Konzept ist dagegen alles andere als einfach, siehe auch Abbildung 3.17. Der folgende Satz gibt eine erste Idee, wie die Definition von zusammenhängenden Teilmengen aussehen könnte:

**Satz 3.4.12 (Zusammenhang von  $[a, b]$ )** Seien  $a, b \in \mathbb{R}$  mit  $a < b$  gegeben. Seien  $O_1, O_2 \subseteq [a, b]$  offene Teilmengen in  $[a, b]$  mit den Eigenschaften

$$[a, b] = O_1 \cup O_2 \quad \text{und} \quad O_1 \cap O_2 = \emptyset. \quad (3.4.11)$$

Dann ist eine der beiden Teilmengen leer und die andere ist  $[a, b]$ .

Abbildung 3.17: Zusammenhängende und nicht zusammenhängende Teilmengen von  $\mathbb{C}$ .

BEWEIS: Wir nehmen an, dass beide Teilmengen nicht leer sind, um einen Widerspruch zu erzielen. Wir finden also Punkte  $x \in O_1$  und  $y \in O_2$  mit  $x \neq y$ . Ohne Einschränkung können wir dann  $x < y$  annehmen, sonst wechseln wir die Rollen von  $O_1$  und  $O_2$ . Da  $O_1$  offen in  $[a, b]$  ist, können wir ohne Einschränkung sogar  $a < x$  erreichen, denn mit  $a \in O_1$  sind auch Punkte etwas größer als  $a$  noch in  $O_1$  enthalten und trotzdem noch kleiner als  $y$ . Gilt dagegen  $a \in O_2$ , ist  $a < x$  automatisch erfüllt, da ja  $x \in O_1$  gilt und daher von  $a$  verschieden sein muss. Genauso können wir  $y < b$  erreichen, so dass wir insgesamt also

$$a < x < y < b$$

mit  $x \in O_1$  und  $y \in O_2$  vorliegen haben. Wir betrachten nun

$$z = \sup\{\xi \mid [x, \xi] \subseteq O_1\}.$$

Da  $x < y \in O_2$  gilt, folgt aus der Offenheit von  $O_2$  auch  $z < y$ . Da auch  $O_1$  offen ist, folgt  $z > x$ . Wir haben also die zwei Möglichkeiten  $z \in O_1$  oder  $z \in O_2$  zu betrachten, da  $z \in [a, b] = O_1 \cup O_2$ . Gilt  $z \in O_1$ , so gibt es ein  $\epsilon > 0$  mit  $z \in B_\epsilon(z) \subseteq O_1$ , da  $O_1$  offen ist. Damit wäre aber auch  $z + \frac{\epsilon}{2}$  ein Punkt mit

$$[x, z + \frac{\epsilon}{2}] = [x, z] \cup [z, \frac{\epsilon}{2}] \subseteq O_1,$$

im Widerspruch zur Supremumseigenschaft von  $z$ . Ist dagegen  $z \in O_2$ , so gibt es ebenfalls ein  $\epsilon > 0$  mit  $z \in B_\epsilon(z) \subseteq O_2$ , da auch  $O_2$  offen ist. Dann ist aber  $[z - \frac{\epsilon}{2}, z] \subseteq O_2$  und folglich kann es kein  $\xi$  mit  $[x, \xi] \subseteq O_1$  und  $\xi \in [z - \frac{\epsilon}{2}, z)$  geben, im Widerspruch zur Supremumseigenschaft von  $z$ . Insgesamt haben wir daher einen Widerspruch erreicht.  $\square$

Wir verwenden diese Eigenschaft nun als Definition einer *zusammenhängenden Teilmenge*:

**Definition 3.4.13 (Zusammenhang)** Sei  $A \subseteq \mathbb{K}$  eine Teilmenge. Dann heißt  $A$  *zusammenhängend*, wenn es keine zwei nicht-leeren offenen Teilmengen  $O_1$  und  $O_2$  in  $A$  gibt, so dass

$$A = O_1 \cup O_2 \quad \text{und} \quad O_1 \cap O_2 = \emptyset \quad (3.4.12)$$

gilt.

Mit dieser Definition sind also die abgeschlossenen Intervalle  $[a, b] \subseteq \mathbb{R}$  zusammenhängend, dies ist gerade die Aussage von Satz 3.4.12. Mit ähnlichen Argumenten zeigt man, dass auch die anderen Typen von Intervallen

$$(a, b], [a, b), (a, b) \subseteq \mathbb{R} \quad (3.4.13)$$

mit  $-\infty \leq a < b \leq +\infty$  zusammenhängend sind. Dies sind aber auch schon alle Möglichkeiten:

**Proposition 3.4.14** Sei  $A \subseteq \mathbb{R}$  zusammenhängend mit  $a < b$  und  $a, b \in A$ . Dann gilt  $[a, b] \subseteq A$ .

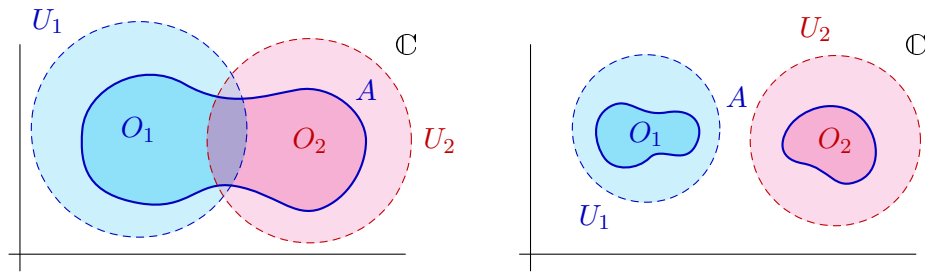


Abbildung 3.18: Im linken Bild überlappen nicht-leere offene Teilmengen  $O_1, O_2 \subseteq A$  in  $A$  notwendigerweise:  $A$  ist zusammenhängend. Im rechten Bild findet man nicht-leere offene Teilmengen  $O_1, O_2 \subseteq A$  in  $A$ , die ganz  $A$  überdecken und trotzdem leeren Durchschnitt besitzen. Man beachte, dass die offenen Teilmengen *in*  $A$  gemeint sind, nicht die offenen Teilmengen  $U_1, U_2 \subseteq \mathbb{K}$ , die  $O_i = U_i \cap A$  liefern.

BEWEIS: Wir nehmen an,  $z \in [a, b]$  sei ein Punkt, der nicht Element von  $A$  ist. Dann sind

$$U_1 = (-\infty, z) \quad \text{und} \quad U_2 = (z, +\infty)$$

offene Teilmengen von  $\mathbb{R}$ . Demnach sind

$$O_1 = U_1 \cap A \quad \text{und} \quad O_2 = U_2 \cap A$$

offene Teilmengen in  $A$ . Als Teilmengen von  $\mathbb{R}$  müssen  $O_1$  und  $O_2$  ja nicht unbedingt offen sein. Es gilt  $a \in O_1$  sowie  $b \in O_2$ . Da  $z \notin A$ , folgt  $A \subseteq U_1 \cup U_2$  also  $A = O_1 \cup O_2$ . Schließlich ist  $O_1 \cap O_2 \subseteq U_1 \cap U_2 = \emptyset$ , womit wir einen Widerspruch zum Zusammenhang von  $A$  erreicht haben.  $\square$

Mit Hilfe dieses Ergebnisses zeigt man nun leicht, dass die Liste (3.4.13) zusammen mit den abgeschlossenen Intervallen bereits alle zusammenhängenden Teilmengen von  $\mathbb{R}$  ausschöpft, siehe Übung 3.9.19. Für  $\mathbb{K} = \mathbb{C}$  gibt es jedoch bereits Beispiele für zusammenhängende Teilmengen, die die Anschauung durchaus strapazieren. Ein recht drastisches Beispiel werden wir in ?? kennenlernen.

## 3.5 Hauptsätze für stetige Funktionen

Mit dem Rüstzeug des letzten Abschnitts können wir nun die wichtigen Sätze zu stetigen Funktionen formulieren und beweisen.

Die erste Situation ist noch recht speziell auf reell-wertige Funktionen einer reellen Variable zugeschnitten. Wie bei Folgen definieren wir auch *Monotonie* für Funktionen:

**Definition 3.5.1 (Monotone Funktion)** Sei  $X \subseteq \mathbb{R}$  und  $f: X \rightarrow \mathbb{R}$  eine Funktion.

i.) Die Funktion  $f$  heißt *monoton wachsend (fallend)*, falls für alle  $x, y \in X$  mit  $x \leq y$  auch

$$f(x) \leq f(y) \quad \text{beziehungsweise} \quad f(x) \geq f(y) \quad (3.5.1)$$

gilt.

ii.) Die Funktion heißt *streng monoton wachsend (fallend)*, falls für  $x, y \in X$  mit  $x < y$  auch

$$f(x) < f(y) \quad \text{beziehungsweise} \quad f(x) > f(y) \quad (3.5.2)$$

gilt.

iii.) Die Funktion heißt *(streng) monoton*, falls  $f$  *(streng) monoton wachsend oder (streng) monoton fallend* ist.

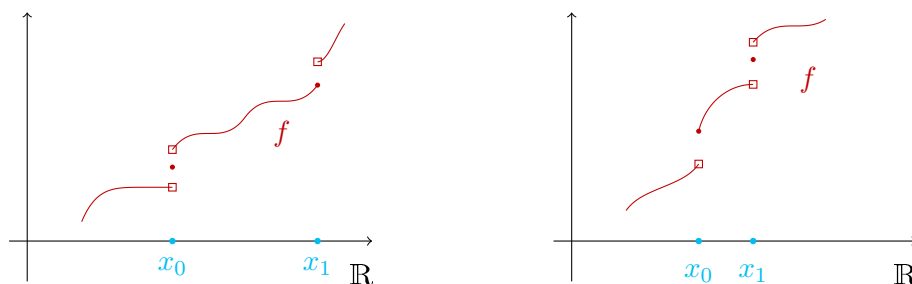


Abbildung 3.19: Monoton wachsende und streng monoton wachsende Funktionen.

Die Situationen in Abbildung 3.19 legen nahe, dass monotone Funktionen zwar nicht stetig sein müssen, aber andererseits auch nicht viele Unstetigkeitsstellen haben können, da es nur Sprünge nach oben beziehungsweise nach unten geben kann, die ihren „Platz“ im Wertebereich erfordern. Der folgende Satz zeigt, dass diese heuristische Vorstellung der mathematischen Realität recht nahe kommt:

**Satz 3.5.2 (Stetigkeit monotoner Funktionen)** Sei  $X \subseteq \mathbb{R}$  eine nicht-leere Teilmenge, und sei  $f: X \rightarrow \mathbb{R}$  eine monotone Funktion.

i.) Für  $x_0 \in X$  existieren die links- und rechtsseitigen Grenzwerte von  $f$ . Für eine monoton wachsende Funktion gilt

$$\lim_{x \nearrow x_0} f(x) = \sup\{f(x) \mid x < x_0\} \quad \text{sowie} \quad \lim_{x \searrow x_0} f(x) = \inf\{f(x) \mid x > x_0\}, \quad (3.5.3)$$

falls  $x_0$  nicht das Minimum beziehungsweise das Maximum von  $X$  ist. Für eine monoton fallende Funktion vertauscht man entsprechend die Rollen von Supremum und Infimum.

ii.) Die Funktion  $f$  ist genau dann stetig bei  $x_0 \in X$ , wenn

$$\lim_{x \searrow x_0} f(x) = \lim_{x \nearrow x_0} f(x). \quad (3.5.4)$$

In diesem Fall stimmen die Grenzwerte mit  $f(x_0)$  überein.

iii.) Die Funktion  $f$  hat höchstens abzählbar viele Unstetigkeitsstellen.

BEWEIS: Sei  $f: X \rightarrow \mathbb{R}$  monoton wachsend. Ist  $x_0 = \min X \in X$  das Minimum, sofern  $X$  überhaupt ein Minimum besitzt, gilt trivialerweise  $\lim_{x \nearrow x_0} f(x) = f(x_0)$ , da es keine anderen Punkte  $x \in X$  mit  $x \leq x_0$  außer  $x = x_0$  gibt, siehe Definition 3.1.8, iv.). Damit existiert also  $\lim_{x \nearrow x_0} f(x)$  aus recht trivialen Gründen. Wir nehmen daher an, dass es Punkte  $x \in X$  mit  $x < x_0$  gibt. In diesem Fall ist das Supremum nicht  $-\infty$ . Da aus Monotoniegründen  $f(x) \leq f(x_0)$  gilt, folgt zunächst

$$\sup\{f(x) \mid x < x_0\} \leq f(x_0). \quad (*)$$

Sei nun  $\epsilon > 0$  vorgegeben und  $(x_n)_{n \in \mathbb{N}}$  eine Folge in  $X$  mit  $x_n < x_0$  sowie  $x_n \rightarrow x_0$ . Dann gibt es zunächst ein  $x_\epsilon \in X$  mit  $x_\epsilon < x_0$  und

$$f(x_\epsilon) > \sup\{f(x) \mid x < x_0\} - \epsilon$$

nach Definition des Supremums als kleinste obere Schranke. Da  $f$  monoton wachsend ist, folgt aber  $f(x') \geq f(x_\epsilon) > \sup\{f(x) \mid x < x_0\} - \epsilon$  für alle  $x_\epsilon \leq x' < x_0$ . Sei also  $n \in \mathbb{N}$  derart, dass  $x_\epsilon \leq x_n \leq x_0$  für alle  $n \geq N$ , was nach  $x_n \rightarrow x_0$  zu finden ist. Dann gilt also

$$f(x_n) > \sup\{f(x) \mid x < x_0\} - \epsilon,$$

oder eben  $|f(x_n) - \sup\{f(x) \mid x < x_0\}| < \epsilon$ . Dies zeigt aber die Konvergenz

$$\lim_{n \rightarrow \infty} f(x_n) = \sup\{f(x) \mid x < x_0\},$$

was die Existenz des ersten Grenzwertes sowie die erste Gleichheit in (3.5.3) klärt. Die zweite Variante erhält man analog. Ist schließlich  $f$  monoton fallend, so kann man die bereits erzielten Ergebnisse auf die monoton wachsende Funktion  $-f$  anwenden. Damit folgt der erste Teil. Da die Existenz der Limites durch den ersten Teil bereits gesichert ist, folgt der zweite Teil sofort, da aus (\*) und der entsprechenden Aussage für das Infimum

$$\lim_{x \nearrow x_0} f(x) = \sup\{f(x) \mid x < x_0\} \leq f(x_0) \leq \inf\{f(x) \mid x > x_0\} = \lim_{x \searrow x_0} f(x) \quad (**)$$

folgt. Stetigkeit liegt vor, wenn die links- und rechtsseitigen Limites gleich sind und mit  $f(x_0)$  übereinstimmen, siehe auch Übung 3.9.12. Mit (\*\*) genügt es aber, dass (3.5.4) gilt, dann stimmen die beiden Limites auch mit  $f(x_0)$  überein. Der letzte Teil ist nun einfach. Ist  $U \subseteq X$  die Menge der Unstetigkeitsstellen von  $f$  und  $x_0 \in U$ , so gilt

$$\lim_{x \nearrow x_0} f(x) < \lim_{x \searrow x_0} f(x).$$

Daher finden wir eine rationale Zahl  $q_{x_0} \in \mathbb{Q}$  mit

$$\lim_{x \nearrow x_0} f(x) < q_{x_0} < \lim_{x \searrow x_0} f(x).$$

Dies definiert eine Abbildung

$$q: U \ni x_0 \mapsto q_{x_0} \in \mathbb{Q},$$

welche nach Konstruktion  $q_x < q_y$  für  $x < y$  erfüllt. Insbesondere ist  $q$  injektiv. Da  $\mathbb{Q}$  abzählbar ist, muss  $U$  entsprechend ebenfalls abzählbar gewesen sein.  $\square$

Dieser Satz über monotone Funktionen benutzt noch sehr spezifische Eigenschaften der reellen Zahlen und hat daher kein unmittelbares Analogon für komplexe Zahlen. Dies ändert sich mit den nächsten Ergebnissen, die sowohl im reellen als auch im komplexen Fall richtig sind. Tatsächlich besitzen die nächsten Sätze eine viel größere Gültigkeit, als wir es zum jetzigen Zeitpunkt absehen können. Wir werden in Abschnitt 7.3 sowie in Abschnitt 7.5 nochmals darauf zurückkommen.

**Satz 3.5.3 (Verkettung stetiger Funktionen)** *Seien  $X$  und  $Y$  nicht-leere Teilmengen der reellen oder komplexen Zahlen und  $f: X \rightarrow Y$  sowie  $g: Y \rightarrow \mathbb{K}$  stetige Funktionen. Dann ist die Verkettung  $g \circ f$  ebenfalls stetig.*

BEWEIS: Sei  $O \subseteq \mathbb{K}$  offen. Da  $g$  stetig ist, ist nach Satz 3.2.7, ii.), das Urbild  $g^{-1}(O) \subseteq Y$  offen in  $Y$ . Verwendet man die Stetigkeit von  $f$ , so zeigt dies, dass  $f^{-1}(g^{-1}(O)) \subseteq X$  offen in  $X$  ist. Da  $f^{-1}(g^{-1}(O)) = (g \circ f)^{-1}(O)$  allgemein gilt, ist  $(g \circ f)^{-1}(O) \subseteq X$  offen in  $X$  für jede offene Teilmenge  $O \subseteq \mathbb{K}$ . Nach Satz 3.2.7, ii.), ist  $g \circ f$  stetig.  $\square$

Dieser fast schon enttäuschend einfache Beweis zeigt, dass die wahre Natur der Stetigkeit in Satz 3.2.7, ii.) liegt. Eine analoge Aussage ist für Funktionen richtig, die nur an einem Punkt stetig sind, siehe Übung 3.9.23. Alternative Beweise von Satz 3.5.3 finden sich in Übung 3.9.21 diskutiert.

Der nächste Satz ist von ähnlicher Natur und kaum zu überschätzender Wichtigkeit. Wir werden im Folgenden viele Anwendungen und Spezialfälle sehen.

**Satz 3.5.4 (Bilder von Kompakta)** *Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen. Ist  $f: X \rightarrow \mathbb{K}$  stetig und  $K \subseteq X$  kompakt, so ist  $f(K)$  ebenfalls kompakt.*

BEWEIS: Sei  $\{O_i\}_{i \in I}$  eine offene Überdeckung von  $f(K)$ , also  $O_i \subseteq \mathbb{K}$  offen mit  $f(K) \subseteq \bigcup_{i \in I} O_i$ . Da  $f$  stetig ist, ist  $f^{-1}(O_i) \subseteq X$  offen in  $X$  für alle  $i \in I$ , womit wir eine offene Überdeckung  $\{f^{-1}(O_i)\}_{i \in I}$  von  $K$  erhalten, da ja

$$K \subseteq f^{-1}(f(K)) \subseteq f^{-1}\left(\bigcup_{i \in I} O_i\right) = \bigcup_{i \in I} f^{-1}(O_i)$$

allgemein gilt. Es gibt also endlich viele  $i_1, \dots, i_n \in I$ , so dass

$$K \subseteq f^{-1}(O_{i_1}) \cup \dots \cup f^{-1}(O_{i_n}) = f^{-1}(O_{i_1} \cup \dots \cup O_{i_n}).$$

Damit folgt aber

$$f(K) \subseteq f(f^{-1}(O_{i_1} \cup \dots \cup O_{i_n})) = O_{i_1} \cup \dots \cup O_{i_n},$$

was die Kompaktheit von  $f(K)$  zeigt. □

Da kompakte Mengen beschränkt sind, betrachtet man allgemein solche Funktionen, deren Bildmengen beschränkt sind:

**Definition 3.5.5 ((Lokal) Beschränkte Funktionen)** Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen.

i.) Eine Funktion  $f: X \rightarrow \mathbb{K}$  heißt beschränkt, falls

$$\|f\|_\infty = \sup_{x \in X} |f(x)| < \infty \quad (3.5.5)$$

gilt. Wir setzen

$$\mathcal{B}(X, \mathbb{K}) = \{f: X \rightarrow \mathbb{K} \mid f \text{ ist beschränkt}\}. \quad (3.5.6)$$

ii.) Eine Funktion  $f: X \rightarrow \mathbb{K}$  heißt lokal beschränkt, falls es zu jedem  $x \in X$  eine Umgebung  $U \subseteq X$  von  $x$  in  $X$  gibt, so dass die Einschränkung  $f|_U$  von  $f$  auf  $U$  beschränkt ist. Wir setzen

$$\mathcal{B}_{\text{loc}}(X, \mathbb{K}) = \{f: X \rightarrow \mathbb{K} \mid f \text{ ist lokal beschränkt}\}. \quad (3.5.7)$$

iii.) Ist  $X$  offen und  $K \subseteq X$  kompakt, so definiert man

$$\|f\|_K = \sup_{x \in K} |f(x)| \in [0, \infty]. \quad (3.5.8)$$

Erste Eigenschaften von  $\mathcal{B}(X, \mathbb{K})$  und  $\mathcal{B}_{\text{loc}}(X, \mathbb{K})$  finden sich in Übung 3.9.25. Man beachte, dass es durchaus Funktionen gibt, die auf jeder offenen Umgebung eines Punktes unbeschränkt sind, siehe erneut Beispiel 3.3.6, iii.), für eine recht absonderliche solche Funktion. Vor diesem Hintergrund ist folgende Eigenschaft von stetigen Funktionen sehr willkommen:

**Korollar 3.5.6** Sei  $f: X \rightarrow \mathbb{K}$  stetig.

i.) Ist  $X$  kompakt, so ist  $f$  beschränkt.

ii.) Ist  $X$  offen, so ist  $f$  lokal beschränkt.

BEWEIS: Mit  $f$  ist auch  $|f|$  stetig, da

$$|f| = \sqrt{\cdot} \circ f\bar{f}$$

und die Wurzelfunktion ebenso wie  $f\bar{f}$  stetig sind. Ist  $X$  kompakt, so ist auch  $|f|(X) = \{|f(x)| \mid x \in X\} \subseteq \mathbb{R}$  nach Satz 3.5.4 wieder kompakt und damit beschränkt. Den zweiten Teil erhält man daraus, da es zu jedem  $x \in X$  eine Umgebung  $B_\epsilon(x)^{\text{cl}} \subseteq X$  gibt. Zuerst findet man  $\epsilon > 0$  und eine offene Kugel  $B_{2\epsilon}(x) \subseteq X$ , da  $X$  offen ist. Die abgeschlossene Kugel  $B_\epsilon(x)^{\text{cl}} \subseteq B_{2\epsilon}(x)$  ist dann auch in  $X$  enthalten und kompakt, womit  $f$  auf  $B_\epsilon(x)^{\text{cl}}$  beschränkt ist. □



**Korollar 3.5.7** *Ist  $X$  kompakt, so gilt*

$$\mathcal{C}(X, \mathbb{K}) \subseteq \mathcal{B}(X, \mathbb{K}). \quad (3.5.9)$$

*Ist  $X$  offen, so gilt*

$$\mathcal{C}(X, \mathbb{K}) \subseteq \mathcal{B}_{\text{loc}}(X, \mathbb{K}). \quad (3.5.10)$$

**Korollar 3.5.8** (*Maximum und Minimum stetiger Funktionen*) *Sei  $X$  eine nicht-leere kompakte Teilmenge der reellen oder komplexen Zahlen und  $f \in \mathcal{C}(X, \mathbb{R})$ . Dann gibt es  $x_{\max}, x_{\min} \in X$  mit*

$$f(x_{\max}) = \sup_{x \in X} f(x) \quad (3.5.11)$$

*und*

$$f(x_{\min}) = \inf_{x \in X} f(x). \quad (3.5.12)$$

BEWEIS: Wir wissen, dass  $f(X) \subseteq \mathbb{R}$  kompakt und daher beschränkt ist. Das Supremum sowie das Infimum über die Funktionswerte sind also endlich. Da das Bild  $f(X) \subseteq \mathbb{R}$  auch abgeschlossen ist, enthält es alle seine Berührungspunkte, wie eben auch das Supremum und das Infimum. Also gilt

$$\sup_{x \in X} f(x), \inf_{x \in X} f(x) \in f(X).$$

Damit gibt es aber Urbilder  $x_{\max}$  und  $x_{\min}$  mit (3.5.11) und (3.5.12).  $\square$

Dies ist eines der wichtigsten Resultate zu reell-wertigen stetigen Funktionen auf kompakten Definitionsbereichen. Sind die Werte stattdessen in  $\mathbb{C}$ , gibt es natürlich kein Supremum oder Infimum. Man kann jedoch die Funktion  $|f|$  anstelle von  $f$  betrachten, für welche dann wieder Korollar 3.5.8 zur Anwendung kommt.

Ebenso wie kompakte Definitionsbereiche verhalten sich auch zusammenhängende Teilmengen gut unter stetigen Funktionen:

**Satz 3.5.9 (Bilder von zusammenhängenden Mengen)** *Sei  $X$  eine nicht-leere Teilmenge von den reellen oder komplexen Zahlen, und sei  $f: X \rightarrow \mathbb{K}$  stetig. Ist  $A \subseteq X$  zusammenhängend, so ist  $f(A) \subseteq \mathbb{K}$  ebenfalls zusammenhängend.*

BEWEIS: Wir nehmen an,  $f(A)$  ist nicht zusammenhängend. Es gibt also zwei nicht-leere offene Mengen  $U_i \subseteq f(A)$  in  $f(A)$  mit  $i = 1, 2$  derart, dass

$$U_1 \cap U_2 = \emptyset \quad \text{und} \quad U_1 \cup U_2 = f(A)$$

gilt. Offen in  $f(A)$  zu sein, heißt, dass es offene Teilmengen  $O_1, O_2 \subseteq \mathbb{K}$  mit  $U_i = f(A) \cap O_i$  gibt. Wir betrachten nun die Teilmengen

$$V_i = f^{-1}(U_i) \cap A = f^{-1}(O_i \cap f(A)) \cap A = f^{-1}(O_i) \cap A, \quad (*)$$

wobei wir zunächst verwenden, dass  $f^{-1}(O_i \cap f(A)) = f^{-1}(O_i) \cap f^{-1}(f(A)) \supseteq f^{-1}(O_i) \cap A$  gilt. Im Allgemeinen ist  $f^{-1}(f(A)) \supseteq A$  ja eine echte Inklusion, aber der folgende Schnitt mit  $A$  zeigt dann die Gleichheit in (\*). Insbesondere sind die  $V_i \subseteq A$  offen in  $A$ , da  $f^{-1}(O_i)$  aufgrund der Stetigkeit von  $f$  offen ist. Weiter gilt  $V_i \neq \emptyset$ , da  $f(V_i) = U_i$  für  $i = 1, 2$ . Schließlich gilt

$$V_1 \cap V_2 = f^{-1}(U_1) \cap A \cap f^{-1}(U_2) \cap A = f^{-1}(U_1 \cap U_2) \cap A = \emptyset$$

ebenso wie

$$V_1 \cup V_2 = (f^{-1}(U_1) \cap A) \cup (f^{-1}(U_2) \cap A)$$

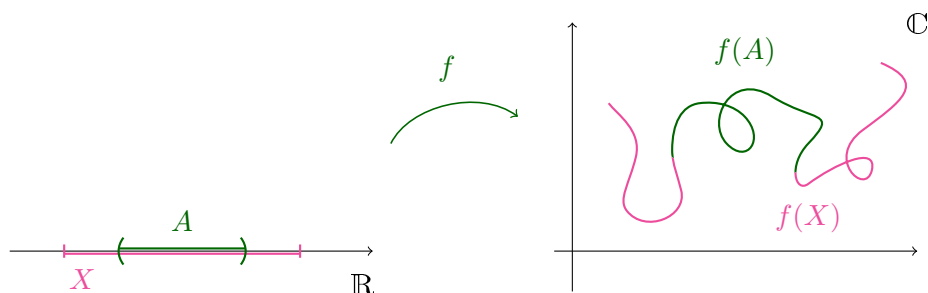


Abbildung 3.20: Das stetige Bild eines zusammenhängenden Intervalls ist wieder zusammenhängend.

$$\begin{aligned}
 &= (f^{-1}(U_1) \cup f^{-1}(U_2)) \cap A \\
 &= f^{-1}(U_1 \cup U_2) \cap A \\
 &= f^{-1}(A) \cap A \\
 &= A.
 \end{aligned}$$

Damit haben wir aber  $A$  durch zwei disjunkte nicht-leere offene Mengen überdeckt, was ein Widerspruch zum Zusammenhang von  $A$  darstellt.  $\square$

**Bemerkung 3.5.10** Dieser Satz trifft vielleicht am Besten die heuristische Vorstellung, dass man den Graphen einer stetigen Funktion ohne Abzusetzen zeichnen können soll, siehe auch Abbildung 3.20.

Eine der wichtigsten Konsequenzen aus Satz 3.5.9 ist der *Zwischenwertsatz*:

**Korollar 3.5.11 (Zwischenwertsatz)** Sei  $X \subseteq \mathbb{R}$  und  $f: X \rightarrow \mathbb{R}$  stetig. Sind  $a, b \in \mathbb{R}$  mit  $a < b$  und  $[a, b] \subseteq X$ , so nimmt  $f$  auf dem Intervall  $[a, b]$  jeden Wert zwischen  $f(a)$  und  $f(b)$  an.

BEWEIS: Das Intervall  $[a, b]$  ist nach Satz 3.4.12 zusammenhängend. Daher ist  $f([a, b]) \subseteq \mathbb{R}$  nach Satz 3.5.9 wieder zusammenhängend. Mit  $f(a), f(b) \in f([a, b])$  gilt nach Proposition 3.4.14 dann  $[f(a), f(b)] \subseteq f([a, b])$  falls  $f(a) \leq f(b)$  beziehungsweise  $[f(b), f(a)] \subseteq f([a, b])$  falls  $f(b) \leq f(a)$ . Dies ist gerade die Behauptung.  $\square$

**Korollar 3.5.12** Sei  $X \subseteq \mathbb{R}$  und  $f: X \rightarrow \mathbb{R}$  stetig. Gilt für  $a, b \in \mathbb{R}$  mit  $a < b$  und  $[a, b] \subseteq X$ , dass  $f(a)$  positiv und  $f(b)$  negativ ist, so besitzt  $f$  eine Nullstelle in  $(a, b)$ . Eine analoge Aussage gilt für  $f(a) < 0$  und  $f(b) > 0$ .

Dieses Korollar ist in vielen Fällen der Garant für die *Existenz* von Nullstellen. Eindeutigkeit der Nullstelle ist typischerweise nicht gegeben, siehe auch Abbildung 3.21. Eine erste wichtige Anwendung dieser Beobachtungen ist folgende Existenz von Nullstellen:

**Proposition 3.5.13** Sei  $p(x) = a_{2n+1}x^{2n+1} + \dots + a_1x + a_0 \in \mathbb{R}[x]$  mit  $a_{2n+1} \neq 0$  ein reelles Polynom ungerader Ordnung  $2n + 1 \geq 1$ . Dann hat  $p$  mindestens eine reelle Nullstelle.

BEWEIS: Die Idee ist natürlich, dass wir den Fundamentalsatz der Algebra *nicht* zum Beweis verwenden wollen. Ohne Einschränkung können wir  $a_{2n+1} = 1$  annehmen, da wir andernfalls  $p$  mit  $\frac{1}{a_{2n+1}}$  reskalieren können, ohne die Nullstellen zu ändern. Für  $x \neq 0$  gilt daher

$$p(x) = x^{2n+1} \left( 1 + \frac{a_{2n}}{x} + \dots + \frac{a_1}{x^{2n}} + \frac{a_0}{x^{2n+1}} \right).$$

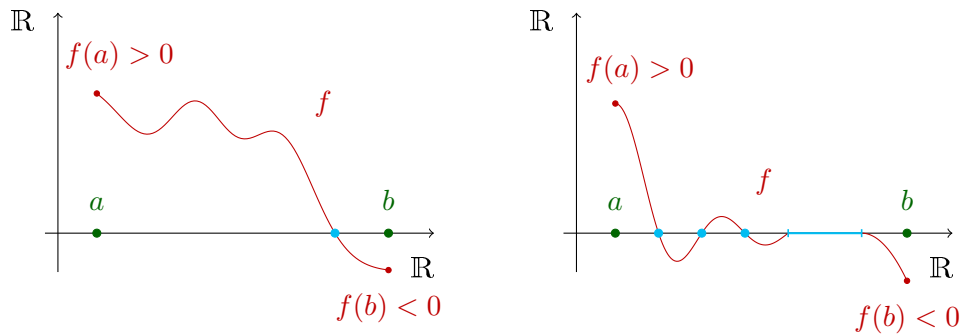


Abbildung 3.21: Korollar 3.5.12 liefert die Existenz aber im Allgemeinen nicht die Eindeutigkeit von Nullstellen.

Jeder der Summanden  $\frac{a_k}{x^{2n+1-k}}$  konvergiert sowohl für  $x \rightarrow +\infty$  als auch für  $x \rightarrow -\infty$  gegen Null. Damit konvergiert die Klammer aber gegen 1. Es gibt also insbesondere ein  $R > 0$  derart, dass

$$1 + \frac{a_{2n}}{x} + \cdots + \frac{a_1}{x^{2n}} + \frac{a_0}{x^{2n+1}} \geq \frac{1}{2} > 0,$$

sofern  $x \geq R$  oder  $x \leq -R$  gilt. Damit ist aber für solche  $x$  der Wert des Polynoms

$$p(x) = x^{2n+1} \left( 1 + \frac{a_{2n}}{x} + \cdots + \frac{a_1}{x^{2n}} + \frac{a_0}{x^{2n+1}} \right) \geq \frac{1}{2} R^{2n+1} > 0,$$

falls  $x \geq R$  beziehungsweise

$$p(x) = x^{2n+1} \left( 1 + \frac{a_{2n}}{x} + \cdots + \frac{a_1}{x^{2n}} + \frac{a_0}{x^{2n+1}} \right) \leq -\frac{1}{2} R^{2n+1} < 0$$

für  $x \leq -R$ . An dieser Stelle benutzen wir entscheidend, dass der Grad von  $p$  ungerade ist. Damit können wir Korollar 3.5.12 anwenden.  $\square$

Diese Beobachtung lässt sich nun in der Algebra nutzen, um einen algebraischen Beweis des Fundamentalsatzes der Algebra zu zeigen. Allein die Hinzunahme der Nullstellen  $\pm i \in \mathbb{C}$  des reellen Polynoms  $p(x) = x^2 + 1$  reicht aus, um einen algebraisch abgeschlossenen Körper zu erreichen, wenn man bereits weiß, dass die ungeraden reellen Polynome eine reelle Nullstelle besitzen. Dieser Zugang zum Fundamentalsatz hat den Charme, in einem komplett algebraischen Kontext gültig zu sein. Einfachere und direktere Beweise von Satz 1.7.8 erhält man dagegen in der Funktionentheorie, wenn man auch weiterhin bereit ist, analytische Techniken einzusetzen.

Der nächste Satz macht erneut Gebrauch von Kompaktheit: Einfache Beispiele zeigen, dass eine injektive stetige Abbildung  $f: X \rightarrow \mathbb{K}$  zwar eine auf ihrem Bild  $f(X) \subseteq \mathbb{K}$  definierte Umkehrabbildung  $f^{-1}: f(X) \rightarrow X$  besitzt, diese aber im Allgemeinen nicht länger stetig zu sein braucht. Ohne die genauen Details zu formulieren, liefert Abbildung 3.23 die nötige Intuition. Ein konkretes Beispiel werden wir erst in Bemerkung 4.5.7 geben können.

Das Problem mit der Situation aus Abbildung 3.23 wird gelöst, wenn man zudem einen kompakten Definitionsbereich fordert:

**Satz 3.5.14 (Stetigkeit der Umkehrabbildung)** *Sei  $X$  eine kompakte Teilmenge der reellen oder komplexen Zahlen und  $f: X \rightarrow \mathbb{K}$  eine stetige injektive Funktion. Damit ist die auf dem Bild  $f(X) \subseteq \mathbb{K}$  definierte Umkehrabbildung  $f^{-1}: f(X) \rightarrow X$  ebenfalls stetig.*

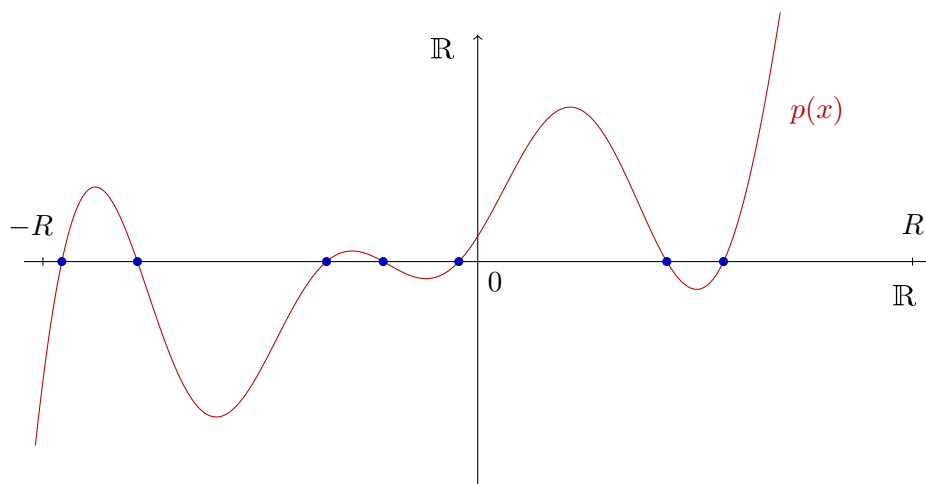


Abbildung 3.22: Ein reelles Polynom ungeraden Grades hat reelle Nullstellen.

BEWEIS: Sei  $A \subseteq X$  eine abgeschlossene Teilmenge in  $X$ . Um die Stetigkeit von  $g = f^{-1}: f(X) \rightarrow X$  zu zeigen, genügt es nach Satz 3.2.7, *iii.*), zu zeigen, dass das Urbild  $g^{-1}(A) \subseteq f(X)$  bezüglich der Umkehrabbildung abgeschlossen in  $f(X)$  ist. Es gilt nun

$$\begin{aligned} g^{-1}(A) &= \{y \in f(X) \mid g(y) \in A\} \\ &= \{y \in f(X) \mid f^{-1}(y) \in A\} \\ &= \{y \in f(X) \mid \exists a \in A \text{ mit } f(a) = y\} \\ &= f(A). \end{aligned}$$

Da  $A \subseteq X$  abgeschlossen in  $X$  ist und  $X$  kompakt ist, folgt nach Proposition 3.4.7, *iii.*), dass  $A$  selbst kompakt ist. Da  $f$  stetig ist, ist  $f(A)$  wieder kompakt gemäß Satz 3.5.4. Als kompakte Teilmenge  $f(A) \subseteq f(X)$  ist  $f(A)$  aber abgeschlossen in  $f(X)$ , siehe erneut Proposition 3.4.7, *ii.*). Also ist  $g^{-1}(A) = f(A)$  abgeschlossen, was die Stetigkeit von  $g = f^{-1}$  nach sich zieht.  $\square$

**Definition 3.5.15 (Homöomorphismus)** Seien  $X, Y$  zwei Teilmengen von reellen oder komplexen Zahlen.

i.) Eine stetige bijektive Abbildung

$$f: X \rightarrow Y \tag{3.5.13}$$

heißt Homöomorphismus, falls die Umkehrabbildung  $f^{-1}: Y \rightarrow X$  ebenfalls stetig ist.

ii.) Die Teilmengen  $X$  und  $Y$  heißen homöomorph, falls es einen Homöomorphismus  $f: X \rightarrow Y$  gibt.

Während injektive stetige Funktionen im Allgemeinen keine stetige Umkehrabbildung besitzen, siehe erneut das Beispiel aus Abbildung 3.25, so ist die Situation für reell-wertige Funktionen auf einem Intervall einfacher:

**Proposition 3.5.16** Sei  $I \subseteq \mathbb{R}$  ein Intervall und  $f: I \rightarrow \mathbb{R}$  eine stetige Funktion.

i.) Ist  $f$  injektiv, so ist  $f$  entweder strikt monoton wachsend oder strikt monoton fallend.

ii.) Ist  $f$  injektiv, so ist  $f(I) \subseteq \mathbb{R}$  ein Intervall vom gleichen Typ (offen, halboffen oder geschlossen) wie  $I$  und  $f^{-1}: f(I) \rightarrow I$  ist stetig.

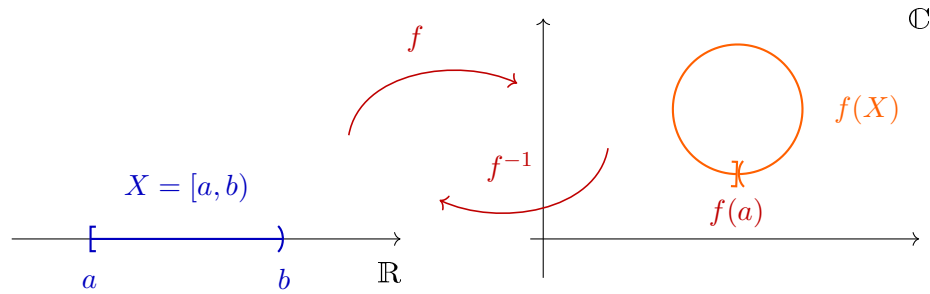


Abbildung 3.23: Die Abbildung  $f$  biegt das halboffene Intervall  $[a, b)$  zu einer Kreislinie in  $\mathbb{C}$ . Diese ist abgeschlossen und beschränkt, also kompakt. Während  $f$  sicherlich stetig ist, kann dies  $f^{-1}$  nicht sein, da ein stetiges Bild der kompakten Kreislinie selbst kompakt wäre,  $[a, b)$  dies aber nicht ist.

BEWEIS: Wir betrachten zunächst den Fall eines abgeschlossenen Intervalls  $I = [a, b]$ . Ist  $I = \{a\}$  nur ein Punkt, so ist  $i.)$  aus trivialen Gründen erfüllt. Daher nehmen wir ohne Einschränkungen  $a < b$  an. Da  $f$  injektiv ist, gilt  $f(a) \neq f(b)$ . Wir betrachten den Fall  $f(a) < f(b)$ , den anderen Fall  $f(a) > f(b)$  erhält man analog. Zunächst bemerken wir, dass  $a$  das Minimum von  $f$  mit Wert  $f(a)$  sein muss: Wäre nämlich  $c \in (a, b]$  das Minimum mit Wert  $f(c) < f(a)$ , so gäbe es im Intervall  $[c, b]$  einen weiteren Punkt  $x \in [c, b]$  mit Funktionswert  $f(x) = f(a)$  nach dem Zwischenwertsatz aus Korollar 3.5.11. Da  $x \neq a$  gilt, kann dies aber aufgrund der Injektivität nicht sein. Wir behaupten, dass  $f$  strikt monoton wachsend ist. Wäre dies nicht der Fall, so gäbe es zwei Punkte  $x_1 < x_2$  in  $[a, b]$  mit  $f(x_2) \leq f(x_1)$ . Es gilt in diesem Fall also

$$f(a) \leq f(x_2) \leq f(x_1),$$

da  $a$  das Minimum von  $f$  ist. Nach dem Zwischenwertsatz gibt es also einen weiteren Punkt  $a \leq z \leq x_1$  mit  $f(z) = f(x_2)$ . Da  $f$  injektiv ist, müsste  $z = x_2$  gelten, was wegen  $x_2 \notin [a, x_1]$  nicht möglich ist. Ist stattdessen  $I = (a, b]$  oder  $[a, b)$  oder  $(a, b)$  ein halboffenes oder offenes Intervall, so ist für je zwei Punkte  $c, d \in I$  mit  $c < d$  die Einschränkung  $f|_{[c, d]}$  immer noch injektiv und stetig, also strikt monoton wachsend oder strikt monoton fallend. Es kann offenbar nur einer der beiden Fälle eintreten, der dann das Verhalten von  $f$  für alle solche Einschränkungen bestimmt. Insgesamt ist  $f$  damit aber strikt monoton, was den ersten Teil zeigt. Bei einer strikt monotonen Funktion kann das Minimum beziehungsweise das Maximum nur an einem Randpunkt von  $I$  angenommen werden. Ist  $I$  also abgeschlossen, so hat  $f$  ein Minimum und ein Maximum, im Fall  $I = (a, b)$  oder  $I = [a, b)$  gibt es entweder ein Minimum oder ein Maximum, je nachdem ob  $f$  wachsend oder fallend ist. Im Falle eines offenen Intervalls  $I = (a, b)$  gibt es weder Minimum noch Maximum. Da in allen Fällen das Bild wieder zusammenhängend ist, siehe nochmal Satz 3.5.9, die einzigen zusammenhängenden Teilmengen von  $\mathbb{R}$  aber Intervalle sind, wie wir anhand von Proposition 3.4.14 und Übung 3.9.19 sehen, folgt, dass  $J = f(I) \subseteq \mathbb{R}$  wieder ein Intervall ist. Nach der vorherigen Definition zur Existenz von Minimum und Maximum ist  $J$  genau dann offen (halboffen, abgeschlossen), wenn  $I$  offen (halboffen, abgeschlossen) ist. Es bleibt die Stetigkeit von  $f^{-1}$  zu zeigen. Ist  $I = [a, b]$  ein abgeschlossenes und daher kompaktes Intervall, so wissen wir die Stetigkeit von  $f^{-1}$  bereits nach Satz 3.5.14. Sei also zunächst  $I = [a, b)$  halboffen. Ist  $f$  beispielsweise wachsend, so ist  $J = [f(a), f(b))$ . Ist  $y \in (f(a), f(b))$ , so gibt es  $c, d \in I$  mit  $f(c) < y < f(d)$ . Die Einschränkung  $f|_{[c, d]}$  bildet daher  $[c, d]$  bijektiv auf  $[f(c), f(d)]$  ab und ist nach wie vor stetig. Nach Satz 3.5.14 ist  $f^{-1}|_{[f(c), f(d)]}$  stetig und somit auch  $f^{-1}|_{(f(c), f(d))}$ . Die Stetigkeit von  $f^{-1}$  auf einer offenen Umgebung  $(f(c), f(d)) \subseteq J$  von  $y$  zeigt aber die Stetigkeit von  $f^{-1}$  bei  $y$ . Ist schließlich  $y = f(a) \in J$ , so gibt es ein  $c \in I$  mit  $f(a) < f(c)$ . Weiter ist  $[a, c]$  kompakt und  $f|_{[a, c]}$  hat eine stetige Umkehrfunktion  $(f|_{[a, c]})^{-1} = f^{-1}|_{[f(a), f(c)]}$ , da  $f([a, c]) = [f(a), f(c)]$ . Nun ist  $[f(a), f(c)] \subseteq J$  eine offene Umgebung von  $f(a) \in J$  in  $J$ , womit die Stetigkeit von  $f^{-1}$  auf

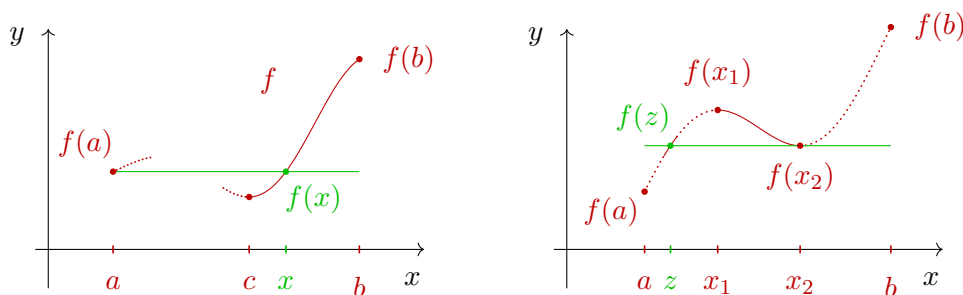


Abbildung 3.24: Eine injektive stetige Funktion  $f: [a, b] \rightarrow \mathbb{R}$  hat ihr Minimum nicht im Inneren  $(a, b)$ . Gilt  $f(x_1) > f(x_2)$  für  $x_1 < x_2$ , so gibt es einen Zwischenpunkt  $a < z < x_1$  mit  $f(z) = f(x_2)$ .

$[f(a), f(c)] \subseteq [f(a), f(c)]$  wieder die Stetigkeit von  $f^{-1}$  bei  $f(a)$  impliziert. Damit ist  $f^{-1}$  überall auf  $J$  stetig. Der Fall einer fallenden Funktion verläuft genauso und ebenso der Fall eines halboffenen Intervalls der Form  $(a, b]$ . Ist schließlich  $(a, b)$  offen, so könne wir wie zuvor für den inneren Punkt  $y \in J$  im halboffenen Fall argumentieren.  $\square$

**Korollar 3.5.17** Seien  $I, J \subseteq \mathbb{R}$  Intervalle. Dann sind  $I$  und  $J$  genau dann homöomorph, wenn beide offen oder beide halboffen oder beide abgeschlossen sind.

BEWEIS: Ein Homöomorphismus  $f: I \rightarrow J$  ist insbesondere stetig und bijektiv, womit Proposition 3.5.16 zum Einsatz kommt und zeigt, dass die Bedingungen an  $I$  und  $J$  notwendig sind. Dass diese hinreichend sind, erfordert nun die Angabe von expliziten Homöomorphismen. Diese werden in Übung 3.9.29 konstruiert.  $\square$

Der letzte wichtige Satz zu stetigen Funktionen in diesem Abschnitt handelt von *stetigen Fortsetzungen*. Oftmals hat man eine einfache Definition einer stetigen Funktion auf einem kleinen Definitionsbereich  $X$ , den man allerdings vergrößern will. Es stellt sich also die Frage, ob man eine andere Funktion  $F$  auf einem größeren Definitionsbereich  $\tilde{X} \supseteq X$  findet, so dass

$$F|_X = f \quad (3.5.14)$$

gilt. Ohne weitere Anforderungen an  $F$  ist dies leicht zu bewerkstelligen, man setzt beispielsweise

$$F(x) = \begin{cases} f(x) & \text{für } x \in X \\ 0 & \text{für } x \in \tilde{X} \setminus X. \end{cases} \quad (3.5.15)$$

Mit solch einer naiven Fortsetzung verliert  $F$  im allgemeinen Eigenschaften, die  $f$  eventuell hatte, wie eben Stetigkeit. Die Frage ist also, ob man eine *stetige* Fortsetzung finden kann. Einfache Beispiele legen nahe, dass dies aus verschiedenen Gründen scheitern kann, siehe Abbildung 3.25. Weiter kann es auch viele stetige Fortsetzungen geben.

**Beispiel 3.5.18** Wir betrachten die rationale Funktion

$$f(x) = \frac{x^2 - 1}{x - 1}, \quad (3.5.16)$$

deren natürlicher Definitionsbereich ganz  $\mathbb{R}$  ohne die Nullstellen des Nenners ist, also

$$X = \mathbb{R} \setminus \{1\}, \quad (3.5.17)$$

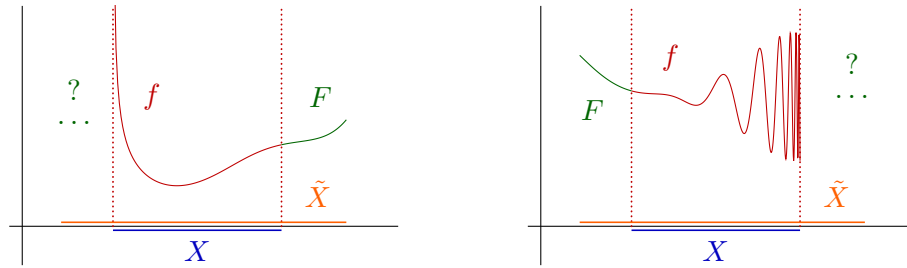


Abbildung 3.25: Für die Funktion auf der linken Seite gibt es stetige Fortsetzungen nach rechts, aber nicht nach links, da  $f$  dort unbeschränkt wird. Im rechten Bild ist  $f$  zwar auf ganz  $X$  beschränkt, oszilliert aber zu sehr, als dass eine stetige Fortsetzung nach rechts möglich wäre. Konkrete Funktionen mit diesem Verhalten werden wir in Beispiel 5.1.3, *iv.*), noch kennenlernen.

siehe nochmals Beispiel 3.1.12. Dort ist  $f$  sicherlich stetig. Da für  $x \neq 1$  aber

$$f(x) = \frac{x^2 - 1}{x - 1} = \frac{(x + 1)(x - 1)}{x - 1} = x + 1 \quad (3.5.18)$$

gilt, können wir  $f$  auf ganz  $\mathbb{R}$  stetig fortsetzen, indem wir

$$F(x) = \begin{cases} f(x) & \text{für } x \neq 1 \\ 2 & \text{für } x = 1 \end{cases} \quad (3.5.19)$$

verwenden, siehe auch Abbildung 3.26. In diesem Beispiel ist  $F$  offenbar auch die einzige stetige Fortsetzung. Jede andere Festlegung des Funktionswertes bei  $x = 1$  führt zu einer unstetigen Fortsetzung.

Die Existenz stetiger Fortsetzungen ist im Allgemeinen ein schwieriges Unterfangen, welches Detailwissen sowohl der geometrischen Eigenschaften des Definitionsbereichs  $X$  als auch der Funktion  $f$  auf  $X$  erfordert. Bei der Eindeutigkeit gibt es eine vergleichsweise einfache Antwort, wenn man eine Fortsetzung auf den Abschluss  $X^{\text{cl}}$  des Definitionsbereichs  $X$  sucht:

**Satz 3.5.19 (Stetige Fortsetzung)** Sei  $X \subseteq \mathbb{K}$  und  $f$  eine stetige Funktion auf  $X$ . Gibt es dann eine stetige Fortsetzung  $F$  von  $f$  auf  $X^{\text{cl}}$ , so ist diese eindeutig bestimmt.

BEWEIS: Der Abschluss  $X^{\text{cl}}$  von  $X$  besteht aus den isolierten Punkten von  $X$  sowie den Häufungspunkten von  $X$ . Da isolierte Punkte bereits zu  $X$  gehören, müssen wir  $f$  hier nicht fortsetzen. Sei also  $x \in X^{\text{cl}}$  ein Häufungspunkt von  $X$ , der nicht zu  $X$  gehört. Dann gibt es eine Folge  $(x_n)_{n \in \mathbb{N}}$  von Punkten  $x_n \in X$  mit  $x_n \rightarrow x$ . Ist  $F$  nun eine stetige Fortsetzung von  $f$ , so gilt

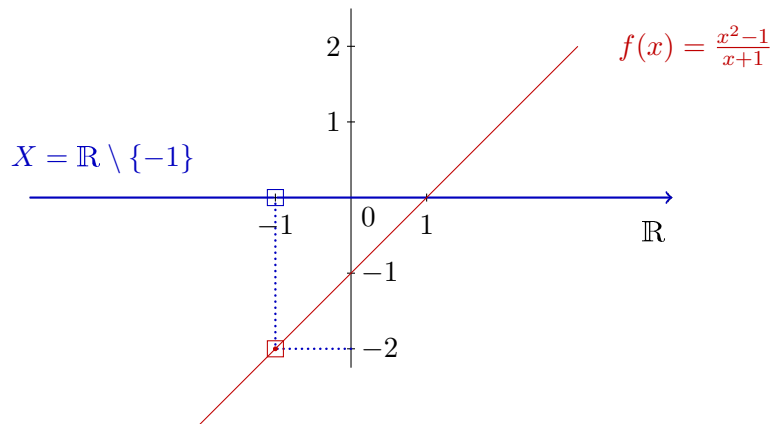
$$F(x) = \lim_{n \rightarrow \infty} F(x_n) = \lim_{n \rightarrow \infty} f(x_n),$$

womit  $F(x)$  bereits eindeutig festliegt. □

**Bemerkung 3.5.20 (Exponentialfunktion)** Sei  $a > 0$ . Aus Beispiel 3.3.5 wissen wir, dass die Exponentialfunktion

$$f(x) = a^x \quad (3.5.20)$$

für  $x \in \mathbb{R}$  stetig ist. Die Definition von  $a^x$  war aber recht kompliziert: Zuerst mussten wir unter Verwendung der Vollständigkeit Wurzeln  $\sqrt[n]{a} = a^{\frac{1}{n}}$  für  $n \in \mathbb{N}$  definieren, siehe Definition 1.6.2, und dann die algebraischen Eigenschaften der Wurzeln nutzen, um rationale Potenzen  $a^q$  für  $q \in \mathbb{Q}$

Abbildung 3.26: Stetige Fortsetzung von  $f(x) = \frac{x^2-1}{x+1}$ .

zu definieren. Dieser zweite Schritt war vergleichsweise einfach, kam er doch ohne die Verwendung von Suprema und damit ohne die Verwendung der Vollständigkeit von  $\mathbb{R}$  aus. Der letzte Schritt, die Ausdehnung auf allgemeine reelle Exponenten  $x \in \mathbb{R}$  war wieder komplizierter, da erneut die Vollständigkeit von  $\mathbb{R}$  verwendet werden musste, siehe Definition 1.6.7. Im Hinblick auf Satz 3.5.19 hatten wir hierbei aber keine Wahl, wenn wir eine *stetige* Exponentialfunktion wollen: Definition 1.6.7 lässt sich als Fortsetzung der Exponentialfunktion

$$f(q) = a^q \quad \text{für } q \in \mathbb{Q} \quad (3.5.21)$$

verstehen, die nach Beispiel 3.3.5 sogar eine *stetige Fortsetzung* ist. Da nach Übung 3.9.5

$$\mathbb{Q}^{\text{cl}} = \mathbb{R} \quad (3.5.22)$$

gilt, ist eine solche Fortsetzung nach Satz 3.5.19, sofern sie überhaupt existiert, eindeutig. Die Schwierigkeiten bei der Konstruktion von  $a^x$  in Definition 1.6.7 und den zugehörigen Eigenschaften aus Satz 1.6.8 können also so verstanden werden, dass wir die Stetigkeit der Fortsetzung gezeigt haben.

## 3.6 Gleichmäßige Stetigkeit

Verwenden wir beispielsweise die  $\epsilon\delta$ -Stetigkeit bei einem Punkt  $x_0$ , so hängt die Wahl von  $\delta > 0$  passend zum vorgegebenen  $\epsilon > 0$  im Allgemeinen von  $x_0$  ab. In der Charakterisierung von Stetigkeit bei einem Punkt  $x_0$  durch Umgebungen ist dies noch klarer: Die Umgebung  $f^{-1}(U)$  zu einer Umgebung  $U$  von  $f(x_0)$  ist eben eine Umgebung von  $x_0$  und daher natürlich von  $x_0$  abhängig. Die reellen wie auch die komplexen Zahlen besitzen nun aber die besondere Eigenschaft, dass Umgebungen von Punkten überall „gleich“ aussehen. Dies wird in folgender Beobachtung präzise formuliert:

**Proposition 3.6.1** Sei  $x, y \in \mathbb{K}$ .

i.) Die Translation  $\tau_x: \mathbb{K} \ni y \mapsto \tau_x(y) = x + y \in \mathbb{K}$  um  $x$  ist ein Homöomorphismus, und es gilt

$$\tau_x \circ \tau_y = \tau_{x+y} \quad \text{sowie} \quad \tau_0 = \text{id}_{\mathbb{K}}. \quad (3.6.1)$$

ii.) Ist  $r > 0$ , so gilt

$$\tau_x(B_r(y)) = B_r(\tau_x(y)). \quad (3.6.2)$$

iii.) Eine Teilmenge  $U \subseteq \mathbb{K}$  ist genau dann eine Umgebung von  $x$ , wenn  $\tau_x(U)$  eine Umgebung von  $\tau_x(y)$  ist.



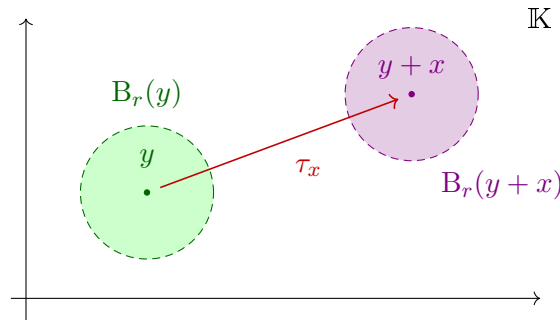


Abbildung 3.27: Translationen von offenen Kugeln bleiben offene Kugeln mit gleichem Radius.

BEWEIS: Die Rechenregel (3.6.1) ist eine direkte Folge der Assoziativität von  $+$ . Die Stetigkeit von  $\tau_x$  ist ein Spezialfall der Stetigkeit der Polynome aus Beispiel 3.3.1. Da  $\tau_x$  bijektiv ist mit Umkehrabbildung  $(\tau_x)^{-1} = \tau_{-x}$ , ist auch die Umkehrabbildung wieder stetig, so dass insgesamt  $\tau_x$  ein Homöomorphismus ist. Die Behauptung (3.6.2) ist klar, da

$$|\tau_x(z) - \tau_x(y)| = |z + x - (y + x)| = |z - y|$$

gilt. Demnach ist  $\tau_x(z) \in B_r(\tau_x(y))$  genau dann, wenn  $z \in B_r(y)$  gilt. Dies ist aber (3.6.2). Der letzte Teil ist eine allgemeine Eigenschaft von Homöomorphismen. Da  $\tau_x$  stetig ist, ist für jede Umgebung  $V$  von  $\tau_x(y)$  auch  $\tau_x^{-1}(V) = \tau_{-x}(V)$  eine Umgebung von  $y$ . Vertauscht man die Rollen von  $y$  und  $\tau_x(y)$  sowie  $x$  und  $-x$ , so erhält man auch die Umkehrung.  $\square$

Wir können also Umgebungen beziehungsweise offene Kugeln mit den Translationen an verschiedenen Punkten vergleichen. Dies erlaubt es nun die Güte der Stetigkeitsabschätzung an verschiedenen Punkten zu vergleichen, was zur Definition der gleichmäßigen Stetigkeit führt:

**Definition 3.6.2 (Gleichmäßige Stetigkeit)** Eine Funktion  $f: X \rightarrow \mathbb{K}$  heißt *gleichmäßig stetig* (auch: *uniform stetig*) auf  $X$ , wenn es zu jedem  $\epsilon > 0$  ein  $\delta > 0$  gibt, so dass für alle  $x_0$

$$y \in B_\delta(x_0) \cap X \implies |f(y) - f(x_0)| < \epsilon \quad (3.6.3)$$

gilt.

Die Idee der gleichmäßigen Stetigkeit ist es also, anders als bei Stetigkeit alleine, dass die Wahl von  $\delta > 0$  *universell* für alle Punkte  $x_0 \in X$  getroffen werden kann. Eine gleichmäßig stetige Funktion ist insbesondere stetig.

Die obige Formulierung der gleichmäßigen Stetigkeit stellt den Begriff der Umgebung in den Vordergrund. Alternativ erhält man folgende Version: Eine Funktion  $f: X \rightarrow \mathbb{K}$  ist gleichmäßig stetig auf  $X$ , falls es zu jedem  $\epsilon > 0$  ein  $\delta > 0$  gibt, so dass für alle  $x, y \in X$

$$|x - y| < \delta \implies |f(x) - f(y)| < \epsilon. \quad (3.6.4)$$

Dies behandelt die beiden Punkte in symmetrischer Weise und zeigt, dass es sich bei gleichmäßiger Stetigkeit um ein Konzept handelt, welches *Paare von Punkten* involviert. Dieser Standpunkt dient dann in der Theorie der uniformen Räume als Ausgangspunkt.

Wir diskutieren nun einige Beispiele:

### Beispiel 3.6.3 (Gleichmäßige Stetigkeit)

i.) Die Funktion

$$f(x) = x^2 \quad (3.6.5)$$

ist nicht gleichmäßig stetig auf  $\mathbb{R}$ . Wäre dies der Fall, so gäbe es zu  $\epsilon = 1$  ein  $\delta > 0$  mit  $|x - y| < \delta \implies |f(x) - f(y)| < 1$  für alle  $x, y \in \mathbb{R}$ . Wir nehmen an, ein solches  $\delta$  gefunden zu haben. Dann betrachten wir  $x = \frac{1}{\delta}$  und  $y = \frac{1}{\delta} + \frac{\delta}{2}$ , womit  $|x - y| = \frac{\delta}{2} < \delta$  erfüllt ist. Es gilt aber

$$|x^2 - y^2| = \left| \frac{1}{\delta^2} - \left( \frac{1}{\delta} + \frac{\delta}{2} \right)^2 \right| = \left| \frac{1}{\delta^2} - \frac{1}{\delta^2} - 1 - \frac{\delta^2}{4} \right| = 1 + \frac{\delta^2}{4} > 1,$$

was einen Widerspruch darstellt. Insbesondere gibt es damit stetige Funktionen, die nicht gleichmäßig stetig sind.

ii.) Ist  $f(x) = ax + b$  mit  $a, b \in \mathbb{C}$ , so ist  $f$  auf  $\mathbb{C}$  gleichmäßig stetig: Wir können zu  $\epsilon > 0$  beispielsweise  $\delta = \frac{\epsilon}{|a|}$  wählen. Dann folgt für  $x, y \in \mathbb{C}$  mit  $|x - y| < \frac{\epsilon}{|a|}$  nämlich

$$|f(x) - f(y)| = |ax + b - ay - b| = |a||x - y| < \epsilon$$

wie gewünscht. Insbesondere sind alle konstanten Funktionen gleichmäßig stetig.

iii.) Sei

$$f(x) = \sqrt{x} \quad (3.6.6)$$

für  $x \in \mathbb{R}_0^+$ . Dann ist  $f$  gleichmäßig stetig auf  $\mathbb{R}_0^+$ . Zu  $\epsilon > 0$  wählen wir nämlich  $\delta = \epsilon^2$  und erhalten für  $x > y \geq 0$  die Abschätzung

$$\begin{aligned} (\sqrt{x} - \sqrt{y})^2 &= x - 2\sqrt{x}\sqrt{y} + y \\ &\leq x - 2\sqrt{y}\sqrt{y} + y \\ &= x - y \\ &= |x - y|, \end{aligned}$$

womit für  $|x - y| < \delta = \epsilon^2$  also  $|\sqrt{x} - \sqrt{y}| < \epsilon$  folgt. Man beachte, dass ohne Einschränkung  $x \geq y$  angenommen werden darf.

Es gibt also durchaus interessante gleichmäßig stetige Funktionen. Gute algebraische Eigenschaften sind aber anders als für stetige Funktionen, nicht zu erwarten, wie das Beispiel 3.6.3, i.) zeigt: Die gleichmäßig stetigen Funktionen sind nicht unter Produktbildung abgeschlossen. Immerhin bilden sie noch einen Vektorraum:

**Proposition 3.6.4** *Sei  $X$  eine nicht-leere Teilmenge der komplexen oder reellen Zahlen. Dann bilden die gleichmäßig stetigen  $\mathbb{K}$ -wertigen Funktionen einen  $\mathbb{K}$ -Vektorraum bezüglich der punktweisen Operationen.*

BEWEIS: Wir müssen nur zeigen, dass sie einen Unterraum aller Funktionen bilden. Zunächst ist klar, dass die Nullfunktion gleichmäßig stetig ist. Sei nun  $f, g: X \rightarrow \mathbb{K}$  gleichmäßig stetig und  $z, w \in \mathbb{K}$ . Zu  $\epsilon > 0$  finden wir also  $\delta > 0$  und  $\delta' > 0$  mit

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{2(|z| + 1)}$$

sowie

$$|x - y| < \delta' \implies |g(x) - g(y)| < \frac{\epsilon}{2(|w| + 1)}$$

für  $x, y \in X$ . Ohne Einschränkung können wir  $\delta = \delta'$  annehmen, indem wir gegebenenfalls zum Minimum übergehen. Dann gilt für  $x, y \in X$  mit  $|x - y| < \delta$

$$|(zf + wg)(x) - (zf + wg)(y)| \leq |z||f(x) - f(y)| + |w||g(x) - g(y)|$$

$$\begin{aligned}
&< \frac{|z|}{2(|z|+1)}\epsilon + \frac{|w|}{2(|w|+1)}\epsilon \\
&< \epsilon,
\end{aligned}$$

was die gleichmäßige Stetigkeit von  $zf + wg$  zeigt.  $\square$

Eine Algebra wie  $\mathcal{C}(X)$  bilden die gleichmäßig stetigen Funktionen dagegen nicht, wie Beispiel 3.6.3, *i.*) zeigt. Die gleichmäßige Stetigkeit ist anders als die *lokale* Eigenschaft der Stetigkeit eine intrinsisch *globale* Eigenschaft der Funktion: Die Beobachtung des ganzen Definitionsbereichs ist entscheidend, siehe auch Übung 3.9.13 als Kontrast zur normalen Stetigkeit. Hätten wir in Beispiel 3.6.3, *i.*), die Größe der Argumente beschränken dürfen, wäre der Nachweis der gleichmäßigen Stetigkeit nicht gescheitert. Es war entscheidend, große  $x$  und  $y$  verwenden zu dürfen. Dies ist tatsächlich ein allgemeines Phänomen, wie der nächste Satz zeigt:

**Satz 3.6.5 (Gleichmäßige Stetigkeit auf Kompakta)** *Sei  $X$  eine kompakte Teilmenge von reellen oder komplexen Zahlen. Dann ist jede stetige Funktion auf  $X$  auch gleichmäßig stetig.*

BEWEIS: Sei  $f: X \rightarrow \mathbb{K}$  stetig und  $\epsilon > 0$  vorgegeben. Zu  $x \in X$  finden wir dann ein  $\delta_x > 0$  derart, dass

$$y \in B_{\delta_x}(x) \cap X \implies |f(x) - f(y)| < \frac{\epsilon}{2} \quad (*)$$

gilt. Dies ist gerade die Stetigkeit bei  $x$ . Man beachte, dass  $\delta_x$  eben von  $x$  abhängt. Die Gesamtheit aller offenen Kugeln  $\{B_{\frac{\delta_x}{2}}(x)\}_{x \in X}$  bilden nun eine offene Überdeckung von  $X$ , da ja jedes  $x \in X$  zumindest in  $B_{\frac{\delta_x}{2}}(x)$  enthalten ist. Da  $X$  kompakt ist, finden wir eine endliche Teilüberdeckung, die wir mit  $B_{\frac{\delta_{x_1}}{2}}(x_1), \dots, B_{\frac{\delta_{x_n}}{2}}(x_n)$  bezeichnen. Sei nun

$$\delta = \min\{\delta_{x_1}, \dots, \delta_{x_n}\} > 0. \quad (**)$$

Sind dann  $x, y \in X$  mit  $|x - y| < \frac{\delta}{2}$  vorgegeben, so gibt es mindestens einen Index  $i \in \{1, \dots, n\}$  mit  $x \in B_{\frac{\delta_{x_i}}{2}}(x_i)$ , da diese offenen Kugeln ja bereits eine Überdeckung bilden. Dann gilt also  $|x - x_i| < \frac{\delta_{x_i}}{2}$ . Weiter gilt für  $y$

$$|y - x_i| \leq |y - x| + |x - x_i| < \frac{\delta}{2} + \frac{\delta_{x_i}}{2} \leq \delta_{x_i},$$

womit also  $x, y \in B_{\delta_{x_i}}(x_i)$  gezeigt ist. Dann gilt aber

$$\begin{aligned}
|f(x) - f(y)| &= |f(x) - f(x_i) + f(x_i) - f(y)| \\
&\leq |f(x) - f(x_i)| + |f(x_i) - f(y)| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon
\end{aligned}$$

nach Wahl von  $\delta_{x_i}$  gemäß (\*). Dies zeigt die gleichmäßige Stetigkeit, da  $\delta$  aus (\*\*) gleichermaßen für alle Punktpaare  $x, y \in X$  seinen Dienst tut.  $\square$

Dieser Satz wird uns gelegentlich wieder begegnen, insbesondere wenn wir Riemann-Integrale in Kapitel 6 kennenlernen werden.

### 3.7 Gleichmäßige Konvergenz

Bislang haben wir Folgen von reellen und komplexen Zahlen auf deren Konvergenzverhalten geprüft. Das soll sich nun drastisch ändern, da wir nicht einzelne Zahlen, sondern ganze Funktionen betrachten wollen. Dies kann man so verstehen, dass wir die Konvergenzeigenschaften von einer Folge  $(f_n(x))_{n \in \mathbb{N}}$  verstehen wollen, die ihrerseits eine Funktion einer Variablen  $x \in X$  ist. Es geht also darum Funktionenfolgen und deren Konvergenzverhalten zu studieren.

Wir haben also eine Folge  $(f_n)_{n \in \mathbb{N}}$  von Funktionen  $f_n \in \text{Abb}(X, \mathbb{K})$  vorliegen, für die wir einen Konvergenzbegriff benötigen. Überraschenderweise gibt es hier viele Optionen, von denen wir hier zwei kennen lernen wollen.

Eine wesentliche Motivation, Funktionenfolgen und deren Konvergenz genauer zu untersuchen, ist ein Wunsch aus der Anwendung: Komplizierte aber relevante Funktionen sollen durch einfachere beschrieben werden, wobei man zwar einen Fehler begeht, dessen Größe man aber durch geeignete Abschätzungen kontrollieren möchte. Einfache Funktionen können dabei beispielsweise stückweise konstante Funktionen oder Polynome sein. Neben diesen ganz praktischen Gründen gibt es natürlich auch eine Vielzahl von konzeptuellen Überlegungen, die das Studium von Funktionenfolgen erfordern.

Wir geben ein erstes Beispiel, wie man die Wurzelfunktion durch Polynome approximieren kann:

**Proposition 3.7.1** Sei  $p_0(x) = 0$  und  $p_n \in \mathbb{R}[x]$  rekursiv durch

$$p_{n+1}(x) = p_n(x) + \frac{1}{2}(x - p_n^2(x)) \quad (3.7.1)$$

definiert.

i.) Es gilt  $p_n(0) = 0$  und  $p_n \in \mathbb{Q}[x]$  für alle  $n \in \mathbb{N}$ .

ii.) Für  $x \in [0, 1]$  gilt für alle  $n \in \mathbb{N}$

$$0 \leq \sqrt{x} - p_n(x) \leq \frac{2\sqrt{x}}{2 + n\sqrt{x}} < \frac{2}{n}. \quad (3.7.2)$$

iii.) Für alle  $x \in [0, 1]$  gilt

$$\lim_{n \rightarrow \infty} p_n(x) = \sqrt{x}. \quad (3.7.3)$$

BEWEIS: Da  $p_0(x) = 0$  bei  $x = 0$  verschwindet, sieht man durch Induktion nach  $n$  sofort, dass  $p_n(x) = 0$  für alle  $n \in \mathbb{N}$  gilt. Die Rekursionsformel enthält nur rationale Parameter, so dass auch  $p_n \in \mathbb{Q}[x]$  durch Induktion folgt. Für den zweiten Teil zeigen wir zuerst die etwas schwächere Abschätzung

$$0 \leq p_n(x) \leq 1 \quad (*)$$

für alle  $x \in [0, 1]$  und  $n \in \mathbb{N}$ . Wieder ist  $(*)$  für  $n = 0$  offensichtlich. Wir nehmen daher an,  $(*)$  gilt für  $p_n$  und betrachten  $p_{n+1}$ . Dazu bemerken wir zunächst, dass

$$\frac{1}{2}\xi^2 - \xi + \frac{1}{2} = \frac{1}{2}(\xi - 1)^2 \geq 0$$

für alle  $\xi \in \mathbb{R}$  gilt. Daher gilt  $\xi - \frac{1}{2}\xi^2 \leq \frac{1}{2}$ . Andererseits gilt für  $\xi \in [0, 1]$  zunächst  $\xi^2 \leq \xi$  und daher

$$\xi - \frac{1}{2}\xi^2 \geq \xi - \frac{1}{2}\xi \geq \frac{1}{2}\xi \geq 0,$$

so dass also insgesamt

$$0 \leq \xi - \frac{1}{2}\xi^2 \leq \frac{1}{2}$$

für alle  $\xi \in [0, 1]$ . Nach Induktionsvoraussetzung  $(*)$  für  $p_n$  erhalten wir also  $0 \leq p_n(x) - \frac{1}{2}p_n^2(x) \leq \frac{1}{2}$  für alle  $x \in [0, 1]$ . Damit gilt zum einen

$$p_{n+1}(x) = p_n(x) + \frac{1}{2}x - \frac{1}{2}p_n^2(x) \geq \frac{1}{2}x \geq 0$$

und zum anderen

$$p_{n+1}(x) = p_n(x) + \frac{1}{2}x - \frac{1}{2}p_n^2(x) \leq \frac{1}{2}x + \frac{1}{2} \leq 1$$

für  $x \in [0, 1]$ , was (\*) durch Induktion beweist. Wir wollen nun die verbesserte Ungleichung (3.7.2) ebenfalls induktiv zeigen. Der Induktionsanfang ist wieder einfach einzusehen. Es gelte daher (3.7.2) für  $p_n$ . Dann gilt für  $p_{n+1}$  zum einen

$$\begin{aligned} \sqrt{x} - p_{n+1}(x) &= \sqrt{x} - p_n(x) - \frac{1}{2}(x - p_n^2(x)) \\ &= \sqrt{x} - p_n(x) - \frac{1}{2}(\sqrt{x} - p_n(x))(\sqrt{x} + p_n(x)) \\ &= (\sqrt{x} - p_n(x)) - \underbrace{\left(1 - \frac{1}{2}(\sqrt{x} + p_n(x))\right)}_{\leq 2} \\ &\geq 0, \end{aligned}$$

da  $\sqrt{x} - p_n(x) \geq 0$  und  $p_n(x) \leq 1$  sowie  $x \in [0, 1]$  und damit auch  $\sqrt{x} \in [0, 1]$ . Zum anderen gilt für  $x \in [0, 1]$

$$\begin{aligned} \sqrt{x} - p_{n+1}(x) &= (\sqrt{x} - p_n(x))\left(1 - \frac{1}{2}\sqrt{x} + \frac{1}{2}p_n(x)\right) \\ &\leq \frac{2\sqrt{x}}{2 + n\sqrt{x}}\left(1 - \frac{1}{2}\sqrt{x}\right) \\ &= \frac{2\sqrt{x} - x}{2 + n\sqrt{x}} \\ &= \frac{2\sqrt{x}}{2 + (n+1)\sqrt{x}} - \frac{2\sqrt{x}}{2 + (n+1)\sqrt{x}} + \frac{2\sqrt{x} - x}{2 + n\sqrt{x}}. \end{aligned} \quad (3.7.4)$$

Nun gilt

$$\begin{aligned} -\frac{2\sqrt{x}}{2 + (n+1)\sqrt{x}} + \frac{2\sqrt{x} - x}{2 + n\sqrt{x}} &= \frac{-2\sqrt{x}(2 + n\sqrt{x}) + (2\sqrt{x} - x)(2 + (n+1)\sqrt{x})}{(2 + (n+1)\sqrt{x})(2 + n\sqrt{x})} \\ &= \frac{-4\sqrt{x} - 2nx + 4\sqrt{x} + 2(n+1)x - 2x - (n+1)\sqrt{x}x}{(2 + (n+1)\sqrt{x})(2 + n\sqrt{x})} \\ &= \frac{-x - (n+1)x\sqrt{x}}{(2 + (n+1)\sqrt{x})(2 + n\sqrt{x})} \\ &\leq 0. \end{aligned}$$

Verwenden wir dies in (3.7.4), folgt aber

$$\sqrt{x} - p_{n+1}(x) \leq \frac{2\sqrt{x}}{2 + (n+1)\sqrt{x}},$$

was den Induktionsbeweis abschließt. Mit dieser Abschätzung ist aber auch (3.7.3) klar, da

$$\lim_{n \rightarrow \infty} \frac{2\sqrt{x}}{2 + n\sqrt{x}} = 0,$$

und somit das Sandwich-Prinzip zum Tragen kommt. □

Wir werden dieses sehr spezielle Resultat überraschenderweise noch öfters benutzen können, um wichtige fundamentale Resultate zu zeigen. Zunächst ist es aber einfach ein Beispiel dafür, wie wir aus einfachen Funktionen, den Polynomen  $p_n$ , eine komplizierte Funktion, die Wurzelfunktion, *punktwise* approximieren können:

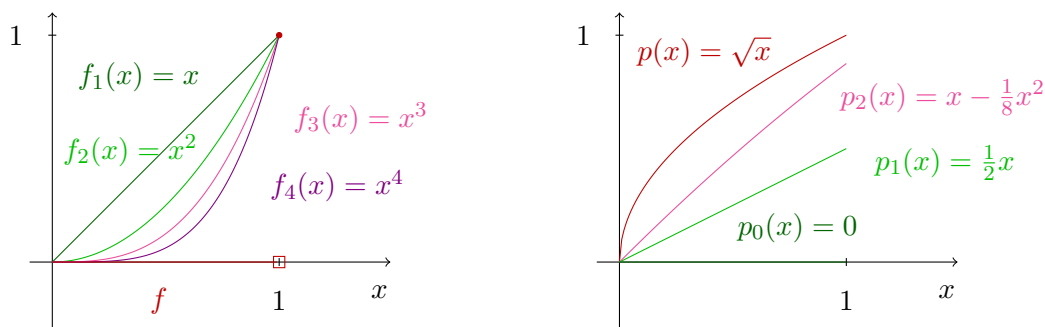


Abbildung 3.28: Die Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  hat eine unstetige Grenzfunktion  $f$ , während die Funktionenfolge  $(p_n)_{n \in \mathbb{N}}$  gegen die stetige Wurzelfunktion konvergiert.

**Definition 3.7.2 (Punktweise Konvergenz)** Sei  $X$  eine Menge und  $f_n \in \text{Abb}(X, \mathbb{K})$  für  $n \in \mathbb{N}$ . Die Folge  $(f_n)_{n \in \mathbb{N}}$  konvergiert punktweise gegen die Funktion  $f \in \text{Abb}(X, \mathbb{K})$ , wenn für jedes  $x \in X$

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad (3.7.5)$$

gilt. Die Funktion  $f$  heißt dann Grenzfunktion der Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$ .

**Bemerkung 3.7.3** Die punktweise Konvergenz der Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  benutzt nur den Konvergenzbegriff in  $\mathbb{K}$ , die Menge  $X$  tritt lediglich zur Buchhaltung auf und trägt keine weitere relevante Struktur. Die üblichen Rechenregeln für Grenzwerte übertragen sich dann direkt auf punktweise Konvergenz von Funktionenfolgen.

**Beispiel 3.7.4 (Punktweise Konvergenz)** Wir betrachten  $X = [0, 1]$  als Definitionsbereich.

i.) Die Funktionen  $f_n(x) = x^n$  für  $x \in [0, 1]$  konvergieren punktweise gegen die Funktion

$$f(x) = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1. \end{cases} \quad (3.7.6)$$

Dieses Beispiel zeigt, dass punktweise Konvergenz von stetigen Funktionen, wie den Polynomen  $f_n$ , zu einer Grenzfunktion  $f$  führen kann, die nicht länger stetig ist, siehe Abbildung 3.28.

ii.) Die Polynome  $p_n$  aus Proposition 3.7.1 konvergieren punktweise auf  $X = [0, 1]$  gegen die Wurzelfunktion, siehe erneut Abbildung 3.28. Diese ist nach wie vor stetig.

Es scheint also einen fundamentalen Unterschied in der Art der punktweisen Konvergenz dieser beiden Beispiele zu geben.

Die Schwierigkeit bei punktweiser Konvergenz ist, dass keinerlei Beziehung der Punkte aus dem Definitionsbereich  $X$  berücksichtigt werden müssen. Die Konvergenz an verschiedenen Punkten  $x, y \in X$  kann daher sehr unterschiedlich schnell erfolgen, so dass selbst bei stetigen Funktionen  $f_n$  die Grenzfunktion „zerreißt“ wie in Beispiel 3.7.4, i.). Der Begriff der *gleichmäßigen Konvergenz* behebt nun dieses Defizit:

**Definition 3.7.5 (Gleichmäßige Konvergenz)** Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen, und sei  $(f_n)_{n \in \mathbb{N}}$  eine Folge von Funktionen  $f_n \in \text{Abb}(X, \mathbb{K})$ .

i.) Die Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  konvergiert gleichmäßig gegen eine Grenzfunktion  $f \in \text{Abb}(X, \mathbb{K})$ , wenn es zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  gibt, so dass für alle  $n \geq N$  und alle  $x \in X$

$$|f_n(x) - f(x)| < \epsilon \quad (3.7.7)$$

gilt.

- ii.) Die Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  konvergiert lokal gleichmäßig auf  $X$  gegen eine Grenzfunktion  $f \in \text{Abb}(X, \mathbb{K})$ , wenn jeder Punkt  $x \in X$  eine offene Umgebung  $U \subseteq X$  besitzt, so dass  $(f_n|_U)_{n \in \mathbb{N}}$  gleichmäßig gegen  $f|_U$  auf  $U$  konvergiert.

**Bemerkung 3.7.6 (Gleichmäßige Konvergenz)** Eine gleichmäßig konvergente Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  konvergiert auch lokal gleichmäßig, jedoch nicht notwendigerweise umgekehrt. Wichtige Beispiele werden wir hierzu noch sehen. Eine lokal gleichmäßig konvergente Folge von Funktionen konvergiert erst recht punktweise. Auch hier gilt die Umkehrung im Allgemeinen nicht.

**Proposition 3.7.7** Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen und  $f_n \in \text{Abb}(X, \mathbb{K})$  für  $n \in \mathbb{N}$ .

- i.) Die Folge  $(f_n)_{n \in \mathbb{N}}$  konvergiert genau dann gleichmäßig gegen eine Grenzfunktion  $f \in \text{Abb}(X, \mathbb{K})$ , wenn

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0. \quad (3.7.8)$$

- ii.) Ist  $X$  zudem offen, so konvergiert die Folge  $(f_n)_{n \in \mathbb{N}}$  genau dann lokal gleichmäßig gegen  $f \in \text{Abb}(X, \mathbb{K})$  wenn für alle kompakten Teilmengen  $K \subseteq X$

$$\lim_{n \rightarrow \infty} \|f_n - f\|_K = 0 \quad (3.7.9)$$

gilt.

BEWEIS: Wir erinnern an die Definition der Supremumsnormen

$$\|f\|_K = \sup_{x \in K} |f(x)|$$

für kompakte  $K \subseteq X$  sowie  $\|f\|_\infty = \sup_{x \in X} |f(x)|$ , siehe Definition 3.5.5, wobei wir den Wert  $+\infty$  für das Supremum hier zulassen. Konvergiert nun  $(f_n)_{n \in \mathbb{N}}$  gleichmäßig gegen  $f$ , so gibt es zu  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit

$$|f_n(x) - f(x)| < \epsilon$$

für alle  $n \geq N$  und alle  $x \in X$ . Daher gilt auch

$$\|f_n - f\|_\infty = \sup_{x \in X} |f_n(x) - f(x)| \leq \epsilon$$

für alle  $n \geq N$ . Insbesondere sind die Suprema  $\|f_n - f\|_\infty$  für fast alle  $n$  endlich. Da wir dies für alle  $\epsilon > 0$  erreichen können, folgt (3.7.8). Gilt umgekehrt (3.7.8), so finden wir zu  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit

$$\sup_{x \in X} |f_n(x) - f(x)| < \epsilon$$

für alle  $n \geq N$ . Da aber

$$|f_n(x_0) - f(x_0)| \leq \sup_{x \in X} |f_n(x) - f(x)|$$

für alle  $x_0 \in X$  gilt, folgt die gleichmäßige Konvergenz, womit i.) gezeigt ist. Für den zweiten Teil wissen wir zunächst, dass es zu jedem  $x_0 \in X$  eine offene Umgebung  $U_{x_0} \subseteq X$  in  $X$  gibt, so dass

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{U_{x_0}} = 0 \quad (*)$$

nach dem ersten Teil. Ist nun  $K \subseteq X$  kompakt, so gibt es endlich viele Punkte  $x_1, \dots, x_n \in X$  mit

$$K \subseteq U_{x_1} \cup \dots \cup U_{x_n},$$

da alle  $\{U_{x_0}\}_{x_0 \in X}$  sicherlich eine offenen Überdeckung von  $K$  bilden. Damit folgt aber

$$\begin{aligned} \|f_n - f\|_K &= \sup_{x \in K} |f_n(x) - f(x)| \\ &\leq \sup_{x \in U_{x_1} \cup \dots \cup U_{x_n}} |f_n(x) - f(x)| \\ &= \max_{i=1}^n \sup_{x \in U_{x_i}} |f_n(x) - f(x)| \\ &= \max_{i=1}^n \|f_n - f\|_{U_{x_i}}. \end{aligned}$$

Wegen (\*) geht dieses Maximum der endlich vielen Beiträge  $\|f_n - f\|_{U_{x_i}}$  immer noch gegen Null, was dann

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

zeigt. Man beachte, dass die Kompaktheit von  $K$  hier entscheidend eingeht. Sei nun umgekehrt (3.7.9) für alle kompakten Teilmengen von  $X$  erfüllt. Da  $X$  selbst als offen angenommen wurde, gibt es zu  $x_0 \in X$  ein  $r > 0$  mit  $B_{2r}(x_0) \subseteq X$ . Dann ist  $B_r(x_0)^{\text{cl}} \subseteq X$  jetzt aber kompakt, womit

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{B_r(x_0)^{\text{cl}}} = 0$$

nach Voraussetzung gilt. Wegen

$$\sup_{x \in B_r(x_0)} |f_n(x) - f(x)| \leq \sup_{x \in B_r(x_0)^{\text{cl}}} |f_n(x) - f(x)|$$

folgt dann, dass auch

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{B_r(x_0)} = 0,$$

womit eine offene Umgebung von  $x_0$ , nämlich  $B_r(x_0)$ , gefunden wurde, auf der  $(f_n)_{n \in \mathbb{N}}$  gleichmäßig gegen  $f$  konvergiert.  $\square$

Die Umformulierung der (lokal) gleichmäßigen Konvergenz mit Hilfe von  $\|\cdot\|_\infty$  beziehungsweise  $\|\cdot\|_K$  legt nahe, von Anfang an nur solche Funktionen zu betrachten, für die  $\|f\|_\infty < \infty$  beziehungsweise  $\|f\|_K < \infty$  für alle  $K$  endlich ist. Dies sind gerade die Unterräume der beschränkten Funktionen  $\mathcal{B}(X, \mathbb{K})$  beziehungsweise der lokal beschränkten Funktionen  $\mathcal{B}_{\text{loc}}(X, \mathbb{K})$  aus Definition 3.5.5. Es zeigt sich, dass diese Eigenschaften unter (lokal) gleichmäßiger Konvergenz erhalten bleiben:

**Proposition 3.7.8** *Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen und  $f_n \in \text{Abb}(X, \mathbb{K})$  für alle  $n \in \mathbb{N}$ .*

i.) *Gilt  $f_n \in \mathcal{B}(X, \mathbb{K})$ , und konvergiert  $(f_n)_{n \in \mathbb{N}}$  gleichmäßig gegen  $f \in \text{Abb}(X, \mathbb{K})$ , so folgt  $f \in \mathcal{B}(X, \mathbb{K})$ .*

ii.) *Gilt  $f_n \in \mathcal{B}(X, \mathbb{K})$ , so konvergiert die Folge  $(f_n)_{n \in \mathbb{N}}$  genau dann gleichmäßig auf  $X$ , wenn sie eine gleichmäßige Cauchy-Folge ist, also zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit*

$$\|f_n - f_m\|_\infty < \epsilon \tag{3.7.10}$$

*für alle  $n, m \geq N$  existiert.*

iii.) *Gilt  $f_n \in \mathcal{B}_{\text{loc}}(X, \mathbb{K})$  und konvergiert  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen  $f \in \text{Abb}(X, \mathbb{K})$ , so folgt  $f \in \mathcal{B}_{\text{loc}}(X, \mathbb{K})$ .*

iv.) *Gilt  $f_n \in \mathcal{B}_{\text{loc}}(X, \mathbb{K})$  und ist  $X$  offen, so konvergiert die Folge  $(f_n)_{n \in \mathbb{N}}$  genau dann lokal gleichmäßig, wenn sie eine lokal gleichmäßige Cauchy-Folge ist, also für alle kompakten Teilmengen  $K \in X$  und alle  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit*

$$\|f_n - f_m\|_K < \epsilon \tag{3.7.11}$$

*für alle  $n, m \geq N$  existiert.*



BEWEIS: Sei  $\epsilon > 0$  vorgegeben und  $N \in \mathbb{N}$  gefunden, so dass  $\|f_n - f\|_\infty < \epsilon$  für  $n \geq N$  gilt. Dann gilt für  $x \in X$

$$\begin{aligned} |f(x)| &= |f_n(x) - f(x) + f_n(x)| \\ &\leq |f_n(x)| + |f_n(x) - f(x)| \\ &\leq \sup_{x \in X} |f_n(x)| + \sup_{x \in X} |f_n(x) - f(x)| \\ &= \|f_n\|_\infty + \|f_n - f\|_\infty \\ &< \|f_n\|_\infty + \epsilon, \end{aligned}$$

sobald  $n \geq N$ . Also gilt auch

$$\|f\|_\infty = \sup_{x \in X} |f(x)| \leq \|f_n\|_\infty + \epsilon < \infty,$$

womit  $f \in \mathcal{B}(X, \mathbb{K})$  gezeigt ist. Als nächstes betrachten wir den Fall, dass  $(f_n)_{n \in \mathbb{N}}$  eine gleichmäßige Cauchy-Folge ist. Sei  $\epsilon > 0$  und  $N \in \mathbb{N}$  passend wie in (3.7.10). Für  $x_0 \in X$  gilt dann

$$|f_n(x_0) - f_m(x_0)| \leq \sup\{|f_n(x) - f_m(x)| \mid x \in X\} = \|f_n - f_m\|_\infty < \epsilon,$$

womit  $(f_n(x_0))_{n \in \mathbb{N}}$  für jedes  $x_0 \in X$  eine Cauchy-Folge ist. Daher ist  $(f_n)_{n \in \mathbb{N}}$  punktweise konvergent gegen eine Funktion  $f: X \rightarrow \mathbb{K}$  via

$$f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

mit  $x \in X$ . Für diese Grenzfunktion gilt dann

$$|f_n(x) - f(x)| = \lim_{m \rightarrow \infty} |f_n(x) - f_m(x)| \leq \epsilon,$$

da im Grenzwert  $m \rightarrow \infty$  die Ungleichung  $m \geq N$  sicher erfüllt wird. Dies bedeutet aber  $\|f_n - f\|_\infty \leq \epsilon$ . Daher konvergiert  $(f_n)_{n \in \mathbb{N}}$  gleichmäßig gegen  $f$  und wir können mit dem ersten Teil wieder  $f \in \mathcal{B}(X, \mathbb{K})$  schließen. Aus der gleichmäßigen Konvergenz von  $(f_n)_{n \in \mathbb{N}}$  gegen  $f \in \mathcal{B}(X, \mathbb{K})$  folgt aber sofort die Cauchy-Bedingung, da wir zu  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit  $\|f_n - f\|_\infty < \frac{\epsilon}{2}$  für alle  $n \geq N$  finden und daher direkt  $\|f_n - f_m\|_\infty \leq \|f_n - f\|_\infty + \|f_m - f\|_\infty < \epsilon$  schließen können, wenn  $n, m \geq N$ . Sei nun  $x_0 \in X$  vorgegeben und  $U \subseteq X$  eine offene Umgebung von  $x_0$  in  $X$ , so dass  $(f_n)_{n \in \mathbb{N}}$  auf  $U$  gleichmäßig gegen  $f$  konvergiert. Zu  $\epsilon > 0$  finden wir wieder  $N \in \mathbb{N}$  mit

$$\sup_{x \in U} |f_n(x) - f(x)| < \epsilon \quad (*)$$

für  $n \geq N$ . Die Funktion  $f_N$  ist lokal beschränkt, also gibt es eine offene Umgebung  $V \subseteq X$  von  $x_0$  in  $X$  mit

$$\|f_N\|_V = \sup_{x \in V} |f_N(x)| < \infty.$$

Mit  $(*)$  gilt

$$\sup_{x \in V \cap U} |f_n(x) - f(x)| < \epsilon$$

für alle  $n \geq N$ . Es folgt für  $x \in V \cap U$

$$|f(x)| \leq |f_N(x) - f(x)| + |f_N(x)| \leq \|f_N - f\|_{V \cap U} + \|f_N\|_{V \cap U} \leq \|f_N\|_V + \epsilon,$$

womit  $f$  auf  $V \cap U$  beschränkt und damit insgesamt lokal beschränkt ist. Der letzte Teil folgt dann wie der zweite, indem man die Supremumsnorm  $\|\cdot\|_\infty$  entsprechend durch  $\|\cdot\|_K$  für jedes kompakte  $K \subseteq X$  ersetzt und benutzt, dass lokal gleichmäßige Konvergenz mit Hilfe der  $\|\cdot\|_K$  gemäß Proposition 3.7.7, ii.), beschrieben werden kann.  $\square$

**Beispiel 3.7.9** Ohne die Annahme der (lokalen) gleichmäßigen Konvergenz ist die Aussage der Proposition im Allgemeinen falsch. Um dies zu sehen betrachten wir die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  aus Beispiel 3.3.6, *iii.*), also

$$f(x) = \begin{cases} q & \text{für } x = \frac{p}{q} \in \mathbb{Q} \setminus \{0\} \text{ mit teilerfremden } p \in \mathbb{Z}, q \in \mathbb{N} \\ 1 & \text{für } x = 0 \\ 0 & \text{sonst.} \end{cases} \quad (3.7.12)$$

Sei nun  $\sigma: \mathbb{N} \rightarrow \mathbb{Q} \setminus \{0\}$  eine Abzählung von  $\mathbb{Q}$ , also eine bijektive Abbildung, die uns Abbildungen  $\sigma': \mathbb{N} \rightarrow \mathbb{Z}$  und  $\sigma'': \mathbb{N} \rightarrow \mathbb{N}$  mit

$$\sigma(n) = \frac{\sigma'(n)}{\sigma''(n)}$$

induziert, wobei wieder verlangt wird, dass  $\sigma'(n)$  und  $\sigma''(n)$  teilerfremd sind. Diese Abbildung verwenden wir nun, um eine Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  zu definieren. Wir setzen

$$f_n(x) = \begin{cases} \sigma''(k) & \text{wenn } x \in \mathbb{Q} \setminus \{0\}, x = \sigma(k) \text{ und } k \leq n \\ 1 & \text{für } x = 0 \\ 0 & \text{sonst.} \end{cases} \quad (3.7.13)$$

Bei irrationalen  $x \in \mathbb{R} \setminus \mathbb{Q}$  stimmt  $f_n(x)$  mit  $f(x)$  überein, womit wir dort trivialerweise Konvergenz vorliegen haben. Ist  $x$  dagegen rational, so gibt es ein  $k \in \mathbb{N}$  mit  $x = \sigma(k) = \frac{\sigma'(k)}{\sigma''(k)}$ . Ist dann  $n \geq k$  so gilt  $f_n(x) = f(x)$  auch für dieses  $x$  womit wir auch bei rationalen Punkten Konvergenz vorliegen haben. Es gilt also punktweise

$$\lim_{n \rightarrow \infty} f_n(x) = f(x). \quad (3.7.14)$$

Da nun  $f_n$  nur endlich viele Werte ungleich Null annimmt, ist  $f_n \in \mathcal{B}(\mathbb{R}, \mathbb{R})$  sicherlich beschränkt. Trotzdem ist der punktweise Limes  $f$  nach Beispiel 3.3.6, *iii.*), nicht einmal lokal beschränkt.

Viel wichtiger als (lokal) beschränkte Funktionen sind natürlich stetige Funktionen. Hier haben wir wieder ein gutes Verhalten bei lokal gleichmäßiger Konvergenz:

**Satz 3.7.10 (Lokal gleichmäßige Konvergenz)** *Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen und  $f_n \in \mathcal{C}(X, \mathbb{K})$  für  $n \in \mathbb{N}$  stetig. Konvergiert die Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen  $f \in \text{Abb}(X, \mathbb{K})$ , so gilt  $f \in \mathcal{C}(X, \mathbb{K})$ .*

BEWEIS: Sei  $x_0 \in X$  vorgegeben. Wir wollen zeigen, dass  $f$  bei  $x_0$  stetig ist. Da die Konvergenz von  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig ist, können wir ein  $\epsilon > 0$  vorgeben und finden eine offene Umgebung  $U \subseteq X$  in  $X$  von  $x_0$  und ein  $N \in \mathbb{N}$ , sodass

$$|f_n(x) - f(x)| < \frac{\epsilon}{3}$$

für alle  $x \in U$  und alle  $n \geq N$  gilt. Insbesondere gilt auch

$$|f_n(x_0) - f(x_0)| < \frac{\epsilon}{3}$$

für alle  $n \geq N$ . Da nun  $f_N$  selbst als stetig vorausgesetzt wurde, gibt es eine weitere offene Umgebung  $\tilde{U} \subseteq X$  in  $X$  von  $x_0$ , sodass

$$|f_N(x) - f_N(x_0)| < \frac{\epsilon}{3}$$

für alle  $x \in \tilde{U}$  gilt. Da  $U \cap \tilde{U}$  immer noch eine offene Umgebung von  $x_0$  in  $X$  ist, erhalten wir für alle  $x \in U \cap \tilde{U}$

$$|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)|$$

$$\begin{aligned}
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \epsilon,
\end{aligned}$$

was die Stetigkeit von  $f$  bei  $x_0 \in X$  zeigt. Diese Beweistechnik nennt man aus naheliegenden Gründen auch ein  $\frac{\epsilon}{3}$ -Argument.  $\square$

**Beispiel 3.7.11** Die Funktionenfolge  $(f_n)_{n \in \mathbb{N}}$  mit  $f_n(x) = x^n$  aus Beispiel 3.7.4, *i.*), konvergiert lokal gleichmäßig auf  $[0, 1)$  aber nicht auf  $[0, 1]$ . Die Funktionenfolge  $(p_n)_{n \in \mathbb{N}}$  aus Beispiel 3.7.4, *ii.*), konvergiert dagegen gleichmäßig auf  $[0, 1]$ . Dies ist gerade die Abschätzung (3.7.2).

## 3.8 Konvergenz von Potenzreihen

Gleichmäßige Konvergenz oder sogar nur lokal gleichmäßige Konvergenz von stetigen Funktionen resultiert in einer stetigen Grenzfunktion. Diese Beobachtung erlaubt es nun, komplizierte und interessante stetige Funktionen zu konstruieren. Ein für die gesamte Mathematik entscheidender Mechanismus hierfür sind die Potenzreihen:

**Definition 3.8.1 (Potenzreihe)** Sei  $(a_n)_{n \in \mathbb{N}_0}$  eine reelle oder komplexe Folge. Die Reihe

$$a(z) = \sum_{n=0}^{\infty} a_n z^n \quad (3.8.1)$$

heißt dann *formale Potenzreihe* zu den Koeffizienten  $(a_n)_{n \in \mathbb{N}_0}$ . Die Menge der formalen Potenzreihen wird mit  $\mathbb{K}[[z]]$  bezeichnet.

Zunächst wird bei einer formalen Potenzreihe also keinerlei Konvergenz vorausgesetzt. Die Menge der formalen Potenzreihen  $\mathbb{K}[[z]]$  ist also nichts anderes als der Vektorraum  $\text{Abb}(\mathbb{N}_0, \mathbb{K})$  aller Abbildungen  $\mathbb{N}_0 \rightarrow \mathbb{K}$ . Dies zeigt, dass wir auch für  $\mathbb{K}[[z]]$  eine Vektorraumstruktur etablieren können, indem wir

$$\sum_{n=0}^{\infty} a_n z^n + \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} (a_n + b_n) z^n \quad (3.8.2)$$

und

$$\alpha \cdot \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} \alpha a_n z^n \quad (3.8.3)$$

für  $\alpha \in \mathbb{K}$  definieren, was gerade den punktweisen Operationen in  $\text{Abb}(\mathbb{N}_0, \mathbb{K})$  entspricht. Überraschender ist daher, dass wir eine interessante Algebrastruktur erhalten, indem wir das Produkt von formalen Potenzreihen als

$$\left( \sum_{n=0}^{\infty} a_n z^n \right) \left( \sum_{m=0}^{\infty} b_m z^m \right) = \sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k b_{n-k} \right) z^n \quad (3.8.4)$$

definieren. Hätte man absolute Konvergenz der beiden Reihen vorliegen, so wäre (3.8.4) gerade das *Cauchy-Produkt* aus Korollar 2.7.16. Man beachte jedoch, dass auch ohne vorliegende Konvergenz die formale Potenzreihe (3.8.4) wohl-definiert ist, da für die Koeffizienten des Produkts immer nur endliche Summen auftreten. Letztlich können wir (3.8.4) für beliebige Folgen von Koeffizienten  $(a_n)_{n \in \mathbb{N}_0}$  und  $(b_n)_{n \in \mathbb{N}_0}$  in einem Ring  $R$  definieren. Man erhält dann wieder einen Ring  $R[[z]]$ , siehe etwa [8, Übung 3.23]. Wir sind nur an  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$  interessiert, wofür wir das allgemeine algebraische Resultat formulieren:

**Proposition 3.8.2** Die formalen Potenzreihen  $\mathbb{K}[[z]]$  bilden bezüglich des Cauchy-Produkts (3.8.4) eine assoziative und kommutative  $\mathbb{K}$ -Algebra mit Einselement  $1 \in \mathbb{K} \subseteq \mathbb{K}[[z]]$ , wobei wir  $\mathbb{K}$  als Teilmenge von  $\mathbb{K}[[z]]$  durch

$$\mathbb{K} \ni a \mapsto a \cdot z^0 + 0 \cdot z^1 + \cdots \in \mathbb{K}[[z]] \quad (3.8.5)$$

auffassen. Die Polynome  $\mathbb{K}[z] \subseteq \mathbb{K}[[z]]$  bilden eine Unteralgebra der formalen Potenzreihen, welche nur endlich viele Koeffizienten ungleich Null besitzen.

Einen Beweis mit Anleitung findet man in Übung 3.9.31.

Da wir hier an den analytischen Eigenschaften von Potenzreihen interessiert sind, wollen wir tatsächliche Konvergenz für möglichst viele  $z \in \mathbb{K}$ . Für  $z = 0$  konvergiert (3.8.1) aus trivialen Gründen und liefert den Wert

$$a(0) = a_0. \quad (3.8.6)$$

Es stellt sich die Frage, ob es auch andere  $z \in \mathbb{K}$  gibt, für die Konvergenz vorliegt. Hier gibt das Wurzelkriterium eine definitive Antwort, die wir gleich für komplexe Potenzreihen und komplexe  $z \in \mathbb{C}$  formulieren, da dieser allgemeine Fall die reelle Situation mit erfasst:

**Satz 3.8.3 (Konvergenz von Potenzreihen)** Sei  $a \in \mathbb{C}[[z]]$  eine komplexe formale Potenzreihe, und sei

$$\alpha = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \in [0, \infty]. \quad (3.8.7)$$

Wir setzen  $R = \frac{1}{\alpha}$  mit den üblichen Konventionen  $\frac{1}{0} = +\infty$  und  $\frac{1}{+\infty} = 0$ .

i.) Für  $z \in \mathbb{C}$  mit  $|z| < R$  konvergiert die Reihe

$$a(z) = \sum_{n=0}^{\infty} a_n z^n \quad (3.8.8)$$

absolut.

ii.) Für  $z \in \mathbb{C}$  mit  $|z| > R$  divergiert die Reihe.

BEWEIS: Sei  $z \in \mathbb{C}$  gegeben. Dann gilt

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n z^n|} = |z| \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = |z| \alpha.$$

Gilt nun  $|z| < R$ , so folgt  $|z| \alpha < 1$  womit (3.8.8) nach Proposition 2.7.5, i.), absolut konvergiert. Gilt dagegen  $|z| > R$ , so ist  $|z| \alpha > 1$  und (3.8.8) divergiert, ebenfalls nach dem Wurzelkriterium aus Proposition 2.7.5, ii.).  $\square$

**Definition 3.8.4 (Konvergenzradius)** Sei  $a(z) = \sum_{n=0}^{\infty} a_n z^n \in \mathbb{C}[[z]]$  eine komplexe formale Potenzreihe. Dann heißt  $R \in [0, \infty]$  aus Satz 3.8.3 der Konvergenzradius von  $a(z)$ , und  $B_R(0)$  heißt Konvergenzkreisscheibe von  $a(z)$ .

**Bemerkung 3.8.5 (Konvergenzradius)** In der offenen Kreisscheibe  $B_R(0) \subseteq \mathbb{C}$  konvergiert die formale Potenzreihe daher absolut, außerhalb von  $B_R(0)^{\text{cl}}$  divergiert die Reihe. Auf dem Rand von  $B_R(0)$ , also für  $|z| = R$ , kann beispielabhängig sowohl Konvergenz als auch Divergenz vorliegen, siehe auch Abbildung 3.29.

### Beispiel 3.8.6 (Konvergenzradius von Potenzreihen)

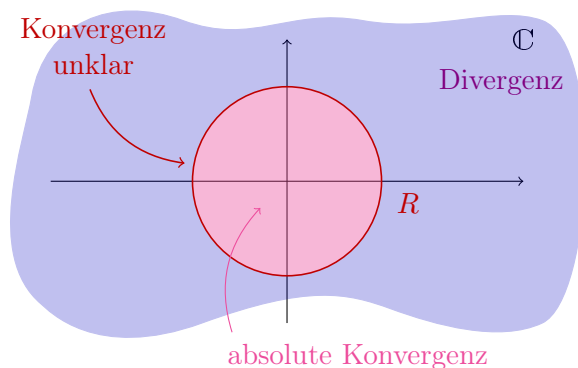


Abbildung 3.29: Der Konvergenzradius einer Potenzreihe

- i.) Die geometrische Reihe  $a(z) = \sum_{n=0}^{\infty} z^n$  hat Konvergenzradius  $R = 1$ . Für  $|z| < 1$  konvergiert sie gegen

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}, \quad (3.8.9)$$

für  $|z| > 1$  divergiert sie. Da für  $|z| = 1$  die Folge  $z^n$  keine Nullfolge ist, divergiert die geometrische Reihe auch auf dem Rand ihrer Konvergenzkreisscheibe.

- ii.) Betrachten wir die formale Potenzreihe

$$a(z) = \sum_{n=0}^{\infty} \frac{z^n}{n}, \quad (3.8.10)$$

so gilt erneut  $R = 1$ , siehe Beispiel 2.7.7, ii.). Jetzt wissen wir aber, dass die Reihe (3.8.10) für  $z = 1$  divergiert, da dies gerade die harmonische Reihe ist. Für  $z = -1$  konvergiert (3.8.10) dagegen bedingt aber nicht absolut, siehe Beispiel 2.6.10.

- iii.) Die formale Potenzreihe

$$a(z) = \sum_{n=0}^{\infty} n! z^n \quad (3.8.11)$$

besitzt Konvergenzradius  $R = 0$ . Ist nämlich  $z \neq 0$ , so ist

$$\lim_{n \rightarrow \infty} \frac{1}{n!|z|^n} = \lim_{n \rightarrow \infty} \frac{1}{n!} \frac{1}{|z|^n} = 0 \quad (3.8.12)$$

nach Beispiel 2.3.12, iv.), falls  $\frac{1}{|z|} > 1$ . Für  $\frac{1}{|z|} \leq 1$  ist (3.8.12) trivialerweise ebenso richtig. Damit ist aber für alle  $z \neq 0$  die Folge  $n!z^n$  keine Nullfolge. Vielmehr divergiert  $n!|z|^n$  gegen  $+\infty$ .

- iv.) Sei  $\epsilon > 0$ . Nach Beispiel 2.7.7, iv.), wissen wir

$$\limsup_{n \rightarrow \infty} \frac{n!^\epsilon}{(n+1)!^\epsilon} = 0, \quad (3.8.13)$$

womit nach Proposition 2.5.17 auch

$$\limsup_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n!^\epsilon}} = 0 \quad (3.8.14)$$

gilt. Daher hat die Potenzreihe

$$a(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!^\epsilon} \quad (3.8.15)$$

Konvergenzradius  $R = +\infty$ , konvergiert also für alle  $z \in \mathbb{C}$  absolut.

Aus den Rechenregeln für absolut konvergente Reihen erhalten wir sofort folgende Rechenregeln für konvergente Potenzreihen:

**Proposition 3.8.7** Seien  $a(z) = \sum_{n=0}^{\infty} a_n z^n$  und  $b(z) = \sum_{n=0}^{\infty} b_n z^n$  komplexe Potenzreihen mit Konvergenzradien  $R_a$  und  $R_b$ .

i.) Die Summe  $\sum_{n=0}^{\infty} (a_n + b_n) z^n$  hat als Konvergenzradius  $R \geq \min(R_a, R_b)$ , und es gilt für  $z \in \mathbb{C}$  mit  $|z| < \min(R_a, R_b)$

$$\sum_{n=0}^{\infty} a_n z^n + \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} (a_n + b_n) z^n. \quad (3.8.16)$$

ii.) Für alle  $w \in \mathbb{C} \setminus \{0\}$  hat  $\sum_{n=0}^{\infty} w a_n z^n$  den gleichen Konvergenzradius  $R_a$ , und es gilt

$$w \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} w a_n z^n. \quad (3.8.17)$$

Für  $w = 0$  ist (3.8.17) ebenfalls richtig, und die Potenzreihe auf der rechten Seite hat trivialerweise den Konvergenzradius  $+\infty$ .

iii.) Das Cauchy-Produkt  $\sum_{n=0}^{\infty} (\sum_{k=0}^n a_k b_{n-k}) z^n$  hat als Konvergenzradius  $R \geq \min(R_a, R_b)$  und für  $z \in \mathbb{C}$  mit  $|z| < \min(R_a, R_b)$  gilt

$$\left( \sum_{n=0}^{\infty} a_n z^n \right) \left( \sum_{n=0}^{\infty} b_n z^n \right) = \sum_{n=0}^{\infty} \left( \sum_{k=0}^n a_k b_{n-k} \right) z^n. \quad (3.8.18)$$

BEWEIS: Die ersten beiden Aussagen folgen aus Proposition 2.6.5. Die dritte Aussage ist eine Konsequenz von Korollar 2.7.16 und der Tatsache, dass Potenzreihen innerhalb ihres Konvergenzkreises absolut konvergieren.  $\square$

Als vorläufig letztes Ergebnis zu konvergenten Potenzreihen zeigen wir ihre Stetigkeit. Später werden wir dies erneut aufgreifen und viele weitere, über die Stetigkeit hinausgehende Eigenschaften von konvergenten Potenzreihen zeigen. An dieser Stelle dienen sie uns zunächst nur zur Illustration des Konzepts der lokal gleichmäßigen Konvergenz:

**Satz 3.8.8 (Stetigkeit von Potenzreihen)** Sei  $a(z) = \sum_{n=0}^{\infty} a_n z^n \in \mathbb{C}[[z]]$  eine komplexe formale Potenzreihe mit Konvergenzradius  $R > 0$ .

i.) Sei  $z \in B_R(0)$  mit  $|z| \leq \rho < R$ . Dann gilt

$$\left| \sum_{n=0}^{\infty} a_n z^n \right| \leq \sum_{n=0}^{\infty} |a_n| \rho^n < \infty. \quad (3.8.19)$$

ii.) Fassen wir  $\sum_{n=0}^{\infty} a_n z^n$  als punktweise konvergente Reihe von Funktionen von  $z$  auf, so konvergiert diese Funktionenreihe lokal gleichmäßig auf  $B_R(0)$ .

iii.) Die Funktion

$$B_R(0) \ni z \mapsto \sum_{n=0}^{\infty} a_n z^n \in \mathbb{C} \quad (3.8.20)$$

ist stetig.

BEWEIS: Nach der Dreiecksungleichung für Reihen gemäß Proposition 2.7.2 erhalten wir den ersten Teil. Man beachte, dass die Reihe auf der rechten Seite in (3.8.19) tatsächlich konvergiert, dies ist gerade die absolute Konvergenz der Potenzreihe. Für den zweiten Teil geben wir  $\epsilon > 0$  und  $\rho < R$  vor und finden zunächst ein  $N \in \mathbb{N}$  derart, dass

$$\sum_{n=N+1}^{\infty} |a_n| \rho^n < \epsilon$$

gilt, da diese Reihe ja konvergiert. Dann gilt aber für  $z \in B_R(0)^{\text{cl}}$

$$\left| \sum_{n=0}^{\infty} a_n z^n - \sum_{n=0}^N a_n z^n \right| = \left| \sum_{n=N+1}^{\infty} a_n z^n \right| \leq \sum_{n=N+1}^{\infty} |a_n| |z|^n \leq \sum_{n=N+1}^{\infty} |a_n| \rho^n < \epsilon. \quad (*)$$

Ist nun  $z_0 \in B_R(0)$  ein beliebiger Punkt, so ist  $B_\rho(0)$  mit  $|z_0| < \rho < R$  eine offene Umgebung von  $z_0$ . Aus dieser haben wir aber gerade die gleichmäßige Abschätzung (\*) der Konvergenz der Partialsummen  $\sum_{n=0}^N a_n z^n$  gegen den Wert der Potenzreihe  $\sum_{n=0}^{\infty} a_n z^n$  gezeigt, da (\*) ja gleichermaßen für alle  $z \in B_\rho(0)$  gilt. Damit liegt lokal gleichmäßige Konvergenz vor. Der dritte Teil folgt dann aus Satz 3.7.10.  $\square$

## 3.9 Übungen

### Übung 3.9.1 (Häufungspunkte von Mengen und Folgen)

### Übung 3.9.2 (Auf $x_0$ zu gerichtete Mengen)

**Übung 3.9.3 (Abschluss)** Sei  $X \subseteq \mathbb{K}$ . Zeigen Sie, dass  $x_0 \in \mathbb{K}$  genau dann im Abschluss  $x_0 \in X^{\text{cl}}$  von  $X$  liegt, wenn es eine Folge  $(x_n)_{n \in \mathbb{N}}$  mit  $x_n \in X$  und  $\lim_{n \rightarrow \infty} x_n = x_0$  gibt.

Hinweis: Ist  $x_0 \in \mathbb{K}$  ein Berührungspunkt, so gilt  $B_{\frac{1}{n}}(x_0) \cap X \neq \emptyset$  für alle  $n \in \mathbb{N}$ . Wie findet man nun eine geeignete Folge?

**Übung 3.9.4 (Abschluss von offenen Kugeln)** Sei  $r > 0$  und  $z \in \mathbb{K}$ . Zeigen Sie, dass der Abschluss der offenen Kugel  $B_r(z)$  gerade die abgeschlossene Kugel  $B_r(z)^{\text{cl}}$  ist, womit diese Notation gerechtfertigt ist.

Hinweis: Übung 3.9.3 ist hier hilfreich.

**Übung 3.9.5 (Abschluss von  $\mathbb{Q}$ )** Zeigen Sie, dass der Abschluss der rationalen Zahlen  $\mathbb{Q} \subseteq \mathbb{R}$  in  $\mathbb{R}$  durch

$$\mathbb{Q}^{\text{cl}} = \mathbb{R} \quad (3.9.1)$$

gegeben ist.

Hinweis: Bestimmen Sie die Menge der Häufungspunkt von rationalen Zahlen, indem Sie geeignete Folgen von rationalen Zahlen betrachten.

### Übung 3.9.6 (Konvergenz von Funktionswerten)

### Übung 3.9.7 (Konvergenz von Funktionswerten mit Umgebungen)

### Übung 3.9.8 (Rechenregeln für Konvergenz von Funktionswerten)

### Übung 3.9.9 (Funktionsgrenzwerte der Exponentialfunktion) Sei $a > 1$ .

i.) Zeigen Sie

$$\lim_{x \rightarrow \infty} a^x = \infty. \quad (3.9.2)$$

Hinweis: Lassen Sie sich von Beispiel 2.2.8, ii.), inspirieren.

ii.) Zeigen Sie

$$\lim_{x \rightarrow -\infty} a^x = 0. \quad (3.9.3)$$

Hinweis: Verwenden Sie (3.9.2), indem Sie zu Kehrwerten übergehen.

### Übung 3.9.10 (Offene Mengen in $\mathbb{R}$ )

**Übung 3.9.11 (Abgeschlossene Teilmengen in abgeschlossenen Mengen)** Sei  $A \subseteq \mathbb{K}$  eine abgeschlossene Menge. Sei weiter  $B \subseteq A$  eine Teilmenge.

- i.) Zeigen Sie, dass  $B$  genau dann in  $A$  abgeschlossen ist, wenn  $B$  als Teilmengen von  $\mathbb{K}$  abgeschlossen ist.
- ii.) Geben Sie einfache Beispiele für abgeschlossene Teilmengen  $C \subseteq X$  in  $X$ , die in  $\mathbb{K}$  nicht abgeschlossen sind, um zu zeigen, dass in i.) die Abgeschlossenheit von  $A$  eine entscheidende Bedingung ist.

**Übung 3.9.12 (Links- und rechtsseitige Stetigkeit)** Sei  $I \subseteq \mathbb{R}$  ein Intervall und  $f: I \rightarrow \mathbb{K}$ . Sei weiter  $x_0 \in I$ . Dann heißt  $f$  *linksseitig* (*rechtsseitig*) *stetig* bei  $x_0$ , falls  $f|_{I \cap (-\infty, x_0]}$  beziehungsweise  $f|_{I \cap [x_0, \infty)}$  stetig ist.

- i.) Zeigen Sie, dass  $f$  genau dann linksseitig stetig bei  $x_0$  ist, wenn  $\lim_{x \nearrow x_0} f(x) = f(x_0)$  gilt.
- ii.) Zeigen Sie, dass  $f$  genau dann linksseitig stetig bei  $x_0$  ist, wenn zu jedem  $\epsilon > 0$  ein  $\delta > 0$  existiert, so dass für alle  $x \in I$  mit  $x_0 - x < \delta$  auch  $|f(x) - f(x_0)| < \epsilon$  gilt.
- iii.) Formulieren Sie analoge Aussagen für die rechtsseitige Stetigkeit sowohl mit Grenzwerten als auch ein  $\epsilon\delta$ -Kriterium.
- iv.) Zeigen Sie, dass  $f$  genau dann stetig bei  $x_0$  ist, wenn  $f$  sowohl links- als auch rechtsseitig stetig bei  $x_0$  ist.

Hinweis: Hier sind die  $\epsilon\delta$ -Kriterien vielleicht die einfachste Möglichkeit.

- v.) Finden Sie Beispiele für links- aber nicht rechtsseitig stetige Funktionen. Finden Sie auch Beispiele für Funktionen, für die beide Limites  $\lim_{x \nearrow x_0} f(x)$  und  $\lim_{x \searrow x_0} f(x)$  existieren, aber nicht gleich sind beziehungsweise gleich sind aber nicht mit  $f(x_0)$  übereinstimmen.

**Übung 3.9.13 (Stetigkeit ist lokal)** Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen. Sei weiter  $\{O_i\}_{i \in I}$  eine offene Überdeckung von  $X$ .

- i.) Zeigen Sie, dass eine Funktion  $f: X \rightarrow \mathbb{K}$  genau dann stetig ist, wenn für alle  $i \in I$  die Einschränkungen  $f|_{O_i \cap X}: O_i \cap X \rightarrow \mathbb{K}$  stetig sind.

Hinweis: Die eine Richtung ist einfach. Für die andere benutzen Sie, dass die Stetigkeit von  $f|_{O_i \cap X}$  die Folgenstetigkeit bei jedem Punkt  $x \in O_i \cap X$  bedeutet. Verwenden Sie dann, dass jedes  $O_i \cap X$  offen in  $X$  ist, um die Folgenstetigkeit von  $f$  bei  $x$  zu zeigen.

- ii.) Seien nun Funktionen  $f_i: O_i \cap X \rightarrow \mathbb{K}$  für alle  $i \in I$  vorgegeben. Zeigen Sie, dass es genau dann eine Funktion  $f: X \rightarrow \mathbb{K}$  mit  $f|_{O_i \cap X} = f_i$  für alle  $i \in I$  gibt, wenn

$$f_i|_{O_i \cap O_j \cap X} = f_j|_{O_i \cap O_j \cap X} \quad (3.9.4)$$

für alle  $i, j \in I$  mit  $O_i \cap O_j \cap X \neq \emptyset$  gilt. Zeigen Sie, dass in diesem Fall  $f$  eindeutig bestimmt ist.



- iii.) Seien nun stetige Funktionen  $f_i: O_i \cap X \rightarrow \mathbb{K}$  für alle  $i \in I$  vorgegeben, welche die Eigenschaft (3.9.4) erfüllen. Zeigen Sie, dass dann die zugehörige eindeutig bestimmte Funktion  $f$  wieder stetig ist.

Hinweis: Hier hilft i.).

Das Resultat dieser Übung ist also, dass Stetigkeit ein lokales Konzept ist, welches in beliebig kleinen offenen Mengen geprüft werden kann.

### Übung 3.9.14 (Was ist $0^0$ ?)

**Übung 3.9.15 (Stetigkeit an einem Punkt)** Seien  $f, g: X \rightarrow \mathbb{K}$  stetig bei einem Punkt  $x_0 \in X$ .

- i.) Zeigen Sie, dass für alle  $z, w \in \mathbb{K}$  auch  $zf + wg$  bei  $x_0$  stetig sind.
- ii.) Zeigen Sie, dass  $fg$  bei  $x_0$  stetig ist.
- iii.) Zeigen Sie, dass im Falle  $\mathbb{K} = \mathbb{C}$  auch  $\bar{f}$  bei  $x_0$  stetig ist.
- iv.) Zeigen Sie, dass falls  $g(x_0) \neq 0$  gilt, auch  $\frac{f}{g}$  bei  $x_0$  stetig ist.
- v.) Es gelte  $f(x_0) > 0$ . Zeigen Sie, dass dann eine Umgebung  $U \subseteq X$  von  $x_0$  in  $X$  existiert, für die  $f|_U > 0$  gilt. Folgern Sie, dass  $f^p$  für  $p > 0$  zumindest auf  $U$  definiert und bei  $x_0$  stetig ist.

Hinweis: Um  $f^p$  definieren zu können, benötigen wir positive Funktionswerte.

**Übung 3.9.16 (Ein offene Überdeckung von  $\mathbb{Q}$ )** Sei  $A \subseteq \mathbb{R}$  eine abzählbare Teilmenge. Geben Sie  $\epsilon > 0$  vor. Konstruieren Sie dann eine abzählbare offene Überdeckung  $\{B_{\epsilon_n}(a_n)\}$  von  $A$  derart, dass die Gesamtlänge der überdeckenden Intervalle kleiner  $\epsilon$  ist. Was sagt dies über die „Größe“ von  $\mathbb{Q} \subseteq \mathbb{R}$  aus?

**Übung 3.9.17 (Endliche Vereinigungen von Kompakta)** Zeigen Sie direkt mit der Definition von Überdeckungskompaktheit, dass endliche Vereinigungen von kompakten Mengen wieder kompakt sind.

**Übung 3.9.18 (Kompaktheit von Rechtecken)** Sei  $K = [a_1, b_1] + i[c_1, d_1] \subseteq \mathbb{C}$  ein abgeschlossenes Rechteck. Ziel dieser Übung ist, in Analogie zu Satz 3.4.3 einen direkten und unabhängigen Beweis der Kompaktheit von  $K$  zu geben. Sei dazu  $\{O_i\}_{i \in I}$  eine offene Überdeckung von  $K$  gegeben, wobei Sie annehmen, dass diese *keine* endliche Teilüberdeckung besitzt.

- i.) Betrachten Sie die abgeschlossenen Rechtecke

$$A_1 = [a_1, \tfrac{1}{2}(a_1 + b_1)] + i[c_1, \tfrac{1}{2}(c_1 + d_1)], \quad B_1 = [a_1, \tfrac{1}{2}(a_1 + b_1)] + i[\tfrac{1}{2}(c_1 + d_1), d_1] \quad (3.9.5)$$

sowie

$$C_1 = [\tfrac{1}{2}(a_1 + b_1), b_1] + i[c_1, \tfrac{1}{2}(c_1 + d_1)], \quad D_1 = [\tfrac{1}{2}(a_1 + b_1), b_1] + i[\tfrac{1}{2}(c_1 + d_1), d_1] \quad (3.9.6)$$

innerhalb von  $K_1 = K$ . Zeigen Sie  $K_1 = A_1 \cup B_1 \cup C_1 \cup D_1$ , und visualisieren Sie dies in einer Skizze.

- ii.) Folgern Sie, dass mindestens eines dieser kleinen Rechtecke ebenfalls keine endliche Teilüberdeckung besitzt. Bezeichnen Sie dieses mit  $K_2$ .
- iii.) Konstruieren Sie induktiv eine Folge  $K_n = [a_n, b_n] + i[c_n, d_n]$  von Teilrechtecken in  $K_1$  derart, dass die Kantenlängen sich in jedem Schritt halbieren und keines der  $K_n$  eine endliche Teilüberdeckung besitzt.

iv.) Zeigen Sie, dass die Folgen der Intervallgrenzen  $(a_n)_{n \in \mathbb{N}}$ ,  $(b_n)_{n \in \mathbb{N}}$ ,  $(c_n)_{n \in \mathbb{N}}$ , und  $(d_n)_{n \in \mathbb{N}}$  konvergieren und  $\alpha = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n$  sowie  $\beta = \lim_{n \rightarrow \infty} c_n = \lim_{n \rightarrow \infty} d_n$  erfüllen. Folgern Sie

$$\{\alpha + i\beta\} = \bigcap_{n \in \mathbb{N}} K_n. \quad (3.9.7)$$

v.) Zeigen Sie, dass es mindestens einen Index  $i \in I$  mit  $\alpha + i\beta \in O_i$  gibt. Wieso ist dies nun der erwünschte Widerspruch, um den Beweis der Kompaktheit von  $K$  abzuschließen?

**Übung 3.9.19 (Zusammenhängende Teilmengen von  $\mathbb{R}$ )** Zeigen Sie, dass die einzigen zusammenhängenden Teilmengen von  $\mathbb{R}$  die Intervalle (offene, abgeschlossene und halboffene) sind.

Hinweis: Benutzen Sie Proposition 3.4.14.

**Übung 3.9.20 (Zusammenhang und Vollständigkeit)**

**Übung 3.9.21 (Alternative Beweise von Satz 3.5.3)** Zeigen Sie Satz 3.5.3 unter Verwendung des Folgenkriteriums für Stetigkeit sowie unter Verwendung des  $\epsilon\delta$ -Kriteriums für Stetigkeit, und vergleichen Sie Ihre Beweise mit dem bei Satz 3.5.3 angegebenen Beweis.

**Übung 3.9.22 (Maximum, Minimum, Positiv- und Negativteil)** Seien  $f, g: M \rightarrow \mathbb{R}$  Funktionen auf einer Menge  $X$ . Dann definiert man das Maximum und Minimum von  $f$  und  $g$  punktweise durch

$$\max(f, g)(x) = \max\{f(x), g(x)\} \quad \text{und} \quad \min(f, g)(x) = \min\{f(x), g(x)\} \quad (3.9.8)$$

für  $x \in X$ . Weiter definiert man den Positivteil  $f_+$  und den Negativteil  $f_-$  von  $f$  durch

$$f_+ = \max(f, 0) \quad \text{und} \quad f_- = -\min(f, 0). \quad (3.9.9)$$

i.) Zeigen Sie, dass  $f_{\pm} \geq 0$  und dass

$$|f| = f_+ + f_- \quad \text{und} \quad f = f_+ - f_- \quad (3.9.10)$$

ebenso wie

$$f_{\pm} = \frac{1}{2}(|f| \pm f) \quad (3.9.11)$$

gilt.

ii.) Zeigen Sie

$$\max(f, g) = \frac{1}{2}(f + g + |f - g|) \quad \text{und} \quad \min(f, g) = \frac{1}{2}(f + g - |f - g|). \quad (3.9.12)$$

iii.) Seien nun  $f, g \in \mathcal{C}(X, \mathbb{R})$  stetige Funktionen auf einer nicht-leeren Teilmenge  $X$  der reellen oder komplexen Zahlen. Folgern Sie, dass dann auch  $f_+$ ,  $f_-$ ,  $\max(f, g)$  und  $\min(f, g)$  stetige Funktionen sind.

**Übung 3.9.23 (Verkettung stetiger Funktionen an einem Punkt)** Seien wieder  $X$  und  $Y$  nicht-leere Teilmengen der reellen oder komplexen Zahlen und  $f: X \rightarrow Y$  sowie  $g: Y \rightarrow \mathbb{K}$ . Sei  $f$  stetig bei  $x_0 \in X$  und  $g$  stetig bei  $y_0 = f(x_0) \in Y$ .

i.) Zeigen Sie analog zu Satz 3.5.3 unter Verwendung geeigneter Umgebungen, dass  $g \circ f$  stetig bei  $x_0$  ist.

ii.) Zeigen Sie die Stetigkeit von  $f \circ g$  bei  $x_0$  unter Verwendung des Folgenkriteriums für Stetigkeit.

iii.) Zeigen Sie die Stetigkeit von  $f \circ g$  bei  $x_0$  unter Verwendung des  $\epsilon\delta$ -Kriteriums für Stetigkeit.

**Übung 3.9.24 (Urbilder von kompakten Mengen)** Finden Sie einfache Beispiele für stetige Funktionen mit der Eigenschaft, dass es kompakte Teilmengen des Bildbereichs gibt, deren Urbilder nicht kompakt sind.

**Übung 3.9.25 (Beschränkte Funktionen)** In dieser Übung verallgemeinern wir die Aussagen zu beschränkten Folgen aus Proposition 2.3.2 für beschränkte Funktionen auf einer beliebigen Menge  $X$ . Sei  $X$  also eine nicht-leere Menge und  $\mathcal{B}(X, \mathbb{C}) \subseteq \text{Abb}(X, \mathbb{C})$  die beschränkten  $\mathbb{C}$ -wertigen Funktionen auf  $X$ . Sei weiter

$$\|f\|_{\infty} = \sup\{|f(x)| \mid x \in X\} \quad (3.9.13)$$

die Supremumsnorm von  $f \in \mathcal{B}(X, \mathbb{C})$ .

- i.) Zeigen Sie, dass genau dann  $\|f\|_{\infty} = 0$  gilt, wenn  $f = 0$  gilt.
- ii.) Zeigen Sie für  $f, g \in \mathcal{B}(X, \mathbb{C})$ , dass  $f + g \in \mathcal{B}(X, \mathbb{C})$  gilt. Zeigen Sie dazu  $\|f + g\|_{\infty} \leq \|f\|_{\infty} + \|g\|_{\infty}$ .
- iii.) Finden Sie zwei beschränkte Funktionen  $f$  und  $g$  mit  $\|f + g\|_{\infty} < \|f\|_{\infty} + \|g\|_{\infty}$ .
- iv.) Zeigen Sie für  $f \in \mathcal{B}(X, \mathbb{C})$  und  $z \in \mathbb{C}$ , dass auch  $zf \in \mathcal{B}(X, \mathbb{C})$  gilt. Zeigen Sie dazu  $\|zf\|_{\infty} = |z| \|f\|_{\infty}$ .
- v.) Folgern Sie, dass  $\mathcal{B}(X, \mathbb{C})$  ein komplexer Vektorraum ist.
- vi.) Zeigen Sie, dass für  $f, g \in \mathcal{B}(X, \mathbb{C})$  auch  $fg \in \mathcal{B}(X, \mathbb{C})$  gilt. Zeigen Sie dazu  $\|fg\|_{\infty} \leq \|f\|_{\infty} \|g\|_{\infty}$ .
- vii.) Zeigen Sie  $1 \in \mathcal{B}(X, \mathbb{C})$ , und bestimmen Sie die Supremumsnorm der Funktion 1.
- viii.) Zeigen Sie, dass genau dann  $f \in \mathcal{B}(X, \mathbb{C})$  gilt, wenn  $\bar{f} \in \mathcal{B}(X, \mathbb{C})$  gilt. Zeigen Sie dann  $\|\bar{f}f\|_{\infty} = \|f\|_{\infty}^2$  und  $\|\bar{f}\|_{\infty} = \|f\|_{\infty}$  für alle  $f \in \mathcal{B}(X, \mathbb{C})$ .

Die Eigenschaften i.), ii.) und iv.) der Supremumsnorm besagen, dass  $\|\cdot\|_{\infty}$  wirklich eine *Norm* ist. Diese Begriffe werden wir in Abschnitt 7.1 noch eingehend studieren. Die Eigenschaft vi.) nennt man auch die *Submultiplikativität* der Supremumsnorm. Damit wird  $\mathcal{B}(X, \mathbb{C})$  unter punktweisen Produkten abgeschlossen und somit eine Unter algebra von allen Funktionen  $\text{Abb}(X, \mathbb{C})$ . Die letzte Eigenschaft viii.) nennt man auch die *C\*-Eigenschaft* der Supremumsnorm.

**Übung 3.9.26 (Nullstellen stetiger Funktionen)** Sei  $f: X \rightarrow \mathbb{K}$  eine stetige Funktion. Sei  $K \subseteq X$  eine kompakte Teilmenge mit der Eigenschaft, dass  $f|_K$  keine Nullstelle hat. Zeigen Sie, dass es dann ein  $\epsilon > 0$  gibt, so dass  $|f|(x) \geq \epsilon$  für alle  $x \in K$  gilt. Zeigen Sie in Beispielen, dass ohne die Voraussetzung der Kompaktheit die Aussage im allgemeinen falsch ist.

**Übung 3.9.27 (Koinzidenzmenge)** Seien  $f, g: X \rightarrow \mathbb{K}$  zwei stetige Funktionen. Zeigen Sie, dass dann die *Koinzidenzmenge*

$$C(f, g) = \{x \in X \mid f(x) = g(x)\} \subseteq X \quad (3.9.14)$$

eine abgeschlossene Teilmenge von  $X$  in  $X$  ist. Folgern Sie, dass die Nullstellenmenge einer stetigen Funktion abgeschlossen in ihrem Definitionsbereich ist.

Hinweis: Hier können Sie entweder mit Folgenstetigkeit argumentieren oder einfacher mit Urbildern von abgeschlossenen Mengen.

**Übung 3.9.28 (Kompaktheit der Einheitskreislinie)** Zeigen Sie, dass die Einheitskreislinie  $\mathbb{S}^1 \subseteq \mathbb{C}$  kompakt ist.

Hinweis: Verwenden Sie Satz 3.4.11, Übung 3.9.27 und eine geeignete stetige Funktion  $f$ , deren Nullstellenmenge gerade  $\mathbb{S}^1$  ist.

**Übung 3.9.29 (Homöomorphismen zwischen Intervallen)**

**Übung 3.9.30 (Gleichmäßige Stetigkeit)** Formulieren Sie die Definition von gleichmäßiger Stetigkeit in Quantorenschreibweise und vergleichen Sie dies mit der Definition von Stetigkeit. Finden Sie auch die Negation von gleichmäßiger Stetigkeit.

**Übung 3.9.31 (Formale Potenzreihen)**

# Kapitel 4

## e und $\pi$

In diesem kurzen Kapitel werden wir nicht nur die verbleibenden zwei wichtigsten Zahlen der Mathematik,  $e$  und  $\pi$ , kennenlernen und ihre Beziehungen, wie beispielsweise die schönste Gleichung der Mathematik

$$e^{i\pi} + 1 = 0, \quad (4.1)$$

zu den anderen drei wichtigsten Zahlen der Mathematik aufzeigen, sondern auch die Exponentialfunktion sowie die trigonometrischen Funktionen definieren. Erste einfache Eigenschaften dieser Funktionen reproduzieren dann heuristische Ergebnisse aus der Schule.

### 4.1 $e$

Die Euler-Zahl  $e$  lässt sich auf viele Weisen als Grenzwert definieren. Die vielleicht einfachste ist mit Hilfe der Exponentialreihe:

**Definition 4.1.1 (Exponentialreihe)** Für  $z \in \mathbb{C}$  definiert man die Exponentialreihe durch

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}. \quad (4.1.1)$$

**Satz 4.1.2 (Exponentialfunktion)** Seien  $z, w \in \mathbb{C}$ .

i.) Die Exponentialreihe  $\exp(z)$  konvergiert absolut und lokal gleichmäßig.

ii.) Die Exponentialfunktion

$$\exp: \mathbb{C} \ni z \mapsto \exp(z) \in \mathbb{C} \quad (4.1.2)$$

ist stetig.

iii.) Es gilt

$$\exp(z) \exp(w) = \exp(z + w) \quad (4.1.3)$$

und

$$\exp(0) = 1. \quad (4.1.4)$$

iv.) Es gilt  $\exp(z) \neq 0$  und

$$\exp(-z) = \frac{1}{\exp(z)}. \quad (4.1.5)$$

v.) Es gilt

$$\overline{\exp(z)} = \exp(\bar{z}). \quad (4.1.6)$$

vi.) Für reelle  $x \in \mathbb{R}$  gilt

$$\exp(x) > 0 \quad (4.1.7)$$

und

$$|\exp(ix)| = 1. \quad (4.1.8)$$

vii.) Für reelle  $x, y \in \mathbb{R}$  gilt

$$\exp(x)^y = \exp(xy). \quad (4.1.9)$$

viii.) Es gilt die Restgliedabschätzung

$$\left| \exp(z) - \sum_{n=0}^N \frac{z^n}{n!} \right| \leq \frac{2|z|^{N+1}}{(N+1)!} \quad (4.1.10)$$

für alle  $N \in \mathbb{N}$  und  $|z| \leq \frac{1}{2}(N+1)$ .

BEWEIS: Die absolute Konvergenz der Exponentialreihe für alle  $z \in \mathbb{C}$  hatten wir in Beispiel 2.7.7, *iv.*), bereits aus dem Quotientenkriterium gewonnen. Jenes Beispiel enthält eine etwas allgemeinere Situation, man erhält die Exponentialreihe für  $\epsilon = 1$ . Aufgefasst als komplexe Potenzreihe hat die Exponentialreihe also Konvergenzradius  $R = +\infty$ , siehe auch Beispiel 3.8.6, *iv.*). Nach Satz 3.8.8 konvergiert sie daher innerhalb ihres Konvergenzkreises lokal gleichmäßig gegen eine stetige Funktion, in unserem Fall also auf ganz  $\mathbb{C}$ . Entscheidend ist nun der dritte Teil. Der Wert  $\exp(0) = 1$  ist klar. Wir zeigen (4.1.3) durch Reihenmultiplikation mit Hilfe des Cauchy-Produkts. Dies dürfen wir nach Satz 2.7.15 beziehungsweise Korollar 2.7.16 in diesem Fall tatsächlich verwenden, da nach *i.*) absolute Konvergenz vorliegt. Es gilt daher

$$\begin{aligned} \exp(z) \exp(w) &= \left( \sum_{n=0}^{\infty} \frac{z^n}{n!} \right) \left( \sum_{m=0}^{\infty} \frac{w^m}{m!} \right) \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^n \frac{z^n}{n!} \frac{w^m}{m!} \frac{1}{(n-m)!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{m=0}^n \frac{n!}{m!(n-m)!} z^m w^{n-m} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n \\ &= \exp(z+w), \end{aligned}$$

wobei wir den Binomischen Lehrsatz verwendet haben. Damit folgt insbesondere  $\exp(z) \exp(-z) = \exp(z-z) = \exp(0) = 1$ , womit  $\exp(z) \neq 0$  gezeigt ist, ebenso wie (4.1.5). Proposition 2.6.5, *ii.*), liefert dann auch

$$\overline{\exp(z)} = \overline{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \sum_{n=0}^{\infty} \frac{\overline{z^n}}{n!} = \sum_{n=0}^{\infty} \frac{(\overline{z})^n}{n!} = \exp(\overline{z}),$$

da die Koeffizienten  $\frac{1}{n!}$  der Exponentialreihe ja reell sind. Es folgt  $\exp(x) \in \mathbb{R}$  für alle  $x \in \mathbb{R}$ . Mit  $\exp(\frac{x}{2}) \exp(\frac{x}{2}) = \exp(\frac{x}{2} + \frac{x}{2}) = \exp(x)$  sehen wir, dass  $\exp(x)$  ein Quadrat einer reellen Zahl  $\exp(\frac{x}{2})$  ist, und somit (4.1.7) gilt. Für (4.1.8) rechnen wir nach, dass

$$\begin{aligned} \exp(ix) \overline{\exp(ix)} &\stackrel{(4.1.7)}{=} \exp(ix) \exp(-ix) \\ &\stackrel{(4.1.3)}{=} \exp(ix - ix) \end{aligned}$$

$$= \exp(0) \\ \stackrel{(4.1.4)}{=} 1.$$

Daher folgt (4.1.8). Die Funktionalgleichung (4.1.3) liefert für  $n \in \mathbb{N}$  zunächst

$$\exp(z)^n = \underbrace{\exp(z) \cdots \exp(z)}_{n\text{-mal}} = \exp(z + \cdots + z) = \exp(nz),$$

noch für beliebige  $z \in \mathbb{C}$ . Es gilt also

$$\exp(z)^n = \exp(nz). \quad (*)$$

Für reelle  $x \in \mathbb{R}$  ist  $\exp(x) > 0$  und wir können die  $n$ -te Wurzel bilden. Es gilt nach (\*)

$$\exp\left(\frac{1}{n}x\right)^n = \exp(x)$$

und  $\exp(\frac{1}{n}x) > 0$ . Nach der Eindeutigkeit der positiven Wurzel folgt

$$\exp\left(\frac{1}{n}x\right) = \sqrt[n]{\exp(x)}.$$

Für  $n \in \mathbb{N}$  gilt schließlich

$$\exp(z)^{-n} = \frac{1}{\exp(z)^n} = \frac{1}{\exp(nz)} = \exp(-nz)$$

nach (\*) und (4.1.5). Insgesamt gilt daher für  $n \in \mathbb{Z}, m \in \mathbb{N}$  und  $x \in \mathbb{R}$

$$\exp(x)^{\frac{n}{m}} = \sqrt[m]{\exp(x)^n} = \sqrt[m]{\exp(nx)} = \exp\left(\frac{n}{m}x\right),$$

womit

$$\exp(x)^q = \exp(qx)$$

für alle  $q \in \mathbb{Q}$  und  $x \in \mathbb{R}$  gezeigt ist. Da  $\exp$  nach ii.) stetig ist, ist die Abbildung

$$\mathbb{R} \ni y \mapsto \exp(yx) \in \mathbb{R} \quad (**)$$

ebenfalls stetig, wobei wir  $x \in \mathbb{R}$  als fest gewählt betrachten. Ebenso ist die Funktion

$$\mathbb{R} \ni y \mapsto \exp(x)^y \in \mathbb{R} \quad (\odot)$$

nach Beispiel 3.3.5 stetig, da ja  $\exp(x) > 0$  gilt. Letztere stimmt mit (\*\*) für  $y = q \in \mathbb{Q}$  überein. Nach Satz 3.5.19 erhalten wir durch (⊙) eine stetige Fortsetzung von (\*\*), welche dann eindeutig bestimmt ist auf  $\mathbb{Q}^{\text{cl}} = \mathbb{R}$ . Dies zeigt (4.1.9). Schließlich betrachten wir die Restgliedabschätzung

$$\begin{aligned} \left| \exp(z) - \sum_{n=0}^N \frac{z^n}{n!} \right| &= \left| \sum_{n=N+1}^{\infty} \frac{z^n}{n!} \right| \\ &\leq \sum_{n=N+1}^{\infty} \frac{|z|^n}{n!} \\ &= \sum_{n=N+1}^{\infty} \frac{|z|^{N+1}}{(N+1)!} \frac{|z|^{n-N-1}}{n(n-1) \cdots (N+2)} \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{n=N+1}^{\infty} \frac{|z|^{N+1}}{(N+1)!} \left(\frac{1}{2}\right)^{n-N-1} \\
&= \frac{|z|^{N+1}}{(N+1)!} \sum_{n=N+1}^{\infty} \left(\frac{1}{2}\right)^{n-N-1} \\
&= \frac{|z|^{N+1}}{(N+1)!} \sum_{n=0}^{\infty} \frac{1}{2^n} \\
&= \frac{2|z|^{N+1}}{(N+1)!},
\end{aligned}$$

wobei wir  $|z| < \frac{1}{2}(N+1)$  und die geometrische Reihe benutzt haben.  $\square$

Mit der Exponentialfunktion  $\exp$  haben wir eine der wichtigsten *transzendenten* Funktionen der Mathematik kennengelernt. Transzendent heißt in diesem Zusammenhang, dass  $\exp$  weder ein Polynom noch eine rationale oder gebrochen rationale Funktion ist. Die Rechenregel (4.1.9) liefert insbesondere für  $x = 1$  die Gleichung

$$\exp(1)^y = \exp(y) \quad (4.1.11)$$

für alle reellen  $y \in \mathbb{R}$ . Das legt nahe, der Zahl  $\exp(1) \in \mathbb{R}$  größere Aufmerksamkeit zu schenken:

**Definition 4.1.3 (Eulerzahl e)** Die Eulerzahl  $e$  ist definiert als

$$e = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}. \quad (4.1.12)$$

Die Exponentialreihe  $\exp$  ist also für reelle Argumente  $x \in \mathbb{R}$  nichts anderes als die Exponentialfunktion

$$\exp(x) = e^x \quad (4.1.13)$$

zur Basis  $e$ . Die wichtige Beobachtung ist nun, dass  $\exp(z)$  für *komplexe* Zahlen  $z \in \mathbb{C}$  ebenso durch die Reihe definiert ist, während wir eine Definition von  $a^z$  für reelles  $a > 0$  und komplexes  $z \in \mathbb{C}$  bislang nicht finden konnten. Der Zugang aus Abschnitt 1.6 basiert ja entscheidend auf der Verwendung von Suprema und damit auf der Anordnung von  $\mathbb{R}$ . Für  $z \in \mathbb{C}$  stehen diese Argumente daher nicht zur Verfügung.

Mit der Exponentialreihe haben wir nun zumindest für die Basis  $e$  die Möglichkeit auch komplexe Exponenten zuzulassen. Wir verwenden (4.1.13) als Definition und schreiben

$$e^z = \exp(z) \quad (4.1.14)$$

für alle  $z \in \mathbb{C}$ . Für reelle  $z = x \in \mathbb{R}$  ist dies mit der Gleichheit (4.1.13) konsistent und stellt somit eine gute Verallgemeinerung dar. Die üblichen Rechenregeln bleiben dann bestehen. Insbesondere gilt also

$$e^z e^w = e^{z+w} \quad (4.1.15)$$

jetzt für  $z, w \in \mathbb{C}$ . Allerdings haben wir momentan noch keine Möglichkeit, für allgemeinere Basen als  $e$  eine Exponentialfunktion  $a^z$  mit komplexen Exponenten zu definieren. Dies werden wir noch schrittweise erreichen müssen.

Wir schließen diesen Abschnitt mit einigen weiteren Eigenschaften zur Eulerschen Zahl  $e$ :

**Proposition 4.1.4** Für die Eulersche Zahl  $e$  gilt

$$2\frac{1}{2} < e < 3. \quad (4.1.16)$$



BEWEIS: Da die definierende Reihe für  $e$  aus positiven Summanden besteht, erhalten wir leicht untere Abschätzungen. Für jedes  $N \in \mathbb{N}$  gilt

$$\sum_{n=0}^N \frac{1}{n!} < \sum_{n=0}^{\infty} \frac{1}{n!} = e.$$

Verwendet man nun beispielsweise  $N = 2$ , so ist die linke Seite

$$\sum_{n=0}^2 \frac{1}{n!} = 1 + 1 + \frac{1}{2} = 2\frac{1}{2},$$

was die linke Ungleichung in (4.1.16) erklärt. Die obere Schranke erhalten wir aus der Restgliedabschätzung in Satz 4.1.2, *viii.*), für  $N = 1$  und  $z = 1$ . Es gilt

$$\begin{aligned} e - 2 &= \exp(1) - \sum_{n=0}^1 \frac{1}{n!} \\ &\stackrel{(a)}{=} \left| \exp(1) - \sum_{n=0}^1 \frac{1}{n!} \right| \\ &\leq \frac{2 \cdot 1^2}{(1+1)!} \\ &= 1, \end{aligned}$$

wobei wir in (a) benutzen, dass die Exponentialreihe  $\exp(1)$  nur positive Summanden besitzt.  $\square$

Die Konvergenz der Reihe (4.1.12) ist tatsächlich sehr gut, wie die Fehlerabschätzung aus Satz 4.1.2, *viii.*) zeigt. Das enorm schnelle Wachstum von  $(N+1)!$  lässt den Fehler sehr schnell klein werden. Die ersten Dezimalstellen von  $e$  lauten

$$e = 2,7182818284590 \dots \quad (4.1.17)$$

Mehr Stellen finden sich beispielsweise in der Online-Datenbank OEIS [4] zu wichtigen Folgen der Mathematik.

**Bemerkung 4.1.5 (Transzendenz von  $e$ )** Es ist relativ einfach zu sehen, dass  $e$  keine rationale Zahl ist, siehe Übung 4.6.1. Deutlich schwieriger ist der Nachweis, dass  $e$  eine transzendente Zahl ist, siehe nochmals Bemerkung 2.1.6. Der Beweis (erstmal von Charles Hermite) dieser Tatsache gilt als Glanzstück der analytischen Zahlentheorie und übersteigt unsere momentanen Möglichkeiten deutlich.

## 4.2 Der natürliche Logarithmus

Die Exponentialabbildung  $\exp$  ist nicht nur stetig sondern auch strikt monoton wachsend, wenn wir nur reelle Argumente betrachten. Daher besitzt sie eine Umkehrabbildung, den natürlichen Logarithmus. Wir zeigen dies nicht nur für die Exponentialabbildung zur Basis  $e$  sondern etwas allgemeiner:

**Proposition 4.2.1** Sei  $a > 1$ .

*i.) Die Funktion*

$$\mathbb{R} \ni x \mapsto a^x \in \mathbb{R} \quad (4.2.1)$$

*ist strikt monoton wachsend.*

ii.) Es gilt

$$\lim_{x \rightarrow -\infty} a^x = 0 \quad \text{und} \quad \lim_{x \rightarrow +\infty} a^x = \infty. \quad (4.2.2)$$

iii.) Die Funktion (4.2.1) ist bijektiv.

iv.) Für  $p > 0$  gilt

$$\lim_{x \rightarrow \infty} \frac{x^p}{a^x} = 0. \quad (4.2.3)$$

BEWEIS: Für  $a > 1$  hatten wir die strikte Monotonie von (4.2.1) bereits in den Rechenregeln zum Exponenzieren in Satz 1.6.8, vii.), gesehen. Ebenso sind die Grenzwerte aus Beispiel 2.2.8, ii.), sowie Übung 3.9.9 bekannt. Der dritte Teil folgt dann direkt, da wegen (4.2.2) jede positive reelle Zahl  $y \in \mathbb{R}^+$  durch Funktionswerte  $a^x$  sowohl unterschritten, als auch überschritten wird. Da das Bild des offenen Intervalls  $\mathbb{R}$  aufgrund der Stetigkeit von (4.2.1) wieder ein offenes Intervall ist, folgt, dass  $y$  im Bild von (4.2.1) liegt. Daher ist (4.2.1) surjektiv. Die Injektivität ist klar aufgrund der strikten Monotonie. Die Idee zum letzten Teil ist aus Beispiel 2.3.12, iii.). Wie dort schreiben wir zunächst

$$\frac{x^p}{a^x} = \left( \frac{x}{a^{\frac{x}{p}}} \right)^p = \left( \frac{x}{(a^{\frac{1}{p}})^x} \right)^p$$

und nutzen, dass für  $a > 1$  auch  $a^{\frac{1}{p}} > 1$  gilt. Es genügt daher  $p = 1$  zu betrachten. Wieder schreiben wir  $a = 1 + b$  mit  $b > 0$ . Sei weiter  $n \in \mathbb{N}$  so gewählt, dass  $n \leq x < n + 1$ . Dann gilt wie in Beispiel 2.3.12, iii.),

$$a^x \geq a^n \geq \frac{1}{2}n(n-1)b^2 \geq \frac{1}{2}x(x-1)b^2$$

und damit

$$\frac{x}{a^x} \leq \frac{2}{(x-a)b^2},$$

womit  $\lim_{x \rightarrow \infty} \frac{x}{a^x} = 0$  folgt. Die Stetigkeit der  $p$ -ten Potenz bei  $x = 0$  zeigt dann auch

$$\lim_{x \rightarrow \infty} \frac{x^p}{(a^{\frac{1}{p}})^x} = \left( \lim_{x \rightarrow \infty} \frac{x}{a^x} \right)^p = 0. \quad \square$$

Ist  $0 < a < 1$ , so gilt  $a^x = \left(\frac{1}{a}\right)^{-x}$  womit die Abbildung  $x \mapsto a^x$  strikt monoton fallend ist und die Rolle der beiden Grenzwerte für  $x \rightarrow \pm\infty$  vertauscht wird. Die Bijektivität bleibt trotzdem gültig. Ist schließlich  $a = 1$ , so ist  $x \mapsto 1^x = 1$  konstant und folglich nicht bijektiv. Die Bijektivität erlaubt es nun, dass Exponenzieren umzukehren:

**Definition 4.2.2 (Logarithmus)** Sei  $a \in \mathbb{R}^+ \setminus \{1\}$ .

i.) Die Umkehrabbildung zu (4.2.1) heißt  $a$ -Logarithmus und wird mit

$$\log_a: \mathbb{R}^+ \ni x \mapsto \log_a(x) \in \mathbb{R} \quad (4.2.4)$$

bezeichnet.

ii.) Für  $a = e$  schreiben wir einfach

$$\ln = \log = \log_e, \quad (4.2.5)$$

und nennen dies den natürlichen Logarithmus.

In der Mathematik werden wir hauptsächlich den natürlichen Logarithmus verwenden und oft einfach von dem Logarithmus sprechen. In technischeren Anwendungen ist die Basis  $a = 10$  von Bedeutung, womit man dann auch vom *dekadischen Logarithmus*  $\log_{10}$  spricht. Zur Vorsicht sei angemerkt, dass in der Literatur das Symbol  $\log$  auch für  $\log_{10}$  stehen kann. In der Informatik spielt der dyadische Logarithmus  $\log_2$  eine wichtige Rolle.

**Korollar 4.2.3** Sei  $a \in \mathbb{R}^+ \setminus \{1\}$ .

- i.) Der  $a$ -Logarithmus ist strikt monoton wachend für  $a > 1$  und strikt monoton fallend für  $0 < a < 1$ .  
 ii.) Der  $a$ -Logarithmus ist bijektiv und erfüllt

$$a^{\log_a(x)} = x \quad (4.2.6)$$

für alle  $x \in \mathbb{R}^+$  sowie

$$\log_a(a^x) = x \quad (4.2.7)$$

für alle  $x \in \mathbb{R}$ . Insbesondere gilt

$$\log_a(1) = 0 \quad \text{und} \quad \log_a(a) = 1. \quad (4.2.8)$$

- iii.) Der  $a$ -Logarithmus ist stetig.

BEWEIS: Der erste Teil folgt aus den Monotonieeigenschaften der Funktion  $x \mapsto a^x$  und der Tatsache, dass  $\log_a$  die entsprechende Umkehrfunktion ist, was in Teil ii.) nochmals ausgeschrieben ist. Dies erklärt auch die speziellen Werte. Nach Proposition 3.5.16 ist  $\log_a$  stetig als Umkehrfunktion der strikt monotonen und stetigen Funktion  $x \mapsto a^x$ .  $\square$

Aus den Rechenregeln für das Potenzieren erhalten wir sofort einige Rechenregeln für das Logarithmieren.

**Proposition 4.2.4** Sei  $a \in \mathbb{R} \setminus \{1\}$ .

- i.) Sei  $x, y \in \mathbb{R}^+$ , dann gilt

$$\log_a(xy) = \log_a(x) + \log_a(y). \quad (4.2.9)$$

- ii.) Sei  $x \in \mathbb{R}^+$  und  $y \in \mathbb{R}$ , dann gilt

$$\log_a(x^y) = y \log_a(x). \quad (4.2.10)$$

- iii.) Für  $x \in \mathbb{R}^+$  gilt

$$\log(x) = \log_a(x) \log(a). \quad (4.2.11)$$

BEWEIS: Zunächst ist  $xy \in \mathbb{R}^+$ , womit die linke Seite in (4.2.9) überhaupt definiert ist. Wir bilden die Potenzen

$$a^{\log_a(xy)} = xy = a^{\log_a(x)} \cdot a^{\log_a(y)} = a^{\log_a(x) + \log_a(y)}.$$

Da die Abbildung  $x \mapsto a^x$  injektiv ist, folgt (4.2.9). Genauso erhält man (4.2.10), denn es gilt

$$a^{\log_a(x^y)} = x^y = (a^{\log_a(x)})^y = a^{y \log_a(x)}$$

nach den Rechenregeln für das Potenzieren. Damit folgt erneut aus der Injektivität des Exponenzierens die Behauptung (4.2.10). Der dritte Teil erlaubt es nun den  $a$ -Logarithmus durch den natürlichen Logarithmus zu ersetzen. Es gilt für alle  $x \in \mathbb{R}^+$  zunächst  $a = e^{\log(a)}$  und daher

$$e^{\log(x)} = x = a^{\log_a(x)} = (e^{\log(a)})^{\log_a(x)} = e^{\log(a) \log_a(x)}. \quad (4.2.12)$$

Entsprechend folgt wieder die Gleichheit der Exponenten (4.2.11).  $\square$

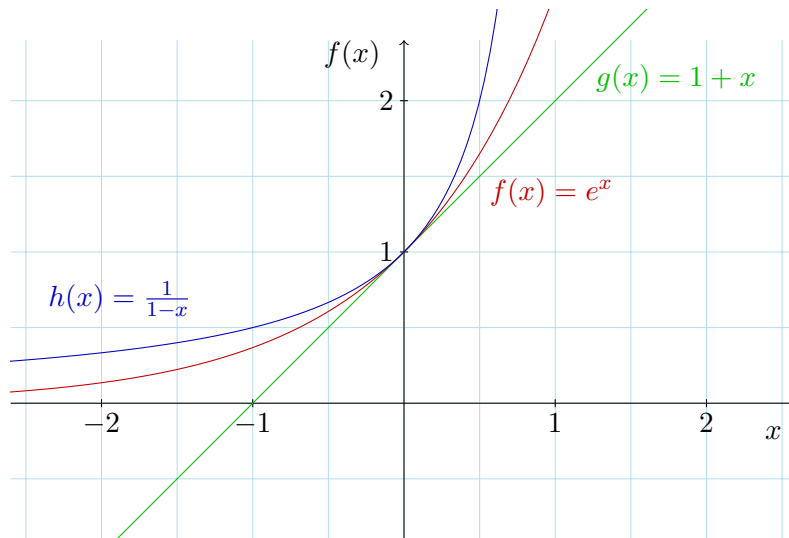


Abbildung 4.1: Der Graph der Exponentialfunktion sowie der Geraden  $g(x) = 1 + x$  und der Hyperbel  $h(x) = \frac{1}{1-x}$  für  $x < 1$ .

Insbesondere die letzte Rechenregel zeigt, dass wir alle  $a$ -Logarithmen durch natürliche Logarithmen ersetzen können, indem wir die einfache Umrechnungsformel (4.2.11) verwenden. Wir werden daher im Folgenden nur noch den natürlichen Logarithmus verwenden und gegebenenfalls Ergebnisse für diesen mittels (4.2.11) auf alle  $a$ -Logarithmen übertragen. Wir wollen nun einige weitere analytische Eigenschaften des Logarithmus aus entsprechenden Eigenschaften der Exponentialfunktion herleiten. Hierzu benötigen wir folgendes einfaches Resultat, welches auch von allgemeiner Bedeutung ist:

**Proposition 4.2.5** Sei  $I \subseteq \mathbb{R}$  ein Intervall und seien  $f, g: I \rightarrow \mathbb{R}$  stetige, strikt monoton wachsende Funktionen mit

$$f(x) < g(x) \quad (4.2.13)$$

für alle  $x \in I$ . Für  $y \in f(I) \cap g(I)$  gilt dann

$$g^{-1}(y) < f^{-1}(y) \quad (4.2.14)$$

für die Umkehrfunktion von  $f$  und  $g$ . Analoge Aussagen gelten für  $f(x) \leq g(x)$ .

BEWEIS: Zunächst wissen wir, dass die Umkehrfunktionen existieren, da  $f$  und  $g$  strikt monoton wachsend sind und entsprechend Bijektionen auf ihre Bilder  $f(I), g(I) \subseteq \mathbb{R}$  liefern. Die Stetigkeit zeigt, dass  $f(I)$  und  $g(I)$  wieder Intervalle sind, womit  $f(I) \cap g(I)$  entweder leer ist, ein Fall in dem wir nichts zu zeigen haben, oder wieder ein Intervall liefert. Sei also  $x \in f(I) \cap g(I)$  im Definitionsbereich von sowohl  $f^{-1}$  als auch  $g^{-1}$ . Sei  $x = g^{-1}(y) \in I$ . Dann gilt nach (4.2.13)

$$f(x) < g(x) = g(g^{-1}(y)) = y.$$

Da mit  $f$  auch  $f^{-1}$  strikt monoton wachsend ist, folgt

$$x = f^{-1}(f(x)) < f^{-1}(y),$$

was  $g^{-1}(y) > f^{-1}(y)$  zeigt. □

Dieser Vergleich von Funktionen und ihrer Umkehrfunktionen hat viele Varianten und Spielarten, siehe auch Übung ??

Wir wenden dieses Resultat nun auf die Exponentialfunktion an, um Ungleichungen für den Logarithmus zu erhalten:

**Proposition 4.2.6** *i.) Für alle  $x \in \mathbb{R}$  gilt die Ungleichung*

$$e^x \geq 1 + x, \quad (4.2.15)$$

*wobei Gleichheit nur für  $x = 0$  erreicht ist.*

*ii.) Für alle  $x < 1$  gilt die Ungleichung*

$$e^x \leq \frac{1}{1-x}, \quad (4.2.16)$$

*wobei Gleichheit nur für  $x = 0$  erreicht wird.*

*iii.) Für alle  $x > 0$  gilt die Ungleichung*

$$\log(x) \leq x - 1, \quad (4.2.17)$$

*wobei Gleichheit nur für  $x = 1$  erreicht wird.*

*iv.) Für alle  $x > 0$  gilt die Ungleichung*

$$\log(x) \geq 1 - \frac{1}{x}, \quad (4.2.18)$$

*wobei Gleichheit nur für  $x = 1$  erreicht wird.*

BEWEIS: Für den ersten Teil betrachten wir zunächst  $x \geq 0$ . Hier gilt

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \sum_{n=2}^{\infty} \frac{x^n}{n!} \geq 1 + x,$$

wobei für  $x \neq 0$  der zusätzliche Term strikt positiv ist, also eine echte Ungleichung vorliegt. Für  $x \leq -1$  ist  $1 + x \leq 0$  und daher  $e^x > 1 - x$  trivialerweise erfüllt, da  $e^x > 0$  für alle  $x \in \mathbb{R}$ . Für  $-1 < x < 0$  betrachten wir

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \sum_{n=4}^{\infty} \frac{x^n}{n!} = 1 + x + \underbrace{\frac{x^2}{2} \left(1 + \frac{x}{3}\right)}_a + \underbrace{\sum_{n=4}^{\infty} \frac{x^n}{n!}}_b.$$

Da  $-1 < x < 0$ , gilt  $\frac{2}{3} < 1 + \frac{x}{3} < 1$ . Folglich ist  $a > \frac{x^2}{3} > 0$ . Für  $b$  erhalten wir wegen  $|x| < 1$

$$|b| = \left| x^2 \sum_{n=4}^{\infty} \frac{x^{n-2}}{n!} \right| < x^2 \sum_{n=4}^{\infty} \frac{1}{n!}.$$

Die verbleibende Reihe ist ein Endstück der definierenden Reihe zu  $e$ , für die wir die Restgliedabschätzung (4.1.10) aus Satz 4.1.2, *viii.*), zum Einsatz bringen können. Es gilt mit  $z = 1$  in (4.1.10)

$$\sum_{n=4}^{\infty} \frac{1}{n!} \leq \frac{2}{4!} = \frac{1}{12}.$$

Damit gilt also  $|b| < \frac{x^2}{12}$  und somit  $a + b > \frac{x^2}{3} - \frac{x^2}{12} = \frac{1}{4}x^2 > 0$ . Dies zeigt die strikte Ungleichung (4.2.15) auch in diesem Bereich, was den Beweis von *i.*) abschließt. Alternativ kann man für  $-1 < x < 0$  auch

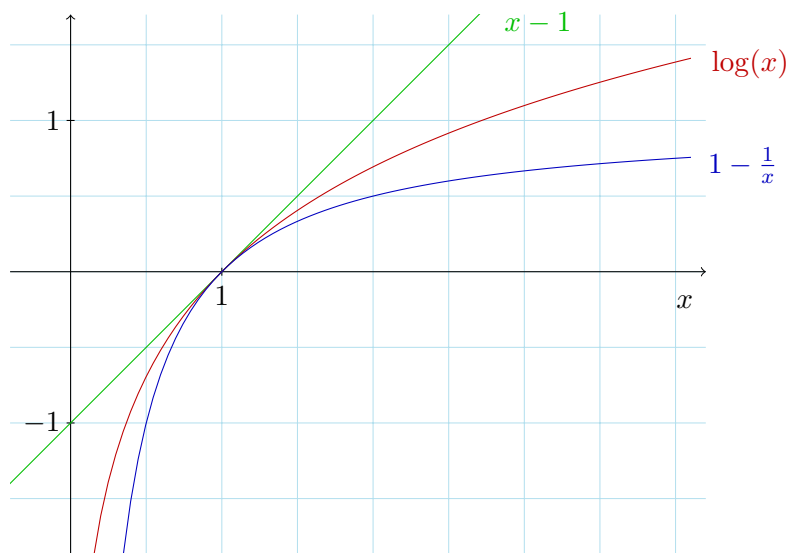


Abbildung 4.2: Die Graphen von  $\log$  sowie den beiden Funktionen aus den Ungleichungen in Proposition 4.2.6.

mit der Restgliedabschätzung aus dem Leibniz-Kriterium für alternierende Reihen argumentieren, siehe Übung ?? . Die zweite Ungleichung erhalten wir für  $0 < x < 1$  durch

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} < \sum_{n=0}^{\infty} x^n = \frac{1}{1-x},$$

unter Verwendung der geometrischen Reihe. Für  $x = 0$  gilt Gleichheit. Für  $x < 0$  gilt

$$e^x = \frac{1}{e^{|x|}} < \frac{1}{1+|x|} = \frac{1}{1-x}$$

unter Verwendung von (4.2.15). Die Ungleichungen für den Logarithmus erhalten wir aus Proposition 4.2.6. Für (4.2.17) verwenden wir  $f(x) = 1 + x$  und  $g(x) = e^x$ , womit nach (4.2.15) sofort (4.2.17) folgt, da  $f^{-1}(y) = y - 1$  die Umkehrfunktion zu  $f$  ist. Für *iv.*) verwenden wir  $g(x) = \frac{1}{1-x}$  und entsprechend  $g^{-1}(y) = 1 - \frac{1}{y}$  sowie  $f(x) = e^x$  mit  $f^{-1}(y) = \log(y)$ . Man macht sich schnell klar, dass Gleichheiten in (4.2.17) beziehungsweise (4.2.18) nur dort gelten, wo Gleichheiten in (4.2.15) beziehungsweise (4.2.16) gelten.  $\square$

Eine überraschende und weitreichende Anwendung der unscheinbaren Ungleichung (4.2.15) ist die folgende Ungleichung zwischen dem *arithmetischen* und *geometrischen Mittel*:

**Korollar 4.2.7 (AM-GM Ungleichung)** Seien  $a_1, \dots, a_n \in \mathbb{R}^+$  positive Zahlen. Dann gilt

$$a_{\text{GM}} = \sqrt[n]{a_1 \cdots a_n} \leq \frac{a_1 + \cdots + a_n}{n} = a_{\text{AM}}. \quad (4.2.19)$$

Gleichheit gilt genau dann, wenn  $a_1 = \cdots = a_n$  gilt.

BEWEIS: Dank der Ungleichung (4.2.15) gilt

$$1 = \exp\left(\sum_{i=1}^n \left(\frac{a_i}{a_{\text{AM}} - 1}\right)\right) = \prod_{i=1}^n \exp\left(\frac{a_i}{a_{\text{AM}} - 1}\right) \geq \prod_{i=1}^n \frac{a_i}{a_{\text{AM}}} = \left(\frac{a_{\text{GM}}}{a_{\text{AM}}}\right)^n.$$

Da Wurzelziehen Ungleichungen erhält, folgt (4.2.19). Da (4.2.15) strikt ist, sobald  $x \neq 0$ , folgt auch die strikte Ungleichung, sobald  $a_i \neq a_{\text{AM}}$  für ein  $i$ .  $\square$

Wir beenden diesen Abschnitt zum Logarithmus mit einigen Grenzwerten, die wir aus den entsprechenden Grenzwerten der Exponentialfunktion gewinnen.

**Proposition 4.2.8** Sei  $\epsilon > 0$ .

i.) Es gilt

$$\lim_{x \rightarrow 0} \log(x) = -\infty \quad (4.2.20)$$

aber

$$\lim_{x \rightarrow 0} x^\epsilon \log(x) = 0. \quad (4.2.21)$$

ii.) Es gilt

$$\lim_{x \rightarrow \infty} \log(x) = +\infty \quad (4.2.22)$$

aber

$$\lim_{x \rightarrow \infty} x^{-\epsilon} \log(x) = 0. \quad (4.2.23)$$

BEWEIS: Für den ersten Teil schreiben wir  $x = e^y$ , dann entspricht der Grenzwert für  $x \rightarrow 0$  dem Grenzwert für  $y \rightarrow -\infty$  nach Proposition 4.2.1, ii.). Da  $\log(e^y) = y$ , folgt (4.2.20). Für (4.2.21) verwenden wir die gleiche Substitution  $x = e^y$ . Dann gilt

$$x^\epsilon \log(x) = (e^y)^\epsilon \log(e^y) = (e^\epsilon)^y y.$$

Da für  $\epsilon > 0$  immer noch  $e^\epsilon > 1$  gilt, folgt (4.2.21) aus Proposition 4.2.1, iv.), indem wir wieder  $y \rightarrow -\infty$  betrachten. Für den zweiten Teil argumentieren wir analog. Mit  $x = e^y$  benötigen wir  $y \rightarrow \infty$  für  $x \rightarrow \infty$ . Damit ist (4.2.22) gezeigt, da  $\log(x) = \log(e^y) = y$  gegen  $+\infty$  divergiert für  $y \rightarrow \infty$ . Ebenso gilt

$$x^{-\epsilon} \log(x) = (e^y)^{-\epsilon} \log(e^y) = \frac{y}{(e^\epsilon)^y}$$

mit  $e^\epsilon > 1$  wie zuvor. Also liefert auch hier Proposition 4.2.1, iv.), den Grenzwert (4.2.23) durch  $y \rightarrow \infty$ .  $\square$

### 4.3 Die trigonometrischen Funktionen

Für reelle  $x \in \mathbb{R}$  ist  $e^{ix} = \exp(ix)$  eine komplexe Zahl mit Betrag 1 und liegt damit auf der Einheitskreislinie  $\mathbb{S}^1 \subseteq \mathbb{C}$ , siehe auch Abbildung 4.3. Wir wissen zum jetzigen Zeitpunkt noch nicht, ob die Abbildung  $x \mapsto e^{ix}$  alle Punkte in  $\mathbb{S}^1$  durchläuft und wie die Zahl  $x$  mit der Lage von  $e^{ix}$  genau zusammenhängt. Trotzdem können wir die Tatsache  $|e^{ix}| = 1$  dazu verwenden, den Real- und Imaginärteil von  $e^{ix}$  zur Definition der Winkelfunktionen heranzuziehen. Für  $x \in \mathbb{R}$  definieren wir daher den *Kosinus* von  $x$  durch

$$\cos(x) = \operatorname{Re}(e^{ix}) = \frac{e^{ix} + e^{-ix}}{2} \quad (4.3.1)$$

und den *Sinus* von  $x$  durch

$$\sin(x) = \operatorname{Im}(e^{ix}) = \frac{e^{ix} - e^{-ix}}{2i}. \quad (4.3.2)$$

Hier beachten wir, dass  $\overline{e^{ix}} = e^{-ix}$  gilt, da  $x \in \mathbb{R}$  reell sein soll. Wir verwenden die zweite Variante zur Definition von  $\cos(z)$  und  $\sin(x)$  für komplexe Argumente  $z \in \mathbb{C}$ :

**Definition 4.3.1 (Kosinus und Sinus)** Sei  $z \in \mathbb{C}$ . Dann ist der Kosinus  $\cos(z) \in \mathbb{C}$  von  $z$  als

$$\cos(z) = \frac{e^{iz} + e^{-iz}}{2} \quad (4.3.3)$$

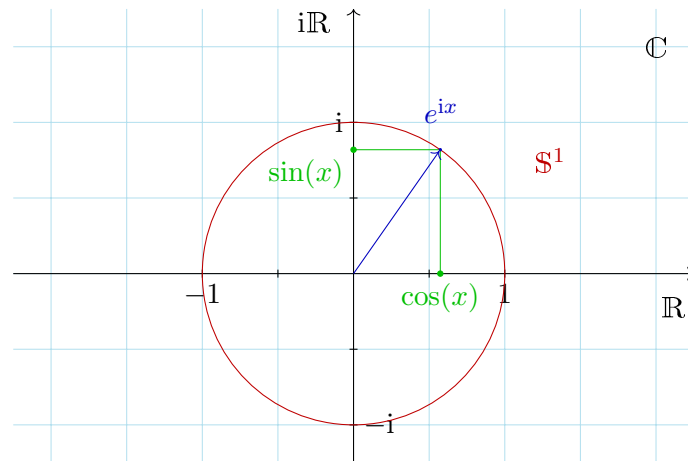


Abbildung 4.3: Die Einheitskreislinie und die komplexe Exponentialfunktion.

und der Sinus  $\sin(z) \in \mathbb{C}$  von  $z$  als

$$\sin(z) = \frac{e^{iz} - e^{-iz}}{2i}. \quad (4.3.4)$$

definiert.

Die Beziehungen

$$\cos(x) = \operatorname{Re}(e^{ix}) \quad \text{und} \quad \sin(z) = \operatorname{Im}(e^{iz}) \quad (4.3.5)$$

gelten dann für reelle Argumente  $x \in \mathbb{R}$ . Die Entscheidung für (4.3.3) und (4.3.4) wird durch folgende einfache Rechenregeln gerechtfertigt:

**Satz 4.3.2 (Kosinus und Sinus)** Seien  $z, w \in \mathbb{C}$ .

i.) Es gilt

$$e^{iz} = \cos(z) + i \sin(z). \quad (4.3.6)$$

ii.) Es gilt

$$\cos(z)^2 + \sin(z)^2 = 1. \quad (4.3.7)$$

iii.) Die Funktionen  $\cos, \sin: \mathbb{C} \rightarrow \mathbb{C}$  sind stetig.

iv.) Die Potenzreihen

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} = \cos(z) \quad (4.3.8)$$

und

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} = \sin(z) \quad (4.3.9)$$

konvergieren für alle  $z \in \mathbb{C}$ .

v.) Es gilt

$$\cos(z) = \cos(-z) \quad \text{sowie} \quad \sin(z) = -\sin(-z) \quad (4.3.10)$$

und

$$\overline{\cos(z)} = \cos(\bar{z}) \quad \text{sowie} \quad \overline{\sin(z)} = \sin(\bar{z}). \quad (4.3.11)$$

vi.) Es gilt

$$\cos(0) = 1 \quad \text{und} \quad \sin(0) = 0. \quad (4.3.12)$$



vii.) Es gelten die Additionstheoreme

$$\cos(z+w) = \cos(z)\cos(w) - \sin(z)\sin(w), \quad (4.3.13)$$

$$\sin(z+w) = \sin(z)\cos(w) + \sin(w)\cos(z), \quad (4.3.14)$$

$$\sin(z) - \sin(w) = 2 \cos\left(\frac{z+w}{2}\right) \sin\left(\frac{z-w}{2}\right) \quad (4.3.15)$$

und

$$\cos(z) - \cos(w) = 2 \sin\left(\frac{z+w}{2}\right) \cos\left(\frac{z-w}{2}\right). \quad (4.3.16)$$

BEWEIS: Der erste Teil ist aufgrund der Definition unmittelbar klar. Den zweiten Teil erhalten wir ebenfalls durch Einsetzen und Verwendung der Eigenschaften der komplexen Exponentialfunktion. Es gilt

$$\begin{aligned} \cos(z)^2 + \sin(z)^2 &= \left(\frac{e^{iz} + e^{-iz}}{2}\right)^2 + \left(\frac{e^{iz} - e^{-iz}}{2i}\right)^2 \\ &= \frac{e^{2iz} + 2 + e^{-2iz}}{4} - \frac{e^{2iz} - 2 + e^{-2iz}}{4} \\ &= 1. \end{aligned}$$

Die Stetigkeit von  $\cos$  und  $\sin$  folgt unmittelbar aus der Stetigkeit von  $\exp$ . Interessanter sind die Darstellungen von  $\cos$  und  $\sin$  als Potenzreihen. Aus der Exponentialreihe erhalten wir

$$\cos(z) = \frac{1}{2} \sum_{n=0}^{\infty} \frac{(iz)^n}{n!} + \frac{1}{2} \sum_{n=0}^{\infty} \frac{(-iz)^n}{n!} = \sum_{n=0}^{\infty} \frac{i^n + (-i)^n}{2n!} z^n.$$

Nun gilt

$$\frac{i^n + (-i)^n}{2} = \begin{cases} 0 & \text{für } n \text{ ungerade} \\ (-1)^k & \text{für } n = 2k \text{ gerade,} \end{cases}$$

da  $i^2 = -1 = (-i)^2$ . Dies liefert

$$\cos(z) = \sum_{n=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k}.$$

Man beachte, dass die Konvergenz von  $\exp(z)$  die von  $\cos(z)$  impliziert, da wir konvergente Reihen unter Beibehaltung der Konvergenz gliedweise addieren dürfen. Alternativ sieht man die Konvergenz von (4.3.8) auch direkt mit dem Quotientenkriterium. Die Darstellung (4.3.9) für  $\sin(z)$  erhält man genauso. Die Gleichungen (4.3.10) und (4.3.11) folgen entweder aus der Definition und (4.1.6) direkt, oder aus den Reihendarstellungen (4.3.8) und (4.3.9). Die speziellen Werte (4.3.12) erhält man aus  $\exp(0) = 1$  oder wieder aus (4.3.8) und (4.3.9). Für die Additionstheoreme im letzten Teil benötigt man die viel einfacheren Additionstheoreme der Exponentialfunktion (4.1.3). Wir zeigen

$$\begin{aligned} &\cos(z+w) + i\sin(z+w) \\ &= e^{i(z+w)} \\ &= e^{iz} e^{iw} \\ &= (\cos(z) + i\sin(z))(\cos(w) + i\sin(w)) \\ &= \cos(z)\cos(w) - \sin(z)\sin(w) + i\sin(z)\cos(w) + i\sin(w)\cos(z). \end{aligned} \quad (*)$$

Ersetzt man  $z+w$  durch  $-(z+w)$ , erhält man

$$\cos(z+w) - i\sin(z+w) = \cos(z)\cos(w) - \sin(z)\sin(w) + i\sin(z)\cos(w) - i\sin(w)\cos(z), \quad (**)$$

da  $\cos$  eine gerade Funktion und  $\sin$  eine ungerade Funktion ist. Addition von (\*) und (\*\*) liefert dann (4.3.13), Subtraktion entsprechend (4.3.14). Ersetzt man statt dessen  $z$  durch  $\frac{z+w}{2}$  und  $w$  durch  $\frac{z-w}{2}$  in (\*) und (\*\*) und bildet wieder die Summe beziehungsweise die Differenz, erhält man die verbleibenden Gleichungen (4.3.15) und (4.3.16).  $\square$

Wie schon die Exponentialreihe konvergieren die Reihen für Kosinus und Sinus sehr schnell. Als detaillierte Fehlerabschätzung erhält man folgendes Resultat:

**Proposition 4.3.3** Sei  $N \in \mathbb{N}_0$  und  $z \in \mathbb{C}$ .

i.) Falls  $|z| \leq \frac{1}{2}(2N+4)$ , gilt

$$\left| \sin(z) - \sum_{k=0}^N \frac{(-1)^k z^{2k+1}}{(2k+1)!} \right| \leq \frac{4}{3} \frac{|z|^{2N+3}}{(2N+3)!}. \quad (4.3.17)$$

ii.) Falls  $|z| \leq \frac{1}{2}(2N+3)$ , gilt

$$\left| \cos(z) - \sum_{k=0}^N \frac{(-1)^k z^{2k}}{(2k)!} \right| \leq \frac{4}{3} \frac{|z|^{2N+2}}{(2N+2)!}. \quad (4.3.18)$$

BEWEIS: Für den ersten Teil erhalten wir

$$\begin{aligned} & \left| \sin(z) - \sum_{k=0}^N \frac{(-1)^k z^{2k+1}}{(2k+1)!} \right| \\ &= \left| \sum_{k=N+1}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!} \right| \\ &\leq \sum_{k=N+1}^{\infty} \frac{|z|^{2N+3}}{(2N+3)!} \frac{|z|^{2k-2N-2}}{(2k+1)(2k)(2k-1) \cdots (2N+4)} \\ &\leq \frac{|z|^{2N+3}}{(2N+3)!} \sum_{k=N+1}^{\infty} \left(\frac{1}{2}\right)^{2k-2N-2} \\ &= \frac{|z|^{2N+3}}{(2N+3)!} \sum_{n=0}^{\infty} \frac{1}{4^n} \\ &= \frac{4}{3} \frac{|z|^{2N+3}}{(2N+3)!}, \end{aligned}$$

da wir für jeden der  $2k-2N-2$  Faktoren  $|z|$  die Abschätzung  $|z| \leq \frac{1}{2}(2N+4)$  verwenden können. Genauso erhält man

$$\begin{aligned} & \left| \cos(z) - \sum_{k=0}^N \frac{(-1)^k z^{2k}}{(2k)!} \right| \\ &= \left| \sum_{k=N+1}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!} \right| \\ &\leq \sum_{k=N+1}^{\infty} \frac{|z|^{2N+2}}{(2N+2)!} \frac{|z|^{2k-2N-2}}{(2k)(2k-1) \cdots (2N+3)} \\ &\leq \frac{|z|^{2N+2}}{(2N+2)!} \sum_{k=N+1}^{\infty} \left(\frac{1}{2}\right)^{2k-2N-2} \end{aligned}$$

$$= \frac{4}{3} \frac{|z|^{2N+2}}{(2N+2)!}.$$

□

Für reelle  $x$  hat man alternierende Vorzeichen vorliegen, womit man das Leibniz-Kriterium zur Reihenkonvergenz und insbesondere die Abschätzung des Restgliedes aus dem Leibniz-Kriterium verwenden kann. Dies liefert geringfügig bessere Abschätzungen, siehe auch Übung ??.

## 4.4 $\pi$

Die Kreiszahl  $\pi$  hat eine bewegte Geschichte und wird zuerst in der Kreisgeometrie kennengelernt. Entweder ist  $\pi$  als halber Umfang des Einheitskreises definiert, oder als Flächeninhalt des Einheitskreises. Beide Sichtweisen kann man zur mathematischen Definition von  $\pi$  heranziehen. Allerdings benötigen beide Optionen einen Integralbegriff, da sowohl die Länge einer Kurve, der Kreislinie, als auch die Fläche eines nicht rechteckigen Gebietes, des Kreisesinneren, mathematisch nicht so naiv zu verstehen sind, wie man das vielleicht zunächst meint. Aus diesem Grunde gehen wir hier einen anderen Weg zur Definition von  $\pi$ , der uns bereits jetzt zur Verfügung steht und keine Integraltheorie erfordert. Der Zugang ist mittlerweile weit verbreitet. Selbstverständlich werden wir zu einem späteren Zeitpunkt den Anschluss an die elementar-geometrischen Ideen zur Kreiszahl  $\pi$  suchen und finden.

Als Vorüberlegung benötigen wir zunächst folgende Monotonieeigenschaft der Kosinusfunktion:

**Lemma 4.4.1** *i.) Für alle  $x \in (0, 2)$  gilt*

$$\sin(x) > 0. \quad (4.4.1)$$

*ii.) Die Kosinusfunktion ist auf dem Intervall  $[0, 2]$  strikt monoton fallend.*

*iii.) Es gilt  $\cos(0) = 1$  und  $\cos(2) < 0$ .*

BEWEIS: Für den ersten Teil verwenden wir die Restgliedabschätzung der Sinusreihe aus Proposition 4.3.3, i.). Für  $N = 0$  erhalten wir

$$|\sin(x) - x| \leq \frac{4}{3} \frac{x^3}{3!} = \frac{2x^3}{9}$$

für  $|x| \leq \frac{1}{2} \cdot 2 = 2$ , also für  $x \in [0, 2]$ . Damit gilt

$$\begin{aligned} \sin(x) &= x + (\sin(x) - x) \\ &\geq x - \frac{2}{9}x^3 \\ &= x \left(1 - \frac{2x^2}{9}\right) \\ &\geq x \left(1 - \frac{8}{9}\right) \\ &= \frac{1}{9}x \\ &> 0 \end{aligned}$$

für  $x \in (0, 2)$ , was den ersten Teil zeigt. Für den zweiten Teil betrachten wir  $0 \leq x < y \leq 2$  und verwenden die Additionstheoreme für die Kosinusfunktion aus Satz 4.3.2, vii.). Es folgt

$$\cos(y) - \cos(x) = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{y-x}{2}\right) < 0, \quad (*)$$

da  $\frac{x+y}{2} \in (0, 2)$  und ebenso  $\frac{y-x}{2} \in (0, 2)$ , womit nach *i.*) die Behauptung (\*) folgt. Dies zeigt die strikte Monotonie von  $\cos$ . Schließlich benötigen wir die Werte von  $\cos(x)$  für  $x = 0$  und  $x = 2$ . Hier ist  $\cos(0) = 1$  anhand der Reihendarstellung klar. Für  $x = 2$  verwenden wir erneut die Restgliedabschätzung aus Proposition 4.3.3, *ii.*), diesmal für die Kosinusreihe. Für  $|x| < \frac{1}{2}(2+3) = \frac{5}{2}$  gilt

$$\left| \cos(x) - 1 + \frac{x^2}{2} \right| \leq \frac{4}{3} \frac{|x|^4}{4!},$$

und daher

$$\cos(x) = 1 - \frac{x^2}{2} + \left( \cos(x) - 1 + \frac{x^2}{2} \right) \leq 1 - \frac{x^2}{2} + \frac{4}{3} \frac{|x|^4}{4!}.$$

Für  $x = 2$  erhalten wir also

$$\cos(2) \leq 1 - \frac{4}{2} + \frac{4}{3} \cdot \frac{16}{24} = -\frac{1}{9} < 0. \quad \square$$

**Korollar 4.4.2** *Auf dem Intervall  $[0, 2]$  besitzt die Kosinusfunktion eine eindeutige Nullstelle.*

BEWEIS: Nach dem Zwischenwertsatz für die stetige Funktion  $\cos$  gibt es Nullstellen, da  $\cos(0) = 1$  und  $\cos(2) < 0$ . Die strikte Monotonie zeigt, dass es nur eine solche geben kann.  $\square$

**Definition 4.4.3 (Kreiszahl  $\pi$ )** *Die Kreiszahl  $\pi$  ist das Doppelte der eindeutig bestimmten Nullstelle der Kosinusfunktion  $\cos$  im Intervall  $[0, 2]$ .*

Es gilt also

$$\cos\left(\frac{\pi}{2}\right) = 0, \quad (4.4.2)$$

und  $\frac{\pi}{2}$  ist die kleinste positive Nullstelle der Kosinusfunktion. Es gilt

$$0 < \pi < 4, \quad (4.4.3)$$

bessere Abschätzungen für  $\pi$  werden wir noch kennenlernen. Mit den Additionstheoremen für die komplexe Exponentialfunktion können wir nun die Periodizität von  $\sin$  und  $\cos$  zeigen.

**Satz 4.4.4 (Periodizität von Sinus und Kosinus)** *Sei  $z \in \mathbb{C}$ .*

*i.) Es gilt folgende Wertetabelle:*

| $x$       | $\frac{\pi}{2}$ | $\pi$ | $\frac{3\pi}{2}$ | $2\pi$ |
|-----------|-----------------|-------|------------------|--------|
| $e^{ix}$  | $i$             | $-1$  | $-i$             | $1$    |
| $\sin(x)$ | $1$             | $0$   | $-1$             | $0$    |
| $\cos(x)$ | $0$             | $-1$  | $0$              | $1$    |

(4.4.4)

*ii.) Es gilt*

$$e^{i\frac{\pi}{2}+z} = ie^z, \quad (4.4.5)$$

$$e^{i\pi+z} = -e^z \quad (4.4.6)$$

und

$$e^{2\pi i+z} = e^z. \quad (4.4.7)$$

iii.) Es gilt

$$\cos\left(\frac{\pi}{2} + z\right) = -\sin(z), \quad (4.4.8)$$

$$\cos(\pi + z) = -\cos(z) \quad (4.4.9)$$

und

$$\cos(2\pi + z) = \cos(z). \quad (4.4.10)$$

iv.) Es gilt

$$\sin\left(\frac{\pi}{2} + z\right) = \cos(z), \quad (4.4.11)$$

$$\sin(\pi + z) = -\sin(z) \quad (4.4.12)$$

und

$$\sin(2\pi + z) = \sin(z). \quad (4.4.13)$$

v.) Es gilt genau dann  $e^z = 1$ , wenn  $z \in 2\pi i\mathbb{Z}$ . Weiter gilt

$$\cos^{-1}(\{0\}) = \left\{k\pi + \frac{\pi}{2} \mid k \in \mathbb{Z}\right\} \quad (4.4.14)$$

und

$$\sin^{-1}(\{0\}) = \{k\pi \mid k \in \mathbb{Z}\}. \quad (4.4.15)$$

BEWEIS: Nach Definition wissen wir  $\cos(\frac{\pi}{2}) = 0$ . Da  $\cos(z)^2 + \sin(z)^2 = 1$  für alle  $z \in \mathbb{C}$  gilt, folgt  $\sin(\frac{\pi}{2}) = 1$ , also  $\sin(\frac{\pi}{2}) = \pm 1$ . Da aber  $0 < \frac{\pi}{2} < 2$  und  $\sin(x)$  auf  $(0, 2)$  positiv ist, folgt  $\sin(\frac{\pi}{2}) = 1$ . Damit folgt auch

$$e^{i\frac{\pi}{2}} = \cos\left(\frac{\pi}{2}\right) + i\sin\left(\frac{\pi}{2}\right) = i,$$

was die erste Spalte in (4.4.4) zeigt. Es folgt

$$e^{in\frac{\pi}{2}} = (e^{i\frac{\pi}{2}})^n = i^n$$

für alle  $n \in \mathbb{Z}$ , womit die zweite Zeile in (4.4.4) gezeigt ist. Betrachten wir erneut

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

so folgen die verbleibenden Einträge in der Tabelle durch Vergleich der Real- und Imaginärteile, da  $\cos(x)$  und  $\sin(x)$  für reelles  $x \in \mathbb{R}$  ja selbst reell sind. Insbesondere erhält man so die schönste Formel der Mathematik (4.1) aus dieser Tabelle. Der zweite Teil folgt nun erneut aus den Additionstheoremen der Exponentialfunktion

$$e^{z+w} = e^z e^w$$

und den speziellen Werten aus der Tabelle. Die entsprechenden Gleichungen für  $\cos$  und  $\sin$  erhält man aus den Additionstheoremen in Satz 4.3.2, vii.), sowie den speziellen Werten aus (4.4.4). Zum Abschluss zeigen wir zunächst (4.4.14) und (4.4.15). Wir wissen  $\cos(x) > 0$  für  $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$  und nach  $\sin(x) = -\sin(-x) = \cos(\frac{\pi}{2} - x)$  erhalten wir  $\sin(x) > 0$  für  $x \in (0, \pi)$ . Bei  $\pi$  hat die Sinusfunktion daher die erste positive Nullstelle. Die  $2\pi$ -Periodizität (4.4.13) zeigt dann, dass alle Zahlen der Form  $\pi k$  mit  $k \in \mathbb{Z}$  Nullstellen sind. Wegen (4.4.12) gilt  $\sin(x) < 0$  für  $x \in (\pi, 2\pi)$ , womit die Sinusfunktion auch keine anderen reellen Nullstellen besitzt. Mit (4.4.11) findet man dann alle reellen Nullstellen der Kosinusfunktion als diejenigen Zahlen  $\pi k + \frac{\pi}{2}$  mit  $k \in \mathbb{Z}$ . Wir betrachten nun  $e^z = 1$  mit  $z = a + ib$  und  $a, b \in \mathbb{R}$ . Da  $1 = |e^z| = |e^a||e^{ib}| = e^a$ , folgt  $a = 0$ , da die *reelle* Exponentialfunktion nach Proposition 4.2.1, iii.), injektiv ist. Also ist  $z = ib$  rein imaginär. Um dann  $e^{ib} = 1$  zu erreichen, muss

also  $\sin(b) = 0$  gelten, was auf  $b = \pi k$  mit  $k \in \mathbb{Z}$  führt. Weiter muss  $\cos(b) = 1$  gelten, was  $b = 2\pi k$  mit  $k \in \mathbb{Z}$  zeigt. Dies liefert also  $z = 2\pi i k$  mit  $k \in \mathbb{Z}$ . Um schließlich (4.4.15) und (4.4.14) zu zeigen, müssen wir ausschließen, dass die Sinus- oder Kosinusfunktion Nullstellen in  $\mathbb{C} \setminus \mathbb{R}$  besitzt. Sei also  $z \in \mathbb{C}$  eine Nullstelle der Sinusfunktion, womit  $0 = \sin(z) = \frac{1}{2i}(e^{iz} - e^{-iz})$  gilt. Daher gilt  $e^{iz} = e^{-iz}$  oder eben  $e^{2iz} = 1$ , was  $z \in 2\pi\mathbb{Z}$  liefert. Dies zeigt nun endgültig die Gleichheit in (4.4.15), womit die Gleichheit in (4.4.14) dank (4.4.11) ebenfalls folgt.  $\square$

Im Beweis haben wir zudem gesehen, dass

$$\sin \Big|_{(0,\pi)} > 0 \quad (4.4.16)$$

und entsprechend

$$\sin \Big|_{(\pi,2\pi)} < 0 \quad (4.4.17)$$

gilt. Mit (4.4.8) sehen wir daher auch

$$\cos \Big|_{(-\frac{\pi}{2},\frac{\pi}{2})} > 0 \quad (4.4.18)$$

und

$$\cos \Big|_{(\frac{\pi}{2},\frac{3\pi}{2})} < 0. \quad (4.4.19)$$

Sowohl bei der Sinusfunktion als auch bei der Kosinusfunktion setzt sich dies  $2\pi$ -periodisch auf alle reellen Argumente fort.

#### Proposition 4.4.5

- i.) Die Kosinusfunktion ist strikt monoton fallend auf  $[0, \pi]$  und strikt monoton wachsend auf  $[\pi, 2\pi]$ .
- ii.) Die Sinusfunktion ist strikt monoton wachsend auf  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  und strikt monoton fallend auf  $[\frac{\pi}{2}, \frac{3\pi}{2}]$ .
- iii.) Bei beiden Funktionen setzt sich dieses Monotonieverhalten  $2\pi$ -periodisch fort.

BEWEIS: Das Argument aus Lemma 4.4.1, ii.), können wir dank der Positivität (4.4.16) auf alle  $x \in (0, \pi)$  ausdehnen und erhalten so die Monotonie von  $\cos$  auf  $[0, \pi]$ . Mit (4.4.8) und (4.4.11) erhalten wir dann die anderen Monotonieaussagen in i.) und ii.). Der letzte Teil ist klar.  $\square$

Mit der Kreiszahl  $\pi$  können wir also insbesondere alle Nullstellen der Sinus- und Kosinusfunktion im Komplexen kodieren. Das erlaubt es uns nun, die verbleibenden trigonometrischen Funktionen, die *Tangens*- und die *Kotangensfunktion*, zu definieren.

**Definition 4.4.6 (Tangens und Kotangens)** Die Tangensfunktion

$$\tan: \mathbb{C} \setminus \left\{ \frac{\pi}{2} + \pi k \mid k \in \mathbb{Z} \right\} \longrightarrow \mathbb{C} \quad (4.4.20)$$

ist definiert durch

$$\tan(z) = \frac{\sin(z)}{\cos(z)}. \quad (4.4.21)$$

Entsprechend definiert man die Kotangensfunktion

$$\cot: \mathbb{C} \setminus \{ \pi k \mid k \in \mathbb{Z} \} \longrightarrow \mathbb{C} \quad (4.4.22)$$

durch

$$\cot(z) = \frac{\cos(z)}{\sin(z)}. \quad (4.4.23)$$

Die Definitionsbereiche sind maximal so gewählt, dass die Nullstellen der Nenner vermieden werden. Diese hatten wir in Satz 4.4.4, v.), bestimmt. Einige erste Eigenschaften fassen wir in folgender Proposition zusammen:

**Proposition 4.4.7** i.) Die Tangens- und Kotangensfunktion sind auf ihrem jeweiligen Definitionsbereich stetig.

ii.) Für alle  $z$  im Definitionsbereich gilt

$$\tan(z + \pi k) = \tan(z) \quad (4.4.24)$$

und

$$\cot(z + \pi k) = \cot(z) \quad (4.4.25)$$

für alle  $k \in \mathbb{Z}$ .

iii.) Es gilt

$$\tan^{-1}(\{0\}) = \{\pi k \mid k \in \mathbb{Z}\} \quad (4.4.26)$$

und

$$\cot^{-1}(\{0\}) = \left\{ \frac{\pi}{2} + \pi k \mid k \in \mathbb{Z} \right\}. \quad (4.4.27)$$

iv.) Es gelten die Additionstheoreme

$$\tan(z + w) = \frac{\tan(z) + \tan(w)}{1 - \tan(z)\tan(w)} \quad (4.4.28)$$

und

$$\cot(z + w) = \frac{\cot(z)\cot(w) - 1}{\cot(z) + \cot(w)}, \quad (4.4.29)$$

sofern  $z, w$  und  $z + w$  in den Definitionsbereichen liegen.

BEWEIS: Da sowohl die Sinus- als auch die Kosinusfunktion stetig auf ganz  $\mathbb{C}$  sind, ist der Quotient (4.4.21) beziehungsweise (4.4.23) immer dort stetig, wo er überhaupt definiert ist, also abseits der Nennernullstellen. Diese sind aber gerade die ausgeschlossenen Punkte der Definitionsbereiche von  $\tan$  und  $\cot$ . Die Periodizität von  $\tan$  und  $\cot$  folgt nun aus der Periodizität von  $\sin$  und  $\cos$  gemäß Satz 4.4.4, iii.) und iv.), wobei wir nun sogar  $\pi$ -Periodizität vorliegen haben, da sich mit (4.4.9) und (4.4.12) in  $\tan$  und  $\cot$  die Vorzeichen gerade aufheben. Die Nullstellen von  $\tan$  sind die von  $\sin$ , also nach (4.4.15) durch (4.4.26) gegeben. Entsprechend erhält man die Nullstellen von  $\cot$  aus (4.4.14). Die Additionstheoreme zeigt man entweder mit den entsprechenden Additionstheoremen für Sinus und Kosinus oder direkt aus den Eigenschaften der komplexen Exponentialfunktion. Für (4.4.28) betrachten wir

$$\begin{aligned} \frac{\tan(z) + \tan(w)}{1 - \tan(z)\tan(w)} &= \frac{\frac{\sin(z)}{\cos(z)} + \frac{\sin(w)}{\cos(w)}}{1 - \frac{\sin(z)}{\cos(z)} \frac{\sin(w)}{\cos(w)}} \\ &= \frac{\sin(z)\cos(w) + \sin(w)\cos(z)}{\cos(z)\cos(w) - \sin(z)\sin(w)} \\ &= \frac{\sin(z+w)}{\cos(z+w)} \\ &= \tan(z+w), \end{aligned}$$

wobei wir nutzen, dass  $\cos(z) \neq 0 \neq \cos(w)$  und daher der Bruch mit  $\cos(z)\cos(w)$  erweitert beziehungsweise gekürzt werden darf. Genauso erhält man (4.4.29).  $\square$

Weitere Eigenschaften von  $\tan$  und  $\cot$  werden wir in den Übungen sowie später diskutieren. Eine graphische Darstellung für reelle Argumente zeigt Abbildung 4.4

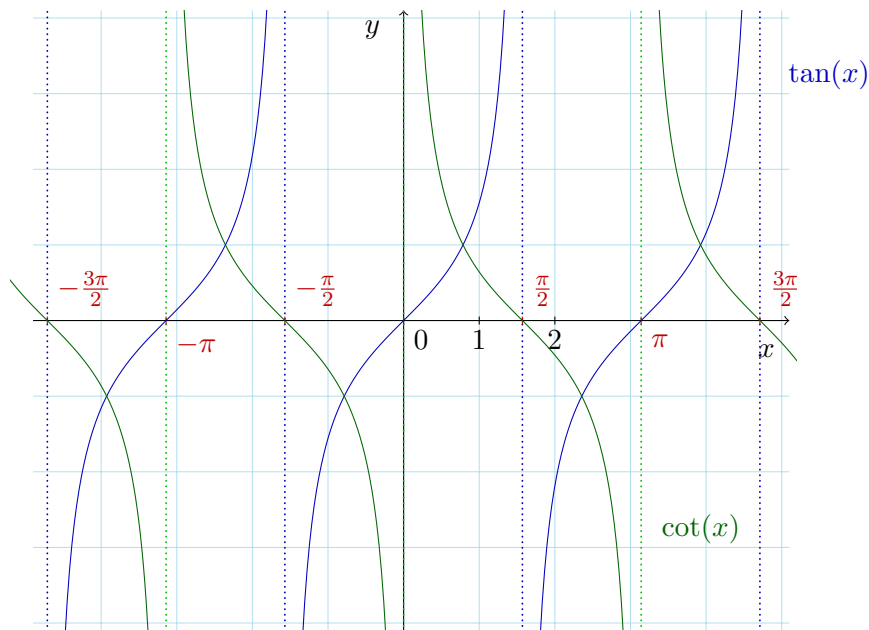


Abbildung 4.4: Tangens- und Kotangensfunktion auf der reellen Achse

## 4.5 Die Arcusfunktionen und Polarkoordinaten

Die trigonometrischen Funktionen sind aufgrund ihrer Periodizität sicherlich nicht injektiv, wenn man sie auf ihrem gesamten Definitionsbereich betrachtet. Schränkt man sie dagegen hinreichend ein, erhält man nach Proposition 4.4.5 beziehungsweise Übung ?? strikt monotone und damit injektive Funktionen. Diese sind bijektiv auf ihr Bild, welches ebenfalls leicht zu bestimmen ist und dann als Definitionsbereich der jeweiligen Umkehrfunktion dient. Insbesondere ist nach Proposition 3.5.16 der Definitionsbereich wieder ein Intervall. Dies führt zu folgenden Resultaten zu den trigonometrischen Funktionen:

### Proposition 4.5.1

i.) Die Einschränkung der Sinusfunktion liefert eine stetige Bijektion

$$\sin: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \longrightarrow [-1, 1]. \quad (4.5.1)$$

ii.) Die Einschränkung der Kosinusfunktion liefert eine stetige Bijektion

$$\cos: [0, \pi] \longrightarrow [-1, 1]. \quad (4.5.2)$$

iii.) Die Tangensfunktion ist auf  $(-\frac{\pi}{2}, \frac{\pi}{2})$  strikt monoton wachsend und liefert eine stetige Bijektion

$$\tan: \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow \mathbb{R}. \quad (4.5.3)$$

iv.) Die Kotangensfunktion ist auf  $(0, \pi)$  strikt monoton fallend und liefert eine stetige Bijektion

$$\cot: (0, \pi) \longrightarrow \mathbb{R}. \quad (4.5.4)$$

BEWEIS: Die ersten beiden Aussagen folgen aus der Monotonie in Proposition 4.4.5 sowie aus den Werten bei 0 und  $\pi$  sowie bei  $-\frac{\pi}{2}$  und  $\frac{\pi}{2}$  gemäß Satz 4.4.4, i.). Man beachte, dass für reelle Argumente sowohl die Sinus- als auch die Kosinusfunktion reell sind und daher  $(\cos(x))^2 + (\sin(x))^2 = 1$  die



Schranken  $|\cos(x)| \leq 1$  sowie  $|\sin(x)| \leq 1$  liefert. Die Monotonie der Tangensfunktion sowie der Kotangensfunktion folgt damit leicht, siehe auch Übung ?? . Da die Kosinusfunktion bei  $\pm \frac{\pi}{2}$  Nullstellen besitzt, die Sinusfunktion aber die Werte  $\pm 1$  annimmt sieht man schnell, dass

$$\lim_{x \nearrow \frac{\pi}{2}} \tan(x) = \infty$$

sowie

$$\lim_{x \searrow -\frac{\pi}{2}} \tan(x) = -\infty$$

für die links- beziehungsweise rechtsseitigen Funktionslimites der Tangensfunktion gilt. Damit ist das Bild aber ganz  $\mathbb{R}$ , was (4.5.3) zeigt. Die Argumentation für die Kotangensfunktion ist analog, siehe erneut Abbildung 4.4.  $\square$

Mit Proposition 3.5.16 erhalten wir daher stetige Umkehrfunktionen zu allen vier trigonometrischen Funktionen, die *Arcusfunktionen*:

**Definition 4.5.2 (Arcusfunktionen)**

i.) Die Arcussinusfunktion

$$\arcsin: [-1, 1] \longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad (4.5.5)$$

ist als Umkehrfunktion  $\arcsin = \sin^{-1}$  zur Einschränkung der Sinusfunktion auf  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  definiert.

ii.) Die Arcuscosinusfunktion

$$\arccos: [-1, 1] \longrightarrow [0, \pi] \quad (4.5.6)$$

ist als Umkehrfunktion  $\arccos = \cos^{-1}$  zur Einschränkung der Kosinusfunktion auf  $[0, \pi]$  definiert.

iii.) Die Arcustangensfunktion

$$\arctan: \mathbb{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \quad (4.5.7)$$

ist als Umkehrfunktion  $\arctan = \tan^{-1}$  zur Einschränkung der Tangensfunktion auf  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  definiert.

iv.) Die Arcuscotangensfunktion

$$\operatorname{arccot}: \mathbb{R} \longrightarrow (0, \pi) \quad (4.5.8)$$

ist als Umkehrfunktion  $\operatorname{arccot} = \cot^{-1}$  zur Einschränkung der Kotangensfunktion auf  $[0, \pi]$  definiert.

In der Literatur werden gelegentlich auch die Symbole  $\operatorname{atan}$  und  $\operatorname{acot}$  für die Arcustangens- und Arcuskotangensfunktion verwendet.

**Korollar 4.5.3** Die Arcusfunktionen sind auf ihrem jeweiligen Definitionsbereich stetig und strikt monoton.

BEWEIS: Dies folgt aus Proposition 3.5.16.  $\square$

Wir können die Arcusfunktionen nun dazu benutzen, Winkel im Einheitskreis zu definieren. Dies basiert auf folgender Überlegung:

**Proposition 4.5.4** Sei  $\mathbb{S}^1 \subseteq \mathbb{C}$  die Einheitskreislinie.

i.) Die Multiplikation von komplexen Zahlen ergibt eine abelsche Gruppenstruktur für  $\mathbb{S}^1$ .

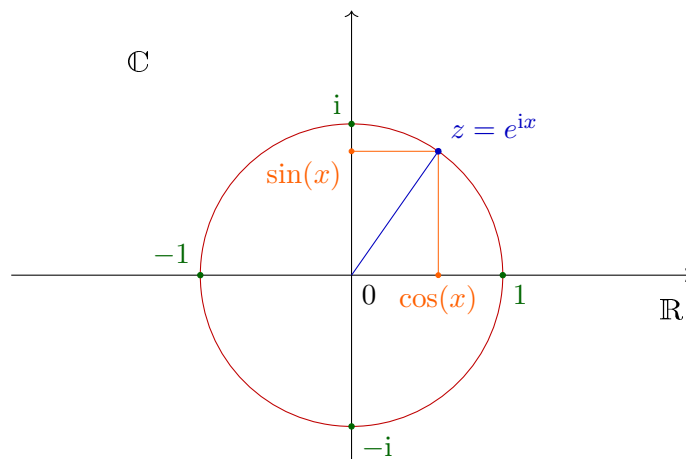


Abbildung 4.5:

ii.) Die Abbildung

$$(\mathbb{R}, +) \ni x \mapsto \exp(ix) \in \mathbb{S}^1 \quad (4.5.9)$$

ist ein stetiger und surjektiver Gruppenmorphismus mit Kern  $2\pi\mathbb{Z} \subseteq \mathbb{R}$ .

BEWEIS: Die Gruppenstruktur folgt direkt aus  $|z||w| = |zw|$  für alle  $z, w \in \mathbb{C}$ , siehe auch Übung 1.8.36. Dass (4.5.9) ein Gruppenmorphismus ist, liegt an der Funktionalgleichung, denn es gilt

$$\exp(i(x+y)) = \exp(ix) \exp(iy)$$

sowie  $\exp(i \cdot 0) = 1$ . Die Stetigkeit von (4.5.9) folgt aus der Stetigkeit von  $\exp$ . Den Kern haben wir bereits in Satz 4.4.4, v.), bestimmt. Es bleibt also die Surjektivität zu zeigen. Sei  $z \in \mathbb{S}^1$  mit  $z = a + ib$  und  $a, b \in \mathbb{R}$  vorgegeben. Dann gilt  $a^2 + b^2 = 1$  und somit  $a, b \in [-1, 1]$ . Wir finden daher genau ein  $x \in [0, \pi]$  mit

$$x = \arccos(a)$$

oder eben  $\cos(x) = a$ . Für  $b$  erhalten wir entsprechend aus  $a^2 + b^2 = 1$  die beiden Möglichkeiten

$$b = +\sqrt{1-a^2} \quad \text{oder} \quad b = -\sqrt{1-a^2},$$

je nachdem ob  $b$  positiv oder negativ ist. Umgekehrt gilt für  $x$

$$1 = \sin(x)^2 + \cos(x)^2 = \sin(x)^2 + a^2,$$

womit  $\sin(x) = \sqrt{1-a^2}$ , da die Sinusfunktion auf  $[0, \pi]$  nicht negativ ist. Ist nun  $b \geq 0$  so gilt  $\sin(x) = b$  und damit

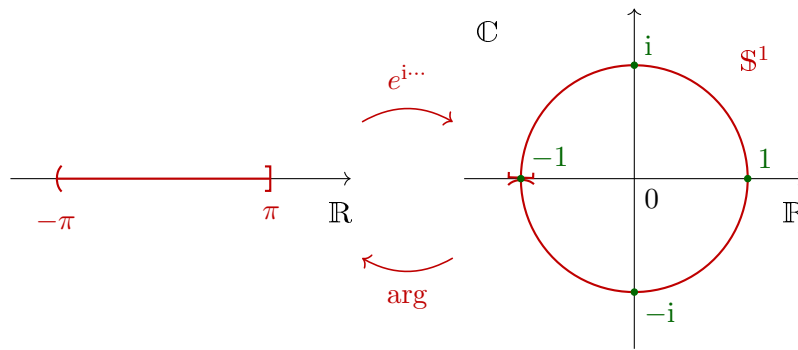
$$e^{ix} = \cos(x) + i \sin(x) = a + ib = z,$$

was die Surjektivität in diesem Fall zeigt. Gilt  $b < 0$ , so setzen wir  $y = -x$  und erhalten  $\sin(y) = \sin(-x) = -\sin(x) = -\sqrt{1-a^2} = b$ . Da  $\cos(x) = \cos(-x) = \cos(y)$ , folgt auch in diesem Fall

$$e^{iy} = \cos(y) + i \sin(y) = a + ib = z,$$

was die Surjektivität insgesamt zeigt. □

Da wir den Kern des Gruppenmorphismus (4.5.9) als die ganzzahligen Vielfachen von  $2\pi$  bestimmt haben, können wir also einer komplexen Zahl  $z \in \mathbb{S}^1$  ein bis auf Addition mit Elementen des Kerns  $2\pi\mathbb{Z}$  eindeutiges Urbild in  $\mathbb{R}$  unter (4.5.9) zuordnen. Dies nennt man auch das *Argument* von  $z$ :

Abbildung 4.6: Unstetigkeit von  $\arg$ .

**Definition 4.5.5 (Argument)** Sei  $z \in \mathbb{S}^1$ . Die eindeutig bestimmte reelle Zahl  $x \in (-\pi, \pi]$  mit  $e^{ix} = z$  heißt Argument von  $z$ , geschrieben als

$$\arg(z) = x. \quad (4.5.10)$$

Alternativ könnte man  $z$  auch ein eindeutiges Argument in  $(0, 2\pi]$  oder  $[0, 2\pi)$  zuordnen. Hier gibt es verschiedenen Konventionen, die sich eben durch Addition eines Elements aus dem Kern von (4.5.9) unterscheiden.

**Bemerkung 4.5.6 (Winkel)** Im Hinblick auf die geometrische Situation im Einheitskreis, siehe Abbildung 4.5, können wir  $\arg(z)$  als Winkel zwischen der positiven  $x$ -Achse und dem Vektor  $z$  verstehen. Mit dieser Interpretation wird die Multiplikation komplexer Zahlen im Einheitskreis gerade die Addition ihrer Winkel modulo  $2\pi$ . Dass dieser Winkel die Länge des Kreissegments von 1 bis  $z$  ist, können wir allerdings noch nicht begründen, da wir noch nicht über eine mathematisch brauchbare Definition von Länge einer Kurve verfügen.

**Bemerkung 4.5.7 (Stetigkeit von  $\arg$ )** Die Abbildung

$$\arg: \mathbb{S}^1 \longrightarrow (-\pi, \pi] \quad (4.5.11)$$

ist eine Bijektion und ihre Umkehrabbildung ist die Einschränkung der stetigen Exponentialabbildung

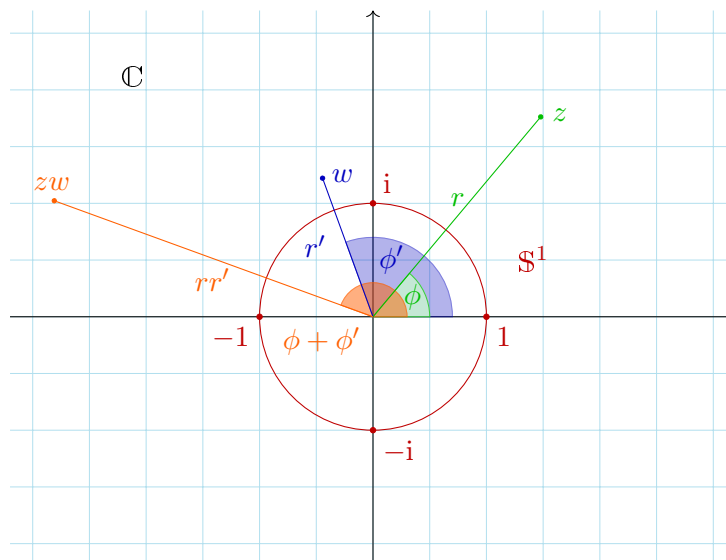
$$(-\pi, \pi] \ni x \mapsto e^{ix} \in \mathbb{S}^1. \quad (4.5.12)$$

Allerdings ist  $\arg$  *nicht* stetig. Dies zeigt man beispielsweise dadurch, dass  $\mathbb{S}^1 \subseteq \mathbb{C}$  kompakt ist, das Bild unter  $\arg$  allerdings nicht, siehe auch Abbildung 4.6. Das passt sehr gut zur anschaulichen Vorstellung, dass die Kreislinie  $\mathbb{S}^1$  bei  $-1$  zerrissen werden muss, um sie auf das Intervall  $(-\pi, \pi]$  abzuwickeln. Wir erhalten hier also das angekündigte Beispiel einer stetigen Bijektion, die kein stetiges Inverses besitzt, siehe nochmal die Diskussion bei Abbildung 3.23. Von diesem Standpunkt aus wird Proposition 3.5.16 nochmals bemerkenswerter: Solange wir mit Intervallen alleine auskommen und nicht den Ausflug in die höher-dimensionale Ebene vornehmen, haben stetige Bijektionen stetige Inverse. Schränkt man  $\arg$  stattdessen auf  $\mathbb{S}^1 \setminus \{-1\}$  ein, erhält man eine stetige Bijektion

$$\arg: \mathbb{S}^1 \setminus \{-1\} \longrightarrow (-\pi, \pi). \quad (4.5.13)$$

**Definition 4.5.8 (Polarkoordinaten)** Sei  $z \in \mathbb{C} \setminus \{0\}$ . Dann bilden die beiden Zahlen  $|z| \in \mathbb{R}^+$  und  $\varphi = \arg(\frac{z}{|z|}) \in (-\pi, \pi]$  die Polarkoordinaten von  $z$ .

Der Null können wir keinen sinnvollen Wert für das Argument zuordnen, weshalb wir diese aus der Definition ausschließen müssen.

Abbildung 4.7: Polarkoordinaten von  $z$  und die komplexe Multiplikation

**Bemerkung 4.5.9** Die komplexe Multiplikation erhält in Polarkoordinaten nun die einfache geometrische Interpretation, dass die Beträge multipliziert und die Winkel addiert werden, siehe auch Abbildung 4.7

## 4.6 Übungen

Übung 4.6.1 (e ist irrational)

Übung 4.6.2 (Algebraische Beziehungen zwischen den Arcusfunktionen)

## Kapitel 5

# Differenzierbarkeit und die Ableitung

Die konzeptuelle Definition der Ableitung einer Funktion kann man mit guter Begründung als Geburtsstunde der modernen Analysis ansehen. Die Geschichte der Differentialrechnung mit ihren beiden Hauptprotagonisten Newton und Leibniz ist eine spannende und wechselhafte Episode der Wissenschaftsgeschichte und gehört unbedingt zu den Sternstunden der Menschheit. Vermutlich gibt es nur wenige Erkenntnisse, die unser Weltbild ebenso verändert haben wie unseren Alltag nachhaltig beeinflussten. In diesem Kapitel wollen wir die Anfangsgründe der Differentialrechnung entwickeln und die ersten Ergebnisse und Rechenregeln dazu vorstellen.

### 5.1 Die Definition der Ableitung

Wir betrachten erneut Funktionen, die auf einem Definitionsbereich  $X$  von reellen oder komplexen Zahlen definiert sein sollen und Werte in  $\mathbb{R}$  oder  $\mathbb{C}$  annehmen dürfen. In der reellen Situation  $X \subseteq \mathbb{R}$  mit einer reellwertigen Funktion ist eine der Motivationen, die Steigungen des Graphen von  $f$  zu definieren und zu bestimmen. Die geometrische Idee ist, dass sich Sekanten zumindest viel einfacher definieren lassen als Tangenten. Die Sekante durch die Punkte  $(x_0, f(x_0))$  und  $(x_0 + \epsilon, f(x_0 + \epsilon))$  des Graphen von  $f$  ist die Gerade

$$g_\epsilon(x) = \frac{f(x_0 + \epsilon) - f(x_0)}{\epsilon}(x - x_0) + f(x_0), \quad (5.1.1)$$

wobei  $x \in \mathbb{R}$ . Hierzu müssen die Punkte  $x_0$  und  $x_0 + \epsilon$  natürlich verschieden sein, also  $\epsilon \neq 0$  erfüllen. Zudem müssen  $x_0$  und  $x_0 + \epsilon$  beide im Definitionsbereich von  $f$  liegen. Es ist daher naheliegend, die Tangente an  $(x_0, f(x_0))$  als Grenzwert der Sekanten  $\{g_\epsilon\}_{\epsilon \neq 0}$  für  $\epsilon \rightarrow 0$  zu definieren. Insbesondere ist die Steigung  $m$  der Tangente dann

$$m = \lim_{\epsilon \rightarrow 0} \frac{f(x_0 + \epsilon) - f(x_0)}{\epsilon}. \quad (5.1.2)$$

Es gilt also die Natur dieses Grenzwertes zu verstehen, siehe auch Abbildung 5.1. Der Quotient  $\frac{1}{\epsilon}(f(x_0 + \epsilon) - f(x_0))$  heißt auch *Differenzenquotient* von  $f$  bei  $x_0$ .

Eine zweite Motivation für die Ableitung ist der Wunsch, eine komplizierte Funktion  $f$  durch eine einfachere zu ersetzen und den dabei entstehenden Fehler gut zu kontrollieren. Diese Ersetzung darf dabei von einem zuvor ausgewählten Punkt  $x_0$  abhängen. Die einfachste Näherung der Funktionswerte  $f(x)$  für  $x$  in der Nähe von  $x_0$  besteht dann darin, die Funktion durch die konstante Funktion  $f(x_0)$  zu ersetzen, siehe auch Abbildung 5.2. Diese Ersetzung ist typischerweise eine sehr grobe und in den meisten Fällen auch unbrauchbare Näherung. Insbesondere gibt es für  $x$  weit von  $x_0$  entfernt keinen Grund, dass die Funktionswerte  $f(x_0)$  und  $f(x)$  nahe beieinander sind. Allerdings können wir die

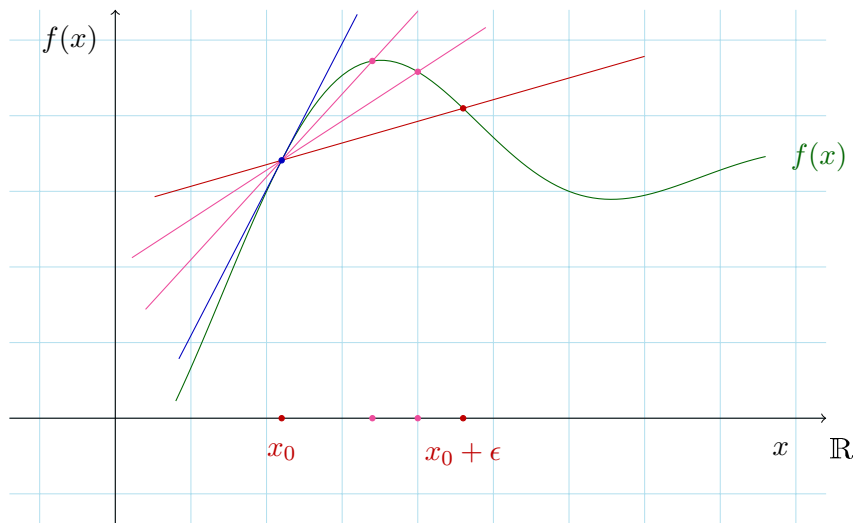
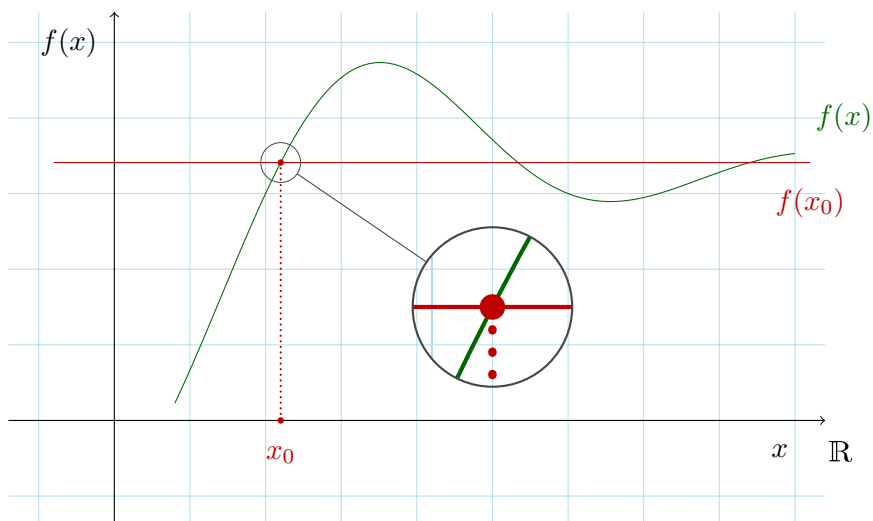


Abbildung 5.1: Die Tangente als Grenzwert von Sekanten.

Abbildung 5.2: Näherung von  $f(x)$  durch  $f(x_0)$ .

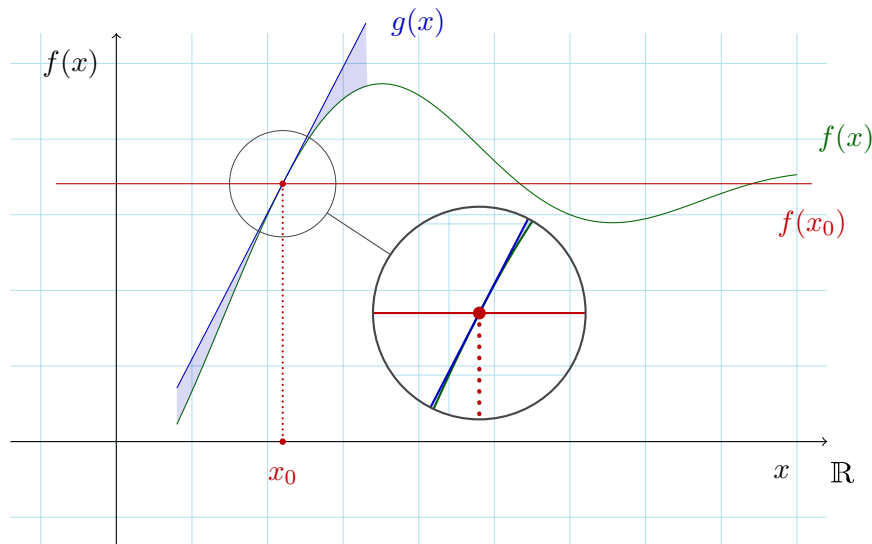


Abbildung 5.3: Die lineare Approximation  $g(x) = m(x - x_0) + f(x_0)$  an  $f(x)$  ist besser als alle anderen linearen Approximationen, wodurch die Steigung  $m$  festgelegt wird.

Stetigkeit von  $f$  bei  $x_0$  gerade so interpretieren, dass die Näherung von  $f(x)$  durch  $f(x_0)$  zumindest im Grenzwert  $x \rightarrow x_0$  brauchbar wird. Stetigkeit ist ja gerade dazu äquivalent, dass

$$f(x_0) = \lim_{x \rightarrow x_0} f(x) \quad (5.1.3)$$

gilt. Daher sind für  $x$  nahe genug bei  $x_0$  die Funktionswerte  $f(x)$  nahe bei  $f(x_0)$ , die  $\epsilon\delta$ -Definition von Stetigkeit quantifiziert genau dies.

Der nächste Schritt besteht darin, die Approximation zu verbessern und die konstante Approximation durch eine *lineare* zu ersetzen. Es wird also eine Gerade  $g(x) = mx + c$  gesucht, die die Funktion  $f$  in der Nähe von  $x_0$  besonders gut approximiert. Der Ansatz ist daher

$$g(x) = m(x - x_0) + f(x_0), \quad (5.1.4)$$

wobei der konstante Term sicherstellt, dass für  $x = x_0$  die Approximation exakt wird. Um nun die Steigung  $m$  dieser Approximationsgeraden bestmöglich zu bestimmen, fordert man, dass der Fehler

$$R(x) = f(x) - g(x) \quad (5.1.5)$$

für  $x \rightarrow x_0$  schneller klein wird als ein linearer Fehler. Mit anderen Worten soll der Fehler relativ zu einem linearen Verhalten immer noch klein werden. Man fordert daher

$$\lim_{x \rightarrow x_0} \frac{R(x)}{x - x_0} = 0, \quad (5.1.6)$$

siehe auch Abbildung 5.3. Mit dieser Idee kann man die komplizierte Funktion  $f$  also durch die einfachere lineare Näherung ersetzen, wenn die Steigung  $m$  so gewählt ist, dass (5.1.6) gilt. Ausgeschrieben bedeutet dies aber

$$m = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}, \quad (5.1.7)$$

was auf den selben Grenzprozess wie für die Tangente (5.1.2) führt. Nun ist aber mit diesem Gesichtspunkt bereits klar, dass man nicht bei einer linearen Approximation stehen bleiben muss. Anstelle dessen kann man ein Polynom

$$g_N(x) = a_N x^N + \dots + a_1 x + a_0 \quad (5.1.8)$$

mit zu bestimmenden Zahlen  $a_N, \dots, a_0$  suchen, so dass die Approximation an  $f$  unter allen solchen Polynomen des Grades  $N$  am besten wird. Man verlangt, dass der Fehler

$$R_N(x) = f(x) - g_N(x) \quad (5.1.9)$$

für  $x \rightarrow x_0$  schneller klein wird als die entsprechende  $N$ -te Potenz, also

$$\lim_{x \rightarrow x_0} \frac{R_N(x)}{(x - x_0)^N} = 0. \quad (5.1.10)$$

Der Fehler ist also von höherer Ordnung als  $N$ . Dies ist der erste Schritt einer Hierarchie nach der Stetigkeit, die den Fall  $N = 0$  charakterisiert. Wir werden auf diese Approximationen höherer Ordnung noch zu sprechen kommen, konzentrieren uns nun aber auf den linearen Fall  $N = 1$ . Es geht also um die Frage, wie die Existenz der Limites (5.1.2) beziehungsweise (5.1.7) zu verstehen ist.

**Definition 5.1.1 (Differenzierbarkeit und Ableitung)** Sei  $X$  eine nicht-leere Teilmenge von reellen oder komplexen Zahlen. Eine Funktion  $f: X \rightarrow \mathbb{K}$  heißt differenzierbar bei  $x_0 \in X$ , wenn  $x_0$  ein Häufungspunkt von  $X$  ist und der Grenzwert

$$f'(x_0) = \lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0} \quad (5.1.11)$$

existiert. In diesem Fall heißt der Wert  $f'(x_0) \in \mathbb{K}$  die Ableitung von  $f$  bei  $x_0$ . Die Funktion  $f$  heißt differenzierbar mit Ableitung  $f': X \rightarrow \mathbb{K}$ , wenn  $f$  an allen Punkten von  $X$  differenzierbar ist.

**Bemerkung 5.1.2** Sei  $f: X \rightarrow \mathbb{K}$  eine Funktion.

- i.) Die Annahme, dass  $x_0$  ein Häufungspunkt von  $X$  sein soll, ist mehr als eine nützliche Konvention. Andernfalls ist der Grenzwert (5.1.11) nicht definierbar, da zum einen  $x \neq x_0$  bei der Auswertung von  $\frac{f(x) - f(x_0)}{x - x_0}$  gelten muss, andererseits ein auf  $x_0$  hin gerichtetes Netz erforderlich ist.
- ii.) Wir werden im Folgenden oftmals einen offenen Definitionsbereich  $X$  voraussetzen. In diesem Fall sind alle Punkte aus  $X$  Häufungspunkte. Allerdings gibt es durchaus interessante Situationen, in denen  $X$  nicht offen ist und wir trotzdem von Differenzierbarkeit sprechen wollen. Als Beispiel sei ein abgeschlossenes Intervall  $X = [a, b] \subseteq \mathbb{R}$  genannt. Differenzierbarkeit bei  $x_0 = a$  bedeutet dann, dass

$$f'(a) = \lim_{\substack{x \rightarrow a \\ x > a}} \frac{f(x) - f(a)}{x - a} \quad (5.1.12)$$

existiert. Entsprechend bedeutet Differenzierbarkeit bei  $x_0 = b$  die Existenz des Grenzwertes

$$f'(b) = \lim_{\substack{x \rightarrow b \\ x < b}} \frac{f(x) - f(b)}{x - b}. \quad (5.1.13)$$

Um diese Situation zu betonen, spricht man auch von *Differenzierbarkeit von rechts* in (5.1.12) beziehungsweise von *Differenzierbarkeit von links* in (5.1.13), siehe auch Abbildung 5.4.

- iii.) Die geometrische Interpretation der Ableitung  $f'(x_0)$  wurde bereits als die Steigung der Tangente bei  $(x_0, f(x_0))$  gegeben. Die analytische Interpretation ist die der *besten linearen Approximation*. Ist  $f$  bei  $x_0$  differenzierbar, so heißt die Gerade

$$T(x) = f'(x_0)(x - x_0) + f(x_0) \quad (5.1.14)$$

die *Tangente* an  $f$  bei  $(x_0, f(x_0))$ . Der entscheidende Punkt ist, dass wir nur dann eine Definition einer Tangente zur Verfügung haben, wenn  $f$  bei  $x_0$  überhaupt differenzierbar ist. Diese geometrische Interpretation setzt natürlich voraus, dass wir sowohl eine reelle Variable als auch reelle Werte betrachten.



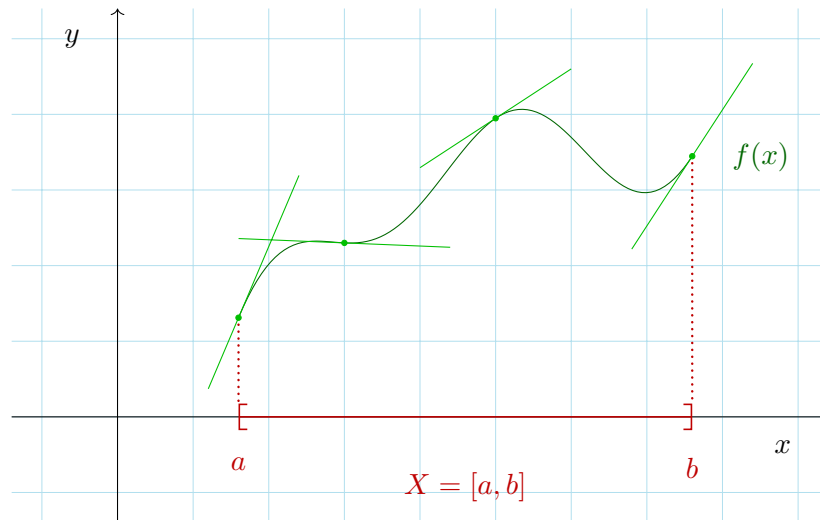


Abbildung 5.4: Geometrische Interpretation der Ableitung als lokale Steigung, insbesondere an den Rändern des Intervalls  $X = [a, b]$ .

Wir diskutieren nun einige Beispiele von differenzierbaren Funktionen, die wir später deutlich erweitern werden, sobald uns die entsprechenden Rechenregeln zur Verfügung stehen.

**Beispiel 5.1.3 (Differenzierbare Funktionen)**

i.) Sei  $c \in \mathbb{C}$ . Dann ist die konstante Funktion

$$f(x) = c \quad (5.1.15)$$

an allen Stellen  $x_0 \in \mathbb{K}$  differenzierbar und es gilt

$$f'(x_0) = 0. \quad (5.1.16)$$

ii.) Sei  $c \in \mathbb{C}$ . Dann ist die Funktion

$$f(x) = cx \quad (5.1.17)$$

an allen Stellen  $x_0 \in \mathbb{K}$  differenzierbar und es gilt

$$f'(x_0) = c. \quad (5.1.18)$$

Dies folgt sofort aus

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{cx - cx_0}{x - x_0} = c \quad (5.1.19)$$

für alle  $x \neq x_0$ .

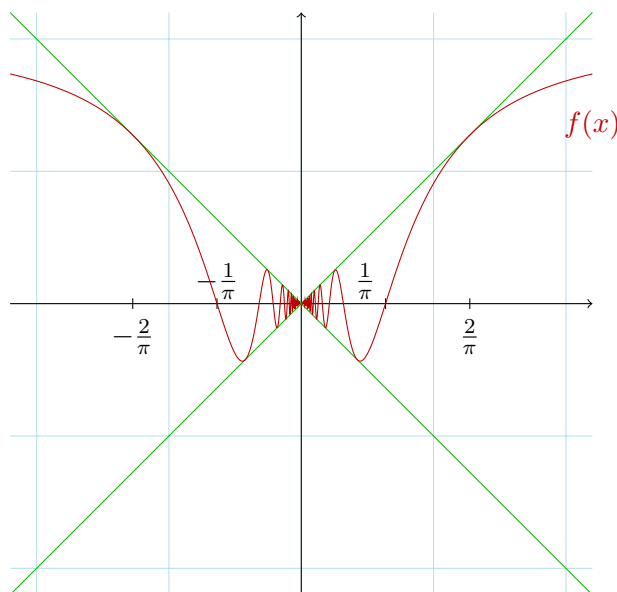
iii.) Die Exponentialfunktion  $\exp$  ist bei  $x_0 = 0$  differenzierbar und erfüllt

$$\exp'(0) = 1, \quad (5.1.20)$$

denn es gilt

$$\lim_{\substack{x \rightarrow 0 \\ x \neq 0}} \frac{\exp(x) - 1}{x} = \lim_{\substack{x \rightarrow 0 \\ x \neq 0}} \sum_{n=1}^{\infty} \frac{x^{n-1}}{n!} = \sum_{n=1}^{\infty} \lim_{\substack{x \rightarrow 0 \\ x \neq 0}} \frac{x^{n-1}}{n!} = 1. \quad (5.1.21)$$

Hier verwenden wir die Rechenregeln für konvergente Reihen, um auf die Konvergenz der Potenzreihen  $\sum_{n=1}^{\infty} \frac{x^{n-1}}{n!}$  zu schließen, und die Stetigkeit konvergenter Potenzreihen, um den Grenzwert zu bestimmen. Wir werden dieses Beispiel in Kürze erneut aufgreifen.

Abbildung 5.5: Die Funktion  $f(x) = x \sin(\frac{1}{x})$  für  $x \neq 0$ .

iv.) Als ein Beispiel einer stetigen aber nicht differenzierbaren Funktion betrachten wir  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0. \end{cases} \quad (5.1.22)$$

siehe Abbildung 5.5. Zuerst ist klar, dass  $f$  abseits von  $x = 0$  stetig ist, als Verkettung von stetigen Funktionen. Da  $|\sin(\frac{1}{x})| \leq 1$  für alle  $x \neq 0$  gilt, folgt, dass

$$\lim_{\substack{x \rightarrow 0 \\ x \neq 0}} f(x) = 0 \quad (5.1.23)$$

gilt, da der Vorfaktor  $x$  den Grenzwert 0 besitzt. Dies erklärt die Stetigkeit bei  $x_0 = 0$ . Für die Nicht-Differenzierbarkeit bei  $x_0 = 0$  betrachten wir den Differenzenquotienten

$$g(x) = \frac{f(x) - f(x_0)}{x - x_0} = \sin\left(\frac{1}{x}\right) \quad (5.1.24)$$

für  $x \neq 0$  bei  $x_0 = 0$ . Die Differenzierbarkeit wäre dann zur Existenz des Grenzwertes  $\lim_{x \rightarrow 0} g(x)$  äquivalent. Dieser Grenzwert existiert aber sicherlich nicht, da beispielsweise  $g(x_n) = 1$  für  $x_n = \frac{1}{\frac{\pi}{2} + 2n\pi}$  aber  $g(y_n) = -1$  für  $y_n = \frac{1}{\frac{3\pi}{2} + 2n\pi}$  und beide Folgen  $(x_n)_{n \in \mathbb{N}}$  und  $(y_n)_{n \in \mathbb{N}}$  gegen Null konvergieren. Der Graph der Funktionen  $f$  und  $g$  ist in Abbildung 5.6 zu sehen und realisiert die heuristische Vorstellung der Funktion  $k$  aus Abbildung 3.5 durch konkrete Funktionen.

Nach diesen Beispielen schließen wir mit einer Umformulierung der Differenzierbarkeit.

**Proposition 5.1.4** Sei  $f: X \rightarrow \mathbb{K}$  und  $x_0 \in X$  ein Häufungspunkt. Dann sind folgende Aussagen äquivalent:

- i.) Die Funktion  $f$  ist bei  $x_0$  differenzierbar.
- ii.) Es gibt eine bei  $x_0$  stetige Funktion  $\varphi: X \rightarrow \mathbb{K}$  mit

$$f(x) = f(x_0) + \varphi(x)(x - x_0). \quad (5.1.25)$$

In diesem Fall ist  $f'(x_0) = \varphi(x_0)$ .

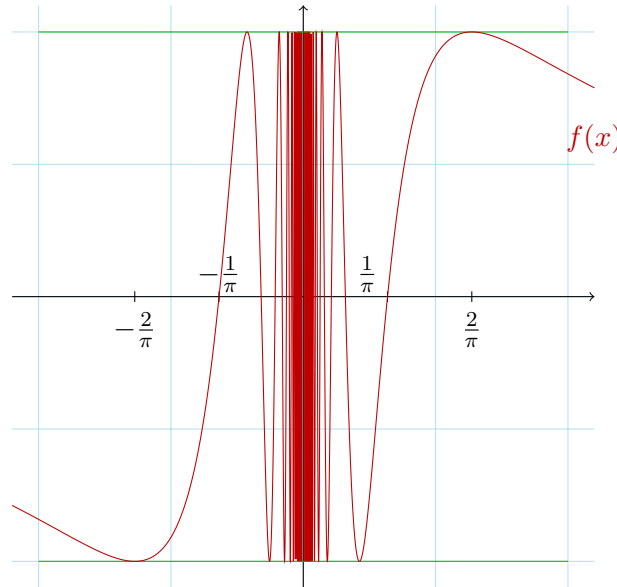


Abbildung 5.6: Die Funktion  $f(x) = x \sin(\frac{1}{x})$  besitzt eine stetige Fortsetzung bei  $x_0 = 0$ . Die Funktion  $g(x) = \sin(\frac{1}{x})$  dagegen oszilliert zu stark, um einen Grenzwert bei  $x_0 = 0$  zu besitzen, bleibt aber trotzdem beschränkt.

iii.) Es gilt

$$f(x) = f(x_0) + m(x - x_0) + R(x) \quad (5.1.26)$$

mit einer Funktion  $R: X \rightarrow \mathbb{K}$ , die

$$\lim_{x \rightarrow x_0} \frac{R(x)}{x - x_0} = 0 \quad (5.1.27)$$

erfüllt. In diesem Fall ist  $f'(x_0) = m$ .

BEWEIS: Ist  $f$  differenzierbar in  $x_0$  mit Ableitung  $f'(x_0)$ , so ist

$$\varphi(x) = \begin{cases} \frac{f(x) - f(x_0)}{x - x_0} & x \neq x_0 \\ f'(x_0) & x = x_0 \end{cases}$$

eine bei  $x = x_0$  stetige Funktion. Dies ist gerade die Definition der Differenzierbarkeit. Ist umgekehrt  $\varphi$  wie in (5.1.25) gegeben, so gilt

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

für  $x \neq x_0$ . Die Stetigkeit von  $\varphi$  bei  $x = x_0$  zeigt dann, dass  $f$  bei  $x_0$  differenzierbar ist. Die Äquivalenz von ii.) und iii.) ist auch klar, da  $\varphi(x) = m + \frac{R(x)}{x - x_0}$  für  $x \neq x_0$  gilt und die Stetigkeit von  $\varphi$  bei  $x = x_0$  gerade (5.1.27) bedeutet.  $\square$

**Korollar 5.1.5** Ist  $f$  bei  $x_0$  differenzierbar, so ist  $f$  bei  $x_0$  stetig.

BEWEIS: Da alle Funktionen auf der rechten Seite von (5.1.25) bei  $x = x_0$  stetig sind, gilt dies nach den Rechenregeln für stetige Funktionen auch für die linke Seite.  $\square$

Beispiel 5.1.3, iv.), zeigt dann, dass die Umkehrung typischerweise falsch ist: Differenzierbarkeit ist eine echt stärkere Eigenschaft als Stetigkeit. Drastischere Beispiele werden wir später noch sehen.

## 5.2 Rechenregeln für differenzierbare Funktionen

Um unser Repertoire an differenzierbaren Funktionen zu erweitern, wollen wir nun einige allgemeine Rechenregeln für differenzierbare Funktionen etablieren, wie wir dies bereits bei stetigen Funktionen getan haben.

**Proposition 5.2.1** *Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen, und  $f, g: X \rightarrow \mathbb{K}$  seien differenzierbar bei  $x_0 \in X$ .*

*i.) Für alle  $\alpha, \beta \in \mathbb{K}$  ist die Funktion  $\alpha f + \beta g: X \rightarrow \mathbb{K}$  bei  $x_0$  differenzierbar mit*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0). \quad (5.2.1)$$

*ii.) Die Funktion  $fg: X \rightarrow \mathbb{K}$  ist differenzierbar bei  $x_0$  mit*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0). \quad (5.2.2)$$

*iii.) Ist  $f(x_0) \neq 0$ , so ist die Funktion  $\frac{1}{f}$ , definiert auf der Umgebung  $U = \{x \in X \mid f(x) \neq 0\} \subseteq X$  von  $x_0$  in  $X$ , ebenfalls bei  $x_0$  differenzierbar mit*

$$\left(\frac{1}{f}\right)'(x_0) = -\frac{f'(x_0)}{f(x_0)^2}. \quad (5.2.3)$$

BEWEIS: Für den ersten Teil betrachten wir den Differenzenquotienten

$$\frac{(\alpha f + \beta g)(x) - (\alpha f + \beta g)(x_0)}{x - x_0} = \alpha \frac{f(x) - f(x_0)}{x - x_0} + \beta \frac{g(x) - g(x_0)}{x - x_0}$$

für  $x \neq x_0$  und  $x \in X$ . Da die rechte Seite nach Voraussetzung und den Rechenregeln für Grenzwerte den Grenzwert  $\alpha f'(x_0) + \beta g'(x_0)$  für  $x \rightarrow x_0$  besitzt, erhalten wir den ersten Teil. Als nächstes betrachten wir

$$\begin{aligned} \frac{(fg)(x) - (fg)(x_0)}{x - x_0} &= \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x)(g(x) - g(x_0)) + (f(x) - f(x_0))g(x_0)}{x - x_0} \\ &= \frac{f(x) - f(x_0)}{x - x_0} g(x_0) + f(x) \frac{g(x) - g(x_0)}{x - x_0} \end{aligned}$$

für  $x \in X \setminus \{x_0\}$ . Da  $f$  bei  $x = x_0$  stetig ist und die Grenzwerte der einzelnen Differenzenquotienten existieren, folgt Teil ii.) mit (5.2.2). Für den letzten Teil bemerken wir zunächst, dass die Stetigkeit von  $f$  bei  $x_0$  zeigt, dass  $U$  eine Umgebung von  $x_0$  in  $X$  ist. Daher können wir von der Frage nach Differenzierbarkeit von  $\frac{1}{f}$  bei  $x_0$  überhaupt erst sprechen. Wir betrachten nun für  $x \in U \setminus \{x_0\}$  den Differenzenquotient

$$\begin{aligned} \frac{\left(\frac{1}{f}\right)(x) - \left(\frac{1}{f}\right)(x_0)}{x - x_0} &= \frac{\frac{1}{f(x)} - \frac{1}{f(x_0)}}{x - x_0} \\ &= \frac{f(x_0) - f(x)}{f(x)f(x_0)(x - x_0)} \\ &= -\frac{f(x) - f(x_0)}{x - x_0} \frac{1}{f(x)f(x_0)}. \end{aligned}$$

Wieder zeigt die Stetigkeit von  $f$  bei  $x_0$  sowie die Rechenregeln für Grenzwerte von Funktionen, dass der Grenzwert der rechten Seite existiert und (5.2.3) liefert.  $\square$

Die Produktregel in (5.2.2) wird auch als *Leibniz-Regel* bezeichnet. Aus (5.2.3) und (5.2.2) erhält man durch Kombination die *Quotientenregel*

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}, \quad (5.2.4)$$

sofern eben  $g(x_0) \neq 0$  und beide Funktionen bei  $x_0$  differenzierbar sind. Umgekehrt impliziert (5.2.4) wieder (5.2.3), indem man  $f = 1$  verwendet und Beispiel 5.1.3, *i.*), zur Berechnung von  $f'(x_0) = 0$  heranzieht.

**Korollar 5.2.2** *Ein Polynom  $p(x) = a_n x^n + \dots + a_1 x + a_0$  ist auf ganz  $\mathbb{K}$  differenzierbar mit*

$$p'(x) = n a_n x^{n-1} + \dots + 2 a_2 x + a_1. \quad (5.2.5)$$

BEWEIS: Nach Proposition 5.2.1, *i.*), genügt es, die Differenzierbarkeit von  $p(x) = x^n$  zu verstehen. Hierfür benutzen wir die Leibniz-Regel aus Proposition 5.2.1, *ii.*), für  $p(x) = x$  sowie Beispiel 5.1.3, *ii.*), und eine Induktion nach  $n$ .  $\square$

**Korollar 5.2.3** *Eine rationale Funktion  $f = \frac{p(x)}{q(x)}$  mit Polynomen  $p, q \in \mathbb{K}[x]$  ist auf ihrem natürlichen Definitionsbereich  $X = \mathbb{K} \setminus \{x \in \mathbb{K} \mid q(x) = 0\}$  differenzierbar, und*

$$f'(x) = \frac{p'(x)q(x) - p(x)q'(x)}{q(x)^2} \quad (5.2.6)$$

*gilt für alle  $x \in X$ .*

BEWEIS: Dies folgt aus Korollar 5.2.2 sowie der Quotientenregel aus (5.2.4).  $\square$

Die Exponentialfunktion  $\exp$  ist ebenso wie die trigonometrischen Funktionen auf ihren jeweiligen maximalen Definitionsbereichen differenzierbar:

**Proposition 5.2.4** *Sei  $a \in \mathbb{C}$ .*

*i.) Die Exponentialfunktion  $\exp: \mathbb{C} \rightarrow \mathbb{C}$  ist überall differenzierbar. Es gilt*

$$\exp'(z) = \exp(z) \quad (5.2.7)$$

*für alle  $z \in \mathbb{C}$ .*

*ii.) Die Funktion  $f(z) = \exp(az)$  ist ebenfalls für alle  $z \in \mathbb{C}$  differenzierbar. Es gilt*

$$f'(z) = a f(z) = a \exp(az). \quad (5.2.8)$$

*iii.) Die Sinus- und Kosinusfunktion sind auf ganz  $\mathbb{C}$  differenzierbar. Es gilt*

$$\sin'(z) = \cos(z) \quad \text{sowie} \quad \cos'(z) = -\sin(z). \quad (5.2.9)$$

*iv.) Die Tangensfunktion ist auf  $\mathbb{C} \setminus \{k\pi + \frac{\pi}{2} \mid k \in \mathbb{Z}\}$  differenzierbar. Es gilt*

$$\tan'(z) = \frac{1}{(\cos(z))^2}. \quad (5.2.10)$$

BEWEIS: Wir zeigen (5.2.8). Dazu betrachten wir den Differenzenquotient

$$\frac{e^{ax} - 1}{x} = \sum_{n=1}^{\infty} \frac{a}{n!} (ax)^{n-1}$$

für  $x \neq 0$ . Da nach den Rechenregeln für konvergente Reihen die rechte Seite immer noch für alle  $x \neq 0$  konvergiert, und die Konvergenz für  $x = 0$  trivialerweise auch gilt, folgt aus der Stetigkeit von konvergenten Potenzreihen

$$\lim_{x \rightarrow 0} \frac{e^{ax} - 1}{x} = a,$$

was (5.2.8) bei  $x = 0$  zeigt. Damit folgt aber auch die Existenz des Grenzwertes

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{e^{ax} - e^{ax_0}}{x - x_0} &= \lim_{x \rightarrow x_0} e^{ax_0} \frac{e^{a(x-x_0)} - 1}{x - x_0} \\ &= \lim_{x \rightarrow 0} e^{ax_0} \frac{e^{ax} - 1}{x} \\ &= e^{ax_0} a, \end{aligned}$$

was den zweiten Teil und damit auch den ersten Teil zeigt. Da wir  $\sin(z)$  und  $\cos(z)$  über die Exponentialfunktion definiert haben, folgt *iii.*) und insbesondere (5.2.9) direkt aus Proposition 5.2.1, *i.*), sowie dem zweiten Teil. Dann ist der letzte Teil eine einfache Anwendung der Quotientenregel (5.2.4) sowie von Satz 4.4.4, *v.*), zur Lage der Nullstellen der Kosinusfunktion. Es gilt

$$\tan'(z) = \left( \frac{\sin(z)}{\cos(z)} \right)' = \frac{\cos(z) \cos(z) + \sin(z) \sin(z)}{(\cos(z))^2} = \frac{1}{(\cos(z))^2}$$

für alle  $z$  im Definitionsbereich von  $\tan$ . □

Die nächste wichtige Rechenregel zur Differenzierbarkeit ist die *Kettenregel*, deren Bedeutung kaum überschätzt werden kann:

**Proposition 5.2.5** *Seien  $X$  und  $Y$  nicht-leere Teilmengen der reellen oder komplexen Zahlen. Sei weiter Sei  $f: X \rightarrow Y$  eine Funktion und  $g: Y \rightarrow \mathbb{K}$ . Ist  $f$  bei  $x_0 \in X$  und  $g$  bei  $f(x_0) \in Y$  differenzierbar, so ist die Verkettung  $g \circ f: X \rightarrow \mathbb{K}$  bei  $x_0$  differenzierbar mit*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0). \quad (5.2.11)$$

BEWEIS: Wir verwenden die Charakterisierung der Differenzierbarkeit aus Proposition 5.1.4. Wir finden daher eine Funktion  $\varphi: X \rightarrow \mathbb{K}$  mit

$$f(x) = f(x_0) + (x - x_0)\varphi(x),$$

so dass  $\varphi$  bei  $x_0$  stetig ist. Ebenso finden wir  $\psi: Y \rightarrow \mathbb{K}$  mit

$$g(y) = g(y_0) + (y - y_0)\psi(y),$$

derart, dass  $\psi$  bei  $y_0 = f(x_0)$  stetig ist. Es folgt

$$\begin{aligned} (g \circ f)(x) - (g \circ f)(x_0) &= g(f(x)) - g(f(x_0)) \\ &= (f(x) - f(x_0))\psi(f(x)) \\ &= (x - x_0)\varphi(x)\psi(f(x)). \end{aligned}$$

Da die Funktion  $x \mapsto \varphi(x)\psi(f(x))$  bei  $x = x_0$  wegen der Stetigkeit von  $\varphi, \psi$  und  $f$  wieder stetig ist, folgt, dass  $g \circ f$  bei  $x = x_0$  differenzierbar ist mit der Ableitung

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

Hier verwenden wir erneut Proposition 5.1.4. □

Mit der Kettenregel können wir nun auch kompliziertere verkettete Funktionen differenzieren. Als erstes Beispiel macht man sich klar, dass Proposition 5.2.4, *ii.*), alternativ aus dem ersten Teil der Proposition und der Kettenregel folgt, da wir  $f(x) = \exp(ax)$  als Verkettung der Exponentialfunktion  $\exp$  mit der linearen Funktion  $x \mapsto ax$  auffassen können. Weitere Beispiele finden sich in Übung ??.

Die vorerst letzte Rechenregel handelt von der Differenzierbarkeit der Umkehrfunktion. Da wir etliche wichtige Funktionen wie den Logarithmus oder die Arcusfunktionen als Umkehrfunktionen definiert haben, wird die folgende Proposition dazu genutzt werden können, deren Ableitungen auszurechnen.

**Proposition 5.2.6** *Seien  $X, Y$  nicht-leere Teilmengen von reellen oder komplexen Zahlen und  $f: X \rightarrow Y$  bijektiv. Ist  $f$  bei  $x_0 \in X$  differenzierbar mit Ableitung  $f'(x) \neq 0$  und  $f^{-1}$  bei  $y_0 = f(x_0)$  stetig, so ist die Umkehrfunktion  $f^{-1}: Y \rightarrow X$  bei  $y_0$  differenzierbar mit Ableitung*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}. \quad (5.2.12)$$

BEWEIS: Die Differenzierbarkeit von  $f$  bei  $x_0$  ist nach Proposition 5.1.4 äquivalent zur Existenz einer stetigen Funktion  $\varphi: X \rightarrow \mathbb{K}$  mit

$$f(x) = f(x_0) + (x - x_0)\varphi(x),$$

wobei dann  $f'(x_0) = \varphi(x_0)$  gilt. Insbesondere gilt  $\varphi(x_0) \neq 0$  nach Voraussetzung. Für  $x \neq x_0$  ist  $\varphi(x) \neq 0$ , da  $f$  injektiv ist, womit  $\varphi \neq 0$  auf ganz  $X$  folgt. Sei  $y \in Y$  mit  $x = f^{-1}(y)$ , dann gilt

$$f^{-1}(y) - f^{-1}(y_0) = x - x_0 = \frac{f(x) - f(x_0)}{\varphi(x)} = \frac{1}{\varphi(f^{-1}(y))}(y - y_0).$$

Da nach Voraussetzung  $f^{-1}$  bei  $y = y_0$  stetig ist und  $\varphi$  ebenfalls stetig ist, ist die Verkettung  $y \mapsto \frac{1}{\varphi(f^{-1}(y))}$  bei  $y = y_0$  stetig. Nach Proposition 5.1.4 ist daher  $f^{-1}$  bei  $y = y_0$  differenzierbar mit Ableitung

$$(f^{-1})'(y_0) = \frac{1}{\varphi(f^{-1}(y_0))} = \frac{1}{\varphi(x_0)} = \frac{1}{f'(x_0)},$$

was den Beweis abschließt. □

Wir nutzen diese Proposition nun, um folgende Ableitungen zu bestimmen:

**Proposition 5.2.7**

*i.) Die Logarithmusfunktion  $\log: \mathbb{R}^+ \rightarrow \mathbb{R}$  ist überall differenzierbar, und es gilt*

$$\log'(x) = \frac{1}{x}. \quad (5.2.13)$$

*ii.) Die Arcussinusfunktion  $\arcsin: (-1, 1) \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$  ist überall differenzierbar, und es gilt*

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}. \quad (5.2.14)$$

*iii.) Die Arcuscosinusfunktion  $\arccos: (-1, 1) \rightarrow (0, \pi)$  ist überall differenzierbar, und es gilt*

$$\arccos'(x) = -\frac{1}{\sqrt{1-x^2}}. \quad (5.2.15)$$

*iv.) Die Arcustangensfunktion  $\arctan: \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$  ist überall differenzierbar, und es gilt*

$$\arctan'(x) = \frac{1}{1+x^2}. \quad (5.2.16)$$

BEWEIS: Wir wissen bereits, dass  $\log$ ,  $\arcsin$ ,  $\arccos$  und  $\arctan$  auf ihren jeweiligen Definitionsbereichen stetig und Umkehrfunktionen der differenzierbaren Funktionen  $\exp$ ,  $\sin$ ,  $\cos$  und  $\tan$  sind. Weiter sind die Ableitungen von  $\exp$ ,  $\sin$ ,  $\cos$  und  $\tan$  dort ungleich Null. Dies folgt aus der bekannten Lage der Nullstellen von  $\sin$ ,  $\cos$  und  $\tan$  und ist der Grund, wieso wir in *ii.*) und *iii.*) die Intervallgrenzen ausschließen müssen. Dann sind die Voraussetzungen von Proposition 5.2.6 für alle Punkte des jeweiligen Definitionsbereichs erfüllt und  $\log$ ,  $\arcsin$ ,  $\arccos$ ,  $\arctan$  sind differenzierbar. Es verbleibt also, die Ableitungen konkret zu bestimmen. Für den Logarithmus gilt mit  $y = \exp(x)$

$$\log'(y) = \frac{1}{\exp'(x)} = \frac{1}{\exp(x)} = \frac{1}{y},$$

da ja  $\exp' = \exp$  gilt. Für den Arcussinus erhalten wir mit  $y = \sin(x)$

$$\arcsin'(y) = \frac{1}{\sin'(x)} = \frac{1}{\cos(x)} = \frac{1}{\sqrt{1 - \sin(x)^2}} = \frac{1}{\sqrt{1 - y^2}},$$

wobei wir zum einen  $\sin' = \cos$  und zum anderen  $\cos(x) = \sqrt{1 - \sin(x)^2}$  verwenden, da für  $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$  eben  $\cos(x) > 0$  gilt, womit wir  $1 = \sin(x)^2 + \cos(x)^2$  nach  $\cos(x)$  auflösen können. Die Ableitung (5.2.15) erhält man genauso oder alternativ aus  $\arccos(x) = \frac{\pi}{2} - \arcsin(x)$ , siehe Übung 4.6.2. Schließlich finden wir für den Arcustangens mit  $y = \tan(x)$

$$\arctan'(y) = \frac{1}{\tan'(x)} = \frac{1}{\cos(x)^2} = \frac{1}{1 + \tan(x)^2} = \frac{1}{1 + y^2},$$

was auch die letzte Behauptung zeigt. □

Als letztes Beispiel wollen wir die erreichten Resultate kombinieren, um in Verallgemeinerung von Korollar 5.2.2 allgemeine Potenzen zu differenzieren:

**Proposition 5.2.8** Sei  $\alpha \in \mathbb{R}$  und  $f: \mathbb{R}^+ \rightarrow \mathbb{R}$  durch

$$f(x) = x^\alpha \tag{5.2.17}$$

definiert. Dann ist  $f$  überall differenzierbar, und es gilt

$$f'(x) = \alpha x^{\alpha-1}. \tag{5.2.18}$$

BEWEIS: Wir schreiben  $f$  etwas um. Da  $x > 0$ , gilt  $x = \exp(\log(x))$  und somit  $x^\alpha = \exp(\log(x))^\alpha = \exp(\alpha \log(x))$ . Nun ist  $f$  eine Verkettung von differenzierbaren Funktionen, also nach Proposition 5.2.5 selbst differenzierbar mit Ableitung

$$\begin{aligned} f'(x) &= \exp(\alpha \log(x)) \cdot (\alpha \log(x))' \\ &= \exp(\alpha \log(x)) \alpha \frac{1}{x} \\ &= \alpha x^\alpha \frac{1}{x} \\ &= \alpha x^{\alpha-1}, \end{aligned}$$

wobei wir  $\exp' = \exp$  und Proposition 5.2.7, *i.*), verwendet haben. □

Natürlich können wir auch die Funktion  $f(x) = x^0 = 1$ , also den Fall für  $\alpha = 0$ , direkt differenzieren und erhalten als Ableitung Null, siehe Beispiel 5.1.3, *i.*).



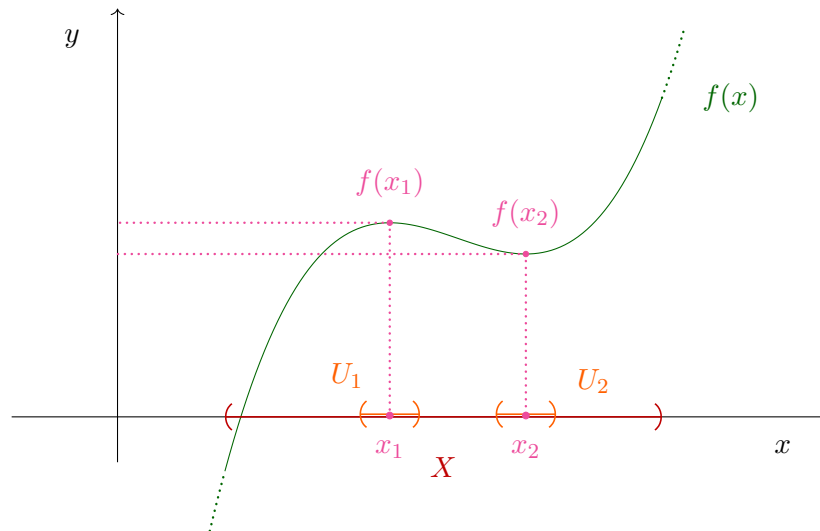


Abbildung 5.7: Der Punkt  $x_1$  ist ein striktes lokales Maximum, da auf der Umgebung  $U_1$  kein Funktionswert von  $f$  größer oder gleich  $f(x_1)$  ist. Entsprechend ist  $x_2$  ein striktes lokales Minimum. Allerdings hat  $f$  weder ein globales Maximum noch ein globales Minimum.

### 5.3 Hauptsätze für differenzierbare Funktionen

In diesem Abschnitt stellen wir nun einige der wichtigsten Eigenschaften von differenzierbaren Funktionen vor. Wir beginnen mit der Frage nach Extrempunkten, die sich so nur für reelle Werte stellen lässt:

**Definition 5.3.1 (Globale Extrema)** Sei  $X$  eine Menge und  $f: X \rightarrow \mathbb{R}$  eine Abbildung.

i.) Ein Punkt  $x_0 \in X$  heißt globales Maximum von  $f$ , falls

$$f(x) \leq f(x_0) \quad (5.3.1)$$

für alle  $x \in X$  gilt. Gilt sogar  $f(x) < f(x_0)$  für alle  $x \in X \setminus \{x_0\}$ , so heißt  $x_0$  striktes globales Maximum.

ii.) Entsprechend definiert man ein globales Minimum und ein striktes globales Minimum.

Ist  $f: X \rightarrow \mathbb{R}$  konstant, so ist jeder Punkt sowohl globales Maximum als auch globales Minimum. Allerdings ist kein Punkt striktes Maximum oder Minimum, sofern  $X$  mindestens zwei Punkte enthält. Interessanter als der rein mengen-theoretische Begriff ist die Frage nach lokalen Extrema:

**Definition 5.3.2 (Lokale Extrema)** Sei  $X \subseteq \mathbb{K}$  und  $f: X \rightarrow \mathbb{R}$  eine Funktion.

i.) Ein Punkt  $x_0 \in X$  heißt (striktes) lokales Maximum von  $f$ , wenn es eine Umgebung  $U$  von  $x_0$  in  $X$  gibt, sodass  $x_0$  ein (striktes) globales Maximum von  $f|_U$  ist.

ii.) Entsprechend definiert man ein (striktes) lokales Minimum von  $f$ .

**Bemerkung 5.3.3 (Lokale Extrema)** Sei  $f: X \rightarrow \mathbb{R}$  mit  $X \subseteq \mathbb{K}$ .

i.) Ein globales Maximum (Minimum) ist auch immer ein lokales Maximum (Minimum). Die Umkehrung muss aber nicht gelten, siehe etwa Abbildung 5.7 für eine heuristische Skizze und Übung ?? für eine konkrete Funktion  $f$  mit diesem Verhalten.

ii.) Man beachte, dass wir einen Umgebungsbegriff benötigen, um von lokalen Extrema zu sprechen.

iii.) Viele mathematischen Fragestellungen lassen sich als Extremwertaufgaben formulieren: Zu einer geeigneten Funktion müssen die lokalen oder globalen Extrema bestimmt werden. Beispiele finden sich in den Übungen, siehe Übung ??.

iv.) Ausgeschrieben ist  $x_0 \in X$  ein lokales Maximum von  $f$ , wenn es ein  $\epsilon > 0$  mit

$$f(x) \leq f(x_0) \quad (5.3.2)$$

für alle  $x \in X$  mit  $|x - x_0| < \epsilon$  gibt.

Im Allgemeinen ist es sehr schwierig, die Lage oder auch nur die Existenz von Extremwerten einer Funktion zu klären. Im Falle einer differenzierbaren Funktion kann dies mit Hilfe der Ableitung zumindest insofern geschehen, indem man es auf eine andere Fragestellung zurückführt:

**Proposition 5.3.4** *Sei  $I = (a, b)$  ein offenes Intervall und  $f: I \rightarrow \mathbb{R}$  eine im Punkt  $x_0 \in (a, b)$  differenzierbare Funktion. Ist  $x_0$  ein lokales Extremum von  $f$ , so gilt*

$$f'(x_0) = 0. \quad (5.3.3)$$

BEWEIS: Sei  $x_0$  beispielsweise ein lokales Maximum und  $\epsilon > 0$  derart gewählt, dass  $f|_U$  bei  $x_0$  ein globales Maximum besitzt, wobei  $U = I \cap (x_0 - \epsilon, x_0 + \epsilon)$ . Durch Verkleinerung von  $\epsilon$  können wir  $U \subseteq I$  also  $(x_0 - \epsilon, x_0 + \epsilon) \subseteq I$  annehmen. Es gilt also  $f(x) \leq f(x_0)$  für alle  $x \in (x_0 - \epsilon, x_0 + \epsilon)$ . Für  $x > x_0$  gilt dann

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0,$$

da  $f(x_0)$  lokales Maximum ist. Entsprechend gilt

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0,$$

für  $x < x_0$ . Da  $f$  aber bei  $x_0$  differenzierbar ist, müssen die Grenzwerte der Differenzenquotienten von links und von rechts übereinstimmen, also

$$f'(x_0) = \lim_{x \nearrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \searrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Da der erste Grenzwert aber  $\geq 0$  ist, der zweite hingegen  $\leq 0$ , bleibt nur  $f'(x_0) = 0$  als einzige Möglichkeit. Der Fall eines lokalen Minimums ist analog.  $\square$

Damit ist  $f'(x_0) = 0$  also eine *notwendige* Bedingung für die Existenz eines lokalen Extremums bei  $x_0$ . Einfache Gegenbeispiele zeigen jedoch, dass aus  $f'(x_0) = 0$  noch nicht notwendigerweise folgt, dass  $x_0$  ein lokales Extremum ist, siehe etwa Abbildung 5.8. Interessanter ist die Situation, wenn wir die Existenz von Extrema aus anderen Gründen garantieren können. Hier kennen wir mit Korollar 3.5.8 bereits ein hinreichendes Szenario: Auf einem kompakten Intervall nimmt eine stetige Funktion Minimum und Maximum an. Dies führt zu folgendem *Satz von Rolle*:

**Satz 5.3.5 (Rolle)** *Sei  $I = [a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall mit  $a < b$ , und sei  $f: I \rightarrow \mathbb{R}$  eine stetige Funktion mit  $f(a) = f(b)$ . Ist  $f$  auf  $(a, b)$  differenzierbar, so gibt es ein  $x_0 \in (a, b)$  mit*

$$f'(x_0) = 0. \quad (5.3.4)$$

BEWEIS: Ist  $f$  auf  $[a, b]$  konstant, so gilt (5.3.4) trivialerweise für alle  $x_0 \in (a, b)$ . Sei also  $f$  nicht konstant. Dann gibt es also ein  $\xi \in (a, b)$  mit entweder  $f(\xi) > f(a)$  oder  $f(\xi) < f(a)$ . Im ersten Fall kann  $a$  nicht das nach Korollar 3.5.8 existierende globale Maximum sein, ebenso wie  $b$ , da  $f(a) = f(b)$ . Also gibt es ein globales Maximum  $x_0 \in (a, b)$ , für welches nach Proposition 5.3.4 dann  $f'(x_0) = 0$  gilt. Der Fall  $f(\xi) < f(a)$  wird genauso behandelt mit dem globalen Minimum anstelle des globalen Maximums.  $\square$

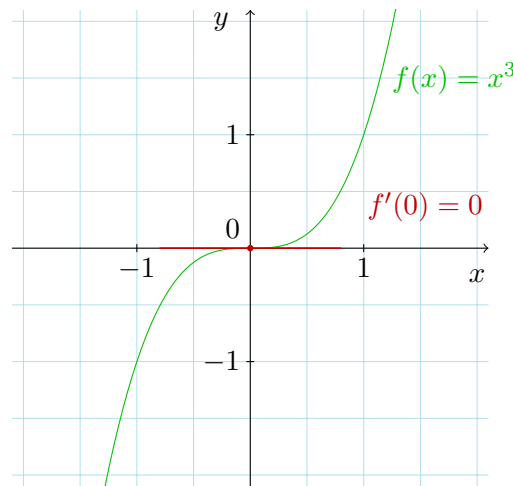


Abbildung 5.8: Die Funktion  $f(x) = x^3$  hat bei  $x = 0$  eine horizontale Tangente,  $f'(0) = 0$ . Trotzdem gilt  $f(x) > 0$  für  $x > 0$  und  $f(x) < 0$  für  $x < 0$ , womit  $x = 0$  kein lokales Extremum sein kann.

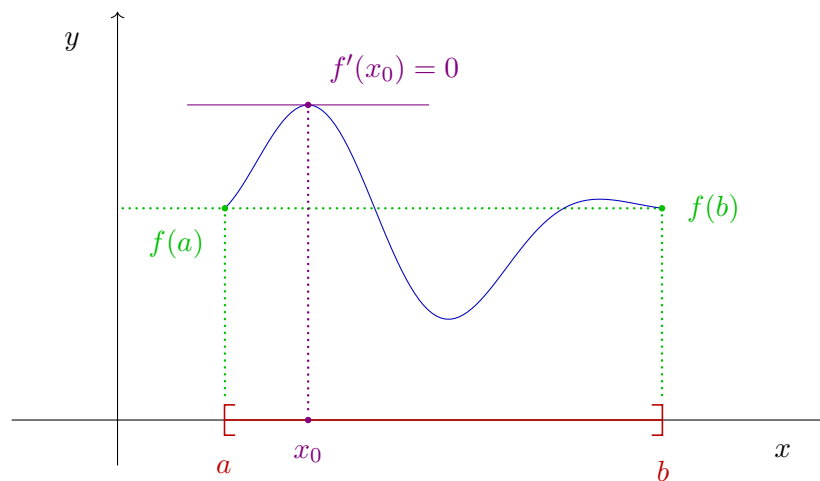


Abbildung 5.9: Situation im Satz von Rolle. Es gibt ein Extremum im offenen Intervall  $(a, b)$ .

Die Lage des Extremums  $x_0 \in (a, b)$  lässt sich im Satz von Rolle jedoch nicht weiter einschränken, ebenso kann über die Anzahl der Extrema nichts weiter gesagt werden. Der Satz von Rolle ist also, wie auch schon Korollar 3.5.8, eine Existenzaussage und kein Verfahren, das Extremum tatsächlich zu finden.

Der nächste wichtige Satz über differenzierbare Funktionen ist der Mittelwertsatz, den wir zunächst folgendermaßen formulieren:

**Satz 5.3.6 (Mittelwertsatz)** *Sei  $I = [a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall mit  $a < b$ . Seien weiter  $f, g: [a, b] \rightarrow \mathbb{R}$  stetige Funktionen, die auf  $(a, b)$  differenzierbar seien.*

*i.) Dann gibt es ein  $x_0 \in (a, b)$  mit*

$$(f(b) - f(a))g'(x_0) = (g(b) - g(a))f'(x_0). \quad (5.3.5)$$

*ii.) Insbesondere gibt es ein  $\xi \in (a, b)$  mit*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}. \quad (5.3.6)$$

BEWEIS: Wir betrachten die Hilfsfunktion

$$\varphi(x) = (f(b) - f(a))g(x) - (g(b) - g(a))f(x)$$

für  $x \in [a, b]$ . Nach den Rechenregeln für stetige und differenzierbare Funktionen ist  $\varphi$  stetig auf  $[a, b]$  und differenzierbar auf  $(a, b)$ . Es gilt

$$\begin{aligned} \varphi(a) &= (f(b) - f(a))g(a) - (g(b) - g(a))f(a) \\ &= f(b)g(a) - g(b)f(a) \\ &= \varphi(b), \end{aligned}$$

womit wir nach dem Satz von Rolle ein  $x_0 \in (a, b)$  finden können, welches  $\varphi'(x_0) = 0$  erfüllt. Da

$$\varphi'(x) = (f(b) - f(a))g'(x) - (g(b) - g(a))f'(x)$$

gilt, folgt (5.3.5). Den zweiten Teil erhält man, indem man  $g(x) = x$  verwendet. □

Wie schon beim Satz von Rolle gibt es im Allgemeinen keine Möglichkeit, mehr über die Lage des Zwischenpunktes  $x_0$  beziehungsweise  $\xi$  zu erfahren. Trotzdem gibt es eine Fülle von nützlichen Konsequenzen, wenn zusätzliche Informationen über die Ableitung  $f'$  zur Verfügung stehen. Wir beginnen mit folgendem Korollar.

**Korollar 5.3.7** *Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig und in  $(a, b)$  differenzierbar mit  $f'(x) = 0$  für alle  $x \in (a, b)$ . Dann ist  $f$  konstant.*

BEWEIS: Sei  $x \in (a, b]$ . Wir betrachten  $f|_{[a, x]}$ , womit wir den Mittelwertsatz anwenden dürfen. Es gibt also ein  $\xi \in (a, x)$  mit

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

Daher folgt  $f(x) = f(a)$  für alle  $x \in (a, b]$ . □

Damit sind die konstanten Funktionen also die Funktionen mit überall verschwindender Ableitung. Wenn die Ableitung nicht verschwindet, aber entweder strikt positiv oder strikt negativ ist, liefert Satz 5.3.6, ii.), einfache Kriterien zur Monotonie:

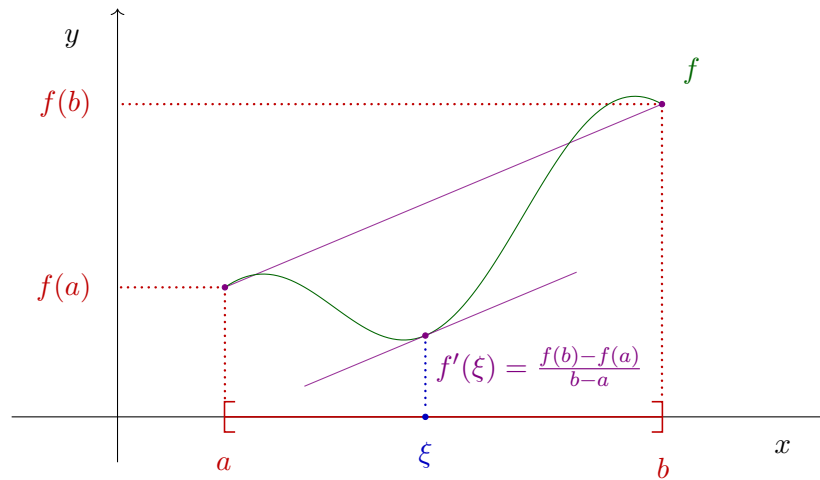


Abbildung 5.10: Die geometrische Interpretation zum Mittelwertsatz. Die *mittlere Steigung* ist durch die Endpunkte als  $\frac{f(b)-f(a)}{b-a}$  festgelegt und wird als tatsächliche Steigung von  $f$  an mindestens einem Punkt  $\xi \in (a, b)$  angenommen.

**Korollar 5.3.8** Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig und auf  $(a, b)$  differenzierbar.

- i.) Gilt  $f' > 0$  für alle  $x \in (a, b)$ , so ist  $f$  strikt monoton wachsend.
- ii.) Gilt  $f' < 0$  für alle  $x \in (a, b)$ , so ist  $f$  strikt monoton fallend.
- iii.) Es gilt genau dann  $f'(x) \geq 0$  für alle  $x \in (a, b)$ , wenn  $f$  monoton wachsend ist.
- iv.) Es gilt genau dann  $f'(x) \leq 0$  für alle  $x \in (a, b)$ , wenn  $f$  monoton fallend ist.

BEWEIS: Für den ersten und zweiten Teil betrachten wir  $x_1, x_2 \in [a, b]$  mit  $x_1 < x_2$  und die Einschränkung  $f|_{[x_1, x_2]}$ . Wir dürfen den Mittelwertsatz anwenden und finden ein  $\xi \in (x_1, x_2)$  mit

$$(x_2 - x_1)f'(\xi) = f(x_2) - f(x_1). \quad (*)$$

Damit entscheidet das Vorzeichen von  $f'(\xi)$  über das Vorzeichen von  $f(x_2) - f(x_1)$ , da  $x_2 - x_1 > 0$ . Dies liefert i.) und ii.). Für iii.) und iv.) sehen wir zunächst, dass für eine monoton wachsende Funktion

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

für alle  $x \neq x_0$  gilt. Daher ist auch der Grenzwert für  $x \rightarrow x_0$  noch größer oder gleich Null. Analog erhält man  $f'(x_0) \leq 0$  für eine monoton fallende Funktion. Die Umkehrung erfolgt dann wieder mit dem Mittelwertsatz in Form von (\*).  $\square$

Man beachte, dass in i.) und ii.) im Allgemeinen keine Äquivalenz zu erwarten ist. Die Funktion

$$f(x) = x^3 \quad (5.3.7)$$

ist strikt monoton steigend auf ganz  $\mathbb{R}$ , besitzt aber eine horizontale Tangente bei  $x = 0$ , siehe erneut Abbildung 5.8.

Die letzte Folgerung aus dem Mittelwertsatz ist eine Wachstumsschranke an die Funktionswerte, wenn wir eine Schranke an die Ableitung kennen:

**Korollar 5.3.9** Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig und auf  $(a, b)$  differenzierbar mit beschränkter Ableitung. Dann sind die Differenzenquotienten von  $f$  ebenfalls beschränkt. Sind nämlich  $m, M \in \mathbb{R}$  mit

$$m \leq f'(x) \leq M \quad (5.3.8)$$

Abbildung 5.11: Eine konvexe kompakte Teilmenge  $K \subseteq X$  des offenen Definitionsbereichs  $X$  von  $f$ .

für alle  $x \in (a, b)$  gegeben, so gilt

$$m \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq M \quad (5.3.9)$$

für alle  $x_1, x_2 \in [a, b]$  mit  $x_1 \neq x_2$ .

BEWEIS: Wir zeigen die zweite, spezifischere Aussage. Zu  $x_1, x_2 \in [a, b]$  mit  $x_1 < x_2$  finden wir einen Zwischenwert  $\xi \in (x_1, x_2)$  mit

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Da  $m \leq f'(\xi) \leq M$  nach Voraussetzung (5.3.8) gilt, folgt die Abschätzung (5.3.9) sofort.  $\square$

Eine etwas weniger spezifische Abschätzung erhalten wir, indem wir die Supremumsnorm  $\|f'\|_\infty = \sup\{|f'(x)| \mid x \in (a, b)\}$  der Ableitung verwenden. Dann gilt

$$\left| \frac{f(x_2) - f(x_1)}{x_2 - x_1} \right| \leq \|f'\|_\infty \quad (5.3.10)$$

für alle  $x_1, x_2 \in [a, b]$  mit  $x_1 \neq x_2$ .

Bislang haben wir vornehmlich Funktionen auf reellen Intervallen mit reellen Werten betrachtet, um Gebrauch von Mittelwertsätzen und Zwischenwertsätzen zu machen. Sind die Werte dagegen komplex, gibt es natürlich stetige Möglichkeiten von einem Funktionswert  $f(a)$  zu einem anderen  $f(b)$  zu gelangen, die alle Werte „zwischen“  $f(a)$  und  $f(b)$  vermeiden: Hier kann man leicht Beispiele zu allen bisherigen Szenarien konstruieren, siehe etwa Übung ???. Diese Sätze sind daher intrinsisch reeller Natur. Um nun trotzdem interessante Konsequenzen für differenzierbare Funktionen im Komplexen zu erhalten, begnügen wir uns mit Abschätzungen wie der folgenden:

**Satz 5.3.10 (Schränkensatz)** Sei  $X \subseteq \mathbb{C}$  eine offene Teilmenge und  $f: X \rightarrow \mathbb{C}$  differenzierbar mit lokal beschränkter Ableitung  $f': X \rightarrow \mathbb{C}$ . Dann gilt für alle konvexen kompakten Teilmengen  $K \subseteq X$  und alle  $z_0, z_1 \in K$  die Abschätzung

$$|f(z_1) - f(z_0)| \leq \|f'\|_K |z_1 - z_0|. \quad (5.3.11)$$

BEWEIS: Konvexität von  $K$  bedeutet, dass mit  $z_0, z_1 \in K$  auch die Punkte

$$z(t) = tz_1 + (1-t)z_0$$

für  $t \in [0, 1]$  in  $K$  liegen, siehe Abbildung 5.11. Es gibt eine komplexe Phase  $c \in \mathbb{S}^1$  mit der Eigenschaft

$$c(f(z_1) - f(z_0)) = |f(z_1) - f(z_0)|.$$

Diese ist eindeutig für  $f(z_1) \neq f(z_0)$  und kann beliebig gewählt werden, wenn  $f(z_1) = f(z_0)$ . Wir betrachten dann die Funktion

$$g: [0, 1] \ni t \mapsto \operatorname{Re}(c(f(z(t)))) \in \mathbb{R}.$$

Da die Funktion  $t \mapsto z(t)$  sicherlich differenzierbar ist, ist  $g$  auf dem offenen Intervall  $(0, 1)$  differenzierbar und stetig auf  $[0, 1]$ . Hier benutzen wir, dass  $f$  differenzierbar ist. Für die Ableitung finden wir mit der Kettenregel

$$g'(t) = \operatorname{Re}(cf'(z(t)) \cdot (z_1 - z_0)).$$

Es gilt für alle  $t \in [0, 1]$

$$\begin{aligned} |g'(t)| &\leq |\operatorname{Re}(cf'(z(t)) \cdot (z_1 - z_0))| \\ &\leq |c| |f'(z(t))| |z_1 - z_0| \\ &= |f'(z(t))| |z_1 - z_0| \\ &\leq \|f'\|_K |z_1 - z_0|, \end{aligned}$$

da alle Punkte  $z(t) \in K$  in diesem Kompaktum liegen und  $f'$  lokal beschränkt ist. Nach (5.3.10) folgt also

$$\left| \frac{g(1) - g(0)}{1 - 0} \right| \leq \|g'\|_\infty \leq \|f'\|_K |z_1 - z_0|.$$

da  $g(1) = \operatorname{Re}(cf(z_1))$  und  $g(0) = \operatorname{Re}(cf(z_0))$ , folgt

$$g(1) - g(0) = \operatorname{Re}(c(f(z_1) - f(z_0))) = \operatorname{Re}(|f(z_1) - f(z_0)|) = |f(z_1) - f(z_0)|$$

nach unserer Wahl der Phase  $c$ . Dies zeigt (5.3.11).  $\square$

Zu einem späteren Zeitpunkt wird man in der Funktionentheorie sehen, dass eine komplex differenzierbare Funktion auf  $X$  automatisch eine stetige Ableitung  $f'$  besitzt, ja sogar unendlich oft stetig differenzierbar ist. Damit ist die Voraussetzung, dass  $f'$  in Satz 5.3.10 lokal beschränkt sein soll, automatisch erfüllt.

Ist die Ableitung  $f'$  einer Funktion  $f$  selbst wieder eine stetige Funktion, so können wir Resultate zu stetigen Funktionen wie beispielsweise den Zwischenwertsatz auf  $f'$  anwenden. Überraschenderweise ist die Stetigkeit von  $f'$  dafür aber nicht notwendig. Allein die Tatsache, dass  $f'$  eine Ableitung ist, liefert bereits folgendes Resultat.

**Satz 5.3.11 (Zwischenwertsatz für Ableitungen)** Sei  $[a, b] \subseteq \mathbb{R}$  ein kompaktes Intervall mit  $a < b$ , und sei  $f: [a, b] \rightarrow \mathbb{R}$  eine differenzierbare Funktion mit

$$f'(a) \neq f'(b). \quad (5.3.12)$$

Dann nimmt  $f'$  jeden Wert zwischen  $f'(a)$  und  $f'(b)$  im offenen Intervall  $(a, b)$  an.

BEWEIS: Ohne Einschränkung können wir  $f'(a) < f'(b)$  annehmen, der andere Fall kann analog bewiesen werden. Zu einem Zwischenwert  $y_0 \in (f'(a), f'(b))$  betrachten wir die Hilfsfunktion

$$\varphi(x) = f(x) - y_0 x$$

auf  $[a, b]$ . Da  $f$  differenzierbar ist, ist  $\varphi$  ebenfalls differenzierbar mit

$$\varphi'(x) = f'(x) - y_0.$$

An den Grenzen erhalten wir

$$\varphi'(a) = f'(a) - y_0 < 0 \quad (*)$$

sowie

$$\varphi'(b) = f'(b) - y_0 > 0. \quad (**)$$

Damit gibt es aber  $\epsilon_1, \epsilon_2 > 0$  derart, dass

$$\varphi(x) < \varphi(a) \quad \text{für alle } x \in (a, a + \epsilon_1) \subseteq [a, b] \quad (\odot)$$

und

$$\varphi(x) < \varphi(b) \quad \text{für alle } x \in (b - \epsilon_2, b) \subseteq [a, b] \quad (\odot)$$

da sonst die entsprechenden Differenzenquotienten nicht gegen die Ableitungen (\*) beziehungsweise (\*\*) konvergieren könnten. Als differenzierbare Funktion ist  $\varphi$  insbesondere stetig und besitzt daher auf dem kompakten Intervall  $[a, b]$  ein Minimum, siehe Korollar 3.5.8. Dieses kann wegen (⊙) und (⊙) aber nur im offenen Intervall  $(a, b)$  und nicht an den Randpunkten  $a$  oder  $b$  angenommen werden. Nach Proposition 5.3.4 gilt für dieses Extremum  $x_0 \in (a, b)$  aber  $\varphi'(x_0) = 0$ , was  $f'(x_0) = y_0$  zeigt.  $\square$

Als letztes wichtiges Resultat zu differenzierbaren Funktionen betrachten wir eine Anwendung: die Regel von l'Hospital. Diese wird es ermöglichen, Grenzwerte von Quotienten der Form  $\frac{f(x)}{g(x)}$  für  $x \rightarrow x_0$  zu berechnen, wenn beide Funktionen  $f(x_0) = 0 = g(x_0)$  dort eine Nullstelle besitzen, der Quotient also zunächst unbestimmt ist. Hierfür werden die Ableitungen von  $f$  und  $g$  herangezogen.

Als Vorüberlegung betrachten wir folgenden Spezialfall:

**Proposition 5.3.12** Seien  $f, g: X \rightarrow \mathbb{K}$  Funktionen, die bei  $x_0 \in X$  differenzierbar sind und

$$f(x_0) = 0 = g(x_0) \quad \text{sowie} \quad g'(x_0) \neq 0 \quad (5.3.13)$$

erfüllen. Dann gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}. \quad (5.3.14)$$

BEWEIS: Da  $f$  und  $g$  bei  $x_0$  differenzierbar sind, sind sie dort erst recht stetig. Es gilt also

$$\lim_{x \rightarrow x_0} f(x) = 0 = \lim_{x \rightarrow x_0} g(x).$$

Da zudem  $g'(x_0) \neq 0$  gelten soll, finden wir eine Umgebung von  $x_0$ , für die  $g(x) \neq 0$  gilt, sobald  $x \in U \setminus \{x_0\}$ . Angenommen, dies wäre nicht der Fall, dann gäbe es zu jeder Umgebung der Form  $X \cap B_{\frac{1}{n}}(x_0)$  mindestens ein  $x_n \in B_{\frac{1}{n}}(x_0) \cap X$  mit  $g(x_n) = 0$ . Für die Ableitung bei  $x_0$  folgt dann aber

$$g'(x_0) = \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = \lim_{n \rightarrow \infty} \frac{g(x_n) - 0}{x_n - x_0} = 0,$$

da bei existierender Ableitung dieser Grenzwert auf einer beliebigen Folge  $x_n \rightarrow x_0$  ausgewertet werden kann. Damit haben wir einen Widerspruch erreicht: Es gibt also doch eine Umgebung  $U \subseteq X$  von  $x_0$  in  $X$  mit  $g(x) \neq 0$  für alle  $x \in U \setminus \{x_0\}$ . Damit ist  $\frac{f}{g}$  zumindest auf  $U \setminus \{x_0\}$  definiert, und es ist sinnvoll, nach dem Grenzwert für  $x \rightarrow x_0$  zu fragen. Für  $x \in U \setminus \{x_0\}$  gilt

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

Da die Grenzwerte der Differenzenquotienten nach Voraussetzung existieren und  $g'(x_0) \neq 0$  gilt, folgt also

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

Man beachte, dass die Stetigkeit von  $f$  und  $g$  bei  $x_0$  den Grenzwert auf der linken Seite diffizil werden lässt. Es gilt eben  $\lim_{x \rightarrow x_0} f(x) = 0 = \lim_{x \rightarrow x_0} g(x)$ , womit eine einfache Auswertung der linken Seite von (5.3.14) nicht möglich ist.

**Beispiel 5.3.13 (Die sinc-Funktion)** Wir betrachten die Funktion  $f(x) = \sin(x)$  sowie  $g(x) = x$ . Beide haben bei  $x_0 = 0$  eine Nullstelle und sind dort differenzierbar. Es gilt  $g'(0) = 1$ , womit wir die Proposition anwenden dürfen und

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \frac{\sin'(0)}{1} = \cos(0) = 1 \quad (5.3.15)$$



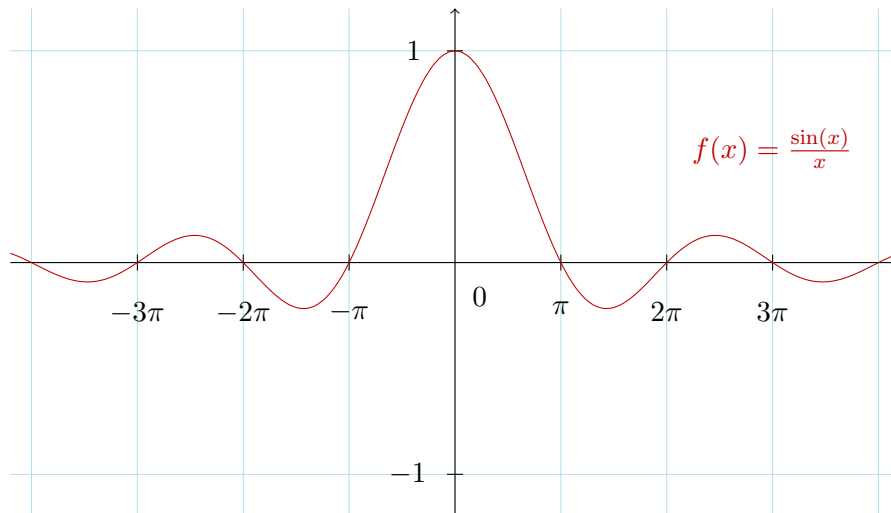


Abbildung 5.12:

erhalten. Insbesondere wird durch die Definition

$$\operatorname{sinc}(x) = \begin{cases} \frac{\sin(x)}{x} & \text{für } x \in \mathbb{C} \setminus \{0\} \\ 1 & \text{für } x = 0 \end{cases} \quad (5.3.16)$$

eine auf ganz  $\mathbb{C}$  stetige Funktion

$$\operatorname{sinc}: \mathbb{C} \longrightarrow \mathbb{C} \quad (5.3.17)$$

definiert, der *Sinus cardinalis*. Diese Funktion spielt in vielen Bereichen der Mathematik eine wichtige Rolle und hat Anwendungen in der Informationstheorie und Signalverarbeitung. Einen alternativen Beweis für (5.3.15) erhalten wir direkter aus der Potenzreihendarstellung sowie der Stetigkeit konvergenter Potenzreihen. Es gilt für  $x \neq 0$

$$\frac{\sin(x)}{x} = \frac{1}{x} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n+1)!}. \quad (5.3.18)$$

Für die rechte Seite ist der Grenzübergang  $x \rightarrow 0$  nun trivial, da die Potenzreihen eben stetig sind. Darüber hinaus erhalten wir für  $\operatorname{sinc}(x)$  die Potenzreihendarstellung

$$\operatorname{sinc}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n+1)!}, \quad (5.3.19)$$

welche für alle  $x \in \mathbb{C}$  konvergiert.

Nach diesen einfachen Beispiel als Motivation können wir die Situation von Proposition 5.3.12 drastisch verallgemeinern: Wir müssen nicht länger annehmen, dass  $x_0$  im Definitionsbereich der Funktionen liegt. Es kann sich vielmehr um einen eventuell sogar uneigentlichen Häufungspunkt handeln. Allerdings benötigen wir dann etwas strengere Differenzierbarkeitsvoraussetzungen und reelle Definitionsbereiche und Wertebereiche.

**Satz 5.3.14 (Regel von l'Hospital)** Sei  $I = (a, b) \subseteq \mathbb{R}$  ein offenes Intervall und  $f, g: I \rightarrow \mathbb{R}$  zwei auf ganz  $I$  differenzierbare reellwertige Funktionen mit

$$g'(x) \neq 0 \quad (5.3.20)$$

für alle  $x \in I$ . Weiter gelte eine der beiden folgenden Annahmen

i.)

$$\lim_{x \searrow a} f(x) = 0 = \lim_{x \searrow a} g(x), \quad (5.3.21)$$

ii.)

$$\lim_{x \searrow a} g(x) = +\infty \quad \text{oder} \quad \lim_{x \searrow a} g(x) = -\infty. \quad (5.3.22)$$

In diesem Fall gilt

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}, \quad (5.3.23)$$

sofern der Grenzwert der Ableitungen in  $\mathbb{R} \cup \{\pm\infty\}$  existiert. Eine entsprechende Aussage gilt für die Grenzwerte  $x \nearrow b$ .

BEWEIS: Wir bezeichnen den Grenzwert der rechten Seite in (5.3.23) mit

$$c = \lim_{x \searrow a} \frac{f'(x)}{g'(x)} \in [-\infty, +\infty],$$

wobei für  $c$  explizit auch  $\pm\infty$  erlaubt sein soll. Sei zunächst  $c \neq +\infty$ . Dann gibt es also zu jedem  $y_0 \in \mathbb{R}$  mit  $c < y_0$  ein  $x_0 \in (a, b)$  derart, dass

$$\frac{f'(x)}{g'(x)} < y_0 \quad \text{für} \quad x \in (a, x_0).$$

Dies folgt aus der Existenz des Grenzwertes  $c$  und schließt den Fall  $c = -\infty$  explizit mit ein. Nach dem Mittelwertsatz für differenzierbare Funktionen finden wir zu  $x_1, x_2 \in (a, x_0)$  mit  $x_2 < x_1$  nach Satz 5.3.6 ein  $\xi \in (x_2, x_1)$ , so dass

$$(f(x_1) - f(x_2))g'(\xi) = (g(x_1) - g(x_2))f'(\xi).$$

Da nach Voraussetzung  $g'$  überall von Null verschieden ist, ist die Funktion  $g$  auf dem zusammenhängenden Intervall  $(a, b)$  streng monoton und daher insbesondere injektiv, siehe Korollar 5.3.8. Daher gilt  $g(x_1) \neq g(x_2)$  und es folgt

$$\frac{f(x_1) - f(x_2)}{g(x_1) - g(x_2)} = \frac{f'(\xi)}{g'(\xi)} < y_0 \quad (*)$$

für das obige  $\xi \in (x_2, x_1)$  passend zu  $a < x_2 < x_1 < x_0$ . Wir bringen nun die Annahme i.) ins Spiel. Gilt (5.3.21), so können wir in (\*) den Grenzwert  $x_2 \searrow a$  bilden und erhalten im Grenzwert zumindest noch die Abschätzung

$$\frac{f(x_1)}{g(x_1)} \leq y_0$$

für alle  $x_1 < x_0$ . Gilt andererseits die Annahme ii.), so finden wir zunächst ein  $x_1 < x_0$  mit der Eigenschaft, dass

$$g(x) > \max\{g(x_0), 0\}$$

im Fall  $\lim_{x \searrow a} g(x) = +\infty$  beziehungsweise

$$g(x) < \max\{g(x_0), 0\}$$

im anderen Fall  $\lim_{x \searrow a} g(x) = -\infty$ , jeweils gültig für alle  $x \in (a, x_1)$ . Wir können für dieses  $x_1$  nach wie vor (\*) verwenden und erhalten daher für alle  $x \in (a, x_1)$

$$\frac{f(x) - f(x_1)}{g(x) - g(x_1)} = \frac{f(x) - f(x_1)}{g(x)} \underbrace{\frac{g(x)}{g(x) - g(x_1)}}_{>0} < y_0,$$

wobei in beiden Fällen  $\frac{g(x)}{g(x)-g(x_1)} > 0$  gilt. Daher folgt

$$\frac{f(x) - f(x_1)}{g(x)} < y_0 \frac{g(x) - g(x_1)}{g(x)}$$

oder eben

$$\frac{f(x)}{g(x)} < y_0 - y_0 \frac{g(x_1)}{g(x)} + \frac{f(x_1)}{g(x)}.$$

Bildet man von dieser Ungleichung den Grenzwert  $x \searrow a$ , so gilt

$$\lim_{x \searrow a} \left( y_0 - y_0 \frac{g(x_1)}{g(x)} + \frac{f(x_1)}{g(x)} \right) = y_0,$$

da nach Voraussetzung ii.) ja  $\lim_{x \searrow a} g(x) = \pm\infty$  gelten soll. Es gibt also auch in diesem Fall ein  $x_2$  mit

$$\frac{f(x)}{g(x)} \leq y_0,$$

sofern  $x \in (a, x_2)$ . Insgesamt erhalten wir unter beiden Annahmen i.) oder ii.) also zu jedem  $y_0 \in \mathbb{R}$  mit  $c < y_0$  ein  $x_0 \in (a, b)$  derart, dass

$$\frac{f(x)}{g(x)} \leq y_0 \quad (\star)$$

für  $x \in (a, x_0)$  gilt. Eine analoge Überlegung zeigt dann, dass es im Fall  $c \in (-\infty, +\infty]$  zu jedem  $z_0 \in \mathbb{R}$  mit  $z_0 < c$  ein  $x_0 \in (a, b)$  mit

$$\frac{f(x)}{g(x)} \geq z_0 \quad (\star\star)$$

für  $x \in (a, x_0)$  gibt. Ist nun  $c = -\infty$ , so zeigt bereits  $(\star)$  alleine, dass  $\lim_{x \searrow a} \frac{f(x)}{g(x)} = -\infty$  gilt. Entsprechend folgt für  $c = +\infty$  aus  $(\star\star)$ , dass  $\lim_{x \searrow a} \frac{f(x)}{g(x)} = +\infty$  gilt. Ist schließlich  $c \in \mathbb{R}$ , so zeigen die beiden Abschätzungen  $(\star)$  und  $(\star\star)$  zusammen, dass nach dem Sandwichprinzip  $\lim_{x \searrow a} \frac{f(x)}{g(x)} = c$  gilt. Damit ist der Beweis erbracht.  $\square$

Der Satz 5.3.14 bietet gegenüber Proposition 5.3.12 nun eine interessante Verallgemeinerung für die Fälle, in denen der Quotient  $\frac{f'(x)}{g'(x)}$  einen Grenzwert besitzt, ohne dass die Ableitungen  $f'(x)$  und  $g'(x)$  einzeln einen Grenzwert für  $x \searrow a$  haben. Trotzdem ist oftmals die viel einfachere Proposition 5.3.12 völlig ausreichend, um Grenzwerte zu bestimmen.

## 5.4 Höhere Ableitungen und $\mathcal{C}^k$ -Funktionen

Ist  $X \subseteq \mathbb{K}$  offen, so können wir für die auf  $X$  differenzierbare Funktion  $f$  an jedem Punkt  $x_0 \in X$  die Ableitung  $f'(x_0)$  bestimmen. Dies definiert eine neue Funktion  $f'$  auf  $X$ , deren Stetigkeit und Differenzierbarkeit wir erneut betrachten können. Induktiv führt dies auf folgende Definition:

**Definition 5.4.1 ( $\mathcal{C}^k$ -Funktionen)** Sei  $X$  eine offene Teilmenge der reellen oder komplexen Zahlen. Sei weiter  $k \geq 1$ .

- i.) Eine Funktion  $f: X \rightarrow \mathbb{K}$  heißt  $k$ -mal stetig differenzierbar, wenn  $f$  differenzierbar ist und  $f': X \rightarrow \mathbb{K}$  eine  $(k-1)$ -mal stetig differenzierbare Funktion ist. Eine 0-mal stetig differenzierbare Funktion ist dabei einfach eine stetige Funktion.
- ii.) Die Menge der  $k$ -mal stetig differenzierbaren Funktionen wird mit

$$\mathcal{C}^k(X, \mathbb{K}) = \{f: X \rightarrow \mathbb{K} \mid f \text{ ist } k\text{-mal stetig differenzierbar}\} \quad (5.4.1)$$

bezeichnet.

iii.) Ist  $f$  eine  $k$ -mal stetig differenzierbare Funktion für alle  $k \in \mathbb{N}$ , so heißt  $f$  *glatt*. Wir setzen

$$\mathcal{C}^\infty(X, \mathbb{K}) = \{f: X \longrightarrow \mathbb{K} \mid f \text{ ist glatt}\}. \quad (5.4.2)$$

Wir wollen also nicht nur die Differenzierbarkeit an einem Punkt sondern auf dem gesamten offenen Definitionsbereich  $X$ . Die Forderung nach stetiger Differenzierbarkeit ist in sofern sinnvoll, als dass eine differenzierbare Ableitungsfunktion  $f': X \longrightarrow \mathbb{K}$  auf einem offenen Definitionsbereich notwendigerweise selbst wieder stetig sein muss. Dies haben wir in Korollar 5.1.5 für eine Funktion  $f$  gesehen und können es entsprechend auf die Funktion  $f'$  anwenden. Wenn  $f$  also zweimal differenzierbar sein soll, so ist  $f'$  entsprechend einmal differenzierbar und damit bei jedem Punkt stetig. Nach Definition gilt

$$\mathcal{C}^\infty(X, \mathbb{K}) \subseteq \mathcal{C}^{k+1}(X, \mathbb{K}) \subseteq \mathcal{C}^k(X, \mathbb{K}) \subseteq \mathcal{C}^0(X, \mathbb{K}) \quad (5.4.3)$$

für alle  $k \in \mathbb{N}$ , wobei  $\mathcal{C}^0(X, \mathbb{K}) = \mathcal{C}(X, \mathbb{K})$  gesetzt wird. Als Bezeichnung für die höheren Ableitungen verwendet man

$$f' = \frac{df}{dx} \quad (5.4.4)$$

und

$$f'' = (f')' = \frac{d^2 f}{dx^2} \quad (5.4.5)$$

sowie

$$f^{(k)} = (f^{(k-1)})' = \frac{d^k f}{dx^k} \quad (5.4.6)$$

für  $k \geq 2$ , immer vorausgesetzt, dass die Ableitungen existieren.

### Beispiel 5.4.2 (Glatte Funktionen)

i.) Polynomiale Funktionen

$$f(x) = a_N x^N + \cdots + a_1 x + a_0 \quad (5.4.7)$$

mit  $a_N, \dots, a_1, a_0 \in \mathbb{K}$  sind unendlich oft stetig differenzierbare Funktionen: Nach Korollar 5.2.2 wissen wir nicht nur, dass  $f$  differenzierbar ist. Vielmehr ist die Ableitung von  $f'$  selbst wieder ein Polynom, womit induktiv folgt, dass  $f \in \mathcal{C}^\infty(\mathbb{K}, \mathbb{K})$  gilt, siehe auch Übung 5.8.2.

ii.) Eine rationale Funktion

$$f(x) = \frac{p(x)}{q(x)} \quad (5.4.8)$$

mit Polynomen  $p, q \in \mathbb{K}[x]$  ist unendlich oft stetig differenzierbar auf ihrem natürlichen Definitionsbereich  $X$ . In der Tat ist nach der Quotientenregel aus Korollar 5.2.3 die Ableitung  $f'$  nicht nur auf  $X$  vorhanden, sondern selbst wieder eine rationale Funktion mit demselben natürlichen Definitionsbereich. Mit Induktion folgt daher  $f \in \mathcal{C}^\infty(X, \mathbb{K})$ .

iii.) Die Exponentialfunktion  $\exp$  ebenso wie die Sinus- und Kosinusfunktion  $\sin$  und  $\cos$  sind auf ganz  $\mathbb{C}$  glatt, da ihre Ableitungen nach Proposition 5.2.4 wieder von dieser Form sind. Es gilt also

$$\exp, \cos, \sin \in \mathcal{C}^\infty(\mathbb{C}, \mathbb{C}). \quad (5.4.9)$$

iv.) Mit ähnlichen Argumenten sieht man, dass auch  $\tan$ ,  $\cot$ ,  $\log$  und die Wurfelfunktionen auf ihren jeweiligen (offenen) Definitionsbereichen glatt sind, siehe auch die Übungen 5.8.3, 5.8.4, 5.8.5, 5.8.6 und 5.8.7.

Wie zu erwarten, kombinieren sich die Rechenregeln von stetigen Funktionen und differenzierbaren Funktionen zu den Rechenregeln für  $\mathcal{C}^\infty$ -Funktionen:

**Proposition 5.4.3** Sei  $X$  eine offene Teilmenge der reellen oder komplexen Zahlen und  $k \in \mathbb{N} \cup \{\infty\}$ .

- i.) Die  $\mathcal{C}^k$ -Funktionen  $\mathcal{C}^k(X, \mathbb{K})$  auf  $X$  bilden eine Unteralgebra aller Funktionen: Für  $f, g \in \mathcal{C}^k(X, \mathbb{K})$  und  $z, w \in \mathbb{K}$  sind  $zf + wg, fg \in \mathcal{C}^k(X, \mathbb{K})$  wieder  $\mathcal{C}^k$ -Funktionen.
- ii.) Ist  $f$  eine  $\mathcal{C}^k$ -Funktion mit  $f(x) \neq 0$  für alle  $x \in X$ , so ist  $\frac{1}{f} \in \mathcal{C}^k(X, \mathbb{K})$ .
- iii.) Ist  $Y$  eine weitere offene Teilmenge der reellen oder komplexen Zahlen und  $f \in \mathcal{C}^k(X, \mathbb{K})$  und  $g \in \mathcal{C}^k(Y, \mathbb{K})$  mit  $f(X) \subseteq Y$ , so ist auch die Verkettung  $g \circ f$  wieder eine  $\mathcal{C}^k$ -Funktion.
- iv.) Ist  $Y$  ebenfalls eine offene Teilmenge der reellen oder komplexen Zahlen und  $f: X \rightarrow Y$  eine bijektive  $\mathcal{C}^k$ -Funktion mit  $f'(x) \neq 0$  für alle  $x \in X$  und stetiger Umkehrfunktion  $f^{-1}: Y \rightarrow X$  so ist  $f^{-1}$  ebenfalls eine  $\mathcal{C}^k$ -Funktion.

BEWEIS: Für den ersten Teil betrachten wir  $f, g \in \mathcal{C}^k(X, \mathbb{K})$  sowie  $z, w \in \mathbb{K}$ . Nach Proposition 5.2.1, i.), ist  $zf + wg$  auf  $X$  differenzierbar mit Ableitung

$$(zf + wg)' = zf' + wg'.$$

Da  $f'$  und  $g'$  stetig sind, ist nach Proposition 3.3.7 auch  $zf' + wg'$  stetig. Dies zeigt die Stetigkeit der Ableitung von  $zf + wg$ , womit  $zf + wg \in \mathcal{C}^1(X, \mathbb{K})$  folgt. Eine einfache Induktion mittels wiederholtem Anwenden von Proposition 5.2.1, i.), und Proposition 3.3.7 zeigt dann  $zf + wg \in \mathcal{C}^k(X, \mathbb{K})$ . Für das Produkt argumentieren wir analog. Zuerst sehen wir

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

nach Proposition 5.2.1, ii.). Da  $f'$  und  $g'$  ebenso wie  $f$  und  $g$  stetig sind, ist  $f'g + fg'$  nach Proposition 3.3.7 wieder stetig. Also ist  $(fg)'$  stetig, was  $fg \in \mathcal{C}^1(X, \mathbb{K})$  zeigt. Induktiv folgt  $fg \in \mathcal{C}^k(X, \mathbb{K})$ , siehe auch Übung 5.8.1 zur Leibniz-Regel für die höheren Ableitungen. Die konstanten Funktionen sind offenbar glatt. Für den zweiten Teil wissen wir zunächst nach Proposition 5.2.1, iii.), dass  $\frac{1}{f}$  wieder differenzierbar ist mit der Ableitung

$$\left(\frac{1}{f}\right)' = -\frac{f'}{f^2}. \quad (*)$$

Nach Proposition 3.3.7 sowie (3.3.15) ist  $-\frac{f'}{f^2}$  wieder stetig, was  $\frac{1}{f} \in \mathcal{C}^1(X, \mathbb{K})$  zeigt. Ist nun  $k \geq 2$ , so wissen wir  $f' \in \mathcal{C}^1(X, \mathbb{K})$  sowie  $\frac{1}{f^2} \in \mathcal{C}^1(X, \mathbb{K})$  nach dem gerade Gezeigten. Also ist  $(\frac{1}{f})'$  erneut  $\mathcal{C}^1(X, \mathbb{K})$ , und wir können induktiv fortfahren, da die höheren Ableitungen mittels der Quotientenregel (\*) wieder auf Ausdrücke in niedrigeren Ableitungen von  $f$  sowie Potenzen von  $\frac{1}{f}$  führen, auf die wir die Induktionsvoraussetzungen anwenden können. Es folgt  $\frac{1}{f} \in \mathcal{C}^k(X, \mathbb{K})$ . Für den dritten Teil benutzen wir die Kettenregel

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

aus Proposition 5.2.5. Nach Satz 3.5.3 ist  $g' \circ f$  als Verkettung der stetigen Funktionen  $g'$  und  $f$  wieder stetig. Damit wird auch das Produkt mit der stetigen Funktion  $f'$  nach Proposition 3.3.7 wieder stetig, was  $g \cdot f' \in \mathcal{C}^1(X, \mathbb{K})$  zeigt. Auch hier können wir induktiv fortfahren, um  $g \circ f \in \mathcal{C}^k(X, \mathbb{K})$  zu schließen, siehe auch Übung 5.8.8. Für den letzten Teil wissen wir nach Proposition 5.2.6, dass die Umkehrabbildung  $f^{-1}$  differenzierbar ist mit Ableitung

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

Da  $f'$  stetig ist, ist die Verkettung mit der stetigen Funktion  $f^{-1}$  ebenfalls stetig. Nach Voraussetzung gilt überall  $f' \neq 0$ , womit auch der Kehrwert  $\frac{1}{f' \circ f^{-1}}$  stetig ist. Es folgt also  $f^{-1} \in \mathcal{C}^1(Y, \mathbb{K})$ . Da wir sowohl Verkettungen als auch Kehrwerte wieder differenzieren können, erhalten wir induktiv auch in diesem letzten Fall  $f^{-1} \in \mathcal{C}^k(Y, \mathbb{K})$ .  $\square$

Abbildung 5.13: Lokale Extrema von  $f$ , Nullstellen von  $f'$  und Nullstellen von  $f''$ . Bei  $x_3$  liegt ein Wendepunkt vor.

Die höheren Ableitungen einer Verkettung oder der Umkehrfunktion explizit auszurechnen ist nicht einfach. Hier muss man wiederholt die Rechenregeln zum Differenzieren in geeigneter Kombination verwenden. Einen ersten Eindruck gibt Übung 5.8.8.

Höhere Ableitungen gestatten es nun, vorherige Resultate zu Extremwerten besser zu verstehen. Hier können wir Proposition 5.3.4 um ein hinreichendes Kriterium erweitern:

**Proposition 5.4.4** Sei  $I = (a, b) \subseteq \mathbb{R}$  ein offenes Intervall und  $f: I \rightarrow \mathbb{R}$  eine zweimal differenzierbare Funktion auf  $I$ . Ist  $x_0 \in (a, b)$  mit  $f'(x_0) = 0$ , so ist  $x_0$

i.) ein striktes lokales Maximum, wenn  $f''(x_0) < 0$  gilt,

oder

ii.) ein striktes lokales Minimum, wenn  $f''(x_0) > 0$  gilt.

BEWEIS: Sei  $f''(x_0) < 0$ . Dann gilt

$$f''(x_0) = \lim_{x \rightarrow x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f'(x)}{x - x_0} < 0.$$

Dies bedeutet aber insbesondere, dass es ein  $\delta > 0$  mit

$$\frac{f'(x)}{x - x_0} < 0$$

für alle  $x \in (a, b) \setminus \{x_0\}$  mit  $|x - x_0| < \delta$  gibt. Ist nun  $x \in (x_0 - \delta, x_0)$ , so ist  $x - x_0 < 0$ , also folgt  $f'(x) > 0$ . Ist  $x \in (x_0, x_0 + \delta)$ , so gilt entsprechend  $f'(x) < 0$ . Daher ist  $f$  auf  $(x_0 - \delta, x_0)$  streng monoton wachsend nach Korollar 5.3.8, i.), und auf  $(x_0, x_0 + \delta)$  streng monoton fallend. Damit ist  $x_0$  aber ein striktes lokales Maximum. Der zweite Fall folgt durch Vertauschen von  $f$  mit  $-f$ .  $\square$

**Bemerkung 5.4.5 (Kurvendiskussion)** Ein lokales Extremum der Ableitung  $f'$  einer differenzierbaren Funktion nennt man auch einen *Wendepunkt*. Anschaulich bedeutet dies, dass der Graph von  $f$  von einer Links- in eine Rechtskurve dreht, wenn das Extremum ein Maximum ist, und von einer Rechtskurve in eine Linkskurve, falls ein Minimum von  $f'$  vorliegt, siehe auch Abbildung 5.13. Die Bestimmung der Nullstellen von  $f$ ,  $f'$  und  $f''$  erlaubt es dann, eine recht aussagekräftige Skizze des Graphen von  $f$  zu erstellen. Zusammen mit dem Verhalten von  $f(x)$  für  $x \searrow a$  beziehungsweise  $x \nearrow b$  erhält man dann eine gute Vorstellung vom Verlauf von  $f$ .

**Beispiel 5.4.6** Die Bedingungen in Proposition 5.4.4 für Extrema sind hinreichend, aber nicht notwendig. Dazu betrachtet man

$$f(x) = x^n \tag{5.4.10}$$

für  $n \in \mathbb{N}$ . Mit

$$f'(x) = nx^{n-1} \quad \text{und} \quad f''(x) = n(n-1)x^{n-2} \tag{5.4.11}$$

sieht man, dass der einzige Kandidat für ein Extremum bei  $x_0 = 0$  liegt. Nun ist dort aber für  $n \geq 2$  auch  $f''(x_0) = 0$ , womit Proposition 5.4.4 nicht zu Anwendung kommt. Tatsächlich gilt, dass  $x_0 = 0$  für gerades  $n$  das globale Minimum ist, da ja  $f(x) \geq 0$  für alle  $x \in \mathbb{R}$  gilt. Für ungerades  $n$  ist  $x_0 = 0$  dagegen ein Wendepunkt, da  $f(x) < 0$  für  $x < 0$  und  $f(x) > 0$  für  $x > 0$ . Dieses Beispiel lässt sich auch in einem allgemeinen Kontext verstehen, siehe Übung 5.8.9.

Abbildung 5.14: Eine konvexe Funktion  $f$ .

**Definition 5.4.7 (Konvexe Funktion)** Sei  $f: (a, b) \rightarrow \mathbb{R}$ . Dann heißt  $f$  konvex, falls für alle  $x_1, x_2 \in (a, b)$  und alle  $\lambda \in (0, 1)$

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad (5.4.12)$$

gilt. Die Funktion  $f$  heißt konkav, falls  $-f$  konvex ist.

Geometrisch bedeutet Konvexität der Funktion  $f$ , dass die Verbindungsstrecke der Punkte  $(x_1, f(x_1))$  und  $(x_2, f(x_2))$  immer oberhalb des Graphen von  $f$  zwischen  $x_1$  und  $x_2$  verläuft, siehe Abbildung 5.14. Bei einer konkaven Funktion verläuft der Graph entsprechend immer oberhalb jeder solchen Sekanten. Mit Induktion sieht man, dass (5.4.12) auch

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i) \quad (5.4.13)$$

impliziert, sofern  $\lambda_1, \dots, \lambda_n \geq 0$  mit  $\lambda_1 + \dots + \lambda_n = 1$  und  $x_1, \dots, x_n \in (a, b)$ , siehe auch Übung 5.8.10.

Im Allgemeinen ist es schwierig zu entscheiden, ob eine Funktion konvex ist oder nicht. Ist  $f$  zudem ein- oder gar zweimal differenzierbar, so vereinfacht sich die Situation erheblich:

**Satz 5.4.8 (Konvexe Funktionen)** Sei  $I = (a, b) \subseteq \mathbb{R}$  ein offenes Intervall, und sei  $f: I \rightarrow \mathbb{R}$  eine Funktion.

i.) Die Funktion  $f$  ist genau dann konvex, wenn zu allen  $x_1, x_2, x \in I$  mit  $x_1 < x < x_2$

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x} \quad (5.4.14)$$

gilt.

ii.) Ist  $f$  auf  $I$  differenzierbar, so ist  $f$  genau dann konvex, wenn  $f'$  monoton wachsend ist.

iii.) Ist  $f$  auf  $I$  zweimal differenzierbar, so ist  $f$  genau dann konvex, wenn

$$f''(x) \geq 0 \quad (5.4.15)$$

für alle  $x \in I$  gilt.

BEWEIS: Sei zunächst  $f$  konvex und  $x_1 < x < x_2$ . Dann gibt es ein eindeutiges  $\lambda \in (0, 1)$  mit

$$x = \lambda x_1 + (1 - \lambda)x_2 \quad (*)$$

nämlich  $\lambda = \frac{x_2 - x}{x_2 - x_1}$ . Für dieses  $\lambda$  erhalten wir aus der Konvexität von  $f$  die Ungleichung

$$\begin{aligned} f(x) &\leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x_2 - x_1 - x_2 + x}{x_2 - x_1} f(x_2) \\ &= \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2). \end{aligned}$$

Umstellen impliziert

$$(x_2 - x_1)f(x) - (x_2 - x_1)f(x_1) \leq (x_1 - x)f(x_1) + (x - x_1)f(x_2).$$

Da wir durch  $(x_2 - x_1) > 0$  und  $(x - x_1) > 0$  teilen können, liefert dies (5.4.14). Die Umkehrung erhält man, indem man für ein gegebenes  $\lambda \in (0, 1)$  den Punkt  $x$  wie in  $(*)$  definiert und  $x \in (x_1, x_2)$

erhält. Dann kann man die Rechenschritte in der umgekehrten Reihenfolge durchlaufen und erhält aus (5.4.14) die Konvexität. Ist nun zudem  $f$  differenzierbar, so folgt aus (5.4.14) für  $x \searrow x_1$

$$\lim_{x \searrow x_1} \frac{f(x) - f(x_1)}{x - x_1} = f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

und für  $x \nearrow x_2$  entsprechend

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \lim_{x \nearrow x_2} \frac{f(x_2) - f(x)}{x_2 - x} = f'(x_2),$$

was zusammen  $f'(x_1) \leq f'(x_2)$  bedeutet. Sei umgekehrt  $f'$  monoton wachsend. Nach dem Mittelwertsatz aus Satz 5.3.6, *ii.*), gibt es dann zu  $x_1 < x < x_2$  Zwischenwerte  $\xi_1 \in (x_1, x)$  und  $\xi_2 \in (x, x_2)$  mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x},$$

womit aus der Monotonie von  $f'$  die Eigenschaft (5.4.14) folgt. Dies zeigt *ii.*). Ist  $f$  sogar zweimal differenzierbar, so ist  $f'$  genau dann monoton wachsend, wenn  $f''(x) \geq 0$  für alle  $x \in I$  gilt, siehe Korollar 5.3.8, *iii.*).  $\square$

Insbesondere der dritte Teil liefert nun ein einfaches Kriterium für Konvexität. Als Anwendung erhalten wir einige der wichtigsten Ungleichungen der Analysis aus der Konkavität der Logarithmusfunktion: Für  $\log: \mathbb{R}^+ \rightarrow \mathbb{R}$  gilt

$$\log'(x) = \frac{1}{x} \quad \text{und} \quad \log''(x) = -\frac{1}{x^2} < 0, \quad (5.4.16)$$

womit nach Satz 5.4.8 die Funktion  $-\log$  konvex ist. Dies liefert folgende Ungleichung:

**Korollar 5.4.9** *Seien  $\lambda_1, \dots, \lambda_n > 0$  mit  $\lambda_1 + \dots + \lambda_n = 1$ . Dann gilt für alle  $a_1, \dots, a_n \in \mathbb{R}^+$  die Ungleichung*

$$a_1^{\lambda_1} \cdots a_n^{\lambda_n} \leq \lambda_1 a_1 + \dots + \lambda_n a_n. \quad (5.4.17)$$

BEWEIS: Wir wissen

$$\lambda_1 \log(a_1) + \dots + \lambda_n \log(a_n) \leq \log(\lambda_1 a_1 + \dots + \lambda_n a_n)$$

nach (5.4.12) und der Tatsache, dass  $\log$  konkav ist. Exponenzieren dieser Ungleichung liefert mit der Monotonie von  $\exp$  dann

$$\begin{aligned} a_1^{\lambda_1} \cdots a_n^{\lambda_n} &= \exp(\lambda_1 \log(a_1)) \cdots \exp(\lambda_n \log(a_n)) \\ &= \exp(\lambda_1 \log(a_1) + \dots + \lambda_n \log(a_n)) \\ &\leq \exp(\log(\lambda_1 a_1 + \dots + \lambda_n a_n)) \\ &= \lambda_1 a_1 + \dots + \lambda_n a_n. \end{aligned} \quad \square$$

Als Spezialfall erhält man aus (5.4.17) für  $\lambda_i = \frac{1}{n}$  die Ungleichung

$$\sqrt[n]{a_1 \cdots a_n} \leq \frac{1}{n}(a_1 + \dots + a_n) \quad (5.4.18)$$

zwischen dem geometrischen und dem arithmetischen Mittel, siehe Korollar 4.2.7 für einen alternativen Beweis. Aus diesem Grunde nennt man (5.4.17) auch die Ungleichung zwischen dem *gewichteten geometrischen und gewichtetem arithmetischem Mittel*.



**Korollar 5.4.10 (Youngsche Ungleichung)** Seien  $p, q > 1$  mit  $\frac{1}{p} + \frac{1}{q} = 1$  gegeben. Dann gilt für alle  $a, b \in \mathbb{R}^+$  die Youngsche Ungleichung

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \quad (5.4.19)$$

BEWEIS: Wir verwenden Korollar 5.4.9 für  $\lambda_1 = \frac{1}{p}$  und  $\lambda_2 = \frac{1}{q}$  sowie für  $a_1 = a^p$  und  $a_2 = b^q$ .  $\square$

**Korollar 5.4.11 (Hölder-Ungleichung)** Seien  $p, q > 1$  mit  $\frac{1}{p} + \frac{1}{q} = 1$  gegeben. Dann gilt für alle  $z_1, \dots, z_n, w_1, \dots, w_n \in \mathbb{K}$  die Hölder-Ungleichung

$$\sum_{k=1}^n |z_k w_k| \leq \left( \sum_{k=1}^n |z_k|^p \right)^{\frac{1}{p}} \left( \sum_{\ell=1}^n |w_\ell|^q \right)^{\frac{1}{q}}. \quad (5.4.20)$$

BEWEIS: Wir setzen

$$a = \sum_{k=1}^n |z_k|^p \quad \text{und} \quad b = \sum_{\ell=1}^n |w_\ell|^q.$$

Die Ungleichung ist trivialerweise erfüllt, wenn  $a = 0$  oder  $b = 0$  gilt, da in diesem Fall alle  $z_k = 0$  oder alle  $w_\ell = 0$  sind. Sei also  $a \neq 0 \neq b$ . Nach der Youngschen Ungleichung erhalten wir

$$\begin{aligned} \frac{|z_k w_k|}{a^{\frac{1}{p}} b^{\frac{1}{q}}} &= \left( \frac{|z_k|^p}{a} \right)^{\frac{1}{p}} \left( \frac{|w_k|^q}{b} \right)^{\frac{1}{q}} \\ &\leq \frac{1}{p} \left( \frac{|z_k|^p}{a} \right)^{\frac{1}{p} \cdot p} + \frac{1}{q} \left( \frac{|w_k|^q}{b} \right)^{\frac{1}{q} \cdot q} \\ &= \frac{1}{p} \frac{|z_k|^p}{a} + \frac{1}{q} \frac{|w_k|^q}{b}. \end{aligned}$$

Summation über alle  $k = 1, \dots, n$  gibt dann

$$\frac{1}{a^{\frac{1}{p}} b^{\frac{1}{q}}} \sum_{k=1}^n |z_k w_k| \leq \frac{1}{p} \frac{1}{a} \sum_{k=1}^n |z_k|^p + \frac{1}{q} \frac{1}{b} \sum_{k=1}^n |w_k|^q = \frac{1}{p} + \frac{1}{q} = 1$$

nach Definition von  $a$  und  $b$ . Umstellen liefert also (5.4.20).  $\square$

**Korollar 5.4.12 (Minkowski-Ungleichung)** Sei  $p \geq 1$ . Dann gilt für alle  $z_1, \dots, z_n, w_1, \dots, w_n \in \mathbb{K}$  die Minkowski-Ungleichung

$$\left( \sum_{k=1}^n |z_k + w_k|^p \right)^{\frac{1}{p}} \leq \left( \sum_{k=1}^n |z_k|^p \right)^{\frac{1}{p}} + \left( \sum_{k=1}^n |w_k|^p \right)^{\frac{1}{p}}. \quad (5.4.21)$$

BEWEIS: Der Fall  $p = 1$  ist die übliche Dreiecksungleichung. Wir nehmen daher  $p > 1$  an und verwenden  $q > 1$  mit  $\frac{1}{p} + \frac{1}{q} = 1$  also  $q = \frac{p}{p-1}$ . Sei dann

$$a = \sum_{k=1}^n |z_k + w_k|^p > 0,$$

der Fall  $a = 0$  liefert eine trivialerweise erfüllte Ungleichung (5.4.21). Dann gilt

$$a = \sum_{k=1}^n |z_k + w_k| |z_k + w_k|^{p-1}$$

$$\begin{aligned}
&\leq \sum_{k=1}^n |z_k| |z_k + w_k|^{p-1} + \sum_{k=1}^n |w_k| |z_k + w_k|^{p-1} \\
&\stackrel{(a)}{\leq} \left( \sum_{k=1}^n |z_k|^p \right)^{\frac{1}{p}} \left( \sum_{k=1}^n |z_k + w_k|^{(p-1)q} \right)^{\frac{1}{q}} + \left( \sum_{k=1}^n |w_k|^p \right)^{\frac{1}{p}} \left( \sum_{k=1}^n |z_k + w_k|^{(p-1)q} \right)^{\frac{1}{q}} \\
&= \left( \sum_{k=1}^n |z_k|^p \right)^{\frac{1}{p}} a^{\frac{1}{q}} + \left( \sum_{k=1}^n |w_k|^p \right)^{\frac{1}{p}} a^{\frac{1}{q}},
\end{aligned}$$

wobei wir in (a) die Hölder-Ungleichung verwenden können, da  $(p-1)q = p$ . Teilt man nun durch  $a^{\frac{1}{q}}$ , erhält man wegen  $aa^{-\frac{1}{q}} = a^{\frac{1}{p}}$  die Ungleichung (5.4.21).  $\square$

Der Spezialfall  $p = 2$  der Hölder-Ungleichung ist die *Cauchy-Schwarz-Ungleichung*, für die man in der linearen Algebra einen gänzlich anderen Beweis findet, siehe etwa [8, Prop. 7.19]. Zum Abschluss dieser Sammlung von Ungleichungen nennen wir noch die folgende Version der *Jensenschen Ungleichung*, auch wenn diese elementar ohne die Konkavität von  $\log$  gezeigt werden kann:

**Korollar 5.4.13 (Jensensche Ungleichung)** Seien  $0 < p < q$  gegeben. Dann gilt für alle  $z_1, \dots, z_n \in \mathbb{K}$  die Jensensche Ungleichung

$$\left( \sum_{k=1}^n |z_k|^q \right)^{\frac{1}{q}} \leq \left( \sum_{k=1}^n |z_k|^p \right)^{\frac{1}{p}}. \quad (5.4.22)$$

BEWEIS: Ohne Einschränkung seien alle Zahlen  $z_1, \dots, z_n$  ungleich Null: Ist nämlich ein  $z_k = 0$ , so trägt es auf beiden Seiten der Ungleichung nicht bei. Wir setzen

$$a = \left( \sum_{k=1}^n |z_k|^p \right)^{\frac{1}{p}},$$

womit  $a > 0$  folgt. Insbesondere gilt dann  $|z_k|^p \leq a^p$  und somit  $\frac{|z_k|}{a} \leq 1$ . Die Monotonie des Exponenzierens liefert damit

$$\frac{|z_k|^q}{a^q} \leq \frac{|z_k|^p}{a^p}.$$

Summation über  $k$  ergibt dann

$$\frac{1}{a^q} \sum_{k=1}^n |z_k|^q \leq \frac{1}{a^p} \sum_{k=1}^n |z_k|^p = 1,$$

also  $\sum_{k=1}^n |z_k|^q \leq a^q$ , was (5.4.22) zeigt.  $\square$

## 5.5 Konvergenz von differenzierbaren Funktionen

Wie schon bei stetigen Funktionen wollen wir nun auch Funktionenfolgen  $(f_n)_{n \in \mathbb{N}}$  von differenzierbaren Funktionen  $f_n: X \rightarrow \mathbb{K}$  betrachten und Kriterien formulieren, unter welchen ein punktwiser Limes

$$f = \lim_{n \rightarrow \infty} f_n \quad (5.5.1)$$

wieder differenzierbar ist. Da differenzierbare Funktionen insbesondere stetig sind, muss also dazu die Grenzfunktion  $f$  zumindest stetig sein. Im Hinblick auf Satz 3.7.10 wird also eine lokal gleichmäßige Konvergenz der  $f_n$  sicherlich nützlich sein.

Das folgende Resultat zeigt nun aber, dass gleichmäßige Konvergenz noch nicht ausreicht, um die Differenzierbarkeit von Grenzfunktionen zu garantieren:

Abbildung 5.15: Die Betragsfunktion und einige der ersten approximierenden Polynome  $q_n$ .

**Proposition 5.5.1** *Seien  $p_n \in \mathbb{R}[X]$  die Polynome aus Proposition 3.7.1. Dann konvergieren die Polynome*

$$q_n(x) = p_n(x^2) \quad (5.5.2)$$

*auf  $[-1, 1]$  gleichmäßig gegen die Betragsfunktion*

$$f(x) = |x|. \quad (5.5.3)$$

BEWEIS: Wir wissen aus Proposition 3.7.1, dass für  $y \in [0, 1]$  die Abschätzung

$$0 \leq \sqrt{y} - p_n(y) \leq \frac{2\sqrt{y}}{2 + n\sqrt{y}} \quad (*)$$

für alle  $n \in \mathbb{N}$  gilt. Dies bedeutet aber für alle  $x \in [-1, 1]$  die Abschätzung

$$0 \leq |x| - q_n(x) \leq \frac{2|x|}{2 + n|x|}, \quad (**)$$

indem wir einfach  $y = x^2$  in  $(*)$  setzten. Damit gilt aber für alle  $x \in [-1, 1]$

$$0 \leq |x| - q_n(x) \leq \frac{2}{n},$$

was die gleichmäßige Konvergenz  $q_n \rightarrow f$  zeigt, siehe auch Abbildung 5.15.  $\square$

Da die Betragsfunktion bei  $x = 0$  sicherlich nicht differenzierbar ist, erlaubt es die gleichmäßige Konvergenz der (sogar glatten) Polynome also nicht, auf die Differenzierbarkeit der Grenzfunktion zu schließen. Nach Satz 3.7.10 wissen wir die Stetigkeit der Grenzfunktion, mehr aber nicht. Wir benötigen stärkere Voraussetzungen als (lokal) gleichmäßige Konvergenz, um auch Differenzierbarkeitseigenschaften zu erhalten. Die folgende technische Proposition liefert hier eine Option:

**Proposition 5.5.2** *Sei  $X$  eine offene und konvexe Teilmenge von reellen oder komplexen Zahlen. Seien weiter  $f_n: X \rightarrow \mathbb{K}$  differenzierbare Funktionen mit folgenden Eigenschaften:*

*i.) Es gibt einen Punkt  $x_0$ , so dass  $(f_n(x_0))_{n \in \mathbb{N}}$  konvergiert.*

*ii.) Die Folge der Ableitungen  $(f'_n)_{n \in \mathbb{N}}$  konvergiert lokal gleichmäßig auf  $X$ .*

*Dann konvergiert auch die Folge  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen eine differenzierbare Funktion*

$$\lim_{n \rightarrow \infty} f_n = f \quad (5.5.4)$$

*auf  $X$  mit*

$$\lim_{n \rightarrow \infty} f'_n = f'. \quad (5.5.5)$$

BEWEIS: Wir zeigen zunächst die lokal gleichmäßige Konvergenz von  $(f_n)_{n \in \mathbb{N}}$ . Sei dazu  $x_1 \in X$  vorgegeben. Wir wählen ein  $r > 0$  derart, dass  $B_r(x_1)^{\text{cl}} \subseteq X$ , was möglich ist, da  $X$  offen ist. Da  $X$  zudem konvex ist, liegt die Verbindungsstrecke von  $x_0$  zu  $x$  ganz in  $X$ , also

$$x(t) = tx_1 + (1 - t)x_0 \in X$$

für  $t \in [0, 1]$ . Wir betrachten nun folgende Teilmenge

$$K = \left\{ tx + (1 - t)x_0 \mid x \in B_r(x_1)^{\text{cl}} \text{ und } t \in [0, 1] \right\}.$$

Abbildung 5.16: Die Menge der Verbindungsstrecken von  $x_0$  zu Punkten aus  $B_r(x)^{\text{cl}}$  ist kompakt.

Abbildung 5.17: Das Kompaktum  $K$  ist konvex.

Da  $X$  konvex ist und  $B_r(x_1)^{\text{cl}} \subseteq X$  gilt, folgt  $K \subseteq X$ . Wir behaupten, dass  $K$  kompakt ist, siehe Abbildung 5.16. Es ist klar, dass  $K$  beschränkt ist, da nur Zahlen und  $x$  mit beschränkter Größe verwendet werden. Etwas schwieriger ist die Abgeschlossenheit. Sind  $y_n \in K$  Punkte in  $K$  mit  $y_n \rightarrow y$ , so gibt es also  $t_n \in [0, 1]$  und  $z_n \in B_r(x_1)^{\text{cl}}$  mit  $y_n = t_n z_n + (1 - t_n)x_0$ . Da  $[0, 1]$  kompakt ist, gibt es eine konvergente Teilfolge  $t_{n_k} \rightarrow t \in [0, 1]$ . Da auch  $B_r(x_1)^{\text{cl}}$  kompakt ist, hat die Teilfolge  $(z_{n_k})_{k \in \mathbb{N}}$  ihrerseits auch eine konvergente Teilfolge  $z_{n_{k_\ell}} \rightarrow z \in B_r(x_1)^{\text{cl}}$ . Dann ist aber  $y_{n_{k_\ell}} = t_{n_{k_\ell}} z_{n_{k_\ell}} + (1 - t_{n_{k_\ell}})x_0$  wieder konvergent gegen  $tz + (1 - t)x_0$ . Insgesamt folgt also  $y = tz + (1 - t)x_0 \in K$ , was die Abgeschlossenheit von  $K$  zeigt. Einen einfacheren konzeptuellen Beweis finden wir später in Übung 7.8.13. Eine weitere geometrische Überlegung zeigt dann, dass  $K$  konvex ist: Sind  $z, y \in K$ , so gibt es Punkte  $z', y' \in B_r(x_1)^{\text{cl}}$  und  $t, s \in [0, 1]$  mit

$$z = tz' + (1 - t)x_0$$

sowie

$$y = sy' + (1 - s)x_0.$$

Da die abgeschlossene Kugel  $B_r(x_1)^{\text{cl}}$  konvex ist, ist für alle  $\tau \in [0, 1]$  auch  $\tau z' + (1 - \tau)y' \in B_r(x_1)^{\text{cl}}$ . Damit zeigt man nun durch elementare geometrische Überlegung, dass die Verbindungsstrecke zwischen  $z$  und  $y$  wieder in  $K$  verläuft, siehe auch Abbildung 5.17. Wir können nun den Mittelwertsatz in Form von (5.3.10) im reellen Fall beziehungsweise den Schrankensatz aus Satz 5.3.10 im komplexen Fall anwenden. Wir betrachten dazu ein  $\epsilon > 0$  und finden ein  $N \in \mathbb{N}$  mit

$$|f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2} \quad (*)$$

sowie

$$\|f'_n - f'_m\|_K < \min \left\{ \frac{\epsilon}{2(|x_1 - x_0| + r)}, \frac{\epsilon}{2} \right\} \quad (**)$$

für  $n, m \geq N$ , indem wir die Konvergenz in *i.*) als Cauchy-Bedingung sowie die gleichmäßige Konvergenz aus *ii.*) auf  $K$  als Cauchy-Bedingung zum Einsatz bringen. Dann gilt für diese  $n, m$  und die differenzierbare Funktion  $f_n - f_m$  auf  $K$

$$|(f_n - f_m)(x) - (f_n - f_m)(y)| \leq \|f'_n - f'_m\|_K |x - y| \stackrel{(**)}{<} \frac{\epsilon}{2} \quad (*)$$

nach (5.3.10) beziehungsweise Satz 5.3.10, wobei wir die elementar-geometrische Abschätzung  $|x - y| \leq |x_1 - x_0| + r$  verwenden. Damit gilt insbesondere für alle  $x \in \mathbb{K}$

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)| < \epsilon,$$

indem wir  $y = x_0$  verwenden und  $(*)$  sowie  $(*)$  benutzen. Also ist  $(f_n)_{n \in \mathbb{N}}$  eine auf  $K$  gleichmäßige Cauchy-Folge. Es folgt die Existenz einer Funktion  $f: X \rightarrow \mathbb{K}$  mit lokal gleichmäßiger Konvergenz (5.5.4), da wir um jeden Punkt  $x_1 \in X$  eine solche Umgebung  $K$  finden. Nach Satz 3.7.10 ist  $f \in \mathcal{C}(X, \mathbb{K})$  stetig. Um zu zeigen, dass  $f$  differenzierbar ist, betrachten wir für einen Punkt  $x_1 \in X$  die Differenzenquotienten

$$F_n(x) = \begin{cases} \frac{f_n(x) - f_n(x_1)}{x - x_1} & x \neq x_1 \\ f'_n(x_1) & x = x_1 \end{cases}$$

sowie

$$F(x) = \frac{f(x) - f(x_1)}{x - x_1} \quad \text{für } x \in X \setminus \{x_1\}.$$

Die Differenzierbarkeit von  $f_n$  bedeutet gerade, dass  $F_n$  stetig auf ganz  $X$  ist. Für  $x \in K$  finden wir wieder

$$|F_n(x) - F_m(x)| = \frac{|f_n(x) - f_n(x_1) - f_m(x) + f_m(x_1)|}{|x - x_1|} \stackrel{(*)}{<} \frac{\epsilon}{2},$$

falls  $n, m \geq N$  gilt und  $x \neq x_1$ . Für  $x = x_1$  gilt dagegen

$$|F_n(x_1) - F_m(x_1)| = |f'_n(x_1) - f'_m(x_1)| \stackrel{(*)}{<} \frac{\epsilon}{2},$$

womit wir für alle  $x \in K$

$$|F_n(x) - F_m(x)| < \frac{\epsilon}{2}$$

gezeigt haben. Damit konvergiert  $(F_n)_{n \in \mathbb{N}}$  auf  $K$  gleichmäßig. Für  $x \neq x_1$  gilt

$$F(x) = \lim_{n \rightarrow \infty} F_n(x),$$

da  $f_n \rightarrow f$ . Für  $x = x_1$  folgt dagegen

$$\lim_{n \rightarrow \infty} F_n(x_1) = \lim_{n \rightarrow \infty} f'_n(x_1).$$

Da die Konvergenz gleichmäßig auf  $K$  ist, gibt es eine stetige Grenzfunktion  $\lim_{n \rightarrow \infty} F_n$  auf  $K$  nach Satz 3.7.10. Auf  $K \setminus \{x_1\}$  stimmt diese mit  $F$  überein. Daher ist

$$F(x) = \begin{cases} \frac{f(x) - f(x_1)}{x - x_1} & x \neq x_1 \\ \lim_{n \rightarrow \infty} f'_n(x_1) & x = x_1 \end{cases}$$

die stetige Grenzfunktion und insbesondere eine stetige Fortsetzung der vormalig nur auf  $X \setminus \{x_1\}$  definierten Funktion  $F$ . Da wir in  $K$  eine offene Kugel  $B_r(x_1) \subseteq K$  enthalten haben, zeigt das die Existenz der Ableitung von  $f$  bei  $x_1$  mit Wert  $f'(x_1) = \lim_{n \rightarrow \infty} f'_n(x_1)$ , dies ist gerade die Stetigkeit von  $F$  bei  $x = x_1$ .  $\square$

Die Voraussetzung *i.)* in Proposition 5.5.2 ist tatsächlich notwendig, wie das triviale Beispiel einer Folge von konstanten Funktionen zeigt: Ist  $f_n(x) = c_n$ , so gilt  $f'_n(x) = 0$  für alle  $x \in X$ . Daher ist die gleichmäßige Konvergenz  $f'_n \rightarrow 0$  trivialerweise erfüllt. Trotzdem muss die Folge  $(c_n)_{n \in \mathbb{N}}$  natürlich nicht konvergieren, womit  $(f_n)_{n \in \mathbb{N}}$  nicht einmal punktweise konvergiert.

Induktiv erhält man daraus folgendes Resultat für  $k$ -mal differenzierbare Funktionen:

**Proposition 5.5.3** *Sei  $k \in \mathbb{N}$ . Sei  $X$  eine offene und konvexe Teilmenge von reellen oder komplexen Zahlen. Seien weiter  $k$ -mal differenzierbare Funktionen  $f_n: X \rightarrow \mathbb{K}$  mit der Eigenschaft gegeben, dass  $f_n^{(k)}$  lokal gleichmäßig auf  $X$  konvergiert und dass es Punkte  $x_0, \dots, x_{k-1} \in X$  gibt, für die die Folgen  $(f_n(x_0))_{n \in \mathbb{N}}, (f'_n(x_1))_{n \in \mathbb{N}}, \dots, (f_n^{(k-1)}(x_{k-1}))_{n \in \mathbb{N}}$  konvergieren. Dann konvergiert  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen eine  $k$ -mal differenzierbare Funktion  $f: X \rightarrow \mathbb{K}$  derart, dass für alle  $\ell = 0, \dots, k$*

$$\lim_{n \rightarrow \infty} f_n^{(\ell)} = f^{(\ell)} \quad (5.5.6)$$

*lokal gleichmäßig auf  $X$ .*

**BEWEIS:** Wir wenden Proposition 5.5.2 auf  $f_n^{(k-1)}$  an: Diese Funktionen sind differenzierbar, konvergieren an einem Punkt  $x_{k-1} \in X$  und besitzen lokal gleichmäßig konvergente Ableitungen  $(f_n^{(k-1)})' = f_n^{(k)}$ . Also konvergieren  $f_n^{(k-1)}$  lokal gleichmäßig auf  $X$  gegen eine differenzierbare Funktion, deren Ableitung der Grenzwert der  $f_n^{(k)}$  ist. So verfährt man nun weiter und zeigt die Behauptung induktiv  $\square$

**Bemerkung 5.5.4** Die Konvexität von  $X$  ist am Ende nicht entscheidend: Solange  $X$  wegzusammenhängend ist, lässt sich der Beweis anpassen. Wir wollen dies nicht tun, da zuerst zu klären wäre, was wegzusammenhängend in diesem Kontext bedeutet. Dies werden wir in Abschnitt 7.6 nachholen. Die Situation  $X \subseteq \mathbb{C}$  mit komplex-differenzierbaren Funktionen benützt unnötig starke Annahmen, die sich aber erst im weiteren Studium der komplex-differenzierbaren Funktionen klären lassen. Dies bleibt der Funktionentheorie vorbehalten.

In vielen Anwendungen sind die Funktionen nicht nur  $k$ -mal differenzierbar sondern sogar  $\mathcal{C}^k$ , also  $k$ -mal stetig differenzierbar. Dann lässt sich die lokal gleichmäßige Konvergenz der Ableitungen  $(f_n^{(k)})_{n \in \mathbb{N}}$  mit Hilfe der lokalen Supremumsnormen kontrollieren. Die Folge  $(f_n^{(k)})_{n \in \mathbb{N}}$  konvergiert genau dann lokal gleichmäßig gegen eine Grenzfunktion  $g: X \rightarrow \mathbb{K}$ , wenn für jedes kompakte  $K \subseteq X$

$$\lim_{n \rightarrow \infty} \|f_n^{(k)} - g\|_K = 0 \quad (5.5.7)$$

gilt, siehe Proposition 3.7.7. Da  $f_n^{(k)}$  als stetig angenommen ist, ist auch  $f_n^{(k)} - f_m^{(k)}$  stetig und folglich lokal beschränkt, so dass  $\|f_n^{(k)} - f_m^{(k)}\|_K < \infty$  endlich ist. Dann ist die Konvergenz von  $(f_n^{(k)})_{n \in \mathbb{N}}$  äquivalent zur Cauchy-Bedingung, dass es zu allen kompakten  $K \subseteq X$  und zu allen  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit

$$\|f_n^{(k)} - f_m^{(k)}\|_K < \epsilon \quad (5.5.8)$$

für alle  $n, m \geq N$  gibt, siehe Proposition 3.7.8, *iv.*). Man betrachte, dass für das Cauchy-Kriterium alle Folgenglieder lokal beschränkt sein müssen, was durch die Forderung nach stetigen Ableitungen erreicht ist. Diese Überlegungen führen daher auf folgende Definition:

**Definition 5.5.5 ( $\mathcal{C}^k$ -Konvergenz)** Sei  $X$  eine offene Teilmenge von reellen oder komplexen Zahlen und sei  $k \in \mathbb{N} \cup \{+\infty\}$ .

- i.) Eine Folge von  $\mathcal{C}^k$ -Funktionen  $(f_n)_{n \in \mathbb{N}}$  mit  $f_n \in \mathcal{C}^k(X, \mathbb{K})$  konvergiert gegen eine Funktion  $f \in \mathcal{C}^k(X, \mathbb{K})$  im  $\mathcal{C}^k$ -Sinne, wenn alle Ableitungen  $f_n^{(\ell)}$  mit  $\ell \leq k$  lokal gleichmäßig gegen  $f^{(\ell)}$  auf  $X$  konvergieren.
- ii.) Eine Folge  $(f_n)_{n \in \mathbb{N}}$  von  $\mathcal{C}^k$ -Funktionen  $f_n \in \mathcal{C}^k(X, \mathbb{K})$  heißt  $\mathcal{C}^k$ -Cauchy-Folge, wenn es für alle  $\epsilon > 0$ , alle  $\ell \in \mathbb{N}$  mit  $0 \leq \ell \leq k$  und alle kompakten Teilmengen  $K \subseteq X$  ein  $N \in \mathbb{N}$  gibt, so dass

$$\|f_n^{(\ell)} - f_m^{(\ell)}\|_K < \epsilon \quad (5.5.9)$$

für  $n, m \geq N$  gilt.

**Satz 5.5.6 (Vollständigkeit von  $\mathcal{C}^k(X, \mathbb{K})$ )** Sei  $X$  eine offene Teilmenge der reellen oder komplexen Zahlen und  $k \in \mathbb{N} \cup \{+\infty\}$ . Eine Folge  $(f_n)_{n \in \mathbb{N}}$  von  $\mathcal{C}^k$ -Funktionen  $f_n \in \mathcal{C}^k(X, \mathbb{K})$  konvergiert genau dann gegen eine  $\mathcal{C}^k$ -Funktion  $f$  im  $\mathcal{C}^k$ -Sinne, wenn  $(f_n)_{n \in \mathbb{N}}$  eine  $\mathcal{C}^k$ -Cauchy-Folge ist.

BEWEIS: Sei  $f \in \mathcal{C}^k(X, \mathbb{K})$  die Grenzfunktion einer  $\mathcal{C}^k$ -konvergenten Folge  $(f_n)_{n \in \mathbb{N}}$  von  $\mathcal{C}^k$ -Funktionen  $f_n \in \mathcal{C}^k(X, \mathbb{K})$ . Sei weiter  $K \subseteq X$  kompakt,  $\ell \in \mathbb{N}$  mit  $0 \leq \ell \leq k$  und  $\epsilon > 0$  vorgegeben. Dann gibt es ein  $N \in \mathbb{N}$  mit

$$\|f_n^{(\ell)} - f^{(\ell)}\|_K < \frac{\epsilon}{2}$$

für  $n \geq N$ . Damit gilt für  $n, m \geq N$  auch

$$\|f_n^{(\ell)} - f_m^{(\ell)}\|_K \leq \|f_n^{(\ell)} - f^{(\ell)}\|_K + \|f^{(\ell)} - f_m^{(\ell)}\|_K < \epsilon,$$

was zeigt, dass  $(f_n)_{n \in \mathbb{N}}$  eine  $\mathcal{C}^k$ -Cauchy-Folge ist. Sei also umgekehrt  $(f_n)_{n \in \mathbb{N}}$  eine  $\mathcal{C}^k$ -Cauchy-Folge. Dann ist für jeden Punkt  $x \in X$  die Folge  $(f_n^{(\ell)}(x))_{n \in \mathbb{N}}$  eine Cauchy-Folge von Zahlen in  $\mathbb{K}$ , wobei

$\ell \in \mathbb{N}$  mit  $0 \leq \ell \leq k$ . Also konvergiert  $(f_n^{(\ell)})_{n \in \mathbb{N}}$  punktweise auf  $X$  wegen der Vollständigkeit von  $\mathbb{K}$ . Für jedes  $\ell \in \mathbb{N}$  mit  $0 \leq \ell \leq k$  ist  $(f_n^{(\ell)})_{n \in \mathbb{N}}$  nach Voraussetzung eine lokal gleichmäßige Cauchy-Folge von stetigen und daher lokal beschränkten Funktionen. Diese konvergiert nach Proposition 3.7.8, *iv.*), gegen eine lokal beschränkte Funktion auf lokal gleichmäßige Weise. Nach Satz 3.7.10 ist dieser Limes wieder eine stetige Funktion. Nach Proposition 5.5.3 folgt auf jeder konvexen offenen Teilmenge die lokal gleichmäßige Konvergenz von  $(f_n)_{n \in \mathbb{N}}$  gegen eine  $k$ -mal differenzierbare Funktion  $f$ , so dass die Ableitungen  $(f_n^{(\ell)})_{n \in \mathbb{N}}$  lokal gleichmäßig gegen  $f^{(\ell)}$  konvergieren. Da  $X$  durch konvexe offene Teilmengen überdeckt wird, beispielsweise durch offene Kugeln, erhält man damit lokal gleichmäßige Konvergenz auf ganz  $X$ . Ist schließlich  $k \in \mathbb{N}$ , so ist auch die  $k$ -te Ableitung von  $f$  wieder stetig, da die  $f_n^{(k)}$  stetig sind und lokal gleichmäßig konvergieren.  $\square$

**Bemerkung 5.5.7** Mit diesem Satz haben wir das erste Mal ein Vollständigkeitsresultat für Funktionenräume formuliert und gezeigt. Wir werden dies gelegentlich benutzen, aber an dieser Stelle nicht weiter vertiefen. Eine Einordnung dieses Resultats in einen konzeptuellen Rahmen geschieht dann in der Funktionalanalysis.

Da wir oftmals Grenzfunktionen in Form einer Reihe erhalten, formulieren wir Satz 5.5.6 auch noch separat für Reihen:

**Korollar 5.5.8** *Sei  $X$  eine offene Teilmenge der reellen oder komplexen Zahlen und  $k \in \mathbb{N} \cup \{+\infty\}$ . Eine Reihe  $\sum_{n=0}^{\infty} f_n$  von  $\mathcal{C}^k$ -Funktionen  $f_n \in \mathcal{C}^k(X, \mathbb{K})$  konvergiert genau dann gegen eine  $\mathcal{C}^k$ -Funktion  $f \in \mathcal{C}^k(X, \mathbb{K})$  im  $\mathcal{C}^k$ -Sinne, wenn  $\sum_{n=0}^{\infty} f_n$  eine  $\mathcal{C}^k$ -Cauchy-Reihe ist, d.h. die Partialsummen bilden eine  $\mathcal{C}^k$ -Cauchy-Folge.*

Dies ist tatsächlich nur eine Umformulierung. Nützlich bei der Entscheidung ist nun folgendes *Majorantenkriterium*:

**Korollar 5.5.9** *Sei  $X$  eine offene Teilmenge der reellen oder komplexen Zahlen und  $k \in \mathbb{N} \cup \{+\infty\}$ . Sei weiter  $\sum_{n=0}^{\infty} f_n$  eine Reihe von  $\mathcal{C}^k$ -Funktionen  $f_n \in \mathcal{C}^k(X, \mathbb{K})$ . Gibt es dann zu jedem Kompaktum  $K \subseteq X$  und jedem  $\ell \in \mathbb{N}$  mit  $0 \leq \ell \leq k$  Zahlen  $c_{K,\ell,n} > 0$ , so dass*

$$\|f_n^{(\ell)}\|_K \leq c_{K,\ell,n} \quad (5.5.10)$$

und

$$\sum_{n=0}^{\infty} c_{K,\ell,n} < \infty, \quad (5.5.11)$$

so konvergiert  $\sum_{n=0}^{\infty} f_n$  im  $\mathcal{C}^k$ -Sinne.

BEWEIS: Sei  $\epsilon > 0$  gegeben und  $N \in \mathbb{N}$  derart, dass

$$\sum_{r=n}^m c_{K,\ell,r} < \epsilon$$

für alle  $n, m \geq N$ . Da (5.5.11) konvergiert, ist dies gerade die Eigenschaft einer Cauchy-Reihe. Damit gilt auch

$$\left\| \sum_{r=n}^m f_r^{(\ell)} \right\|_K \leq \sum_{r=n}^m \|f_r^{(\ell)}\|_K \leq \sum_{r=n}^m c_{K,\ell,r} < \epsilon,$$

womit die Reihe  $\sum_{n=0}^{\infty} f_n^{(\ell)}$  eine Cauchy-Reihe im  $\mathcal{C}^k$ -Sinne ist. Wir können nun Korollar 5.5.8 zum Einsatz bringen.  $\square$

**Beispiel 5.5.10 (Die Riemannsche  $\zeta$ -Funktion)** Wir wollen nun erneut die  $\zeta$ -Funktion

$$\zeta(z) = \sum_{n=0}^{\infty} \frac{1}{n^z} \quad (5.5.12)$$

aus Beispiel 2.6.8 betrachten. Dort hatten wir die absolute Konvergenz für reelle  $z \in \mathbb{R}$  mit  $z > 1$  gezeigt. Wir erweitern dieses Konvergenzresultat zunächst auf alle  $z \in \mathbb{C}$  mit  $\operatorname{Re}(z) > 1$ . Dazu schreiben wir

$$\frac{1}{n^z} = n^{-z} = (e^{\log(n)})^{-z} = e^{-z \log(n)}, \quad (5.5.13)$$

womit wir jeden einzelnen Summanden auch für  $z \in \mathbb{C}$  verwenden können. Sei also  $z = \alpha + i\beta \in \mathbb{C}$  mit  $\alpha, \beta \in \mathbb{R}$ . Dann gilt

$$\left| \frac{1}{n^z} \right| = |e^{-(\alpha + i\beta) \log(n)}| = |e^{-\alpha \log(n)} e^{-i\beta \log(n)}| = e^{-\alpha \log(n)},$$

da  $e^{-i\beta \log(n)} \in \mathbb{S}^1$  vom Betrag 1 ist. Der Imaginärteil von  $z$  beeinflusst die absolute Konvergenz daher nicht, was zeigt, dass

$$\zeta: X = \{z \in \mathbb{C} \mid \operatorname{Re}(z) > 1\} \longrightarrow \mathbb{C} \quad (5.5.14)$$

auf dieser offenen Halbebene  $X \subseteq \mathbb{C}$  definiert ist. Jeder einzelne Summand ist glatt, und

$$(n^{-z})' = (e^{-z \log(n)})' = -\log(n) n^{-z} \quad (5.5.15)$$

zeigt, dass die Ableitung der Ordnung  $k$  durch

$$(n^{-z})^{(k)} = (-\log(n))^k n^{-z} \quad (5.5.16)$$

gegeben sind. Sei nun  $\epsilon > 0$  gewählt und  $z \in X$  mit  $\operatorname{Re}(z) > 1 + 2\epsilon$ . Nach Proposition 4.2.8, *i.*), gilt

$$\lim_{n \rightarrow \infty} \frac{\log(n)}{n^\epsilon} = 0 \quad (5.5.17)$$

und damit auch

$$\lim_{n \rightarrow \infty} \frac{(\log(n))^k}{n^\epsilon} = 0, \quad (5.5.18)$$

da  $\epsilon > 0$  beliebig war. Es folgt, dass

$$\begin{aligned} |(n^{-z})^{(k)}| &= |\log(n)^k n^{-z}| \\ &\leq |\log(n)|^k n^{-(1+2\epsilon)} \\ &= \frac{|\log(n)|^k}{n^\epsilon} n^{-(1+\epsilon)} \\ &\leq c_k \cdot n^{-(1+\epsilon)} \end{aligned}$$

mit einer Konstanten  $c_k = \sup_{n \in \mathbb{N}} \frac{|\log(n)|^k}{n^\epsilon} < \infty$ , da  $\frac{\log(n)^k}{n^\epsilon}$  ja eine Nullfolge ist. Damit gilt aber, dass

$$\sum_{n=1}^{\infty} |(n^{-z})^{(k)}| \leq c_k \sum_{n=1}^{\infty} n^{-(1+\epsilon)} < \infty \quad (5.5.19)$$

durch eine absolut konvergente Reihe majorisiert wird, und zwar durch dieselbe Reihe für alle  $z$  mit  $\operatorname{Re}(z) > 1 + 2\epsilon$ . Dies genügt nach Korollar 5.5.9, um die gleichmäßige Cauchy-Eigenschaft von  $\sum_{n=1}^{\infty} (n^{-z})^{(k)}$  auf dem offenen Teilgebiet mit  $\operatorname{Re}(z) > 1 + 2\epsilon$  zu zeigen. Da  $\epsilon > 0$  beliebig war, liegt eine lokal gleichmäßige Cauchy-Reihe auf ganz  $X$  vor und wir können Korollar 5.5.9 beziehungsweise



Korollar 5.5.8 anwenden. Damit ist die Grenzfunktion  $\zeta$  eine unendlich oft differenzierbare Funktion mit Ableitung

$$\zeta'(z) = - \sum_{n=1}^{\infty} \frac{\log(n)}{n^z} \quad (5.5.20)$$

und entsprechend für höhere Ableitungen. Wir haben so zu einem den Definitionsbereich drastisch vergrößert und zum anderen wichtige analytische Eigenschaften von  $\zeta$  gezeigt. Mit Methoden der Funktionentheorie lässt sich  $\zeta$  weiter fortsetzen, so dass der endgültige Definitionsbereich von  $\zeta$  dann  $\mathbb{C} \setminus \{1\}$  wird.

## 5.6 Potenzreihen und Taylor-Entwicklung

Für eine (komplexe) Potenzreihe

$$a(z) = \sum_{n=0}^{\infty} a_n z^n \quad (5.6.1)$$

mit positivem Konvergenzradius  $R > 0$  wissen wir nach Satz 3.8.8, dass die Reihe in  $B_R(0) \subseteq \mathbb{C}$  lokal gleichmäßig konvergiert und daher Dank der Stetigkeit der einzelnen Summanden wieder eine stetige Funktion  $a: B_R(0) \rightarrow \mathbb{C}$  liefert. Nun sind die Summanden aber nicht nur stetig sondern glatt, womit sich die Frage stellt, ob die Konvergenz (5.6.1) sogar für alle Ableitungen lokal gleichmäßig erfolgt, also im  $\mathcal{C}^\infty$ -Sinne vorliegt. In diesem Fall könnten wir Satz 5.5.6 zum Einsatz bringen und auf die Glattheit von  $a$  schließen. Das ist tatsächlich immer der Fall:

**Satz 5.6.1 (Glattheit von Potenzreihen)** Sei  $a(z) = \sum_{n=0}^{\infty} a_n z^n$  eine komplexe Potenzreihe mit Konvergenzradius  $R > 0$ .

i.) Auf  $B_R(0)$  konvergiert  $\sum_{n=0}^{\infty} a_n z^n$  lokal gleichmäßig in allen Ableitungen, also im  $\mathcal{C}^\infty$ -Sinne, gegen  $a(z)$ .

ii.) Für  $0 < r < R$  gilt die uniforme Abschätzung

$$\sum_{n=0}^{\infty} \|(a_n z^n)^{(\ell)}\|_{B_r(0)^{\text{cl}}} \leq \sum_{n=0}^{\infty} n(n-1) \cdots (n-\ell+1) |a_n| r^{n-\ell} < \infty \quad (5.6.2)$$

iii.) Für alle  $\ell \in \mathbb{N}_0$  gilt

$$a^{(\ell)}(z) = \sum_{n=\ell}^{\infty} n(n-1) \cdots (n-\ell+1) a_n z^{n-\ell} \quad (5.6.3)$$

und insbesondere

$$a^{(\ell)}(0) = \ell! a_\ell. \quad (5.6.4)$$

iv.) Es gilt

$$a(z) = \sum_{n=0}^{\infty} \frac{a^{(n)}(0)}{n!} z^n. \quad (5.6.5)$$

BEWEIS: Zunächst ist klar, dass  $f_n(z) = a_n z^n$  eine glatte Funktion mit Ableitungen

$$f_n^{(\ell)}(z) = n(n-1) \cdots (n-\ell+1) a_n z^{n-\ell}$$

für  $0 \leq \ell \leq n$  und  $f_n^{(\ell)} = 0$  für  $\ell > n$  ist. Für  $z \in B_r(0)^{\text{cl}}$  gilt daher

$$|f_n^{(\ell)}(z)| \leq n(n-1) \cdots (n-\ell+1) |a_n| r^{n-\ell} \quad (*)$$

für  $0 \leq \ell \leq n$  und wieder  $|f_n^{(\ell)}(z)| = 0$  für  $\ell > n$ . Damit ist aber

$$\|f_n^{(\ell)}(z)\|_{B_r(z)^{\text{cl}}} \leq n(n-1) \cdots (n-\ell+1) |a_n| r^{n-\ell}$$

gezeigt. Wir zeigen, dass dies eine konvergente Reihe liefert. Dazu bestimmen wir den Konvergenzradius der Reihe mit Koeffizienten  $n(n-1) \cdots (n-\ell+1)a_n$  anstelle der  $a_n$ . Wir erhalten

$$\limsup_{n \in \mathbb{N}} \sqrt[n]{|a_n|} \leq \limsup_{n \in \mathbb{N}} \sqrt[n]{n(n-1) \cdots (n-\ell+1) |a_n|} \leq \limsup_{n \in \mathbb{N}} \sqrt[n]{n^\ell |a_n|} = \limsup_{n \in \mathbb{N}} \sqrt[n]{|a_n|},$$

da  $\lim_{n \rightarrow \infty} \sqrt[n]{n^\ell} = 1$  konvergiert und daher den Limes superior nicht weiter verändert, siehe Beispiel 2.3.12, *ii.*). Damit hat diese Reihe den selben Konvergenzradius  $R$ . Nach Satz 3.8.3, *i.*), konvergiert die rechte Seite von (5.6.2) absolut für  $r < R$ . Damit ist *ii.*) gezeigt. Wir können nun Korollar 5.5.9 anwenden und erhalten direkt *i.*) ebenso wie *iii.*) nach Proposition 5.5.3. Da wir die Ableitung bei 0 einfach bestimmen können, indem wir jeden Summanden in (5.6.3) bei  $z = 0$  auswerten, folgt (5.6.4) beziehungsweise (5.6.5) sofort.  $\square$

**Bemerkung 5.6.2** Mit diesem Satz erhalten wir eine konzeptuelle und unabhängige Erklärung für die Glattheit vieler zuvor definierter Funktionen wie beispielsweise  $\exp$ ,  $\sin$  und  $\cos$ . Oftmals werden wir weitere interessante Funktionen in der Analysis durch Potenzreihen definieren, womit wir also direkt ihre Glattheit sichergestellt haben.

Bislang haben wir formale Potenzreihen  $a \in \mathbb{C}[[z]]$  und ihre Konvergenz in Kreisscheiben um  $z_0 = 0$  betrachtet. Diese Einschränkung ist weder nötig noch sinnvoll, da wir neue Funktionen leicht durch Translation erhalten können. Ist also  $z_0 \in \mathbb{C}$  ein beliebiger Punkt und  $a \in \mathbb{C}[[z]]$  eine Potenzreihe mit Konvergenzradius  $R > 0$ , so liefert

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad (5.6.6)$$

eine konvergente Reihe für  $z - z_0 \in B_R(0)$ . Dies ist aber gleichbedeutend zu

$$z \in B_R(z_0), \quad (5.6.7)$$

was durch eine einfache Verschiebung um  $z_0$  zu sehen ist. Die Aussagen von Satz 5.6.1 übertragen sich auf diesen Fall ohne weitere Schwierigkeiten. Dazu überlegen wir uns zuerst folgende Translationsinvarianz des Ableitungsbegriffs:

**Proposition 5.6.3** Sei  $X \subseteq \mathbb{K}$  und  $z_0 \in \mathbb{K}$  gegeben. Zu einer Funktion  $f: X \rightarrow \mathbb{K}$  ist die translatierte Funktion  $\tau_{-z_0}^* f = f \circ \tau_{-z_0}$  auf  $\tau_{z_0}(X) \subseteq X$  definiert.

- i.) Die Funktion  $f$  ist genau dann bei  $x_0 \in X$  stetig, wenn  $\tau_{-z_0}^* f$  bei  $\tau_{z_0}(x_0)$  stetig ist.*
- ii.) Die Funktion  $f$  ist genau dann bei  $x_0 \in X$  differenzierbar wenn  $\tau_{-z_0}^* f$  bei  $\tau_{z_0}(x)$  differenzierbar ist. In diesem Fall gilt*

$$(\tau_{-z_0}^* f)'(\tau_{z_0}(x_0)) = f'(x_0). \quad (5.6.8)$$

**BEWEIS:** Der erste Teil folgt aus der Tatsache, dass die Translationen Homöomorphismen sind, siehe erneut Proposition 3.6.1. Die zweite Aussage ist die Kettenregel, da  $\tau_{-z_0}$  offensichtlich überall differenzierbar ist und eine ebenso differenzierbare Umkehrfunktion  $\tau_{z_0}$  besitzt.  $\square$

**Korollar 5.6.4** Sei  $a \in \mathbb{C}[[z]]$  eine Potenzreihe mit Konvergenzradius  $R > 0$  und  $z_0 \in \mathbb{C}$ .

i.) Auf  $B_R(z_0)$  konvergiert

$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n \quad (5.6.9)$$

absolut und lokal gleichmäßig in allen Ableitungen, also im  $\mathcal{C}^\infty$ -Sinne.

ii.) Es gilt

$$f^{(\ell)}(z) = \sum_{n=\ell}^{\infty} n(n-1)\cdots(n-\ell+1)a_n(z - z_0)^{n-\ell} \quad (5.6.10)$$

sowie

$$f(z) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(z_0)(z - z_0)^n. \quad (5.6.11)$$

BEWEIS: Wir können die Eigenschaften von

$$a(z) = \sum_{n=0}^{\infty} a_n z^n$$

aus Satz 5.6.1 mit Hilfe von Proposition 5.6.3 an die Stelle  $z_0$  transportieren, da  $f = \tau_{-z_0}^* a$ .  $\square$

Selbstverständlich bleiben die Aussagen von Satz 5.6.1 beziehungsweise Korollar 5.6.4 richtig, wenn wir die Potenzreihe nur auf den reellen Punkten  $B_R(0) \cap \mathbb{R}$  beziehungsweise  $B_R(z_0) \cap \mathbb{R}$  für  $z_0 \in \mathbb{R}$  betrachten. Wie bereits zuvor wird aber die wahre Natur der Potenzreihen und insbesondere ihr Konvergenzverhalten erst in der komplexen Ebene sichtbar. Es stellt sich nun also die Frage, ob nicht etwa jede  $\mathcal{C}^\infty$ -Funktion als Potenzreihe wie in (5.6.11) dargestellt werden kann. Dies ist nicht der Fall wie folgendes Beispiel zeigt:

**Beispiel 5.6.5** Wir betrachten die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & x > 0 \\ 0 & x \leq 0. \end{cases} \quad (5.6.12)$$

Eine einfache Induktion zeigt, dass es Polynome  $p_n \in \mathbb{R}[x]$  gibt, so dass die  $n$ -te Ableitung von  $f$  auf  $\mathbb{R}^+$  als

$$f^{(n)}(x) = p_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}} \quad (5.6.13)$$

gegeben ist. Man kann die ersten  $p_1, p_2, \dots$  direkt mit der Kettenregel bestimmen und für den allgemeinen Fall eine Rekursionsformel herleiten, siehe auch Übung 5.8.15. Mit den bekannten Grenzwerten aus Proposition 4.2.1, *iv.*), sieht man dann, dass

$$\lim_{x \searrow 0} f^{(n)}(x) = \lim_{x \searrow 0} p_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}} = \lim_{y \rightarrow \infty} p_n(y) e^{-y} = 0 \quad (5.6.14)$$

für alle  $n \in \mathbb{N}_0$  gilt. Da  $f^{(n)}(x) = 0$  für  $x < 0$ , folgt, dass die Funktion  $f$  glatt auf ganz  $\mathbb{R}$  ist. Da  $f^{(n)}(0) = 0$  für alle  $n \in \mathbb{N}_0$ , kann die Funktion  $f$  nicht als

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) x^n \quad (5.6.15)$$

dargestellt werden. Die linke Seite ist strikt positiv für  $x > 0$ , die rechte Seite konvergiert aus trivialen Gründen und ist Null für alle  $x \in \mathbb{R}$ . Diese Funktion  $f$  wird uns noch oft begegnen, siehe auch Abbildung 5.18.

Abbildung 5.18: Die Funktion  $f$  verschwindet bei  $x = 0$  mit allen Ableitungen, ist aber positiv für alle  $x > 0$ . Asymptotisch geht  $f(x)$  gegen 1 für  $x \rightarrow \infty$ .

Bei einer Verallgemeinerung von der Darstellung (5.6.11) ist also eine gewisse Vorsicht angebracht. Um diesen Fragen angemessen nachgehen zu können, definiert man die *formale Taylor-Reihe* einer glatten Funktion:

**Definition 5.6.6 (Formale Taylor-Reihe)** Sei  $X$  eine offene Teilmenge von  $\mathbb{R}$  oder  $\mathbb{C}$  und  $f: X \rightarrow \mathbb{K}$  eine glatte Funktion. Sei weiter  $x_0 \in X$ .

i.) Für  $n \in \mathbb{N}$  heißt das Polynom

$$T_{n,x_0}(f) = f(x_0) + f'(x_0)h + \cdots + \frac{1}{n!}f^{(n)}(x_0)h^n \in \mathbb{K}[h] \quad (5.6.16)$$

das  $n$ -te Taylor-Polynom von  $f$  an der Stelle  $x_0$ .

ii.) Die formale Potenzreihe  $T_{x_0}(f) \in \mathbb{K}[[h]]$  mit

$$T_{x_0}(f) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0) h^n \quad (5.6.17)$$

heißt formale Taylor-Reihe von  $f$  bei  $x_0$ .

Wir können nun mit den Taylor-Polynomen eine am Anfang von diesem Kapitel gestellte Frage erneut aufgreifen: Kann die Funktion  $f$  in der Nähe von  $x_0$  gut durch ein Polynom von der Ordnung  $n$  approximiert werden, so dass der Fehler höherer Ordnung als  $n$  ist, siehe (5.1.8) und (5.1.10). Das Taylor-Polynom einer  $n$ -mal differenzierbaren Funktion besitzt diese Eigenschaft. Hierzu bemerken wir, dass das Taylor-Polynom  $T_{n,x_0}(f) \in \mathbb{K}[[x]]$  bereits dann definiert werden kann, wenn  $f$  in einer offenen Umgebung von  $x_0$  zumindest  $n$ -mal differenzierbar ist.

Bevor wir dieses Resultat formulieren, sei hier angemerkt, dass für komplex differenzierbare Funktionen  $f$  auf einem offenen Definitionsbereich  $X$  sogar viel stärkere Resultate existieren. Eine einmal komplex differenzierbare Funktion ist automatisch  $\mathcal{C}^\infty$  auf  $X$  und lässt sich als Potenzreihe darstellen. Von diesem Standpunkt aus ist daher der folgende Satz im Komplexen eher uninteressant, da er deutlich hinter dem zurück bleibt, was man tatsächlich zeigen kann. Da dies aber in der Funktionentheorie geschieht, wollen wir nicht vorgreifen und nur die reelle Situation betrachten:

**Satz 5.6.7 (Taylor-Entwicklung mit Restglied)** Sei  $X \subseteq \mathbb{R}$  ein offenes Intervall und  $x_0 \in X$ . Sei weiter  $f: X \rightarrow \mathbb{K}$  eine  $k$ -mal differenzierbare Funktion mit  $k \in \mathbb{N} \cup \{\infty\}$ . Sei schließlich  $n \in \mathbb{N}$  mit  $n \leq k$  gegeben.

i.) Für das  $n$ -te Taylor-Restglied

$$R_{n,x_0}(f)(h) = f(x_0 + h) - T_{n,x_0}(f)(h) \quad (5.6.18)$$

von  $f$  bei  $x_0$  gilt

$$\lim_{h \rightarrow 0} \frac{R_{n,x_0}(f)(h)}{h^n} = 0. \quad (5.6.19)$$

ii.) Für alle  $h \in \mathbb{R}$  mit  $x_0 + h \in X$  gilt

$$|R_{n,x_0}(f)(h)| \leq \sup_{t \in [0,1]} |f^{(n)}(x_0 + th) - f^{(n)}(x_0)| \frac{|h|^n}{(n-1)!}. \quad (5.6.20)$$

iii.) Sei  $f$  zudem reell-wertig. Dann gibt es für alle  $h \in \mathbb{R} \setminus \{0\}$  mit  $x_0 + h \in X$  ein  $t \in (0, 1)$  derart, dass

$$R_{n-1,x_0}(f)(h) = \frac{f^{(n)}(x_0 + th)}{n!} h^n. \quad (5.6.21)$$

BEWEIS: Den ersten Teil zeigen wir durch wiederholte Anwendung von der Regel von l'Hospital nach Satz 5.3.14. Da  $\lim_{h \rightarrow 0} h^n = 0$  und auch  $\lim_{h \rightarrow 0} R_{n,x_0}(f)(h) = 0$  aus trivialen Gründen für die stetige Funktion  $f$  gilt, liegt also tatsächlich eine Situation vor, in der wir Satz 5.3.14 zur Anwendung bringen können. Das Problem ist, dass die Ableitung der Funktion im Nenner durch  $nh^{n-1}$  gegeben ist, und damit bei  $h = 0$  immer noch eine Nullstelle liefert, wenn  $n \geq 2$ . Für die Ableitung des Zählers erhalten wir

$$\begin{aligned} \frac{d}{dh} R_{n,x_0}(f)(h) &= \frac{d}{dh} \left( f(x_0 + h) - \sum_{m=0}^n \frac{1}{m!} f^{(m)}(x_0) h^m \right) \\ &= f'(x_0 + h) - \sum_{m=1}^n \frac{n}{m!} f^{(m)}(x_0) h^{m-1} \\ &= f'(x_0 + h) - \sum_{m=0}^{n-1} \frac{1}{m!} f^{(m+1)}(x_0) h^m \\ &= R_{n-1,x_0}(f')(h) \end{aligned}$$

und induktiv

$$\frac{d^\ell}{dh^\ell} R_{n,x_0}(f) = R_{n-\ell,x_0}(f^{(\ell)})$$

für alle  $1 \leq \ell \leq n$ , wobei  $R_{0,x_0}(f^{(n)})(h) = f^{(n)}(x_0 + h) - f^{(n)}(x_0)$ . Insbesondere sehen wir, dass die  $\ell$ -te Ableitung von  $R_{n,x_0}(f)$  für  $1 \leq \ell \leq n-1$  immer noch stetig ist und

$$\lim_{h \rightarrow 0} \frac{d^\ell}{dh^\ell} R_{n,x_0}(f)(h) = \lim_{h \rightarrow 0} R_{n-\ell,x_0}(f^{(\ell)})(h) = 0$$

erfüllt. Nach Satz 5.3.14 gilt also

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{R_{n,x_0}(f)(h)}{h^n} &= \lim_{h \rightarrow 0} \frac{R_{n-1,x_0}(f')(h)}{nh^{n-1}} \\ &\vdots \\ &= \lim_{h \rightarrow 0} \frac{R_{1,x_0}(f^{(n-1)})(h)}{n!h} \\ &= \lim_{h \rightarrow 0} \frac{1}{n!h} (f^{(n-1)}(x_0 + h) - f^{(n-1)}(x_0) - f^{(n)}(x_0)h) \\ &= 0, \end{aligned}$$

da  $f^{(n-1)}$  immer noch differenzierbar ist. Dies zeigt den ersten Teil. Der zweite Teil gibt nun eine Abschätzung für das Verhalten des Restglieds jenseits des Entwicklungspunkts  $x_0$ . Um diese Abschätzung und auch (5.6.21) zu zeigen, betrachten wir folgende Hilfsfunktion

$$F_\rho(y) = f(x_0 + h) - \sum_{m=0}^{n-1} \frac{f^{(m)}(y)}{m!} (x_0 + h - y)^m - \frac{\rho(x_0 + h - y)^n}{n!}$$

für einen Parameter  $\rho \in \mathbb{R}$ . Es gilt immer

$$F_\rho(x_0 + h) = 0.$$

Für die Ableitung von  $F_\rho$  erhalten wir

$$\begin{aligned}
 F'_\rho(y) &= - \sum_{m=0}^{n-1} \frac{f^{(m+1)}(y)}{m!} (x_0 + h - y)^m \\
 &\quad + \sum_{m=0}^{n-1} \frac{f^{(m)}(y)}{m!} m (x_0 + h - y)^{m-1} \\
 &\quad + \frac{\rho(x_0 + h - y)^{n-1}}{(n-1)!} \\
 &= - \frac{f^{(n)}(y)}{(n-1)!} (x_0 + h - y)^{n-1} + \frac{\rho(x_0 + h - y)^{n-1}}{(n-1)!}, \tag{*}
 \end{aligned}$$

da sich die übrigen Terme der beiden Summen paarweise herausheben. Wir spezifizieren den Wert von  $\rho$  nun auf zwei Weisen:

i.) Zuerst setzen wir  $\rho = f^{(n)}(x_0)$ . Dann gilt

$$F_\rho(x_0) = f(x_0 + h) - \sum_{m=0}^{n-1} \frac{f^{(m)}(x_0)}{m!} h^m - \frac{f^{(n)}(x_0)}{n!} h^n = R_{n,x_0}(f)(h).$$

Da  $F_\rho(x_0 + h) = 0$  gilt, können wir den Schrankensatz, siehe Satz 5.3.10, auf die Funktion  $F_\rho$  anwenden und erhalten somit

$$\begin{aligned}
 &|R_{n,x_0}(f)(h)| \\
 &= |F_\rho(x_0 + h) - F_\rho(x_0)| \\
 &\leq \sup_{t \in [0,1]} |F'_\rho(x_0 + th)| |h| \\
 &= \sup_{t \in [0,1]} \left| - \frac{f^{(n)}(x_0 + th)}{(n-1)!} (x_0 + h - x_0 - th)^{n-1} + \frac{f^{(n)}(x_0)}{(n-1)!} (x_0 + h - x_0 - th)^{n-1} \right| |h| \\
 &= \sup_{t \in [0,1]} |f^{(n)}(x_0 + th) - f^{(n)}(x_0)| (1-t)^{n-1} \frac{|h|^n}{(n-1)!} \\
 &\leq \sup_{t \in [0,1]} |f^{(n)}(x_0 + th) - f^{(n)}(x_0)| \frac{|h|^n}{(n-1)!},
 \end{aligned}$$

was (5.6.20) zeigt.

ii.) Da  $h \neq 0$  gilt, gibt es ein eindeutiges  $\rho \in \mathbb{R}$ , so dass

$$F_\rho(x_0) = 0$$

gilt. Wir wählen nun dieses  $\rho$  und finden wegen  $F_\rho(x_0 + h) = 0 = F_\rho(x_0)$  nach dem Satz von Rolle, siehe Satz 5.3.5, einen Zwischenpunkt  $x_0 + th$ , wobei  $t \in (0, 1)$ , so dass  $F'_\rho(x_0 + th) = 0$  gilt. Mit (\*) erlaubt es uns dieser Zwischenpunkt,  $\rho$  zu bestimmen. Es folgt

$$0 = - \frac{f^{(n)}(x_0 + th)}{(n-1)!} (x_0 + h - x_0 - th)^{n-1} + \frac{\rho(x_0 + h - x_0 - th)}{(n-1)!},$$

also

$$\rho = f^{(n)}(x_0 + th).$$

Setzen wir dies nun in die Bedingung  $F_\rho = 0$  ein, erhalten wir

$$0 = F_\rho(x_0)$$

$$\begin{aligned}
&= f(x_0 + h) - \sum_{m=0}^{n-1} \frac{f^{(m)}(x_0)}{m!} h^m - \frac{f^{(n)}(x_0 + th)}{n!} h^n \\
&= R_{n-1, x_0}(f)(h) - \frac{f^{(n)}(x_0 + th)}{n!} h^n,
\end{aligned}$$

was dann auch den dritten Teil zeigt.  $\square$

### Bemerkung 5.6.8 (Taylor-Entwicklung mit Restglied)

- i.) Die Abschätzung (5.6.20) liefert noch nicht das Resultat (5.6.19), da wir nicht die Stetigkeit der  $n$ -ten Ableitung vorausgesetzt haben. Ist diese aber gegeben, folgt (5.6.19) aus (5.6.20). Genauso ist in (5.6.21) der Zwischenwert nur für das vorherige Restglied zu finden.
- ii.) Die Form (5.6.21) heißt auch das *Lagrange-Restglied*. Die genaue Position des Zwischenwerts  $x_0 + th$  ist nicht bekannt. Von daher ist (5.6.21) vor allem für Abschätzungen des Restglieds interessant, wenn man etwas über die Größe der  $n$ -ten Ableitung sagen kann. Wir werden später noch weitere Formen des Restglieds finden.
- iii.) Bei einer  $n$ -mal differenzierbaren Funktion  $f$  können wir den Verlauf von  $f$  zumindest in der Nähe eines Punktes  $x_0$  gut durch ein Polynom  $n$ -ten Grades annähern, dem Taylor-Polynom  $T_{n, x_0}(f)$ , so dass der Fehler schneller als  $h^n$  gegen Null geht, wenn  $x = x_0 + h \rightarrow x_0$ . Damit löst das Taylor-Polynom also das in (5.1.8) beziehungsweise in (5.1.10) gestellte Problem. Allerdings darf man nicht zu viel von (5.6.19) erwarten: Der tatsächliche Fehler  $R_{n, x_0}(f)(h)$  kann für  $h \neq 0$  durchaus groß sein.

Die interessante Frage ist nun, ob wir die Approximation immer weiter verbessern können, so dass die formale Taylor-Reihe von  $f$ , im Falle, dass  $f$  glatt ist, sogar gegen  $f$  konvergiert. Dann hätten wir  $f$  als konvergente Potenzreihe geschrieben. Für jede  $\mathcal{C}^\infty$ -Funktion wird dies nicht möglich sein, wie uns das Beispiel 5.6.5 lehrt. Es gibt aber nützliche Kriterien, wann dieses Programm erfolgreich ist.

**Proposition 5.6.9** Sei  $f \in \mathcal{C}^\infty(X, \mathbb{K})$  eine glatte Funktion auf einem offenen Intervall  $X \subseteq \mathbb{R}$ . Sei weiter  $x_0 \in X$ .

- i.) Die formale Taylor-Reihe  $T_{x_0}(f)$  konvergiert für  $h \in \mathbb{R}$  genau dann gegen den Funktionswert  $f(x_0 + h)$ , wenn

$$\lim_{n \rightarrow \infty} R_{n, x_0}(f)(h) = 0 \quad (5.6.22)$$

gilt.

- ii.) Erfüllen die Ableitungen  $f^{(n)}$  von  $f$  die Abschätzungen

$$\|f^{(n)}\|_{B_r(x_0)^{\text{cl}}} \leq c \alpha^n \quad (5.6.23)$$

auf einem abgeschlossenen Ball  $B_r(x_0)^{\text{cl}} \subseteq X$ , so hat die formale Taylor-Reihe einen Konvergenzradius  $R \geq r$  und konvergiert auf  $B_r(x_0)^{\text{cl}}$  gegen  $f$ , also

$$f(x_0 + h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} h^n \quad (5.6.24)$$

für  $|h| \leq r$ .

BEWEIS: Die Konvergenz der formalen Taylor-Reihe gegen  $f$  bedeutet gerade

$$f(x_0 + h) = \lim_{n \rightarrow \infty} T_{n, x_0}(f)(h),$$

was nach Definition des Restglieds in (5.6.18) zu (5.6.22) äquivalent ist. Wir nehmen nun an, dass die Größe der  $n$ -ten Ableitung auf dem kompakten Ball  $B_r(x_0)^{\text{cl}}$  durch (5.6.23) beschränkt ist. Wichtig ist hier nicht, dass  $f^{(n)}$  auf  $B_r(x_0)^{\text{cl}}$  beschränkt ist. Dies ist für die stetige Ableitung  $f^{(n)}$  gemäß Korollar 3.5.6, *i.*), sowieso automatisch der Fall. Entscheidend ist die Abhängigkeit der Schranke von  $n$ . Mit (5.6.23) finden wir nach Satz 5.6.7, *iii.*), dann für  $h \in B_r(x_0)^{\text{cl}}$

$$|R_{n-1,x_0}(f)(h)| \leq \|f^{(n)}\|_{B_r(x_0)^{\text{cl}}} \frac{|h|^n}{n!} \leq c \frac{\alpha^n |h|^n}{n!},$$

da der Zwischenwert in  $B_r(x_0)^{\text{cl}}$  liegt und entsprechend  $|f^{(n)}(x_0 + th)|$  durch das Supremum über alle Punkte im Ball  $B_r(x_0)^{\text{cl}}$  abgeschätzt werden kann. Dank des dominierenden Wachstums der Fakultät  $n!$  gilt dann

$$\lim_{n \rightarrow \infty} |R_{n-1,x_0}(f)(h)| = 0,$$

was den zweiten Teil zeigt. □

Zur Illustration der Taylor-Entwicklung betrachten wir nun konvergente Potenzreihen. Hier reproduziert die Taylor-Entwicklung die Funktion aus trivialen Gründen, und das Restglied lässt sich ebenfalls leicht verstehen. Wir formulieren das folgende Resultat im Komplexen:

**Proposition 5.6.10** *Sei  $a \in \mathbb{C}[[z]]$  eine formale Potenzreihe mit Konvergenzradius  $R > 0$ . Sei*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad (5.6.25)$$

für  $z \in B_R(z_0)$ .

*i.) Das  $n$ -te Taylor-Polynom von  $f$  mit Entwicklungspunkt  $z_0$  ist*

$$T_{n,z_0}(f)(h) = a_n h^n + a_{n-1} h^{n-1} + \cdots + a_1 h + a_0. \quad (5.6.26)$$

*ii.) Das  $n$ -te Taylor-Restglied von  $f$  mit Entwicklungspunkt  $z_0$  ist*

$$R_{n,z_0}(f)(h) = \sum_{k=n+1}^{\infty} a_k h^k = h^{n+1} \sum_{k=0}^{\infty} a_{k+n+1} h^k. \quad (5.6.27)$$

*iii.) Die Reihe*

$$\rho_n(h) = \sum_{k=0}^{\infty} a_{k+n+1} h^k \quad (5.6.28)$$

*hat Konvergenzradius  $R$  und ist insbesondere selbst eine glatte Funktion auf  $B_R(0)$ .*

BEWEIS: Nach Satz 5.6.1 wissen wir  $f \in \mathcal{C}^\infty(B_R(z_0))$  und

$$f^{(n)}(z_0) = n! a_n.$$

Damit ist *i.*) klar und *ii.*) ebenso. Man beachte, dass die Reihe (5.6.27) tatsächlich konvergiert und wir für  $h \neq 0$  mit  $h^{-n-1}$  multiplizieren dürfen, ohne die Konvergenz zu beeinträchtigen. Dies zeigt dann (5.6.27) und (5.6.28). □

Für Potenzreihen stimmt die Taylor-Reihe also mit der Potenzreihe überein. Wir bestimmen nun einige (formale) Taylor-Reihen explizit und zeigen deren Konvergenz.

### Beispiel 5.6.11 (Taylor-Entwicklungen)



i.) Wir betrachten die Logarithmusfunktion und entwickeln diese um  $x_0 = 1$ . Dazu sei

$$f(x) = \log(1+x) \quad (5.6.29)$$

für  $x \in (-1, \infty)$  gegeben. Für die Ableitung erhalten wir

$$f'(x) = \frac{1}{1+x} \quad (5.6.30)$$

und induktiv daher

$$f^{(n)}(x) = \frac{(-1)^{n-1}}{(1+x)^n} (n-1)! \quad (5.6.31)$$

für  $n \in \mathbb{N}$ . Bei  $x = 0$ , was also dem Entwicklungspunkt  $x_0 = 1$  von  $\log(x)$  entspricht, gilt daher

$$f^{(n)}(0) = (-1)^{n-1} (n-1)!. \quad (5.6.32)$$

Damit ist die (formale) Taylor-Reihe von  $f$  durch

$$T_{x_0=0}(f)(h) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} h^n \quad (5.6.33)$$

gegeben, da der nullte Term wegen  $\log(1) = 0$  verschwindet. Um die Konvergenz zu studieren, betrachten wir  $x \in (-1, 1]$ . Dann gilt für das Restglied

$$R_{n-1, x_0=0}(f)(x) = \frac{f^{(n)}(tx)}{n!} x^n = \frac{(-1)^{n-1} x^n}{n(1+tx)^{n+1}} \quad (5.6.34)$$

mit einem  $t \in (0, 1)$  nach Satz 5.6.7, iii.). Ist  $0 \leq x \leq 1$  so gilt  $\frac{x^n}{(1+tx)^{n+1}} < 1$ , da ja  $t \in (0, 1)$ . Also folgt

$$|R_{n-1,0}(f)(x)| < \frac{1}{n} \longrightarrow 0, \quad (5.6.35)$$

was zeigt, dass die formale Taylor-Reihe gegen  $\log(1+x)$  konvergiert, wenn  $0 \leq x \leq 1$  gilt. Insbesondere erhalten wir

$$\log(2) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \quad (5.6.36)$$

als Grenzwert der alternierenden harmonischen Reihe. Deren Konvergenz hatten wir mit Hilfe des Leibniz-Kriteriums in Beispiel 2.6.10 gesehen, ihren Wert aber noch nicht bestimmen können. Man beachte jedoch, dass für eine praktische Approximation von  $\log(2)$  die Reihendarstellung (5.6.36) ungeeignet ist, da die Konvergenz sehr langsam ist. Für  $x \in (-1, 0)$  ist die Abschätzung des Restglieds etwas komplizierter. Einen direkteren Weg erhalten wir so: Zunächst ist klar, dass die formale Potenzreihe (5.6.33) einen Konvergenzradius  $R \geq 1$  besitzt, da sie ja zumindest für  $h = 1$  konvergiert. Man kann sich schnell überlegen, dass tatsächlich  $R = 1$  der Konvergenzradius ist, siehe auch Übung 5.8.16. Insbesondere konvergiert

$$a(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n \quad (5.6.37)$$

für alle (sogar komplexen)  $x$  mit  $|x| < 1$ . Nach Satz 5.6.1 ist  $a$  eine glatte Funktion mit Ableitung

$$a'(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} n x^{n-1} = \sum_{n=0}^{\infty} (-x)^n = \frac{1}{1+x}. \quad (5.6.38)$$

Es gilt also für alle  $x \in (-1, 1)$  die Gleichheit

$$f'(x) = a'(x) \quad (5.6.39)$$

und damit  $f(x) = a(x) + c$  für eine Konstante  $c \in \mathbb{R}$  nach Korollar 5.3.7. Da  $a(0) = 0 = f(0)$ , folgt  $c = 0$ , also  $a(x) = \log(1+x)$  für alle  $x \in (-1, 1)$ . Dies zeigt daher etwas indirekt, dass

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n \quad (5.6.40)$$

für alle  $x \in (-1, 1]$  konvergiert.

ii.) Das zweite Beispiel ist der verallgemeinerte Binomialsatz. Wieder für  $x \in (-1, 1]$  betrachten wir die Funktion

$$f(x) = (1+x)^\alpha \quad \text{mit} \quad \alpha \in \mathbb{R}. \quad (5.6.41)$$

Da  $1+x > 0$ , ist diese Potenz definiert und  $f$  ist glatt. Wir bestimmen die Ableitungen als

$$\begin{aligned} f'(x) &= \alpha(1+x)^{\alpha-1}, \\ f''(x) &= \alpha(\alpha-1)(1+x)^{\alpha-2} \end{aligned}$$

und allgemein

$$f^{(n)}(x) = \alpha(\alpha-1) \cdots (\alpha-n+1)(1+x)^{\alpha-n} \quad (5.6.42)$$

nach Proposition 5.2.8. Es können nun zwei Fälle auftreten. Entweder ist  $\alpha = n \in \mathbb{N}_0$ , dann gilt  $f^{(n+1)}(x) = 0$  und wir erhalten die endliche Taylor-Reihe

$$(1+x)^n = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(0) x^k = \sum_{k=0}^n \binom{n}{k} x^k, \quad (5.6.43)$$

was nichts anderes als der Binomialsatz ist. Oder aber es gilt  $\alpha \notin \mathbb{N}_0$ , dann ist  $f^{(n)}(0) \neq 0$  für alle  $n \in \mathbb{N}_0$ , was zu einer interessanten (formalen) Taylor-Reihe

$$T_{x_0=0}(f)(x) = \sum_{n=0}^{\infty} \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n \quad (5.6.44)$$

führt. Als Abkürzung definiert man die *verallgemeinerten Binomialkoeffizienten*

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} \quad (5.6.45)$$

und kann die (formale) Taylor-Reihe daher als

$$T_{x_0=0}(f)(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \quad (5.6.46)$$

schreiben. Wieder stellt sich die Frage, ob diese sogar konvergent ist und  $f(x)$  reproduziert. Wir zeigen zunächst die Konvergenz nach dem Quotientenkriterium, da

$$\left| \frac{\binom{\alpha}{n+1} x^{n+1}}{\binom{\alpha}{n} x^n} \right| = \frac{\frac{|\alpha(\alpha-1) \cdots (\alpha-n+1)(\alpha-n)|}{(n+1)n!}}{\frac{|\alpha(\alpha-1) \cdots (\alpha-n+1)|}{n!}} |x| = \left| \frac{\alpha-n}{n+1} \right| |x| \quad (5.6.47)$$

und

$$\lim_{n \rightarrow \infty} \left| \frac{\alpha-n}{n+1} \right| |x| = |x| < 1 \quad (5.6.48)$$

unabhängig vom tatsächlichen Wert von  $\alpha$ . Damit ist der Konvergenzradius von (5.6.44) also  $R \geq 1$ . Wieder kann man zeigen, dass sogar  $R = 1$  gilt, siehe auch Übung 5.8.17. Wir wollen als nächstes zeigen, dass (5.6.44) tatsächlich auf  $(-1, 1)$  gegen  $f$  konvergiert. Für  $x \geq 0$  ist dies wieder einfach mit Satz 5.6.7, *iii.*), denn wir finden ein  $t \in (0, 1)$  mit

$$|R_{n-1, x_0=0}(f)(x)| = \left| \binom{\alpha}{n} (1+tx)^{\alpha-n} x^n \right| \leq \left| \binom{\alpha}{n} x^n \right| \quad (5.6.49)$$

für  $n \geq |\alpha|$ , da dann  $(1+tx)^{\alpha-n} \leq 1$  gilt. Es folgt  $\lim_{n \rightarrow \infty} |R_{n-1, x_0=0}(f)(x)| = 0$  für  $0 \leq x \leq 1$ , da wir zuvor in (5.6.48) gesehen haben, dass  $\left(\binom{\alpha}{n} x^n\right)_{n \in \mathbb{N}}$  eine Nullfolge ist. Es gilt daher

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \quad (5.6.50)$$

für  $0 \leq x \leq 1$ . Für  $-1 < x < 0$  ist das Argument wieder komplizierter, da der Zwischenpunkt auf  $1+tx < 1$  führt und daher  $(1+tx)^{\alpha-n}$  nicht so leicht abgeschätzt werden kann. Wie schon beim Logarithmus bestimmen wir die Ableitung. Da (5.6.44) für alle  $x$  mit  $|x| < 1$  konvergiert, erhalten wir eine glatte Funktion

$$a(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \quad (5.6.51)$$

für  $x \in (-1, 1)$ , deren Ableitung wir nach Satz 5.6.1 bestimmen können. Es folgt

$$a'(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} n x^{n-1} = \sum_{n=0}^{\infty} \alpha \binom{\alpha-1}{n} x^n. \quad (5.6.52)$$

Wie schon bei den normalen Binomialkoeffizienten gilt die Identität

$$\binom{\alpha-1}{n} + \binom{\alpha-1}{n-1} = \binom{\alpha}{n}, \quad (5.6.53)$$

siehe nochmals Übung 5.8.17. Damit folgt aber

$$\begin{aligned} (1+x)a'(x) &= \sum_{n=0}^{\infty} \alpha \binom{\alpha-1}{n} x^n + \sum_{n=0}^{\infty} \alpha \binom{\alpha-1}{n} x^{n+1} \\ &= \alpha + \sum_{n=1}^{\infty} \left( \alpha \binom{\alpha-1}{n} + \alpha \binom{\alpha-1}{n-1} \right) x^n \\ &= \alpha + \sum_{n=1}^{\infty} \alpha \binom{\alpha}{n} x^n \\ &= \alpha a(x). \end{aligned}$$

Da umgekehrt  $f'(x) = \alpha(1+x)^{\alpha-1}$  gilt, folgt auch

$$(1+x)f'(x) = \alpha f(x).$$

Für  $x \in (-1, 0)$  ist  $(1+x)^\alpha$  invertierbar, womit

$$\begin{aligned} \left( \frac{a(x)}{(1+x)^\alpha} \right)' &= \frac{a'(x)(1+x)^\alpha - a(x)\alpha(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} \\ &= \frac{a'(x)(1+x)^\alpha - (1+x)a'(x)(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} \end{aligned}$$

$$= 0.$$

Damit ist nach Korollar 5.3.7 die Funktion  $\frac{a(x)}{(1+x)^\alpha}$  konstant und gleich  $\frac{a(0)}{1} = 1$ . Dies zeigt schließlich  $a(x) = (1+x)^\alpha$  auch für  $x \in (-1, 0)$ , womit

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \quad (5.6.54)$$

für alle  $x \in (-1, 1)$  folgt. Für  $\alpha = -1$  findet man so insbesondere die geometrische Reihe wieder. Weitere interessante Spezialfälle und Anwendungen werden in den Übungen diskutiert, siehe Übung ??.

## 5.7 Asymptotische Entwicklungen

Nicht jede glatte Funktion lässt sich in eine Potenzreihe entwickeln. Trotzdem ist es interessant, dass Verhalten einer (glatten) Funktion  $f$  mit dem eines Polynoms zu vergleichen. Die Taylor-Formel mit Restglied leistet genau das unter der Voraussetzung, dass  $f$  entsprechende Differenzierbarkeitseigenschaften auf einem offenen Intervall um den betrachteten Punkt  $x_0$  besitzt. In vielen Situationen ist  $f$  jedoch nicht stetig differenzierbar auf einem offenen Intervall um  $x_0$  und  $x_0$  ist eventuell nur ein Häufungspunkt des Definitionsbereichs ohne im Definitionsbereich zu sein. In diesen Fällen ist man auf folgende Definition einer asymptotischen Entwicklung angewiesen, die wir hier nur für reelle Definitionsbereiche ansprechen:

**Definition 5.7.1 (Asymptotische Entwicklung)** Sei  $f: I \rightarrow \mathbb{K}$  eine Funktion auf einem offenen Intervall  $I = (0, b) \subseteq \mathbb{R}$ .

- i.) Ein Polynom  $p_n \in \mathbb{K}[x]$  der Ordnung  $n \in \mathbb{N}_0$  heißt asymptotische Entwicklung der Ordnung  $n$  von  $f$  bei  $x = 0$  von rechts, falls

$$\lim_{x \searrow 0} \frac{f(x) - p_n(x)}{x^n} = 0 \quad (5.7.1)$$

gilt.

- ii.) Eine formale Potenzreihe  $a = \sum_{n=0}^{\infty} a_n x^n \in \mathbb{K}[[x]]$  heißt asymptotische Potenzreihenentwicklung von  $f$  bei  $x = 0$  von rechts, falls für alle  $n \in \mathbb{N}_0$

$$\lim_{x \searrow 0} \frac{f(x) - a_n x^n - a_{n-1} x^{n-1} - \dots - a_1 x - a_0}{x^n} = 0 \quad (5.7.2)$$

gilt. In diesem Fall schreiben wir

$$f(x) \underset{x \searrow 0}{\sim} \sum_{n=0}^{\infty} a_n x^n. \quad (5.7.3)$$

Analog definiert man asymptotische Entwicklungen von links.

Eine asymptotische Entwicklung der Ordnung Null ist also die Angabe einer Zahl  $a_0 \in \mathbb{K}$ , so dass

$$\lim_{x \searrow 0} (f(x) - a_0) = 0, \quad (5.7.4)$$

was nichts anderes als

$$a_0 = \lim_{x \searrow 0} f(x) \quad (5.7.5)$$

bedeutet. Damit gibt es also genau dann eine asymptotische Entwicklung von rechts von  $f$  der Ordnung Null, wenn der Grenzwert der Funktionswerte von  $f$  für  $x \searrow 0$  von rechts existiert. Dieser Grenzwert ist dann auch das gesuchte Polynom  $p_0(x) = a_0$  der Ordnung Null. Insbesondere gibt es nur ein solches Polynom  $p_0$ , welches dieser Anforderung genügt:  $a_0$  ist eindeutig durch (5.7.5) bestimmt. Dieses Verhalten setzt sich nun auch für asymptotische Entwicklungen höherer Ordnung fort:

**Proposition 5.7.2** *Sei  $f: I \rightarrow \mathbb{K}$  eine Funktion auf einem offenen Intervall  $I = (0, b)$ .*

*i.) Sei  $p_n \in \mathbb{K}[[x]]$  eine asymptotische Entwicklung von  $f$  bei  $x = 0$  von rechts der Ordnung  $n \in \mathbb{N}$ . Dann ist*

$$p_{n-1}(x) = a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \quad (5.7.6)$$

*eine asymptotische Entwicklung der Ordnung  $n - 1$ .*

*ii.) Wenn  $f$  eine asymptotische Entwicklung bei  $x = 0$  von rechts der Ordnung  $n$  besitzt, so ist diese eindeutig bestimmt.*

*iii.) Wenn  $f$  eine asymptotische Potenzreihenentwicklung bei  $x = 0$  von rechts besitzt, so ist diese eindeutig.*

Analoge Aussagen gelten auch für asymptotische Entwicklungen von links.

BEWEIS: Sei  $p_n(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$  eine asymptotische Entwicklung von  $f$  bei  $x = 0$  oder Ordnung  $n$  von rechts. Es gilt also

$$\begin{aligned} 0 &= \lim_{x \searrow 0} \frac{f(x) - a_nx^n - a_{n-1}x^{n-1} - \cdots - a_1x - a_0}{x^n} \\ &= \lim_{x \searrow 0} \frac{f(x) - a_{n-1}x^{n-1} - \cdots - a_1x - a_0}{x^n} - a_n, \end{aligned}$$

und damit existiert der Grenzwert

$$\lim_{x \searrow 0} \frac{f(x) - a_{n-1}x^{n-1} - \cdots - a_1x - a_0}{x^n} = a_n. \quad (*)$$

Mit  $\lim_{x \searrow 0} x = 0$  folgt aus den Rechenregeln zu Grenzwerten von Funktionen

$$\begin{aligned} &\lim_{x \searrow 0} \frac{f(x) - a_{n-1}x^{n-1} - \cdots - a_1x - a_0}{x^{n-1}} \\ &= \lim_{x \searrow 0} x \cdot \frac{f(x) - a_{n-1}x^{n-1} - \cdots - a_1x - a_0}{x^n} \\ &= \lim_{x \searrow 0} x \lim_{x \searrow 0} \frac{f(x) - a_{n-1}x^{n-1} - \cdots - a_1x - a_0}{x^n} \\ &= 0 \cdot a_n \\ &= 0, \end{aligned}$$

wobei wir entscheidend benutzen, dass der Grenzwert  $(*)$  existiert. Dies zeigt den ersten Teil. Die Eindeutigkeit folgt durch Induktion nach  $n$ . Für  $n = 0$  haben wir  $a_0$  bereits in (5.7.5) eindeutig bestimmt. Seien daher asymptotische Entwicklungen Ordnung  $n - 1$  eindeutig und  $p_n$  eine solche der Ordnung  $n$ . Nach  $(*)$  ist der führende Term  $a_n$  ebenfalls eindeutig und  $p_{n-1}(x) = p_n(x) - a_nx^n$  ist eine asymptotische Entwicklung der Ordnung  $n - 1$  nach *i.)*, also eindeutig nach Induktion. Damit ist  $p_n$  auch eindeutig. Dies zeigt dann auch den dritten Teil.  $\square$

Wir können daher im Folgenden immer von *der* asymptotischen Entwicklung sprechen, sofern diese überhaupt existiert.

Abbildung 5.19: Die lieblos geklebte Funktion  $f$  hat bei  $x = 0$  sowohl eine asymptotische Entwicklung von links als auch von rechts, die nicht übereinstimmen müssen.

**Bemerkung 5.7.3** Dass wir als Entwicklungspunkt  $x = 0$  gewählt haben, ist nur eine nützliche Konvention. Ist man an der asymptotischen Entwicklung einer Funktion  $f: (a, b) \rightarrow \mathbb{K}$  für  $x \searrow a$  interessiert, kann man stattdessen die Funktion  $\tau_{-a}^* f = f \circ \tau_{-a}$  mit der Translation  $\tau_{-a}(x) = x - a$  um  $x = 0$  entwickeln, siehe auch Übung ??.

Die folgende Proposition gibt nun eine große Klasse von Funktionen, die eine asymptotische Entwicklung besitzen:

**Proposition 5.7.4** Sei  $f: (a, b) \rightarrow \mathbb{K}$  eine  $k$ -mal differenzierbare Funktion für  $k \in \mathbb{N} \cup \{\infty\}$  auf einem offenen Intervall mit  $0 \in (a, b)$ . Dann besitzt  $f$  eine asymptotische Entwicklung bei  $x = 0$  der Ordnung  $k$ , sowohl von links als auch von rechts. Diese ist durch das  $k$ -te Taylor-Polynom

$$T_{k,x_0}(f) = \sum_{i=0}^k \frac{f^{(i)}(0)}{i!} x^i \quad (5.7.7)$$

gegeben.

BEWEIS: Dass  $T_{k,x_0=0}(f)$  die definierende Eigenschaft

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

einer asymptotischen Entwicklung erfüllt, ist gerade die Aussage von Satz 5.6.7, i.).  $\square$

Interessanterweise sehen wir mit Proposition 5.7.2, iii.), dass das Taylor-Polynom das *einzige* Polynom ist, welches die Grenzwerteigenschaft aus der Taylor-Entwicklung in Satz 5.6.7, i.), erfüllt. Zudem ist in diesem Fall die asymptotische Entwicklung von links gleich der von rechts. Dies ist keineswegs selbstverständlich, wie folgende Beispiele zeigen:

### Beispiel 5.7.5 (Asymptotische Entwicklungen)

i.) Sei  $g: (0, b) \rightarrow \mathbb{K}$  eine Funktion mit einer asymptotischen Entwicklung  $p_n$  bei  $x = 0$  von rechts der Ordnung  $n \in \mathbb{N}_0 \cup \{\infty\}$ , und sei  $h: (a, 0) \rightarrow \mathbb{K}$  eine Funktion mit asymptotischer Entwicklung  $q_n$  bei  $x_0$  von links der gleichen Ordnung  $n \in \mathbb{N}_0 \cup \{\infty\}$ . Dann definieren wir die Funktion

$$f(x) = \begin{cases} g(x) & x > 0 \\ 0 & x = 0 \\ h(x) & x < 0, \end{cases} \quad (5.7.8)$$

siehe auch Abbildung 5.19. Da die Definition der asymptotischen Entwicklung von links beziehungsweise von rechts nur Funktionswerte  $f(x)$  von  $f$  für  $x < 0$  beziehungsweise  $x > 0$  verwendet, folgt, dass  $f$  eine asymptotische Entwicklung bei  $x = 0$  von links besitzt, nämlich die von  $h$  also  $q_n$ . Ebenso besitzt  $f$  eine asymptotische Entwicklung von rechts, nämlich die von  $g$ , also  $p_n$ . Insbesondere mit Proposition 5.7.4 lassen sich leicht Beispiele konstruieren, für die

$$p_n \neq q_n \quad (5.7.9)$$

gilt.

- ii.) Ein deutlich extremeres Beispiel zeigt, dass das Verhalten von  $f$  für Punkte außerhalb  $x = 0$  sehr singulär sein kann. Wir betrachten die Funktion

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & x \in \mathbb{Q} \setminus \{0\} \\ 0 & \text{sonst.} \end{cases} \quad (5.7.10)$$

als kleine Variante der Funktion aus Beispiel 5.6.5. Die Funktion ist bei  $x = 0$  stetig aber bei allen anderen Punkten unstetig. Weiter hat  $f$  bei  $x = 0$  die zweiseitige asymptotische Potenzreihenentwicklung

$$f(x) \underset{x \rightarrow 0}{\sim} 0. \quad (5.7.11)$$

Dies zeigt man wie die Glattheit der Funktion aus Beispiel 5.6.5, wobei wir für irrationale Punkte direkt  $f(x) = 0$  verwenden.

**Bemerkung 5.7.6 (Glattheit an einem Punkt)** Gerade das letzte Beispiel zeigt, dass die Anforderung, eine asymptotische Entwicklung zu besitzen, sehr mild sind. Insbesondere muss die Funktion keineswegs glatt sein: Unser bisheriger Begriff von Glattheit erfordert die Existenz von Ableitungen auf einem offenen Intervall, um überhaupt über die nächste Ableitung sprechen zu können. Dies ist bei einer asymptotischen Entwicklung nicht erforderlich und, nach Beispiel 5.7.5, *ii.*), im Allgemeinen auch nicht gegeben. Im Hinblick auf Proposition 5.7.4 kann man die Existenz einer asymptotischen Entwicklung bei  $x = 0$  der Ordnung  $k$  als  $k$ -fache Differenzierbarkeit *am Punkt*  $x = 0$  interpretieren, auch wenn die Funktion außerhalb dieses Punktes nicht einmal stetig sein muss. Wie zuvor liefert eine Translation eine entsprechende Begriffsbildung bei anderen Punkten.

Das nächste Resultat klärt die algebraischen Eigenschaften von asymptotischen Entwicklungen. Wir formulieren sie direkt für asymptotische Potenzreihenentwicklungen, endliche Ordnungsergebnisse lassen sich ähnlich erzielen, siehe Übung ??.

**Proposition 5.7.7** Sei  $I = (0, c)$  ein offenes reelles Intervall, und seien  $f, g: I \rightarrow \mathbb{K}$  zwei Funktionen, die beide asymptotische Potenzreihenentwicklungen  $f \underset{x \searrow 0}{\sim} a$  und  $g \underset{x \searrow 0}{\sim} b$  mit  $a, b \in \mathbb{K}[[x]]$  bei  $x = 0$  von rechts besitzen.

- i.) Für alle  $z, w \in \mathbb{K}$  hat die Funktion  $zf + wg$  die asymptotische Entwicklung

$$zf + wg \underset{x \searrow 0}{\sim} za + wb. \quad (5.7.12)$$

- ii.) Das Produkt  $fg$  hat die asymptotische Entwicklung

$$fg \underset{x \searrow 0}{\sim} ab. \quad (5.7.13)$$

BEWEIS: Beide Punkte beinhalten zwei Aussagen, nämlich dass  $zf + wg$  und  $fg$  zum einen überhaupt asymptotische Entwicklungen besitzen und zum anderen, dass diese durch (5.7.12) beziehungsweise (5.7.13) gegeben ist. Um (5.7.12) zu zeigen, betrachten wir für  $n \in \mathbb{N}_0$

$$\begin{aligned} 0 &= z \lim_{x \searrow 0} \frac{f(x) - a_n x^n - \cdots - a_1 x - a_0}{x^n} + w \lim_{x \searrow 0} \frac{g(x) - b_n x^n - \cdots - b_1 x - b_0}{x^n} \\ &= \lim_{x \searrow 0} \frac{zf(x) + wg(x) - z(a_n x^n + \cdots + a_1 x - a_0) - w(b_n x^n + \cdots + b_1 x + b_0)}{x^n} \\ &= \lim_{x \searrow 0} \frac{(zf + wg)(x) - (za_n + wb_n)x^n - \cdots - (za_1 + wb_1)x - (za_0 + wb_0)}{x^n}, \end{aligned}$$

wobei wir verwenden, dass  $f$  und  $g$  asymptotische Entwicklungen besitzen. Dies zeigt den ersten Teil. Für den zweiten Teil wählen wir wieder ein  $n \in \mathbb{N}_0$  und definieren

$$A_n(x) = a_n x^n + \cdots + a_1 x + a_0$$

sowie

$$B_n(x) = b_n x^n + \cdots + b_1 x + b_0$$

und

$$C_n(x) = A_n(x)B_n(x) - \sum_{i=0}^n \sum_{j=0}^i a_j b_{i-j} x^i.$$

Dann gilt  $\lim_{x \rightarrow 0} \frac{1}{x^n} C_n(x) = 0$ , da in  $C_n(x)$  nur Potenzen  $i \geq n+1$  von  $x$  auftreten. Wir erhalten für  $x > 0$

$$\begin{aligned} & \frac{1}{x^n} \left( f(x)g(x) - \sum_{i=0}^n \sum_{j=0}^i a_j b_{i-j} x^i \right) \\ &= \frac{1}{x^n} (f(x)g(x) - A_n(x)B_n(x) + C_n(x)) \\ &= \frac{1}{x^n} ((f(x) - A_n(x))g(x) + A_n(x)(g(x) - B_n(x)) + C_n(x)) \\ &= \frac{f(x) - A_n(x)}{x^n} g(x) + A_n(x) \frac{g(x) - B_n(x)}{x^n} + \frac{C_n(x)}{x^n}. \end{aligned}$$

Für  $x \searrow 0$  erhalten wir  $g(x) \rightarrow b_0$ , womit der erste Summand wegen  $\lim_{x \searrow 0} \frac{f(x) - A_n(x)}{x^n} = 0$  gegen Null geht. Aus dem selben Grund liefert auch der zweite Summand im Grenzwert  $x \searrow 0$  Null, da wieder  $A_n(x) \rightarrow a_0$  stetig ist. Der dritte Summand liefert direkt Null für  $x \searrow 0$ . Insgesamt zeigt dies

$$\lim_{x \searrow 0} \frac{1}{x^n} \left( f(x)g(x) - \sum_{i=0}^n \sum_{j=0}^i a_j b_{i-j} x^i \right) = 0,$$

was ii.) zeigt. □

**Korollar 5.7.8** Die Funktionen auf  $(0, b)$ , welche eine asymptotische Potenzreihenentwicklung für  $x \searrow 0$  besitzen, bilden eine Unterálgebra von  $\text{Abb}((0, b), \mathbb{K})$ . Die Abbildung, die einer solchen Funktion ihre asymptotische Potenzreihenentwicklung zuordnet, ist ein einserhaltender Algebramorphismus in die Algebra  $\mathbb{K}[[x]]$  der formalen Potenzreihen.

**Bemerkung 5.7.9** Mit asymptotischen Entwicklungen kann man Aussagen wie „in erster Näherung“ oder „bis auf quadratische Korrekturen“ und dergleichen einen mathematischen Sinn geben: Eine komplizierte Funktion wird durch ihre asymptotische Entwicklung ersetzt, meist sogar nur bis zu einer kleinen Ordnung. Um dies als einen Mechanismus zur Näherung zu verstehen, ist es wichtig, sich die Stellung der Quantoren abermals klar zu machen: Die Existenz einer asymptotischen Potenzreihenentwicklung  $f(x) \stackrel{x \searrow 0}{\sim} a$  bedeutet, dass es zu jedem  $\epsilon > 0$  und zu jedem  $n \in \mathbb{N}$  ein  $\delta_n > 0$  gibt, so dass

$$\frac{1}{x^n} |f(x) - a_n x^n - \cdots - a_1 x - a_0| < \epsilon \quad (5.7.14)$$

für alle  $x \in (0, \delta_n)$  gilt. Insbesondere ist  $\delta_n$  nicht nur von  $\epsilon$  sondern eben auch von  $n$  abhängig. In vielen interessanten Situationen lässt sich daher kein uniformes  $\delta > 0$  unabhängig von  $n$  mit (5.7.14)



finden. Gilt nämlich (5.7.14) für  $x \in (0, \delta)$  mit einem  $n$ -unabhängigen  $\delta > 0$ , so bedeutet dies die Konvergenz der Potenzreihe

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad (5.7.15)$$

für alle  $x \in (0, \delta)$ . Hier haben wir aber viele Beispiele gesehen, wo (5.7.15) sicherlich falsch ist. Realistischer ist daher,

$$\inf_{n \in \mathbb{N}_0} \delta_n = 0 \quad (5.7.16)$$

zu erwarten.

## 5.8 Übungen

**Übung 5.8.1 (Leibniz-Regel)** Seien  $f, g: X \rightarrow \mathbb{K}$  zwei  $k$ -mal differenzierbare Funktionen mit  $k \in \mathbb{N} \cup \{\infty\}$ , und sei  $n \in \mathbb{N}$  mit  $n \leq k$  gegeben. Zeigen Sie die Leibniz-Regel

$$(fg)^{(n)} = \sum_{m=0}^n \binom{n}{m} f^{(m)} g^{(n-m)} \quad (5.8.1)$$

für die  $n$ -te Ableitung.

Hinweis: Induktion nach  $n$ .

**Übung 5.8.2 (Charakterisierung von Polynomen)** Sei  $f \in \mathcal{C}^\infty(\mathbb{K}, \mathbb{K})$  eine glatte Funktion. Zeigen Sie, dass  $f$  genau dann ein Polynom der Ordnung höchstens  $k \in \mathbb{N}$  ist, wenn  $f^{(k+1)} = 0$ .

**Übung 5.8.3 (Glattheit von  $\log$ )**

**Übung 5.8.4 (Glattheit von  $\tan$  und  $\cot$ )**

**Übung 5.8.5 (Glattheit der Potenzfunktionen)**

**Übung 5.8.6 (Glattheit der Arcusfunktionen)**

**Übung 5.8.7 (Glattheit der Hyperbelfunktionen)**

**Übung 5.8.8 (Kettenregel für höhere Ableitungen)**

**Übung 5.8.9 (Maxima, Minima und Wendepunkte)**

**Übung 5.8.10 (Konvexe Funktionen)**

**Übung 5.8.11 (Verkettung konvexer und monotoner Funktionen)**

**Übung 5.8.12 (Wieso  $\frac{1}{p} + \frac{1}{q} = 1$ ?)**

**Übung 5.8.13 (Youngsche Ungleichung für  $p = q = 2$ )** Geben Sie einen direkten und elementaren Beweis der Youngschen Ungleichung für den Fall  $p = q = 2$ .

**Übung 5.8.14 (Bessel-Funktionen)**

**Übung 5.8.15 (Eine glatte Funktion)**

**Übung 5.8.16 (Konvergenzradius der Taylor-Reihe von  $\log(1+x)$ )** Zeigen Sie, dass der Konvergenzradius der Taylor-Reihe von  $\log(1+x)$  um  $x_0 = 0$  genau  $R = 1$  ist.

**Übung 5.8.17 (Verallgemeinerte Binomialkoeffizienten)**



## Kapitel 6

# Das Riemann-Integral

Integration stellt in gewisser Hinsicht eine inverse Operation zur Differentiation dar. Ziel dieses Kapitels ist es daher, einen ersten Eindruck zu Integrationstheorie zu geben. Auch wenn mit dem Lebesgue-Integral und der allgemeinen Maßtheorie eine in vielen Aspekten leistungsfähigere Integration vorhanden ist, bietet das Riemann-Integral nicht zuletzt seiner Einfachheit wegen viele Vorteile, die weit über ein rein historisches Interesse hinausgehen.

Nach seiner geometrischen Motivation und der Definition werden wir uns den Eigenschaften Riemann-integrierbarer Funktionen im Detail zuwenden. Ein zentrales Resultat wird die Riemann-Integrierbarkeit stetiger Funktionen sein, welches im Hauptsatz der Differential- und Integralrechnung eine entscheidende Rolle spielt. Zudem werden wir verschiedene fundamentale Eigenschaften des Riemann-Integrals, wie den Mittelwertsatz, die Substitutionsregel und die partielle Integration, kennenlernen, die es in vielen Beispielen erlauben, Integrale konkret zu berechnen oder zumindest abzuschätzen. Uneigentliche Riemann-Integrale sowie die Beziehungen zur gleichmäßigen Konvergenz schließen dieses Kapitel dann ab.

### 6.1 Definition des Riemann-Integrals

Eine der gängigen Motivationen zur Definition des Riemann-Integrals ist es, einen Flächeninhalt unter dem Graphen einer hinreichend „gutartigen“ Funktion  $f: [a, b] \rightarrow \mathbb{R}$  definieren zu wollen. Hierzu kann man das Intervall  $[a, b]$  in kleine Teilintervalle zerlegen, also  $N + 1$  Punkte

$$a = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = b \quad (6.1.1)$$

wählen, was zu  $N$  Teilintervallen

$$I_1 = [t_0, t_1], \dots, I_N = [t_{N-1}, t_N] \quad (6.1.2)$$

mit

$$I = I_1 \cup \cdots \cup I_N \quad (6.1.3)$$

führt. Derartige Teilintervalle  $I_1, \dots, I_N$  bilden nun eine *Zerlegung* oder auch *Partition* des ursprünglichen Intervalls  $I$ . Wichtig ist hier, dass die Intervallgrenzen die Eigenschaft (6.1.1) erfüllen, was gleichbedeutend damit ist, dass die Durchschnitte

$$I_i \cap I_{i+1} = \{t_i\} \quad (6.1.4)$$

nur aus einem Punkt bestehen.

Im zweiten Schritt wählt man in jedem Teilintervall einen Punkt

$$\xi_i \in I_i \quad (6.1.5)$$

Abbildung 6.1: Eine Riemann-Summe zur Funktion  $f$  approximiert den Flächeninhalt unter dem Graphen von  $f$ .

für  $i = 1, \dots, N$  aus. Die elementar-geometrische Vorstellung legt dann nahe, die Flächeninhalte der Rechtecke der Breite  $t_i - t_{i-1}$  und Höhe  $f(\xi_i)$  zu betrachten und deren Summe als Approximation an den Flächeninhalt unter dem Graphen von  $f$  anzusehen, siehe Abbildung 6.1. Die Definition der Fläche eines Rechtecks wird elementar-geometrisch vorgenommen. Die Idee ist nun, die Teilintervalle immer kleiner zu wählen, und so eine immer bessere Approximation an den Flächeninhalt zu erreichen. Die konzeptuelle Schwierigkeit dabei ist, dass wir noch nicht über eine mathematisch tragfähige Definition des Flächeninhaltes, den wir approximieren wollen, verfügen, außer eben für den naive definierten Flächeninhalt von Rechtecken. Deshalb werden wir den obigen Approximationsgedanken schlichtweg dazu verwenden, eine *Definition* des Flächeninhaltes als Grenzwert der Approximationen zu geben. Diese Überlegungen führen daher zu folgender Begriffsbildung:

**Definition 6.1.1 (Riemann-Summen)** Sei  $I = [a, b]$  ein kompaktes Intervall, und sei  $f: [a, b] \rightarrow \mathbb{K}$  eine Funktion.

- i.) Eine Zerlegung (auch Partition)  $\mathcal{J} = \{I_1, \dots, I_N\}$  von  $I$  ist eine Auswahl von  $N + 1$  Punkten  $a = t_0 < t_1 < \dots < t_N = b$  in  $[a, b]$  mit zugehörigen Teilintervallen

$$I_i = [t_{i-1}, t_i]. \quad (6.1.6)$$

- ii.) Eine markierte Zerlegung ist Paar  $(\mathcal{J}, \Xi)$  bestehend aus einer Zerlegung  $\mathcal{J}$  zusammen mit einer Auswahl  $\Xi = (\xi_1, \dots, \xi_N)$  von Punkten  $\xi_i \in I_i$  für  $i = 1, \dots, N$ .
- iii.) Zu einer markierten Zerlegung  $(\mathcal{J}, \Xi)$  mit  $\mathcal{J} = \{I_1, \dots, I_N\}$  und einer Auswahl  $\Xi = \{\xi_1, \dots, \xi_N\}$  von Zwischenpunkten  $\xi_i \in I_i$  für  $i = 1, \dots, N$  definiert man die Riemann-Summe von  $f$  als

$$\sum_{(\mathcal{J}, \Xi)}(f) = \sum_{i=1}^N f(\xi_i)(t_i - t_{i-1}). \quad (6.1.7)$$

**Bemerkung 6.1.2** Die Riemann-Summe hängt also von zwei Daten, der Zerlegung  $\mathcal{J}$  und den Zwischenpunkten  $\Xi$  ab. Wir lassen in (6.1.7) ausdrücklich auch  $\mathbb{C}$ -wertige Funktionen  $f$  zu. Die Interpretation als Fläche der rechteckigen Streifen unterhalb von  $\text{graph}(f)$  wie in Abbildung 6.1 ist natürlich nur für  $\mathbb{R}^+$ -wertige Funktionen  $f$  möglich. Die Formel (6.1.7) dagegen ist allgemeiner sinnvoll. Tatsächlich ist dies einer der Vorteile der Riemannschen Integrationstheorie gegenüber der Lebesgue-Theorie, dass wir nicht nur komplexwertige Integrale, sondern sogar vektorwertige Integrale direkt und einfach definieren könnten. Wir werden dies jedoch hier nicht weiter verfolgen.

Um einen Grenzwert von Riemann-Summen definieren zu können, nehmen wir den Standpunkt der Netzkonvergenz ein: Die Indexmenge zur Indizierung der Riemann-Summen besteht aus den Paaren  $(\mathcal{J}, \Xi)$  von Zerlegungen und Zwischenpunkten des Intervalls  $I$ . Es gilt also, eine Richtung auf dieser Menge zu etablieren, die dem „Feinerwerden“ Rechnung tragen soll. Hier bieten sich (mindestens) zwei Varianten an:

**Definition 6.1.3 (Verfeinerung)** Sei  $I = [a, b]$  ein kompaktes Intervall. Seien weiter  $(\mathcal{J}, \Xi)$  und  $(\mathcal{J}', \Xi')$  markierte Zerlegungen von  $I$  mit zugehörigen Zwischenpunkten  $\Xi$  und  $\Xi'$ .

- i.) Die Feinheit von  $(\mathcal{J}, \Xi)$  ist definiert als

$$|(\mathcal{J}, \Xi)| = \max \{|t_1 - t_0|, \dots, |t_N - t_{N-1}|\}. \quad (6.1.8)$$

- ii.) Gilt  $|(\mathcal{J}, \Xi)| \leq |(\mathcal{J}', \Xi')|$ , so heißt  $(\mathcal{J}, \Xi)$  feiner als  $(\mathcal{J}', \Xi')$ , und wir schreiben  $(\mathcal{J}, \Xi) \succ (\mathcal{J}', \Xi')$ .

Abbildung 6.2: Durch Hinzunahme von neuen Intervallgrenzen wird die maximale Intervalllänge nicht größer.

Abbildung 6.3: Die Vereinigung der Zerlegungen liefert eine gemeinsame Verfeinerung.

iii.) Gilt  $\{t_0, \dots, t_N\} \supseteq \{t'_0, \dots, t'_{N'}\}$ , so heißt  $(\mathcal{J}, \Xi)$  Verfeinerung von  $(\mathcal{J}', \Xi')$ , und wir schreiben  $(\mathcal{J}, \Xi) \gg (\mathcal{J}', \Xi')$ .

**Lemma 6.1.4** Seien  $I = [a, b]$  ein kompaktes Intervall und  $(\mathcal{J}, \Xi)$  sowie  $(\mathcal{J}', \Xi')$  markierte Zerlegungen von  $I$  mit zugehörigen Zwischenpunkten  $\Xi$  und  $\Xi'$ .

i.) Gilt  $(\mathcal{J}, \Xi) \gg (\mathcal{J}', \Xi')$ , so folgt  $(\mathcal{J}, \Xi) \succ (\mathcal{J}', \Xi')$ .

ii.) Sowohl  $\gg$  als auch  $\succ$  definieren eine Richtung auf der Menge

$$\mathfrak{Z}(I) = \{(\mathcal{J}, \Xi) \mid \mathcal{J} \text{ ist eine Zerlegung von } I \text{ mit Zwischenpunkten } \Xi\} \quad (6.1.9)$$

der markierten Zerlegungen von  $I$ .

BEWEIS: Sind die Intervallgrenzen  $t'_1, \dots, t'_{N'-1}$  eine Teilmenge der Intervallgrenzen  $t_1, \dots, t_{N-1}$ , so ist der maximale Abstand der benachbarten  $t'_i$  auf jeden Fall größer oder gleich des maximalen Abstandes der benachbarten  $t_i$ , siehe auch Abbildung 6.2. Damit ist definitionsgemäß  $|(\mathcal{J}, \Xi)| \leq |(\mathcal{J}', \Xi')|$ , was  $(\mathcal{J}, \Xi) \succ (\mathcal{J}', \Xi')$  bedeutet. Dies zeigt den ersten Teil. Für den zweiten Teil sieht man schnell, dass sowohl  $\gg$  als auch  $\succ$  reflexiv und transitiv sind. Es bleibt noch die dritte Eigenschaft iii.) aus Definition 2.8.1 zu prüfen. Seien also zwei Elemente  $(\mathcal{J}, \Xi)$  und  $(\mathcal{J}', \Xi')$  von  $\mathfrak{Z}(I)$  gegeben. Wir definieren dann die neuen Intervallgrenzen

$$\{t''_0, \dots, t''_M\} = \{t_0, \dots, t_N\} \cup \{t'_0, \dots, t'_{N'}\},$$

siehe auch Abbildung 6.3. Als Zwischenpunkte können wir irgendwelche Punkte wählen, womit wir ein neues Element  $(\mathcal{J}'', \Xi'')$  von  $\mathfrak{Z}(I)$  erhalten. Nach Konstruktion ist dies eine Verfeinerung sowohl von  $(\mathcal{J}, \Xi)$  als auch von  $(\mathcal{J}', \Xi')$ . Es gilt also  $(\mathcal{J}'', \Xi'') \gg (\mathcal{J}, \Xi)$  ebenso wie  $(\mathcal{J}'', \Xi'') \gg (\mathcal{J}', \Xi')$ , was zeigt, dass  $\gg$  eine Richtung definiert. Nach i.) folgt aber auch  $(\mathcal{J}'', \Xi'') \succ (\mathcal{J}, \Xi)$  und  $(\mathcal{J}'', \Xi'') \succ (\mathcal{J}', \Xi')$ , was zeigt, dass auch  $\succ$  eine Richtung liefert.  $\square$

Bemerkenswerterweise spielt die Wahl der Zwischenpunkte  $\Xi = \{\xi_1, \dots, \xi_N\}$  in den Paaren  $(\mathcal{J}, \Xi)$  keine Rolle bei der Richtung von  $\mathfrak{Z}(I)$ . Aus diesem Grunde erhalten wir auch gerichtete Mengen von (unmarkierten) Zerlegungen durch die gleichen Definitionen. Dies ist ein Beispiel für die Konstruktion von gerichteten Mengen wie in Beispiel 2.8.3, vi.). Die Markierungen benötigen wir jedoch entscheidend, um die Riemann-Summen (6.1.7) zu definieren. Da wir nun eine gerichtete Menge  $\mathfrak{Z}(I)$  haben, mit deren Elementen wir die Riemann-Summen indizieren, erhalten wir insgesamt ein Netz von Riemann-Summen. Als Riemann-Integral von  $f$  über  $[a, b]$  wollen wir dann den Netzlimes der Riemann-Summen definieren, sofern dieser überhaupt existiert. Die erste Schwierigkeit ist jedoch, dass wir zwei Richtungen  $\ll$  und  $\preceq$  zur Wahl haben. Wir werden später sehen, dass sich diese Wahl nicht auswirkt. Für die Definition des Riemann-Integrals entscheiden wir uns für die stärkere Richtung  $\ll$ :

**Definition 6.1.5 (Riemann-Integral)** Sei  $f: [a, b] \rightarrow \mathbb{K}$  eine Funktion.

i.) Die Funktion  $f$  heißt Riemann-integrierbar, falls das Netz  $(\sum_{(\mathcal{J}, \Xi)}(f))_{(\mathcal{J}, \Xi) \in (\mathfrak{Z}(I), \ll)}$  konvergiert.

ii.) Ist  $f$  Riemann-integrierbar, so definiert man das Riemann-Integral von  $f$  über  $[a, b]$  als

$$\int_a^b f(x) dx = \lim_{(\mathcal{J}, \Xi) \in (\mathfrak{Z}(I), \ll)} \sum_{(\mathcal{J}, \Xi)}(f) \quad (6.1.10)$$

iii.) Die Menge der Riemann-integrablen Funktionen auf  $[a, b]$  wird mit

$$\mathcal{R}([a, b], \mathbb{K}) = \{f: [a, b] \longrightarrow \mathbb{K} \mid f \text{ ist Riemann-integrierbar}\} \quad (6.1.11)$$

bezeichnet.

Die Definition ist leider nicht sofort in konkrete Kriterien zur Riemann-Integrierbarkeit umzusetzen. Hierzu benötigen wir erst weiterführende Resultate. Trotzdem erhalten wir aus der Linearität der Riemann-Summen bereits an dieser Stelle folgendes einfache Resultat:

**Proposition 6.1.6** Sei  $[a, b] \subseteq \mathbb{R}$  ein kompaktes Intervall.

i.) Die konstanten Funktionen auf  $[a, b]$  sind Riemann-integrierbar. Es gilt

$$\int_a^b c \, dx = c(b - a) \quad (6.1.12)$$

für alle  $c \in \mathbb{K}$ .

ii.) Die Riemann-integrierbaren Funktionen

$$\mathcal{R}([a, b], \mathbb{K}) \subseteq \text{Abb}([a, b], \mathbb{K}) \quad (6.1.13)$$

bilden einen Unterraum aller Funktionen.

iii.) Das Riemann-Integral ist ein lineares Funktional

$$\int_a^b \cdot \, dx: \mathcal{R}([a, b], \mathbb{K}) \longrightarrow \mathbb{K}. \quad (6.1.14)$$

BEWEIS: Sei  $c \in \mathbb{K}$  vorgegeben und  $f(x) = c$  die zugehörige konstante Funktion auf  $[a, b]$ . Dann gilt für eine Zerlegung  $\mathcal{J}$  von  $I$  mit Zwischenpunkten  $\Xi$  für die Riemann-Summe

$$\sum_{(\mathcal{J}, \Xi)}(f) = \sum_{i=1}^N f(\xi_i)(t_i - t_{i-1}) = \sum_{i=1}^N c(t_i - t_{i-1}) = c(b - a),$$

womit das Netz der Riemann-Summen das konstante Netz mit Wert  $c(b - a)$  ist. Dieses konvergiert trivialerweise mit Grenzwert wie in (6.1.12) angegeben. Insbesondere gilt  $0 \in \mathcal{R}([a, b], \mathbb{K})$ . Seien nun  $f, g \in \mathcal{R}([a, b], \mathbb{K})$  sowie  $z, w \in \mathbb{K}$ . Für eine Zerlegung  $\mathcal{J}$  von  $[a, b]$  mit Zwischenpunkten  $\Xi$  gilt dann

$$\begin{aligned} \sum_{(\mathcal{J}, \Xi)}(zf + wg) &= \sum_{i=1}^N (zf(\xi_i) + wg(\xi_i))(t_i - t_{i-1}) \\ &= z \sum_{i=1}^N f(\xi_i)(t_i - t_{i-1}) + w \sum_{i=1}^N g(\xi_i)(t_i - t_{i-1}) \\ &= z \sum_{(\mathcal{J}, \Xi)}(f) + w \sum_{(\mathcal{J}, \Xi)}(g). \end{aligned}$$

Da die Netze  $(\sum_{(\mathcal{J}, \Xi)}(f))_{(\mathcal{J}, \Xi) \in \mathfrak{Z}(I)}$  und  $(\sum_{(\mathcal{J}, \Xi)}(g))_{(\mathcal{J}, \Xi) \in \mathfrak{Z}(I)}$  beide nach Voraussetzung konvergieren, gilt dies auch für das Netz

$$(\sum_{(\mathcal{J}, \Xi)}(zf + wg))_{(\mathcal{J}, \Xi)} = z(\sum_{(\mathcal{J}, \Xi)}(f))_{(\mathcal{J}, \Xi)} + w(\sum_{(\mathcal{J}, \Xi)}(g))_{(\mathcal{J}, \Xi)}. \quad (*)$$

Also folgt  $zf + wg \in \mathcal{R}([a, b], \mathbb{K})$ . Zudem können wir den Grenzwert der linken Seite von  $(*)$  durch die Grenzwerte der Netze auf der rechten Seite ausrechnen, was dann den dritten Teil zeigt.  $\square$

Der Vektorraum  $\mathcal{R}([a, b], \mathbb{K})$  wird also der Ausgangspunkt aller Überlegungen zum Riemann-Integral sein. Momentan haben wir weder viele Beispiele noch konkrete Möglichkeiten, Integrale auch zu berechnen. Die nächste Proposition zeigt einige weitere Eigenschaften von  $\mathcal{R}([a, b], \mathbb{K})$  und dem Integral auf.

**Proposition 6.1.7** *Sei  $[a, b] \subseteq \mathbb{R}$  ein kompaktes Intervall.*

i.) *Es gilt genau dann  $f \in \mathcal{R}([a, b], \mathbb{C})$ , wenn  $\bar{f} \in \mathcal{R}([a, b], \mathbb{C})$ . In diesem Fall gilt*

$$\int_a^b \bar{f}(x) dx = \overline{\int_a^b f(x) dx}. \quad (6.1.15)$$

ii.) *Es gilt genau dann  $f \in \mathcal{R}([a, b], \mathbb{C})$ , wenn  $\operatorname{Re}(f), \operatorname{Im}(f) \in \mathcal{R}([a, b], \mathbb{R})$ . In diesem Fall gilt*

$$\int_a^b \operatorname{Re}(f)(x) dx = \operatorname{Re} \left( \int_a^b f(x) dx \right) \quad (6.1.16)$$

und

$$\int_a^b \operatorname{Im}(f)(x) dx = \operatorname{Im} \left( \int_a^b f(x) dx \right). \quad (6.1.17)$$

iii.) *Seien  $f, g \in \mathcal{R}([a, b], \mathbb{R})$  mit  $f \leq g$  gegeben. Dann gilt*

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx. \quad (6.1.18)$$

iv.) *Sei  $f \in \mathcal{R}([a, b], \mathbb{R})$  mit  $f \geq 0$  gegeben. Dann gilt*

$$\int_a^b f(x) dx \geq 0. \quad (6.1.19)$$

BEWEIS: Für die Riemann-Summen gilt

$$\sum_{(j, \Xi)} (\bar{f}) = \sum_{i=1}^N \bar{f}(\xi_i)(t_i - t_{i-1}) = \overline{\sum_{i=1}^N f(\xi_i)(t_i - t_{i-1})} = \overline{\sum_{(j, \Xi)} (f)}, \quad (*)$$

womit der Limes für  $f$  genau dann existiert, wenn der Limes für  $\bar{f}$  existiert. In diesem Fall sind beide Limes in  $(*)$  gleich. Dies folgt aus den Rechenregeln für Netzlimits in Proposition 2.8.7, iii.). Mit der Linearität des Integrals und  $\operatorname{Re}(f) = \frac{1}{2}(f + \bar{f})$  sowie  $\operatorname{Im}(f) = \frac{1}{2i}(f - \bar{f})$  sowie  $f = \operatorname{Re}(f) + i\operatorname{Im}(f)$  erhalten wir dann aus Proposition 2.8.7, i.), den zweiten Teil. Der dritte Teil folgt aus Proposition 2.8.7, v.). Der vierte ist eine Konsequenz davon, dass die Nullfunktion Riemann-integrierbar mit Integral Null ist.  $\square$

## 6.2 Riemann-integrierbare Funktionen

Es stellt sich nun also die entscheidende Frage, welche Funktionen tatsächlich Riemann-integrierbar sind. Als erstes notwendiges Kriterium wollen wir zeigen, dass Riemann-integrierbare Funktionen notwendigerweise beschränkt sind. Hierzu benötigen wir folgendes unabhängiges Resultat:

**Proposition 6.2.1** *Sei  $X$  eine nicht-leere kompakte Teilmenge der reellen oder komplexen Zahlen und  $f: X \rightarrow \mathbb{R}$ . Dann gibt es ein  $x_0 \in X$ , so dass für jede Umgebung  $U \subseteq X$  von  $x_0$  in  $X$*

$$\sup(f|_U) = \sup(f) \quad (6.2.1)$$

*gilt. Eine analoge Aussage gilt für das Infimum von  $f$ .*

BEWEIS: Sei  $\alpha = \sup f \in \mathbb{R} \cup \{\infty\}$ , wobei  $\alpha \neq -\infty$ , da  $K$  als nicht-leer angenommen wurde. Der Wert  $+\infty$  ist dagegen explizit erlaubt und möglich. Nach Definition des Supremums der Menge  $f(X)$  gibt es zu jedem  $n \in \mathbb{N}$  ein  $x_n \in X$  mit  $f(x_n) > \alpha - \frac{1}{n}$  im Falle, dass  $\alpha < \infty$ . Für  $\alpha = \infty$  gibt es zu jedem  $n \in \mathbb{N}$  ein  $x_n \in X$  mit  $f(x_n) \geq n$ . In beiden Fällen erhalten wir eine Folge  $(x_n)_{n \in \mathbb{N}}$  in  $X$ , welche nach dem Satz von Heine-Borel, siehe Satz 3.4.11, eine konvergente Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  mit Grenzwert  $x_0 = \lim_{k \rightarrow \infty} x_{n_k}$  besitzt. Wir behaupten, dieses  $x_0$  erfüllt die Eigenschaft (6.2.1). Sei also  $U \subseteq X$  eine Umgebung von  $x_0$  in  $X$  und  $\alpha_U = \sup(f|_U)$  das Supremum der eingeschränkten Funktion. Es gilt trivialerweise  $\alpha_U \leq \alpha$ . Ist nun  $\alpha = \infty$ , so sind alle bis auf endlich viele Folgenglieder von  $(x_{n_k})_{k \in \mathbb{N}}$  in  $U$  enthalten. Dies zeigt, dass  $\alpha_U = \sup(f|_U) \geq f(x_{n_k}) \geq n_k$  für fast alle  $k \in \mathbb{N}$ , also  $\alpha_U = \infty = \alpha$ , was (6.2.1) in diesem Fall zeigt. Ist  $\alpha < \infty$ , so gilt wieder  $\alpha_U = \sup(f|_U) \geq f(x_{n_k}) \geq \alpha - \frac{1}{n_k}$  für fast alle  $k$ . Auch in diesem Fall folgt daher  $\alpha_U = \alpha$ .  $\square$

Man beachte, dass diese Aussage für beliebige reell-wertige Funktionen gültig ist, also mehr eine Aussage über Kompakta als über Funktionen ist. Wir nutzen diese Proposition nun, um folgenden Satz zu zeigen:

**Satz 6.2.2 (Riemann-integriable Funktionen sind beschränkt)** Sei  $[a, b] \subseteq \mathbb{R}$  ein kompaktes Intervall. Dann gilt

$$\mathcal{R}([a, b], \mathbb{K}) \subseteq \mathcal{B}([a, b], \mathbb{K}). \quad (6.2.2)$$

BEWEIS: Ist  $\mathbb{K} = \mathbb{C}$  und  $f$  unbeschränkt, so ist  $\operatorname{Re}(f)$  oder  $\operatorname{Im}(f)$  unbeschränkt. Da mit  $f \in \mathcal{R}([a, b], \mathbb{C})$  auch  $\operatorname{Re}(f)$  und  $\operatorname{Im}(f)$  Riemann-integrierbar sind, können wir ohne Einschränkung  $\mathbb{K} = \mathbb{R}$  annehmen. Wir nehmen an,  $f \in \operatorname{Abb}([a, b], \mathbb{R})$  sei unbeschränkt. Dann gilt entweder  $\sup f = +\infty$  oder  $\inf f = -\infty$ . Wir betrachten den ersten Fall, da der zweite durch  $f \leftrightarrow -f$  auf den ersten zurückgeführt werden kann. Wir finden nach Proposition 6.2.1 also ein  $x_0 \in [a, b]$ , so dass  $\sup(f|_U) = \infty$  für jede Umgebung  $U \subseteq [a, b]$  von  $x_0$  in  $[a, b]$ . Wir wählen eine Zahl  $\alpha \in \mathbb{R}^+$  und eine beliebige markierte Zerlegung  $(\mathcal{J}, \Xi)$ . Dann gibt es einen Index  $i \in \{1, \dots, N\}$  mit

$$x_0 \in [t_{i-1}, t_i]$$

Nun gibt es zwei prinzipielle Möglichkeiten:  $x_0 \in (t_{i-1}, t_i)$  oder  $x_0$  ist einer der Randpunkte.

i.) Ist  $x_0 \in (t_{i-1}, t_i)$ , so betrachten wir die zu  $(\mathcal{J}, \Xi)$  gehörige Riemann-Summe

$$S = \sum_{\substack{j=1 \\ i \neq j}}^N f(\xi_j)(t_j - t_{j-1})$$

ohne den  $i$ -ten Term. Nach Proposition 6.2.1 finden wir ein  $\xi \in (t_{i-1}, t_i)$  mit

$$f(\xi) \geq \frac{\alpha - S}{t_i - t_{i-1}} \quad (*)$$

oder eben

$$\sum_{\substack{j=1 \\ i \neq j}}^N f(\xi_j)(t_j - t_{j-1}) + f(\xi)(t_i - t_{i-1}) \geq \alpha.$$

Ersetzen wir den Zwischenpunkt  $\xi_i$  durch  $\xi$ , so erhalten wir eine Zerlegung  $(\mathcal{J}, \Xi')$  mit

$$\sum_{(\mathcal{J}, \Xi')} (f) \geq \alpha \quad (**)$$



ii.) Ist  $x_0 = t_{i-1}$ , so funktioniert i.) nur dann, wenn  $x_0 = t_0 = a$  ist, da dann  $[x_0, t_1)$  eine offene Umgebung von  $x_0$  in  $[a, b]$  ist und wir nach wie vor  $\xi \in [x_0, t_1)$  mit (\*) finden. Ist dagegen  $i \geq 2$ , so müssen wir die größere offene Umgebung  $(t_{i-2}, t_i)$  von  $x_0$  betrachten. Wir setzen dann

$$S' = \sum_{\substack{j=1 \\ j \neq i}}^N f(\xi_j)(t_j - t_{j-1}) \quad \text{und} \quad S'' = \sum_{\substack{j=1 \\ j \neq i-1}}^N f(\xi_j)(t_j - t_{j-1})$$

sowie  $S = \max\{S', S''\}$ . Entsprechend finden wir nach Proposition 6.2.1 ein  $\xi \in (t_{i-2}, t_i)$  mit

$$f(\xi) \geq \frac{\alpha - S}{\min\{t_i - t_{i-1}, t_{i-1} - t_{i-2}\}} \quad (*)$$

Die Schwierigkeit ist, dass wir nicht wissen, auf welcher Seite von  $x_0$  dieses  $\xi$  liegt, weshalb wir mit (\*) beide Situationen abdecken müssen. Ist  $\xi$  bestimmt, können wir, je nachdem auf welcher Seite von  $x_0$  dieser Punkt liegt, entweder  $\xi_{i-1}$  oder  $\xi_i$  durch  $\xi$  ersetzen. Für die resultierende markierte Zerlegung  $(\mathcal{J}, \Xi')$  erhalten wir dann erneut (\*\*). Ist schließlich  $x_0 = t_N = b$ , so vereinfacht sich die Situation wieder, da  $\xi$  auf jeden Fall links von  $x_0$  liegt und wir  $\xi_N$  ersetzen müssen.

In beiden Fällen haben wir aber zu  $\alpha \in \mathbb{R}^+$  und einer beliebigen markierten Zerlegung eine markierte Zerlegung  $(\mathcal{J}, \Xi') \gg (\mathcal{J}, \Xi)$  gefunden, für die (\*\*) gilt. Damit kann  $f$  aber nicht Riemann-integrierbar sein.  $\square$

Als erste Konsequenz erhalten wir aus Satz 6.2.2 die folgende fundamentale Abschätzung für Riemann-Integrale:

**Korollar 6.2.3** Sei  $f \in \mathcal{R}([a, b], \mathbb{K})$ . Dann gilt

$$\left| \int_a^b f(x) \, dx \right| \leq \|f\|_\infty (b - a). \quad (6.2.3)$$

BEWEIS: Zunächst wissen wir nach Satz 6.2.2, dass  $\|f\|_\infty < \infty$  ist. Für alle Punkte  $\xi \in [a, b]$  gilt daher  $|f(\xi)| \leq \|f\|_\infty$ . Ist nun  $(\mathcal{J}, \Xi)$  eine beliebige Zerlegung, so gilt

$$\begin{aligned} |\sum_{(\mathcal{J}, \Xi)}(f)| &= \left| \sum_{i=1}^N f(\xi_i)(t_i - t_{i-1}) \right| \\ &\leq \sum_{i=1}^N |f(\xi_i)|(t_i - t_{i-1}) \\ &\leq \sum_{i=1}^N \|f\|_\infty (t_i - t_{i-1}) \\ &= \|f\|_\infty (b - a), \end{aligned}$$

was (6.2.3) zeigt.  $\square$

Die Beschränktheit der Riemann-integrierbaren Funktionen erlaubt nun eine alternative Konstruktion, wenn wir uns auf Funktionen mit Werten in  $\mathbb{R}$  einschränken. Im Hinblick auf Proposition 6.1.7 ist dies keine schwerwiegende Einschränkung.

**Definition 6.2.4 (Ober- und Unterintegral)** Sei  $I = [a, b] \subseteq \mathbb{R}$  ein kompaktes Intervall, und sei  $f: [a, b] \rightarrow \mathbb{R}$  eine beschränkte Funktion.

i.) Für eine Zerlegung  $\mathcal{J} = \{t_0 = a < t_1 < \cdots < t_N = b\}$  von  $[a, b]$  definiert man die Obersumme von  $f$  als

$$\mathcal{O}_{\mathcal{J}}(f) = \sum_{i=1}^N \sup \left( f|_{[t_{i-1}, t_i]} \right) (t_i - t_{i-1}) \quad (6.2.4)$$

und die Untersumme von  $f$  durch

$$\mathcal{U}_{\mathcal{J}}(f) = \sum_{i=1}^N \inf \left( f|_{[t_{i-1}, t_i]} \right) (t_i - t_{i-1}). \quad (6.2.5)$$

ii.) Das Oberintegral von  $f$  ist definiert als

$$\int_a^{\bar{b}} f(x) \, dx = \inf \{ \mathcal{O}_{\mathcal{J}}(f) \mid \mathcal{J} \text{ ist Zerlegung von } I \}, \quad (6.2.6)$$

und das Unterintegral von  $f$  ist entsprechend

$$\int_a^b f(x) \, dx = \sup \{ \mathcal{U}_{\mathcal{J}}(f) \mid \mathcal{J} \text{ ist Zerlegung von } I \}. \quad (6.2.7)$$

iii.) Die Funktion  $f$  heißt Darboux-integrierbar, falls

$$\int_a^b f(x) \, dx = \int_a^{\bar{b}} f(x) \, dx \quad (6.2.8)$$

gilt. In diesem Fall nennt man den Wert (6.2.8) das Darboux-Integral von  $f$  über  $[a, b]$ .

Der Vorteil der Betrachtung von Ober- und Untersummen ist, dass wir allein mit Zerlegungen und nicht mit markierten Zerlegungen arbeiten können. Der Nachteil ist, dass wir auf reelle Funktionswerte angewiesen sind. Wie schon erwähnt, verwenden wir die beiden Richtungen  $\preccurlyeq$  und  $\ll$  auch für (unmarkierte) Zerlegungen.

**Proposition 6.2.5** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine beschränkte Funktion auf einem kompakten Intervall  $I = [a, b] \subseteq \mathbb{R}$ .

i.) Sind  $\mathcal{J}, \mathcal{J}'$  Zerlegungen von  $I$  mit  $\mathcal{J} \ll \mathcal{J}'$ , so gilt

$$- \|f\|_{\infty}(b-a) \leq \mathcal{U}_{\mathcal{J}}(f) \leq \mathcal{U}_{\mathcal{J}'}(f) \leq \mathcal{O}_{\mathcal{J}'}(f) \leq \mathcal{O}_{\mathcal{J}}(f) \leq \|f\|_{\infty}(b-a). \quad (6.2.9)$$

ii.) Das Netz der Obersummen  $(\mathcal{O}_{\mathcal{J}}(f))_{\mathcal{J}}$  ist bezüglich  $\ll$  monoton fallend und nach unten beschränkt.

iii.) Das Netz der Untersummen  $(\mathcal{U}_{\mathcal{J}}(f))_{\mathcal{J}}$  ist bezüglich  $\ll$  monoton wachsend und nach oben beschränkt.

iv.) Es gilt

$$\lim_{\mathcal{J}} \mathcal{O}_{\mathcal{J}}(f) = \int_a^{\bar{b}} f(x) \, dx \quad (6.2.10)$$

und

$$\lim_{\mathcal{J}} \mathcal{U}_{\mathcal{J}}(f) = \int_a^b f(x) \, dx. \quad (6.2.11)$$

v.) Für jede Wahl von Zwischenpunkten  $\Xi$  zu einer Zerlegung  $\mathcal{J}$  gilt

$$\mathcal{U}_{\mathcal{J}}(f) \leq \sum_{(\mathcal{J}, \Xi)}(f) \leq \mathcal{O}_{\mathcal{J}}(f). \quad (6.2.12)$$

vi.) Die Funktion  $f$  ist genau dann Darboux-integrierbar, wenn es zu jedem  $\epsilon > 0$  eine Zerlegung  $\mathcal{J}$  von  $I$  gibt, sodass

$$\mathcal{O}_{\mathcal{J}}(f) - \mathcal{U}_{\mathcal{J}}(f) < \epsilon. \quad (6.2.13)$$

BEWEIS: Seien  $\mathcal{J} = \{a = t_0 < t_1 < \dots < t_N = b\}$  und  $\mathcal{J}' = \{a = t'_0 < t'_1 < \dots < t'_M = b\}$  zwei Zerlegungen von  $[a, b]$  mit  $\mathcal{J} \subseteq \mathcal{J}'$  also  $\mathcal{J} \ll \mathcal{J}'$ . Wir nehmen zunächst an, dass  $M = N + 1$  gilt. Dann gibt es also ein Teilintervall  $[t_{i-1}, t_i]$  von  $\mathcal{J}$ , welches einen zusätzlichen Punkt  $t'_i \in \mathcal{J}'$  enthält. Somit gilt

$$\begin{aligned} & \sup \left( f|_{[t_{i-1}, t'_i]} \right) (t'_i - t_{i-1}) + \sup \left( f|_{[t'_i, t_i]} \right) (t_i - t'_i) \\ & \leq \sup \left( f|_{[t_{i-1}, t_i]} \right) (t'_i - t_{i-1}) + \sup \left( f|_{[t_{i-1}, t_i]} \right) (t_i - t'_i) \\ & = \sup \left( f|_{[t_{i-1}, t_i]} \right) (t_i - t_{i-1}). \end{aligned}$$

Da alle anderen Terme der Obersumme gleich bleiben, gilt in diesem Fall

$$\mathcal{O}_{\mathcal{J}'}(f) \leq \mathcal{O}_{\mathcal{J}}(f). \quad (*)$$

Durch Induktion zeigt man dann, dass  $(*)$  auch für beliebige  $N$  und  $M$  gilt. Die Abschätzung

$$\mathcal{U}_{\mathcal{J}}(f) \leq \mathcal{U}_{\mathcal{J}'}(f)$$

erfolgt genauso. Da das Infimum von  $f$  über ein Teilintervall  $[t_{i-1}, t_i]$  immer kleiner oder gleich dem Supremum von  $f$  über  $[t_{i-1}, t_i]$  ist, folgt auch die Ungleichung  $\mathcal{U}_{\mathcal{J}'}(f) \leq \mathcal{O}_{\mathcal{J}'}(f)$ . Für die triviale Zerlegung  $\mathcal{J}_0 = \{a = t_0 < t_1 = b\}$  gilt  $\mathcal{J}_0 \ll \mathcal{J}$  für jede andere Zerlegung  $\mathcal{J}$ . Es gilt weiter

$$-\|f\|_{\infty}(b-a) \leq \inf(f)(b-a) = \mathcal{U}_{\mathcal{J}_0}(f)$$

und

$$\mathcal{O}_{\mathcal{J}_0}(f) = \sup(f)(b-a) \leq \|f\|_{\infty}(b-a),$$

womit *i.)* gezeigt ist. Damit ist *ii.)* und *iii.)* aber klar. Der vierte Teil gilt nun ganz allgemein: Ein monoton wachsendes Netz von reellen Zahlen konvergiert gegen das Supremum seiner Werte. Dies zeigt man wie für Folgen in Proposition 2.4.9, siehe auch Übung 2.9.19. Da für jeden Punkt  $\xi_i \in [t_{i-1}, t_i]$  die Ungleichung

$$\inf \left( f|_{[t_{i-1}, t_i]} \right) \leq f(\xi_i) \leq \sup \left( f|_{[t_{i-1}, t_i]} \right)$$

nach Definition von Infimum und Supremum gilt, folgt (6.2.12) direkt, siehe Abbildung 6.4. Auch der letzte Teil ist einfach. Zu  $\epsilon > 0$  gibt es eine Zerlegung  $\mathcal{J}'$  mit

$$\mathcal{O}_{\mathcal{J}'}(f) \leq \int_a^b f(x) dx + \frac{\epsilon}{2}$$

und eine Zerlegung  $\mathcal{J}''$  mit

$$\mathcal{U}_{\mathcal{J}''}(f) \geq \int_a^b f(x) dx - \frac{\epsilon}{2}.$$

Also gilt  $\mathcal{O}_{\mathcal{J}'}(f) - \mathcal{U}_{\mathcal{J}''}(f) < \epsilon$ . Ist nun  $\mathcal{J}$  eine gemeinsame Verfeinerung von  $\mathcal{J}'$  und  $\mathcal{J}''$ , so gilt wegen  $\mathcal{O}_{\mathcal{J}}(f) \leq \mathcal{O}_{\mathcal{J}'}(f)$  und  $\mathcal{U}_{\mathcal{J}''}(f) \leq \mathcal{U}_{\mathcal{J}}(f)$  erst recht  $\mathcal{O}_{\mathcal{J}}(f) - \mathcal{U}_{\mathcal{J}}(f) < \epsilon$ , was die eine Richtung zeigt. Findet man umgekehrt zu jedem  $\epsilon > 0$  eine Zerlegung  $\mathcal{J}$  mit (6.2.13), so gilt für jede Verfeinerung  $\mathcal{J}'$  von  $\mathcal{J}$  die Ungleichung

$$\mathcal{O}_{\mathcal{J}}(f) < \mathcal{U}_{\mathcal{J}}(f) + \epsilon \leq \mathcal{U}_{\mathcal{J}'} + \epsilon,$$

Abbildung 6.4: Zu einer (markierten) Zerlegung  $(\mathcal{J}, \Xi)$  die Ober- und Untersummen sowie die Riemann-Summe.

womit

$$\mathcal{O}_{\mathcal{J}}(f) \leq \sup\{\mathcal{U}_{\mathcal{J}'} \mid \mathcal{J}' \gg \mathcal{J}\} + \epsilon = \int_a^b f(x) dx + \epsilon$$

folgt, da die Zerlegungen, die nicht feiner als  $\mathcal{J}'$  sind, gemeinsame Verfeinerungen mit  $\mathcal{J}'$  besitzen, die größere Untersummen liefern. Daher ist das obige Supremum bereits ausreichend, um das Unterintegral zu erhalten. Weil nun aber für jede Verfeinerung  $\mathcal{J}'$  von  $\mathcal{J}$  auch  $\mathcal{O}_{\mathcal{J}'}(f) \leq \mathcal{O}_{\mathcal{J}}(f)$  gilt, folgt für alle solchen  $\mathcal{J}'$

$$\mathcal{O}_{\mathcal{J}'}(f) \leq \int_a^b f(x) dx + \epsilon.$$

Wieder können wir das Infimum über alle  $\mathcal{J}'$  bilden, und erhalten

$$\int_a^{\bar{b}} f(x) dx \leq \int_a^b f(x) dx + \epsilon$$

für alle  $\epsilon > 0$ . Da die Ungleichung

$$\int_a^{\bar{b}} f(x) dx \geq \int_a^b f(x) dx$$

trivialerweise immer gilt, folgt also die Gleichheit, was *vi.*) zeigt.  $\square$

Nach diesen Vorbereitungen können wir nun zeigen, dass Riemann-Integrierbarkeit dasselbe ist wie Darboux-Integrierbarkeit. Dies vereinfacht die Riemann-Integrierbarkeit deutlich:

**Satz 6.2.6 (Riemann- und Darboux-Integral)** *Sei  $I = [a, b] \subseteq \mathbb{R}$  ein kompaktes Intervall. Eine Funktion  $f: [a, b] \rightarrow \mathbb{R}$  ist genau dann Riemann-integrierbar, wenn  $f$  beschränkt und Darboux-integrierbar ist. In diesem Fall gilt*

$$\int_a^b f(x) dx = \int_a^b f(x) dx = \int_a^{\bar{b}} f(x) dx. \quad (6.2.14)$$

BEWEIS: Sei zunächst  $f$  Riemann-integrierbar. Nach Satz 6.2.2 ist  $f$  beschränkt. Für  $\epsilon > 0$  finden wir dann eine Zerlegung  $\mathcal{J}_0$  von  $I$  mit der Eigenschaft, dass

$$\left| \sum_{(\mathcal{J}, \Xi)}(f) - \int_a^b f(x) dx \right| < \frac{\epsilon}{2}$$

für alle markierten Zerlegungen  $(\mathcal{J}, \Xi)$  mit  $\mathcal{J} \gg \mathcal{J}_0$ . Es gilt also

$$\int_a^b f(x) dx - \frac{\epsilon}{2} < \sum_{(\mathcal{J}, \Xi)}(f) < \int_a^b f(x) dx + \frac{\epsilon}{2}$$

Da die Zwischenpunkte  $\Xi = \{\xi_1, \dots, \xi_N\}$  beliebig sind, gilt immer noch

$$\int_a^b f(x) dx - \frac{\epsilon}{2} \leq \inf_{\Xi} \sum_{(\mathcal{J}, \Xi)}(f) = \mathcal{U}_{\mathcal{J}}(f) \quad (*)$$

sowie

$$\mathcal{O}_{\mathcal{J}}(f) = \sup_{\Xi} \sum_{(\mathcal{J}, \Xi)}(f) \leq \int_a^b f(x) dx + \frac{\epsilon}{2} \quad (**)$$

Man beachte, dass generell

$$\mathcal{U}_{\mathcal{J}}(f) = \inf_{\Xi} \sum_{(\mathcal{J}, \Xi)}(f)$$

und entsprechend

$$\mathcal{O}_{\mathcal{J}}(f) = \sup_{\Xi} \sum_{(\mathcal{J}, \Xi)}(f)$$

nach den Rechenregeln für Suprema und Infima bei endlichen Summen aus Proposition 1.5.13, *iii.*), gilt. Damit folgt aber

$$\mathcal{O}_{\mathcal{J}}(f) - \mathcal{U}_{\mathcal{J}}(f) < \epsilon$$

für alle solchen Zerlegungen  $\mathcal{J}$ . Nach Proposition 6.2.5, *vi.*), ist  $f$  Darboux-integrierbar. Der Grenzwert der Obersummen ist nach  $(**)$  durch  $\int_a^b f(x) dx$  nach oben beschränkt, der Grenzwert der Untersummen nach  $(*)$  durch  $\int_a^b f(x) dx$  nach unten. Damit stimmen die Ober- und Unterintegrale mit dem Riemann-Integral überein.

Sei umgekehrt  $f$  Darboux-integrierbar. Dann gilt für jede markierte Zerlegung  $(\mathcal{J}, \Xi)$  die Ungleichung

$$\mathcal{U}_{\mathcal{J}}(f) \leq \sum_{(\mathcal{J}, \Xi)}(f) \leq \mathcal{O}_{\mathcal{J}}(f).$$

Da die Richtung der markierten Zerlegungen  $(\mathcal{J}, \Xi)$  nur von der Richtung  $\ll$  der (unmarkierten) Zerlegungen  $\mathcal{J}$  Gebrauch macht und  $\Xi$  ignoriert, können wir die Netz-Konvergenz von den Unter- und Obersummen verwenden, um nach dem Sandwich-Prinzip, diesmal für Netze, die Konvergenz der Riemann-Summen gegen den gleichen Grenzwert zu folgern. Dies zeigt die verbliebene Implikation.  $\square$

**Bemerkung 6.2.7** Es stellt sich natürlich die berechtigte Frage, wieso man hier zwei (äquivalente) Integralbegriffe zur Auswahl hat und welchen man bevorzugen sollte. Das Darboux-Integral hat sicher den großen Vorteil, in vielen technischen Belangen einfacher zu sein. Man kommt eben mit einfachen Zerlegungen aus und kann auf die markierten Zerlegungen verzichten. Aus diesem Grunde wird in vielen Lehrbüchern dieser Zugang gewählt. Andererseits bietet das Riemann-Integral unmittelbarer eine Definition auch für komplex-wertige Funktionen. Dies ist tatsächlich von enormer Tragweite, da man zwar das Darboux-Integral einfach separat für Real- und Imaginärteil definieren kann, das Riemann-Integral aber leicht auf allgemeine vektor-wertige Funktionen ausgedehnt werden kann. Diese Option erweist sich später als die tatsächlich entscheidende Stärke des Riemann-Integrals.

**Korollar 6.2.8** Sei  $f: [a, b] \rightarrow \mathbb{R}$  monoton. Dann gilt  $f \in \mathcal{R}([a, b], \mathbb{R})$ .

BEWEIS: Wir wählen eine äquidistante Zerlegung  $\mathcal{J}$  mit  $a = t_0 < t_1 < \dots < t_N = b$  mit  $t_i = a + \frac{i}{N}(b-a)$ . Weiter dürfen wir ohne Einschränkung annehmen, dass  $f$  monoton wachsend ist, andernfalls betrachten wir  $-f$  stattdessen. Dann gilt

$$\sup \left( f|_{[t_{i-1}, t_i]} \right) = f(t_i)$$

sowie

$$\inf \left( f|_{[t_{i-1}, t_i]} \right) = f(t_{i-1})$$

aufgrund der Monotonie. Wir berechnen

$$\begin{aligned} \mathcal{O}_{\mathcal{J}}(f) - \mathcal{U}_{\mathcal{J}}(f) &= \sum_{i=1}^N \left( \sup \left( f|_{[t_{i-1}, t_i]} \right) - \inf \left( f|_{[t_{i-1}, t_i]} \right) \right) (t_i - t_{i-1}) \\ &= \sum_{i=1}^N (f(t_i) - f(t_{i-1})) \frac{b-a}{N} \end{aligned}$$

$$= (f(b) - f(a)) \frac{b-a}{N}$$

mittels eines Teleskopsummenarguments. Wählt man  $N$  groß genug, so folgt, dass diese Differenz beliebig klein wird. Nach Proposition 6.2.5, *vi.*), ist  $f$  Darboux-integrierbar und damit nach Satz 6.2.6 Riemann-integrierbar.  $\square$

Mit diesem Korollar haben wir viele Funktionen gefunden, die Riemann-integrierbar sind. Es gibt aber auch Funktionen, die nicht Riemann-integrierbar sind, und nicht-monotone Funktionen, die trotzdem Riemann-integrierbar sind.

**Beispiel 6.2.9 (Riemann-integrierbare Funktionen)** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall mit  $a < b$ .

- i.)* Wir betrachten erneut die Dirichlet-Funktion  $\chi_{\mathbb{Q}}$  aus Beispiel 3.3.6, *ii.*), eingeschränkt auf  $[a, b]$ . Da  $\chi_{\mathbb{Q}}(x) = 1$  für  $x \in \mathbb{Q}$ , ist in jedem Teilintervall  $[t_{i-1}, t_i]$  zu jeder Zerlegung  $\mathcal{J} = \{a = t_0 < t_1 < \dots < t_N = b\}$  von  $[a, b]$  immer

$$\sup(\chi_{\mathbb{Q}}|_{[t_{i-1}, t_i]}) = 1, \quad (6.2.15)$$

weil es immer rationale Punkte in  $[t_{i-1}, t_i]$  gibt. Mit der gleichen Argumentation finden wir

$$\inf(\chi_{\mathbb{Q}}|_{[t_{i-1}, t_i]}) = 0 \quad (6.2.16)$$

für jede Zerlegung, da  $[t_{i-1}, t_i]$  immer auch irrationale Punkte enthält. Es folgt

$$\mathcal{O}_{\mathcal{J}}(\chi_{\mathbb{Q}}|_{[a, b]}) = b - a \quad (6.2.17)$$

und

$$\mathcal{U}_{\mathcal{J}}(\chi_{\mathbb{Q}}|_{[a, b]}) = 0 \quad (6.2.18)$$

für jede Zerlegung. Nach Proposition 6.2.5, *vi.*), und Satz 6.2.6 ist  $\chi_{\mathbb{Q}}$  daher nicht Riemann-integrierbar.

- ii.)* Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  aus Beispiel 3.3.6, *iv.*), mit

$$f(x) = \begin{cases} \frac{1}{q} & x = \frac{p}{q} \in \mathbb{Q} \setminus \{0\}, p \in \mathbb{Z}, q \in \mathbb{N} \text{ teilerfremd} \\ 0 & \text{sonst} \end{cases} \quad (6.2.19)$$

ist dagegen auf  $[a, b]$  Riemann-integrierbar. Dies ist nicht ganz so einfach zu sehen. Wir wählen  $\epsilon > 0$ . Dann gibt es in  $[a, b]$  nur endlich viele rationale Zahlen  $\frac{p}{q}$  mit  $p \in \mathbb{Z}$  und  $q \in \mathbb{N}$  teilerfremd, sodass  $\epsilon < \frac{1}{q}$  gilt. Diese Anzahl bezeichnen wir mit  $n_{\epsilon} \in \mathbb{N}$ . Wir betrachten dann eine äquidistante Zerlegung  $\mathcal{J} = \{a = t_0 < t_1 < \dots < t_N = b\}$  mit

$$t_i = a + \frac{i}{N}(b-a), \quad (6.2.20)$$

wie bereits schon öfters gesehen. Einer der obigen rationalen Punkte  $\frac{p}{q}$  mit  $\epsilon < \frac{1}{q}$  kann dann maximal in zwei Intervallen der Zerlegung liegen, wenn er mit einem der  $t_i$  übereinstimmt, ansonsten in nur einem. Es gibt also höchstens  $2n_{\epsilon}$  Teilintervalle, in denen die obigen Ausnahmepunkte liegen. Auf einem solchen Teilintervall ist das Supremum von  $f$  dann das Maximum der  $\frac{1}{q}$ , und somit höchstens 1. Auf allen anderen Intervallen ist das Supremum von  $f$  höchstens

$\epsilon$ , da bei den dortigen rationalen Punkten  $\frac{p}{q}$  eben  $f(\frac{p}{q}) = \frac{1}{q} \leq \epsilon$  gilt und bei den irrationalen Punkten  $f(x) = 0$  sowieso verschwindet. Für die Obersummen erhalten wir also

$$\mathcal{O}_{\mathcal{J}}(f) \leq 2n_{\epsilon} \frac{b-a}{N} + \epsilon(b-a). \quad (6.2.21)$$

Die Untersumme ist dagegen immer

$$\mathcal{U}_{\mathcal{J}}(f) = 0, \quad (6.2.22)$$

da in jedem Teilintervall irrationale Zahlen liegen. Wählen wir  $N$  groß genug, so erreichen wir für die zugehörige Zerlegung  $\mathcal{J}$  mit (6.2.21)

$$\mathcal{O}_{\mathcal{J}}(f) \leq 2\epsilon(b-a). \quad (6.2.23)$$

Da  $\epsilon > 0$  beliebig war, finden wir also immer eine Zerlegung mit  $\mathcal{O}_{\mathcal{J}}(f) < \epsilon$ , womit wir aus Proposition 6.2.5, *vi.*), und Satz 6.2.6 die Riemann-Integrierbarkeit folgern. Es gilt entsprechend

$$\int_a^b f(x) dx = 0. \quad (6.2.24)$$

Man beachte, dass  $f$  sicherlich nicht monoton ist. Es gibt vielmehr keine Zahlen  $c, d \in [a, b]$  mit  $c < d$ , so dass  $f|_{[c,d]}$  monoton ist. Damit ist die Riemann-Integrierbarkeit von  $f$  keine Folge aus Korollar 6.2.8. Man beachte, dass die Funktion  $\chi_{\mathbb{Q}}$  nirgends stetig ist, während  $f$  nur abzählbar viele Unstetigkeitsstellen, nämlich bei  $\mathbb{Q} \setminus \{0\}$ , besitzt.

Als vorerst letztes allgemeines Resultat zu Riemann-Integralen wollen wir die Integrabilität über Teilintervalle studieren. Hier erhält man folgendes wichtige Resultat:

**Proposition 6.2.10** *Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $f: [a, b] \rightarrow \mathbb{K}$  eine Funktion.*

*i.) Die Funktion  $f$  ist genau dann Riemann-integrierbar, wenn die Einschränkungen  $f|_{[\alpha, \beta]}: [\alpha, \beta] \rightarrow \mathbb{K}$  für alle  $\alpha, \beta \in [a, b]$  mit  $\alpha < \beta$  Riemann-integrierbar sind.*

*ii.) Ist  $c \in (a, b)$  und  $f \in \mathcal{R}([a, b], \mathbb{K})$ , so gilt*

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (6.2.25)$$

BEWEIS: Ist  $f|_{[\alpha, \beta]} \in \mathcal{R}([\alpha, \beta], \mathbb{K})$  für alle  $\alpha, \beta$ , so auch für  $\alpha = a$  und  $\beta = b$ , womit eine Richtung in *i.)* trivial ist. Sei also  $f \in \mathcal{R}([a, b], \mathbb{K})$ . Ohne Einschränkung können wir  $\mathbb{K} = \mathbb{R}$  betrachten. Zu  $\epsilon > 0$  wählen wir eine Zerlegung  $\mathcal{J}$  von  $[a, b]$  mit

$$\mathcal{O}_{\mathcal{J}}(f) - \mathcal{U}_{\mathcal{J}}(f) < \epsilon.$$

Dann verfeinern wir  $\mathcal{J}$ , indem wir als neue Zwischenpunkte  $\alpha$  und  $\beta$  hinzunehmen. Damit gilt also

$$a = t_0 < t_1 < \cdots < t_n = \alpha < t_{n+1} < \cdots < t_{n+M} = \beta < \cdots < t_N = b$$

mit geeigneten  $n, M \in \mathbb{N}_0$ . Insbesondere erhalten wir eine Dreiteilung

$$\begin{aligned} \mathcal{J}' &= \{t_0 < \cdots < t_n\}, \\ \mathcal{J}'' &= \{t_n < \cdots < t_{n+M}\}, \\ \mathcal{J}''' &= \{t_{n+M} < \cdots < t_N\}, \end{aligned}$$

wobei wir im Falle  $\alpha = a$  oder  $\beta = b$  auch leere Stücke  $\mathcal{J}'$  beziehungsweise  $\mathcal{J}'''$  zulassen. Auf diese Weise erhalten wir Zerlegungen der Intervalle  $[a, \alpha]$ ,  $[\alpha, \beta]$  und  $[\beta, b]$  mit der Eigenschaft

$$\mathcal{O}_{\mathcal{J}}(f) = \mathcal{O}_{\mathcal{J}'}(f|_{[a, \alpha]}) + \mathcal{O}_{\mathcal{J}''}(f|_{[\alpha, \beta]}) + \mathcal{O}_{\mathcal{J}'''}(f|_{[\beta, b]})$$

und

$$\mathcal{U}_{\mathcal{J}}(f) = \mathcal{U}_{\mathcal{J}'}(f|_{[a, \alpha]}) + \mathcal{U}_{\mathcal{J}''}(f|_{[\alpha, \beta]}) + \mathcal{U}_{\mathcal{J}'''}(f|_{[\beta, b]})$$

Da die Obersummen immer größer oder gleich den zugehörigen Untersummen sind, folgt

$$\begin{aligned} & \underbrace{\mathcal{O}_{\mathcal{J}'}(f|_{[a, \alpha]}) - \mathcal{U}_{\mathcal{J}'}(f|_{[a, \alpha]})}_{\geq 0} + \underbrace{\mathcal{O}_{\mathcal{J}''}(f|_{[\alpha, \beta]}) - \mathcal{U}_{\mathcal{J}''}(f|_{[\alpha, \beta]})}_{\geq 0} + \underbrace{\mathcal{O}_{\mathcal{J}'''}(f|_{[\beta, b]}) - \mathcal{U}_{\mathcal{J}'''}(f|_{[\beta, b]})}_{\geq 0} \\ &= \mathcal{O}_{\mathcal{J}}(f) - \mathcal{U}_{\mathcal{J}}(f) \\ &< \epsilon. \end{aligned}$$

Damit gilt aber insbesondere auch  $\mathcal{O}_{\mathcal{J}''}(f|_{[\alpha, \beta]}) - \mathcal{U}_{\mathcal{J}''}(f|_{[\alpha, \beta]}) < \epsilon$ , womit nach Satz 6.2.6 die Einschränkung  $f|_{[\alpha, \beta]}$  Riemann-integrierbar ist. Der zweite Teil ist nun einfach: Zu jeder Zerlegung  $\mathcal{J}_0$  von  $[a, b]$  können wir eine Verfeinerung  $\mathcal{J}$  bilden, indem wir den Punkt  $c$  als Zwischenpunkt hinzunehmen. Dies liefert wieder eine Zerlegung  $\mathcal{J}'$  von  $[a, c]$  und eine Zerlegung  $\mathcal{J}''$  von  $[c, b]$  mit  $\mathcal{J} = \mathcal{J}' \cup \mathcal{J}''$ . Sind nun  $\Xi = \Xi' \cup \Xi''$  Zwischenpunkte in den entsprechenden Teilintervallen, so gilt

$$\sum_{(\mathcal{J}, \Xi)}(f) = \sum_{(\mathcal{J}', \Xi')}(f|_{[a, c]}) + \sum_{(\mathcal{J}'', \Xi'')}(f|_{[c, b]}). \quad (*)$$

Da wir nach *i.*) wissen, dass  $f|_{[a, c]}$  und  $f|_{[c, b]}$  wieder Riemann-integrierbar sind, können wir in  $(*)$  Limites bilden und erhalten direkt (6.2.25). Hierbei ist wichtig, dass wir zu jeder markierten Zerlegung  $(\mathcal{J}_0, \Xi_0)$  solche Verfeinerungen finden können.  $\square$

Die Rechenregel aus (6.2.25) legt nun folgende Konvention nahe. Für  $f \in \mathcal{R}([a, b], \mathbb{K})$  und  $\alpha, \beta \in [a, b]$  definiert man

$$\int_{\alpha}^{\beta} f(x) \, dx = \begin{cases} \int_{\alpha}^{\beta} f(x) \, dx & \text{für } \alpha < \beta \\ 0 & \text{für } \alpha = \beta \\ -\int_{\beta}^{\alpha} f(x) \, dx & \text{für } \beta < \alpha \end{cases} \quad (6.2.26)$$

Sind nun  $\alpha, \beta, \gamma \in [a, b]$ , so folgt

$$\int_{\alpha}^{\beta} f(x) \, dx = \int_{\alpha}^{\gamma} f(x) \, dx + \int_{\gamma}^{\beta} f(x) \, dx \quad (6.2.27)$$

uneingeschränkt, egal welche der Zahlen größer ist. Dies überlegt man sich durch eine Fallunterscheidung aus (6.2.25). Wichtig ist natürlich, dass  $f$  überhaupt auf allen Intervallen Riemann-integrierbar ist, was durch Proposition 6.2.10 gewährleistet ist, siehe auch Übung 6.8.2.

Auch wenn die Netzkonvergenz die elegantere Weise ist, Riemann-Integrale zu definieren, bieten konkretere Folgen von Riemann-Summen ein oftmals einfacheres Bestimmen von Riemann-Integralen. Wir zeigen daher folgende Proposition:

**Proposition 6.2.11** *Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall.*

*i.) Ist  $f \in \mathcal{R}([a, b], \mathbb{K})$  und  $(\mathcal{J}_n)_{n \in \mathbb{N}}$  eine Folge von Zerlegungen von  $[a, b]$  mit  $|\mathcal{J}_n| \rightarrow 0$ , so gilt für jede Wahl  $(\Xi_n)_{n \in \mathbb{N}}$  von Markierungen*

$$\lim_{n \rightarrow \infty} \sum_{(\mathcal{J}_n, \Xi_n)}(f) = \int_a^b f(x) \, dx. \quad (6.2.28)$$



ii.) Ist umgekehrt  $f \in \text{Abb}([a, b], \mathbb{K})$  eine Funktion mit der Eigenschaft, dass für jede Folge  $(J_n)_{n \in \mathbb{N}}$  von Zerlegungen von  $[a, b]$  mit  $|J_n| \rightarrow 0$  die Folge der Riemann-Summen konvergiert, so gilt  $f \in \mathcal{R}([a, b], \mathbb{K})$ .

BEWEIS: Wie zuvor genügt es,  $\mathbb{K} = \mathbb{R}$  zu betrachten. Sei  $\epsilon > 0$  vorgegeben und  $J$  eine Zerlegung von  $[a, b]$  mit

$$\mathcal{O}_J(f) - \mathcal{U}_J(f) < \frac{\epsilon}{3}.$$

nach Proposition 6.2.5, vi.). Wir benutzen erneut Satz 6.2.6 zur Charakterisierung von Riemann-integrierbaren Funktionen als Darboux-integrierbare Funktionen. Die Anzahl der Teilintervalle in  $J$  sei  $N$ . Wir betrachten dann

$$\delta = \frac{\epsilon}{3N\|f\|_\infty},$$

wobei wir  $\|f\|_\infty \neq 0$  annehmen dürfen, um Trivialitäten zu vermeiden. Sei nun  $\tilde{J}$  eine Zerlegung mit

$$|\tilde{J}| < \delta.$$

Wir wählen zu  $\tilde{J}$  und  $J$  die gemeinsame Verfeinerung  $J' = \tilde{J} \cup J$ . Diese hat höchstens  $N$  zusätzliche Intervalle gegenüber  $\tilde{J}$ . Im Beweis von Proposition 6.2.5, i.), haben wir gesehen, dass das Einfügen eines Zwischenpunkts die Obersumme um höchstens  $\|f\|_\infty$  mal die Intervalllänge verringert, siehe auch Abbildung 6.5. Da die Intervalllängen von  $\tilde{J}$  alle kleiner  $\delta$  sind, und wir höchstens  $N$  neue Zwischenpunkte einfügen, folgt

$$\mathcal{O}_{\tilde{J}}(f) - \mathcal{O}_{J'}(f) < N\|f\|_\infty\delta = \frac{\epsilon}{3}.$$

Mit dem gleichen Argument für die Untersummen finden wir auch

$$\mathcal{U}_{J'}(f) - \mathcal{U}_{\tilde{J}}(f) < \frac{\epsilon}{3}.$$

Insgesamt zeigt dies

$$\mathcal{O}_{\tilde{J}}(f) - \mathcal{U}_{\tilde{J}}(f) = \mathcal{O}_{\tilde{J}}(f) - \mathcal{O}_{J'}(f) + \mathcal{O}_{J'}(f) - \mathcal{U}_{J'}(f) + \mathcal{U}_{J'}(f) - \mathcal{U}_{\tilde{J}}(f) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Ist also  $(J_n, \Xi_n)_{n \in \mathbb{N}}$  eine Folge von markierten Zerlegungen mit  $|J_n| \rightarrow 0$ , so gibt es ein  $n_0 \in \mathbb{N}$  mit  $|J_n| < \delta$  für  $n \geq n_0$ . Für diese  $n$  gilt dann

$$\mathcal{O}_{J_n}(f) - \mathcal{U}_{J_n}(f) < \epsilon, \quad (*)$$

und daher auch

$$\mathcal{U}_{J_n}(f) \leq \sum_{(J_n, \Xi_n)}(f) \leq \mathcal{O}_{J_n}(f). \quad (**)$$

Da  $f$  Riemann-integrierbar ist, gilt für jede Zerlegung

$$\mathcal{U}_J(f) \leq \int_a^b f(x) \, dx \leq \mathcal{O}_J(f),$$

womit aus (\*) und (\*\*) sofort

$$\left| \sum_{(J_n, \Xi_n)}(f) - \int_a^b f(x) \, dx \right| < \epsilon$$

folgt. Dies zeigt aber die Konvergenz (6.2.28).

Sei umgekehrt  $f$  eine Abbildung, die den Anforderungen aus ii.) genügt. Da die Folge der Riemann-Summen für jede Folge von markierten Zerlegungen  $(J_n, \Xi_n)_{n \in \mathbb{N}}$  mit  $|J_n| \rightarrow 0$  konvergiert, konvergieren

Abbildung 6.5: Differenz der Obersummen bei einem neuen Zwischenpunkt  $\tau$  ist durch den blauen Bereich gegeben und höchstens  $\|f\|_\infty(t_{i+1} - t_i)$ .

diese Folgen notwendigerweise gegen denselben Grenzwert, da man zwei solche Folgen mischen kann: Sind nämlich  $(J_n)_{n \in \mathbb{N}}$  und  $(\tilde{J}_n)_{n \in \mathbb{N}}$  zwei solche Folgen, so erfüllt auch die gemischte Folge

$$\bar{J}_n = \begin{cases} J_k & \text{falls } n = 2k \\ \tilde{J}_k & \text{falls } n = 2k + 1 \end{cases}$$

die Anforderung  $|\bar{J}_n| \rightarrow 0$ . Also konvergieren auch hier die zugehörigen Riemann-Summen. Da die vormaligen Folgen aber Teilfolgen dieser sind, konvergieren sie gegen den gleichen Grenzwert. Damit folgt, dass für diesen Grenzwert  $S$  gilt, dass zu jedem  $\epsilon > 0$  ein  $\delta > 0$  existiert, so dass

$$\left| \sum_{(J, \Xi)}(f) - S \right| < \epsilon \quad (\star)$$

gilt, wenn  $|J| < \delta$  ist. Dies ist aber gerade die Netzkonvergenz

$$\lim_{(J, \Xi)} \sum_{(J, \Xi)}(f) = S$$

bezüglich der Richtung  $\succcurlyeq$  aus Definition 6.1.3, siehe auch Übung ?? zur Charakterisierung von Netzkonvergenz mittels Folgenkonvergenz. Sei also nun  $\epsilon > 0$  vorgegeben und  $\delta > 0$  mit  $(\star)$  gefunden. Ist  $J_0$  eine Zerlegung mit  $|J_0| < \delta$ , so ist für jede andere Zerlegung  $J$  mit  $J \succcurlyeq J_0$  erst recht  $|J| \leq |J_0|$  und damit auch  $J \succcurlyeq J_0$  wie schon in Lemma lemma:FeinheitUndVerfeinerungLiefernRichtungen,  $i.$ ), gezeigt. Daher gilt  $(\star)$  für  $J$  ebenfalls, was gerade die Existenz des Grenzwertes  $S$  der markierten Riemann-Summen bezüglich der Richtung  $\succcurlyeq$  zeigt. Also ist  $f \in \mathcal{R}([a, b], \mathbb{K})$ .  $\square$

### 6.3 Riemann-Integral stetiger Funktionen

Im letzten Beispiel haben wir gesehen, dass sehr unstetige Funktionen durchaus noch Riemann-integrierbar sein können. In diesem Abschnitt wollen wir uns den stetigen Funktionen zuwenden, deren Riemann-Integrierbarkeit allgemein zu zeigen noch aussteht. Man beachte jedoch, dass wir mit Korollar 6.2.8 bereits viele wichtige Funktionen als Riemann-integrierbar identifiziert haben.

Mit dem folgenden Satz werden wir nicht nur die Riemann-Integrierbarkeit der stetigen Funktionen allgemein zeigen, sondern viele weitere Eigenschaften von  $\mathcal{R}([a, b], \mathbb{K})$  folgern können.

**Satz 6.3.1** ( $\mathcal{C} \circ \mathcal{R} \subseteq \mathcal{R}$ ) Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall. Sei weiter  $f \in \mathcal{R}([a, b], \mathbb{R})$  mit  $f([a, b]) \subseteq [c, d]$  und  $g \in \mathcal{C}([c, d], \mathbb{R})$  gegeben. Dann gilt  $g \circ f \in \mathcal{R}([a, b], \mathbb{R})$ .

BEWEIS: Wir geben ein  $\epsilon > 0$  vor. Da  $g$  auf dem kompakten Intervall nach Satz 3.6.5 sogar gleichmäßig stetig ist, finden wir ein  $\delta > 0$  mit

$$|g(x) - g(y)| < \epsilon$$

für alle  $x, y \in [c, d]$  mit  $|x - y| < \delta$ . Ohne Einschränkung können wir  $\delta < \epsilon$  annehmen. Da  $f \in \mathcal{R}([a, b], \mathbb{R})$  Riemann-integrierbar ist, finden wir eine Zerlegung

$$J = \{a = t_0 < t_1 < \dots < t_N = b\}$$

von  $[a, b]$  mit der Eigenschaft

$$\mathcal{O}_J(f) - \mathcal{U}_J(f) < \delta^2,$$

siehe Satz 6.2.6. Zur Abkürzung schreiben wir

$$\Delta_i(f) = \sup \left( f|_{[t_{i-1}, t_i]} \right) - \inf \left( f|_{[t_{i-1}, t_i]} \right)$$

und ebenso  $\Delta_i(g \circ f)$ . Insbesondere gilt dann

$$\mathcal{O}_J(f) - \mathcal{U}_J(f) = \sum_{i=1}^N \Delta_i(f)(t_i - t_{i-1})$$

und ebenso für  $g \circ f$ . Wir teilen die Zerlegung nun in

$$A = \{i \in \{1, \dots, N\} \mid \Delta_i(f) < \delta\}$$

und

$$B = \{i \in \{1, \dots, N\} \mid \Delta_i(f) \geq \delta\}$$

auf. Ist also  $i \in A$ , so gilt für alle  $x, y \in [t_{i-1}, t_i]$

$$|f(x) - f(y)| \leq \sup \left( f|_{[t_{i-1}, t_i]} \right) - \inf \left( f|_{[t_{i-1}, t_i]} \right) < \delta$$

und damit

$$|g(f(x)) - g(f(y))| < \epsilon$$

aufgrund der gleichmäßigen Stetigkeit von  $g$ . Damit folgt aber

$$\Delta_i(g \circ f) \leq \epsilon.$$

Ist dagegen  $i \in B$ , so gilt zumindest die grobe Abschätzung

$$\Delta_i(g \circ f) \leq 2\|g\|_\infty$$

durch die Supremumsnorm von  $g$ . Man beachte, dass  $g$  als stetige Funktion auf dem kompakten Intervall  $[c, d]$  ja beschränkt ist. Von diesen Indizes kann es aber nicht allzu viele geben, denn

$$\delta \sum_{i \in B} (t_i - t_{i-1}) \leq \sum_{i \in B} \Delta_i(f)(t_i - t_{i-1}) \leq \sum_{i=1}^N \Delta_i(f)(t_i - t_{i-1}) = \mathcal{O}_J(f) - \mathcal{U}_J(f) < \delta^2,$$

womit also

$$\sum_{i \in B} (t_i - t_{i-1}) < \delta$$

folgt. Für die Ober- und Untersummen von  $g \circ f$  erhalten wir daher

$$\begin{aligned} \mathcal{O}_J(g \circ f) - \mathcal{U}_J(g \circ f) &= \sum_{i \in A} \Delta_i(g \circ f)(t_i - t_{i-1}) + \sum_{i \in B} \Delta_i(g \circ f)(t_i - t_{i-1}) \\ &\leq \epsilon \sum_{i \in A} (t_i - t_{i-1}) + 2\|g\|_\infty \sum_{i \in B} (t_i - t_{i-1}) \\ &\leq \epsilon(b - a) + 2\|g\|_\infty \delta \\ &< \epsilon((b - a) + 2\|g\|_\infty). \end{aligned}$$

Da  $\epsilon > 0$  beliebig war, können wir Satz 6.2.6 zum Einsatz bringen, und aus der gezeigten Darboux-Integrabilität von  $g \circ f$  dann auch  $g \circ f \in \mathcal{R}([a, b], \mathbb{R})$  folgern.  $\square$

Wir wissen, dass  $f$  als Riemann-integrierbare Funktion auch beschränkt ist. Daher können wir ein Intervall  $[c, d]$  mit  $f([a, b]) \subseteq [c, d]$  immer finden, womit dies keine wesentliche Einschränkung an die Nützlichkeit des Satzes darstellt sondern lediglich sicherstellt, dass wir  $f$  mit  $g$  verknüpfen können. Wir erhalten aus Satz 6.3.1 eine Fülle von nützlichen Konsequenzen:

**Korollar 6.3.2** Für  $f \in \mathcal{R}([a, b], \mathbb{R})$  und  $g \in \mathcal{C}([c, d], \mathbb{C})$  mit  $f([a, b]) \subseteq [c, d]$  gilt  $g \circ f \in \mathcal{R}([a, b], \mathbb{C})$ .

BEWEIS: Nach Satz 6.3.1 gilt  $\operatorname{Re}(g) \circ f, \operatorname{Im}(g) \circ f \in \mathcal{R}([a, b], \mathbb{R})$ . Nach Proposition 6.1.7, ii.), folgt dann  $g \circ f \in \mathcal{R}([a, b], \mathbb{C})$ .  $\square$

**Korollar 6.3.3** Für jedes abgeschlossene Intervall  $[a, b] \subseteq \mathbb{R}$  gilt

$$\mathcal{C}([a, b], \mathbb{K}) \subseteq \mathcal{R}([a, b], \mathbb{K}). \quad (6.3.1)$$

BEWEIS: Die Funktion  $f: [a, b] \rightarrow [a, b]$  mit  $f(x) = x$  ist monoton, also nach Korollar 6.2.8 Riemann-integrierbar. Für  $g \in \mathcal{C}([a, b], \mathbb{K})$  gilt dann  $g = g \circ f \in \mathcal{R}([a, b], \mathbb{K})$  nach Satz 6.3.1 beziehungsweise Korollar 6.3.2.  $\square$

**Korollar 6.3.4** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall.

- i.) Für  $f \in \mathcal{R}([a, b], \mathbb{R})$  gilt  $f^2 \in \mathcal{R}([a, b], \mathbb{R})$ .
- ii.) Für  $f, g \in \mathcal{R}([a, b], \mathbb{K})$  gilt  $fg \in \mathcal{R}([a, b], \mathbb{K})$ .
- iii.) Für  $f \in \mathcal{R}([a, b], \mathbb{K})$  gilt  $|f| \in \mathcal{R}([a, b], \mathbb{R})$ .
- iv.) Für  $f, g \in \mathcal{R}([a, b], \mathbb{R})$  gilt  $\max(f, g), \min(f, g) \in \mathcal{R}([a, b], \mathbb{R})$ .
- v.) Ist  $f \in \mathcal{R}([a, b], \mathbb{R})$ , so gilt  $f_+, f_- \in \mathcal{R}([a, b], \mathbb{R})$ .
- vi.) Ist  $\alpha > 0$  und  $f \in \mathcal{R}([a, b], \mathbb{R})$  mit  $f(x) \geq 0$  für alle  $x \in [a, b]$ , so ist  $f^\alpha \in \mathcal{R}([a, b], \mathbb{R})$ .

BEWEIS: Da  $g(x) = x^2$  eine auf ganz  $\mathbb{R}$  stetige Funktion  $g \in \mathcal{C}(\mathbb{R}, \mathbb{R})$  definiert, folgt der erste Teil aus Satz 6.3.1. Für den zweiten Teil betrachtet man zunächst  $\mathbb{K} = \mathbb{R}$  und die *Polarisierungsidentität*

$$fg = \frac{1}{4}((f+g)^2 - (f-g)^2).$$

Da  $f+g, f-g \in \mathcal{R}([a, b], \mathbb{R})$  nach Proposition 6.1.6, ii.), und Quadrate nach i.) wieder in  $\mathcal{R}([a, b], \mathbb{R})$  liegen, gilt  $fg \in \mathcal{R}([a, b], \mathbb{R})$ . Ist dagegen  $\mathbb{K} = \mathbb{C}$ , wendet man den reellen Fall auf die vier Produkte der Real- und Imaginärteile von  $f$  und  $g$  an. Der dritte Teil folgt damit auch, da zunächst  $f\bar{f} \in \mathcal{R}([a, b], \mathbb{R})$  nach ii.), und  $\sqrt{f\bar{f}} = |f| \in \mathcal{R}([a, b], \mathbb{R})$  nach Satz 6.3.1 aus der Stetigkeit der Wurzel auf  $[0, \infty)$  folgt. Teil iv.) und v.) folgen dann aus den algebraischen Darstellungen von  $\max(f, g), \min(f, g), f_+$  und  $f_-$  mittels der Betragfunktion, siehe erneut Übung 3.9.22. Der letzte Teil folgt aus der Stetigkeit der Potenzfunktion  $g(x) = x^\alpha$  für  $x \in [0, \infty)$ .  $\square$

**Bemerkung 6.3.5** Man findet relativ leicht Gegenbeispiele von Riemann-integrierbaren Funktionen  $f$  und  $g$ , sodass  $g \circ f$  nicht länger Riemann-integrierbar ist. Etwas schwieriger ist es, eine stetige Funktion  $f$  und eine Riemann-integrierbare Funktion  $g$  zu finden, so dass  $g \circ f$  *nicht* Riemann-integrierbar ist. In diesem Sinne ist also das Resultat von Satz 6.3.1 nicht zu verbessern, und die Reihenfolge in „ $\mathcal{C} \circ \mathcal{R} \subseteq \mathcal{R}$ “ entscheidend.

Da mit  $f \in \mathcal{R}([a, b], \mathbb{K})$  auch  $|f| \in \mathcal{R}([a, b], \mathbb{R})$  gilt, können wir die Ungleichung aus Korollar 6.2.3 verbessern:

**Korollar 6.3.6** Sei  $f \in \mathcal{R}([a, b], \mathbb{K})$ . Dann gilt

$$\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f|(x) \, dx \leq \|f\|_\infty (b-a) \quad (6.3.2)$$

BEWEIS: Entscheidend ist, dass  $|f|$  überhaupt Riemann-integrierbar ist, siehe Korollar 6.3.4, *iii.*). Dann ist für jede markierte Zerlegung  $(\mathcal{J}, \Xi)$

$$|\sum_{(\mathcal{J}, \Xi)}(f)| = \left| \sum_{i=1}^N f(\xi_i)(t_i - t_{i-1}) \right| \leq \sum_{i=1}^N |f(\xi_i)|(t_i - t_{i-1}) = \sum_{(\mathcal{J}, \Xi)}(|f|).$$

Da die Limites beider Riemann-Summen nach Voraussetzung existieren, folgt die erste Ungleichung in (6.3.2). Die zweite ist dann einfach wieder Korollar 6.2.3 für  $|f|$ , da  $\|f\|_\infty = \||f|\|_\infty$  gilt.  $\square$

## 6.4 Der Hauptsatz und Stammfunktionen

Nachdem wir nun eine gute Vorstellung mit vielen Kriterien und Beispielen zur Riemann-Integrierbarkeit entwickelt haben, stellt sich nun die entscheidende Frage, wie man Riemann-Integrale effektiv berechnen kann: Es ist klar, dass die ursprüngliche Definition ebenso wie das Darboux-Integral wenig geeignet sind, die Integrale für konkret gegebene Funktionen auch explizit zu bestimmen, siehe etwa Übung 6.8.1 für einige einfache Fälle.

Wir benötigen also eine effektivere Methode. Dazu betrachten wir zunächst Riemann-Integrale mit variabler oberer Grenze:

**Proposition 6.4.1** *Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall.*

*i.) Ist  $f \in \mathcal{R}([a, b], \mathbb{K})$ , so ist die Funktion  $F: [a, b] \rightarrow \mathbb{K}$  mit*

$$F(x) = \int_a^x f(t) dt \tag{6.4.1}$$

*stetig mit*

$$|F(x) - F(y)| \leq \|f\|_\infty |x - y| \tag{6.4.2}$$

*für alle  $x, y \in [a, b]$ .*

*ii.) Ist  $f \in \mathcal{R}([a, b], \mathbb{K})$  bei  $x_0 \in [a, b]$  stetig, so ist die Funktion  $F$  aus (6.4.1) bei  $x_0$  differenzierbar mit Ableitung*

$$F'(x_0) = f(x_0). \tag{6.4.3}$$

BEWEIS: Zunächst ist  $f|_{[a, x]} \in \mathcal{R}([a, b], \mathbb{K})$  nach Proposition 6.2.10, *i.*), womit  $F$  überhaupt wohldefiniert ist. Wenn wir (6.4.2) zeigen, ist  $F$  sicherlich stetig: Diese Eigenschaft ist stärker als Stetigkeit und wird uns als Lipschitz-Stetigkeit in Abschnitt 7.3 noch im Detail begegnen. Nach Proposition 6.2.10, *ii.*), folgt dann mit (6.2.27)

$$F(x) - F(y) = \int_a^x f(t) dt - \int_a^y f(t) dt = \int_y^x f(t) dt,$$

womit (6.4.2) aus Korollar 6.2.3 folgt. Tatsächlich wird in (6.4.2) sogar nur das Supremum von  $|f|$  über  $[x, y]$  benötigt, die Ungleichung (6.4.2) besagt also, dass die Schranke  $\|f\|_\infty$  unabhängig von  $x$  und  $y$  gewählt ist. Für den zweiten Teil nehmen wir an,  $f$  sei bei  $x_0$  zudem stetig. Sei  $\epsilon > 0$  und  $\delta > 0$  passend gewählt, so dass

$$|f(x) - f(x_0)| < \epsilon$$

für  $|x - x_0| < \delta$ , also

$$\sup_{x \in B_\delta(x_0)} |f(x) - f(x_0)| \leq \epsilon.$$

Wir betrachten  $x \neq x_0$  mit  $x \in B_\delta(x_0)$ . Dann gilt

$$\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| = \frac{1}{|x - x_0|} \left| \int_a^x f(t) dt - \int_a^{x_0} f(t) dt - f(x_0)(x - x_0) \right|$$

$$\begin{aligned}
&= \frac{1}{|x - x_0|} \left| \int_{x_0}^x (f(t) - f(x_0)) \, dt \right| \\
&= \frac{1}{|x - x_0|} \int_{x_0}^x |f(t) - f(x_0)| \, dt \\
&\leq \frac{1}{|x - x_0|} \epsilon |x - x_0| \\
&= \epsilon
\end{aligned}$$

nach Korollar 6.2.3 sowie Proposition 6.1.6, i.), für die konstante Funktion  $f(x_0)$ . Damit ist aber die Existenz und der Wert des Limes

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

gezeigt, was die Differenzierbarkeit von  $F$  bei  $x_0$  mit  $F'(x_0) = f(x_0)$  bedeutet.  $\square$

Um nun den zweiten Teil der Proposition nutzbar zu machen, benötigen wir also zu einer Funktion  $f$  eine andere Funktion  $F$  mit  $F' = f$ . Diese nennt man eine Stammfunktion:

**Definition 6.4.2 (Stammfunktion)** Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen. Eine Funktion  $F: X \rightarrow \mathbb{K}$  heißt dann Stammfunktion einer Funktion  $f: X \rightarrow \mathbb{K}$ , falls  $F$  auf  $X$  differenzierbar ist und

$$F' = f \tag{6.4.4}$$

gilt.

**Proposition 6.4.3** Sei  $X = [a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $f: X \rightarrow \mathbb{K}$ . Besitzt  $f$  eine Stammfunktion  $F$  auf  $[a, b]$ , so ist jede andere Stammfunktion von  $f$  von der Form  $F + c$  mit  $c \in \mathbb{K}$ , und jede solche Funktion ist eine Stammfunktion.

BEWEIS: Es gilt zunächst  $(F + c)' = F' = f$ , da die Ableitung einer konstanten Funktion  $c$  verschwindet. Ist  $\mathbb{K} = \mathbb{C}$ , so ist  $\operatorname{Re}(F)$  eine Stammfunktion von  $\operatorname{Re}(f)$  und  $\operatorname{Im}(F)$  eine Stammfunktion von  $\operatorname{Im}(f)$ , da Ableiten linear ist. Es genügt also  $\mathbb{K} = \mathbb{R}$  zu betrachten. Sind dann  $F$  und  $\tilde{F}$  Stammfunktionen von  $f$ , so gilt  $(F - \tilde{F})' = F' - \tilde{F}' = f - f = 0$ , womit nach Korollar 5.3.7 die Funktion  $F - \tilde{F}$  konstant ist.  $\square$

Damit sind Stammfunktionen also eindeutig bis auf konstante Funktionen, sofern es überhaupt eine Stammfunktion gibt. Man beachte, dass die Annahme  $X = [a, b]$  tatsächlich wichtig ist, siehe auch Übung ???. Mit einer kontrollierten Sprachungenauigkeit werden wir also in Zukunft von *der* Stammfunktion  $F$  von  $f$  sprechen, auch wenn diese nur eindeutig bis auf eine additive Konstante ist.

Aus den Rechenregeln für Ableitungen erhalten wir sofort entsprechende Rechenregeln für Stammfunktionen:

**Proposition 6.4.4** Sei  $X$  eine nicht-leere Teilmenge der reellen oder komplexen Zahlen, und  $f, g: X \rightarrow \mathbb{K}$  Funktionen mit Stammfunktionen  $F, G: X \rightarrow \mathbb{K}$ .

- i.) Für alle  $z, w \in \mathbb{K}$  besitzt auch  $zf + wg$  eine Stammfunktion, nämlich  $zF + wG$ .
- ii.) Die Funktion  $fG + Fg$  besitzt eine Stammfunktion, nämlich  $FG$ .

BEWEIS: Dies ist eine direkte Folgerung aus Proposition 5.2.1.  $\square$

Auch die Kettenregel liefert eine entsprechende Regel für Stammfunktionen.

| Funktion                 |   | Stammfunktion                     |
|--------------------------|---|-----------------------------------|
| $z^n$                    | für $n \in \mathbb{Z} \setminus \{-1\}, z \in \mathbb{C}$ | $\frac{1}{n+1} z^{n+1}$           |
| $x^\alpha$               | für $\alpha \neq -1, x > 0$                               | $\frac{1}{\alpha+1} x^{\alpha+1}$ |
| $\frac{1}{x}$            | für $x \in \mathbb{R} \setminus \{0\}$                    | $\log( x )$                       |
| $e^z$                    | für $z \in \mathbb{C}$                                    | $e^z$                             |
| $\cos(z)$                | für $z \in \mathbb{C}$                                    | $\sin(z)$                         |
| $\sin(z)$                | für $z \in \mathbb{C}$                                    | $-\cos(z)$                        |
| $\frac{1}{1+x^2}$        | für $x \in \mathbb{R}$                                    | $\arctan(x)$                      |
| $\frac{1}{\sqrt{1-x^2}}$ | für $ x  < 1$   | $\arcsin(x)$                      |
| $\frac{1}{\cos(x)^2}$    | für $\cos(x) \neq 0$                                      | $\tan(x)$                         |
| $\frac{1}{\sin(x)^2}$    | für $\sin(x) \neq 0$                                      | $-\cot(x)$                        |
| $\log(x)$                | für $x > 0$   | $x \log(x) - x$                   |

Tabelle 6.1: Einige der wichtigsten Stammfunktionen

**Proposition 6.4.5** Sei  $f: X \rightarrow \mathbb{K}$  eine Funktion mit Stammfunktion  $F: \rightarrow Y$ , und  $g: Y \rightarrow \mathbb{K}$  eine Funktion mit Stammfunktion  $G: Y \rightarrow \mathbb{K}$ . Dann besitzt auch das Produkt  $(g \circ F)f: X \rightarrow \mathbb{K}$  eine Stammfunktion, nämlich  $G \circ F$ .

BEWEIS: Dies ist gerade Proposition 5.2.5. □

**Bemerkung 6.4.6** Das konkrete Auffinden einer Stammfunktion  $F$  zu einer gegebenen Funktion  $f$  ist oft schwierig. Hier ist *Raten* ein völlig legitimes Mittel, da die obigen Techniken aus Proposition 6.4.4 und Proposition 6.4.5 selten direkt zum Ziel führen. Hat man einen guten Kandidaten  $F$  geraten, ist die Verifikation  $F' = f$  mittels den Ableitungsregeln aus Abschnitt 5.2 typischerweise sehr einfach. Verschiedene Mathematikprogramme bieten hier gute Hilfestellungen an. Zudem gibt es Bücher mit tabellierten Stammfunktionen, wie etwa [1], siehe auch Tabelle 6.1 für die Stammfunktionen einiger wichtiger Funktionen.

Wir kommen nun zum Hauptsatz der Differential- und Integralrechnung, den wir aus Proposition 6.4.1 leicht erhalten:

**Satz 6.4.7 (Hauptsatz der Differential- und Integralrechnung)** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $f \in \mathcal{C}([a, b], \mathbb{K})$ . Dann besitzt  $f$  eine Stammfunktion  $F \in \mathcal{C}^1([a, b], \mathbb{K})$  und jede solche ist durch

$$F(x) = F(a) + \int_a^x f(t) dt \quad (6.4.5)$$

gegeben.

BEWEIS: Nach Proposition 6.4.1, ii.), ist (6.4.5) eine Stammfunktion. Hier ist wichtig, dass  $f$  als stetige Funktion nach Korollar 6.3.3 überhaupt Riemann-integrierbar ist. Dies klärt die Existenz. Ist nun  $\tilde{F}$  eine Stammfunktion und  $F(x) = \int_a^x f(t) dt$ , so gilt  $\tilde{F} = F + c$  mit einer Konstanten  $c \in \mathbb{K}$  nach Proposition 6.4.3. Da  $F(a) = 0$ , folgt  $\tilde{F}(a) = c$ , was  $\tilde{F}(x) = \tilde{F}(a) + \int_a^x f(t) dt$  zeigt. □

**Bemerkung 6.4.8** Die Voraussetzung der Stetigkeit von  $f$  im Hauptsatz ist wichtig, wie die folgenden beiden Beispiele zeigen.

Abbildung 6.6: Die stückweise stetige Funktion  $f$  ist Riemann-integrierbar, hat aber keine Stammfunktion.

Abbildung 6.7: Die differenzierbare Funktion  $F = x^2 \sin(\frac{1}{x^2})$  und ihre unbeschränkte Ableitung  $f = F'$ .

- i.) Es gibt Riemann-integrierbare Funktionen, die keine Stammfunktion besitzen, wie beispielsweise die stückweise stetige Funktion

$$f(x) = \begin{cases} 0 & \text{für } x \in [-1, 0] \\ 1 & \text{für } x \in (0, 1], \end{cases} \quad (6.4.6)$$

siehe auch Abbildung 6.6. Dass  $f$  Riemann-integrierbar ist, lässt sich leicht mit Ober- und Untersummen sehen, und kann auch aus Korollar 6.2.8 direkt gefolgert werden. Wäre  $F: [-1, 1] \rightarrow \mathbb{R}$  nun eine differenzierbare Funktion mit  $F' = f$ , so müsste die Ableitung  $F'(x) = f(x)$  nach dem Zwischenwertsatz für Ableitungen, siehe Satz 5.3.11, jeden Wert zwischen  $F'(-1) = 0$  und  $F'(1) = 1$  im Intervall  $(-1, 1)$  annehmen, was offenbar nicht der Fall ist. Also besitzt  $f$  keine Stammfunktion.

- ii.) Eine Funktion  $f$  mit Stammfunktion  $F$  muss nicht Riemann-integrierbar sein. Die Funktion

$$F(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \quad (6.4.7)$$

ist überall differenzierbar: Für  $x \in \mathbb{R} \setminus \{0\}$  ist dies mit den bekannten Differentiationsregeln klar. Für  $x_0 = 0$  betrachtet man

$$\frac{F(x) - F(0)}{x - 0} = \frac{x^2 \sin\left(\frac{1}{x^2}\right)}{x} = x \sin\left(\frac{1}{x^2}\right) \quad (6.4.8)$$

Da die Sinusfunktion beschränkt ist, folgt

$$\lim_{x \rightarrow 0} \frac{F(x) - F(0)}{x - 0} = 0. \quad (6.4.9)$$

Damit gilt also

$$F'(x) = f(x) = \begin{cases} 2x \sin\left(\frac{1}{x^2}\right) + 2\frac{1}{x} \cos\left(\frac{1}{x^2}\right) & x \neq 0 \\ 0 & \end{cases} \quad (6.4.10)$$

nach den üblichen Ableitungsregeln. Die Funktion  $f$  ist nun auf jedem Intervall  $[-a, a]$  mit  $a > 0$  unbeschränkt, und daher sicherlich nicht Riemann-integrierbar, siehe auch Abbildung ??.

Auch wenn der Hauptsatz stetige Funktionen  $f$  als Voraussetzung benötigt, deckt dieser Fall doch eine sehr große Bandbreite an Anwendungsmöglichkeiten ab. Der Hauptsatz wird so zum wichtigsten Werkzeug, um Integrale zu berechnen:

**Korollar 6.4.9** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $f \in \mathcal{C}([a, b], \mathbb{K})$ . Für jede Stammfunktion  $F \in \mathcal{C}^1([a, b], \mathbb{K})$  von  $f$  gilt

$$\int_{\alpha}^{\beta} f(x) dx = F(\beta) - F(\alpha) \quad (6.4.11)$$

für alle  $\alpha, \beta \in [a, b]$  mit  $\alpha < \beta$ .



BEWEIS: Ist  $F$  eine (nach Satz 6.4.7 existierende) Stammfunktion, so gilt

$$F(x) = F(a) + \int_a^x f(t) \, dt$$

und daher

$$F(\beta) - F(\alpha) = \int_a^\beta f(t) \, dt - \int_a^\alpha f(t) \, dt = \int_\alpha^\beta f(t) \, dt$$

nach Proposition 6.2.10, beziehungsweise (6.2.27).  $\square$

Mit diesem Ergebnis sind wir also in der Lage, Integrale auf einfache Weise zu berechnen, und müssen nicht länger auf die komplizierten Grenzübergänge, wie etwa in Übung 6.8.1, zurückgreifen. Wir schreiben auch

$$F(a) - F(b) = F(x) \Big|_a^b \quad (6.4.12)$$

zur Vereinfachung in (6.4.11). Aus der Produktregel für Stammfunktionen in Proposition 6.4.4 und der Kettenregel in Proposition 6.4.5 erhalten wir für Integrale die beiden wichtigen Regeln der *partiellen Integration* und der *Substitutionsregel*:

**Proposition 6.4.10** *Seien  $[a, b] \subseteq \mathbb{R}$  und  $f, g \in \mathcal{C}^1([a, b], \mathbb{K})$  gegeben. Dann gilt*

$$\int_a^b f(x)g'(x) \, dx = f(x)g(x) \Big|_a^b - \int_a^b f'(x)g(x) \, dx. \quad (6.4.13)$$

BEWEIS: Wir integrieren die Identität

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

über  $[a, b]$  und erhalten (6.4.13) aus Korollar 6.4.9, da alle Integranden voraussetzungsgemäß stetig sind.  $\square$

**Proposition 6.4.11** *Seien  $[a, b] \subseteq \mathbb{R}$  und  $[c, d] \subseteq \mathbb{R}$  mit  $f \in \mathcal{C}^1([a, b], \mathbb{R})$  derart, dass  $f([a, b]) \subseteq [c, d]$ , und  $g \in \mathcal{C}([c, d], \mathbb{K})$  gegeben. Dann gilt*

$$\int_a^b (g \circ f)(x)f'(x) \, dx = \int_{f(a)}^{f(b)} g(y) \, dy. \quad (6.4.14)$$

BEWEIS: Die Funktion  $g$  besitzt eine Stammfunktion  $G \in \mathcal{C}^1([c, d], \mathbb{K})$ . Es gilt

$$(G \circ f)'(x) = (g \circ f)(x)f'(x),$$

was wir über  $[a, b]$  integrieren können. Dies liefert

$$\int_a^b (g \circ f)(x)f'(x) \, dx = \int_a^b (G \circ f)'(x) \, dx = (G \circ f)(x) \Big|_a^b = G(f(b)) - G(f(a)) = \int_{f(a)}^{f(b)} g(y) \, dy$$

nach Korollar 6.4.9.  $\square$

Wir illustrieren diese Rechenregeln nun durch einige weitere Beispiele:

### Beispiel 6.4.12 (Explizite Integralberechnung)

i.) Wir betrachten die Funktionen  $f(x) = x$  und  $g'(x) = \cos(x)$ , also  $g(x) = \sin(x)$ . Dann gilt

$$\int_a^b x \cos(x) \, dx = x \sin(x) \Big|_a^b - \int_a^b \sin(x) \, dx = b \sin(b) - a \sin(a) - \sin(b) + \sin(a). \quad (6.4.15)$$

mittels partieller Integration nach Proposition 6.4.10. Genauso erhält man

$$\int_a^b x \sin(x) \, dx = -x \cos(x) \Big|_a^b + \int_a^b \cos(x) \, dx. \quad (6.4.16)$$

Etwas allgemeiner gilt

$$\int_a^b x^n \cos(x) \, dx = x^n \sin(x) \Big|_a^b - \int_a^b n x^{n-1} \sin(x) \, dx \quad (6.4.17)$$

und

$$\int_a^b x^n \sin(x) \, dx = -x^n \cos(x) \Big|_a^b + \int_a^b n x^{n-1} \cos(x) \, dx \quad (6.4.18)$$

für  $n \in \mathbb{N}$ . Induktiv lassen sich dann alle Integrale der Form (6.4.17) beziehungsweise (6.4.18) auf (6.4.15) und (6.4.16) zurückführen, da in jedem neuen Schritt  $n$  um eins verringert wird, siehe auch Übung 6.8.7.

ii.) Mit derselben Idee der partiellen Integration erhält man

$$\int_a^b x e^x \, dx = x e^x \Big|_a^b - \int_a^b e^x \, dx \quad (6.4.19)$$

und allgemeiner

$$\int_a^b x^n e^x \, dx = x^n e^x \Big|_a^b - \int_a^b n x^{n-1} e^x \, dx \quad (6.4.20)$$

für alle  $n \in \mathbb{N}$ , womit wieder alle solchen Integrale rekursiv gelöst werden können, siehe erneut Übung 6.8.7.

iii.) Zur Illustration der Substitutionsregel aus Proposition 6.4.11 betrachtet man eine Funktion  $f \in \mathcal{C}^1([a, b], \mathbb{R})$  mit  $f(x) \neq 0$  für alle  $x \in [a, b]$ . Da  $[a, b]$  zusammenhängend ist, gilt entweder  $f([a, b]) \subseteq \mathbb{R}^+$  oder  $f([a, b]) \subseteq \mathbb{R}^-$ . Wir betrachten dann die Funktion  $g(x) = \frac{1}{x}$  für  $x \in \mathbb{R} \setminus \{0\}$  und

$$\int_a^b \frac{f'(x)}{f(x)} \, dx = \int_{f(a)}^{f(b)} \frac{1}{y} \, dy = \log(|y|) \Big|_{f(a)}^{f(b)}, \quad (6.4.21)$$

wobei nur entweder  $f(b), f(a) \in \mathbb{R}^+$  oder  $f(b), f(a) \in \mathbb{R}^-$  auftritt, der Betrag in  $\log(|y|)$  daher entweder überflüssig ist, oder durch  $\log(-y)$  gegeben ist. Diese *logarithmische Ableitung* lässt sich also mit der Substitutionsregel schnell integrieren.

iv.) Da wir die Integrale als Fläche unter dem Graphen interpretieren wollen, können wir nun die Fläche einer Kreisscheibe zum ersten mal konkret berechnen: Wir betrachten  $r > 0$  sowie die Funktion

$$f(x) = \sqrt{r^2 - x^2} \quad (6.4.22)$$

auf dem Intervall  $[-r, r]$ . Der Graph von  $f$  ist dann ein Halbkreis mit Radius  $r$ , siehe Abbildung 6.8. Wir bestimmen das Integral durch zwei Substitutionen. Zuerst reskalieren wir  $x$  zu  $\xi = \frac{x}{r}$

$$\int_{-r}^r \sqrt{r^2 - x^2} \, dx = \int_{-1}^1 \sqrt{r^2 - r^2 \xi^2} r \, d\xi = r^2 \int_{-1}^1 \sqrt{1 - \xi^2} \, d\xi. \quad (6.4.23)$$

Abbildung 6.8: Die Fläche eines Halbkreises und dem Graph von  $f(x) = \sqrt{r^2 - x^2}$ .

Dies kann man geometrisch so verstehen, dass die Fläche sich unter Reskalieren des Radius mit dem Quadrat des Reskalierungsfaktors ändert, weil es eben eine Fläche ist. Es ist allgemein immer eine gute Strategie, bei Integralen zu „dimensionslosen“ Variablen überzugehen, wenn eine geometrische Interpretation der Parameter dies zulässt. Um das verbleibende Integral zu vereinfachen, betrachtet man die weitere Substitution  $\xi(\varphi) = \cos(\varphi)$ . Dann benötigt man die Grenzen  $-1 = \xi(\varphi_0) = \cos(\varphi_0)$ , also  $\varphi_0 = \pi$ , sowie  $1 = \xi(\varphi_1) = \cos(\varphi_1)$ , also  $\varphi_1 = 0$ . Weiter ist  $\xi'(\varphi) = -\sin(\varphi)$ , womit wir

$$\int_{-1}^1 \sqrt{1 - \xi^2} d\xi = \int_{\pi}^0 \sqrt{1 - \cos(\varphi)^2} (-\sin(\varphi)) d\varphi = \int_0^{\pi} \sin(\varphi)^2 d\varphi \quad (6.4.24)$$

erhalten, da auf  $[0, \pi]$  eben  $\sqrt{1 - \cos(\varphi)^2} = \sin(\varphi)$  gilt. Mit den Additionstheoremen für die Sinusfunktion finden wir

$$\sin(\varphi)^2 = \left(\frac{1}{2i}(e^{i\varphi} - e^{-i\varphi})\right)^2 = -\frac{1}{4}(e^{2i\varphi} - 2 + e^{-2i\varphi}) = \frac{-\cos(2\varphi) + 1}{2} \quad (6.4.25)$$

als Spezialfall der allgemeinen Additionstheoreme aus Satz 4.3.2, *vii.*). Damit erhalten wir

$$\int_0^{\pi} \sin(\varphi)^2 d\varphi = \int_0^{\pi} \frac{1 - \cos(2\varphi)}{2} d\varphi = \left(\frac{\varphi}{2} - \frac{1}{2}\sin(2\varphi)\right)\Big|_0^{\pi} = \frac{\pi}{2}, \quad (6.4.26)$$

womit die Fläche  $A(r)$  einer Kreisscheibe vom Radius  $r$  also erwartungsgemäß

$$A(r) = \pi r^2 \quad (6.4.27)$$

ist. Diese Rechnung zeigt insbesondere, dass die recht unanschauliche Definition der Sinus- und Kosinusfunktion durch die komplexe Exponentialfunktion, und die anschließende Definition von  $\pi$  über die Nullstellen von  $\cos$ , tatsächlich die „richtige“ Wahl war und die elementaren Vorstellungen der Kreisgeometrie erfüllt.

Abschließend sollte man sagen, dass eine gewisse Vertrautheit mit den Tricks der Integrations-techniken nützlich ist, aber bei weitem nicht mehr den Stellenwert in der Mathematik besitzt wie noch vor wenigen Jahrzehnten. Computerprogramme haben hier die umfangreichen Tabellenwerke [] abgelöst und sind ihnen mittlerweile so weit überlegen, dass im Zweifelsfall ihre Hilfe angenommen werden sollte, anstatt Integrale selbst von Hand zu berechnen.

Als vorerst letzte Anwendung des Hauptsatzes wollen wir einen alternativen Beweis der Taylor-Formel mit Restglied geben, wobei wir allerdings etwas stärkere Voraussetzungen an die Stetigkeit der Ableitungen stellen müssen. In typischen Anwendungen ist dies aber selten ein wesentliches Problem, hat aber den Vorteil, erheblich einfacher in der Handhabung zu sein. Als Motivation betrachten wir wie in Satz 5.6.7 ein offenes Intervall  $X \subseteq \mathbb{R}$  mit  $x_0 \in X$  und eine Funktion  $f \in \mathcal{C}^1(X, \mathbb{K})$ . Dann gilt

$$f(x_0 + h) = f(x_0) + \int_0^h \frac{df}{dx}(x_0 + \xi) d\xi \quad (6.4.28)$$

nach dem Hauptsatz der Differential- und Integralrechnung. Damit ist aber das nullte Restglied der Taylor-Entwicklung durch

$$R_{0,x_0}(f)(h) = \int_0^h \frac{df}{dx}(x_0 + \xi) d\xi \quad (6.4.29)$$

gegeben. Dies liefert eine explizitere Form des Restgliedes für eine  $\mathcal{C}^1$ -Funktion als die allgemeinere Version, die nur  $\mathcal{C}^0$  benötigt. Die Stetigkeitseigenschaft

$$\lim_{h \rightarrow 0} R_{0,x_0}(f)(h) = 0 \quad (6.4.30)$$

ist dafür nun leichter zu sehen: Da  $f'$  noch stetig ist, können wir das Integral als

$$|R_{0,x_0}(f)(h)| = \left| \int_0^h f'(x_0 + \xi) d\xi \right| \leq \|f'\|_\infty |h| \quad (6.4.31)$$

abschätzen, womit (6.4.30) trivial wird. Die Idee der folgenden Version der Taylor-Entwicklung ist nun, durch fortlaufende partielle Integration auch für die höheren Restglieder Integralformeln zu finden:

**Satz 6.4.13 (Taylor-Entwicklung mit Restglied II)** *Sei  $X \subseteq \mathbb{R}$  ein offenes Intervall mit  $x_0 \in X$ . Sei weiter  $f \in \mathcal{C}^k(X, \mathbb{K})$  mit  $k \in \mathbb{N} \cup \{\infty\}$  und  $n \in \mathbb{N}_0$  mit  $n+1 \leq k$ . Dann gilt für das  $n$ -te Restglied  $R_{n,x_0}(f)(h)$  der Taylor-Entwicklung von  $f$  um  $x_0$*

$$R_{n,x_0}(f)(h) = \frac{1}{n!} \int_0^h f^{(n+1)}(x_0 + \xi)(h - \xi)^n d\xi. \quad (6.4.32)$$

BEWEIS: Den Fall  $n = 0$  haben wir bereits in (6.4.29) gezeigt. Wir zeigen (6.4.32) daher durch Induktion nach  $n$ . Es gilt

$$\begin{aligned} f(x_0 + h) &= f(x_0) + \cdots + \frac{1}{n!} f^{(n)}(x_0) h^n + R_{n,x_0}(f)(h) \\ &\stackrel{(a)}{=} f(x_0) + \cdots + \frac{1}{n!} f^{(n)}(x_0) h^n + \frac{1}{n!} \int_0^h f^{(n+1)}(x_0 + \xi)(h - \xi)^n d\xi \\ &\stackrel{(b)}{=} f(x_0) + \cdots + \frac{1}{n!} f^{(n)}(x_0) h^n + \frac{1}{n!} f^{(n+1)}(x_0 + \xi) \frac{-1}{n+1} (h - \xi)^{n+1} \Big|_0^h \\ &\quad - \frac{1}{n!} \int_0^h f^{(n+2)}(x_0 + \xi) \frac{-1}{n+1} (h - \xi)^{n+1} d\xi \\ &= f(x_0) + \cdots + \frac{1}{n!} f^{(n)}(x_0) h^n + \frac{1}{(n+1)!} f^{(n+1)}(x_0) h^{n+1} \\ &\quad + \frac{1}{(n+1)!} \int_0^h f^{(n+2)}(x_0 + \xi)(h - \xi)^{n+1} d\xi, \end{aligned}$$

wobei wir in (a) die Induktionsvoraussetzung, und in (b) eine partielle Integration verwendet haben. Damit ist (6.4.32) aber gezeigt.  $\square$

Wieder können wir den Grenzwert (5.6.19) leichter aus (6.4.32) erhalten. Es gilt

$$\begin{aligned} |R_{n,x_0}(f)(h)| &= \frac{1}{n!} \left| \int_0^h f^{(n+1)}(x_0 + \xi)(h - \xi)^n d\xi \right| \\ &= \frac{1}{n!} \left| \int_0^1 f^{(n+1)}(x_0 + h\tau)(h - h\tau)^n d\tau \right| \\ &\leq \frac{|h|^{n+1}}{n!} \int_0^1 |f^{(n+1)}(x_0 + h\tau)|(1 - \tau)^n d\tau \end{aligned} \quad (6.4.33)$$

dank der einfachen Substitution  $\xi = h\tau$  durch Reskalierung. Da  $|1 - \tau| \leq 1$  für  $\tau \in [0, 1]$  können wir das Restglied weiter als

$$|R_{n,x_0}(f)(h)| \leq \frac{|h|^{n+1}}{n!} \|f^{(n+1)}\|_\infty \quad (6.4.34)$$

abschätzen, wobei wir für die Supremumsnorm von  $f^{(n+1)}$  nur die Werte  $x_0 + h\tau$  mit  $\tau \in [0, 1]$  verwenden müssen. Insbesondere bilden diese Punkte ein Kompaktum, weshalb  $\|f^{(n+1)}\|_\infty$  für eine  $\mathcal{C}^{n+1}$ -Funktion tatsächlich endlich ist. Dies liefert eine weitere Abschätzung des Restglied, die etwas stärkere Voraussetzungen benötigt als die in Satz 5.6.7, *ii.*), dafür aber die stärkere Abschätzung (6.4.34) liefert.

## 6.5 Ungleichungen und Mittelwertsätze für Riemann-Integrale

Die Dreiecksungleichung

$$\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx \quad (6.5.1)$$

für Riemann-integrierbare Funktionen  $f \in \mathcal{R}([a, b], \mathbb{K})$  aus Korollar 6.3.6 kann man als Analogon und direkte Folge der Dreiecksungleichung für Summen verstehen. Es stellt sich daher die Frage, ob wir auch die anderen fundamentalen Ungleichungen für Summen aus Abschnitt 5.4 auf Riemann-Integrale übertragen können. Dies ist tatsächlich der Fall. Wir beginnen mit der *Hölder-Ungleichung*.

**Proposition 6.5.1** *Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $p, q \in (0, \infty)$ , so dass  $\frac{1}{p} + \frac{1}{q} = 1$ . Dann gilt die Hölder-Ungleichung*

$$\int_a^b |f(x)g(x)| \, dx \leq \left( \int_a^b |f(x)|^p \, dx \right)^{\frac{1}{p}} \left( \int_a^b |g(x)|^q \, dx \right)^{\frac{1}{q}} \quad (6.5.2)$$

für alle  $f, g \in \mathcal{R}([a, b], \mathbb{K})$ .

BEWEIS: Aus Satz 6.3.1 beziehungsweise Korollar 6.3.4 folgt, dass alle beteiligten Integranden tatsächlich wieder Riemann-integrierbar sind. Sei  $(\mathcal{J}, \Xi)$  eine markierte Zerlegung von  $[a, b]$ . Dann gilt

$$\begin{aligned} \sum_{(\mathcal{J}, \Xi)} (|fg|) &= \sum_{i=1}^N |f(\xi_i)g(\xi_i)|(t_i - t_{i-1}) \\ &= \sum_{i=1}^N |f(\xi_i)|(t_i - t_{i-1})^{\frac{1}{p}} |g(\xi_i)|(t_i - t_{i-1})^{\frac{1}{q}} \\ &\leq \left( \sum_{i=1}^N |f(\xi_i)|^p (t_i - t_{i-1}) \right)^{\frac{1}{p}} \left( \sum_{i=1}^N |g(\xi_i)|^q (t_i - t_{i-1}) \right)^{\frac{1}{q}} \\ &= \left( \sum_{(\mathcal{J}, \Xi)} (|f|^p) \right)^{\frac{1}{p}} \left( \sum_{(\mathcal{J}, \Xi)} (|g|^q) \right)^{\frac{1}{q}} \end{aligned}$$

nach der Hölder-Ungleichung für Summen, wobei wir entscheidend  $\frac{1}{p} + \frac{1}{q} = 1$  benutzt haben. Bilden der existierenden Netz-Limites unter Verwendung der Stetigkeit der Abbildung  $y \mapsto y^{\frac{1}{p}}$  sowie  $y \mapsto y^{\frac{1}{q}}$  auf der rechten Seite zeigt dann die gewünschte Ungleichung für die Integrale.  $\square$

Der Fall  $p = q = 2$  ist wie immer von besonderer Bedeutung und liefert die Cauchy-Schwarz-Ungleichung

$$\left( \int_a^b |f(x)g(x)| \, dx \right)^2 \leq \int_a^b |f(x)|^2 \, dx \int_a^b |g(x)|^2 \, dx \quad (6.5.3)$$

für Riemann-integrierbare Funktionen  $f, g \in \mathcal{R}([a, b], \mathbb{K})$ . Diese können wir auch auf alternative Weise erhalten und interpretieren:

**Proposition 6.5.2** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall. Durch

$$\langle f, g \rangle = \int_a^b \overline{f(x)} g(x) \, dx \quad (6.5.4)$$

für  $f, g \in \mathcal{R}([a, b], \mathbb{K})$  wird ein positiv semidefinites inneres Produkt auf  $\mathcal{R}([a, b], \mathbb{K})$  definiert, welches auf  $\mathcal{C}([a, b], \mathbb{K})$  sogar positiv definit ist.

BEWEIS: Wieder garantiert Korollar 6.3.4, dass  $\overline{f}g$  überhaupt Riemann-integrierbar ist. Damit ist (6.5.4) definiert. Die  $\mathbb{K}$ -Linearität im zweiten Argument ist klar nach der Linearität des Integrals. Ebenso erhalten wir die Symmetrie  $\langle f, g \rangle = \overline{\langle g, f \rangle}$  aus der Kompatibilität des Integrals mit komplexer Konjugation. Interessanter ist die Positivität. Hier verwenden wir Proposition 6.1.7, *iv.*), für

$$\langle f, f \rangle = \int_a^b |f(x)|^2 \, dx \geq 0,$$

da  $|f(x)|^2 \geq 0$  für alle  $x \in [a, b]$ . Sei nun zudem  $f \in \mathcal{C}([a, b], \mathbb{K})$  und ungleich der Nullfunktion. Aufgrund der Stetigkeit gibt es dann nicht nur einen Punkt  $x_0 \in [a, b]$  mit  $f(x_0) \neq 0$  sondern sogar ein ganzes Intervall  $[c, d] \subseteq [a, b]$  mit  $c < d$  und  $f(x) \neq 0$  für  $x \in [c, d]$ . Da die stetige Funktion  $|f|^2$  auf dem kompakten Intervall  $[c, d]$  ein Minimum annimmt, kann dieses wegen  $f(x) \neq 0$  für  $x \in [c, d]$  nicht 0 sein. Es gibt daher ein  $\epsilon > 0$  mit  $|f(x)|^2 \geq \epsilon > 0$  für alle  $x \in [c, d]$ . Die stückweise definierte, stückweise konstante, Funktion

$$g(x) = \begin{cases} \epsilon & x \in [c, d] \\ 0 & x \in [a, b] \setminus [c, d] \end{cases}$$

ist Riemann-integrierbar mit

$$\int_a^b g(x) \, dx = \epsilon(d - c) > 0.$$

Es folgt daher aus Proposition 6.1.7, *iii.*), dass

$$\int_a^b |f(x)|^2 \, dx \geq \int_a^b g(x) \, dx = \epsilon(d - c) > 0,$$

da punktweise  $|f(x)|^2 \geq g(x)$  nach Konstruktion gilt. Dies zeigt aber  $\langle f, f \rangle > 0$ , womit (6.5.4) positiv definit auf  $\mathcal{C}([a, b], \mathbb{K})$  ist.  $\square$

Wir können daher die Ungleichung (6.5.3) tatsächlich als Cauchy-Schwarz-Ungleichung

$$\langle f, g \rangle \langle g, f \rangle \leq \langle f, f \rangle \langle g, g \rangle \quad (6.5.5)$$

des positiv semidefiniten inneren Produktes  $\langle \cdot, \cdot \rangle$  auf  $\mathcal{R}([a, b], \mathbb{K})$  interpretieren, siehe beispielsweise [8, Prop. 7.10 und Prop. 7.19].

Ebenso wie die Hölder-Ungleichung gilt auch die Minkowski-Ungleichung für Riemann-Integrale:

**Proposition 6.5.3** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $p \geq 1$ . Dann gilt die Minkowski-Ungleichung

$$\left( \int_a^b |f(x) + g(x)|^p \, dx \right)^{\frac{1}{p}} \leq \left( \int_a^b |f(x)|^p \, dx \right)^{\frac{1}{p}} + \left( \int_a^b |g(x)|^p \, dx \right)^{\frac{1}{p}} \quad (6.5.6)$$

für alle  $f, g \in \mathcal{R}([a, b], \mathbb{K})$ .

BEWEIS: Zunächst bemerken wir wie zuvor, dass alle Integranden nach Korollar 6.3.4 tatsächlich Riemann-integrierbar sind. Wie bei der Hölder-Ungleichung betrachten wir eine markierte Zerlegung  $(\mathcal{J}, \Xi)$  von  $[a, b]$ . Dann gilt

$$\begin{aligned} (\sum_{(\mathcal{J}, \Xi)} (|f + g|^p))^{\frac{1}{p}} &= \left( \sum_{i=1}^N |f(\xi_i) + g(\xi_i)|^p (t_i - t_{i-1}) \right)^{\frac{1}{p}} \\ &= \left( \sum_{i=1}^N \left| f(\xi_i)(t_i - t_{i-1})^{\frac{1}{p}} + g(\xi_i)(t_i - t_{i-1})^{\frac{1}{p}} \right|^p \right)^{\frac{1}{p}} \\ &\leq \left( \sum_{i=1}^N |f(\xi_i)|^p (t_i - t_{i-1}) \right)^{\frac{1}{p}} + \left( \sum_{i=1}^N |g(\xi_i)|^p (t_i - t_{i-1}) \right)^{\frac{1}{p}} \\ &= (\sum_{(\mathcal{J}, \Xi)} (|f|^p))^{\frac{1}{p}} + (\sum_{(\mathcal{J}, \Xi)} (|g|^p))^{\frac{1}{p}}, \end{aligned}$$

nach der Minkowski-Ungleichung für Summen. Da die Netz-Limites alle existieren und die Riemann-Integrale liefern, folgt (6.5.6).  $\square$

Als vorerst letzte Ungleichung für das Riemann-Integral betrachten wir das Analogon zur Jensen-Ungleichung aus Korollar 5.4.13. Hier wird das Verhalten geringfügig komplizierter, da die Länge des Integrationsintervalls explizit Eingang in die Ungleichung findet, und die Rolle der Exponenten vertauscht wird:

**Proposition 6.5.4** *Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $0 < p < q$ . Dann gilt die Jensen-Ungleichung*

$$\left( \frac{1}{b-a} \int_a^b |f(x)|^p dx \right)^{\frac{1}{p}} \leq \left( \frac{1}{b-a} \int_a^b |f(x)|^q dx \right)^{\frac{1}{q}} \quad (6.5.7)$$

für alle  $f \in \mathcal{R}([a, b], \mathbb{K})$ .

BEWEIS: Zunächst sind beide Integranden nach Korollar 6.3.4 wieder Riemann-integrierbar. Zudem ist die konstante Funktion 1 ebenfalls integrierbar. Wir wenden die Hölder-Ungleichung für die konjugierten Exponenten  $p' = \frac{q}{p}$  und  $\frac{1}{q'} = 1 - \frac{1}{p'} = \frac{q-p}{q}$  an. Da  $q > p > 0$  ist  $p' > 1$ , und ebenso ist  $q' = \frac{q}{q-p} > 1$ . Dann gilt

$$\begin{aligned} \int_a^b |f(x)|^p dx &= \int_a^b 1 |f(x)|^p dx \\ &\leq \left( \int_a^b 1^{q'} dx \right)^{\frac{1}{q'}} \left( \int_a^b (|f(x)|^p)^{p'} dx \right)^{\frac{1}{p'}} \\ &= \left( \int_a^b 1 dx \right)^{\frac{1}{q'}} \left( \int_a^b |f(x)|^q dx \right)^{\frac{p}{q}} \\ &= (b-a)^{\frac{q-p}{q}} \left( \int_a^b |f(x)|^q dx \right)^{\frac{p}{q}} \end{aligned}$$

nach der Hölder-Ungleichung aus Proposition 6.5.1. Nun bildet man die  $\frac{1}{p}$ -te Potenz dieser Ungleichung und erhält

$$\left( \int_a^b |f(x)|^p dx \right)^{\frac{1}{p}} \leq (b-a)^{\frac{q-p}{qp}} \left( \int_a^b |f(x)|^q dx \right)^{\frac{1}{q}}.$$

Schließlich gilt  $\frac{q-p}{qp} = \frac{1}{p} - \frac{1}{q}$ , womit (6.5.7) gezeigt ist.  $\square$

Man beachte, dass eine naive Ersetzung der Summen in Korollar 5.4.13 nicht die richtige Jensen-Ungleichung für Integrale liefert: Die Rolle der Indizes ist vertauscht und man hat die zusätzlichen Vorfaktoren abhängig von der Länge des Integrationsintervalls.

Wir zeigen als nächstes einen Mittelwertsatz für Riemann-Integrale:

**Proposition 6.5.5** *Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall. Für Riemann-integrierbare Funktionen  $f, g \in \mathcal{R}([a, b], \mathbb{R})$  mit  $g \geq 0$  gibt es ein  $\mu \in \mathbb{R}$  mit*

$$\int_a^b f(x)g(x) \, dx = \mu \int_a^b g(x) \, dx \quad (6.5.8)$$

mit  $\inf_{x \in [a, b]} f(x) \leq \mu \leq \sup_{x \in [a, b]} f(x)$ . Ist  $f$  zudem stetig, gibt es ein  $\xi \in [a, b]$  mit

$$\mu = f(\xi) \quad (6.5.9)$$

BEWEIS: Sei  $\alpha = \inf_{x \in [a, b]} f(x)$  und  $\beta = \sup_{x \in [a, b]} f(x)$ , dann gilt  $\alpha \leq f(x) \leq \beta$  für alle  $x \in [a, b]$  und entsprechend  $\alpha g(x) \leq f(x)g(x) \leq \beta g(x)$ , da nach Voraussetzung  $g \geq 0$ . Die Monotonie des Riemann-Integrals aus Proposition 6.1.7, iii.), zeigt dann

$$\alpha \int_a^b g(x) \, dx \leq \int_a^b f(x)g(x) \, dx \leq \beta \int_a^b g(x) \, dx. \quad (*)$$

Ist  $\int_a^b g(x) \, dx = 0$ , zeigen diese Ungleichungen, dass auch das Integral in der Mitte verschwindet. In diesem Fall erfüllt jede Zahl  $\mu$  die Bedingung, und im Fall  $f \in \mathcal{C}([a, b], \mathbb{R})$  kann diese als Funktionswert von  $f$  gewählt werden. Ist  $\int_a^b g(x) \, dx \neq 0$ , so ist

$$\mu = \frac{\int_a^b f(x)g(x) \, dx}{\int_a^b g(x) \, dx}$$

die eindeutige Lösung von (6.5.8). Die Ungleichung (\*) zeigt dann  $\mu \in [\alpha, \beta]$  wie gewünscht. Ist  $f \in \mathcal{C}([a, b], \mathbb{R})$ , so wird nach dem Zwischenwertsatz aus Korollar 3.5.11 jeder Funktionswert zwischen  $\alpha = \min_{x \in [a, b]} f(x)$  und  $\beta = \max_{x \in [a, b]} f(x)$  auch als Funktionswert angenommen. Hier verwenden wir zuerst, dass nach Korollar 3.5.8 das Infimum und Supremum als Minimum und Maximum angenommen werden.  $\square$

Einen wichtigen Spezialfall erhält man für  $g = 1$ . Wir nennen

$$\mu(f) = \frac{1}{b-a} \int_a^b f(x) \, dx \quad (6.5.10)$$

den *Mittelwert* von  $f$  auf  $[a, b]$ . Als Motivation für diese Begriffsbildung betrachtet man das arithmetischen Mittel

$$\text{AM}(f(\xi_1), \dots, f(\xi_N)) = \frac{1}{N} \sum_{i=1}^N f(\xi_i) = \frac{1}{b-a} \sum_{i=1}^N f(\xi_i) \frac{b-a}{N} = \frac{1}{b-a} \sum_{(\mathcal{J}, \Xi)}(f). \quad (6.5.11)$$

der Funktionswerte  $f(\xi_i)$  für eine äquidistante markierte Zerlegung  $(\mathcal{J}, \Xi)$  und bildet dann den Grenzwert  $N \rightarrow \infty$ , was auf der rechten Seite das Riemann-Integral liefert.

**Korollar 6.5.6** *Sei  $f \in \mathcal{R}([a, b], \mathbb{R})$ . Dann ist der Mittelwert  $\mu(f)$  die eindeutig bestimmte Zahl mit*

$$\int_a^b f(x) \, dx = \mu(f)(b-a). \quad (6.5.12)$$

Im Falle  $f \in \mathcal{C}([a, b], \mathbb{R})$  gibt es ein  $\xi \in [a, b]$  mit

$$\mu(f) = f(\xi). \quad (6.5.13)$$



Abbildung 6.9: Der Mittelwertsatz für Riemann-Integrale.

Die Aussage (6.5.12) ist natürlich gerade die Definition von  $\mu(f)$  aus (6.5.10). Trotzdem erkennt man hier die Verallgemeinerung von (6.5.8). Eine graphische Idee von (6.5.12) beziehungsweise (6.5.13) gibt Abbildung 6.9.

## 6.6 Uneigentliche Riemann-Integrale

Bislang ist unsere Integrationstheorie aus zwei Gründen erheblich limitiert. Zum einen sind Riemann-integrierbare Funktionen auf einem abgeschlossenen Intervall notwendigerweise beschränkt. Es sind aber durchaus Situationen denkbar, in denen der Graph einer unbeschränkten Funktion eine endliche Fläche einschließt. Beispiele werden wir konkret sehen. Die Idee ist, dass die Spitzen, wo  $f$  unbeschränkt wird, nur schmal genug sein müssen. Es gilt also diese Einschränkung durch eine Erweiterung des Integralbegriffs zu überwinden. Die andere Einschränkung ist die, dass wir nur über abgeschlossene Intervalle integrieren können. Auch hier besteht der Wunsch, unbeschränkte Integrationsintervalle der Form  $[a, \infty)$  oder  $(-\infty, b]$ , sowie  $(-\infty, \infty)$  zuzulassen. Beide Beschränkungen sind in der Tat verwandt und können auf folgende Weise überwunden werden:

**Definition 6.6.1 (Uneigentliches Riemann-Integral)** Sei  $(a, b) \subseteq \mathbb{R}$  ein offenes Intervall, wobei  $a = -\infty$  und  $b = +\infty$  erlaubt sind. Sei weiter  $f: (a, b) \rightarrow \mathbb{K}$  eine Abbildung.

i.) Ist  $c \in (a, b)$  und  $f|_{[c, d]} \in \mathcal{R}([c, d], \mathbb{K})$  für alle  $c < d < b$ , und existiert der Grenzwert

$$\int_c^b f(t) dt = \lim_{x \nearrow b} \int_c^x f(t) dt, \quad (6.6.1)$$

so heißt der Wert das uneigentliche Riemann-Integral von  $f$  über  $[c, b)$ .

ii.) Analog definiert man das uneigentliche Riemann-Integral

$$\int_a^c f(t) dt = \lim_{x \searrow a} \int_x^c f(t) dt \quad (6.6.2)$$

von  $f$  über  $(a, c]$ .

iii.) Gibt es schließlich ein  $c \in (a, b)$  derart, dass beide Grenzwerte (6.6.1) und (6.6.2) existieren, so heißt

$$\int_a^b f(t) dt = \lim_{x \nearrow b} \int_c^x f(t) dt + \lim_{x' \searrow a} \int_{x'}^c f(t) dt \quad (6.6.3)$$

das uneigentliche Riemann-Integral von  $f$  über  $(a, b)$ .

Es gilt nun zuerst einige Konsistenzprüfungen vorzunehmen:

**Proposition 6.6.2** Sei  $(a, b) \subseteq \mathbb{R}$  ein offenes Intervall, und seien  $f, g: (a, b) \rightarrow \mathbb{K}$  Funktionen.

i.) Sei  $c \in (a, b)$  derart, dass das uneigentliche Riemann-Integral von  $f$  über  $[c, b)$  existiert. Dann existiert das uneigentliche Riemann-Integral von  $f$  auch über  $[c', b)$  für jedes  $c' \in (c, b)$  und

$$\int_c^b f(t) dt = \int_c^{c'} f(t) dt + \int_{c'}^b f(t) dt. \quad (6.6.4)$$

ii.) Das uneigentliche Riemann-Integral ist linear: Existieren die uneigentlichen Riemann-Integrale von  $f$  und  $g$  über  $[c, b)$  für ein  $c \in (a, b)$ , so existiert auch das uneigentliche Riemann-Integral von  $zf + wg$  für  $z, w \in \mathbb{K}$  über  $[c, b)$  mit

$$\int_c^b (zf + wg)(t) dt = z \int_c^b f(t) dt + w \int_c^b g(t) dt. \quad (6.6.5)$$

- iii.) Ist  $b \in \mathbb{R}$  und  $f \in \mathcal{R}([c, b], \mathbb{K})$ , so existiert das uneigentliche Riemann-Integral von  $f$  über  $[c, b)$  und stimmt mit dem Riemann-Integral von  $f$  über  $[c, b]$  überein.
- iv.) Analoge Aussagen gelten für uneigentliche Riemann-Integrale über  $(a, c]$  und  $(a, b)$ .
- v.) Existiert das uneigentliche Riemann-Integral von  $f$  über  $(a, b)$  bezüglich eines Wertes  $c \in (a, b)$ , so auch bezüglich jeden anderen Wertes, und die Grenzwerte (6.6.3) stimmen überein.

BEWEIS: Sei  $c' \in (c, b)$ . Dann gilt für jedes  $x \in (c', b)$  zunächst

$$\int_c^x f(t) dt = \int_c^{c'} f(t) dt + \int_{c'}^x f(t) dt$$

nach Proposition 6.2.10, ii.). Da der erste Summand auf der rechten Seite unabhängig von  $x$  ist, existiert der Limes  $\lim_{x \nearrow b}$  der linken Seite genau dann, wenn der entsprechende Limes  $\lim_{x \nearrow b} \int_{c'}^x f(t) dt$  existiert. Zudem gilt dann

$$\begin{aligned} \int_c^b f(t) dt &= \lim_{x \nearrow b} \int_c^x f(t) dt \\ &= \lim_{x \nearrow b} \left( \int_c^{c'} f(t) dt + \int_{c'}^x f(t) dt \right) \\ &= \int_c^{c'} f(t) dt + \lim_{x \nearrow b} \int_{c'}^x f(t) dt \\ &= \int_c^{c'} f(t) dt + \int_{c'}^b f(t) dt, \end{aligned}$$

was (6.6.4) zeigt. Der zweite Teil ist klar, da Grenzwerte von Funktionswerten nach den Rechenregeln zur Konvergenz von Funktionswerten aus Bemerkung 3.1.10 linear sind. Der dritte Teil ist die Stetigkeitsaussage von Proposition 6.4.1, i.). Den vierten Teil erhält man mit den gleichen Argumenten. Wir betrachten nun also eine Funktion mit existierendem uneigentlichen Riemann-Integral (6.6.3). Hier wird die Existenz eines Punktes  $c \in (a, b)$  gefordert, womit das uneigentliche Riemann-Integral also zunächst von  $c$  abhängt. Dass dies nicht so ist, folgt nun aus Teil i.) und iii.). Sei also zunächst  $c' \in (c, b)$  größer als  $c$ . Dann existiert  $\int_{c'}^b f(t) dt$  und es gilt

$$\int_c^b f(t) dt = \int_c^{c'} f(t) dt + \int_{c'}^b f(t) dt.$$

Mit der analogen Aussage für  $\int_a^c f(t) dt$  sehen wir, dass auch  $\int_a^{c'} f(t) dt$  existiert und

$$\int_a^{c'} f(t) dt = \int_a^c f(t) dt + \int_c^{c'} f(t) dt$$

erfüllt. Zusammen existieren also die beiden uneigentlichen Riemann-Integrale  $\int_a^{c'} f(t) dt$  und  $\int_{c'}^b f(t) dt$  ebenfalls, und es gilt

$$\int_a^{c'} f(t) dt + \int_{c'}^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt.$$

Den Fall  $c' < c$  zeigt man genauso. □

Wir illustrieren das Konzept nun durch einige wichtige Beispiele:

### Beispiel 6.6.3 (Uneigentliche Riemann-Integrale)

i.) Wir betrachten die Funktion  $f: \mathbb{R}^+ \ni x \mapsto \frac{1}{x^\alpha} \in \mathbb{R}^+$  für  $\alpha > 0$ . Da  $f$  stetig ist, ist  $f$  über jedes abgeschlossene Intervall  $[a, b] \subseteq \mathbb{R}^+$  Riemann-integrierbar. Wir kennen die Stammfunktion explizit, womit

$$\int_a^b \frac{1}{x^\alpha} dx = \begin{cases} \log(x) \Big|_a^b & \text{für } \alpha = 1 \\ \frac{1}{1-\alpha} x^{1-\alpha} \Big|_a^b & \text{für } \alpha \neq 1. \end{cases} \quad (6.6.6)$$

Für  $\alpha = 1$  erhalten wir daher

$$\lim_{b \rightarrow +\infty} \int_a^b \frac{1}{x} dx = \lim_{b \rightarrow +\infty} (\log(b) - \log(a)) = +\infty, \quad (6.6.7)$$

womit das zugehörige uneigentliche Riemann-Integral divergiert. Ebenso gilt

$$\lim_{a \searrow 0} \int_a^b \frac{1}{x} dx = \lim_{a \searrow 0} (\log(b) - \log(a)) = +\infty, \quad (6.6.8)$$

womit auch dieses uneigentliche Riemann-Integral divergiert. Lässt man den Wert  $+\infty$  als (uneigentlichen) Grenzwert zu, erhalten wir also für alle  $a$  und  $b$

$$\int_a^\infty \frac{1}{x} dx = +\infty = \int_0^b \frac{1}{x} dx. \quad (6.6.9)$$

Für  $\alpha \neq 1$  erhalten wir entsprechend

$$\lim_{b \rightarrow +\infty} \int_a^b \frac{1}{x^\alpha} dx = \lim_{b \rightarrow +\infty} \frac{1}{1-\alpha} (b^{1-\alpha} - a^{1-\alpha}) = \begin{cases} +\infty & \text{für } \alpha < 1 \\ \frac{a^{1-\alpha}}{\alpha-1} & \text{für } \alpha > 1, \end{cases} \quad (6.6.10)$$

womit

$$\int_a^\infty \frac{1}{x^\alpha} dx = \begin{cases} +\infty & \text{für } \alpha < 1 \\ \frac{a^{1-\alpha}}{\alpha-1} & \text{für } \alpha > 1. \end{cases} \quad (6.6.11)$$

Für die untere Grenze erhalten wir entsprechend

$$\lim_{a \searrow 0} \int_a^b \frac{1}{x^\alpha} dx = \lim_{a \searrow 0} \frac{1}{1-\alpha} (b^{1-\alpha} - a^{1-\alpha}) = \begin{cases} \frac{b^{1-\alpha}}{1-\alpha} & \text{für } \alpha < 1 \\ +\infty & \text{für } \alpha > 1 \end{cases} \quad (6.6.12)$$

und damit

$$\int_0^b \frac{1}{x^\alpha} dx = \begin{cases} \frac{b^{1-\alpha}}{1-\alpha} & \text{für } \alpha < 1 \\ +\infty & \text{für } \alpha > 1, \end{cases} \quad (6.6.13)$$

womit wir alle Fälle  $\alpha > 0$  abgedeckt haben. Insbesondere sieht man, dass das doppelt uneigentliche Riemann-Integral

$$\int_0^\infty \frac{1}{x^\alpha} dx = +\infty \quad (6.6.14)$$

immer divergiert, während einige der uneigentlichen Riemann-Integrale bei einer Grenze zwischen 0 und  $+\infty$  konvergieren. Eine Illustration der entsprechenden endlichen beziehungsweise unendlichen Flächen findet sich in Abbildung 6.10, Abbildung 6.11 und Abbildung 6.12.

ii.) Die Funktion  $f: \mathbb{R} \ni x \mapsto \frac{1}{1+x^2} \in \mathbb{R}^+$  ist stetig und daher Riemann-integrierbar über jedes abgeschlossene Intervall. Wir kennen auch hier eine Stammfunktion, nämlich  $\arctan$ , siehe nochmals Tabelle ???. Daher gilt

$$\int_a^b \frac{1}{1+x^2} dx = \arctan(b) - \arctan(a). \quad (6.6.15)$$

Abbildung 6.10: Ein endlicher Flächeninhalt nach rechts, aber ein unendlicher unter dem Graph von 0 bis  $a$ .

Abbildung 6.11: Ein endlicher Flächeninhalt unter dem Graph von 0 bis  $b$ , ein unendlicher nach rechts.

Der Grenzwert  $b \rightarrow +\infty$  existiert und liefert

$$\lim_{b \rightarrow +\infty} \arctan(b) = \frac{\pi}{2}, \quad (6.6.16)$$

siehe Proposition 4.5.1 zum entsprechenden Limes

$$\lim_{x \rightarrow \frac{\pi}{2}} \tan(x) = +\infty. \quad (6.6.17)$$

Genauso erhält man

$$\lim_{a \rightarrow -\infty} \arctan(a) = -\frac{\pi}{2}, \quad (6.6.18)$$

womit insgesamt

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \pi \quad (6.6.19)$$

folgt.

Bei Riemann-Integralen über abgeschlossene Intervalle war für eine Riemann-integrierbare Funktion  $f$  der Betrag  $|f|$  ebenfalls Riemann-integrierbar, siehe Korollar 6.3.4. Im Falle uneigentlicher Riemann-Integrale muss dies nicht länger der Fall sein, da im Grenzwert  $\lim_{x \rightarrow b} \int_a^x f(t) dt$  unterschiedliche Vorzeichen zu besserem Konvergenzverhalten führen können, vergleichbar zur Konvergenz von Reihen nach dem Leibniz-Kriterium. Aus diesem Grunde ist folgende Begriffsbildung interessant:

**Definition 6.6.4 (Absolut konvergente uneigentliche Riemann-Integrale)** Sei  $(a, b) \subseteq \mathbb{R}$  ein offenes Intervall mit  $c \in (a, b)$ . Für eine uneigentlich Riemann-integrierbare Funktion  $f: [c, b) \rightarrow \mathbb{R}$  heißt das uneigentliche Riemann-Integral  $\int_c^b f(t) dt$  absolut konvergent, wenn auch  $|f|$  über  $[c, b)$  uneigentlich Riemann-integrierbar ist. Analog definiert man absolute Konvergenz von  $\int_a^c f(t) dt$ .

Hier ist natürlich wichtig, dass wir voraussetzen, dass die Riemann-Integrale von  $f$  über  $[c, x]$  für alle  $x \in [c, b)$  existieren: Aus der Existenz der Riemann-Integrale von  $|f|$  kann nicht geschlossen werden, dass  $f$  auch Riemann-integrierbar ist, siehe Übung 6.8.4.

Die bekannten Rechenregeln und Konvergenzkriterien für Netze ergeben nun schnell entsprechende Resultate für uneigentliche Riemann-Integrale. Wir erwähnen hier exemplarisch einige:

**Proposition 6.6.5** Sei  $(a, b) \subseteq \mathbb{R}$  ein offenes Intervall mit  $c \in (a, b)$ . Sei weiter  $f: (a, b) \rightarrow \mathbb{K}$  eine Funktion mit  $f \in \mathcal{R}([c, x], \mathbb{K})$  für alle  $x \in [c, b)$ . Das uneigentliche Riemann-Integral existiert genau dann, wenn es zu jedem  $\epsilon > 0$  ein  $x_0 \in [c, d)$  gibt, sodass für alle  $x_0 \leq x \leq x'$

$$\left| \int_x^{x'} f(t) dt \right| < \epsilon \quad (6.6.20)$$

gilt. Eine analoge Aussage gilt für uneigentliche Riemann-Integrale der Form  $\int_a^c f(t) dt$ .

Abbildung 6.12: Für  $f(x) = \frac{1}{x}$  sind beide uneigentlichen Riemann-Integrale divergent. Aufgrund der Spiegelungssymmetrie des Graphen an der ersten Winkelhalbierenden ist heuristisch insofern zu erwarten, dass entweder beide Flächen endlich oder beide Flächen unendlich sein sollten.

Dies ist gerade das Cauchy-Kriterium für Netz-Konvergenz, siehe Übung ??, angewendet auf das Netz der Riemann-Integrale  $(\int_c^x f(t) dt)_{x \in [c, b)}$  mit Richtung „näher bei  $b$ “.

**Beispiel 6.6.6** Wir betrachten die stetige Funktion

$$\text{sinc}: \mathbb{R} \ni x \mapsto \text{sinc}(x) = \frac{\sin(x)}{x} \in \mathbb{R} \quad (6.6.21)$$

aus Beispiel 5.3.13, wobei  $\text{sinc}(0) = 1$  die stetige Fortsetzung bewirkt. Insbesondere ist  $\text{sinc} \in \mathcal{R}([0, x], \mathbb{R})$  für alle  $x \geq 0$ . Sei  $0 < x_0 \leq x \leq x'$ . Dann gilt

$$\int_x^{x'} \text{sinc}(t) dt = \int_x^{x'} \frac{\sin(t)}{t} dt = -\frac{\cos(t)}{t} \Big|_x^{x'} - \int_x^{x'} \frac{\cos(t)}{t^2} dt$$

nach partieller Integration. Wir können dies nun folgendermaßen abschätzen

$$\begin{aligned} \left| \int_x^{x'} \text{sinc}(t) dt \right| &\leq \left| -\frac{\cos(t)}{t} \Big|_x^{x'} \right| + \int_x^{x'} \frac{|\cos(t)|}{t^2} dt \\ &\leq \frac{1}{x'} + \frac{1}{x} + \int_x^{x'} \frac{1}{t^2} dt \\ &= \frac{1}{x'} + \frac{1}{x} + \left( -\frac{1}{t} \right) \Big|_x^{x'} \\ &= \frac{1}{x'} + \frac{1}{x} - \frac{1}{x'} + \frac{1}{x} \\ &= \frac{2}{x}, \end{aligned}$$

unabhängig von der oberen Integrationsgrenze. Wählt man daher  $x_0$  groß genug, folgt mit  $x \geq x_0$ , dass das Integral  $|\int_x^{x'} \text{sinc}(t) dt|$  beliebig klein wird. Das Cauchy-Kriterium ist erfüllt. Daher existiert das uneigentliche Riemann-Integral von  $\text{sinc}$  über  $[0, \infty)$ . Wir werden später sehen, dass

$$\int_0^\infty \frac{\sin(t)}{t} dt = \frac{\pi}{2} \quad (6.6.22)$$

gilt. Man kann nun weiter zeigen, dass dieses uneigentliche Riemann-Integral *nicht* absolut konvergiert, siehe Übung 6.8.8.

Für absolute Konvergenz ist das Majorantenkriterium hier sehr nützlich.

**Proposition 6.6.7** Sei  $(a, b) \subseteq \mathbb{R}$  ein offenes Intervall mit  $c \in (a, b)$ . Sei weiter  $f: (a, b) \rightarrow \mathbb{K}$  eine Funktion mit  $f \in \mathcal{R}([c, x], \mathbb{K})$  für alle  $x \in [c, b)$ .

i.) Konvergiert das uneigentliche Riemann-Integral von  $f$  über  $[c, b)$  absolut, so auch im bisherigen Sinne, und es gilt

$$\left| \int_c^b f(t) dt \right| \leq \int_c^b |f(t)| dt. \quad (6.6.23)$$

ii.) Ist  $g: (a, b) \rightarrow \mathbb{R}$  positiv und  $|f(t)| \leq g(t)$  für alle  $t$ , und existiert das uneigentliche Riemann-Integral von  $g$  über  $[c, b)$ , so konvergiert das uneigentliche Riemann-Integral von  $f$  über  $[c, b)$  absolut und es gilt

$$\left| \int_c^b f(t) dt \right| \leq \int_c^b g(t) dt. \quad (6.6.24)$$

BEWEIS: Für den ersten Teil betrachten wir  $c \leq x \leq x' < b$ . Dann gilt

$$\left| \int_x^{x'} f(t) dt \right| \leq \int_x^{x'} |f(t)| dt \quad (*)$$

nach der Dreiecksungleichung für Riemann-Integrale aus Korollar 6.3.6. Da das uneigentliche Riemann-Integral von  $|f|$  nach Voraussetzung existiert, folgt aus dem Cauchy-Kriterium von Proposition 6.6.5, dass es zu jedem  $\epsilon > 0$  ein  $x_0 \in [c, b)$  gibt, so dass

$$\left| \int_x^{x'} |f(t)| dt \right| = \int_x^{x'} |f(t)| dt < \epsilon$$

gilt, sobald  $x_0 \leq x \leq x' < b$  gilt. Zusammen mit  $(*)$  zeigt dies, dass das Cauchy-Kriterium auch für  $\int_c^b f(t) dt$  erfüllt ist, was den ersten Teil nach Proposition 6.6.5 zeigt. Der zweite ist dann klar, da

$$0 \leq \int_x^{x'} |f(t)| dt \leq \int_x^{x'} g(t) dt$$

und die Konvergenz des uneigentlichen Riemann-Integrals von  $g$  über  $[c, b)$  nach dem Cauchy-Kriterium die rechte Seite für  $x_0 \leq x \leq x' < b$  mit  $x_0$  groß genug klein werden lässt. Damit wird das Cauchy-Kriterium auch für  $|f(t)|$  erfüllt.  $\square$

Uneigentliche Riemann-Integrale über ein unbeschränktes Intervall  $[a, \infty)$ , also für  $b = +\infty$ , verhalten sich in vielerlei Hinsicht ähnlich wie unendliche Reihen. Diese Ähnlichkeit ist tatsächlich viel tiefergehend, wie folgendes Resultat zeigt:

**Proposition 6.6.8** Sei  $f: [1, \infty) \rightarrow \mathbb{R}^+$  eine monoton fallende Funktion. Sei weiter  $(a_n)_{n \in \mathbb{N}}$  die Folge mit

$$a_n = \sum_{k=1}^n f(k) - \int_1^{n+1} f(t) dt. \quad (6.6.25)$$

i.) Die Folge  $(a_n)_{n \in \mathbb{N}}$  konvergiert, und es gilt

$$0 \leq \lim_{n \rightarrow \infty} a_n \leq f(1) \quad (6.6.26)$$

ii.) Die Reihe  $\sum_{k=1}^{\infty} f(k)$  konvergiert genau dann, wenn das uneigentliche Riemann-Integral  $\int_1^{\infty} f(t) dt$  existiert.

BEWEIS: Zunächst gilt  $f \in \mathcal{R}([1, x], \mathbb{R})$  nach Korollar 6.2.8 aufgrund der Monotonie, womit die Folge  $(a_n)_{n \in \mathbb{N}}$  wohldefiniert ist. Die Monotonie

$$f(k) \geq f(x) \geq f(k+1)$$

für  $x \in [k, k+1]$  zeigt dann

$$f(k) = \int_k^{k+1} f(k) dt \geq \int_k^{k+1} f(t) dt \geq \int_k^{k+1} f(k+1) dt = f(k+1) \quad (*)$$

nach der Monotonie des Riemann-Integrals gemäß Proposition 6.1.7, iii.). Summiert man diese Ungleichung über  $k$  erhält man

$$\sum_{k=1}^n f(k) \geq \sum_{k=1}^n \int_k^{k+1} f(t) dt = \int_1^{n+1} f(t) dt \geq \sum_{k=1}^n f(k+1) = \sum_{k=2}^{n+1} f(k).$$

Abbildung 6.13: Die Ungleichungen aus dem Beweis von Proposition 6.6.8 sperren das uneigentliche Integral zwischen Ober- und Untersummen ein.

Für die Folgenglieder  $a_n$  bedeutet dies zum einen

$$a_n = \sum_{k=1}^n f(k) - \int_1^{n+1} f(t) dt \geq 0$$

und zum anderen

$$a_n = f(1) - f(n+1) + \sum_{k=2}^{n+1} f(k) - \int_1^{n+1} f(t) dt \leq f(1) - f(n+1) \leq f(1),$$

da  $f(n+1) \geq 0$  nach Voraussetzung. Insgesamt gilt also

$$0 \leq a_n \leq f(1)$$

für alle  $n$ . Für die Differenz  $a_{n+1} - a_n$  gilt

$$a_{n+1} - a_n = \sum_{k=1}^{n+1} f(k) - \int_1^{n+2} f(t) dt - \left( \sum_{k=1}^n f(k) - \int_1^{n+1} f(t) dt \right) = f(n+1) - \int_{n+1}^{n+2} f(t) dt \geq 0$$

nach (\*). Also ist  $(a_n)_{n \in \mathbb{N}}$  monoton wachsend und nach oben durch  $f(1)$  beschränkt. Damit konvergiert  $(a_n)_{n \in \mathbb{N}}$ , und der Grenzwert erfüllt (6.6.26), womit der erste Teil gezeigt ist. Der zweite Teil ist nun einfach: Konvergiert beispielsweise die Reihe, so existiert

$$\lim_{n \rightarrow \infty} \int_1^{n+1} f(t) dt = - \lim_{n \rightarrow \infty} a_n + \sum_{k=1}^{\infty} f(k).$$

Da die Funktion  $f$  monoton fallend ist, genügt dies, um das Cauchy-Kriterium für das uneigentliche Riemann-Integral zu zeigen. Ist umgekehrt das uneigentliche Riemann-Integral konvergent, so auch die Reihe, da

$$\sum_{k=1}^{\infty} f(k) = \lim_{n \rightarrow \infty} a_n + \int_1^{\infty} f(t) dt.$$

Dies zeigt den zweiten Teil. □

Geometrisch lässt sich die Konvergenzaussage so verstehen, dass das uneigentliche Integral zwischen Ober- und Untersumme „eingespart“ wird, siehe auch Abbildung 6.13.

**Beispiel 6.6.9 (Harmonische Reihe)** Wir betrachten nochmals die harmonische Reihe mittels der Funktion

$$f(x) = \frac{1}{x} \tag{6.6.27}$$

für  $x \in \mathbb{R}^+$ . Diese genügt den Anforderungen der Proposition 6.6.8. Damit erhalten wir zum einen, dass die harmonische Reihe  $\sum_{k=1}^{\infty} f(k) = \sum_{k=1}^{\infty} \frac{1}{k}$  genau dann konvergiert, wenn das uneigentliche Riemann-Integral

$$\int_1^x \frac{1}{t} dt = \log(x) \tag{6.6.28}$$

für  $x \rightarrow \infty$  konvergiert. Da der Logarithmus unbeschränkt ist, gilt aber

$$\int_1^{\infty} \frac{dt}{t} = +\infty, \tag{6.6.29}$$

was nun einen weiteren, unabhängigen, Beweis für die Divergenz

$$\sum_{k=1}^{\infty} \frac{1}{k} = +\infty \quad (6.6.30)$$

liefert, der im Vergleich zum Vorgehen in Beispiel 2.6.2 deutlich klarer ist. Die Proposition zeigt aber zudem, dass

$$\gamma = \lim_{n \rightarrow \infty} \left( \sum_{k=1}^n \frac{1}{k} - \int_1^{n+1} \frac{1}{x} dx \right) = \lim_{n \rightarrow \infty} \left( \sum_{k=1}^n \frac{1}{k} - \log(n+1) \right) \quad (6.6.31)$$

konvergiert. Wegen

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

und  $\log\left(\frac{n+1}{n}\right) \rightarrow 0$  aufgrund der Stetigkeit von  $\log$  bei  $x = 1$ , können wir dies auch als

$$\gamma = \lim_{n \rightarrow \infty} \left( \sum_{k=1}^n \frac{1}{k} - \log(n) \right) \quad (6.6.32)$$

schreiben.

**Definition 6.6.10 (Euler-Mascheroni-Konstante)** Der Grenzwert  $\gamma$  aus (6.6.32) heißt *Euler-Mascheroni-Konstante*.

Numerisch findet man [5]

$$\gamma = 0,5772156649 \dots \quad (6.6.33)$$

Anders als bei  $\pi$  und  $e$  ist die Natur von  $\gamma$  weitgehend mysteriös und unbekannt. Es ist beispielsweise nicht bekannt, ob  $\gamma$  rational oder irrational beziehungsweise algebraisch oder transzendent ist, auch wenn vermutet wird, dass  $\gamma$  zumindest irrational ist.

Uneigentliche Riemann-Integrale erlauben es uns, eine weitere Funktion zu konstruieren, deren Wichtigkeit für die Mathematik kaum zu überschätzen ist: die *Gamma-Funktion*.

**Proposition 6.6.11** Sei  $z \in \mathbb{C}$  mit  $\operatorname{Re}(z) > 0$ .

i.) Das uneigentliche Riemann-Integral

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt \quad (6.6.34)$$

konvergiert absolut.

ii.) Es gilt  $\Gamma(1) = 1$  und

$$z\Gamma(z) = \Gamma(z+1). \quad (6.6.35)$$

iii.) Für alle  $n \in \mathbb{N}_0$  gilt

$$\Gamma(n+1) = n!. \quad (6.6.36)$$

BEWEIS: Zunächst gilt für  $t \in \mathbb{R}^+$  und  $z = a + ib$

$$|t^{z-1}| = |t^{a+ib-1}| = t^{a-1} |e^{\log(t)ib}| = t^{a-1}.$$

Damit ist die Frage nach der absoluten Konvergenz des uneigentlichen Riemann-Integrals auf den Fall  $z \in \mathbb{R}$  zurückgeführt. Hier gilt nun für  $x > 0$

$$\int_0^1 t^{x-1} e^{-t} dt \leq \int_0^1 t^{x-1} dt = \frac{1}{x} t^x \Big|_0^1 = \frac{1}{x},$$



Abbildung 6.14: Der Graph der Gamma-Funktion für reelle Argumente.

womit  $\int_0^1 t^{x-1} e^{-t} dt$  nach dem Majorantenkriterium konvergiert. Genauso erhält man mit den Grenzwerten aus Proposition 4.2.1, *iv.*), dass es ein  $t_0 \geq 1$  gibt mit

$$t^{x-1} e^{-t} = \frac{1}{t^2} t^{x+1} e^{-t} \leq \frac{1}{t^2}$$

für alle  $t \geq t_0$ , da  $\lim_{t \rightarrow \infty} t^{x+1} e^{-t} = 0$ . Damit folgt aber

$$\int_{t_0}^{\infty} t^{x-1} e^{-t} dt \leq \int_{t_0}^{\infty} \frac{1}{t^2} dt = -\frac{1}{t} \Big|_{t_0}^{\infty} = \frac{1}{t_0},$$

womit wir auch hier eine Majorante gefunden haben. Da das Integral dazwischen, also  $\int_1^{t_0} t^{x-1} e^{-t} dt$ , unzweifelhaft existiert, haben wir insgesamt die absolute Konvergenz des (doppelt) uneigentlichen Riemann-Integrals gezeigt. Für  $z = 1$  erhalten wir

$$\Gamma(1) = \int_0^{\infty} t^0 e^{-t} dt = \int_0^{\infty} e^{-t} dt = -e^{-t} \Big|_0^{\infty} = 1,$$

wobei hier nur die obere Grenze zu einem uneigentlichen Riemann-Integral führt. Für beliebiges  $z$  verwenden wir eine partielle Integration in

$$\begin{aligned} \int_{\epsilon}^R t^z e^{-t} dt &= -t^z e^{-t} \Big|_{\epsilon}^R - z \int_{\epsilon}^R t^{z-1} e^{-t} dt \\ &= -R^z e^{-R} + \epsilon^z e^{-\epsilon} - z \int_{\epsilon}^R t^{z-1} e^{-t} dt. \end{aligned}$$

Für  $R \rightarrow \infty$  erhalten wir aus Proposition 4.2.1, *iv.*),

$$\lim_{R \rightarrow \infty} -R^z e^{-R} = 0$$

wie bereits zuvor argumentiert auch für  $z \in \mathbb{C}$ . Für  $\epsilon \rightarrow 0$  gilt entsprechend auch

$$\lim_{\epsilon \rightarrow 0} \epsilon^z e^{-\epsilon} = 0,$$

wieder für  $z \in \mathbb{C}$  mit  $\operatorname{Re}(z) > 1$  aus demselben Grunde wie zuvor. Damit folgt aber

$$\int_0^{\infty} t^z e^{-t} dt = z \int_0^{\infty} t^{z-1} e^{-t} dt,$$

was (6.6.35) zeigt. Zusammen mit  $\Gamma(1) = 1$  ergibt dies (6.6.36) durch eine einfache Induktion nach  $n$ . □

Die Gamma-Funktion interpoliert also die diskreten Werte der Fakultät  $n!$  unter Erhaltung der Funktionalgleichung (6.6.35). Eine genauere Analyse liefert eine graphische Darstellung der Gamma-Funktion in Abbildung 6.14 für reelle  $x \in \mathbb{R}^+$ . Mit etwas größerem Aufwand kann man zeigen, dass  $\Gamma$  eine stetige und sogar glatte Funktion von  $z \in \mathbb{C}$  mit  $\operatorname{Re}(z) > 0$  ist. Zudem lässt sich  $\Gamma$  auf  $\mathbb{C} \setminus \{0, -1, -2, \dots\}$  fortsetzen und liefert dort eine komplex differenzierbare Funktion. Diese und viele weitere Eigenschaften der Gamma-Funktion werden in Kursen zur Funktionentheorie im Detail vorgestellt.

Abbildung 6.15: Eine Treppenfunktion  $\phi$  zu einer Zerlegung  $\mathcal{J}$  von  $[a, b]$ .

## 6.7 Integration und gleichmäßige Konvergenz

Wie schon bei Stetigkeit und Differenzierbarkeit wollen wir auch für die Riemann-Integration einen geeigneten Konvergenzbegriff finden, der es uns erlaubt, Integration und Grenzwertbildung zu vertauschen. Es zeigt sich, dass auch hier gleichmäßige Konvergenz ein hinreichendes Kriterium liefert.

Zunächst benötigen wir folgende Neuinterpretation von Satz 6.2.6: Wir sagen,

$$\phi: [a, b] \longrightarrow \mathbb{R} \quad (6.7.1)$$

ist eine *Treppenfunktion*, falls es eine Zerlegung  $\mathcal{J} = \{a = t_0 < t_1 < \cdots < t_N = b\}$  von  $[a, b]$  gibt, so dass

$$\phi|_{(t_{i-1}, t_i)} = c_i \quad (6.7.2)$$

für eine Zahl  $c_i \in \mathbb{R}$  und alle  $i = 1, \dots, N$  gilt. Damit sind Treppenfunktionen also stückweise konstant, siehe auch Abbildung 6.15. Treppenfunktionen sind immer Riemann-integrierbar mit

$$\int_a^b \phi(x) \, dx = \sum_{i=1}^N c_i (t_i - t_{i-1}). \quad (6.7.3)$$

Dies zeigt man durch explizite Angabe einer Ober- und Untersumme: Die einzige Schwierigkeit ist zu sehen, dass die Funktionswerte an den Punkten  $t_i$  der Zerlegung keine Rolle spielen. Da dies nur endlich viele Werte sind, kann man Zerlegungen wählen, die sehr kurze Intervalle um die Punkte  $t_i$  enthalten, so dass der Gesamtbeitrag zu den Ober- und Untersummen klein bleibt, siehe auch Übung 6.8.5. Mit Treppenfunktionen können wir die Riemann-Integrierbarkeit folgendermaßen charakterisieren:

**Proposition 6.7.1** *Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $f: [a, b] \longrightarrow \mathbb{R}$  eine Funktion. Die Funktion  $f$  ist genau dann Riemann-integrierbar, wenn es zu jedem  $\epsilon > 0$  Treppenfunktionen  $\phi, \psi: [a, b] \longrightarrow \mathbb{R}$  gibt, so dass*

$$\phi(x) \leq f(x) \leq \psi(x) \quad (6.7.4)$$

für alle  $x \in [a, b]$  und

$$\int_a^b \psi(x) \, dx - \int_a^b \phi(x) \, dx < \epsilon \quad (6.7.5)$$

gilt.

BEWEIS: Ist  $f \in \mathcal{R}([a, b], \mathbb{R})$ , so ist  $f$  beschränkt und zu  $\epsilon > 0$  finden wir eine Zerlegung  $\mathcal{J} = \{a = t_0 < t_1 < \cdots < t_N = b\}$  mit der Eigenschaft

$$\mathcal{O}_{\mathcal{J}}(f) - \mathcal{U}_{\mathcal{J}}(f) < \epsilon.$$

Nach Definition der Ober- und Untersummen gilt also

$$\mathcal{O}_{\mathcal{J}}(f) = \sum_{i=1}^N \sup \left( f|_{[t_{i-1}, t_i]} \right) (t_i - t_{i-1})$$

und

$$\mathcal{U}_{\mathcal{J}}(f) = \sum_{i=1}^N \inf \left( f|_{[t_{i-1}, t_i]} \right) (t_i - t_{i-1}).$$

Wir definieren  $\phi$  und  $\psi$  durch

$$\phi(x) = \begin{cases} \inf \left( f|_{[t_{i-1}, t_i]} \right) & \text{für } x \in (t_{i-1}, t_i) \\ f(t_i) & \text{für } x = t_i \end{cases}$$

sowie durch

$$\psi(x) = \begin{cases} \sup \left( f|_{[t_{i-1}, t_i]} \right) & \text{für } x \in (t_{i-1}, t_i) \\ f(t_i) & \text{für } x = t_i, \end{cases}$$

wobei  $i = 1, \dots, N$ . Nach Konstruktion gilt dann (6.7.4). Weiter gilt mit (6.7.3)

$$\int_a^b \phi(x) \, dx = \sum_{i=1}^N \inf \left( f|_{[t_{i-1}, t_i]} \right) (t_i - t_{i-1}) = \mathcal{U}_J(f)$$

und

$$\int_a^b \psi(x) \, dx = \sum_{i=1}^N \sup \left( f|_{[t_{i-1}, t_i]} \right) (t_i - t_{i-1}) = \mathcal{O}_J(f),$$

was dann (6.7.5) zeigt. Finden wir zu  $\epsilon > 0$  umgekehrt solche Treppenfunktionen mit (6.7.4) und (6.7.5), so gilt zum einen

$$\int_a^b \phi(x) \, dx = \int_a^b \phi(x) \, dx \leq \int_a^b f(x) \, dx$$

und

$$\int_a^b f(x) \, dx \leq \int_a^b \psi(x) \, dx = \int_a^b \psi(x) \, dx$$

aufgrund von (6.7.4) und der Riemann-Integrierbarkeit von  $\phi$  und  $\psi$ . Mit (6.7.5) folgt dann

$$\int_a^b f(x) \, dx - \int_a^b f(x) \, dx < \epsilon$$

für alle  $\epsilon > 0$ , womit Ober- und Unterintegral übereinstimmen. Dies ist nach Satz 6.2.6 zu  $f \in \mathcal{R}([a, b], \mathbb{R})$  äquivalent.  $\square$

Mit diesem Kriterium können wir nun schnell folgenden Satz zeigen: gleichmäßige Konvergenz vertauscht mit Riemann-Integration. Der schwierigste Teil dabei ist, die Riemann-Integrierbarkeit der Grenzfunktion sicherzustellen:

**Satz 6.7.2 (Gleichmäßige Konvergenz und Riemann-Integral)** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall und  $(f_n)_{n \in \mathbb{N}}$  eine Folge von Riemann-integrierbaren Funktionen  $f_n \in \mathcal{R}([a, b], \mathbb{K})$ . Konvergiert  $(f_n)_{n \in \mathbb{N}}$  gleichmäßig auf  $[a, b]$  gegen  $f \in \text{Abb}([a, b], \mathbb{K})$ , so gilt  $f \in \mathcal{R}([a, b], \mathbb{K})$  und

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx. \quad (6.7.6)$$

BEWEIS: Wir betrachten zunächst  $\mathbb{K} = \mathbb{R}$ . Da  $f_n \in \mathcal{R}([a, b], \mathbb{R})$  beschränkt ist, folgt  $f \in \mathcal{B}([a, b], \mathbb{R})$  nach Proposition 3.7.8, i.). Sei  $\epsilon > 0$  gegeben und  $\delta = \frac{\epsilon}{4(b-a)}$ . Dann finden wir ein  $N \in \mathbb{N}$  mit  $\|f_n - f\|_\infty < \delta$  für alle  $n \geq N$ . Nach Proposition 6.7.1 finden wir Treppenfunktionen  $\phi, \psi: [a, b] \rightarrow \mathbb{R}$  mit  $\phi \leq f_N \leq \psi$  auf  $[a, b]$  und

$$\int_a^b (\psi(x) - \phi(x)) \, dx < \frac{\epsilon}{2}.$$

Abbildung 6.16: Die Funktionen  $f$  und  $f_n$ .

Verschieben wir eine Treppenfunktion um eine Konstante, so erhalten wir erneut eine Treppenfunktion. Insbesondere sind  $\phi - \delta$  und  $\psi + \delta$  ebenfalls Treppenfunktionen. Dann gilt

$$\int_a^b (\psi(x) + \delta - (\phi(x) - \delta)) \, dx = \int_a^b (\psi(x) - \phi(x)) \, dx + \int_a^b 2\delta \, dx < \frac{\epsilon}{2} + 2\delta(b-a) = \epsilon$$

nach unserer Wahl von  $\delta$ . Weiter gilt wegen  $\|f - f_N\|_\infty < \delta$  auch

$$f_N - \delta \leq f \leq f_N + \delta,$$

siehe auch Abbildung 6.16. Damit folgt aber

$$\phi - \delta \leq f \leq \psi + \delta,$$

womit wir nach Proposition 6.7.1 die Riemann-Integrabilität von  $f$  gezeigt haben. Der komplexe Fall  $\mathbb{K} = \mathbb{C}$  folgt durch Betrachtung von  $\operatorname{Re}(f_n)$  und  $\operatorname{Im}(f_n)$ . Mit  $f$  und  $f_n$  ist aber auch  $f - f_n \in \mathcal{R}([a, b], \mathbb{K})$ , und wir erhalten

$$\left| \int_a^b f(x) \, dx - \int_a^b f_n(x) \, dx \right| = \left| \int_a^b (f - f_n)(x) \, dx \right| \leq (b-a) \|f - f_n\|_\infty$$

aus Korollar 6.3.6. Damit ist (6.7.6) aber klar.  $\square$

Die eigentliche Konvergenzaussage (6.7.6) ist eine triviale Konsequenz von Korollar 6.3.6. Es ist die dazu notwendige Eigenschaft  $f \in \mathcal{R}([a, b], \mathbb{K})$ , welche nicht ganz offensichtlich war.

Als Warnung sei hier aber bemerkt, dass ein analoges Resultat für uneigentliche Riemann-Integrale nicht gültig ist. Hier gibt es recht einfache Gegenbeispiele:

**Beispiel 6.7.3** Wir betrachten das Intervall  $I = [0, \infty)$  mit Funktionen  $f_n: [0, \infty) \rightarrow \mathbb{R}$  mit

$$f_n(x) = \begin{cases} \frac{1}{n} & x \in [n, 2n] \\ 0 & \text{sonst,} \end{cases} \quad (6.7.7)$$

siehe Abbildung 6.17. Es gilt  $\|f_n\|_\infty = \frac{1}{n}$ , womit  $f_n \rightarrow 0$  gleichmäßig konvergiert. Als Treppenfunktion ist jedes  $f_n$  integrierbar und

$$\int_0^b f_n(x) \, dx = \frac{1}{n} \cdot (2n - n) = 1 \quad (6.7.8)$$

gilt, sobald  $b \geq 2n$  erfüllt ist. Damit existieren auch die uneigentlichen Riemann-Integrale

$$\int_0^\infty f_n(x) \, dx = 1 \quad (6.7.9)$$

für alle  $n \in \mathbb{N}$ . Daher gilt in diesem Fall

$$\lim_{n \rightarrow \infty} \int_0^\infty f_n(x) \, dx = 1 \neq 0 = \int_0^\infty \left( \lim_{n \rightarrow \infty} f_n \right)(x) \, dx, \quad (6.7.10)$$

obwohl gleichmäßige Konvergenz vorliegt.

Eine wichtige Situation, wo wir Satz 6.7.2 anwenden können, sind konvergente Potenzreihen. Hier gibt es folgendes Resultat:

Abbildung 6.17: Die Treppenfunktionen  $f_n$  konvergieren auf  $[0, \infty)$  gleichmäßig gegen Null.

**Proposition 6.7.4** Sei  $a(z) = \sum_{n=0}^{\infty} a_n z^n \in \mathbb{C}[[z]]$  eine Potenzreihe mit Konvergenzradius  $R > 0$ .

i.) Die Potenzreihe

$$A(z) = \sum_{n=1}^{\infty} \frac{1}{n} a_{n-1} z^n \quad (6.7.11)$$

hat ebenfalls den Konvergenzradius  $R$ .

ii.) Die zugehörige Funktion  $A$  ist die Stammfunktion von  $a$  mit  $A(0) = 0$ .

iii.) Für jedes Intervall  $[a, b] \subseteq (-R, R)$  gilt

$$\int_a^b a(x) dx = A(b) - A(a) = \sum_{n=0}^{\infty} \int_a^b a_n x^n dx. \quad (6.7.12)$$

BEWEIS: Wir wissen, dass  $\sum_{n=0}^{\infty} a_n z^n$  lokal gleichmäßig auf  $B_R(0)$  konvergiert, also insbesondere gleichmäßig auf dem kompakten Intervall  $[0, \rho]$ . Damit dürfen wir den Grenzwert  $\lim_{n \rightarrow \infty} \sum_{k=0}^n a_k z^k$  mit der Integration  $\int_0^\rho$  vertauschen und erhalten die Konvergenz von

$$\int_0^\rho \sum_{n=0}^{\infty} a_n x^n dx = \sum_{n=0}^{\infty} a_n \int_0^\rho x^n dx = \sum_{n=0}^{\infty} \frac{1}{n+1} a_n \rho^{n+1} = \sum_{n=1}^{\infty} \frac{1}{n} a_{n-1} \rho^n.$$

Dies zeigt, dass  $A(z)$  Konvergenzradius mindestens  $\rho$  besitzt, für alle  $\rho < R$ . Also ist der Konvergenzradius mindestens  $R$ . Wäre er echt größer als  $R$ , so dürften wir  $A(z)$  termweise differenzieren und erhielten eine konvergente Potenzreihe mit dem gleichen Konvergenzradius, echt größer als  $R$ , nach Satz 5.6.1. Diese Reihe ist aber  $a(z)$ , womit der Konvergenzradius doch  $R$  gewesen ist. Wir dürfen  $A(z)$  termweise differenzieren, was sofort  $A' = a$  zeigt. Also ist  $A$  eine Stammfunktion von  $a$ . Da  $A(0) = 0$  gilt, folgt auch ii.). Den dritten Teil erhält man nun mit dem gleichen Argument wie für die Integration über  $[0, \rho]$  zuvor.  $\square$

Als erste kleine Anwendung können wir eine Potenzreihendarstellung des Arcustangens finden:

**Beispiel 6.7.5** Für den Arcustangens gilt

$$\arctan'(x) = \frac{1}{1+x^2} \quad (6.7.13)$$

nach Proposition 5.2.7, iv.). Diese Funktion können wir als geometrische Reihe

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-x^2)^n = \sum_{n=0}^{\infty} (-1)^n x^{2n} \quad (6.7.14)$$

schreiben, wobei  $|x^2| < 1$  gelten muss. Dies ist gleichbedeutend mit  $|x| < 1$ , womit (6.7.14) also auch Konvergenzradius 1 besitzt. Termweises Integrieren liefert die konvergente Potenzreihe

$$\arctan(x) = \sum_{n=1}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \quad (6.7.15)$$

für  $|x| < 1$ , welche wegen  $\arctan(0) = 0$  tatsächlich mit  $\arctan$  übereinstimmt. Mit ähnlichen Überlegungen erhält man Potenzreihen für  $\arcsin$  und  $\arccos$ , siehe auch Übung ???. Dies ist gelegentlich einfacher, als alle Taylor-Koeffizienten bei  $x = 0$  auszurechnen und die Konvergenz der Taylor-Entwicklung von Hand zu zeigen.

## 6.8 Übungen

### Übung 6.8.1 (Darboux-Integrale)

### Übung 6.8.2 (Stückweise Riemann-integrierbare Funktionen)

### Übung 6.8.3 (Verkettung Riemann-integrierbarer Funktionen)

### Übung 6.8.4 (Riemann-Integrierbarkeit von $f$ und $|f|$ )

**Übung 6.8.5 (Riemann-Integrierbarkeit und endlicher Träger)** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall.

- i.) Sei zuerst  $f: [a, b] \rightarrow \mathbb{K}$  eine Funktion, die nur an endlich vielen Punkten  $x_1, \dots, x_n \in [a, b]$  von Null verschieden ist. Man sagt auch, dass  $f$  einen endlichen Träger besitzt. Zeigen Sie, dass  $f$  Riemann-integrierbar ist und dass

$$\int_a^b f(x) \, dx = 0 \quad (6.8.1)$$

gilt.

Hinweis: Betrachten Sie zunächst den Fall  $n = 1$ .

- ii.) Seien nun  $f, g: [a, b] \rightarrow \mathbb{K}$  zwei Funktionen, so dass  $f \in \mathcal{R}([a, b], \mathbb{K})$  gilt und  $g$  sich von  $f$  nur an endlich vielen Punkten unterscheidet. Zeigen Sie, dass dann  $g \in \mathcal{R}([a, b], \mathbb{K})$  und

$$\int_a^b f(x) \, dx = \int_a^b g(x) \, dx \quad (6.8.2)$$

gilt.

**Übung 6.8.6 (Abschätzungen und Konvergenz mit Riemann-Summen)** Seien  $n, m \in \mathbb{N}$ . Betrachten Sie dann die Folge  $(a_m)_{m \in \mathbb{N}}$  mit

$$a_m = \frac{1}{m} \sum_{k=0}^{m-1} \left(\frac{k}{m}\right)^n. \quad (6.8.3)$$

- i.) Zeigen Sie, dass  $(a_m)_{m \in \mathbb{N}}$  monoton wachsend ist.

- ii.) Zeigen Sie

$$a_m \leq \frac{1}{n+1}. \quad (6.8.4)$$

- iii.) Zeigen Sie, dass  $\frac{1}{n+1}$  der Grenzwert der Folge  $(a_m)_{m \in \mathbb{N}}$  ist.

Hinweis: Alle drei Aussagen werden sehr einfach, wenn man die Summe als Riemann-Summe einer geeigneten Funktion versteht.

### Übung 6.8.7 (Partielle Integrationen)

### Übung 6.8.8 (Das Integral $\int_0^\infty \text{sinc}(x) \, dx$ )

## Kapitel 7

# Normierte, metrische und topologische Räume

Bislang hatten wir Funktionen einer reellen oder komplexen Variablen betrachtet. Dies stellt für viele Belange eine zu restriktive Sichtweise dar. Wir benötigen daher Funktionen mehrerer reeller oder komplexer Variablen, was Definitionsbereiche

$$X \subseteq \mathbb{K}^n \tag{7.1}$$

für ein geeignetes  $n \in \mathbb{N}$ , die Zahl der Variablen, erfordert. Um dann wieder von Konvergenz, Stetigkeit sowie Differenzierbarkeit und Integrierbarkeit sprechen zu können, müssen wir diese Konzepte vom eindimensionalen Fall  $X \subseteq \mathbb{K}$  auf (7.1) verallgemeinern. Neben der Zahl der Variablen müssen wir auch allgemeinere Werte als nur reelle oder komplexe Zahlen zulassen, was allgemeinere Wertebereiche als  $\mathbb{K}$  erfordert.

Um in dieser allgemeinen Situation die bisherigen Konzepte der Analysis wiederzufinden, lohnt es sich, einen deutlich allgemeineren Rahmen zu betrachten, als nur Definitions- und Wertebereiche in  $\mathbb{K}^n$  anstelle von  $\mathbb{K}$  zuzulassen. Das soll in diesem Kapitel geschehen, indem wir die Begriffe von normierten (Vektor-)Räumen, metrischen Räumen und topologischen Räumen kennenlernen und mit ersten Eigenschaften vorstellen wollen. Das tatsächliche Studium metrischer und topologischer Räume bleibt allerdings weiterführenden Diskussionen vorbehalten.

Überraschenderweise haben wir in Kapitel 2 und Kapitel 3 an den Beispielen von  $\mathbb{R}$  und  $\mathbb{C}$  bereits die wesentlichen Eigenschaften von metrischen und topologischen Räumen kennengelernt, wenn auch noch nicht unter diesem Namen. Viele dieser Sätze und ihre Beweise können wir leicht und wörtlich übertragen.

### 7.1 Normierte Vektorräume und metrische Räume

Den Begriff der Norm haben wir implizit schon an verschiedenen Stellen gesehen. Hier geben wir nun die allgemeine Definition:

**Definition 7.1.1 (Norm)** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Eine Norm  $\|\cdot\|$  auf  $V$  ist eine Abbildung

$$\|\cdot\|: V \longrightarrow \mathbb{R} \tag{7.1.1}$$

mit folgenden Eigenschaften:

- i.) Es gilt  $\|v\| \geq 0$  für alle  $v \in V$  und  $\|v\| = 0$  genau dann, wenn  $v = 0$ .
- ii.) Es gilt  $\|\lambda v\| = |\lambda| \|v\|$  für alle  $v \in V$  und  $\lambda \in \mathbb{K}$ .
- iii.) Es gilt die Dreiecksungleichung  $\|v + w\| \leq \|v\| + \|w\|$  für alle  $v, w \in V$ .

Ein Vektorraum  $(V, \|\cdot\|)$  mit einer Norm heißt auch *normierter Vektorraum* oder kurz auch *normierter Raum*.

Die Liste der Beispiele ist lang:

### Beispiel 7.1.2 (Normierte Vektorräume)

- i.) Der Absolutbetrag  $|\cdot|$  für den eindimensionalen  $\mathbb{K}$ -Vektorraum  $\mathbb{K}$  ist eine Norm. Dies ist gerade die Aussage von Proposition 1.3.10 beziehungsweise Proposition 1.7.6. Dieses fundamentale Beispiel ist letztlich die Motivation für die Definition einer Norm, da wir die bisherige Analysis auf Abschätzungen mittels des Absolutbetrags aufgebaut haben. Es stellt sich nun heraus, dass diese drei definierenden Eigenschaften einer Norm alles sind, was wir benötigen, um analytische Konzepte auch auf allgemeine normierte Vektorräume zu übertragen.
- ii.) Sei  $p \geq 1$ . Auf  $V = \mathbb{K}^n$  definiert man die  $p$ -Norm oder auch  $\ell_p$ -Norm durch

$$\|v\|_p = \sqrt[p]{\sum_{k=1}^n |v_k|^p}, \quad (7.1.2)$$

wobei  $v \in \mathbb{K}^n$  und  $v_k \in \mathbb{K}$  die  $k$ -te Komponente von  $v$  bezüglich der Standardbasis  $e_1, \dots, e_n \in \mathbb{K}^n$  ist. Der Nachweis der ersten beiden Normeigenschaften ist einfach, die Dreiecksungleichung für  $\|\cdot\|$  ist gerade die Minkowski-Ungleichung aus Korollar 5.4.12. Der Fall  $p = 2$  führt dann auf die *Euklidische Norm*. Ist  $V$  ein beliebiger endlich-dimensionaler Vektorraum, so können wir die  $p$ -Norm auch bezüglich einer gewählten Basis von  $V$  genauso anhand der Koeffizienten der Basis definieren. Diese  $p$ -Norm hängt dann nach Definition von der gewählten Basis ab.

- iii.) Wir dehnen die Definition der  $p$ -Norm aus (7.1.2) für  $p = \infty$  aus, indem wir

$$\|v\|_\infty = \max\{|v_1|, \dots, |v_n|\} \quad (7.1.3)$$

für  $v \in \mathbb{K}^n$  setzten. Die Eigenschaften einer Norm sind elementar nachzuprüfen. Man nennt  $\|\cdot\|_\infty$  entsprechend die *Maximumsnorm* oder *Supremumsnorm*. Ebenso kann man die Maximumsnorm auch bezüglich einer Basis eines beliebigen endlich-dimensionalen Vektorraums definieren.

- iv.) Interessantere Beispiele erhalten wir für die Funktionenräume. Diese sind unendlich-dimensional und besitzen ebenfalls ausgezeichnete Normen. Für  $V = \mathcal{B}(X, \mathbb{K})$ , die beschränkten Funktionen auf einer Menge  $X$  mit Werten in  $\mathbb{K}$ , haben wir die *Supremumsnorm*

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in X\} \quad (7.1.4)$$

bereits öfters verwendet. Die Normeigenschaften wurden in Übung 3.9.21 nachgewiesen. Ein genauerer Blick zeigt, dass die Supremumsnorm aus (7.1.3) tatsächlich ein Spezialfall der Supremumsnorm aus (7.1.4) ist, wenn man die Vektoren  $v$  eines endlichdimensionalen Vektorraums als Funktionen auf der Basis als endliche Menge interpretiert.

**Bemerkung 7.1.3 (Unterräume normierter Räume)** Ist  $(V, \|\cdot\|)$  ein normierter Vektorraum über  $\mathbb{K}$  und  $U \subseteq V$  ein Unterraum, so ist die Einschränkung von  $\|\cdot\|$  auf  $U$  immer noch eine Norm, da die Eigenschaften einer Norm ja für alle Vektoren aus  $V$ , also insbesondere für die aus  $U$  gelten. Dies liefert einige weitere Beispiele von normierten Vektorräumen:

- i.) Die stetigen Funktionen  $\mathcal{C}(K, \mathbb{K})$  auf einer nicht-leeren kompakten Teilmenge  $K$  der reellen oder komplexen Zahlen sind bezüglich der Supremumsnorm  $\|\cdot\|_\infty$  von  $\mathcal{B}(K, \mathbb{K})$  wieder ein normierter Vektorraum. Hierfür ist die Inklusion  $\mathcal{C}(K, \mathbb{K}) \subseteq \mathcal{B}(K, \mathbb{K})$  aus Korollar 3.5.7 entscheidend.



- ii.) Ist  $I = [a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall, so sind auch die Riemann-integrierbaren Funktionen nach Satz 6.2.2 beschränkt. Damit erhält auch  $\mathcal{R}([a, b], \mathbb{K}) \subseteq \mathcal{B}([a, b], \mathbb{K})$  die Supremumsnorm  $\|\cdot\|_\infty$  von  $\mathcal{B}([a, b], \mathbb{K})$ .

In beiden Fällen ist es natürlich wichtig, vorher zu zeigen, dass die stetigen beziehungsweise die Riemann-integrierbaren Funktionen überhaupt Unterräume sind.

**Bemerkung 7.1.4** Wie immer in der Mathematik will man neue Beispiele aus alten konstruieren, indem man verschiedene kanonische Verfahren anwendet. Im Falle normierter Vektorräume kann man versuchen, die üblichen Konstruktionen der linearen Algebra dazu zu verwenden. Hier gibt es unter anderem folgende Optionen:

- i.) Sind  $(V, \|\cdot\|_V)$  und  $(W, \|\cdot\|_W)$  normierte Vektorräume, so können wir auf der direkten Summe  $V \oplus W$  eine Norm durch

$$\|v + w\|_p = \sqrt[p]{\|v\|_V^p + \|w\|_W^p} \quad (7.1.5)$$

für  $p \in [1, \infty)$  und

$$\|v + w\|_\infty = \max\{\|v\|_V, \|w\|_W\} \quad (7.1.6)$$

definieren. Der Nachweis der Normeigenschaften erfolgt wie bei der  $p$ -Norm für  $\mathbb{K}^n$  in Beispiel 7.1.2, ii.) beziehungsweise iii.), siehe auch Übung 7.8.2.

- ii.) Ist  $(V, \|\cdot\|_V)$  ein normierter Vektorraum, so möchte man auf seinem Dualraum  $V^*$  ebenfalls eine Norm etablieren. Im allgemeinen ist dies nicht so einfach möglich. Statt dessen betrachtet man den *stetigen Dualraum*

$$V' = \{\varphi \in V^* \mid \sup\{\|v\|_V \mid v \in V \text{ mit } |\varphi(v)| \leq 1\} < \infty\}. \quad (7.1.7)$$

Es zeigt sich, dass  $V' \subseteq V^*$  ein im allgemeinen echter Unterraum ist. Für  $\varphi \in V'$  ist

$$\|\varphi\|_{V'} = \sup\{|\varphi(v)| \mid v \in V \text{ mit } \|v\|_V \leq 1\} \quad (7.1.8)$$

dann eine Norm, die sogenannte *Funktionalnorm*. Die Details und eine konzeptuelle Einordnung werden wir später noch kennenlernen. Man beachte jedoch, dass in der Literatur die Bezeichnungen leider nicht einheitlich sind und gelegentlich  $V^*$  der topologische Dualraum und  $V'$  der algebraische Dualraum ist.

- iii.) Sind  $(V, \|\cdot\|_V)$  und  $(W, \|\cdot\|_W)$  normierte Vektorräume, so kann man ii.) dahingehend verallgemeinern, dass man allgemeine lineare Abbildungen  $\text{Hom}(V, W)$  betrachtet. Wieder ist es nicht leicht möglich, eine sinnvolle Norm auf  $\text{Hom}(V, W)$  zu konstruieren. Analog zu (7.1.7) betrachtet man

$$L(V, W) = \{A \in \text{Hom}(V, W) \mid \sup\{\|A(v)\|_W \mid v \in V \text{ mit } \|v\|_V \leq 1\} < \infty\}. \quad (7.1.9)$$

und

$$\|A\| = \sup\{\|A(v)\|_W \mid v \in V \text{ mit } \|v\|_V \leq 1\} \quad (7.1.10)$$

für  $A \in L(V, W)$ . Es zeigt sich, dass  $L(V, W) \subseteq \text{Hom}(V, W)$  ein im allgemeinen echter Unterraum ist und (7.1.10) auf diesem eine Norm liefert, die *Operatornorm*. Die Details werden wir auch hierzu später sehen. Man beachte jedoch, dass der Spezialfall  $W = \mathbb{K}$  uns zunächst zu ii.) zurückführt.

- iv.) Das Tensorprodukt  $V \otimes W$  von normierten Vektorräumen kann auf sehr viele Weisen wieder mit einer Norm versehen werden, deren Konstruktion aber erheblich schwieriger und weniger kanonisch ist.

Wir wollen in normierten Räumen nun alle bisher gefundenen Konzepte der Analysis zum Einsatz bringen. Das ist naheliegend, da wir den Absolutbetrag von  $\mathbb{K}$  als Spezialfall einer Norm verstehen können. Die wichtigste Beobachtung dabei ist, dass eine Norm einen Abstandsbegriff liefert:

**Proposition 7.1.5** *Sei  $(V, \|\cdot\|_V)$  ein normierter Vektorraum. Die Abbildung*

$$d: V \times V \ni (v, w) \mapsto d(v, w) = \|v - w\| \in \mathbb{R} \quad (7.1.11)$$

*erfüllt folgende Eigenschaften:*

- i.) Für alle  $v, w \in V$  gilt  $d(v, v) = 0$ , und  $d(v, w) = 0$  gilt nur für  $v = w$ .*
- ii.) Für alle  $v, w \in V$  gilt  $d(v, w) = d(w, v)$ .*
- iii.) Für alle  $v, u, w \in V$  gilt die Dreiecksungleichung*

$$d(v, w) \leq d(v, u) + d(u, w). \quad (7.1.12)$$

BEWEIS: Alle behaupteten Aussagen folgen direkt aus den entsprechenden definierenden Eigenschaften der Norm. Im Detail gilt  $d(v, w) = \|v - w\| = 0$  genau dann, wenn  $v - w = 0$  also  $v = w$  gilt. Weiter gilt  $\|v - w\| = \|(-1)(w - v)\| = |-1| \cdot \|w - v\| = \|w - v\|$  sowie schließlich

$$\begin{aligned} d(v, w) &= \|v - w\| \\ &= \|v - u + u - w\| \\ &\leq \|v - u\| + \|u - w\| \\ &= d(v, u) + d(u, w), \end{aligned}$$

womit der Beweis erbracht ist. □

Die Eigenschaften *i.) – iii.)* für  $d$  aus (7.1.11) legen die geometrische Interpretation von  $d(v, w)$  als *Abstand* von  $v$  und  $w$  nahe. Insbesondere wird kein Bezug mehr auf die lineare Struktur von  $V$  genommen, womit wir folgende, drastische Verallgemeinerung vornehmen können:

**Definition 7.1.6 (Metrischer Raum)** *Sei  $M$  eine nicht-leere Menge und*

$$d: M \times M \longrightarrow \mathbb{R} \quad (7.1.13)$$

*eine Abbildung mit*

$$d(x, y) = 0 \iff x = y, \quad (7.1.14)$$

$$d(x, y) = d(y, x) \quad (7.1.15)$$

*und*

$$d(x, y) \leq d(x, z) + d(z, y) \quad (7.1.16)$$

*für alle  $x, y, z \in M$ . Dann heißt  $d$  eine Metrik (oder Distanzfunktion) für  $M$ , und  $(M, d)$  heißt metrischer Raum. Eine Abbildung*

$$\Phi: (M_1, d_1) \longrightarrow (M_2, d_2) \quad (7.1.17)$$

*zwischen zwei metrischen Räumen heißt isometrisch, falls*

$$d_2(\Phi(x), \Phi(y)) = d_1(x, y) \quad (7.1.18)$$

*für alle  $x, y \in M_1$  gilt.*

Abbildung 7.1: Die elementargeometrische Dreiecksungleichung in der Ebene.

Insbesondere wird ein normierter Vektorraum  $(V, \|\cdot\|)$  durch Proposition 7.1.5 zu einem metrischen Raum. Im folgenden werden wir immer diese Metrik für einen normierten Vektorraum verwenden.

**Bemerkung 7.1.7** Ist  $(M, d)$  ein metrischer Raum, so gilt  $d(x, y) \geq 0$  für alle  $x, y \in M$ . Dies folgt aus der Dreiecksungleichung, da

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y). \quad (7.1.19)$$

Abstände sind eben positiv oder null. Die Dreiecksungleichung bezieht ihren Namen nun aus der elementargeometrischen Situation, dass die Summe zweier Seitenlängen in einem Dreieck mindestens die verbleibende Seitenlänge ist, siehe Abbildung 7.1. In einem metrischen Raumes wird diese Vorstellung Teil der Definition.

Neben den normierten Vektorräumen mit ihrer kanonischen Metrik aus Proposition 7.1.5 gibt es viele weitere Beispiele. Insbesondere wird eine zugrundeliegende Vektorraumstruktur *nicht* benötigt:

**Beispiel 7.1.8 (Metrische Räume)** Sei  $M$  eine nicht-leere Menge.

i.) Mit

$$d(x, y) = \begin{cases} 1 & \text{für } x \neq y \\ 0 & \text{für } x = y \end{cases} \quad (7.1.20)$$

für  $x, y \in M$  wird auf  $M$  eine Metrik definiert, die *diskrete Metrik*. Der Nachweis der Eigenschaften einer Metrik ist elementar. Insbesondere können wir so jede Menge zu einem diskreten metrischen Raum aufwerten.

ii.) Ist  $d$  eine Metrik auf  $M$  und  $U \subseteq M$  eine nicht-leere Teilmenge, so ist die Einschränkung

$$d_U = d|_{U \times U} : U \times U \longrightarrow \mathbb{R} \quad (7.1.21)$$

eine Metrik auf  $U$ . Dies ist klar, da die gewünschten Eigenschaften für alle Punkte aus  $M$  gelten, also insbesondere für die aus  $U$ . Damit können wir also die Metrik an jede Teilmenge eines metrischen Raumes vererben. Wir werden dies stillschweigend immer tun und die eingeschränkte Metrik weiterhin mit  $d$  bezeichnen. Etwas konzeptueller kann man  $d_U$  so verstehen, dass die Inklusionsabbildung  $i: U \longrightarrow M$  der Teilmenge isometrisch ist.

iii.) Ist  $d$  eine Metrik auf  $M$ , so wird durch

$$d'(x, y) = \frac{d(x, y)}{1 + d(x, y)} \quad (7.1.22)$$

für  $x, y \in M$  eine neue Metrik  $d'$  definiert, siehe Übung 7.8.4. Interessant ist nun, dass bezüglich  $d'$  alle Abstände kleiner als 1 sind. Die genaue Beziehung der Metriken  $d$  und  $d'$  werden wir noch untersuchen.

## 7.2 Konvergenz in metrischen Räumen

Die Definition einer konvergenten Folge aus Definition 2.2.1 können wir direkt auf Folgen in metrischen Räumen übertragen:

**Definition 7.2.1 (Konvergenz von Folgen II)** Sei  $(M, d)$  ein metrischer Raum. Eine Folge  $(p_n)_{n \in \mathbb{N}}$  von Punkten  $p_n \in M$  heißt konvergent gegen einen Punkt  $p \in M$ , wenn es zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  mit

$$d(p_n, p) < \epsilon \quad (7.2.1)$$

für alle  $n \geq N$  gibt. Wir schreiben in diesem Fall

$$p = \lim_{n \rightarrow \infty} p_n \quad (7.2.2)$$

und nennen  $p$  den Grenzwert von  $(p_n)_{n \in \mathbb{N}}$ . Analog definiert man die Konvergenz von Netzen in  $M$ .

**Bemerkung 7.2.2** Sei  $(M, d)$  ein metrischer Raum.

- i.) Für  $M = \mathbb{K}$  mit der vom Absolutbetrag induzierten Metrik reproduziert Definition 7.2.1 gerade unsere vorherige Definition 2.2.1. Dies ist wegen  $d(p_n, p) = |p_n - p|$  unmittelbar klar.
- ii.) Der Beweis von Proposition 2.2.3 überträgt sich sofort auf den allgemeinen Fall, womit der Grenzwert einer konvergenten Folge *eindeutig* ist.
- iii.) Mit dem gleichen Argument lassen sich Resultate übertragen, die nur die Eigenschaften des Absolutbetrags verwenden. Insbesondere ist für eine konvergente Folge  $(p_n)_{n \in \mathbb{N}}$  auch jede Teilfolge und jede Umordnung konvergent, mit gleichem Grenzwert.
- iv.) Die Rechenregeln für konvergente Folgen von reellen oder komplexen Zahlen aus Abschnitt 2.3 benutzen explizit die algebraische Struktur von  $\mathbb{K}$ . In einem metrischen Raum stehen diese im Allgemeinen nicht zur Verfügung, womit die Frage nach Rechenregeln wie in Abschnitt 2.3 hinfällig ist.

Wie schon in Abschnitt 2.5 können wir die Konvergenz auch mit Hilfe von Umgebungen formulieren. Dazu übertragen wir die Konstruktion von Umgebungen und offenen Mengen in allgemeine metrische Räume. Der Schlüssel hierzu ist, dass wir nach wie vor von *Kugeln* sprechen können, da wir über einen Abstandsbegriff verfügen:

**Definition 7.2.3 (Umgebungen, offene und abgeschlossene Teilmengen)** Sei  $(M, d)$  ein metrischer Raum.

- i.) Für  $r > 0$  und  $p \in M$  heißt

$$B_r(p) = \{q \in M \mid d(p, q) < r\} \subseteq M \quad (7.2.3)$$

die offene Kugel um  $p$  mit Radius  $r$ . Entsprechend heißt

$$B_r(p)^{\text{cl}} = \{q \in M \mid d(p, q) \leq r\} \subseteq M \quad (7.2.4)$$

die abgeschlossene Kugel um  $p$  mit Radius  $r$ .

- ii.) Eine Teilmenge  $U \subseteq M$  heißt Umgebung von  $p \in M$ , wenn es ein  $r > 0$  mit

$$B_r(p) \subseteq U \quad (7.2.5)$$

gibt.

- iii.) Eine Teilmenge  $O \subseteq M$  heißt offen, wenn sie Umgebung aller Punkte  $p \in O$  ist.
- iv.) Eine Teilmenge  $A \subseteq M$  heißt abgeschlossen, wenn  $M \setminus A$  offen ist.

Je nach metrischem Raum ist die Form einer Kugel nicht sonderlich kugelförmig, siehe etwa Abbildung 7.2. Entscheidend für das weitere Vorgehen ist dagegen, dass die Eigenschaften offener und abgeschlossener Mengen in Abschnitt 2.5 allein aus den metrischen Eigenschaften offener und abgeschlossener Kugeln gewonnen wurden. Dies liefert sofort folgendes Resultat:

Abbildung 7.2: Die Kugeln um  $0 \in \mathbb{R}^2$  mit Radius 1 bezüglich der Metriken zu  $p$ -Normen  $\|\cdot\|_p$  für  $p = 1, 2$  und  $\infty$ .

Abbildung 7.3: Verschiedene Punkte lassen sich durch disjunkte offene Kugeln trennen.

**Proposition 7.2.4** *Sei  $(M, d)$  ein metrischer Raum.*

- i.) Ist  $U \subseteq M$  eine Umgebung von  $p \in M$  und  $U \subseteq U'$  so ist auch  $U'$  eine Umgebung von  $p$ .*
- ii.) Eine offene Kugel ist offen.*
- iii.) Die leere Menge  $\emptyset \subseteq M$  und  $M$  sind offen.*
- iv.) Sind  $\{O_i\}_{i \in I}$  offene Teilmengen von  $M$ , so ist auch  $\bigcup_{i \in I} O_i \subseteq M$  wieder offen.*
- v.) Sind  $O_1, \dots, O_n \subseteq M$  offen, so ist auch  $O_1 \cap \dots \cap O_n \subseteq M$  offen.*
- vi.) Eine abgeschlossene Kugel ist abgeschlossen. Insbesondere ist auch  $\{p\} \subseteq M$  abgeschlossen.*
- vii.) Ist  $U \subseteq M$  eine Umgebung von  $p \in M$ , so gibt es eine offene Umgebung  $V \subseteq U$  von  $p$ .*
- viii.) Zu  $p, q \in M$  mit  $p \neq q$  gibt es offene Teilmengen  $O_1, O_2 \subseteq M$  mit  $p \in O_1$  und  $q \in O_2$  sowie*

$$O_1 \cap O_2 = \emptyset. \quad (7.2.6)$$

BEWEIS: Der Beweis von *i.) – vii.)* ist völlig parallel zum Beweis der entsprechenden Aussagen von Proposition 2.5.2 und beruht letztendlich darin, die Verwendung des Absolutbetrags in Proposition 2.5.2 durch eine allgemeine Metrik zu ersetzen. Lediglich der letzte Teil bedarf eines kleinen Arguments. Ist  $p \neq q$ , so ist  $r = d(p, q) > 0$ . Dann gilt

$$B_{\frac{r}{2}}(p) \cap B_{\frac{r}{2}}(q) = \emptyset,$$

denn ist  $z \in B_{\frac{r}{2}}(p) \cap B_{\frac{r}{2}}(q)$ , so gilt  $d(p, z) < \frac{r}{2}$  sowie  $d(q, z) < \frac{r}{2}$ , also nach der Dreiecksungleichung

$$d(p, q) < d(p, z) + d(q, z) < \frac{r}{2} + \frac{r}{2} = r,$$

ein Widerspruch. Da offene Kugeln offen sind, können wir  $O_1 = B_{\frac{r}{2}}(p)$  und  $O_2 = B_{\frac{r}{2}}(q)$  setzen, siehe auch Abbildung 7.3.  $\square$

Wie zuvor für  $\mathbb{K}$  können wir einen Häufungspunkt einer Folge in einem metrischen Raum definieren:

**Definition 7.2.5 (Häufungspunkt einer Folge II)** *Sei  $(M, d)$  ein metrischer Raum und  $(p_n)_{n \in \mathbb{N}}$  eine Folge in  $M$ . Ein Punkt  $p \in M$  heißt dann Häufungspunkt von  $(p_n)_{n \in \mathbb{N}}$ , wenn jede Umgebung von  $p$  unendlich viele Folgenglieder von  $(p_n)_{n \in \mathbb{N}}$  enthält. Die Menge der Häufungspunkte von  $(p_n)_{n \in \mathbb{N}}$  wird mit  $\text{HP}((p_n)_{n \in \mathbb{N}})$  bezeichnet,*

Mit dieser Definition analog zu Definition 2.5.6 erhalten wir dann alle Konvergenzeigenschaften von Folgen und ihren Teilfolgen wie in Abschnitt 2.5. Wir fassen die Ergebnisse in einer Proposition zusammen:

**Proposition 7.2.6** *Sei  $(M, d)$  ein metrischer Raum.*

- i.) Eine Folge  $(p_n)_{n \in \mathbb{N}}$  in  $M$  konvergiert genau dann gegen  $p \in M$ , wenn jede Umgebung von  $p$  fast alle Folgenglieder von  $(p_n)_{n \in \mathbb{N}}$  enthält.*
- ii.) Eine Teilmenge  $A \subseteq M$  ist genau dann abgeschlossen, wenn für jede konvergente Folge  $(p_n)_{n \in \mathbb{N}}$  in  $M$  mit  $p_n \in A$  auch  $\lim_{n \rightarrow \infty} p_n \in A$  gilt.*

Abbildung 7.4: Die Funktion  $f(x) = \frac{x}{1+x}$  und ihre Umkehrfunktion  $f^{-1}(\xi) = \frac{\xi}{1-\xi}$ .

- iii.) Eine konvergente Folge besitzt genau einen Häufungspunkt, ihren Grenzwert.
- iv.) Ein Punkt  $p \in M$  ist genau dann ein Häufungspunkt einer Folge  $(p_n)_{n \in \mathbb{N}}$  in  $M$ , wenn es eine Teilfolge von  $(p_n)_{n \in \mathbb{N}}$  gibt, die gegen  $p$  konvergiert.
- v.) Die Menge der Häufungspunkte  $\text{HP}((p_n)_{n \in \mathbb{N}}) \subseteq M$  einer Folge  $(p_n)_{n \in \mathbb{N}}$  in  $M$  ist abgeschlossen.

BEWEIS: Der Beweis beruht darin, die Beweise von Proposition 2.5.4, Proposition 2.5.5, Proposition 2.5.7, Proposition 2.5.9 sowie Satz 2.5.11, i.), so zu interpretieren, dass der Absolutbetrag von  $\mathbb{K}$  durch die Metrik von  $M$  ersetzt wird. Dann können die Argumente wörtlich übernommen werden, da die dortigen Beweise lediglich mit Eigenschaften offener Kugeln argumentieren, die wir nach Proposition 7.2.4 allgemein zu Verfügung haben.  $\square$

**Bemerkung 7.2.7** Auch wenn Proposition 7.2.6 nahelegt, dass die Verallgemeinerung von  $\mathbb{K}$  mit Absolutbetrag zu einem beliebigen metrischen Raum  $(M, d)$  zwar weitreichend aber letztlich harmlos ist, gibt es sehr wohl Resultate aus Kapitel 2 und Kapitel 3, die sich keineswegs so naiv übertragen lassen. Hierzu zählt beispielsweise Korollar 2.5.10. In einem metrischen Raum braucht eine beschränkte Folge, also eine Folge, die sich innerhalb einer groß genug gewählten Kugel bewegt, keinen Häufungspunkt zu besitzen. Ein Gegenbeispiel ist schnell zu Hand. Wir wählen  $M = \mathbb{N}$  mit der diskreten Metrik: Alle Abstände sind 1. Die Folge  $(n)_{n \in \mathbb{N}}$  ist dann in dieser Metrik beschränkt, da beispielsweise

$$n \in B_1(1)^{\text{cl}} \quad (7.2.7)$$

wegen  $d(1, n) = 1$  für alle  $n \in \mathbb{N}$  gilt. Gibt es nun einen Häufungspunkt  $n_0 \in \mathbb{N}$ , so müssten insbesondere in der  $\frac{1}{2}$ -Kugel um  $n_0$  unendlich viele Folgenglieder sein. Dies ist aber sicher nicht der Fall, da

$$B_{\frac{1}{2}}(n_0) = \{n \in \mathbb{N} \mid d(n, n_0) < \frac{1}{2}\} = \{n_0\} \quad (7.2.8)$$

nur aus dem Punkt  $n_0$  besteht. Es ist also durchaus Vorsicht angebracht, wenn man bekannte Resultate von  $\mathbb{K}$  zu einem beliebigen metrischen Raum verallgemeinern will.

Die Begriffe von Umgebungen, offenen und abgeschlossenen Teilmengen sind recht robust und hängen nicht stark von der verwendeten Metrik ab. Die Metrik selbst beinhaltet deutlich mehr Information, wie folgende Beobachtung zeigt:

**Proposition 7.2.8** Sei  $(M, d)$  ein metrischer Raum und

$$d'(p, q) = \frac{d(p, q)}{1 + d(p, q)} \quad (7.2.9)$$

die Metrik aus Beispiel 7.1.8, iii.). Dann stimmen die offenen Mengen bezüglich  $d$  und  $d'$  überein.

BEWEIS: Die Funktion  $f: [0, \infty) \rightarrow [0, 1)$  definiert durch

$$f(x) = \frac{x}{1+x}$$

ist strikt monoton wachsend und surjektiv, also eine Bijektion, siehe auch Abbildung 7.4. Die Umkehrfunktion ist schnell als

$$f^{-1}(\xi) = \frac{\xi}{1-\xi}$$

berechnet. Nach Konstruktion von  $d'$  gilt  $d'(p, q) = f(d(p, q))$  und damit

$$\begin{aligned} B_{r,d'}(p) &= \{q \in M \mid d'(p, q) < r\} \\ &= \{q \in M \mid f(d(p, q)) < r\} \\ &= \{q \in M \mid d(p, q) < f^{-1}(r)\} \\ &= B_{f^{-1}(r),d}(p), \end{aligned}$$

womit bis auf Reskalierung der Radien mittels  $f$  beziehungsweise  $f^{-1}$  die offenen Kugeln von  $d$  und  $d'$  übereinstimmen. Da aus den offenen Kugeln die übrigen offenen Mengen nach Definition 7.2.3, *iii.*), konstruiert werden, folgt die Behauptung.  $\square$

### 7.3 Stetigkeit und Lipschitz-Stetigkeit

In einem nächsten Schritt geht es nun darum, Stetigkeit auch für Abbildungen zwischen metrischen Räumen zu etablieren. Ausgehend von der bisherigen skalaren Situation haben wir verschiedene Optionen: Folgenstetigkeit,  $\epsilon\delta$ -Stetigkeit oder die Charakterisierung mittels offener Mengen. Für metrische Räume können wir alle drei Konzepte übertragen und als nach wie vor äquivalent nachweisen:

**Satz 7.3.1 (Stetigkeit bei metrischen Räumen)** *Sei  $f: (M, d) \rightarrow (M', d')$  eine Abbildung zwischen metrischen Räumen und  $p \in M$ . Dann sind folgende Aussagen äquivalent:*

*i.) Für jede konvergente Folge  $(p_n)_{n \in \mathbb{N}}$  in  $M$  mit  $\lim_{n \rightarrow \infty} p_n = p$  ist  $(f(p_n))_{n \in \mathbb{N}}$  konvergent in  $M'$  mit*

$$\lim_{n \rightarrow \infty} f(p_n) = f(p) = f\left(\lim_{n \rightarrow \infty} p_n\right). \quad (7.3.1)$$

*ii.) Für alle  $\epsilon > 0$  gibt es ein  $\delta > 0$  derart, dass für  $q \in M$  mit  $d(p, q) < \delta$  auch  $d'(f(p), f(q)) < \epsilon$  gilt, also*

$$B_\delta(p) \subseteq f^{-1}(B_\epsilon(f(p))). \quad (7.3.2)$$

*iii.) Für jede Umgebung  $U' \subseteq M'$  von  $f(p)$  ist  $U = f^{-1}(U') \subseteq M$  eine Umgebung von  $p$ .*

BEWEIS: Wir können wieder die Argumente aus Abschnitt 3.2 wiederholen und auf metrische Räume übertragen. Es gelte *i.*) und wir nehmen an, *ii.*) sei falsch. Dann finden wir also ein  $\epsilon > 0$  derart, dass es zu jedem  $\delta > 0$  ein  $p_\delta \in B_\delta(p)$  mit  $d(f(p), f(p_\delta)) \geq \epsilon$  gibt. Wir setzen  $\delta = \frac{1}{n}$  und erhalten so eine Folge  $p_n \in B_{\frac{1}{n}}(p)$ . Da  $d(p, p_n) < \frac{1}{n}$ , konvergiert diese Folge gegen  $p$ . Es gilt aber  $d(f(p), f(p_n)) \geq \epsilon$  für alle  $n$  ein Widerspruch zu *i.*). Wir nehmen *ii.*) an. Ist  $U' \subseteq M'$  eine Umgebung von  $f(p)$ , so gibt es ein  $\epsilon > 0$  mit  $B_\epsilon(f(p)) \subseteq U'$  nach Definition einer Umgebung. Nach *ii.*) finden wir ein passendes  $\delta > 0$  mit (7.3.2). Dann gilt also

$$B_\delta(p) \subseteq f^{-1}(B_\epsilon(f(p))) \subseteq f^{-1}(U') = U,$$

da Urbilder inklusionserhaltend sind. Also ist  $U$  eine Umgebung, was *iii.*) zeigt. Gelte schließlich *iii.*). Sei dann  $(p_n)_{n \in \mathbb{N}}$  eine Folge in  $M$  mit  $p_n \rightarrow p$ . Sei weiter  $U' \subseteq M'$  eine beliebige Umgebung von  $f(p)$ , so dass nach *iii.*) also  $U = f^{-1}(U') \subseteq M$  eine Umgebung von  $p$  ist. Nach Proposition 7.2.6, *i.*), liegen fast alle  $p_n$  in  $U$ . Damit liegen ihre Bilder aber in  $f(U) = f(f^{-1}(U')) \subseteq U'$ , womit fast alle  $f(p_n)$  in  $U'$  liegen. Dies ist gleichbedeutend mit  $f(p_n) \rightarrow f(p)$ , wieder nach Proposition 7.2.6, *i.*).  $\square$

Wir nehmen diese Äquivalenz nun als Motivation für die Definition von Stetigkeit an einem Punkt. Alternativ kann man auch eine Netz-basierte Version hinzufügen, siehe Übung ??, um die Vorgehensweise in Kapitel 3 auch hier komplett umzusetzen.

**Definition 7.3.2 (Stetigkeit II)** *Sei  $f: (M, d) \rightarrow (M', d')$  eine Abbildung zwischen metrischen Räumen.*

- i.) Die Abbildung  $f$  heißt stetig bei  $p \in M$ , wenn sie eine der Eigenschaften aus Satz 7.3.1 erfüllt.
- ii.) Die Abbildung heißt stetig, falls sie stetig an jedem Punkt von  $M$  ist. Die Menge der stetigen Abbildungen wird mit

$$\mathcal{C}(M, M') = \{f: M \longrightarrow M' \mid f \text{ ist stetig}\} \quad (7.3.3)$$

bezeichnet.

**Korollar 7.3.3** Sei  $f: (M, d) \longrightarrow (M', d')$  eine Abbildung zwischen metrischen Räumen. Die Abbildung  $f$  ist genau dann stetig, wenn für jede offene Teilmenge  $O' \subseteq M'$  das Urbild  $f^{-1}(O') = O \subseteq M$  wieder offen ist.

BEWEIS: Dies zeigt man wortwörtlich genauso wie in Satz 3.2.7, ii.). □

Da Urbilder und Komplemente kompatibel sind, können wir Stetigkeit alternativ auch mit abgeschlossenen Mengen anstelle offener Mengen charakterisieren: Eine Abbildung ist genau dann stetig, wenn das Urbild jeder abgeschlossenen Teilmenge wieder abgeschlossen ist. Da wir stetige Funktionen nochmal in einem allgemeineren Kontext diskutieren werden, wollen wir an dieser Stelle die (gültigen) Eigenschaften aus Kapitel 3 nicht weiter verfolgen. Dies holen wir in Abschnitt 7.5 nach. Stattdessen verschärfen wir die Definition von Stetigkeit, um expliziter Gebrauch von der Metrik zu machen: Die obige Stetigkeit verwendet letztlich ja nur die offenen Mengen, während Proposition 7.2.8 zeigt, dass sehr unterschiedliche Metriken zu den gleichen offenen Mengen und damit zum gleichen Stetigkeitsbegriff führen können. Dies ist bei der *Lipschitz-Stetigkeit* nicht der Fall:

**Definition 7.3.4 (Lipschitz-Stetigkeit)** Sei  $f: (M, d) \longrightarrow (M', d')$  eine Abbildung zwischen metrischen Räumen.

- i.) Die Abbildung  $f$  heißt Lipschitz-stetig mit Lipschitz-Konstante  $\lambda \geq 0$ , falls

$$d'(f(p), f(q)) \leq \lambda d(p, q) \quad (7.3.4)$$

für alle  $p, q \in M$  gilt.

- ii.) Eine Lipschitz-stetige Funktion  $f$  mit Lipschitz-Konstante  $\lambda < 1$  heißt Kontraktion.
- iii.) Die Abbildung  $f$  heißt lokal Lipschitz-stetig, falls es zu jedem  $p \in M$  eine Umgebung  $U \subseteq M$  gibt, so dass  $f|_U: U \longrightarrow M'$  Lipschitz-stetig ist. Die Menge der lokal Lipschitz-stetigen Abbildungen wird auch mit

$$\mathcal{Lip}(M, M') = \{f: M \longrightarrow M' \mid f \text{ ist lokal Lipschitz-stetig}\} \quad (7.3.5)$$

bezeichnet.

Lipschitz-stetige Abbildungen sind insbesondere stetig und verhalten sich auch in anderer Hinsicht gut:

**Proposition 7.3.5**

- i.) Die Verkettung von (lokal) Lipschitz-stetigen Abbildungen ist (lokal) Lipschitz-stetig.
- ii.) Eine lokal Lipschitz-stetige Abbildung ist stetig.
- iii.) Eine Lipschitz-stetige Abbildung ist lokal Lipschitz-stetig.

BEWEIS: Sei  $f: (M, d) \longrightarrow (M', d')$  lokal Lipschitz-stetig. Für  $p_0 \in M$  finden wir dann eine Umgebung  $U \subseteq M$  von  $p_0$  und eine lokale Lipschitz-Konstante  $\lambda \geq 0$  mit

$$d'(f(p), f(q)) \leq \lambda d(p, q) \quad (*)$$



für alle  $p, q \in U$ . Ist nun  $(p_n)_{n \in \mathbb{N}}$  eine Folge in  $M$  mit  $p_n \rightarrow p$ , so liegen alle bis auf endlich viele Folgenglieder in  $U$ . Wir wenden daher  $(*)$  an und erhalten

$$d'(f(p_n), f(p_0)) \leq \lambda d(p_n, p_0)$$

für fast alle  $n \in \mathbb{N}$ . Da nun  $d(p_n, p_0) \rightarrow 0$  für  $n \rightarrow \infty$  gilt, folgt  $d'(f(p_n), f(p_0)) \rightarrow 0$ , was die Konvergenz  $f(p_n) \rightarrow f(p_0)$  zeigt. Also ist  $f$  stetig bei  $p_0$ . Da  $p_0 \in M$  beliebig war, ist  $f$  stetig auf ganz  $M$ . Dies zeigt *ii.*). Sei nun weiter  $g: (M', d') \rightarrow (M'', d'')$  ebenfalls lokal Lipschitz-stetig. Dann finden wir zudem eine Umgebung  $U' \subseteq M'$  von  $f(p_0)$  und eine zugehörige lokale Lipschitz-Konstante  $\lambda' \geq 0$  mit

$$d''(g(x), g(y)) \leq \lambda' d'(x, y) \quad (**)$$

für alle  $x, y \in U'$ . Da  $f$  stetig ist, ist  $f^{-1}(U')$  eine Umgebung von  $p_0$  und somit auch  $f^{-1}(U') \cap U$ . Für diese verkleinerte Umgebung gilt dann  $f(q) \in U'$  für alle  $q \in f^{-1}(U') \cap U$  und daher

$$d''(g(f(p)), g(f(q))) \leq \lambda' d'(f(p), f(q)) \leq \lambda' \lambda d(p, q)$$

für alle  $p, q \in U$ . Damit ist  $g \circ f$  auf der Umgebung  $f^{-1}(U') \cap U$  von  $p_0$  Lipschitz-stetig mit lokaler Lipschitz-Konstante  $\lambda\lambda' > 0$ , was den ersten Teil zeigt. Der letzte Teil ist klar.  $\square$

Wir kommen nun zu einigen ersten Beispielen für Lipschitz-stetige Abbildungen:

### Beispiel 7.3.6 (Lipschitz-stetige Abbildungen)

i.) Sei  $(V, \|\cdot\|)$  ein normierter Vektorraum. Dann ist die Norm  $\|\cdot\|: V \rightarrow \mathbb{R}$  Lipschitz-stetig mit Lipschitz-Konstante  $\lambda = 1$ , denn

$$\left| \|v\| - \|w\| \right| \stackrel{(a)}{\leq} \|v - w\| = d(v, w) \quad (7.3.6)$$

für  $v, w \in V$ . Hier haben wir in (a) die (umgekehrte) Dreiecksungleichung für  $\|\cdot\|$  verwendet, siehe auch Übung 7.8.1. Genauso zeigt man, dass die Metrik eines metrischen Raumes Lipschitz-stetig ist, siehe Übung ??.

ii.) Die Abbildung  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f(x) = x^2 \quad (7.3.7)$$

ist lokal Lipschitz-stetig aber nicht Lipschitz-stetig. Zunächst gilt für  $x, y \in [-R, R]$  nämlich

$$|x^2 - y^2| = |(x + y)(x - y)| \leq 2R|x - y|, \quad (7.3.8)$$

womit wir auf  $[-R, R]$  eine lokale Lipschitz-Konstante  $\lambda = 2R$  gefunden haben. Da jeder Punkt in einem solchen Intervall enthalten ist, ist  $f$  lokal Lipschitz-stetig. Eine globale Lipschitz-Konstante kann es nicht geben, da

$$\sup_{x \neq y} \frac{x^2 - y^2}{x - y} = \sup_{x \neq y} |x + y| = \infty \quad (7.3.9)$$

unbeschränkt ist.

iii.) Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f(x) = \sqrt{|x|} \quad (7.3.10)$$

ist stetig aber nicht lokal Lipschitz-stetig in der Nähe von  $x = 0$ , siehe auch Abbildung 7.5. Es gilt nämlich auf jeder Umgebung  $U$  von 0

$$\sup_{\substack{x \neq y \\ x, y \in U}} \frac{|f(x) - f(y)|}{|x - y|} = \sup_{\substack{x \neq y \\ x, y \in U}} \frac{\sqrt{|x|} - \sqrt{|y|}}{|x - y|}$$

Abbildung 7.5: Die Funktion  $f(x) = \sqrt{|x|}$

$$\begin{aligned} &\geq \sup_{\substack{x \neq 0 \\ x \in U}} \frac{\sqrt{|x|}}{|x|} \\ &= \sup_{\substack{x \neq 0 \\ x \in U}} \frac{1}{\sqrt{|x|}} \\ &= \infty. \end{aligned}$$

Damit kann es keine lokale Lipschitz-Konstante um  $x = 0$  geben. Auf  $\mathbb{R} \setminus \{0\}$  ist  $f$  dagegen lokal Lipschitz-stetig, siehe auch Übung ??.

*iv.)* Die Schwierigkeit in *iii.)* lässt sich geometrisch mit der divergenten Ableitung von  $f$  für  $x \rightarrow 0$  verstehen. Es gilt nämlich allgemein folgendes Resultat: Sei  $f \in \mathcal{R}([a, b], \mathbb{K})$  Riemann-integrierbar und

$$F(x) = \int_a^x f(t) dt. \quad (7.3.11)$$

für  $x \in [a, b]$ . Dann ist  $F$  auf  $[a, b]$  Lipschitz-stetig mit Lipschitz-Konstante  $\|f\|_\infty$ . Dies ist gerade die Aussage von Proposition 6.4.1, *i.)*.

*v.)* Wir betrachten  $V = \mathbb{K}^n$  und  $W = \mathbb{K}^m$  mit den jeweiligen Euklidischen Normen  $\|\cdot\|_2$  und den zugehörigen Metriken. Ist dann  $A \in \text{Hom}(V, W)$  eine lineare Abbildung, so gilt für alle  $v, v' \in V$

$$\begin{aligned} \|Av - Av'\|_2 &= \|A(v - v')\|_2 \\ &= \sqrt{\sum_{i=1}^m |A(v - v')_i|^2} \\ &\leq \sum_{i=1}^m |A(v - v')_i| \\ &= \sum_{i=1}^m \left| \sum_{j=1}^n A_{ij}(v_j - v'_j) \right| \\ &\leq \sum_{i=1}^m \sum_{j=1}^n |A_{ij}| |v_j - v'_j| \\ &\leq \sum_{j=1}^n \sum_{i=1}^m |A_{ij}| \|v - v'\|_2, \end{aligned}$$

womit  $A$  Lipschitz-stetig ist. Insbesondere ist

$$\lambda = \sum_{j=1}^n \sum_{i=1}^m |A_{ij}| = \|A\|_1 \quad (7.3.12)$$

eine Lipschitz-Konstante. Hier identifizieren wir  $\text{Hom}(\mathbb{K}^n, \mathbb{K}^m) = M_{n \times m}(\mathbb{K}) = \mathbb{K}^{nm}$ , womit (7.3.12) die 1-Norm auf  $\mathbb{K}^{nm}$  ist. Im Allgemeinen gibt es sogar eine bessere (kleinere) Lipschitz-Konstante. Die kleinste ist die Operatornorm

$$\|A\| = \sup_{v \neq 0} \frac{\|Av\|_2}{\|v\|_2} \leq \|A\|_1 \quad (7.3.13)$$

aus Bemerkung 7.1.4, *iii.*). Damit haben wir bezüglich der Euklidischen Normen insbesondere

$$L(\mathbb{K}^n, \mathbb{K}^m) = \text{Hom}(\mathbb{K}^n, \mathbb{K}^m) \quad (7.3.14)$$

gezeigt. Zudem haben wir gezeigt, dass in endlichen Dimensionen jede lineare Abbildung bezüglich der Euklidischen Norm *stetig* ist, ein Resultat, dass entscheidend von der Endlichdimensionalität Gebrauch macht.

*vi.)* Als vorerst letztes Beispiel betrachten wir einen normierten Vektorraum  $(V, \|\cdot\|)$ . Dann versehen wir  $V \times V$  mit der 1-Norm  $\|\cdot\|_1$  aus Bemerkung 7.1.4, *i.*), also

$$\|(v, w)\|_1 = \|v\| + \|w\|. \quad (7.3.15)$$

Ebenso verwenden wir für  $\mathbb{K} \times V$  die zugehörige 1-Norm  $\|(z, v)\|_1 = |z| + \|v\|$ . Dann gilt

$$\|v + w - (v' + w')\| \leq \|v - v'\| + \|w - w'\| = \|(v, w) - (v', w')\|_1,$$

womit die Addition

$$+ : V \times V \ni (v, w) \mapsto v + w \in V \quad (7.3.16)$$

Lipschitz-stetig mit Lipschitz-Konstante  $\lambda = 1$  ist. Sei weiter  $z_n \in \mathbb{K}$  mit  $z_n \rightarrow z$  und  $v_n \in V$  mit  $v_n \rightarrow v$  gegeben. Dann konvergiert auch  $(z_n, v_n) \rightarrow (z, v)$  bezüglich der obigen 1-Norm in  $\mathbb{K} \times V$ . Weiter gilt

$$\begin{aligned} \|z_n v_n - z v\| &= \|z_n v_n - z v_n + z v_n - z v\| \\ &\leq \|z_n v_n - z v_n\| + \|z v_n - z v\| \\ &= |z_n - z| \|v_n\| + |z| \|v_n - v\|. \end{aligned}$$

Da die Norm stetig ist, folgt  $\|v_n\| \rightarrow \|v\|$ , sowie  $\|v_n - v\| \rightarrow 0$  nach Definition von Konvergenz bezüglich  $\|\cdot\|$  in  $V$ . Da zudem  $|z_n - z| \rightarrow 0$ , folgt insgesamt  $\|z_n v_n - z v\| \rightarrow 0$ , was die (Folgen-) Stetigkeit der Multiplikationsabbildung

$$\cdot : \mathbb{K} \times V \ni (z, v) \mapsto z v \in V \quad (7.3.17)$$

zeigt. Damit sind also die Vektorraumoperationen eines normierten Vektorraums stetig. Man kann zeigen, dass die Multiplikation mit Skalaren (7.3.17) nicht Lipschitz-stetig, aber lokal Lipschitz-stetig ist, siehe auch Übung ???. Die Idee ist dabei analog zu *ii.*).

Zum Abschluss dieses Abschnitts wollen wir nochmals die Stetigkeit linearer Abbildungen aufgreifen und den Zusammenhang zur Operatornorm herausarbeiten.

**Satz 7.3.7 (Stetigkeit linearer Abbildungen)** *Seien  $(V, \|\cdot\|_V)$  und  $(W, \|\cdot\|_W)$  normierte Vektorräume sowie  $A \in \text{Hom}(V, W)$  eine lineare Abbildung. Dann sind folgende Eigenschaften äquivalent:*

- i.) Es gilt  $A \in L(V, W)$ .*
- ii.) Die Abbildung  $A$  ist Lipschitz-stetig.*
- iii.) Die Abbildung  $A$  ist stetig.*
- iv.) Die Abbildung  $A$  ist stetig bei  $0 \in V$ .*

*In diesem Fall gilt für die Operatornorm  $\|A\|$  von  $A$  die alternative Charakterisierung*

$$\|A\| = \sup_{v \in V \setminus \{0\}} \frac{\|Av\|_W}{\|v\|_V}. \quad (7.3.18)$$

BEWEIS: Wir zeigen  $i.) \implies ii.) \implies iii.) \implies iv.) \implies i.)$ . Sei also zunächst  $A \in L(V, W)$  mit endlicher Operatornorm

$$\|A\| = \sup\{\|Av\|_W \mid v \in V \text{ mit } \|v\|_V \leq 1\} \quad (*)$$

wie in (7.1.10). Insbesondere gilt für  $v \in V$  mit  $\|v\|_V = 1$  die Ungleichung  $\|Av\|_W \leq \|A\|$ . Ist  $v \in V \setminus \{0\}$  beliebig, so hat  $\frac{v}{\|v\|_V}$  Norm 1, womit

$$\frac{v}{\|v\|_V} \|Av\|_W = \left\| A \frac{v}{\|v\|_V} \right\|_W \leq \|A\|, \quad \text{was} \quad \|Av\|_W \leq \|A\| \|v\|_V$$

für alle solchen  $v \in V \setminus \{0\}$  zeigt. Also ist  $\|A\| \leq \frac{\|Av\|_W}{\|v\|_V}$ , was „ $\leq$ “ in (7.3.18) zeigt. Umgekehrt gilt

$$\sup_{v \in V \setminus \{0\}} \frac{\|Av\|_W}{\|v\|_V} = \sup_{v \in V \setminus \{0\}} \left\| A \frac{v}{\|v\|_V} \right\|_W \leq \sup_{v \in V \setminus \{0\}} \|A\| = \|A\|,$$

womit die verbleibende Ungleichung und damit (7.3.18) gezeigt ist. Mit dieser Vorüberlegung folgt auch

$$\|Av\|_W \leq \|A\| \|v\|_V$$

für alle  $v \in V$ , da die Ungleichung für  $v = 0$  trivialerweise erfüllt ist. Es folgt

$$d_W(Av, Av') = \|Av - Av'\|_W = \|A(v - v')\|_W \leq \|A\| \|v - v'\|_V = \|A\| d_V(v, v'),$$

womit  $A$  Lipschitz-stetig mit Lipschitz-Konstante  $\|A\|$  ist. Da jede Lipschitz-stetige Abbildung insbesondere stetig ist und jede stetige Abbildung stetig an jedem Punkt ist, sind die Implikationen  $ii.) \implies iii.)$  und  $iii.) \implies iv.)$  trivial. Sei also schließlich  $A$  stetig bei  $0 \in V$ . Zu  $\epsilon = 1$  finden wir daher ein  $\delta > 0$  mit der Eigenschaft, dass

$$d_V(v, 0) < \delta \implies d_W(A(v), A(0)) < 1.$$

Da  $A(0) = 0$  gilt, bedeutet dies, dass für  $\|v\|_V < \delta$  auch  $\|Av\|_W < 1$  gilt. Ist nun  $v \in V \setminus \{0\}$  beliebig, so gilt für  $\tilde{v} = \frac{\delta}{2\|v\|_V} v$ , dass  $\|\tilde{v}\|_V = \frac{\delta}{2\|v\|_V} \|v\|_V < \delta$ , womit

$$\|Av\|_W = \left\| A \left( \frac{2\|v\|_V}{\delta} \tilde{v} \right) \right\|_W = \frac{2\|v\|_V}{\delta} \|A\tilde{v}\|_W \leq \frac{2\|v\|_V}{\delta}$$

folgt. Da diese Ungleichung für  $v = 0$  ebenfalls aus trivialen Gründen richtig ist, ist das Supremum (7.3.18) endlich mit  $\|A\| \leq \frac{2}{\delta}$ . Daher ist  $i.)$  gezeigt.  $\square$

Mit anderen Worten haben wir

$$\mathcal{C}(V, W) \cap \text{Hom}(V, W) = L(V, W) \quad (7.3.19)$$

gezeigt. Die Operatornorm besitzt nun tatsächlich die Eigenschaften einer Norm:

**Proposition 7.3.8** *Seien  $(V_i, \|\cdot\|_i)$  normierte Vektorräume für  $i = 1, 2, 3$ .*

- i.) Die stetigen linearen Abbildungen  $L(V_1, V_2)$  bilden einen Unterraum von  $\text{Hom}(V, W)$ .*
- ii.) Die Operatornorm ist eine Norm auf  $L(V_1, V_2)$ .*
- iii.) Für  $A \in L(V_1, V_2)$  und  $B \in L(V_2, V_3)$  gilt*

$$\|BA\| \leq \|B\| \|A\|. \quad (7.3.20)$$

BEWEIS: Sei  $v \in V_1$  und  $\lambda, \lambda' \in \mathbb{K}$  sowie  $A, A' \in L(V_1, V_2)$ . Dann gilt mit den Normeigenschaften von  $\|\cdot\|_2$

$$\begin{aligned} \|(\lambda A + \lambda' A')(v)\|_2 &= \|\lambda A(v) + \lambda' A'(v)\|_2 \\ &\leq |\lambda| \|A(v)\|_2 + |\lambda'| \|A'(v)\|_2 \\ &\leq |\lambda| \|A\|_2 \|v\|_1 + |\lambda'| \|A'\|_2 \|v\|_1 \\ &= (|\lambda| \|A\| + |\lambda'| \|A'\|) \|v\|_1. \end{aligned}$$

Dies zeigt, dass das nötige Supremum für die Operatornorm endlich ist und

$$\|\lambda A + \lambda' A'\| \leq |\lambda| \|A\| + |\lambda'| \|A'\|$$

erfüllt. Zum einen ist damit  $\lambda A + \lambda' A' \in L(V_1, V_2)$  gezeigt und entsprechend *i.*) bewiesen, d.  $0 \in L(V_1, V_2)$  klar ist. Zum anderen können wir *ii.*) zeigen. Es verbleibt lediglich  $\|A\| > 0$  für  $A \neq 0$  nachzuweisen. Dies ist aber klar, da für  $A \neq 0$  ein  $v \in V_1$  mit  $Av \neq 0$  existiert, womit  $\|Av\|_2 \neq 0$ . Damit muss in der Abschätzung  $\|Av\|_2 \leq \|A\| \|v\|_1$  die Operatornorm aber strikt positiv sein. Es verbleibt *iii.*). Sei also  $v \in V_1$ , dann gilt

$$\|B(A(v))\|_3 \leq \|B\| \|A(v)\|_2 \leq \|B\| \|A\| \|v\|_1,$$

womit (7.3.20) folgt. □

Wir können diese Resultate nun nutzen, um vorherige in einem neuen Licht zu sehen.

**Beispiel 7.3.9 (Stetige  $\delta$ -Funktionale)** Sei  $X$  eine Menge und  $\mathcal{B}(X, \mathbb{K})$  der normierte Raum der beschränkten Funktionen mit der Supremumsnorm. Das Funktional

$$\delta_x: \mathcal{B}(X, \mathbb{K}) \ni f \mapsto f(x) \in \mathbb{K} \quad (7.3.21)$$

heißt  $\delta$ -Funktional bei  $x$  oder *Auswertungsfunktional* bei  $x$ . Es gilt

$$|\delta_x(f)| = |f(x)| \leq \|f\|_\infty, \quad (7.3.22)$$

womit  $\delta_x$  stetig ist mit Funktionalnorm höchstens 1. Da es eine Funktion  $f$  mit  $f(x) = \|f\|_\infty$  gibt, beispielsweise die konstante Funktion 1, ist die Funktionalnorm sogar gleich 1, also

$$\|\delta_x\| = 1. \quad (7.3.23)$$

Die Stetigkeit von  $\delta_x$  bedeutet dann beispielsweise, dass gleichmäßig konvergente Folgen, also Konvergenz bezüglich  $\|\cdot\|_\infty$ , punktweise Konvergenz nach sich zieht.

**Beispiel 7.3.10 (Stetigkeit des Riemann-Integrals)** Sei  $[a, b] \subseteq \mathbb{R}$  ein abgeschlossenes Intervall. Die Riemann-integrierbaren Funktionen  $\mathcal{R}([a, b], \mathbb{K})$  auf  $[a, b]$  sind wie immer mit der Supremumsnorm  $\|\cdot\|_\infty$  versehen. Mit Korollar 6.2.3 sehen wir, dass das Riemann-Integral

$$\int_a^b \cdot dx: \mathcal{R}([a, b], \mathbb{K}) \ni f \mapsto \int_a^b f(x) dx \in \mathbb{K} \quad (7.3.24)$$

ein stetiges lineares Funktional ist, denn

$$\left| \int_a^b f(x) dx \right| \leq (b - a) \|f\|_\infty \quad (7.3.25)$$

zeigt, dass die Funktionalnorm höchstens  $b - a$  ist. Die Abschätzung ist auch optional, was man beispielsweise wieder mit der konstanten Funktion 1 sieht. Es gilt also

$$\left\| \int_a^b \cdot dx \right\| = b - a. \quad (7.3.26)$$

Die Konvergenzaussage (6.7.6) von Satz 6.7.2 ist nichts anderes als die (Folgen-) Stetigkeit. Man beachte jedoch, dass die eigentliche Schwierigkeit von Satz 6.7.2 darin liegt, zu zeigen, dass ein gleichmäßiger Limes von Funktionen aus  $\mathcal{R}([a, b], \mathbb{K})$  selbst wieder Riemann-integrierbar ist.

Neben der Lipschitz-Stetigkeit können wir auch den Begriff der gleichmäßigen Stetigkeit für Abbildungen zwischen metrischen Räumen analog zu Definition 3.6.2 etablieren.

**Definition 7.3.11 (Gleichmäßige Stetigkeit II)** Seien  $(M, d)$  und  $(M', d')$  metrische Räume und  $f: M \rightarrow M'$  eine Abbildung. Dann heißt  $f$  gleichmäßig stetig, falls zu jedem  $\epsilon > 0$  ein  $\delta > 0$  existiert, so dass für  $p, q \in M$  mit  $d(p, q) < \delta$  auch

$$d'(f(p), f(q)) < \epsilon \quad (7.3.27)$$

gilt.

Man beachte, dass dies zum einen die Verallgemeinerung aus Definition 3.6.2 darstellt, wenn wir als Metriken die durch den Absolutbetrag induzierte verwenden. Zum anderen bemerken wir, dass anders als Stetigkeit die gleichmäßige Stetigkeit ein globales Konzept ist, dass alle Punktepaare  $(p, q) \in M \times M$  betrifft.

**Proposition 7.3.12** Sei  $f: (M, d) \rightarrow (M', d')$  eine Abbildung zwischen metrischen Räumen.

i.) Ist  $f$  Lipschitz-stetig, so ist  $f$  gleichmäßig stetig.

ii.) Ist  $f$  gleichmäßig stetig, so ist  $f$  stetig.

BEWEIS: Sei zunächst  $f$  Lipschitz-stetig mit Lipschitz-Konstante  $\lambda > 0$ . Dann gilt  $d'(f(p), f(q)) \leq \lambda d(p, q)$ . Setzt man zu  $\epsilon > 0$  entsprechend  $\delta = \frac{\epsilon}{\lambda} > 0$ , so erhält man direkt die gleichmäßige Stetigkeit von  $f$ . Die zweite Implikation ist, wie schon im Spezialfall in Abschnitt 3.6 klar.  $\square$

Es zeigt sich, dass keine der beiden Implikationen im Allgemeinen umkehrbar werden kann: Stetige aber nicht gleichmäßig stetige Funktionen haben wir bereits in Beispiel 3.6.3, i.) kennengelernt. Eine gleichmäßig stetige aber nicht Lipschitz-stetige Funktion ist beispielsweise die Wurzelfunktion aus Beispiel 7.3.6, iii.), wenn man sie auf  $[-1, 1]$  einschränkt, da sie auf dem kompakten Intervall nach Satz 3.6.5 sicherlich gleichmäßig stetig ist.

## 7.4 Vollständigkeit und der Banachsche Fixpunktsatz

Die Vollständigkeit von  $\mathbb{R}$  hat sich als Schlüssel zur Analysis schlechthin erweisen. Für die komplexen Zahlen hatten wir gesehen, dass Vollständigkeit im Sinne von existierenden Grenzwerten von Cauchy-Folgen ebenfalls gegeben ist. Diese Sichtweise auf Vollständigkeit lässt sich nun auf metrische Räume übertragen:

**Definition 7.4.1 (Vollständigkeit metrischer Räume)** Sei  $(M, d)$  ein metrischer Raum.

i.) Eine Folge  $(p_n)_{n \in \mathbb{N}}$  in  $M$  heißt Cauchy-Folge, falls zu jedem  $\epsilon > 0$  ein  $N \in \mathbb{N}$  existiert, so dass

$$d(p_n, p_m) < \epsilon \quad (7.4.1)$$

für alle  $n, m \geq N$  gilt.

ii.) Konvergiert jede Cauchy-Folge in  $M$ , so heißt  $(M, d)$  vollständiger metrischer Raum.

Viele interessante Beispiele von metrischen Räumen sind normierte Räume. Vollständige normierte Räume spielen in der Analysis eine derart wichtige Rolle, dass sie einen eigenen Namen verdienen:

**Definition 7.4.2 (Banach-Raum)** Ein vollständiger normierter Raum heißt Banach-Raum.

**Bemerkung 7.4.3 (Cauchy-Folgen)**

- i.) Eine konvergente Folge in einem metrischen Raum ist immer auch eine Cauchy-Folge. Dies zeigt man genauso wie für  $\mathbb{K}$  mit der Dreiecksungleichung. In einem vollständigen metrischen Raum sind daher konvergente Folgen und Cauchy-Folgen identisch.
- ii.) In einem normierten Raum  $V$  wird die Cauchy-Bedingung für eine Folge  $(v_n)_{n \in \mathbb{N}}$  zur Bedingung, dass für alle  $\epsilon > 0$  ein  $N \in \mathbb{N}$  existiert, so dass

$$\|v_n - v_m\| < \epsilon \quad (7.4.2)$$

für alle  $n, m > N$  gilt.

Es zeigt sich, dass wir bereits viele Beispiele für vollständige metrische Räume und insbesondere für Banach-Räume kennen:

**Beispiel 7.4.4 (Vollständige metrische Räume)**

- i.) Für alle  $n \in \mathbb{N}$  ist  $\mathbb{K}^n$  bezüglich der Maximumsnorm  $\|\cdot\|_\infty$  aus Beispiel 7.1.2, iii.), ein Banach-Raum. Dies liegt daran, dass Konvergenz bezüglich  $\|\cdot\|_\infty$  gerade komponentenweise Konvergenz ist und ebenso die Cauchy-Bedingung bezüglich  $\|\cdot\|_\infty$  der komponentenweisen Cauchy-Bedingung entspricht. Da aber alle Komponenten Kopien von  $\mathbb{K}$  sind, folgt die Vollständigkeit von  $\mathbb{K}^n$  aus der von  $\mathbb{K}$ .
- ii.) Ist  $X$  eine beliebige nicht-leere Menge, so ist  $\mathcal{B}(X, \mathbb{K})$  bezüglich  $\|\cdot\|_\infty$  ein Banach-Raum. Für eine Teilmenge  $X \subset \mathbb{K}$  hatten wir die Vollständigkeit in Proposition 3.7.8, i.) und ii.), gezeigt. Das dortige Argument verwendet aber nicht, dass  $X \subseteq \mathbb{K}$  gilt, sondern ist für eine beliebige Menge gleichermaßen gültig.
- iii.) Die Riemann-integrierbaren Funktionen  $\mathcal{R}([a, b], \mathbb{K})$  auf einem abgeschlossenen Intervall sind ein Banach-Raum bezüglich der Supremumsnorm. Zunächst wissen wir, dass die Riemann-integrierbaren Funktionen  $\mathcal{R}([a, b], \mathbb{K}) \subseteq \mathcal{B}([a, b], \mathbb{K})$  ein Unterraum des Banach-Raums der beschränkten Funktionen ist. Eine Cauchy-Folge in  $\mathcal{R}([a, b], \mathbb{K})$  hat also nach ii.) zunächst einen Grenzwert in  $\mathcal{B}([a, b], \mathbb{K})$ . Nach Satz 6.7.2 ist dieser Grenzwert aber selbst wieder Riemann-integrierbar, womit  $\mathcal{R}([a, b], \mathbb{K})$  vollständig ist.
- iv.) Sei nun  $K \subseteq \mathbb{K}$  kompakt. Die stetigen Funktionen  $\mathcal{C}(K, \mathbb{K})$  sind dann ebenfalls ein Banach-Raum bezüglich der Supremumsnorm. Wir wissen zunächst, dass auf einer kompakten Teilmenge lokal gleichmäßige Konvergenz mit gleichmäßiger Konvergenz übereinstimmt, siehe Übung ???. Nach der Vollständigkeit von  $\mathcal{B}(K, \mathbb{K})$  und der Inklusion  $\mathcal{C}(K, \mathbb{K}) \subseteq \mathcal{B}(K, \mathbb{K})$  konvergiert eine Cauchy-Folge in  $\mathcal{C}(K, \mathbb{K})$  daher gegen eine beschränkte Funktion. Nach Satz 3.7.10 ist diese wieder stetig, womit  $\mathcal{C}(K, \mathbb{K})$  selbst vollständig und damit ein Banach-Raum ist.

Die Beispiele  $\mathcal{R}([a, b], \mathbb{K})$  und  $\mathcal{C}(K, \mathbb{K})$  legen nahe, allgemein Teilmengen von vollständigen metrischen Räumen auf ihren Vollständigkeit hin zu untersuchen. Hier gibt es ein einfaches Kriterium:

**Proposition 7.4.5** Sei  $(M, d)$  ein vollständiger metrischer Raum. Eine Teilmenge  $A \subseteq M$  ist bezüglich der eingeschränkten Metrik genau dann selbst wieder vollständig, wenn  $A$  abgeschlossen ist.

BEWEIS: Sei zunächst  $A$  abgeschlossen und  $(p_n)_{n \in \mathbb{N}}$  eine Cauchy-Folge von Punkten  $p_n \in A$ . Da  $p_n \in M$  und  $M$  vollständig ist, konvergiert diese Folge gegen einen Punkt  $p \in M$ . Wir müssen zeigen, dass  $p$  zu  $A$  gehört. Wir nehmen das Gegenteil an, womit  $p \in M \setminus A$  folgt. Nun ist  $M \setminus A$  offen, also eine Umgebung aller Punkte in  $M \setminus A$  und damit insbesondere eine Umgebung von  $p$ . Die Konvergenz  $p_n \rightarrow p$  zeigt dann nach Proposition 7.2.6, i.), dass  $p_n \in M \setminus A$  für fast alle  $n \in \mathbb{N}$  gilt, ein Widerspruch. Sei umgekehrt  $A$  vollständig. Wir müssen zeigen, dass  $A$  abgeschlossen also  $M \setminus A$  offen ist. Angenommen  $M \setminus A$  ist nicht offen, dann gibt es einen Punkt  $p \in M \setminus A$  derart, dass jede Umgebung und insbesondere jeder offene Ball  $B_{\frac{1}{n}}(p)$  auch Punkte von  $A$  enthält, es darf eben nicht  $B_{\frac{1}{n}}(p) \subseteq M \setminus A$  gelten. Wir wählen einen Punkt  $p_n \in B_{\frac{1}{n}}(p) \cap A$  für jedes  $n \in \mathbb{N}$ . Dann konvergiert  $p_n \rightarrow p$  nach Konstruktion, ist also insbesondere eine Cauchy-Folge in  $A$ . Da  $A$  vollständig ist, konvergiert  $(p_n)_{n \in \mathbb{N}}$  in  $A$ , womit  $p \in A$  folgt, ein Widerspruch.  $\square$

**Korollar 7.4.6** *Ein Unterraum eines Banach-Raums ist genau dann selbst ein Banach-Raum, wenn er abgeschlossen ist.*

Wir kommen nun zum vorerst wichtigsten Satz zu vollständigen metrischen Räumen, dem *Banachschen Fixpunktsatz*. Der eigentliche Sachverhalt ist nicht weiter schwierig, die Anwendungsmöglichkeiten sind dagegen derart vielseitig, dass der Banachsche Fixpunktsatz zu einem mächtigen Verbündeten auch weit jenseits der Analysis wird.

**Satz 7.4.7 (Banachscher Fixpunktsatz)** *Sei  $(M, d)$  ein vollständiger metrischer Raum und*

$$\Phi: M \longrightarrow M \quad (7.4.3)$$

*eine Kontraktion. Dann hat  $\Phi$  einen eindeutigen Fixpunkt  $p_* \in M$ , also*

$$\Phi(p_*) = p_*. \quad (7.4.4)$$

*Ist  $p_0 \in M$  ein beliebiger Startpunkt, so konvergiert die Folge  $(p_n)_{n \in \mathbb{N}}$  mit*

$$p_n = \Phi^n(p_0) \quad (7.4.5)$$

*gegen  $p_*$ .*

BEWEIS: Wir zeigen zuerst, dass die so definierte Folge  $(p_n)_{n \in \mathbb{N}}$  eine Cauchy-Folge ist. Sei dazu  $\epsilon > 0$  vorgegeben. Für alle  $n, k \in \mathbb{N}$  gilt dann nach Konstruktion

$$\begin{aligned} d(p_n, p_{n+k}) &= d(\Phi^n(p_0), \Phi^{n+k}(p_0)) \\ &\leq \lambda d(\Phi^{n-1}(p_0), \Phi^{n+k-1}(p_0)) \\ &\vdots \\ &\leq \lambda^n d(p_0, \Phi^k(p_0)) \\ &= \lambda^n d(p_0, p_k), \end{aligned}$$

wenn  $\lambda < 1$  die Lipschitz-Konstante von  $\Phi$  ist. Mit der Dreiecksungleichung finden wir weiter

$$\begin{aligned} d(p_0, p_k) &\leq d(p_0, p_1) + d(p_1, p_2) + \cdots + d(p_{k-1}, p_k) \\ &\leq d(p_0, p_1) + \lambda d(p_0, p_1) + \cdots + \lambda^{k-1} d(p_0, p_1) \\ &= (1 + \lambda + \lambda^2 + \cdots + \lambda^{k-1}) d(p_0, p_1) \\ &< \frac{1}{1 - \lambda} d(p_0, p_1), \end{aligned}$$



da nach Voraussetzung  $\lambda < 1$  und somit die geometrische Reihe konvergiert. Da umgekehrt  $\lambda \geq 0$  gilt, konvergiert die Reihe monoton wachsend gegen ihren Grenzwert. Zusammen folgt also für alle  $n$  und  $k$

$$d(p_n, p_{n+k}) < \frac{\lambda^n}{1-\lambda} d(p_0, p_1).$$

Mit  $0 \leq \lambda < 1$  folgt  $\lambda^n \rightarrow 0$ , so dass  $(p_n)_{n \in \mathbb{N}}$  eine Cauchy-Folge ist. Die Vollständigkeit liefert daher einen Grenzwert  $p_* = \lim_{n \rightarrow \infty} p_n$ . Dieser erfüllt

$$\Phi(p_*) = \Phi\left(\lim_{n \rightarrow \infty} p_n\right) = \lim_{n \rightarrow \infty} \Phi(p_n) = \lim_{n \rightarrow \infty} p_{n+1} = p_*,$$

wobei wir verwenden, dass  $\Phi$  als Lipschitz-stetige Abbildung erst recht stetig ist. Also ist  $p_*$  ein Fixpunkt. Ist  $p' \in M$  nun ein weiterer Fixpunkt, so gilt

$$d(p_*, p') = d(\Phi(p_*), \Phi(p')) \leq \lambda d(p_*, p')$$

mit  $\lambda < 1$ . Dies ist nur für  $d(p_*, p') = 0$  und damit  $p_* = p'$  möglich, was die Eindeutigkeit zeigt.  $\square$

Aus dem Beweis entnehmen wir die folgende Abschätzung zur Güte der Konvergenz:

**Korollar 7.4.8** *Sei  $(M, d)$  ein vollständiger metrischer Raum und  $\Phi: M \rightarrow M$  eine Kontraktion mit Lipschitz-Konstante  $\lambda < 1$ . Ist  $p_0 \in M$  beliebig, so gilt für den eindeutigen Fixpunkt  $p_*$  von  $\Phi$  und die Folge  $(p_n)_{n \in \mathbb{N}}$  mit  $p_n = \Phi^n(p_0)$  die Abschätzung*

$$d(p_n, p_*) \leq \frac{\lambda^n}{1-\lambda} d(p_0, p_1). \quad (7.4.6)$$

BEWEIS: Die Metrik von  $M$  ist stetig, womit für alle  $n \in \mathbb{N}$

$$\begin{aligned} d(p_n, p_*) &= d\left(p_n, \lim_{k \rightarrow \infty} p_k\right) \\ &= \lim_{k \rightarrow \infty} d(p_n, p_k) \\ &\leq \limsup_{k \rightarrow \infty} d(p_n, p_k) \\ &= \limsup_{k \rightarrow \infty} d(p_n, p_{n+k}) \\ &\leq \frac{\lambda^n}{1-\lambda} d(p_0, p_1), \end{aligned}$$

wobei wir die letzte Abschätzung aus dem Beweis von Satz 7.4.7 entnehmen und zuvor verwenden, dass der größte Häufungspunkt einer Folge nicht von den ersten  $n$  Folgengliedern abhängt.  $\square$

In den Übungen findet man einige erste Anwendungen des Banachschen Fixpunktsatzes, siehe Übung ??.

## 7.5 Topologische Räume und Stetigkeit

Wir sind nun an dem Punkt angelangt, wo wir endgültig der Frage nachgehen sollten, welche Eigenschaften und Konzepte metrischer Räume tatsächlich die Metrik benötigen und welche eventuell mit gröberer Information auskommen. In Proposition 7.2.8 hatten wir gesehen, wie man von einer gegebenen Metrik zu einer anderen wechseln kann, ohne das System der offenen Kugeln und damit das System der offenen Mengen zu ändern. Die neue Metrik hat dann typischerweise trotzdem sehr verschiedenen *metrische* Eigenschaften. So waren beispielsweise alle Abstände kleiner 1 und damit alle Teilmengen beschränkt. Umgekehrt benutzten Konzepte wie Konvergenz und Stetigkeit nur die offenen Teilmengen, weshalb wir uns auf deren Eigenschaften fokussieren wollen. Dies führt durch eine direkte Abstraktion ausgehend von Proposition 7.2.4 zu folgender Definition einer *Topologie*:

**Definition 7.5.1 (Topologischer Raum)** Sei  $M$  eine Menge.

i.) Eine Teilmenge  $\mathcal{M} \subseteq 2^M$  der Potenzmenge von  $M$  heißt *Topologie auf  $M$* , falls folgende Eigenschaften erfüllt sind:

(a) Es gilt  $\emptyset, M \in \mathcal{M}$ .

(b) Für  $O_i \in \mathcal{M}$  mit  $i \in I$  und einer beliebigen Indexmenge  $I$  gilt

$$\bigcup_{i \in I} O_i \in \mathcal{M}. \quad (7.5.1)$$

(c) Für endlich viele  $O_i, \dots, O_n \in \mathcal{M}$  gilt

$$O_1 \cap \dots \cap O_n \in \mathcal{M}. \quad (7.5.2)$$

ii.) Ein topologischer Raum  $(M, \mathcal{M})$  ist ein Paar bestehend aus einer Menge  $M$  und einer Topologie  $\mathcal{M}$ .

iii.) Ist  $(M, \mathcal{M})$  ein topologischer Raum, so heißen die Teilmengen von  $M$  aus  $\mathcal{M}$  *offene Teilmengen*. Eine Teilmenge  $A \subseteq M$  heißt *abgeschlossen*, falls  $M \setminus A \in \mathcal{M}$ .

Die metrischen Räume liefern eine große Klasse von Beispielen. Es gibt aber auch interessante Beispiele jenseits der metrischen Räume:

**Beispiel 7.5.2 (Topologische Räume)**

- i.) Ist  $(M, d)$  ein metrischer Raum, so bilden die offenen Mengen im Sinne von Definition 7.2.3, iii.), tatsächlich eine Topologie auf  $M$ . Dies ist gerade die Aussage von Proposition 7.2.4, iii.), iv.) und v.). In diesem Zusammenhang sprechen wir von der zugrundeliegenden *metrischen Topologie*. Nach Proposition 7.2.8 wissen wir, dass unterschiedliche Metriken durchaus die gleiche zugrundeliegende Topologie besitzen können.
- ii.) Einen besonders drastischen Fall erhält man, indem man schlicht alle Teilmengen von  $M$  als offen erklärt, also  $\mathcal{M} = 2^M$  setzt. Dies nennt man die *diskrete Topologie*. Die Eigenschaften einer Topologie sind trivialerweise erfüllt. Es zeigt sich, dass dieses Beispiel metrisch ist: Die diskrete Metrik aus Beispiel 7.1.8, i.), besitzt alle Teilmengen als offene Mengen, siehe auch Übung 7.8.5.
- iii.) Gewissermaßen am anderen Ende der Skala liegt die *indiskrete Topologie*

$$\mathcal{M}_{\text{indiskret}} = \{\emptyset, M\}, \quad (7.5.3)$$

die also nur aus der leeren und der ganzen Menge  $M$  besteht. Auch hier ist der Nachweis der definierenden Eigenschaften trivial. Diese Topologie ist für Mengen mit mindestens zwei Punkten nicht metrisch.

- iv.) Ein erstes nicht-triviales Beispiel erhält man mit der *kofiniten Topologie*  $\mathcal{M}_{\text{cofinite}} \subseteq 2^M$ , die aus allen denjenigen Teilmengen von  $M$  besteht, deren Komplemente endliche Teilmengen sind. Ist  $M$  unendlich, so ist dies eine von ii.) verschiedene Topologie, siehe auch Übung 7.8.8.
- v.) Ein vorerst letztes Beispiel ist die *Zariski-Topologie* für  $\mathbb{C}^n$ . Man betrachtet eine Teilmenge  $A \subseteq \mathbb{C}^n$  als abgeschlossen, wenn es Polynome  $f_1, \dots, f_k \in \mathbb{C}[x_1, \dots, x_n]$  in  $n$  Variablen gibt, so dass

$$A = \{x \in \mathbb{C}^n \mid f_i(x) = 0 \text{ für } i = 1, \dots, k\} \quad (7.5.4)$$

gilt. Entsprechend heißt  $O \subseteq \mathbb{C}^n$  *offen*, wenn das Komplement  $\mathbb{C}^n \setminus O$  abgeschlossen im Sinne von (7.5.4) ist. Es zeigt sich, dass dies eine Topologie ist, siehe Übung 7.8.9. Sie spielt die entscheidende Rolle in der algebraischen Geometrie.

Für die Analysis sind vor allem Topologien auf Funktionenräumen wichtig, die von Normen und damit von metrischen Räumen kommen. Es gibt aber wichtige Ausnahmen, wo Vektorräume mit nicht-metrischen Topologien in der Analysis eine entscheidende Rolle spielen. Diese werden wir aber zunächst nicht benötigen.

**Korollar 7.5.3** *Sei  $(M, \mathcal{M})$  ein topologischer Raum. Dann gilt:*

i.) *Die Mengen  $M$  und  $\emptyset$  sind abgeschlossen.*

ii.) *Sind  $A_i \subseteq M$  abgeschlossene Teilmengen für  $i \in I$  mit einer beliebigen Indexmenge  $I$ , so ist der Durchschnitt*

$$\bigcap_{i \in I} A_i \subseteq M \quad (7.5.5)$$

*ebenfalls abgeschlossen.*

iii.) *Sind  $A_1, \dots, A_n \subseteq M$  endlich viele abgeschlossene Teilmengen, so ist auch*

$$A_1 \cap \dots \cap A_n \subseteq M \quad (7.5.6)$$

*wieder abgeschlossen.*

Mit der üblichen Mengenarithmetik ist dies nun unmittelbar klar. Abgeschlossene und offene Mengen verhalten sich also komplementär zueinander. Wir hätten alternativ einen topologischen Raum auch durch die Angabe von einem System abgeschlossener Mengen mit den Eigenschaften aus Korollar 7.5.3 charakterisieren können und offene Mengen als Komplemente von abgeschlossenen Mengen definieren können.

Wie zuvor können wir offenen Mengen dazu benutzen, Umgebungen zu definieren. Dies verallgemeinert die Situation aus metrischen Räumen:

**Definition 7.5.4 (Umgebung)** *Sei  $(M, \mathcal{M})$  ein topologischer Raum. Eine Teilmenge  $U \subseteq M$  heißt Umgebung von  $p \in M$ , wenn es eine offene Teilmenge  $O \subseteq M$  mit*

$$p \in O \subseteq U \quad (7.5.7)$$

*gibt.*

**Proposition 7.5.5** *Sei  $(M, \mathcal{M})$  ein topologischer Raum mit  $p \in M$ .*

i.) *Ist  $U \subseteq M$  eine Umgebung von  $p$  und  $U \subseteq U'$ , so ist  $U'$  auch eine Umgebung von  $p$ .*

ii.) *Sind  $U_1, \dots, U_n \subseteq M$  endlich viele Umgebungen von  $p$ , so ist auch  $U_1 \cap \dots \cap U_n \subseteq M$  eine Umgebung von  $p$ .*

iii.) *Zu jeder Umgebung  $U$  von  $p$  existiert eine Umgebung  $V \subseteq U$  von  $p$ , die eine Umgebung aller ihrer Punkte  $q \in V$  ist.*

iv.) *Eine Teilmenge  $O \subseteq M$  ist genau dann offen, wenn sie Umgebung aller ihrer Punkte ist.*

BEWEIS: Die Argumente sind uns schon in Proposition 7.2.4 beziehungsweise in Proposition 2.5.2 begegnet. Der Vollständigkeit wegen zeigen wir die Proposition hier trotzdem im Detail. Der erste Teil ist nach Definition klar. Sind  $U_1, \dots, U_n$  endlich viele Umgebungen von  $p$ , so gibt es offenen Teilmengen  $O_1, \dots, O_n$  mit

$$p \in O_i \subseteq U_i$$

für  $i = 1, \dots, n$ . Dann ist zum einen  $O = O_1 \cap \dots \cap O_n$  wieder offen, zum anderen gilt

$$p \in O = O_1 \cap \dots \cap O_n \subseteq U_1 \cap \dots \cap U_n,$$

womit  $U_1 \cap \dots \cap U_n$  eine Umgebung von  $p$  ist. Ist  $O \subseteq M$  offen und  $p \in O$ , so gilt  $p \in O \subseteq O$ , womit  $O$  eine Umgebung aller Punkte aus  $O$  ist. Ist umgekehrt  $U \subseteq M$  eine Teilmenge, welche eine Umgebung aller  $p \in U$  ist, so gibt es zu jedem  $p \in U$  eine offene Teilmenge  $O_p \subseteq M$  mit  $p \in O_p \subseteq U$ . Es folgt

$$U = \bigcup_{p \in U} O_p,$$

womit  $U$  offen ist. Dies zeigt *iv.*), womit *iii.*) ebenfalls klar ist, da zu einer Umgebung  $U$  von  $p$  die definitionsgemäß existierende offene Menge  $V = O \subseteq M$  mit  $p \in O \subseteq U$  das Gewünschte leistet.  $\square$

Mit Hilfe von Umgebungen können wir auch Berührungspunkte sowie den Abschluss einer Teilmenge definieren, wie wir dies bereits für Teilmengen von  $\mathbb{K}$  in Definition 3.1.2 und Definition 3.1.4 getan haben:

**Definition 7.5.6 (Berührungspunkte und Abschluss)** Sei  $(M, \mathcal{M})$  ein topologischer Raum mit einer Teilmenge  $X \subseteq M$ .

i.) Ein Punkt  $p \in M$  heißt Berührungspunkt von  $X$ , wenn für jede Umgebung  $U \subseteq M$  von  $p$

$$U \cap X \neq \emptyset$$

gilt.

ii.) Der Abschluss  $X^{\text{cl}}$  von  $X$  ist die kleinste abgeschlossene Teilmenge von  $M$ , die  $X$  enthält.

Mit anderen Worten ist  $X^{\text{cl}}$  durch

$$X^{\text{cl}} = \bigcap_{\substack{X \subseteq A \\ A \text{ abgeschlossen}}} A \quad (7.5.8)$$

gegeben. Dieser Durchschnitt ist nämlich abgeschlossen, enthält  $X$  und ist nach Konstruktion auch die kleinste solche abgeschlossene Menge. Man beachte, dass mit  $M$  als abgeschlossener Menge auf jeden Fall eine abgeschlossene Menge existiert, die  $X$  enthält. Wie schon in Proposition 3.1.3, *iv.*), besteht der Abschluss  $X^{\text{cl}}$  von  $X$  gerade aus den Berührungspunkten:

**Proposition 7.5.7** Sei  $(M, \mathcal{M})$  ein topologischer Raum mit einer Teilmenge  $X \subseteq M$ . Dann ist der Abschluss  $X^{\text{cl}}$  die Menge der Berührungspunkte von  $X$ .

BEWEIS: Sei zunächst  $p \in M \setminus X^{\text{cl}}$ . Da  $M \setminus X^{\text{cl}}$  offen ist, ist diese Menge eine Umgebung von  $p$  mit

$$(M \setminus X^{\text{cl}}) \cap X = \emptyset,$$

womit  $p$  kein Berührungspunkt sein kann. Daher sind die Berührungspunkte in  $X^{\text{cl}}$  enthalten. Ist  $p$  kein Berührungspunkt, so gibt es eine Umgebung  $U \subseteq M$  von  $p$  mit  $U \cap X = \emptyset$ . Insbesondere gibt es eine offene Teilmenge  $O \subseteq M$  mit  $p \in O \subseteq M \setminus X$ . Damit sind aber auch alle anderen Punkte  $q \in O$  keine Berührungspunkte, da  $O$  ebenfalls eine Umgebung von  $q$  ist. Es folgt, dass die Menge derjenigen Punkte, die keine Berührungspunkte sind, selbst offen ist. Daher ist die Menge der Berührungspunkte aber abgeschlossen und enthält mit  $X$  auch  $X^{\text{cl}}$ . Zusammen folgt die Behauptung.  $\square$

Wie immer in der Mathematik will man aus den vorhandenen Objekten eines gewissen Typs neue konstruieren. Bei topologischen Räumen gibt es eine Vielzahl solcher Konstruktionen, von denen zwei von besonderer Bedeutung für uns sind:

**Proposition 7.5.8** Ist  $(M, \mathcal{M})$  ein topologischer Raum mit einer Teilmenge  $N \subseteq M$ , so definiert

$$\mathcal{N} = \mathcal{M}|_N = \{O \cap N \mid O \in \mathcal{M}\} \quad (7.5.9)$$

eine Topologie auf  $N$ .

Abbildung 7.6: Durchschnitte von kartesischen Produkten sind die Produkte der Durchschnitte.

BEWEIS: Die Definition ist gänzlich analog zur Definition 3.2.4 von offenen Mengen in einer Teilmenge  $X \subseteq \mathbb{K}$ . Entsprechend können wir den Beweis von Proposition 3.2.5 wörtlich übernehmen, wo mit unseren gegenwärtigen Begriffen gezeigt wurde, dass  $\mathcal{N} = \mathcal{M}|_N$  eine Topologie ist. Wir wiederholen die entscheidenden Schritte: Zunächst ist  $\emptyset \in \mathcal{N}$ , was man mit  $O = \emptyset$  erreicht. Weiter ist  $N \in \mathcal{N}$ , da wir  $O = M$  verwenden können. Ist nun  $U_i \in \mathcal{N}$  für  $i \in I$ , so gibt es  $O_i \in \mathcal{M}$  mit  $U_i = N \cap O_i$  für alle  $i \in I$ . Dann gilt aber

$$\bigcup_{i \in I} U_i = \bigcup_{i \in I} (N \cap O_i) = N \cap \bigcup_{i \in I} O_i \in \mathcal{N},$$

da  $\bigcup_{i \in I} O_i$  offen ist, Genauso zeigt man für  $U_1 = N \cap O_1, \dots, U_n = N \cap O_n$  mit offenen  $O_1, \dots, O_n$ , dass

$$U_1 \cap \dots \cap U_n = (N \cap O_1) \cap \dots \cap (N \cap O_n) = N \cap (O_1 \cap \dots \cap O_n)$$

wieder in  $\mathcal{N}$  ist, da  $O_1 \cap \dots \cap O_n \in \mathcal{M}$ . □

**Definition 7.5.9 (Unterraumtopologie)** Sei  $(M, \mathcal{M})$  ein topologischer Raum mit einer Teilmenge  $N \subseteq M$ . Die Topologie  $\mathcal{M}|_N$  aus (7.5.9) heißt dann die Unterraumtopologie oder Relativtopologie von  $N$ .

Damit erweist sich also unsere vorherige Konstruktion von offenen Mengen in einer Teilmenge  $X \subseteq \mathbb{K}$  aus Definition 3.2.4 als ein Spezialfall der Unterraumtopologie. Im Folgenden wird stillschweigend für Teilmengen eines topologischen Raumes immer die Unterraumtopologie verwendet.

Die zweite Konstruktion ist die eines kartesischen Produkts von endlich vielen topologischen Räumen. Es gibt darüber hinaus auch eine Definition für beliebige kartesische Produkte, die allerdings an dieser Stelle zu weit führen würde.

**Proposition 7.5.10** Seien  $(M_i, \mathcal{M}_i), \dots, (M_n, \mathcal{M}_n)$  topologische Räume und  $M = M_1 \times \dots \times M_n$  ihr kartesisches Produkt. Dann ist

$$\mathcal{M} = \{O \subseteq M \mid O \text{ ist Vereinigung von Mengen der Form } O_1 \times \dots \times O_n \text{ mit } O_i \in \mathcal{M}_i \text{ für } i = 1, \dots, n\} \quad (7.5.10)$$

eine Topologie auf  $M$ .

BEWEIS: Zunächst sind  $\emptyset$  und  $M$  in  $\mathcal{M}$  als Produkt von  $\emptyset = \emptyset \times \dots \times \emptyset$  sowie  $M = M_1 \times \dots \times M_n$ . Seien weiter  $O^{(1)}, \dots, O^{(k)} \in \mathcal{M}$  endlich viele Teilmengen aus  $\mathcal{M}$ . Diese lassen sich daher als Vereinigung von Produkten

$$O^{(j)} = \bigcup_{i \in I_j} O_{1,i}^{(j)} \times \dots \times O_{n,i}^{(j)}$$

mit  $j = 1, \dots, k$  und  $O_{1,i}^{(j)} \in \mathcal{M}_1, \dots, O_{n,i}^{(j)} \in \mathcal{M}_n$  für alle  $i \in I_j$  schreiben. Für deren Durchschnitt gilt

$$\begin{aligned} & O^{(1)} \cap \dots \cap O^{(k)} \\ &= \bigcup_{i_1 \in I_1} \left( O_{1,i_1}^{(1)} \times \dots \times O_{n,i_1}^{(1)} \right) \cap \dots \cap \bigcup_{i_k \in I_k} \left( O_{1,i_k}^{(k)} \times \dots \times O_{n,i_k}^{(k)} \right) \\ &= \bigcup_{i_1 \in I_1} \dots \bigcup_{i_k \in I_k} \left( O_{1,i_1}^{(1)} \cap \dots \cap O_{1,i_k}^{(k)} \right) \times \dots \times \left( O_{n,i_1}^{(1)} \cap \dots \cap O_{n,i_k}^{(k)} \right), \end{aligned}$$

siehe auch Abbildung 7.6. Damit sieht man, dass die endlichen Durchschnitte solcher Mengen wieder von der gewünschten Form sind, da die endlichen Durchschnitte  $O_{1,i_1}^{(1)} \cap \dots \cap O_{1,i_k}^{(k)} \in \mathcal{M}$  etc. wieder offen sind. Bildet man schließlich Vereinigungen von Teilmengen aus  $\mathcal{M}$ , so sind diese nach Konstruktion automatisch wieder in  $\mathcal{M}$ . □

Im Folgenden werden wir kartesische Produkte immer mit dieser *Produkttopologie* versehen. Glücklicherweise ist diese Konstruktion im Falle metrischer Räume wieder metrisch. Wir formulieren das Resultat für zwei metrische Räume, der allgemeine Fall endlich vieler metrischer Räume erhält man genauso, siehe auch Übung ??.

**Proposition 7.5.11** *Seien  $(M_1, d_1)$  und  $(M_2, d_2)$  metrische Räume mit ihrer metrischen Topologie versehen. Dann ist die Produkttopologie von  $M = M_1 \times M_2$  eine metrische Topologie zur Metrik*

$$d((x_1, x_2), (y_1, y_2)) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}, \quad (7.5.11)$$

wobei  $x = (x_1, x_2), y = (y_1, y_2)$  mit  $x_1, y_1 \in M_1$  und  $x_2, y_2 \in M_2$ .

BEWEIS: Seien  $O_1 \subseteq M_1$  und  $O_2 \subseteq M_2$  offen und  $p_1 \in O_1$  sowie  $p_2 \in O_2$ . Um zu zeigen, dass  $O_1 \times O_2$  bezüglich der Metrik  $d$  offen ist, wählen wir zunächst  $r > 0$  derart, dass

$$B_r(p_1) \subseteq O_1 \quad \text{und} \quad B_r(p_2) \subseteq O_2$$

gilt, was nach der Offenheit von  $O_1$  und  $O_2$  ja möglich ist. Man beachte, dass wir ohne Einschränkung den gleichen Radius für beide Kugeln verwenden können, da wir anderenfalls zum kleineren der beiden Radien übergehen. Nun gilt

$$\begin{aligned} B_r(p_1) \times B_r(p_2) &= \{(q_1, q_2) \in M \mid d_1(q_1, p_1) < r \text{ und } d_2(q_2, p_2) < r\} \\ &= \{(q_1, q_2) \in M \mid \max\{d_1(q_1, p_1), d_2(q_2, p_2)\} < r\} \\ &= B_r((p_1, p_2)) \end{aligned}$$

nach der Wahl der Metrik  $d$ . Also ist  $B_r(p_1) \times B_r(p_2) = B_r((p_1, p_2))$  eine offene Kugel bezüglich der Produktmetrik  $d$ , womit

$$B_r((p_1, p_2)) \subseteq O_1 \times O_2$$

folgt. Da dies für jeden Punkt  $p = (p_1, p_2) \in O_1 \times O_2$  möglich ist, folgt, dass  $O_1 \times O_2$  bezüglich der metrischen Topologie von  $(M, d)$  offen ist. Hat man nun eine beliebige offene Teilmenge  $O \subseteq M$  bezüglich der Produkttopologie, so ist diese eine Vereinigung von Teilmengen der Form  $O_1 \times O_2$ , die alle in der metrischen Topologie offen sind. Daher ist  $O$  bezüglich der metrischen Topologie offen. Sei nun umgekehrt  $O \subseteq M$  offen bezüglich der metrischen Topologie. Dann gibt es zu jedem Punkt  $p = (p_1, p_2) \in O$  eine offene Kugel  $B_{r_p}(p) \subseteq O$  bezüglich  $d$  und damit wieder

$$O = \bigcup_{p \in O} B_{r_p}(p).$$

Nach dem gerade Gezeigten gilt aber

$$B_{r_p}(p) = B_{r_p}(p_1) \times B_{r_p}(p_2),$$

womit  $O$  eine Vereinigung offener Mengen der Produkttopologie ist und daher offen bezüglich der Produkttopologie ist. Damit stimmen also beide Topologien überein.  $\square$

Nachdem wir nun über interessante Beispiele von topologischen Räumen verfügen, wollen wir Stetigkeit für Abbildungen zwischen ihnen definieren. Dies geschieht nun am einfachsten folgendermaßen:

**Definition 7.5.12 (Stetigkeit)** *Sei  $f: (M, \mathcal{M}) \rightarrow (N, \mathcal{N})$  eine Abbildung zwischen topologischen Räumen.*

i.) *Die Abbildung  $f$  heißt stetig bei  $p \in M$ , falls für jede Umgebung  $U \subseteq N$  von  $f(p)$  auch das Urbild  $f^{-1}(U) \subseteq M$  eine Umgebung von  $p$  ist.*

ii.) Die Abbildung  $f$  heißt stetig, falls das Urbild jeder offenen Menge in  $N$  offen in  $M$  ist.

**Proposition 7.5.13** Seien  $f: (M_1, \mathcal{M}_1) \rightarrow (M_2, \mathcal{M}_2)$  und  $g: (M_2, \mathcal{M}_2) \rightarrow (M_3, \mathcal{M}_3)$  Abbildungen zwischen topologischen Räumen.

i.) Die Abbildung  $f$  ist genau dann stetig, wenn sie stetig an jedem Punkt ist.

ii.) Ist  $f$  stetig bei  $p \in M_1$  und  $g$  stetig bei  $f(p) \in M_2$ , so ist  $g \circ f$  stetig bei  $p$ .

iii.) Sind  $f$  und  $g$  stetig, so auch  $g \circ f$ .

BEWEIS: Sei zunächst  $f$  stetig und  $p \in M_1$ . Ist  $U \subseteq M_2$  eine Umgebung von  $f(p) \in M_2$ , so enthält diese eine offene Umgebung  $O \subseteq U$  mit  $f(p) \in O$ . Dann ist  $f^{-1}(O)$  offen und es gilt  $p \in f^{-1}(O) \subseteq f^{-1}(U)$ . Also ist  $f^{-1}(U)$  eine Umgebung von  $p$ , womit  $f$  stetig bei  $p$  ist. Sei umgekehrt  $f$  stetig in jedem Punkt  $p \in M_1$ . Ist  $O \subseteq M_2$  offen, so ist für  $p \in f^{-1}(O)$  die Menge  $O$  eine Umgebung von  $p$ . Da dies für alle  $p$  gilt, ist  $f^{-1}(O) \subseteq M_1$  nach Proposition 7.5.5, iv.), eine offene Menge. Damit ist i.) gezeigt. Da generell für  $A \subseteq M_3$

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

gilt, folgen ii.) und iii.) unmittelbar aus den Definitionen. □

Wir sehen, dass sich auch hier die Resultate zur Verkettung von stetigen Funktionen aus Abschnitt 3.5 sich inklusive der Beweise auf beliebige topologische Räume verallgemeinern lassen. Wir werden noch weitere derartige Schlussfolgerungen sehen.

Im Falle metrischer Räume hatten wir zudem Konvergenz von Folgen und eine Charakterisierung von Stetigkeit mittels konvergenter Folgen gefunden, siehe nochmals Abschnitt 7.2 sowie Satz 7.3.1. Im Falle topologischer Räume kann man *Konvergenz von Folgen* oder auch allgemeinen Netzen ebenfalls definieren, indem man fordert, dass für jede Umgebung des Grenzwertes alle bis auf endlich viele Folgenglieder in dieser Umgebung liegen. Bei Netzen verlangt man entsprechend, dass alle Netzglie der mit Indizes später als einem festen, von der Umgebung abhängigen Index, in der Umgebung liegen. Dann zeigen sich, anders als bei metrischen Räumen, jedoch auch neue Phänomene. In bestimmten topologischen Räumen gibt es sehr wenige Umgebungen, was zu recht trivialen konvergenten Folgen oder Netzen führt. Insbesondere können Folgen oder Netze gegen mehrere Grenzwerte gleichzeitig konvergieren, siehe etwa Übung ?? . Um dies zu verhindern, benötigt man zusätzliche Anforderungen an die Topologie, die *Trennungseigenschaften*:

**Definition 7.5.14 (Hausdorff-Topologie)** Ein topologischer Raum  $(M, \mathcal{M})$  heißt Hausdorffsch oder  $T_2$ -Raum, falls es zu  $p, q \in M$  mit  $p \neq q$  offene Mengen  $O_1, O_2 \subseteq M$  mit  $p \in O_1, q \in O_2$  und

$$O_1 \cap O_2 = \emptyset \tag{7.5.12}$$

gibt.

**Korollar 7.5.15** Ein metrischer Raum ist Hausdorffsch.

BEWEIS: Dies ist gerade Proposition 7.2.4, viii.). □

Das andere wichtige und neue Phänomen ist, dass sich Stetigkeit von Abbildungen im Allgemeinen nicht länger durch Folgen charakterisieren lässt. Man muss allgemeinere Netze hierzu verwenden. Bei metrischen Räumen ist dies nicht notwendig. Daher werden wir diese Aspekte hier nicht weiter verfolgen sondern verweisen auf die weiterführende Literatur wie etwa [2, 6] für umfassende Darstellungen sowie auf [7] für eine kurze Einführung zur mengentheoretischen Topologie.

## 7.6 Kompaktheit und Zusammenhang in topologischen Räumen

Wie schon in Abschnitt 3.4 wollen wir nun auch für allgemeine topologische Räume die Begriffe von kompakten und von zusammenhängenden Teilmengen etablieren.

In Satz 3.4.11 hatten wir drei äquivalente aber konzeptuell sehr verschiedene Versionen von Kompaktheit für Teilmengen von  $\mathbb{K}$  gesehen. Es zeigt sich, dass die folgenbasierte Definition in metrischen Räumen einen tragfähigen Begriff liefert, aber für allgemeine topologische Räume modifiziert werden muss: Folgen müssen durch Netze ersetzt werden und Teilfolgen durch Teilnetze, ein nicht ganz unproblematischer Begriff, der unintuitive Eigenschaften an den Tag legt. Die Variante, dass abgeschlossene und beschränkte Teilmengen als die kompakten angesehen werden sollen, ist für allgemeine topologische Räume nicht möglich, da es keinen Begriff für Beschränktheit gibt. Im Falle metrischer Räume wäre es möglich, von beschränkten Teilmengen zu sprechen. Da wir aber generell nach Proposition 7.2.8 die Metrik so abändern können, dass alle Abstände kleiner 1 und damit alle Teilmengen beschränkt werden, *ohne* die zugrundeliegende Topologie zu ändern, ist diese Formulierung selbst für metrische Räume wenig aussichtsreich. Es gibt eine Abwandlung, indem man „beschränkt“ durch den sehr viel technischeren Begriff „total beschränkt“ abwandelt. Dies erscheint aber für eine Definition wenig konzeptuell. Es verbleibt die Option der *Überdeckungskompaktheit*, die sofort für beliebige topologische Räume nutzbar ist:

**Definition 7.6.1 (Kompakte Teilmenge)** Sei  $(M, \mathcal{M})$  ein topologischer Raum. Eine Teilmenge  $K \subseteq M$  heißt *kompakt*, wenn jede offene Überdeckung von  $K$  eine endliche Teilüberdeckung besitzt.

Wie schon in Proposition 3.4.6 erhalten wir folgende äquivalente Umformulierung. Der Beweis ist identisch und beruht lediglich auf Komplementbildung:

**Proposition 7.6.2** Ein topologischer Raum  $(M, \mathcal{M})$  ist genau dann kompakt, wenn es für jedes System  $\{A_i\}_{i \in I}$  abgeschlossener Mengen  $A_i \subseteq M$  mit  $\bigcap_{i \in I} A_i = \emptyset$  bereits endlich viele  $A_{i_1}, \dots, A_{i_n}$  mit  $A_{i_1} \cap \dots \cap A_{i_n} = \emptyset$  gibt.

Interessanter wird die Übertragung des nächsten Resultats aus Proposition 3.4.7. Hier benötigen wir teilweise die zusätzliche Hausdorff-Eigenschaft.

**Proposition 7.6.3** Sei  $(M, \mathcal{M})$  ein topologischer Raum.

- i.) Ist  $K \subseteq M$  kompakt und  $A \subseteq K$  abgeschlossen in  $K$ , so ist  $A$  ebenfalls kompakt.
- ii.) Sei  $M$  zudem Hausdorffsch. Ist  $K \subseteq M$  kompakt, und  $p \in M \setminus K$ , so existieren offene Teilmengen  $O_1, O_2 \subseteq M$  mit  $O_1 \cap O_2 = \emptyset$  und

$$p \in O_1 \quad \text{sowie} \quad K \subseteq O_2. \quad (7.6.1)$$

- iii.) Ist  $M$  zudem Hausdorffsch, so ist jede kompakte Teilmenge von  $M$  abgeschlossen.

BEWEIS: Ist  $A \subseteq M$  abgeschlossen und  $B \subseteq A$ , so ist  $B$  genau dann abgeschlossen in  $A$ , also abgeschlossen bezüglich der Unterraumtopologie  $\mathcal{M}|_A$  von  $A$ , wenn  $B$  als Teilmenge von  $M$  abgeschlossen ist. Dies zeigt man allgemein mit den gleichen Argumenten wie in Proposition 3.4.7. Dann können wir Proposition 7.6.2 verwenden, um wie in Proposition 3.4.7, *iii.*), den ersten Teil zu zeigen. Der zweite und dritte Teil kann nicht ganz wörtlich aus Proposition 3.4.7 übernommen werden, wenn auch die grundlegende Idee die gleiche ist: Wir müssen die Hausdorff-Eigenschaft nun explizit zum Einsatz bringen. Sei also  $K \subseteq M$  kompakt und  $p \in M \setminus K$ . Zu jedem  $q \in K$  finden wir daher offene Teilmengen  $O_q, U_q \subseteq M$ , derart, dass

$$q \in O_q \quad \text{und} \quad p \in U_q \quad \text{sowie} \quad O_q \cap U_q = \emptyset,$$



Abbildung 7.7: Zur Trennung von  $p$  und  $K$  analog zu Abbildung 7.3.

siehe auch Abbildung 7.7. Dann gilt  $K \subseteq \bigcup_{q \in K} O_q$ , womit wir endlich viele  $q_1, \dots, q_n$  finden, so dass

$$K \subseteq O_{q_1} \cup \dots \cup O_{q_n} = O_2.$$

Wir setzen

$$O_1 = U_{q_1} \cap \dots \cap U_{q_n},$$

was eine offene Teilmenge  $O_1$  liefert, so dass zusammen mit  $O_2$  nun alle Anforderungen in *ii.*) erfüllt sind. Nach *ii.*) können wir zu jedem  $p \in M \setminus K$  eine offene Teilmenge  $O_p \subseteq M$  mit  $K \cap O_p = \emptyset$  wählen. Dann folgt  $O_p \subseteq M \setminus K$  und daher

$$M \setminus K = \bigcup_{p \in M \setminus K} O_p,$$

was zeigt, dass  $M \setminus K$  offen ist. □

Als nächstes wollen wir das Verhalten von kompakten Teilmengen unter stetigen Abbildungen verstehen, wie wir dies auch schon in Abschnitt 3.5 thematisiert haben. Mit dem identischen Beweis wie zu Satz 3.5.4 erhalten wir auch die allgemeine Situation: Das stetige Bild eines Kompaktums ist kompakt.

**Proposition 7.6.4** *Sei  $f: (M, \mathcal{M}) \rightarrow (N, \mathcal{N})$  eine stetige Abbildung zwischen topologischen Räumen. Ist  $K \subseteq M$  kompakt, so ist  $f(K) \subseteq N$  ebenfalls kompakt.*

Da die kompakten Teilmengen von  $\mathbb{R}$  und  $\mathbb{C}$  gerade die abgeschlossenen und beschränkten Teilmengen sind, können wir auch Korollar 3.5.6 sowie Korollar 3.5.8 unmittelbar übernehmen:

**Korollar 7.6.5** *Sei  $(M, \mathcal{M})$  ein kompakter topologischer Raum.*

*i.) Ist  $f \in \mathcal{C}(M, \mathbb{K})$ , so ist  $f$  beschränkt.*

*ii.) Ist  $f \in \mathcal{C}(M, \mathbb{R})$ , so nimmt  $f$  Minimum und Maximum an: Es gibt  $p_{\max}, p_{\min} \in M$  mit*

$$f(p_{\max}) = \sup_{p \in M} f(p) \tag{7.6.2}$$

und

$$f(p_{\min}) = \inf_{p \in M} f(p). \tag{7.6.3}$$

Wir kommen nun zur allgemeinen Formulierung des Satzes von Tikhonov für endliche kartesische Produkte von kompakten topologischen Räumen. Wir formulieren den Satz für das Produkt zweier kompakter Räume. Den allgemeinen Fall endlicher Produkte erhält man durch einen analogen aber notationsbedingt aufwändigeren Beweis oder durch eine Induktion wie in Übung ??.

**Satz 7.6.6 (Tikhonov II)** *Seien  $(M_1, \mathcal{M}_1)$  und  $(M_2, \mathcal{M}_2)$  kompakte topologische Räume. Dann ist  $M_1 \times M_2$  bezüglich der Produkttopologie ebenfalls kompakt.*

BEWEIS: Sei  $\{O_i\}_{i \in I}$  eine offene Umgebung von  $M_1 \times M_2$ . Nach Definition der Produkttopologie ist jedes  $O_i$  eine (beliebig große) Vereinigung von Produkten der Form  $U \times V$  mit offenen  $U \subseteq M_1$  und  $V \subseteq M_2$ . Auf diese Weise erhalten wir eine offene Überdeckung von  $M_1 \times M_2$  durch Produkte von offenen Mengen. Können wir für diese zeigen, dass bereits endlich viele der Produkte ausreichen, um  $M_1 \times M_2$  zu überdecken, dann reichen endlich viele der ja größeren ursprünglichen offenen Mengen

Abbildung 7.8: Das kompakte Bild  $i_p(M_2) = \{p\} \times M_2$  wird durch endlich viele Rechtecke überdeckt.

erst recht. Mit dieser Beobachtung können wir also direkt annehmen, dass jedes  $O_i = U_i \times V_i$  ein Produkt ist. Wir betrachten nun einen Punkt  $p \in M_1$  sowie die Abbildung

$$i_p: M_2 \ni q \mapsto (p, q) \in M_1 \times M_2,$$

siehe auch Abbildung 7.8. Da die Abbildung  $i_p$  stetig ist, ist das Bild  $i_p(M_2) = \{p\} \times M_2 \subseteq M_1 \times M_2$  des kompakten Raums  $M_2$  wieder kompakt. Es genügen also endlich viele  $O_{i_1,p}, \dots, O_{i_n,p}$ , um  $\{p\} \times M_2$  zu überdecken. Diese sind von der Form

$$O_{i_j,p} = U_{j,p} \times V_{j,p}$$

mit offenen Teilmengen  $U_{1,p}, \dots, U_{n_p,p} \subseteq M_1$  und  $V_{1,p}, \dots, V_{n_p,p} \subseteq M_2$ . Da für jede der nötigen offenen Mengen ein Punkt der Form  $(p, q)$  mit geeignetem  $q \in M_2$  enthalten ist, folgt, dass

$$p \in U_{1,p} \cap \dots \cap U_{n_p,p} = U_p$$

gilt. Damit ist  $U_p$  zum einen offen, zum anderen eine Umgebung von  $p$ . Die  $O_{i_1,p}, \dots, O_{i_{n_p},p}$  überdecken dann sogar den Streifen  $U_p \times M_2$ . Es folgt weiter

$$M_1 = \bigcup_{p \in M_1} U_p,$$

womit wir wieder eine offene Überdeckung  $\{U_p\}_{p \in M_1}$ , diesmal von  $M_1$  gefunden haben. Es genügen daher endlich viele  $U_{p_1}, \dots, U_{p_m}$ , um das Kompaktum  $M_1$  zu überdecken. Die endlich vielen zugehörigen Streifen  $U_{p_1} \times M_2, \dots, U_{p_m} \times M_2$  überdecken dann  $M_1 \times M_2$  und werden ihrerseits jeweils von endlich vielen der ursprünglichen  $O_i$  überdeckt.  $\square$

Man beachte, dass der Beweis von Satz 7.6.6 nach entsprechender Interpretation wörtlich der selbe wie zum Spezialfall in Satz 3.4.9 ist. Für beliebige endliche kartesische Produkte gilt die entsprechende Aussage ebenfalls, siehe auch Übung ??:

**Korollar 7.6.7** Sind  $(M_1, \mathcal{M}_1), \dots, (M_n, \mathcal{M}_n)$  topologische Räume, so ist auch  $M_1 \times \dots \times M_n$  in der Produkttopologie kompakt.

Die letzte Übertragung, die wir aus Kapitel 3 übernehmen wollen, ist der Begriff der zusammenhängenden Teilmenge. Hier basierte unsere bisherige Definition 3.4.13 bereits ausschließlich auf der Verwendung von offenen Mengen, womit wir unmittelbar die folgende Formulierung auch für topologische Räume erhalten:

**Definition 7.6.8 (Zusammenhang)** Sei  $(M, \mathcal{M})$  ein topologischer Raum. Dann heißt  $M$  zusammenhängend, falls es keine zwei offenen Mengen  $O_1, O_2 \subseteq M$  mit  $O_1 \cup O_2 = M$  und  $O_1 \cap O_2 = \emptyset$  außer  $M$  und  $\emptyset$  gibt. Eine Teilmenge von  $M$  heißt zusammenhängend, wenn sie bezüglich der Unterraumtopologie zusammenhängend ist.

Da wir aus Satz 3.4.12 wissen, dass die abgeschlossenen Intervalle zusammenhängend sind, können wir diese dazu verwenden, um die Zusammenhangseigenschaften anderer topologischer Räume zu testen. Dies führt auf die folgende Begriffsbildungen:

**Definition 7.6.9 (Wegzusammenhang)** Sei  $(M, \mathcal{M})$  ein topologischer Raum.

Abbildung 7.9: Stetige Bilder von wegzusammenhängenden Teilmengen sind wegzusammenhängend.

- i.) Eine stetige Abbildung  $\gamma: [a, b] \rightarrow M$  heißt Weg in  $M$  mit Startpunkt  $\gamma(a)$  und Zielpunkt  $\gamma(b)$ . In diesem Fall sagt man,  $\gamma$  verbindet  $\gamma(a)$  und  $\gamma(b)$ .
- ii.) Gibt es zu je zwei Punkten  $p, q \in M$  einen Weg, der  $p$  und  $q$  verbindet, so heißt  $M$  wegzusammenhängend.

Wir zeigen nun zwei nützliche Eigenschaften von zusammenhängenden Teilmengen:

**Proposition 7.6.10** Sei  $(M, \mathcal{M})$  ein topologischer Raum.

- i.) Ist  $M$  wegzusammenhängend, so auch zusammenhängend.
- ii.) Ist  $C \subseteq M$  eine zusammenhängende Teilmenge und  $B \subseteq M$  mit  $C \subseteq B \subseteq C^{\text{cl}}$ , so ist auch  $B$  zusammenhängend. Insbesondere ist der Abschluss  $C^{\text{cl}}$  von  $C$  selbst wieder zusammenhängend.

BEWEIS: Sei  $M$  wegzusammenhängend und  $O_1, O_2 \subseteq M$  offen mit  $O_1 \cap O_2 = \emptyset$  sowie  $O_1 \cup O_2 = M$ , und beide Teilmengen seien nicht leer. Dann wählen wir  $p \in O_1$  und  $q \in O_2$  und finden einen Weg  $\gamma: [a, b] \rightarrow M$  mit  $\gamma(a) = p$  sowie  $\gamma(b) = q$ . Da  $\gamma$  stetig ist, sind  $\gamma^{-1}(O_1), \gamma^{-1}(O_2) \subseteq [a, b]$  offen in  $[a, b]$ , nach wie vor disjunkt und beide nicht leer, da  $a \in \gamma^{-1}(O_1)$  und  $b \in \gamma^{-1}(O_2)$ . Da zudem  $\gamma^{-1}(O_1) \cup \gamma^{-1}(O_2) = \gamma^{-1}(O_1 \cup O_2) = \gamma^{-1}(M) = [a, b]$  gilt, haben wir einen Widerspruch zum Zusammenhang von  $[a, b]$  erzielt, was den ersten Teil zeigt. Für den zweiten nehmen wir an,  $B$  sei nicht zusammenhängend. Es gibt also offene  $O_1, O_2 \subseteq M$  mit

- $B \subseteq O_1 \cup O_2$ ,
- $B \cap O_1 \neq \emptyset \neq B \cap O_2$ ,
- $(B \cap O_1) \cap (B \cap O_2) = \emptyset$ .

Da  $C \subseteq B$ , folgt ebenso  $C \subseteq O_1 \cup O_2$  und  $(C \cap O_1) \cap (C \cap O_2) = \emptyset$ . Da  $B \subseteq C^{\text{cl}}$ , gilt für alle  $p \in B$  und alle offenen  $O$  mit  $p \in O$

$$C \cap O \neq \emptyset$$

nach Proposition 7.5.7. Insbesondere für  $p \in B \cap O_1$  und  $q \in B \cap O_2$  liefert dies

$$C \cap O_1 \neq \emptyset \neq C \cap O_2.$$

Dies ist der Widerspruch zum Zusammenhang von  $C$ . □

Mit den gleichen Argumenten wie in Satz 3.5.9 erhalten wir folgendes Resultat:

**Proposition 7.6.11** Sei  $f: (M, \mathcal{M}) \rightarrow (N, \mathcal{N})$  eine stetige Abbildung zwischen topologischen Räumen.

- i.) Ist  $X \subseteq M$  zusammenhängend, so auch  $f(X) \subseteq N$ .
- ii.) Ist  $X \subseteq M$  wegzusammenhängend, so auch  $f(X) \subseteq N$ .

BEWEIS: Den ersten Teil erhält man wörtlich wie im Beweis zu Satz 3.5.9. Der zweite ist viel einfacher, da mit einem Weg  $\gamma: [a, b] \rightarrow M$  die Verkettung  $f \circ \gamma: [a, b] \rightarrow N$  ein Weg ist, der nun  $f(\gamma(a))$  mit  $f(\gamma(b))$  verbindet, siehe auch Abbildung 7.9.

Dieses Resultat könne wir also als eine abstrakte Version und drastische Verallgemeinerung des Zwischenwertsatzes aus Korollar 3.5.11 verstehen.

Wir schließen diesen Abschnitt mit einem überraschenden Beispiel einer zusammenhängenden aber nicht wegzusammenhängenden Teilmenge des  $\mathbb{R}^2 \cong \mathbb{C}$ . Insbesondere gilt die Umkehrung von Proposition 7.6.10, i.), im Allgemeinen nicht:

Abbildung 7.10: Der Graph von  $f(x) = \sin(\frac{1}{x})$ .

**Beispiel 7.6.12 (Sinuskurve des Topologen)** Wir versehen  $\mathbb{R}^2 \cong \mathbb{C}$  mit der üblichen metrischen Topologie. Sei dann  $f: (0, 1) \rightarrow \mathbb{R}$  durch

$$f(x) = \sin\left(\frac{1}{x}\right) \quad (7.6.4)$$

gegeben. Wir betrachten dann den Graph

$$X = \text{graph}(f) = \left\{ \begin{pmatrix} x \\ \sin\left(\frac{1}{x}\right) \end{pmatrix} \mid x \in (0, 1) \right\} \subseteq \mathbb{R}^2 \quad (7.6.5)$$

als Teilmenge von  $\mathbb{R}^2$ , siehe auch Abbildung 7.10. Nach Konstruktion ist  $\text{graph}(f)$  wegzusammenhängend, da wir die stetige Funktion  $f$  als Weg verwenden können: Ist nämlich  $p, q \in \text{graph}(f)$  vorgegeben, so gibt es eindeutige  $a, b \in (0, 1)$  mit

$$p = \begin{pmatrix} a \\ \sin\left(\frac{1}{a}\right) \end{pmatrix} \quad \text{und} \quad q = \begin{pmatrix} b \\ \sin\left(\frac{1}{b}\right) \end{pmatrix}, \quad (7.6.6)$$

womit

$$\gamma(t) = \begin{pmatrix} t \\ \sin\left(\frac{1}{t}\right) \end{pmatrix}, \quad t \in [a, b] \quad (7.6.7)$$

der gesuchte Weg ist. Allgemein erhält man dies auch aus Proposition 7.6.11, *ii.*). Damit ist  $X$  nach Proposition 7.6.10, *i.*), auch zusammenhängend und  $X^{\text{cl}}$  ist immer noch zusammenhängend nach Proposition 7.6.10, *ii.*). Der Abschluss von  $X$  ist

$$X^{\text{cl}} = \{0\} \times [-1, 1] \cup X, \quad (7.6.8)$$

siehe erneut Abbildung 7.10. Es zeigt sich nun, dass Punkte aus  $X^{\text{cl}} \setminus X = \{0\} \times [-1, 1]$  nicht durch einen stetigen Weg mit Punkten aus  $X$  verbunden werden können, siehe Übung ?? . Damit ist  $X^{\text{cl}}$  anders als  $X$  nicht länger wegzusammenhängend.

Bringt man den Begriff der gleichmäßigen Stetigkeit für Abbildungen zwischen metrischen Räumen mit Kompaktheit zusammen, so erhält man ohne weitere Schwierigkeiten folgende Verallgemeinerung von Satz 3.6.5:

**Proposition 7.6.13** *Sei  $f: (M, d) \rightarrow (M', d')$  eine stetige Abbildung zwischen metrischen Räumen. Ist  $K \subseteq M$  kompakt, so ist  $f|_K$  gleichmäßig stetig.*

BEWEIS: Der Beweis von Satz 3.6.5 kann wörtlich übernommen werden, wenn der Absolutbetrag durch die jeweiligen Metriken von  $M$  und  $M'$  ersetzt wird.  $\square$

## 7.7 Die Standardtopologie auf $\mathbb{R}^n$

Da  $\mathbb{R}^n$  das  $n$ -fache kartesische Produkt

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}} \quad (7.7.1)$$

ist, werden wir  $\mathbb{R}^n$  mit der Produkttopologie versehen, wobei  $\mathbb{R}$  immer die übliche, durch die Standardmetrik gegebene Topologie trägt. Diese Topologie ist die Standardtopologie von  $\mathbb{R}^n$ :

**Definition 7.7.1 (Standardtopologie)** Sei  $n \in \mathbb{N}$ . Die Standardtopologie von  $\mathbb{R}^n$  ist die Produkttopologie gemäß (7.7.1) mit der üblichen metrischen Topologie auf  $\mathbb{R}$ .

Wir fassen einige der wichtigen Eigenschaften dieser Topologie zusammen:

**Proposition 7.7.2** Sei  $n \in \mathbb{N}$ .

i.) Die Standardtopologie von  $\mathbb{R}^n$  wird durch die Metrik

$$\tilde{d}(x, y) = \max\{|x_1 - y_1|, \dots, |x_n - y_n|\} \quad (7.7.2)$$

induziert, wobei  $x, y \in \mathbb{R}^n$ .

ii.) Die Standardtopologie von  $\mathbb{R}^n$  ist Hausdorffsch.

iii.) Eine Folge  $(x_m)_{m \in \mathbb{N}}$  in  $\mathbb{R}^n$  konvergiert genau dann gegen  $x \in \mathbb{R}^n$ , wenn ihre Komponentenfolgen  $(x_{m,i})_{m \in \mathbb{N}}$  für alle  $i = 1, \dots, n$  gegen  $x_i$  in  $\mathbb{R}$  konvergieren.

iv.) Eine Folge  $(x_m)_{m \in \mathbb{N}}$  in  $\mathbb{R}^n$  ist genau dann eine Cauchy-Folge, wenn ihre Komponentenfolgen  $(x_{m,i})_{m \in \mathbb{N}}$  für alle  $i = 1, \dots, n$  Cauchy-Folgen in  $\mathbb{R}$  sind.

v.) Die Metrik (7.7.2) ist vollständig.

BEWEIS: Der erste Teil folgt aus Proposition 7.5.11 beziehungsweise Übung 7.8.11. Damit ist der zweite Teil klar nach Korollar 7.5.15. Nach Definition von  $\tilde{d}$  in (7.7.2) gilt genau dann  $\tilde{d}(x_m, x) < \epsilon$ , wenn

$$|x_{m,i} - x_i| < \epsilon$$

für alle  $i = 1, \dots, n$  gilt. Damit folgt der dritte Teil, siehe auch erneut Übung 7.8.11. Den vierten Teil erhält man genauso. Da in  $\mathbb{R}$  jede Cauchy-Folge dank der Vollständigkeit von  $\mathbb{R}$  konvergiert, folgt der letzte Teil aus iii.) und iv.).  $\square$

Die bisherige metrische Topologie von  $\mathbb{C}$  ist damit ein Spezialfall der Standardtopologie für  $n = 2$ , wenn man  $\mathbb{C} \cong \mathbb{R}^2$  identifiziert. Damit erhalten wir auch die topologischen Identifikationen

$$\mathbb{C}^n \cong \mathbb{R}^{2n} \quad (7.7.3)$$

für alle  $n \in \mathbb{N}$ , siehe auch Übung ?? . Aus diesem Grunde genügt es für topologische Fragen, nur den reellen Fall zu betrachten.

Als nächste Eigenschaft benötigen wir Charakterisierungen von stetigen Abbildungen. Die Produktstruktur erlaubt hier folgendes Resultat:

**Proposition 7.7.3** Sei  $n \in \mathbb{N}$ .

i.) Eine Abbildung  $f: (M, \mathcal{M}) \rightarrow \mathbb{R}^n$  von einem topologischen Raum  $(M, \mathcal{M})$  nach  $\mathbb{R}^n$  ist genau dann stetig, wenn jede Komponente

$$f_i = \text{pr}_i \circ f: M \rightarrow \mathbb{R} \quad (7.7.4)$$

stetig ist, wobei  $\text{pr}_i: \mathbb{R}^n \rightarrow \mathbb{R}$  für  $i = 1, \dots, n$  die  $i$ -te Komponente bezeichnet.

ii.) Für alle  $i = 1, \dots, n$  und  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n \in \mathbb{R}$  ist die Abbildung

$$j_i: \mathbb{R} \ni t \mapsto (x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n) \in \mathbb{R}^n \quad (7.7.5)$$

stetig.

BEWEIS: Sei  $p \in M$  und  $U \subseteq \mathbb{R}^n$  eine Umgebung von  $f(p) \in \mathbb{R}^n$ . Diese enthält eine offene Kugel  $\tilde{B}_r(f(p)) \subseteq U \subseteq \mathbb{R}^n$  um  $f(p)$  bezüglich der Metrik  $\tilde{d}$  aus (7.7.2). Sei nun  $q \in f^{-1}(\tilde{B}_r(f(p)))$ , dann gilt also  $\tilde{d}(f(q), f(p)) < r$  und damit

$$|f(q) - f_i(p)| < r$$

für alle  $i = 1, \dots, n$ . Es folgt

$$f^{-1}(\tilde{B}_r(f(p))) = \bigcap_{i=1}^n f_i^{-1}(\tilde{B}_r(f(p))). \quad (*)$$

Ist nun  $f$  stetig bei  $p$ , so ist  $f^{-1}(\tilde{B}_r(f(p)))$  eine Umgebung von  $p$ . Mit (\*) folgt

$$f^{-1}(\tilde{B}_r(f(p))) \subseteq f_i^{-1}(\tilde{B}_r(f(p)))$$

für alle  $i = 1, \dots, n$ , womit auch  $f_i^{-1}(\tilde{B}_r(f(p)))$  eine Umgebung von  $p$  ist. Damit folgt aus der Stetigkeit von  $f$  bei  $p$  die von  $f_i$  für alle  $i = 1, \dots, n$ . Sind umgekehrt alle Komponenten  $f_1, \dots, f_n$  stetig bei  $p$ , so sind alle  $f_1^{-1}(\tilde{B}_r(f(p))), \dots, f_n^{-1}(\tilde{B}_r(f(p)))$  Umgebungen von  $p$ . Damit ist aber auch ihr endlicher Durchschnitt (\*) eine Umgebung, was die Stetigkeit von  $f$  bei  $p$  bedeutet. Dies zeigt *i.*). Damit ist der zweite Teil aber klar, da  $\text{pr}_i \circ j_j$  für  $i = j$  schlichtweg die Identitätsabbildung ist.  $\square$

Im nächsten Schritt wollen wir die kompakten Teilmengen von  $\mathbb{R}^n$  bestimmen und so ein Analogon von Satz 3.4.11 finden. Mit dem Satz von Tikhonov können wir genauso argumentieren wie für Satz 3.4.11:

**Satz 7.7.4 (Heine-Borel II)** *Sei  $K \subseteq \mathbb{R}^n$ . Dann sind folgende Aussagen äquivalent:*

- i.) Die Teilmenge  $K$  ist kompakt.*
- ii.) Die Teilmenge  $K$  ist abgeschlossen und beschränkt.*
- iii.) Jede Folge in  $K$  besitzt eine in  $K$  konvergente Teilfolge.*

BEWEIS: In einem metrischen Raum sind kompakte Teilmengen nach Proposition 7.6.3, *iii.*), und Korollar 7.5.15 immer abgeschlossen. Die Beschränktheit einer kompakten Menge zeigt man dann analog zu Satz 3.4.11, da sie immer von endlich vielen offenen Kugeln von Radius 1 überdeckt wird. Somit ist die Implikation *i.)  $\implies$  ii.)* also in jedem metrischen Raum gültig. Ist  $K$  abgeschlossen und beschränkt, so gilt  $K \subseteq \tilde{B}_r(0)^{\text{cl}}$  für einen geeigneten Radius  $r > 0$ . Da  $\tilde{B}_r(0)^{\text{cl}} = [-r, r] \times \dots \times [-r, r]$  ein Produkt von abgeschlossenen Intervallen ist, ist  $\tilde{B}_r(0)^{\text{cl}}$  nach Satz 7.6.6 kompakt. Als abgeschlossene Teilmenge der kompakten Kugel  $\tilde{B}_r(0)^{\text{cl}}$  ist  $K$  nach Proposition 7.6.3, *i.*), ebenfalls kompakt. Sei nun  $K$  kompakt und  $(p_\ell)_{\ell \in \mathbb{N}}$  eine Folge in  $K$ . Dann ist das Bild

$$K_i = \text{pr}_i(K) \subseteq \mathbb{R}$$

nach Proposition 7.6.4 erneut kompakt. Die  $i$ -te Komponentenfolge  $(p_{n,i})_{n \in \mathbb{N}}$  liegt in  $K_i$  und hat nach dem 1-dimensionalen Satz von Heine-Borel eine konvergente Teilfolge für alle  $i = 1, \dots, n$ . Damit erhält man zunächst eine Teilfolge, so dass die ersten Komponenten dieser Teilfolge konvergieren. In dieser Teilfolge findet man dann eine weitere Teilfolge, so dass zudem die zweiten Komponenten konvergieren. Induktiv erhält man nach  $n$  Schritten insgesamt eine Teilfolge  $(p_{\ell_k})_{k \in \mathbb{N}}$ , für die alle  $n$  Komponentenfolgen konvergieren. Dies bedeutet aber die Konvergenz von  $(p_{\ell_k})_{k \in \mathbb{N}}$  in  $K$ , was *iii.*) zeigt. Gilt umgekehrt *iii.*), so hat auch jede Folge  $(x_\ell)_{\ell \in \mathbb{N}}$  in  $K_i = \text{pr}_i(K)$  eine konvergente Teilfolge: Wir können die anderen Koordinaten konstant lassen und erhalten so eine Folge  $(p_\ell)_{\ell \in \mathbb{N}}$  in  $K$  mit  $i$ -ter Komponentenfolge  $(x_\ell)_{\ell \in \mathbb{N}}$ . Also sind alle  $K_i \subseteq \mathbb{R}$  kompakt nach dem eindimensionalen Satz von Heine-Borel. Nun gilt sicherlich

$$K \subseteq K_1 \times \dots \times K_n,$$

womit  $K$  eine Teilmenge einer kompakten Teilmenge  $K_1 \times \dots \times K_n \subseteq \mathbb{R}$  wird. Da jede Folge in  $K$  eine konvergente Teilfolge in  $K$  besitzt, ist  $K$  abgeschlossen, also selbst kompakt.  $\square$

Abbildung 7.11: Die Kugel  $\tilde{B}_r(p)$  ist wegzusammenhängend.

Abbildung 7.12: Die Kugel  $B_1(0)$  bezüglich  $\|\cdot\|_1$  enthält eine Kugel  $\tilde{B}_2(0)$  bezüglich  $\|\cdot\|_2$ .

Man kann an verschiedenen Stellen mit alternativen Argumenten den Nachweis der Äquivalenzen führen. Wir haben mit diesem Satz nun ein einfach handhabbares Kriterium, um kompakte Teilmengen von  $\mathbb{R}^n$  zu charakterisieren.

**Korollar 7.7.5** Sei  $p \in \mathbb{R}^n$ .

- i.) Jede Umgebung  $U$  von  $p$  enthält eine kompakte Umgebung von  $p$ .
- ii.) Jede Umgebung  $U$  von  $p$  enthält eine wegzusammenhängende Umgebung von  $p$ .

BEWEIS: Wenn  $U$  eine Umgebung von  $p$  ist, so gibt es einen Radius  $r > 0$  mit

$$\tilde{B}_r(p) \subseteq U. \quad \square$$

Dann ist  $\tilde{B}_{\frac{r}{2}}(p)^{\text{cl}} \subseteq \tilde{B}_r(p) \subseteq U$  eine kompakte Umgebung von  $p$  in  $U$ . Ebenso ist  $\tilde{B}_r(p)$  wegzusammenhängend, was man durch Angabe expliziter stetiger Wege, die einen Punkt  $q \in B_r(p)$  mit  $p$  verbinden, direkt sieht, siehe auch Abbildung 7.11.

Tatsächlich ist  $\tilde{B}_r(p)$  sogar konvex: Je zwei Punkte in  $\tilde{B}_r(p)$  können durch die Verbindungsstrecken innerhalb von  $\tilde{B}_r(p)$  verbunden werden. Hier verwendet man allerdings die zusätzliche Struktur, dass  $\mathbb{R}^n$  ein Vektorraum ist. Konvexität ist kein Konzept metrischer oder gar topologischer Räume.

Wir kommen nun zum Hauptresultat dieses Abschnitts, mit dem wir die Standardtopologie von  $\mathbb{R}^n$  als die „einzig sinnvolle“ rechtfertigen wollen. Wir haben bereits viele Beispiele für Normen auf  $\mathbb{R}^n$  gesehen, die alle dazu herangezogen werden können, über die zugehörige Metriken Topologien zu definieren. Die Metriken sind dabei sehr verschieden. So sind die offenen „Kugeln“  $\tilde{B}_r(0)$  bezüglich der Metrik  $\tilde{d}$  aus (7.7.2) eben offene  $n$ -dimensionale Würfel und keineswegs kugelförmig, wie dies für die Kugeln  $B_r(0)$  bezüglich der euklidischen Metrik

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} \quad (7.7.6)$$

der Fall ist, siehe erneut Abbildung 7.2. Es zeigt sich aber, dass die Topologien für alle diese Normen mit der Standardtopologie übereinstimmen. Dies wollen wir nun zeigen, wofür wir folgendes vorbereitende Resultat benötigen:

**Proposition 7.7.6** Seien  $\|\cdot\|_1$  und  $\|\cdot\|_2$  Normen auf einem  $\mathbb{K}$ -Vektorraum  $V$ . Die zugehörigen Topologien  $\mathcal{V}_1$  und  $\mathcal{V}_2$  stimmen genau dann überein, wenn es Konstanten  $c_1 > 0$  und  $c_2 > 0$  gibt, so dass

$$\|v\|_1 \leq c_1 \|v\|_2 \quad (7.7.7)$$

und

$$\|v\|_2 \leq c_2 \|v\|_1 \quad (7.7.8)$$

für alle  $v \in V$  gilt.

BEWEIS: Wir nehmen zunächst an, dass die Topologien übereinstimmen. Dann ist  $B_r(0)$  bezüglich  $\|\cdot\|_1$  eine offene Umgebung von  $0$  bezüglich  $\mathcal{V}_1 = \mathcal{V}_2$  und enthält daher eine offene Kugel  $\tilde{B}_\epsilon(0)$  bezüglich  $\|\cdot\|_2$  mit einem geeigneten Radius  $\epsilon > 0$ , siehe Abbildung 7.12. Es gilt also

$$\|v\|_2 < \epsilon \implies \|v\|_1 < 1.$$

Abbildung 7.13: Die Kugel  $B_r(v)$  ist offen bezüglich der Topologie von  $\|\cdot\|_2$ .

Sei nun  $v \neq 0$  und damit  $\|v\|_2 \neq 0 \neq \|v\|_1$ . Dann gilt für  $w = \frac{\epsilon}{\|v\|_2}v$  die Ungleichung  $\|w\|_2 < \epsilon$  und damit

$$1 > \|w\|_1 = \left\| \frac{\epsilon}{\|v\|_2} v \right\|_1 = \frac{\epsilon}{\|v\|_2} \|v\|_1$$

was

$$\|v\|_1 < \frac{1}{\epsilon} \|v\|_2$$

zeigt. Damit ist (7.7.7) mit  $c_1 = \frac{1}{\epsilon}$  erfüllt. Aus Symmetriegründen folgt auch (7.7.8), da  $\tilde{B}_1(0)$  eine offene Umgebung von 0 bezüglich  $\|\cdot\|_1$  ist. Es gelte nun umgekehrt (7.7.7) und (7.7.8). Seien  $u, v \in V$  und  $r > 0$  mit  $w \in B_r(0)$  gegeben, also

$$\|v - w\|_1 < r.$$

Wir betrachten nun  $u \in \tilde{B}_\epsilon(w)$ , wobei

$$\epsilon = \frac{r - \|v - w\|_1}{c_1} > 0.$$

Dann gilt

$$\begin{aligned} \|u - v\|_1 &\leq \|u - w\|_1 + \|w - v\|_1 \\ &< c_1 \|u - w\|_2 + \|w - v\|_1 \\ &< r - \|v - w\|_1 + \|w - v\|_1 \\ &= r, \end{aligned}$$

was  $u \in B_r(v)$  zeigt. Es folgt

$$\tilde{B}_\epsilon(w) \subseteq B_r(v).$$

Da  $w \in B_r(v)$  beliebig ist und da wir immer solch ein  $\epsilon_w > 0$  passend zu  $w$  finden können, ist

$$B_r(v) = \bigcup_{w \in B_r(v)} \tilde{B}_{\epsilon_w}(w)$$

eine Vereinigung von offenen Kugeln  $\tilde{B}_{\epsilon_w}(w)$  bezüglich  $\|\cdot\|_2$  und damit eine offene Menge in  $\mathcal{V}_2$ . Da jede offene Menge in  $\mathcal{V}_1$  eine Vereinigung von offenen Kugeln  $B_r(v)$  ist, folgt, dass jede offene Menge in  $\mathcal{V}_1$  auch offen in  $\mathcal{V}_2$  ist. Es gilt also

$$\mathcal{V}_1 \subseteq \mathcal{V}_2.$$

Mit der anderen Abschätzung (7.7.8) zeigt man analog  $\mathcal{V}_2 \subseteq \mathcal{V}_1$ , womit insgesamt die Gleichheit folgt.  $\square$

**Definition 7.7.7 (Äquivalente Normen)** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Zwei Normen  $\|\cdot\|_1$  und  $\|\cdot\|_2$  auf  $V$  heißen äquivalent, falls es  $c_1 > 0$  und  $c_2 > 0$  gibt, so dass

$$\|v\|_1 \leq c_1 \|v\|_2 \tag{7.7.9}$$

und

$$\|v\|_2 \leq c_2 \|v\|_1 \tag{7.7.10}$$

für alle  $v \in V$  gilt.



Mit anderen Worten sind zwei Normen äquivalent, falls ihre Topologien übereinstimmen. Die beiden Ungleichungen lassen sich auch in eine Ungleichung zusammenfassen. Die Normen sind äquivalent, falls es  $c, C > 0$  gibt, so dass

$$c\|v\|_1 \leq \|v\|_2 \leq C\|v\|_1 \quad (7.7.11)$$

für alle  $v \in V$  gilt.

In unendlichen Dimensionen können Vektorräume viele inäquivalente Normen besitzen. Ist  $V$  dagegen endlich-dimensional, ist dies nicht der Fall:

**Satz 7.7.8 (Äquivalenz von Normen)** Sei  $n \in \mathbb{N}$ . Dann sind je zwei Normen auf  $\mathbb{K}^n$  äquivalent.

BEWEIS: Sei  $\|\cdot\|$  eine beliebige Norm auf  $\mathbb{K}^n$  und  $\|\cdot\|_\infty$  die Maximumsnorm bezüglich der kanonischen Basis, die wir zur Definition der Produktmetrik  $\tilde{d}$  in (7.7.2) herangezogen haben. Wir entwickeln  $v = \sum_{i=1}^n v_i e_i \in \mathbb{K}^n$  bezüglich der Standardbasis und erhalten

$$\|v\| \leq \sum_{i=1}^n |v_i| \|e_i\| \leq c \max\{|v_1|, \dots, |v_n|\} = c\|v\|_\infty$$

nach der Dreiecksungleichung, wobei wir  $c = \sum_{i=1}^n \|e_i\|$  verwenden können. Dies zeigt eine der beiden notwendigen Abschätzungen. Wir behaupten, dass damit  $\|\cdot\|$  bezüglich der Standardtopologie stetig ist. Ist nämlich  $(v_m)_{m \in \mathbb{N}}$  eine Folge, die gegen  $v \in \mathbb{K}^n$  konvergiert, so gilt

$$\|\|v_m\| - \|v\|\| \leq \|v_m - v\| \leq c\|v_m - v\| \rightarrow 0$$

für  $m \rightarrow \infty$ , womit  $\lim_{m \rightarrow \infty} \|v_m\| = \|v\|$  gezeigt ist. Dies ist die (Folgen-)Stetigkeit von  $\|\cdot\|$ . Sei nun

$$K = \{v \in \mathbb{K}^n \mid \|v\|_\infty = 1\}$$

die Oberfläche des Einheitswürfels. Da  $\|\cdot\|_\infty$  ebenfalls stetig ist, ist  $K = \|\cdot\|_\infty^{-1}(\{1\})$  eine abgeschlossene Teilmenge, die zudem beschränkt ist. Also ist  $K$  kompakt bezüglich der Standardtopologie. Da auch  $\|\cdot\|$  stetig ist, gibt es auf  $K$  ein Minimum

$$\tilde{c} = \min\{\|v\| \mid v \in K\}$$

nach Korollar 7.6.5, ii.). Da  $0 \notin K$ , folgt  $\tilde{c} > 0$ . Für beliebiges  $v \neq 0$  ist  $\frac{v}{\|v\|_\infty} \in K$  und damit

$$\left\| \frac{v}{\|v\|_\infty} \right\| \geq \tilde{c}, \text{ was}$$

$$\|v\| \geq \tilde{c}\|v\|_\infty$$

zeigt. □

**Bemerkung 7.7.9** Damit ist die Standardtopologie also die eindeutig bestimmte Topologie auf  $\mathbb{K}^n$ , die sich durch Normen beschreiben lässt. Welche Norm man verwendet, ist dank Satz 7.7.8 egal, solange wir an topologischen Konzepten und Fragestellungen interessiert sind. Wir werden daher von dieser Freiheit gelegentlich Gebrauch machen, da je nach Fragestellung eine Norm besser geeignet sein kann als eine andere. Oft sind die euklidische Norm

$$\|v\|_2 = \sqrt{\sum_{i=1}^n |v_i|^2} \quad (7.7.12)$$

oder eben die Maximumsnorm

$$\|v\|_\infty = \max\{|v_1|, \dots, |v_n|\} \quad (7.7.13)$$

bezüglich der Standardbasis sehr nützlich. Wir werden im Folgenden hauptsächlich die euklidische Norm verwenden und daher kurz  $\|v\| = \|v\|_2$  schreiben. Die offenen Kugeln  $B_r(v) \subseteq \mathbb{K}$  seien von nun an immer bezüglich der euklidischen Norm gemeint, sofern nichts anderes verabredet wurde. Die euklidische Metrik wird mit  $d$  bezeichnet.

**Bemerkung 7.7.10 (Vollständigkeit von  $\mathbb{K}$ )** Die Frage nach Cauchy-Folgen ist keine topologische Fragestellung. Daher ist, anders als bei Stetigkeit etc. zunächst nicht klar, ob  $\mathbb{K}^n$  auch bezüglich anderer Normen vollständig ist. Mit den Abschätzungen (7.7.9) und (7.7.10) sieht man aber direkt, dass nicht nur Konvergenz (ein topologisches Konzept), sondern auch Cauchy-Folgen für eine äquivalente Norm wieder Cauchy-Folgen sind. Daher ist  $\mathbb{K}^n$  bezüglich aller Normen vollständig.

Es sei jedoch gewarnt, dass  $\mathbb{K}^n$  durchaus andere Topologien außer der Standardtopologie tragen kann. Die diskrete Topologie ist eine ebenfalls metrische Topologie, die von der Standardtopologie grundverschieden ist. Die indiskrete Topologie ist nicht einmal Hausdorffsch, geschweige denn metrisch. Wir werden  $\mathbb{K}^n$  aber immer mit der Standardtopologie betrachten.

Als kleine Anwendung unserer neuen Künste wollen wir auch für Teilmengen  $X \subseteq \mathbb{R}^n$  zeigen, dass die stetigen Funktionen gute algebraische Eigenschaften besitzen. Dies lässt sich einerseits völlig analog zu Proposition 3.3.7 unter Verwendung von Resultaten zu konvergenten Folgen (und Netzen) erreichen. Es gibt allerdings einen konzeptionellen und einfacheren Weg, der das Entscheidende klarer aufzeigt.

**Beispiel 7.7.11 (Stetigkeit der Addition und Multiplikation)** Die Additionsabbildung

$$+: \mathbb{K} \times \mathbb{K} \ni (x, y) \mapsto x + y \in \mathbb{K} \quad (7.7.14)$$

und die Multiplikationsabbildung

$$\cdot: \mathbb{K} \times \mathbb{K} \ni (x, y) \mapsto xy \in \mathbb{K} \quad (7.7.15)$$

sind stetig. Man beachte, dass wir hierzu die Produkttopologie auf  $\mathbb{K} \times \mathbb{K}$  verwenden, um dieser Aussage Leben einzuhauchen. Der Nachweis ist nun einfach, da es genügt, Folgenstetigkeit zu zeigen, weil  $\mathbb{K} \times \mathbb{K}$  ja metrisch ist. Seien also  $(x_n, y_n) \in \mathbb{K} \times \mathbb{K}$  für  $n \in \mathbb{N}$  vorgegeben, so dass  $(x_n, y_n)_{n \in \mathbb{N}} \rightarrow (x, y)$  bezüglich der Standardmetrik von  $\mathbb{K} \times \mathbb{K}$  konvergiert. Nach Proposition 7.7.2, *iii.*), bedeutet dies einfach

$$x_n \rightarrow x \quad \text{und} \quad y_n \rightarrow y. \quad (7.7.16)$$

Die Rechenregeln für konvergente Folgen aus Proposition 2.3.5 zeigen dann

$$x_n + y_n \rightarrow x + y \quad \text{sowie} \quad x_n y_n \rightarrow xy, \quad (7.7.17)$$

was gerade die Folgenstetigkeit von (7.7.14) und (7.7.15) ist. Etwas allgemeiner haben wir diese Stetigkeitsaussage bereits in Beispiel 7.3.6, *vi.*), kennengelernt.

Mit einer ähnlichen Argumentation lassen sich die anderen Rechenregeln aus Proposition 2.3.5 als Stetigkeitsaussagen von Abbildungen  $\mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$  etc. interpretieren. Wichtig an diesem neuen Gesichtspunkt ist, dass wir erst jetzt über die adequate Sprache der Produkttopologie verfügen.

**Beispiel 7.7.12 (Stetigkeit der Diagonalen)** Die Abbildung

$$\Delta: \mathbb{K} \ni x \mapsto (x, x) \in \mathbb{K} \times \mathbb{K} \quad (7.7.18)$$

ist stetig, da beide Komponenten die Identität sind und daher Proposition 7.7.3, *i.*), zum Einsatz kommen kann. Entsprechend sind die Diagonalen

$$\Delta: \mathbb{K} \ni x \mapsto (x, \dots, x) \in \mathbb{K}^n \quad (7.7.19)$$

ebenfalls stetig. Die Resultate bleiben auch für Teilmengen  $X \subseteq \mathbb{K}$  richtig. Alternativ können wir mit Beispiel 7.3.6, *v.*), argumentieren, da  $\Delta$  sicherlich linear ist.

Ein letztes Beispiel sind die Produktabbildungen:

**Beispiel 7.7.13 (Produktabbildungen)** Seien

$$f_1, \dots, f_m: \mathbb{R}^n \longrightarrow \mathbb{R} \quad (7.7.20)$$

stetige Abbildungen. Dann ist auch  $F: \mathbb{R}^n \longrightarrow \mathbb{R}^m$  mit

$$F(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)) \quad (7.7.21)$$

stetig. Die Argumentation ist wie in Beispiel 7.7.12, das wir hier verallgemeinern. Insbesondere ist für stetige Funktionen

$$f_1, \dots, f_n: \mathbb{R} \longrightarrow \mathbb{R} \quad (7.7.22)$$

die Produktabbildung  $F: \mathbb{R}^n \longrightarrow \mathbb{R}^n$  mit

$$F(x_1, \dots, x_n) = (f_1(x_1), \dots, f_n(x_n)) \quad (7.7.23)$$

stetig. Hierzu beachte man, dass jedes einzelne  $f_i$  als Abbildung

$$\tilde{f}_i: \mathbb{R}^n \ni x \mapsto f_i(x_i) \in \mathbb{R} \quad (7.7.24)$$

stetig ist: Es gilt  $\tilde{f}_i = f_i \circ \text{pr}_i$  mit der stetigen Projektion  $\text{pr}_i: \mathbb{R}^n \ni x \mapsto x_i \in \mathbb{R}$  auf die  $i$ -te Komponente. Daher ist  $\tilde{f}_i$  als Verkettung wieder stetig und  $F$  als Produktabbildung der  $\tilde{f}_1, \dots, \tilde{f}_n$  ebenso. Analoge Aussagen gelten für Teilmengen von  $\mathbb{R}^n$  als Definitionsbereiche.

Man beachte, dass wir zu diesem Zeitpunkt für die Stetigkeitsnachweise keinerlei Abschätzungen mehr verwenden müssen, sondern sehr viel einfachere abstrakte Argumente zur Verfügung haben. Zusammen liefern diese Beispiele nun folgendes Resultat:

**Proposition 7.7.14** *Sei  $X \subseteq \mathbb{R}^n$  eine nicht-leere Teilmenge. Dann bilden die stetigen Funktionen  $\mathcal{C}(X, \mathbb{K})$  auf  $X$  eine Unteralgebra aller Funktionen  $\text{Abb}(X, \mathbb{K})$ . Ist  $X \subseteq \mathbb{R}^n$  offen, so gilt  $\mathcal{C}(X, \mathbb{K}) \subseteq \mathcal{B}_{\text{loc}}(X, \mathbb{K})$ .*

BEWEIS: Seien  $f, g \in \mathcal{C}(X, \mathbb{K})$  stetig. Dann gilt

$$f + g = + \circ (f, g) \circ \Delta$$

und

$$fg = \cdot \circ (f, g) \circ \Delta$$

mit  $+$  und  $\cdot$  aus Beispiel 7.7.11, der Produktabbildung  $(f, g): X \times X \longrightarrow \mathbb{K} \times \mathbb{K}$  aus Beispiel 7.7.13 sowie der Diagonalen  $\Delta: X \longrightarrow X \times X$  aus Beispiel 7.7.12. Die Verkettung stetiger Funktionen ist dann stetig, was  $f + g, fg \in \mathcal{C}(X, \mathbb{K})$  zeigt. Dass stetige Funktionen lokal beschränkt sind, folgt aus Korollar 7.6.5, *ii.*), und der Beobachtung in Korollar 7.7.5, *i.*).  $\square$

## 7.8 Übungen

**Übung 7.8.1 (Umgekehrte Dreiecksungleichung)** Sei  $V$  ein normierter Vektorraum. Zeigen Sie die umgekehrte Dreiecksungleichung

$$|\|v\| - \|w\|| \leq \|v - w\| \quad (7.8.1)$$

für alle  $v, w \in V$ .

**Übung 7.8.2 ( $p$ -Normen auf direkten Summen)** Seien  $V_1, \dots, V_n$  endlich viele normierte Vektorräume und  $p \in [1, \infty]$ . Dann definiert man die  $p$ -Norm auf der direkten Summe  $V = V_1 \oplus \dots \oplus V_n$  durch

$$\|v\|_p = \begin{cases} \sqrt[p]{\|v_1\|_{V_1}^p + \dots + \|v_n\|_{V_n}^p} & \text{für } p \in [1, \infty) \\ \max\{\|v_1\|_{V_1}, \dots, \|v_n\|_{V_n}\} & \text{für } p = \infty, \end{cases} \quad (7.8.2)$$

wobei  $v = v_1 + \dots + v_n \in V_1 \oplus \dots \oplus V_n$  die jeweiligen Komponenten sind.

- i.) Zeigen Sie, dass  $\|\cdot\|_p$  für alle Werte von  $p$  eine Norm auf  $V$  liefert.
- ii.) Betrachten Sie  $1 \leq q \leq p \leq \infty$  und finden Sie wechselseitige Ungleichungen zwischen  $\|\cdot\|_q$  und  $\|\cdot\|_p$ . Folgern Sie, dass alle diese Normen auf  $V$  äquivalent sind.

**Übung 7.8.3 (Isometrien)** Sei  $\Phi: (M, d) \longrightarrow (M', d')$  eine Isometrie zwischen metrischen Räumen.

- i.) Zeigen Sie, dass  $\Phi$  injektiv ist. Finden Sie Beispiele für injektive aber nicht isometrische Abbildungen.
- ii.) Zeigen Sie, dass die Verkettung zweier Isometrien wieder eine Isometrie ist.
- iii.) Zeigen Sie, dass eine bijektive Isometrie eine isometrische Umkehrabbildung besitzt.

**Übung 7.8.4 (Die Metrik  $d'$ )** Sei  $(M, d)$  ein metrischer Raum. Zeigen Sie, dass durch (7.1.22) eine neue Metrik auf  $M$  definiert wird.

**Übung 7.8.5 (Die diskrete Metrik)** Sei  $M$  eine nicht-leere Menge mit der diskreten Metrik  $d$ .

- i.) Zeigen Sie, dass die offenen Kugeln  $B_r(p)$  um einen Punkt  $p \in M$  je nach Wert des Radius  $r > 0$  entweder  $B_r(p) = \{p\}$  oder  $B_r(p) = M$  erfüllen.
- ii.) Zeigen Sie, dass eine Folge  $(p_n)_{n \in \mathbb{N}}$  in  $M$  genau dann gegen  $p \in M$  konvergiert, wenn es ein  $N \in \mathbb{N}$  mit  $p_n = p$  für  $n \geq N$  gibt.
- iii.) Zeigen Sie, dass jede Teilmenge von  $M$  offen und abgeschlossen ist.
- iv.) Zeigen Sie, dass eine nicht-leere Teilmenge von  $M$  genau dann zusammenhängend ist, wenn sie aus nur einem Punkt besteht.
- v.) Sei  $\Phi: (M, d) \longrightarrow (M', d')$  eine Abbildung in eine ebenfalls mit der diskreten Metrik  $d'$  versehene Menge. Zeigen Sie, dass  $\Phi$  genau dann isometrisch ist, wenn  $\Phi$  injektiv ist.

**Übung 7.8.6 (Offene Mengen aus offenen Kugeln)** Sei  $(M, d)$  ein metrischer Raum. Zeigen Sie, dass eine offene Menge  $O \subseteq M$  als geeignete Vereinigung offener Kugeln geschrieben werden kann.

**Übung 7.8.7 (Direkte Summe von Banach-Räumen)**

**Übung 7.8.8 (Kofinite Topologie)** Sei  $M$  eine nicht-leere Menge.

- i.) Zeigen Sie, dass die kofinite Topologie aus Beispiel 7.5.2, iv.), tatsächlich eine Topologie ist.
- ii.) Welche Teilmengen sind abgeschlossen bezüglich der kofiniten Topologie?
- iii.) Zeigen Sie, dass die kofinite Topologie genau dann mit der diskreten Topologie auf  $M$  übereinstimmt, wenn  $M$  eine endliche Menge ist.

**Übung 7.8.9 (Zariski-Topologie)** Sei  $n \in \mathbb{N}$ .

- i.) Zeigen Sie, dass die Zariski-Topologie aus Beispiel 7.5.2, v.), auf  $\mathbb{C}^n$  tatsächlich eine Topologie definiert.
- ii.) Zeigen Sie, dass für  $n = 1$  die Zariski-Topologie mit der kofiniten Topologie übereinstimmt.

**Übung 7.8.10 (Metrische Relativtopologie)** Sei  $(M, d)$  ein metrischer Raum mit einer Teilmenge  $U \subseteq M$ . Zeigen Sie, dass die Relativtopologie von  $U$  bezüglich der metrischen Topologie von  $M$  mit der metrischen Topologie der auf  $U$  eingeschränkten Metrik übereinstimmt.

**Übung 7.8.11 (Produktmetrik)** Seien  $(M_1, d_1), \dots, (M_k, d_k)$  metrische Räume. Definieren Sie dann

$$d(p, q) = \max\{d_1(p_1, q_1), \dots, d_k(p_k, q_k)\} \quad (7.8.3)$$

für  $p = (p_1, \dots, p_k)$  und  $q = (q_1, \dots, q_k) \in M = M_1 \times \dots \times M_k$ .

- i.) Zeigen Sie, dass  $d$  eine Metrik auf  $M$  liefert.
- ii.) Sei  $r > 0$ . Zeigen Sie, dass die offene Kugel  $B_r(p)$  um  $p \in M$  mit Radius  $r$  das kartesische Produkt der offenen Kugeln  $B_r(p_1), \dots, B_r(p_k)$  ist.
- iii.) Seien nun  $r_1, \dots, r_k \in \mathbb{R}^+$ . Zeigen Sie, dass für alle Punkte  $p_1 \in M_1, \dots, p_k \in M_k$  das kartesische Produkt  $O = B_{r_1}(p_1) \times \dots \times B_{r_k}(p_k)$  im allgemeinen zwar keine offene Kugel bezüglich der Produktmetrik ist, aber zumindest eine offene Teilmenge von  $M$  liefert.
- iv.) Zeigen Sie, dass eine Folge  $(p^{(n)})_{n \in \mathbb{N}}$  in  $M$  mit  $p_n = (p_1^{(n)}, \dots, p_k^{(n)})$  genau dann gegen einen Punkt  $p = (p_1, \dots, p_k)$  in  $M$  konvergiert, wenn die Komponentenfolgen  $(p_i^{(n)})$  für alle  $i = 1, \dots, k$  gegen  $p_i \in M_i$  konvergieren.
- v.) Zeigen Sie schließlich die Verallgemeinerung von Proposition 7.5.11 für beliebige endliche Produkte von metrischen Räumen: die metrische Topologie von  $(M, d)$  stimmt mit der Produkttopologie der einzelnen Faktoren überein.

**Übung 7.8.12 (Konvergenz in der indiskreten Topologie)** Sei  $M$  eine Menge mit der indiskreten Topologie aus Beispiel 7.5.2, iii.). Zeigen Sie, dass eine Folge  $(p_n)_{n \in \mathbb{N}}$  immer konvergiert und jeder Punkt ein Grenzwert ist.

**Übung 7.8.13 (Kompaktheit der Eistüte)**



## Kapitel 8

# Differenzierbarkeit in $\mathbb{R}^n$

Bislang hatten wir Funktionen einer reellen oder komplexen Variable mit Werten in  $\mathbb{R}$  oder  $\mathbb{C}$  auf ihre Differenzierbarkeit untersucht. Dies stellt selbstverständlich einen wichtigen Spezialfall mit besonderen Eigenschaften dar, ist aber für viele Problemstellungen deutlich zu einschränkend. In diesem Kapitel wollen wir daher unseren Blickwinkel erweitern und Funktionen betrachten, die auf geeigneten Teilmengen  $X \subseteq \mathbb{R}^n$  definiert sind und auch Werte in höherdimensionalen Vektorräumen annehmen dürfen. Wir betrachten daher Abbildungen

$$f: X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m. \quad (8.1)$$

Es gibt nun zwei prinzipielle Sichtweisen um für Funktionen der Form (8.1) Differenzierbarkeit und andere analytische Konzepte einzuführen. Zum einen kann man die kanonischen linearen Koordinaten  $(x_1, \dots, x_n)$  von  $\mathbb{R}^n$  verwenden und erhält so eine Funktion von  $n$  reellen Variablen. Im Bild verwendet man ebenfalls die kanonischen Koordinaten, womit (8.1) dann als  $m$  reell-wertige Komponentenfunktionen jeweils von  $n$  Variablen aufgefasst werden kann.

Diese Sichtweise ist zwar naheliegend aber für viele Fragestellungen nicht angemessen. Schon bei komplex-wertigen Funktionen haben wir die Werte nicht in Real- und Imaginärteil zerlegt und getrennt betrachtet. Vielmehr ergeben sich weitere Einsichten, wenn man die komplexe Natur ernst nimmt und die Werte in  $\mathbb{C}$  als Einheit versteht.

Dies führt damit zum zweiten Standpunkt für Abbildungen der Form (8.1). Wir verwenden als Ziel einen Vektorraum  $W = \mathbb{R}^m$ , ohne explizit auf die Koordinaten Bezug zu nehmen. Dies erlaubt es auch andere (lineare) Koordinaten zu betrachten. Für den Definitionsbereich verfahren wir genauso und verstehen  $X \subseteq V = \mathbb{R}^n$  ebenfalls als Teilmenge des Vektorraums  $V = \mathbb{R}^n$  ohne die kanonischen Koordinaten zu spezifizieren. Wir wollen lediglich von der umgebenden Vektorraumstruktur Gebrauch machen. Da  $X$  eventuell als Teilmenge von  $\mathbb{R}^n$  selbst kein Vektorraum mehr zu sein braucht, kann es sehr wohl nützlich sein, auch andere Koordinaten als lineare zu verwenden, wenn diese der Geometrie von  $X$  besser angepasst sind.

Aus diesem Grunde betrachten wir also Abbildungen

$$f: X \subseteq V \longrightarrow W \quad (8.2)$$

auf einem Definitionsbereich  $X \subseteq V$  in einem reellen Vektorraum  $V$  mit Werten in einem reellen (oder auch komplexen) Vektorraum  $W$ . Um jedoch Analysis und insbesondere Differenzierbarkeit formulieren zu können, benötigen wir sowohl auf  $X$  als auch auf  $W$  einen Konvergenzbegriff. Daher spezifizieren wir  $V$  und  $W$  weiter als normierte Räume oder sogar als *Banach-Räume*. Im Falle endlich-dimensionaler Banach-Räume wissen wir zum einen  $V \simeq \mathbb{R}^n$  als Vektorraum. Zum anderen garantiert Satz 7.7.8, dass die Wahl der Norm auf  $V$  keine weiteren Auswirkungen auf die Topologie hat, da ja alle Normen in endlichen Dimensionen äquivalent sind. Nach Wahl eines linearen Isomorphismus und einer Norm sind wir von (8.2) wieder zurück bei (8.1).

Allerdings haben wir in (8.2) einen deutlich allgemeineren Standpunkt gewonnen. Im Folgenden sind wir zwar meistens an der endlich-dimensionalen Situation interessiert, wo wir frei zwischen (8.1) und (8.2) wechseln können. Etliche Definitionen und Resultate formulieren wir allerdings trotzdem direkt für normierte Räume beziehungsweise für Banach-Räume, wenn die Vollständigkeit wichtig wird.

## 8.1 Totale Differenzierbarkeit

In Abschnitt 5.1 hatten wir verschiedene äquivalente Definitionen und Charakterisierungen von Differenzierbarkeit in einer Dimension kennengelernt. In höheren Dimensionen wollen wir die Approximierbarkeit durch lineare Abbildungen in den Vordergrund stellen, was zu folgender Definition führt:

**Definition 8.1.1 (Totale Differenzierbarkeit)** Seien  $V$  und  $W$  normierte Vektorräume und  $X \subseteq V$  eine nicht-leere offene Teilmenge. Eine Abbildung  $f: X \rightarrow W$  heißt (total) differenzierbar oder auch Fréchet-differenzierbar bei  $x_0 \in X$ , wenn es eine stetige lineare Abbildung  $Df(x_0) \in L(V, W)$  gibt, so dass

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0) - Df(x_0)(x - x_0)}{\|x - x_0\|_V} = 0 \quad (8.1.1)$$

gilt. In diesem Fall heißt  $Df(x_0)$  die (totale) Ableitung oder auch Fréchet-Ableitung von  $f$  bei  $x_0$ .

Im Folgenden werden wir einfach von Differenzierbarkeit und Ableitung sprechen. Nur wenn die Unterscheidung zu anderen (schwächeren) Formulierungen benötigt wird, betonen wir (8.1.1) als totale Differenzierbarkeit.

### Bemerkung 8.1.2 (Totale Differenzierbarkeit)

- i.) Sind  $V$  und  $W$  endlich-dimensional, so ist nach Beispiel 7.3.6, v.), jede lineare Abbildung stetig, ja sogar Lipschitz-stetig, da wir ja  $V$  und  $W$  mit  $\mathbb{K}^n$  und  $\mathbb{K}^m$ , versehen mit der Standardtopologie identifizieren können. Hier verwendet man, wie schon so oft die Äquivalenz von Normen nach Satz 7.7.8. Damit ist die Bedingung, dass die Ableitung  $Df(x_0)$  eine stetige lineare Abbildung sein soll, automatisch erfüllt: Lediglich die Linearität sowie das Grenzverhalten (8.1.1) ist dann zu prüfen. Tatsächlich lässt sich zeigen, dass der Wertebereich  $W$  unerheblich ist und lineare Abbildungen automatisch stetig sind, sobald lediglich der Definitionsbereich  $V$  endlich-dimensional ist.
- ii.) Wie schon im eindimensionalen Fall gibt es prinzipiell einen wesentlichen Unterschied zwischen komplexer und reeller Differenzierbarkeit. Bei komplexer Differenzierbarkeit müssen beide normierte Vektorräume komplexe Vektorräume sein, damit die Ableitung als  $\mathbb{C}$ -lineare Abbildung überhaupt definiert werden kann. Man beachte, dass unter der kanonischen Identifikation  $\mathbb{R}^{2n} \cong \mathbb{C}^n$  nicht jede  $\mathbb{R}$ -lineare Abbildung  $A: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2m}$  automatisch auch  $\mathbb{C}$ -linear wird. Bei reellem Definitionsbereich  $X \subseteq V$  sind dagegen sowohl reelle als auch komplexe Vektorräume  $W$  als Bildbereich von Interesse, was keine Schwierigkeiten verursacht, da ein komplexer Vektorraum ja auch als reeller aufgefasst werden kann.
- iii.) Für die Definition ist die Vollständigkeit von  $V$  oder  $W$  noch nicht wichtig, weshalb wir hier noch mit beliebigen normierten Räumen auskommen. Dies wird sich allerdings ändern, wenn wir beispielsweise wieder Funktionenfolgen betrachten wollen.
- iv.) Äquivalent zu (8.1.1) ist die Bedingung, dass die Restabbildung  $R: X \rightarrow W$  mit

$$R(x) = f(x) - f(x_0) - Df(x_0)(x - x_0) \quad (8.1.2)$$

die Eigenschaft

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{R(x)}{\|x - x_0\|_V} = 0 \quad (8.1.3)$$



erfüllt, was äquivalent zu

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{\|R(x)\|_W}{\|x - x_0\|_V} = 0 \quad (8.1.4)$$

ist. Die Approximation von  $f(x)$  durch den konstanten Term  $f(x_0)$  und den linearen Term  $Df(x_0)(x - x_0)$  soll also im dem Sinne optimal sein, dass der Restterm schneller als linear verschwindet: Diese Vorstellung implementiert man durch den Vergleich von  $R(x)$  mit der Norm  $\|x - x_0\|_V$ . Es folgt aus (8.1.3) beziehungsweise (8.1.4) direkt

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} R(x) = 0, \quad (8.1.5)$$

womit wir  $R$  durch (8.1.5) in  $x = x_0$  stetig fortsetzen können. Allerdings ist (8.1.3) eine viel stärkere Forderung als (8.1.5), wie wir das ja im eindimensionalen Fall bereits gesehen haben.

v.) Gilt  $V = \mathbb{K} = W$ , so reproduziert Definition 8.1.1 die eindimensionale Situation. Dies ist gerade die Aussage von Proposition 5.1.4, *iii.*).

Um wieder von „der“ Ableitung sprechen zu können, müssen wir zeigen, dass  $Df(x_0)$  notwendigerweise eindeutig ist, sobald Differenzierbarkeit vorliegt:

**Proposition 8.1.3** *Seien  $V$  und  $W$  normierte Vektorräume und  $X \subseteq V$  eine nicht-leere offene Teilmenge.*

*i.) Ist  $f: X \rightarrow W$  bei  $x_0 \in X$  differenzierbar, so gilt*

$$Df(x_0)v = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} \quad (8.1.6)$$

*für alle  $v \in V$ .*

*ii.) Ist  $f: X \rightarrow W$  bei  $x_0 \in X$  differenzierbar, so ist die Ableitung  $Df(x_0)$  eindeutig bestimmt.*

BEWEIS: Da  $X \subseteq V$  offen ist, gibt es für jeden Vektor  $v \in V$  ein  $t_0 > 0$  derart, dass

$$x_0 + tv \in X$$

für alle  $|t| < t_0$ . In diesem Sinne ist (8.1.6) dann zu verstehen, da wir ja  $f$  auf den Punkt  $x_0 + tv$  überhaupt erst definieren müssen. Mit dem Restterm  $R(x_0 + tv)$  aus (8.1.2) gilt dann für  $t \neq 0$

$$\frac{f(x_0 + tv) - f(x_0)}{t} = Df(x_0)v + \frac{R(x_0 + tv)}{t},$$

da  $Df(x_0)$  ja linear ist. Nun gilt mit (8.1.4) aber für  $v \neq 0$

$$\lim_{\substack{t \rightarrow t_0 \\ t \neq 0}} \frac{\|R(x_0 + tv)\|_W}{|t|} = \lim_{\substack{t \rightarrow t_0 \\ t \neq 0}} \frac{\|R(x_0 + tv)\|_W}{\|x_0 + tv - x_0\|_V} \|v\|_V = \lim_{\substack{t \rightarrow t_0 \\ t \neq 0}} \frac{1}{|t|} \|R(x_0 + tv)\|_W = 0.$$

Da für  $v = 0$  der Restterm sowieso verschwindet, folgt (8.1.6). Damit ist *ii.*) aber auch klar.  $\square$

Wir können nun einige erste Beispiele von differenzierbaren Abbildungen finden. Später werden wir unseren Vorrat an Beispielen durch interessantere Konstruktionen drastisch aufstocken.

**Beispiel 8.1.4 (Total differenzierbare Abbildungen)** Seien  $V$  und  $W$  normierte Vektorräume.

i.) Jede konstante Abbildung  $f: X = V \longrightarrow W$  mit

$$f(x) = w_0 \quad (8.1.7)$$

für ein festes  $w_0 \in W$  ist differenzierbar mit Ableitung

$$Df(x_0) = 0 \quad (8.1.8)$$

für alle  $x_0 \in X$ . Dies ist unmittelbar aus der Definition klar, da  $f(x) - f(x_0) = 0$  für alle Paare  $x, x_0 \in X = V$ .

ii.) Jede stetige lineare Abbildung  $A \in L(V, W)$  ist an jeder Stelle  $x_0 \in X = V$  differenzierbar mit

$$DA(x_0) = A. \quad (8.1.9)$$

Auch dies ist klar, da für alle  $x \in X$  eben

$$A(x) - A(x_0) = A(x - x_0) \quad (8.1.10)$$

aufgrund der Linearität von  $A$  gilt. Der Restterm ist also identisch 0. Man beachte, dass dieses Beispiel letztlich der Grund dafür ist, die Stetigkeit der Ableitung explizit zu fordern, da sonst jede lineare Abbildung differenzierbar wäre. In unendlichen Dimensionen gibt es dann aber unstetige lineare Abbildungen, die trotzdem differenzierbar wären. Diese Situation möchte man vermeiden.

iii.) Wir betrachten  $V = W \oplus W$  mit der Maximumsnorm

$$\|(w, w')\|_V = \max\{\|w\|, \|w'\|\}. \quad (8.1.11)$$

Dann ist  $V$  wieder ein normierter Vektorraum und sogar ein Banach-Raum, sobald  $W$  ein Banach-Raum war. Als Abbildung verwenden wir die Addition

$$+: W \oplus W \ni (w, w') \mapsto w + w' \in W, \quad (8.1.12)$$

welche linear ist. Bezüglich (8.1.11) ist auch einfach zu sehen, dass  $+$  stetig ist. Es gilt

$$\|w + w'\|_W \leq \|w\|_W + \|w'\|_W \leq 2\|(w, w')\|_V \quad (8.1.13)$$

nach der Dreiecksungleichung. Damit erhalten wir als Operatornorm der Additionsabbildung

$$\|+\| = 2, \quad (8.1.14)$$

da für  $w = w'$  offenbar ein Gleichheitszeichen in (8.1.13) erreicht wird. Nach ii.) ist  $+$  also differenzierbar mit Ableitung

$$(D+)(w_0, w'_0)(u, u') = u + u' \quad (8.1.15)$$

für alle Punkte  $(w_0, w'_0) \in W \oplus W = V$ .

iv.) Als zweite Vektorraumoperation betrachten wir die Multiplikation mit Skalaren als Abbildung

$$\mu: \mathbb{K} \times V \ni (\alpha, v) \mapsto \alpha v \in V. \quad (8.1.16)$$

Auf  $\mathbb{K} \times V$  verwenden wir beispielsweise wieder die Maximumsnorm

$$\|(\alpha, v)\| = \max\{|\alpha|, \|v\|_V\}. \quad (8.1.17)$$

Dann ist  $\mu$  an jedem Punkt  $(\alpha_0, v_0) \in \mathbb{R} \times V$  differenzierbar mit der Ableitung

$$D\mu(\alpha_0, v_0)(\beta, u) = \alpha_0 u + \beta v_0. \quad (8.1.18)$$

Um dies zu zeigen, betrachten wir den Restterm

$$\begin{aligned} R(\alpha, v) &= \mu(\alpha, v) - \mu(\alpha_0, v_0) - (D\mu)(\alpha_0, v_0)((\alpha, v) - (\alpha_0, v_0)) \\ &= \alpha v - \alpha_0 v_0 - \alpha_0(v - v_0) - (\alpha - \alpha_0)v_0 \\ &= (\alpha - \alpha_0)(v - v_0). \end{aligned}$$

Um die Differenzierbarkeit zu zeigen, betrachten wir

$$\begin{aligned} \|R(\alpha, v)\|_V &= \|(\alpha - \alpha_0)(v - v_0)\|_V \\ &= |\alpha - \alpha_0| \|v - v_0\|_V \\ &\leq \max\{|\alpha - \alpha_0|, \|v - v_0\|_V\}^2 \\ &= \|(\alpha, v) - (\alpha_0, v_0)\|^2. \end{aligned}$$

Damit folgt aber  $\lim_{(\alpha, v) \rightarrow (\alpha_0, v_0)} \frac{R(\alpha, v)}{\|(\alpha, v) - (\alpha_0, v_0)\|} = 0$  sofort. Insbesondere ist das Produkt von Zahlen

$$\mu: \mathbb{K} \times \mathbb{K} \ni (\alpha, \beta) \mapsto \alpha\beta \in \mathbb{K} \quad (8.1.19)$$

eine überall differenzierbare Abbildung mit Ableitung

$$D\mu(\alpha_0, \beta_0)(\alpha\beta) = \alpha_0\beta + \beta_0\alpha. \quad (8.1.20)$$

Bevor wir weitere Beispiele konstruieren, benötigen wir noch einige allgemeine Resultate zur Differenzierbarkeit, die aus der eindimensionalen Situation bekannt sind:

**Proposition 8.1.5** *Seien  $V$  und  $W$  normierte Vektorräume mit einer nicht-leeren offenen Teilmenge  $X \subseteq V$  und einer Abbildung  $f: X \rightarrow W$ . Ist  $f$  bei  $x_0 \in X$  differenzierbar, so ist  $f$  bei  $x_0$  stetig.*

BEWEIS: Sei  $f$  differenzierbar bei  $x_0$  mit Ableitung  $Df(x_0)$ . Mit dem Restterm  $R(x)$  aus Bemerkung 8.1.2, *iv.*) folgt dann

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + R(x).$$

Insbesondere gilt  $\lim_{x \rightarrow x_0} R(x) = 0$  nach (8.1.5). Da  $Df(x_0)$  ebenfalls stetig vorausgesetzt ist, folgt auch

$$\lim_{x \rightarrow x_0} Df(x_0)(x - x_0) = Df(x_0) \lim_{x \rightarrow x_0} (x - x_0) = 0,$$

womit insgesamt

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

und damit die Stetigkeit bei  $x_0$  gezeigt ist. □

Man beachte, dass diese Aussage bereits für lineare Abbildungen falsch wäre, hätten wir in der Definition von Differenzierbarkeit auf die Stetigkeit der Ableitung  $Df(x_0)$  verzichtet, siehe nochmals Beispiel 8.1.4, *ii.*)

Als nächstes formulieren und zeigen wir die Kettenregel in nahezu der gleichen Weise wie schon im eindimensionalen Fall. Das Produkt der (skalaren) Ableitungen in Proposition 5.2.5 wird nun durch die Hintereinanderausführung ersetzt:

**Proposition 8.1.6** *Seien  $V, W$  und  $U$  normierte Vektorräume mit nicht-leeren offenen Teilmengen  $X \subseteq V$  und  $Y \subseteq W$  sowie Abbildungen  $f: X \rightarrow W$  und  $g: Y \rightarrow U$ , so dass  $f(X) \subseteq Y$  gilt. Ist dann  $f$  bei  $x_0$  und  $g$  bei  $y_0 = f(x_0)$  differenzierbar, so ist die Verkettung  $g \circ f: X \rightarrow U$  bei  $x_0$  differenzierbar mit der Ableitung*

$$D(g \circ f)(x_0) = Dg(f(x_0)) \circ Df(x_0). \quad (8.1.21)$$

BEWEIS: Wir schreiben

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + R(x)$$

sowie

$$g(y) = g(y_0) + Dg(y_0)(y - y_0) + \tilde{R}(y)$$

mit Resttermen  $R: X \rightarrow W$  und  $\tilde{R}: Y \rightarrow U$ . Dann wissen wir

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{R(x)}{\|x - x_0\|_V} = 0 \quad (*)$$

und

$$\lim_{\substack{y \rightarrow y_0 \\ y \neq y_0}} \frac{\tilde{R}(y)}{\|y - y_0\|_W} = 0. \quad (**)$$

Für die Verkettung erhalten wir dann

$$\begin{aligned} g(f(x)) &= g(f(x_0) + Df(x_0)(x - x_0) + R(x)) \\ &= g(f(x_0)) + Dg(f(x_0))(Df(x_0)(x - x_0) + R(x)) + \tilde{R}(f(x_0) + Df(x_0)(x - x_0) + R(x)) \\ &= g(f(x_0)) + (Dg(f(x_0)) \circ Df(x_0))(x - x_0) + Dg(f(x_0))R(x) + \tilde{R}(f(x)), \end{aligned}$$

wobei wir die Linearität von  $Dg(f(x_0))$  ausgenutzt haben. Die Verkettung  $Dg(f(x_0)) \circ Df(x_0)$  ist nun sowohl linear als auch wieder stetig, womit wir den Kandidaten für die Ableitung von  $g \circ f$  gefunden haben. Der zugehörige Restterm ist also

$$\hat{R}(x) = Dg(f(x_0))R(x) + \tilde{R}(f(x)),$$

für den wir  $\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{\hat{R}(x)}{\|x - x_0\|_V} = 0$  zeigen müssen. Der erste Summand ist dabei leicht zu verstehen, da  $Dg(f(x_0))$  linear und stetig ist, womit

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{Dg(f(x_0))R(x)}{\|x - x_0\|_V} = Dg(f(x_0)) \lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{R(x)}{\|x - x_0\|_V} = 0$$

nach (\*). Für den zweiten Teil betrachten wir zunächst nur solche  $x \in X$  mit  $f(x) \neq f(x_0)$ . Dann gilt

$$\begin{aligned} \frac{\|\tilde{R}(f(x))\|_U}{\|x - x_0\|_V} &= \frac{\|\tilde{R}(f(x))\|_U}{\|f(x) - f(x_0)\|_W} \frac{\|f(x) - f(x_0)\|_W}{\|x - x_0\|_V} \\ &= \frac{\|\tilde{R}(f(x))\|_U}{\|f(x) - f(x_0)\|_W} \frac{\|Df(x_0)(x - x_0) + R(x)\|_W}{\|x - x_0\|_V} \\ &\leq \frac{\|\tilde{R}(f(x))\|_U}{\|f(x) - f(x_0)\|_W} \left( \frac{\|Df(x_0)(x - x_0)\|_W}{\|x - x_0\|_V} + \frac{\|R(x)\|_W}{\|x - x_0\|_V} \right) \\ &\leq \frac{\|\tilde{R}(f(x))\|_U}{\|f(x) - f(x_0)\|_W} \left( \|Df(x_0)\| + \frac{\|R(x)\|_W}{\|x - x_0\|_V} \right) \end{aligned}$$

wobei wir ausnutzen, dass auch  $Df(x_0)$  stetig ist und daher eine endliche Operatornorm besitzt. Für  $x \rightarrow x_0$  konvergiert der zweite Summand gegen Null nach (\*) und der erste Faktor konvergiert ebenfalls gegen Null nach (\*\*), da ja  $f$  stetig bei  $x_0$  ist und daher  $f(x) \rightarrow f(x_0)$ . Insgesamt erhalten wir Konvergenz gegen Null unter der Voraussetzung  $f(x) \neq f(x_0)$ . Da für  $f(x) = f(x_0)$  aber sowieso schon  $\tilde{R}(f(x_0)) = 0$  gilt, stellt die obige Einschränkung keine weitere Schwierigkeit dar und wir haben

$$\lim_{x \rightarrow x_0} \frac{\hat{R}(x)}{\|x - x_0\|_V} = 0$$

gezeigt. □

Bei Abbildungen in einen endlich-dimensionalen Banach-Raum können wir Differenzierbarkeit komponentenweise prüfen. Dies stellt oft eine entscheidende Vereinfachung dar:

**Proposition 8.1.7** *Seien  $V$  und  $W$  normierte Vektorräume mit  $\dim W = m < \infty$ , und sei  $X \subseteq V$  eine nicht-leere offene Teilmenge. Für eine Abbildung  $f: X \rightarrow W$  sind folgende Aussagen äquivalent:*

- i.) *Die Funktion  $f$  ist bei  $x_0 \in X$  differenzierbar.*
- ii.) *Für jedes (stetige) lineare Funktional  $\alpha \in W^*$  ist  $\alpha \circ f: X \rightarrow \mathbb{R}$  bei  $x_0 \in X$  differenzierbar.*
- iii.) *Für jede Basis  $e_1, \dots, e_m \in W$  mit dualer Basis  $e^1, \dots, e^m \in W^*$  sind die Koeffizientenfunktionen  $f^i = e^i \circ f: X \rightarrow \mathbb{R}$  bei  $x_0 \in X$  für  $i = 1, \dots, m$  differenzierbar.*
- iv.) *Es gibt eine Basis  $e_1, \dots, e_m \in W$  mit dualer Basis  $e^1, \dots, e^m \in W^*$ , so dass die Koeffizientenfunktionen  $f^i = e^i \circ f: X \rightarrow \mathbb{R}$  bei  $x_0 \in X$  für  $i = 1, \dots, m$  differenzierbar sind.*

In diesem Fall gilt

$$Df(x_0) = \sum_{i=1}^m Df^i(x_0)e_i. \quad (8.1.22)$$

BEWEIS: Da  $W$  endlich-dimensional ist, ist jedes lineare Funktional automatisch stetig nach Beispiel 7.3.6, v.), und der Tatsache, dass in endlichen Dimensionen alle Normen äquivalent sind. Nach Beispiel 8.1.4, ii.), und der Kettenregel aus Proposition 8.1.6 folgt damit die Implikation i.)  $\implies$  ii.). Die Implikationen ii.)  $\implies$  iii.) sowie iii.)  $\implies$  iv.) sind aus trivialen Gründen wahr, da die Elemente der dualen Basis eben lineare Funktionale sind. Es bleibt iv.)  $\implies$  i.) zu zeigen. Seien also die Koeffizientenfunktionen  $f^i = e^i \circ f: X \rightarrow \mathbb{R}$  bei  $x_0$  differenzierbar mit Ableitungen  $Df^i(x_0) \in L(V, \mathbb{R})$  und entsprechenden Resttermen  $R^i: X \rightarrow \mathbb{R}$  mit

$$f^i(x) = f^i(x_0) + Df^i(x_0)(x - x_0) + R^i(x)$$

und

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{R^i(x)}{\|x - x_0\|_V} = 0. \quad (*)$$

Nach Definition der dualen Basis gilt  $f(x) = \sum_{i=1}^m f^i(x)e_i$  für alle  $x \in X$  und damit

$$f(x) = f(x_0) + \sum_{i=1}^m Df^i(x_0)(x - x_0)e_i + \sum_{i=1}^m R^i(x)e_i.$$

Der Kandidat für die Ableitung von  $f$  ist daher die lineare Abbildung (8.1.22) mit entsprechendem Restterm

$$R(x) = \sum_{i=1}^m R^i(x)e_i.$$

Man beachte, dass für stetige  $Df^i(x_0)$  auch die lineare Abbildung (8.1.22) wieder stetig ist, da  $W$  ja die Produkttopologie von  $\mathbb{R}^m \simeq W$  trägt. Es bleibt also, die geforderte Grenzwerteigenschaft von  $R$

zu zeigen. Da auf  $W$  sowieso alle Normen äquivalent sind, können wir die Maximumsnorm bezüglich der Basis  $e_1, \dots, e_m$  benutzen. Dann gilt

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{\|R(x)\|_W}{\|x - x_0\|_V} = \lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{\max\{|R^1(x)|, \dots, |R^m(x)|\}}{\|x - x_0\|_V} = 0$$

nach (\*). Dies zeigt *iv.*)  $\implies$  *i.*) ebenso wie (8.1.22).  $\square$

Es sei an dieser Stelle aber gewarnt, dass zwar die Implikation *i.*)  $\implies$  *ii.*) auch für beliebige normierte Vektorräume bestehen bleibt, wenn man stetige lineare Funktionale  $\alpha \in L(W, \mathbb{R})$  in *ii.*) verwendet, aber die umgekehrte Implikation *ii.*)  $\implies$  *i.*) typischerweise falsch werden kann, sobald  $\dim W = \infty$ . Trotzdem stellt Proposition 8.1.7 bei endlich-dimensionalem Bild eine wichtige Vereinfachung dar.

Aus den vorangegangenen Überlegungen erhalten wir nun einige Rechenregeln für die Ableitung, die von der eindimensionalen Situation vertraut sind.

**Korollar 8.1.8** *Seien  $V$  und  $W$  normierte Vektorräume mit einer nicht-leeren offenen Teilmenge  $X \subseteq V$ .*

*i.) Sind  $f, g: X \rightarrow W$  differenzierbar bei  $x_0 \in X$ , so ist  $\alpha f + \beta g: X \rightarrow W$  für alle  $\alpha, \beta \in \mathbb{R}$  bei  $x_0$  ebenfalls differenzierbar mit*

$$D(\alpha f + \beta g)(x_0) = \alpha Df(x_0) + \beta Dg(x_0). \quad (8.1.23)$$

*ii.) Sind  $f: X \rightarrow \mathbb{K}$  und  $g: X \rightarrow W$  differenzierbar bei  $x_0 \in X$ , so ist  $fg: X \rightarrow W$  bei  $x_0$  differenzierbar mit*

$$D(fg)(x_0) = g(x_0) Df(x_0) + f(x_0) Dg(x_0). \quad (8.1.24)$$

BEWEIS: Mit  $f$  ist auch  $\alpha f$  differenzierbar bei  $x_0$  mit Ableitung  $\alpha Df(x_0)$ , was unmittelbar aus der Definition folgt. Mit dem selben Argument sieht man, dass die Summe  $f + g$  wieder differenzierbar ist. Alternativ kann man auch mit Beispiel 8.1.4, *iii.*), und der Kettenregel argumentieren. Für den zweiten Teil verwendet man Beispiel 8.1.4, *iv.*), und die Kettenregel. Hierzu beachte man, dass die rechte Seite als lineare Abbildung

$$(g(x_0) Df(x_0) + f(x_0) Dg(x_0))(v) = g(x_0) \cdot Df(x_0)(v) + f(x_0) \cdot Dg(x_0)v$$

aufgefasst wird. Hier ist  $Df(x_0)v \in \mathbb{K}$  eine Zahl, die an den Vektor  $g(x_0)$  multipliziert wird und  $Dg(x_0)$  ist eine lineare Abbildung, die wir auf den Vektor  $v$  anwenden und anschließend mit der Zahl  $f(x_0)$  multiplizieren.  $\square$

Damit ist Differenzieren auch in beliebigen Dimensionen *linear*. Die Rechenregel (8.1.22) verallgemeinert die Leibniz-Regel aus Proposition 5.2.1, *ii.*), und wird entsprechend ebenfalls *Leibniz-Regel* oder *Produktregel* genannt, siehe auch Übung ?? für weitere Varianten der Produktregel.

**Bemerkung 8.1.9 (Kurven und Funktionen)** Als Spezialfälle betrachten wir wieder einen normierten Vektorraum  $V$

*i.) Eine stetige Abbildung*

$$\gamma: I \rightarrow V \quad (8.1.25)$$

auf einem Intervall  $I \subseteq \mathbb{R}$  hatten wir allgemein schon eine Kurve oder einen Weg genannt, siehe erneut Definition 7.6.9, *i.*). Ist nun  $\gamma$  bei  $t_0 \in I$  sogar differenzierbar, so erhalten wir als Ableitung  $D\gamma(t_0) \in L(\mathbb{R}, V)$  einen Vektor

$$\dot{\gamma}(t_0) = D\gamma(t_0)1 \in V \quad (8.1.26)$$

Abbildung 8.1: Der Tangentialvektor an eine Kurve.

Abbildung 8.2: Der Graph von  $f: X \rightarrow \mathbb{R}$  als Höhenprofil mit einigen Koordinatenlinien.

in  $V$ , indem wir die lineare Abbildung  $D\gamma(t_0): \mathbb{R} \rightarrow V$  auf  $1 \in \mathbb{R}$  anwenden. Ausgeschrieben ist dies nach Proposition 8.1.3, i.), gerade

$$\dot{\gamma}(t_0) = \lim_{t \rightarrow 0} \frac{\gamma(t_0 + t) - \gamma(t_0)}{t}. \quad (8.1.27)$$

Geometrisch entspricht  $\dot{\gamma}(t_0)$  also dem *Tangentialvektor* von  $\gamma$  bei  $t_0$ , siehe auch Abbildung 8.1. Alternativ kann man  $\dot{\gamma}(t_0)$  als die *Momentangeschwindigkeit* interpretieren, wenn  $\gamma$  die Bewegung eines Punktes zur Zeit  $t$  beschreibt. Die Notation  $\dot{\gamma}(t_0)$  ist der Newtonschen Tradition entlehnt und wird bevorzugt bei Kurven verwendet. Alternativ schreibt man

$$\gamma'(t_0) = \frac{d\gamma}{dt}(t_0) = \dot{\gamma}(t_0). \quad (8.1.28)$$

ii.) Der zweite Spezialfall von besonderem Interesse sind Abbildungen der Form

$$f: X \rightarrow \mathbb{R}, \quad (8.1.29)$$

wobei  $X \subseteq V$  nicht-leer und offen ist. Solche speziellen Abbildungen werden wir auch *Funktionen* nennen beziehungsweise reell-wertige Funktionen, um den Wertebereich  $\mathbb{R}$  zu betonen. Analog können wir komplex-wertige Funktionen betrachten. Ist nun  $f$  bei  $x_0 \in X$  differenzierbar, so ist die Ableitung  $Df(x_0) \in L(X, \mathbb{R})$  ein Element des Dualraums  $V^* = \text{Hom}(V, \mathbb{R})$ . Da wir aber zudem Stetigkeit gefordert haben, ist  $Df(x_0)$  sogar ein stetiges lineares Funktional. Für Ableitung  $Df(x_0)$  gilt also

$$Df(x_0) \in V'. \quad (8.1.30)$$

Die Ableitung  $Df(x_0)$  wird auch das *Differential* der Funktion  $f$  genannt. Geometrisch kann man sich  $f$  als die Höhenfunktion eines Gebirges vorstellen, siehe Abbildung 8.2. Eine geometrische Interpretation von  $Df(x_0)$  werden wir später noch kennenlernen.

## 8.2 Richtungsableitungen und partielle Ableitungen

Sind  $V$  und  $W$  normierte Vektorräume, und ist  $X \subseteq V$  offen und nicht-leer, so können wir für eine bei  $x_0$  total differenzierbare Abbildung  $f: X \rightarrow W$  die Verkettung

$$\mathbb{R} \supseteq I \ni t \mapsto f(x_0 + tv) \in W \quad (8.2.1)$$

betrachten, wobei  $v \in V$  ein beliebiger Vektor ist und  $I \subseteq \mathbb{R}$  ein offenes Intervall um  $0 \in I$ , welches wir klein genug wählen, so dass  $x_0 + tv \in X$  für alle  $t \in I$ . Da  $X$  offen ist, ist eine solche Wahl immer möglich, siehe auch Abbildung 8.3. Da die Gerade  $\gamma_v: \mathbb{R} \rightarrow V$  mit

$$\gamma_v(t) = x_0 + tv \quad (8.2.2)$$

Abbildung 8.3: Die Abbildung  $f$  entlang einer Geraden  $\gamma_v$ .

an jeder Stelle aus  $I$  differenzierbar ist, ist die Verkettung  $f \circ \gamma_v$  bei  $t = 0$  nach der Kettenregel aus Proposition 8.1.6 differenzierbar. Als Ableitung nach  $t$  erhalten wir

$$\frac{d}{dt}(f \circ \gamma_v)(0) = Df(x_0) \cdot \frac{d\gamma}{dt}(0) = Df(x_0)v, \quad (8.2.3)$$

da  $\dot{\gamma}_v(t) = v$  für alle  $t \in \mathbb{R}$  gilt, womit wir eine neue geometrische Interpretation des Resultats aus Proposition 8.1.3, i.) erhalten.

**Definition 8.2.1 (Richtungsableitung)** Seien  $V$  und  $W$  normierte Räume mit einer nicht-leeren offenen Teilmenge  $X \subseteq V$ . Eine Abbildung  $f: X \rightarrow W$  heißt bei  $x_0 \in X$  in Richtung  $v \in V \setminus \{0\}$  differenzierbar (oder auch Gateaux-differenzierbar), wenn der Grenzwert

$$D_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} \quad (8.2.4)$$

existiert. In diesem Fall heißt  $D_v(f)(x_0)$  die Richtungsableitung oder auch die Gateaux-Ableitung.

Die Voraussetzung  $v \neq 0$  ist zweckmäßig, um Trivialitäten zu vermeiden.

**Proposition 8.2.2** Seien  $V$  und  $W$  normierte Vektorräume und  $X \subseteq V$  eine nicht-leere offene Teilmenge. Ist eine Abbildung  $f: X \rightarrow W$  bei  $x_0 \in X$  total differenzierbar, so existieren alle Richtungsableitungen von  $f$  bei  $x_0$ . Genauer gilt

$$D_v f(x_0) = Df(x_0)v \quad (8.2.5)$$

für alle  $v \in V \setminus \{0\}$ .

Dies ist gerade die Aussage von Proposition 8.1.3, i.), im Gewande der neuen Begriffe. Überraschenderweise ist die Existenz aller Richtungsableitungen an einem Punkt  $x_0 \in X$  bereits ab Dimension zwei nicht länger äquivalent zur totalen Differenzierbarkeit, wie folgendes einfache Beispiel zeigt:

**Beispiel 8.2.3** Wir betrachten die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } x = 0 = y. \end{cases} \quad (8.2.6)$$

Sei nun  $v = (v_1, v_2) \in \mathbb{R}^2 \setminus \{0\}$  eine Richtung, dann gilt für  $t \neq 0$

$$\frac{f(0 + tv) - f(0)}{t} = \frac{1}{t} \frac{t^3 v_1^2 v_2}{t^2(v_1^2 + v_2^2)} = \frac{v_1^2 v_2}{v_1^2 + v_2^2} = f(v),$$

womit die Existenz der Richtungsableitung

$$D_v f(0) = \lim_{t \rightarrow 0} \frac{f(0 + tv) - f(0)}{t} = f(v) \quad (8.2.7)$$

sofort folgt. Allerdings ist  $f$  bei 0 nicht total differenzierbar. Wäre dies der Fall, gäbe es eine lineare Abbildung  $Df(0) \in L(\mathbb{R}^2, \mathbb{R})$ , welche nach Proposition 8.2.2

$$Df(0)v = (D_v f)(0) = f(v) \quad (8.2.8)$$

erfüllte. Daher wäre  $Df(0) = f$ . Nun ist aber die Abbildung  $f$  alles andere als linear, womit  $f$  bei 0 nicht total differenzierbar sein kann.



Abbildung 8.4: Die Funktion  $f$  ist bei 0 unstetig, hat aber alle Richtungsableitungen.

Es ist tatsächlich so, dass die Existenz aller Richtungsableitungen bei einem Punkt  $x_0$  noch nicht einmal die Stetigkeit bei  $x_0$  impliziert. Dies illustriert folgendes Beispiel:

**Beispiel 8.2.4** Wir betrachten  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$f(x, y) = \begin{cases} 0 & \text{für } y \leq 0 \text{ oder } y \geq x^2 \\ 1 & \text{für } 0 < y < x^2, \end{cases} \quad (8.2.9)$$

siehe auch Abbildung 8.4. Da es Punkte  $x_n \in \mathbb{R}^2$  mit  $x_n \rightarrow 0$  sowohl im Gebiet mit  $y \leq 0$  als auch im Gebiet mit  $0 < y < x^2$  gibt, sieht man, dass  $f$  bei 0 nicht stetig ist. Trotzdem existieren alle Richtungsableitungen: Ist  $v = (v_1, v_2) \in \mathbb{R}^2 \setminus \{0\}$  mit  $v_2 \leq 0$ , so ist  $f(tv) = 0$  für alle  $t$ , womit  $D_v f(0) = 0$  existiert. Ist nun  $v_2 > 0$ , so gibt es ein  $t_0 > 0$  derart, dass der Vektor  $tv$  wieder im Bereich liegt, wo  $f$  den Wert  $f(tv) = 0$  für alle  $t$  mit  $|t| < t_0$  hat. Daher existiert auch in diesem Fall der Grenzwert  $D_v f(0)$  und liefert den Wert 0. Es gilt daher

$$D_v f(0) = 0 \quad (8.2.10)$$

für alle  $v \in \mathbb{R}^2 \setminus \{0\}$ .

Diese zunächst etwas enttäuschenden Beispiele zeigen, dass Gateaux-Differenzierbarkeit ein strikt schwächeres Konzept als totale Differenzierbarkeit ist, ja noch nicht einmal Stetigkeit impliziert. Trotzdem ist gerade in unendlichen Dimensionen die Gateaux-Differenzierbarkeit in vielen Situationen zum einen viel einfacher in der Handhabung, zum anderen aber trotzdem nützlich genug, um bei vielen Problemstellungen wie beispielsweise in der Variationsrechnung zu Lösungen zu führen.

In endlichen Dimensionen ist die Verwendung von (linearen) Koordinaten oftmals die Methode der Wahl. Dies wollen wir nun auch für die Richtungsableitungen tun. Zuvor verabreden wir aber folgende Notation.

**Konvention 8.2.5 (Einsteinsche Summenkonvention)** Ist  $V$  endlich-dimensional, so wählen wir eine Basis  $e_1, \dots, e_n \in V$ , im Falle von  $V = \mathbb{R}^n$  typischerweise die kanonische Basis. Der Dualraum  $V^* = \text{Hom}(V, \mathbb{R})$  erbt dann die duale Basis, die wir mit  $e^1, \dots, e^n \in V^*$  bezeichnen, wobei die duale Basis durch

$$e^i(e_j) = \delta_j^i \quad (8.2.11)$$

für  $i, j = 1, \dots, n$  festgelegt ist. Die Koordinatendarstellung eines Vektors ist damit

$$v = \sum_{i=1}^n e^i(v) e_i = \sum_{i=1}^n v^i e_i = v^i e_i \quad (8.2.12)$$

mit  $v^i = e^i(v) \in \mathbb{R}$ . Hier ist zum einen auf die Stellung der Indizes zu achten. Zum anderen verabreden wir, die Summenzeichen wegzulassen, wenn über alle Koordinaten summiert wird. Da dies in *Paaren* von Indizes eigentlich immer gewünscht ist, stellt diese *Einsteinsche Summenkonvention* eine deutliche Entlastung dar, ohne zu Missverständnissen zu führen. Obere Indizes entsprechen Indizes der dualen Basis, untere dagegen Indizes der Basis von  $V$ . Summiert wird also, wenn beide zusammentreffen, siehe auch [9, Abschnitt 3.6] für eine weiterführende Diskussion und insbesondere die Rechtfertigung dieser Unterscheidung anhand des unterschiedlichen Transformationsverhaltens unter Basiswechsel.

Richtungsableitungen in  $\mathbb{R}^n$  in Richtung der kanonischen Basisvektoren verdienen aus naheliegenden Gründen nun besondere Aufmerksamkeit:

Abbildung 8.5: Die Funktion  $f$  ist Null auf den Koordinatenachsen und 1 sonst.

**Definition 8.2.6 (Partielle Ableitung)** Sei  $X \subseteq \mathbb{R}^n$  eine nicht-leere offene Teilmenge und sei  $W$  ein normierter Vektorraum. Eine Funktion  $f: X \rightarrow W$  heißt bei  $x_0 \in X$  partiell differenzierbar in  $i$ -te Richtung, falls die Richtungsableitung  $D_{e_i}f(x_0)$  existiert. In diesem Fall schreibt man

$$\partial_i f(x_0) = \frac{\partial f}{\partial x^i}(x_0) = D_{e_i}f(x_0) \quad (8.2.13)$$

und nennt (8.2.13) auch die  $i$ -te partielle Ableitung von  $f$  bei  $x_0$  für  $i = 1, \dots, n$ .

Existieren alle Richtungsableitungen, dann erst recht alle partiellen Ableitungen, da diese ja spezielle Richtungsableitungen sind. Wenig überraschend ist, dass die Umkehrung im Allgemeinen nicht zu gelten braucht:

**Beispiel 8.2.7** Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$f(x, y) = \begin{cases} 0 & \text{für } x = 0 \text{ oder } y = 0 \\ 1 & \text{sonst,} \end{cases} \quad (8.2.14)$$

siehe auch Abbildung 8.5. Dann existieren die partiellen Ableitungen

$$\frac{\partial f}{\partial x}(0) = 0 = \frac{\partial f}{\partial y}(0) \quad (8.2.15)$$

da

$$\frac{f(te_1) - f(0)}{t} = 0 = \frac{f(te_2) - f(0)}{t} \quad (8.2.16)$$

für alle  $t \neq 0$ . Der Grenzwert  $t \rightarrow 0$  ist daher einfach zu bestimmen. Ist umgekehrt  $v \in \mathbb{R}^2 \setminus \{0\}$  eine Richtung ungleich der Koordinatenrichtung, so gilt

$$\frac{f(tv) - f(0)}{t} = \frac{1}{t} \quad (8.2.17)$$

für alle  $t \neq 0$ . Dies zeigt, dass alle anderen Richtungsableitungen nicht existieren. Insbesondere ist  $f$  bei 0 nicht stetig.

Richtungsableitungen können wir nun dazu verwenden, auch für Funktionen mehrerer Variablen Minima und Maxima zu finden. Hierzu betrachtet man folgende Situation einer skalaren Funktion

$$f: V \supseteq X \rightarrow \mathbb{R}, \quad (8.2.18)$$

wobei wie zuvor  $V$  ein normierter Vektorraum sei und  $X \subseteq V$  offen und nicht-leer ist. Ist  $x_0 \in X$  ein (lokales) Extremum von  $f$ , und ist

$$\gamma: I \rightarrow X \quad (8.2.19)$$

eine beliebige Kurve durch  $x_0$ , wobei  $I \subseteq \mathbb{R}$  ein offenes Intervall mit  $0 \in I$  und  $\gamma(0) = x_0$  sei, so ist für die Verkettung

$$f \circ \gamma: I \rightarrow \mathbb{R} \quad (8.2.20)$$

der Punkt  $0 \in I$  erst recht ein lokales Extremum, siehe Abbildung 8.6. Die Überlegung führt sofort zu folgendem notwendigen Kriterium für die Existenz von Extrema:

**Proposition 8.2.8** Sei  $V$  ein normierter Vektorraum mit einer nicht-leeren offenen Teilmenge  $X \subseteq V$  und einer Funktion  $f: X \rightarrow \mathbb{R}$ . Sei weiter  $x_0 \in X$  ein lokales Extremum von  $f$ .

Abbildung 8.6: Das lokale Extremum von  $f$  ist erst recht ein lokales Maximum längs der Kurve  $\gamma$ .

i.) Ist  $f$  bei  $x_0$  differenzierbar, so gilt

$$Df(x_0) = 0. \quad (8.2.21)$$

ii.) Ist  $f$  bei  $x_0$  in Richtung  $v \in V \setminus \{0\}$  differenzierbar, so gilt

$$D_v f(x_0) = 0. \quad (8.2.22)$$

iii.) Ist  $V = \mathbb{R}^n$  und  $f$  bei  $x_0$  partiell in  $i$ -te Richtung differenzierbar, so gilt

$$\frac{\partial f}{\partial x^i}(x_0) = 0. \quad (8.2.23)$$

BEWEIS: Teil iii.) ist ein Spezialfall von ii.). Ist  $f$  total differenzierbar bei  $x_0$ , so ist  $f$  auch in jede Richtung differenzierbar mit  $Df(x_0)v = D_v f(x_0)$ . Gilt also ii.) so folgt i.), da  $Df(x_0)$  durch die Richtungsableitungen bestimmt ist. Wir müssen also ii.) zeigen. Dazu betrachten wir die Kurve

$$\gamma(t) = x_0 + tv$$

für kleine  $t \in I \subseteq \mathbb{R}$  derart, dass  $\gamma(t) \in X$ . Dann ist  $0 \in I$  ein Extremum der skalaren Funktion  $f \circ \gamma: I \rightarrow \mathbb{R}$  einer Variable, womit wir Proposition 5.3.4 zum Einsatz bringen können. Es gilt also

$$0 = \left. \frac{d}{dt} \right|_{t=0} (f \circ \gamma)(t) = D_v f(x_0)$$

nach der Kettenregel beziehungsweise der Definition der Richtungsableitung.  $\square$

Es gibt allerdings auch hier neue Effekte, womit das Verschwinden der Richtungsableitungen in jede Richtung noch nicht hinreichend für die Existenz eines lokalen Extremums ist. Beispiele kennen wir zum einen aus der eindimensionalen Theorie, zum anderen werden neue Phänomene wie etwa Sattelpunkte in Übung ?? diskutiert.

In endlichen Dimensionen mit  $X \subseteq \mathbb{R}^n$  und  $W = \mathbb{R}^m$  können wir zum einen die Komponentenfunktionen  $f^j = \tilde{e}^j \circ f: X \rightarrow \mathbb{R}$  der Abbildung  $f$  bezüglich der Standardbasis  $\tilde{e}_1, \dots, \tilde{e}_m$  von  $\mathbb{R}^m$  betrachten. Zum anderen ist bei einer differenzierbaren Abbildung die Ableitung als lineare Abbildung durch die Werte auf einer Basis bestimmt. Wir wissen daher

$$Df(x_0)e_i = D_{e_i} f(x_0) = \frac{\partial f}{\partial x^i}(x_0) \quad (8.2.24)$$

und

$$\frac{\partial f}{\partial x^i}(x_0) = \tilde{e}^j \left( \frac{\partial f}{\partial x^i}(x_0) \right) \tilde{e}_j = \frac{\partial f^j}{\partial x^i}(x_0) \tilde{e}_j \quad (8.2.25)$$

da die linearen Funktionale  $\tilde{e}^j$  ja differenzierbar sind mit Ableitung wieder gleich  $\tilde{e}^j$  und deshalb die Kettenregel zum Einsatz kommen kann. Damit ist  $Df(x_0)$  also durch die partiellen Ableitungen  $\frac{\partial f^j}{\partial x^i}(x_0)$  der Komponentenfunktionen bestimmt. Es folgt

$$Df(x_0) = \frac{\partial f^j}{\partial x^i}(x_0) \tilde{e}_j \otimes e^i, \quad (8.2.26)$$

was gerade die Basisdarstellung von  $Df(x_0)$  bezüglich der Standardbasis  $e_1, \dots, e_n$  von  $\mathbb{R}^n$  beziehungsweise  $\tilde{e}_1, \dots, \tilde{e}_m$  von  $\mathbb{R}^m$  ist. Hier ist

$$\tilde{e}_j \otimes e^i: \mathbb{R}^n \rightarrow \mathbb{R}^m \quad (8.2.27)$$

diejenige eindeutig bestimmte lineare Abbildung mit

$$(\tilde{e}_j \otimes e^i)(e_k) = \tilde{e}_j \delta_k^i, \quad (8.2.28)$$

siehe auch [9, Abschnitt 3.6] für eine detaillierte Darstellung dieser Tensorschreibweise. Die Matrixdarstellung der Ableitung verdient nun einen eigenen Namen:

**Definition 8.2.9 (Jacobi-Matrix und Jacobi-Determinante)** Sei  $X \subseteq \mathbb{R}^n$  eine nicht-leere offene Teilmenge mit einer bei  $x_0 \in X$  differenzierbaren Abbildung  $f: X \rightarrow \mathbb{R}^m$ . Die Matrixdarstellung der Ableitung  $Df(x_0)$  bezüglich der Standardbasen heißt dann die *Jacobi-Matrix*  $J_f(x_0) \in M_{m \times n}(\mathbb{R})$  von  $f$  bei  $x_0$ . Gilt zudem  $n = m$ , so heißt die Determinante von  $Df(x_0)$  die *Jacobi-Determinante* von  $f$  bei  $x_0$ .

**Bemerkung 8.2.10** Die Jacobi-Matrix  $J_f(x_0)$  ist also ausgeschrieben

$$J_f(x_0) = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x_0) & \cdots & \frac{\partial f^1}{\partial x^n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x^1}(x_0) & \cdots & \frac{\partial f^m}{\partial x^n}(x_0) \end{pmatrix} \quad (8.2.29)$$

und die Jacobi-Determinante entsprechend

$$\det(Df(x_0)) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \frac{\partial f^1}{\partial x^{\sigma(1)}}(x_0) \cdots \frac{\partial f^n}{\partial x^{\sigma(n)}}(x_0) \quad (8.2.30)$$

im Falle einer quadratischen Jacobi-Matrix, also für  $n = m$ . Man beachte, dass  $\det(Df(x_0)) \in \mathbb{R}$  unabhängig von der Wahl der Basis ist. Wir hätten anstelle der Standardbasis von  $\mathbb{R}^n$  auch jede andere Basis verwenden können. Die Jacobi-Matrix nimmt dagegen expliziten Bezug auf die Standardbasen von  $\mathbb{R}^n$  und  $\mathbb{R}^m$ .

### 8.3 Stetig differenzierbare Funktionen

Die Differenzierbarkeit einer Funktion an einer Stelle ist typischerweise wenig hilfreich. Ist dagegen an allen Punkten des Definitionsbereich Differenzierbarkeit gewährleistet, so lassen sich wie schon im eindimensionalen Fall sehr viel interessantere Schlussfolgerungen treffen.

Als erstes zeigen wir folgende Version des Mittelwertsatzes oder Schrankensatzes: Für vektorwertige Funktionen kann der Mittelwertsatz nur noch in Form einer Ungleichung wie schon beim eindimensionalen Schrankensatz, siehe Satz 5.3.10, gezeigt werden.

**Satz 8.3.1 (Schrankensatz II)** Sei  $X \subseteq V$  eine nicht-leere offene Teilmenge eines normierten Vektorraums. Sei weiter  $f: X \rightarrow W$  eine differenzierbare Abbildung mit Werten in einem weiteren normierten Vektorraum  $W$ . Ist  $K \subseteq X$  kompakt und konvex, so gilt für alle  $x, x' \in K$  die Abschätzung

$$\|f(x) - f(x')\|_W \leq \sup_{\xi \in K} \|Df(\xi)\| \|x - x'\|_V. \quad (8.3.1)$$

BEWEIS: Hier bezeichnet  $\|Df(\xi)\|$  wie immer die Operatornorm von  $Df(\xi) \in L(V, W)$ . Konvexität ist wie im eindimensionalen Fall definiert und benutzt die Vektorraumstruktur von  $V$ . Wir betrachten die Strecke

$$\gamma(t) = tx' + (1-t)x$$

von  $x$  nach  $x'$ , welche für  $t \in [0, 1]$  ganz in  $K$  liegt. Da  $K \subseteq X$  kompakt ist, finden wir ein  $\epsilon > 0$  derart, dass  $\gamma(t) \in X$  für  $t \in (-\epsilon, 1 + \epsilon) = I$  immer noch in  $X$  liegt. Damit ist

$$\gamma: I \rightarrow V$$

eine differenzierbare Kurve auf dem offenen Definitionsbereich  $I$  mit  $\gamma([0, 1]) \subseteq K$ . Nach der Kettenregel können wir die Ableitung der Verkettung  $g = f \circ \gamma$  als

$$D(f \circ \gamma)(t) = Df(\gamma(t))\dot{\gamma}(t)$$

schreiben. Es gilt  $\dot{\gamma}(t) = x' - x$  für alle  $t \in I$  und daher

$$\|Dg(t)\| = \|D(f \circ \gamma)\| \leq \|Df(\gamma(t))\| \|x - x'\|_V.$$

Schließlich ist  $g(0) = f(x)$  und  $g(1) = f(x')$ , womit wir also

$$\|g(1) - g(0)\| \leq \|Dg(t)\| \quad (*)$$

für alle  $t \in [0, 1]$  zeigen wollen, um (8.3.1) zu schließen. Sei daher

$$\sup_{t \in [0, 1]} \|Dg(t)\| < \infty$$

um Trivialitäten zu vermeiden. Wir wollen nun für jedes  $\alpha > \sup_{t \in [0, 1]} \|Dg(t)\|$  zeigen, dass

$$\|g(1) - g(0)\|_W \leq \alpha \quad (\odot)$$

gilt, was dann  $(*)$  zeigt und somit auch (8.3.1). Dazu betrachten wir

$$\tau = \sup\{t \in [0, 1] \mid \|g(t) - g(0)\|_W \leq t\alpha\}.$$

Da offenbar  $\|g(0) - g(0)\|_W = 0 \leq 0 \cdot \alpha$ , folgt, dass die Menge aller solchen  $t$  nicht leer ist. Nach Definition ist die Menge zudem beschränkt, womit das Supremum  $\tau \in [0, 1]$  existiert. Wir nehmen an, dass  $\tau < 1$  gilt und wollen dies zum Widerspruch führen. Die Differenzierbarkeit von  $g$  bei  $\tau$  liefert einen Restterm  $R: I \rightarrow W$  mit

$$g(t) = g(\tau) + Dg(\tau)(t - \tau) + R(t)$$

und

$$\lim_{\substack{t \rightarrow \tau \\ t \neq \tau}} \frac{R(t)}{t - \tau} = 0. \quad (**)$$

Daher gilt also für  $t > \tau$

$$\frac{\|g(t) - g(\tau)\|_W}{t - \tau} = \frac{\|Dg(\tau)(t - \tau) + R(t)\|_W}{t - \tau} \leq \|Dg(\tau)\| + \frac{\|R(t)\|_W}{t - \tau}.$$

Mit  $(**)$  sehen wir, dass es ein  $\delta > 0$  gibt, so dass für  $t - \tau < \delta$  nach Annahme zu  $\alpha$

$$\frac{\|g(t) - g(\tau)\|_W}{t - \tau} < \alpha$$

gilt. Es folgt also  $\|g(t) - g(\tau)\|_W < \alpha(t - \tau)$  und damit

$$\begin{aligned} \|g(t) - g(0)\|_W &\leq \|g(t) - g(\tau)\|_W + \|g(\tau) - g(0)\|_W \\ &\leq \alpha(t - \tau) + \alpha\tau \\ &= t\alpha \end{aligned}$$

für alle  $t > \tau$  mit  $t - \tau < \delta$ . Nach Definition von  $\tau$  ist dies aber ein Widerspruch, was also  $(\odot)$  und damit die Behauptung (8.3.1) zeigt.  $\square$

**Bemerkung 8.3.2** Der Schrankensatz hat viele Spezialfälle und Anwendungen. Eine interessante Situation erhält man für eine differenzierbare Kurve  $\gamma: I \rightarrow V$ , wobei  $I \subseteq \mathbb{R}$  ein offenes Intervall ist. Sind dann  $a, b \in I$  mit  $a < b$ , so besagt der Schrankensatz

$$\|\gamma(b) - \gamma(a)\| \leq \sup_{t \in [a, b]} \|\dot{\gamma}(t)\| (b - a). \quad (8.3.2)$$

Geometrisch bedeutet dies, dass bei vorhandener Schranke

$$5 = \sup_{t \in [a, b]} \|\dot{\gamma}(t)\| < \infty \quad (8.3.3)$$

an die Geschwindigkeit in einer gewissen Zeit  $t = b - a$  höchstens eine Strecke der Länge

$$\|\gamma(b) - \gamma(a)\| \leq v \cdot t \quad (8.3.4)$$

zurückgelegt werden kann.

Einen alternativen Beweis des Schrankensatzes erhält man in endlichen Dimensionen mit Hilfe des kanonischen Skalarprodukts, siehe Übung ???. Eine gänzlich andere Argumentation beruht auf der Verwendung des eindimensionalen Schrankensatzes und (einer Version) des Satzes von Hahn-Banach, siehe Übung ??, den wir hier nicht zur Verfügung haben.

Ist  $f: X \rightarrow W$  an allen Punkten von  $X$  differenzierbar, so erhalten wir eine neue Abbildung

$$Df: X \rightarrow L(V, W) \quad (8.3.5)$$

mit Werten in den stetigen linearen Abbildungen  $L(V, W)$  anstelle von  $W$ . Da für normierte Vektorräume  $V$  und  $W$  die stetigen linearen Abbildungen wieder ein durch die Operatornorm normierter Vektorraum sind, können wir erneut nach Stetigkeit von  $Df$  fragen. Dies führt auf folgende Definition:

**Definition 8.3.3 ( $\mathcal{C}^1$ -Abbildung)** Seien  $V$  und  $W$  normierte Vektorräume mit einer nicht-leeren offenen Teilmenge  $X \subseteq V$ . Eine Abbildung  $f: X \rightarrow W$  heißt stetig differenzierbar oder  $\mathcal{C}^1$ , wenn  $f$  an jeder Stelle von  $X$  differenzierbar ist und die Abbildung  $Df$  wieder eine stetige Abbildung

$$Df \in \mathcal{C}(X, L(V, W)) \quad (8.3.6)$$

ist. Die Menge der  $\mathcal{C}^1$ -Abbildungen von  $X$  nach  $W$  wird mit

$$\mathcal{C}^1(X, W) = \{f: X \rightarrow W \mid f \text{ ist } \mathcal{C}^1\} \quad (8.3.7)$$

bezeichnet.

Da eine bei  $x_0 \in X$  differenzierbare Funktion automatisch stetig bei  $x_0$  ist, folgt also insbesondere

$$\mathcal{C}^1(X, W) \subseteq \mathcal{C}^0(X, W) = \mathcal{C}(X, W), \quad (8.3.8)$$

wobei wir wieder  $\mathcal{C}^0$  für die stetigen Abbildungen schreiben, wie wir das im eindimensionalen Fall bereits getan haben.

Es gilt also zunächst die Natur von  $L(V, W)$  etwas genauer zu beleuchten. Im Allgemeinen hängt die Operatornorm von  $L(V, W)$  sehr intensiv von den auf  $V$  und  $W$  verwendeten Normen ab. ist jedoch  $V$  endlich-dimensional, so erhalten wir folgende äquivalente Beschreibungen der Topologie von  $L(V, W)$ .

**Proposition 8.3.4** Seien  $V$  und  $W$  normierte Vektorräume.

- i.) Sind  $\|\cdot\|'_V$  und  $\|\cdot\|'_W$  zwei äquivalente Normen auf  $V$  beziehungsweise  $W$ , so ist eine lineare Abbildung  $A: V \rightarrow W$  genau dann stetig, wenn  $A$  bezüglich  $\|\cdot\|'_V$  und  $\|\cdot\|'_W$  stetig ist. Die Operatornorm  $\|\cdot\|'$  bezüglich  $\|\cdot\|'_V$  und  $\|\cdot\|'_W$  ist äquivalent zu ursprünglichen Operatornorm.
- ii.) Ist  $V$  endlich-dimensional, so gilt

$$L(V, W) = \text{Hom}(V, W), \quad (8.3.9)$$

und die Normen

$$\|A\|_\infty = \max\{\|Ae_i\|_W, \dots, \|Ae_n\|_W\} \quad (8.3.10)$$

und

$$\|A\|_p = \sqrt[p]{\sum_{i=1}^n \|Ae_i\|_W^p} \quad (8.3.11)$$

für  $p \in [1, \infty)$  und eine beliebige Basis  $e_1, \dots, e_n \in V$  sind äquivalent zur Operatornorm.

BEWEIS: Für den ersten Teil haben wir Konstanten  $c, c', d, d' > 0$  mit

$$\|v\|_V \leq c\|v\|'_V \quad \text{und} \quad \|v\|'_V \leq c'\|v\|_V$$

sowie

$$\|w\|_W \leq d\|w\|'_W \quad \text{und} \quad \|w\|'_W \leq d'\|w\|_W$$

für alle  $v \in V$  und  $w \in W$ . Sei nun  $A: V \rightarrow W$  linear und stetig bezüglich  $\|\cdot\|_V$  und  $\|\cdot\|_W$ . Dann gilt also

$$\|Av\|_W \leq \|A\|\|v\|_V$$

für alle  $v \in V$  nach Definition der Operatornorm und Satz 7.3.7. Für die alternative Wahl der Normen erhalten wir dann

$$\|Av\|'_W \leq d'\|Av\|_W \leq d'\|A\|\|v\|_V \leq d'c\|A\|\|v\|'_V,$$

womit nach Satz 7.3.7 die Abbildung  $A$  auch stetig bezüglich der äquivalenten Normen ist. Zudem sehen wir

$$\|A\|' \leq d'c\|A\|.$$

Durch Vertauschen der Rollen erhält man die Umkehrung sowie die Abschätzung

$$\|A\| \leq dc'\|A\|',$$

womit der erste Teil gezeigt ist. Für den zweiten Teil erinnern wir zunächst daran, dass auf  $V$  alle Normen nach Satz 7.7.8 äquivalent sind. Nach dem ersten Teil ist daher  $L(V, W)$  unabhängig von der Wahl der Norm auf  $V$ . Wir verwenden daher beispielsweise die Maximumsnorm  $\|\cdot\|_\infty$  bezüglich einer Basis. Für eine lineare Abbildung  $A: V \rightarrow W$  gilt dann

$$\begin{aligned} \|Av\|_W &= \left\| \sum_{i=1}^n (Ae_i)v^i \right\|_W \\ &\leq \sum_{i=1}^n |v^i| \|Ae_i\|_W \\ &\leq \sum_{i=1}^n \|Ae_i\|_W \max\{|v^1|, \dots, |v^n|\} \\ &= \|A\|_1 \|v\|_\infty, \end{aligned} \quad (*)$$

womit  $A$  stetig bezüglich  $\|\cdot\|_\infty$  und zugehöriger Operatornorm  $\|A\|_1$  ist. Man überlegt sich schnell, dass  $\|A\|_1$  wirklich die beste Konstante in (\*) und damit die Operatornorm ist. Dies zeigt (8.3.9), da die Wahl von  $\|\cdot\|_\infty$  nach *i.*) unerheblich ist. Wir zeigen nun die Äquivalenz der Normen aus (8.3.10) sowie (8.3.11). Für  $p, q \in [1, \infty)$  mit  $p < q$  gilt nach Korollar 5.4.13

$$\|A\|_q = \sqrt[q]{\sum_{i=1}^n \|Ae_i\|_W^q} \leq \sqrt[p]{\sum_{i=1}^n \|Ae_i\|_W^p} = \|A\|_p,$$

also  $\|A\|_q \leq \|A\|_p$ . Weiter gilt

$$\|A\|_\infty = \max\{\|Ae_1\|_W, \dots, \|Ae_n\|_W\} \leq \sqrt[q]{\sum_{i=1}^n \|Ae_i\|_W^q} = \|A\|_q,$$

und somit  $\|A\|_\infty \leq \|A\|_q$ . Schließlich erhalten wir

$$\|A\|_1 = \sum_{i=1}^n \|Ae_i\|_W \leq n \cdot \max\{\|Ae_1\|_W, \dots, \|Ae_n\|_W\} = n \cdot \|A\|_\infty,$$

was insgesamt die Kette von Ungleichungen

$$\|A\|_\infty \leq \|A\|_q \leq \|A\|_p \leq \|A\|_1 \leq n\|A\|_\infty$$

für alle  $p, q \in [0, \infty)$  mit  $q > p$  zeigt. Damit folgt aber die behauptete Äquivalenz.  $\square$

Man beachte, dass die Dimension  $n = \dim V$  in der Abschätzung der Normen

$$\|A\|_1 \leq n\|A\|_\infty \tag{8.3.12}$$

entscheidend eingeht, weshalb eine ähnliche Ungleichung in unendlichen Dimensionen auf keinen Fall zu erwarten ist.

Nach diesen Vorüberlegungen können wir nun das zentrale Resultat dieses Abschnitts formulieren. Die totale Differenzierbarkeit ist typischerweise die interessantere, weil im Allgemeinen stärkere, Eigenschaft im Vergleich zur Existenz von Richtungsableitungen oder partiellen Ableitungen. Letztere sind andererseits in vielen Situationen deutlich leichter zu bestimmen. Aus diesem Grunde wird folgende Charakterisierung von stetig differenzierbaren Funktionen ungemein wichtig:

**Satz 8.3.5 (Stetige Differenzierbarkeit)** *Sei  $X \subseteq \mathbb{R}^n$  eine nicht-leere offene Teilmenge mit einer Abbildung  $f: X \rightarrow W$  in einen normierten Vektorraum  $W$ . Dann sind folgende Aussagen äquivalent:*

- i.) Es gilt  $f \in \mathcal{C}^1(X, W)$ .*
- ii.) Alle Richtungsableitungen von  $f$  existieren und sind stetig.*
- iii.) Alle partiellen Ableitungen von  $f$  existieren und sind stetig.*

BEWEIS: Wir zeigen  $i.) \implies ii.) \implies iii.) \implies i.)$ . Sei also zuerst  $f$  stetig differenzierbar mit Ableitung  $Df: X \rightarrow L(\mathbb{R}^n, W)$ . Ist  $v \in \mathbb{R}^n \setminus \{0\}$  eine Richtung, so gilt für  $x \in X$

$$D_v f(x) = Df(x)v$$

nach Proposition 8.2.2. Insbesondere existieren die Richtungsableitungen an jedem Punkt. Weiter ist das Auswerten auf einem festen Vektor

$$\text{ev}_v: L(V, W) \ni A \mapsto \text{ev}_v(A) = Av \in W$$



Abbildung 8.7: Die Hilfspunkte  $x_1, x_2, x_3$  mit den zugehörigen, dazwischenliegenden Verbindungsstrecken.

eine stetige (lineare) Abbildung bezüglich der Operatornorm auf  $L(V, W)$ , da

$$\|\mathrm{ev}_v(A)\| = \|Av\|_W \leq \|A\| \|v\|_V.$$

Damit ist aber die Verkettung

$$X \ni x \mapsto D_v f(x) = (\mathrm{ev}_v \circ Df)(x) \in W$$

eine stetige Abbildung, was *i.)*  $\implies$  *ii.)* zeigt. Da partielle Ableitungen selbst besondere Richtungsableitungen sind, ist *ii.)*  $\implies$  *iii.)* trivial. Es bleibt *iii.)*  $\implies$  *i.)* zu zeigen. Für  $x \in X$  und  $v \in \mathbb{R}^n$  definieren wir die lineare Abbildung  $A(x): \mathbb{R}^n \rightarrow W$  durch

$$A(x)v = \sum_{i=1}^n \frac{\partial f}{\partial x^i} v^i$$

wobei wie immer  $v^i = e^i(v)$  die linearen Koordinaten von  $v$  bezüglich der Standardbasis sind. Nach Proposition 8.3.4 ist  $A(x) \in \mathrm{Hom}(\mathbb{R}^n, W) = L(V, W)$  automatisch stetig und damit ein guter Kandidat für die Ableitung von  $f$ . Für die Norm (8.3.10) erhalten wir weiter

$$\|A(x) - A(x')\|_\infty = \max_{i=1}^n \{ \|(A(x) - A(x'))e_i\|_W \} = \max_{i=1}^n \left\{ \left\| \frac{\partial f}{\partial x^i}(x) - \frac{\partial f}{\partial x^i}(x') \right\|_W \right\}$$

für alle  $x, x' \in X$ . Damit und mit der Stetigkeit der partiellen Ableitungen zeigt man sofort die Stetigkeit der Abbildung

$$X \ni x \mapsto A(x) \in L(\mathbb{R}^n, W)$$

bezüglich der Norm  $\|\cdot\|_\infty$ , welche nach Proposition 8.3.4, *ii.)* zur Operatornorm äquivalent ist. Es bleibt also lediglich zu zeigen, dass  $f$  bei jedem Punkt  $x_0 \in X$  differenzierbar mit Ableitung  $Df(x_0) = A(x_0)$  ist. Sei also  $x_0 \in X$  und  $r > 0$  derart, dass  $x_0 + v \in X$  für  $v \in \mathbb{R}^n$  mit  $\|v\|_\infty < r$ . Dank der Offenheit von  $X$  finden wir immer ein solches  $r$ . Wir betrachten nun die Hilfspunkte

$$\begin{aligned} x_1 &= x_0 + v^1 e_1 \\ x_2 &= x_0 + v^1 e_1 + v^2 e_2 = x_1 + v^2 e_2 \\ &\vdots \\ x_n &= x_0 + v^1 e_1 + \cdots + v^n e_n = x_{n-1} + v^n e_n \end{aligned}$$

sowie die zugehörigen Verbindungsstrecken

$$x_k(t) = x_{k-1} + tv^k e_k = x_0 + v^1 e_1 + \cdots + v^{k-1} e_{k-1} + tv^k e_k$$

für  $k = 1, \dots, n$  (ohne Summenkonvention) für  $t \in [0, 1]$ , siehe auch Abbildung 8.7. Längs der Koordinatenachsen ist  $f$  ja partiell differenzierbar mit Ableitungen  $\frac{\partial f}{\partial x^k}$ . Wir betrachten nun den Restterm für die totale Differenzierbarkeit von  $f$  bei  $x_0$  und fügen die Zwischenpunkte und die partiellen Ableitungen geeignet ein. Dies liefert

$$\begin{aligned} &\|f(x_0 + v) - f(x_0) - A(x)v\|_W \\ &= \|f(x_n) - f(x_{n-1}) + f(x_{n-1}) - \cdots + f(x_1) - f(x_0) - A(x)v\| \end{aligned}$$

$$\begin{aligned}
&\leq \left\| f(x_n) - f(x_{n-1}) - \frac{\partial f}{\partial x^n}(x_{n-1})v^n \right\|_W + \cdots + \left\| f(x_1) - f(x_0) - \frac{\partial f}{\partial x^1}(x_0)v^1 \right\|_W \\
&\quad + \left\| \left( \frac{\partial f}{\partial x^n}(x_{n-1}) - \frac{\partial f}{\partial x^n}(x_0) \right) v^n \right\|_W + \cdots + \left\| \left( \frac{\partial f}{\partial x^2}(x_1) - \frac{\partial f}{\partial x^2}(x_0) \right) v^2 \right\|_W \\
&= \left\| f(x_{n-1} + v^n e_n) - f(x_{n-1}) - \frac{\partial f}{\partial x^n}(x_{n-1})v^n \right\|_W + \cdots + \left\| f(x_0 + v^1 e_1) - f(x_0) - \frac{\partial f}{\partial x^1}(x_0)v^1 \right\|_W \\
&\quad + \left\| \frac{\partial f}{\partial x^n}(x_{n-1}) - \frac{\partial f}{\partial x^n}(x_0) \right\| |v^n| + \cdots + \left\| \frac{\partial f}{\partial x^2}(x_1) - \frac{\partial f}{\partial x^2}(x_0) \right\| |v^2|.
\end{aligned}$$

Damit erhalten wir für  $v \neq 0$

$$\begin{aligned}
&\frac{1}{\|v\|_\infty} \|f(x_0 + v) - f(x_0) - A(x_0)v\|_W \\
&\leq \frac{1}{\|v\|_\infty} \left\| f(x_{n-1} + v^n e_n) - f(x_{n-1}) - \frac{\partial f}{\partial x^n}(x_{n-1})v^n \right\|_W \\
&\quad + \cdots + \frac{1}{\|v\|_\infty} \left\| f(x_0 + v^1 e_1) - f(x_0) - \frac{\partial f}{\partial x^1}(x_0)v^1 \right\|_W \\
&\quad + \left\| \frac{\partial f}{\partial x^n}(x_{n-1}) - \frac{\partial f}{\partial x^n}(x_0) \right\|_W \frac{|v^n|}{\|v\|_\infty} \\
&\quad + \cdots + \left\| \frac{\partial f}{\partial x^2}(x_1) - \frac{\partial f}{\partial x^2}(x_0) \right\|_W \frac{|v^2|}{\|v\|_\infty}.
\end{aligned}$$

Wir behaupten, dass für  $v \rightarrow 0$  jeder einzelne Summand gegen Null konvergiert. Für  $k = 1, \dots, n$  ist nämlich  $|v^k| \leq \|v\|_\infty$  und die partielle Differenzierbarkeit von  $f$  bei  $x_{k-1}$  zeigt dann

$$\frac{1}{\|v\|_\infty} \left\| f(x_{k-1} + v^k e^k) - f(x_{k-1}) - \frac{\partial f}{\partial x^k}(x_{k-1})v^k \right\|_W \rightarrow 0.$$

Weiter ist die partielle Ableitung  $\frac{\partial f}{\partial x^k}$  bei  $x_0$  stetig, so dass

$$\left\| \frac{\partial f}{\partial x^k}(x_{k-1}) - \frac{\partial f}{\partial x^k}(x_0) \right\|_W \rightarrow 0$$

für  $x_{k-1} \rightarrow x_0$ , was durch  $v \rightarrow 0$  gewährleistet wird. Schließlich ist  $\frac{|v^k|}{\|v\|_\infty} \leq 1$ , womit alle Summanden im Grenzwert  $v \rightarrow 0$  gegen Null konvergieren. Dies zeigt die Differenzierbarkeit von  $f$  bei  $x_0$  mit der Ableitung  $A(x)$  wie gewünscht.  $\square$

Mit diesem Satz, der Kettenregel sowie der Produktregel und dem Wissen über die Verträglichkeit von Stetigkeit und Differenzierbarkeit haben wir nun alle wichtigen Kriterien versammelt, um Differenzierbarkeitseigenschaften auch für Abbildungen zwischen normierten Räumen zu verstehen. Es sei jedoch an dieser Stelle angemerkt, dass der Satz 8.3.5 entscheidend von der Endlichdimensionalität des Definitionsbereichs Gebrauch macht.

**Proposition 8.3.6** *Seien  $V$  und  $W$  normierte Vektorräume mit einer nicht-leeren offenen Teilmenge  $X \subseteq V$ .*

*i.) Die stetig differenzierbaren Abbildungen von  $X$  nach  $W$  bilden einen Untervektorraum*

$$\mathcal{C}^1(X, W) \subseteq \mathcal{C}^0(X, W). \tag{8.3.13}$$

*ii.) Die stetig differenzierbaren Funktionen auf  $X$  bilden eine Unteralgebra*

$$\mathcal{C}^1(X, \mathbb{R}) \subseteq \mathcal{C}^0(X, \mathbb{R}). \tag{8.3.14}$$

iii.) Die Abbildungen  $\mathcal{C}^1(X, W)$  bilden einen Modul über  $\mathcal{C}^1(X, \mathbb{R})$  mit Modulmultiplikation

$$\mathcal{C}^1(X, \mathbb{R}) \times \mathcal{C}^1(X, W) \ni (f, g) \mapsto f \cdot g \in \mathcal{C}^1(X, W), \quad (8.3.15)$$

definiert punktweise durch

$$(f \cdot g)(x) = f(x) \cdot g(x). \quad (8.3.16)$$

Es gilt die Produktregel

$$D(fg) = Df \cdot g + f Dg \quad (8.3.17)$$

wie in Korollar 8.1.8, ii.).

iv.) Die Ableitung ist eine lineare Abbildung

$$D: \mathcal{C}^1(X, W) \longrightarrow \mathcal{C}^0(X, L(V, W)). \quad (8.3.18)$$

BEWEIS: Dies ist im Wesentlichen eine Zusammenfassung von Ergebnissen, die wir schon erzielt haben. Seien  $f, g \in \mathcal{C}^1(X, W)$  und  $\alpha, \beta \in \mathbb{R}$ . Dann ist nach Korollar 8.1.8, i.) die Linearkombination  $\alpha f + \beta g$  an jedem Punkt differenzierbar mit  $D(\alpha f + \beta g) = \alpha Df + \beta Dg$ . Damit ist aber die Ableitung  $D(\alpha f + \beta g)$  als Linearkombination stetiger Abbildungen selbst wieder stetig. Für  $V = \mathbb{R}^n$  ist dies gerade die Aussage von Proposition 7.7.14. Der dortige Beweis ist aber allgemein gültig, siehe Übung ???. Dies zeigt i.) und ebenso iv.). Das Produkt von einer skalaren Funktion  $f \in \mathcal{C}^1(X, \mathbb{R})$  mit einer vektorwertigen Funktion  $g \in \mathcal{C}^1(X, W)$  wieder differenzierbar sind, ist die Aussage von Korollar 8.1.8, ii.). Die Formel aus der Produktregel zeigt dann, dass die Ableitung (8.3.17) auch wieder stetig ist, siehe nochmals Übung ???. Dies zeigt dann auch ii.), indem man  $W = \mathbb{R}$  setzt. Dass es sich algebraisch um eine Modulstruktur handelt ist klar, siehe auch Übung ???.  $\square$

## 8.4 $\mathcal{C}^k$ -Funktionen

Da für eine differenzierbare Abbildung  $f$  die Ableitung wieder eine vektorwertige Abbildung, nun eben im normierten Vektorraum der stetigen linearen Abbildungen ist, können wir auch nach der Differenzierbarkeit der Ableitung fragen. Da Differenzierbarkeit Stetigkeit impliziert, benötigen wir also wie schon im eindimensionalen Fall eine mindestens stetig differenzierbare Abbildung  $f$ . Dies könne wir induktiv fortsetzen, womit wir die  $k$ -malige Differenzierbarkeit definieren.

**Definition 8.4.1 (Höhere Ableitungen und  $\mathcal{C}^k$ -Funktionen)** Seien  $V$  und  $W$  normierte Vektorräume mit einer nicht-leeren offenen Teilmenge  $X \subseteq V$ .

- i.) Eine Abbildung  $f: X \longrightarrow W$  heißt  $k$ -mal differenzierbar für  $k \in \mathbb{N}$ , wenn  $f$  differenzierbar ist und  $Df: X \longrightarrow L(V, W)$  eine  $(k-1)$ -mal stetig differenzierbare Abbildung ist.
- ii.) Eine Abbildung  $f: X \longrightarrow W$  heißt  $k$ -mal stetig differenzierbar oder  $\mathcal{C}^k$ , wenn  $f$  differenzierbar ist und  $Df: X \longrightarrow L(V, W)$  eine  $(k-1)$ -mal stetig differenzierbare Abbildung ist. Die Menge der  $k$ -mal stetig differenzierbaren Abbildungen bezeichnet man mit

$$\mathcal{C}^k(X, W) = \left\{ f: X \longrightarrow W \mid f \text{ ist } \mathcal{C}^k \right\}. \quad (8.4.1)$$

- iii.) Eine Abbildung  $f: X \longrightarrow W$  heißt glatt oder  $\mathcal{C}^\infty$ , falls  $f \in \mathcal{C}^k(X, W)$  für alle  $k \in \mathbb{N}$ . Die Menge der glatten Abbildungen wird mit

$$\mathcal{C}^\infty(X, W) = \bigcap_{k=1}^{\infty} \mathcal{C}^k(X, W) \quad (8.4.2)$$

bezeichnet.

Der Fall  $k = 0$  wird hier mit in der Notation erfasst, wenn man wie schon zuvor

$$\mathcal{C}^0(X, W) = \mathcal{C}(X, W) \quad (8.4.3)$$

setzt.

Um einen besseren Überblick zu erlangen, schreiben wir die Differenzierbarkeitsbedingungen für  $K = 2$  aus. Zunächst ist die Ableitung  $Df$  von  $f$  eine Abbildung

$$Df: X \longrightarrow L(X, W). \quad (8.4.4)$$

Damit wird die Ableitung  $D(Df)$  der Ableitung  $Df$  eine Abbildung

$$D(Df): X \longrightarrow L(V, L(V, W)). \quad (8.4.5)$$

Damit ist also  $D(Df)(x): V \longrightarrow L(V, W)$  eine stetige lineare Abbildung, mit Werten in den stetigen linearen Abbildungen  $L(V, W)$ , wobei also für  $v \in V$

$$D(Df)(x)v \in L(V, W) \quad (8.4.6)$$

und dann für  $u \in V$

$$(D(Df)(x)v)u \in W. \quad (8.4.7)$$

Da die Vektorraumstruktur von  $L(V, W)$  punktweise erklärt ist, ist die Abbildung

$$V \otimes V \ni (v, u) \mapsto (D(Df)(x)v)u \in W \quad (8.4.8)$$

nicht nur in  $u$  linear, sondern eben auch in  $v$  bei festem  $u$ . Insgesamt ist (8.4.8) also eine bilineare Abbildung. Dieser Überlegung liegt folgender allgemeine Isomorphismus aus der linearen Algebra zugrunde:

**Proposition 8.4.2** *Seien  $V_1, V_2$  und  $W$  Vektorräume über einem Körper  $\mathbb{k}$ . Die bilinearen Abbildungen*

$$\text{Hom}(V_1, V_2; W) = \left\{ \tilde{\Phi}: V_1 \otimes V_2 \longrightarrow W \mid \tilde{\Phi} \text{ bilinear} \right\} \quad (8.4.9)$$

*sind dann kanonisch in linearer Bijektion zu  $\text{Hom}(V_1, \text{Hom}(V_2, W))$  via*

$$\text{Hom}(V_1, \text{Hom}(V_2, W)) \ni \Phi \mapsto ((v_1, v_2) \mapsto \tilde{\Phi}(v_1, v_2) = \Phi(v_1)v_2) \in \text{Hom}(V_1, V_2; W). \quad (8.4.10)$$

BEWEIS: Auch wenn dies eigentlich ein Resultat der linearen Algebra ist, geben wir hier einen Beweis. Sei also  $\phi \in \text{Hom}(V_1, \text{Hom}(V_2, W))$  gegeben. Dann ist  $\Phi(v_1) \in \text{Hom}(V_2, W)$ , womit  $\Phi(v_1)v_2 \in W$  wie gewünscht. Da  $\Phi(v_1)$  linear ist, ist  $\tilde{\Phi}$  im zweiten Argument  $v_2$  linear bei festem  $v_1$ . Weiter ist  $v_1 \mapsto \Phi(v_1) \in \text{Hom}(V_2, W)$  nach Voraussetzung linear, womit für festes  $v_2 \in V_2$  die Abbildung

$$V_1 \ni v_1 \mapsto \tilde{\Phi}(v_1, v_2) = \Phi(v_1)v_2 \in W$$

linear ist, da die lineare Struktur von  $\text{Hom}(V_2, W)$  punktweise in  $V_2$  erklärt ist. Dies zeigt  $\tilde{\Phi} \in \text{Hom}(V_1, V_2; W)$  wie gewünscht. Da die lineare Struktur sowohl von  $\text{Hom}(V_1, \text{Hom}(V_2, W))$  als auch von  $\text{Hom}(V_1, V_2; W)$  punktweise in  $V_1$  beziehungsweise in  $V_1 \times V_2$  erklärt ist, folgt, dass (8.4.10) eine lineare Abbildung ist. Es bleibt also zu zeigen, dass (8.4.10) bijektiv ist. Dazu betrachten wir für eine bilineare Abbildung  $\Psi \in \text{Hom}(V_1, V_2; W)$  die Abbildung  $\hat{\Psi} \in \text{Hom}(V_1, \text{Hom}(V_2, W))$  mit

$$\hat{\Psi}(v_1) = (v_2 \mapsto \hat{\Psi}(v_1)v_2 = \Psi(v_1, v_2)).$$

Die Bilinearität von  $\Psi$  zeigt zunächst, dass  $\hat{\Psi}(v_1)$  für festes  $v_1 \in V$  tatsächlich in  $\text{Hom}(V_2, W)$  ist. Weiter folgt man, dass  $\hat{\Psi}(v_1)$  auch linear von  $v_1$  abhängt, daher ist  $\hat{\Psi}$  tatsächlich ein Element von  $\text{Hom}(V_1, \text{Hom}(V_2, W))$  wie behauptet. Anhand der Formeln sieht man nun, dass  $\Phi \mapsto \tilde{\Phi}$  und  $\Psi \mapsto \hat{\Psi}$  zueinander invers sind, was den Beweis abschließt.  $\square$

In der linearen Algebra wird die Existenz eines speziellen Vektorraums, des Tensorprodukts  $V_1 \otimes V_2$  von  $V_1$  und  $V_2$  gezeigt, der die universelle Eigenschaft hat, dass

$$\text{Hom}(V_1, V_2; W) \cong \text{Hom}(V_1 \otimes V_2; W) \quad (8.4.11)$$

gilt. Mit dieser Theorie der Tensorprodukte wird der Isomorphismus (8.4.10) die sogenannte *T- $\text{Hom}$ -Adjunktion*. Wir werden dies nicht weiter benötigen, verweisen daher aber auf [ , ]. Es sollte ebenfalls darauf hingewiesen werden, dass es iterierte Versionen von (8.4.10) mit entsprechenden multilinearen Abbildungen mit mehr als zwei Argumenten gibt, siehe Übung ???. Da wir nicht an einer rein linear-algebraischen Situation interessiert sind, sondern normierte Vektorräume vorliegen haben und stetige Abbildungen untersuchen wollen, benötigen wir eine etwas verfeinerte Version von Proposition 8.4.2. Sind  $V_1, V_2$  und  $W$  normierte Vektorräume, so bezeichnen wir mit

$$L(V_1, V_2; W) = \{ \Phi \in \text{Hom}(V_1, V_2; W) \mid \Phi \text{ ist stetig} \} \quad (8.4.12)$$

die stetigen bilinearen Abbildungen  $\Phi: V_1 \times V_2 \rightarrow W$ . Hier verstehen wir  $V_1 \times V_2$  mit der Produkttopologie, die beispielsweise von der Norm aus Bemerkung 7.1.4, *i.*) induziert wird. Es gilt also, Kriterien für die Stetigkeit von bilinearen Abbildungen zu finden und mit dem Isomorphismus aus Proposition 8.4.2 in Einklang zu bringen.

**Proposition 8.4.3** *Seien  $V_1, V_2$  und  $W$  normierte Vektorräume, und sei  $\Phi: V_1 \times V_2 \rightarrow W$  eine bilineare Abbildung. Dann sind folgende Aussagen äquivalent:*

- i.) Die bilineare Abbildung  $\Phi$  ist stetig.*
- ii.) Die bilineare Abbildung  $\Phi$  ist stetig bei  $(0, 0)$ .*
- iii.) Es gibt ein  $c > 0$  mit*

$$\|\Phi(v_1, v_2)\|_W \leq c \|v_1\|_{V_1} \|v_2\|_{V_2}. \quad (8.4.13)$$

- iv.) Es gilt*

$$\|\Phi\| = \sup \{ \|\Phi(v_1, v_2)\|_W \mid v_1 \in V_1, v_2 \in V_2 \text{ mit } \|v_1\|_{V_1}, \|v_2\|_{V_2} \leq 1 \} < \infty \quad (8.4.14)$$

BEWEIS: Wir zeigen zunächst die Äquivalenz von *iii.*) und *iv.*). Gilt (8.4.14) und sind  $v_1 \in V_1 \setminus \{0\}$  und  $v_2 \in V_2 \setminus \{0\}$ , so haben die Vektoren  $\frac{v_1}{\|v_1\|_{V_1}}$  und  $\frac{v_2}{\|v_2\|_{V_2}}$  Norm eins. Daher folgt

$$\begin{aligned} \|\Phi\| &\geq \left\| \Phi \left( \frac{v_1}{\|v_1\|_{V_1}}, \frac{v_2}{\|v_2\|_{V_2}} \right) \right\|_W \\ &= \left\| \frac{1}{\|v_1\|_{V_1} \|v_2\|_{V_2}} \Phi(v_1, v_2) \right\|_W \\ &= \frac{1}{\|v_1\|_{V_1} \|v_2\|_{V_2}} \|\Phi(v_1, v_2)\|_W, \end{aligned}$$

da  $\Phi$  bilinear ist. Dies zeigt (8.4.13) mit  $c = \|\Phi\|$ , da für  $v_1$  oder  $v_2$  gleich Null die Ungleichung (8.4.13) trivialerweise gilt. Gilt umgekehrt (8.4.13), so folgt

$$\|\Phi(v_1)v_2\|_W \leq c$$

für alle  $v_1 \in V_1$  und  $v_2 \in V_2$  mit  $\|v_1\|_{V_1}, \|v_2\|_{V_2} \leq 1$  was sofort

$$\|\Phi\| = \sup \{ \|\Phi(v_1, v_2)\|_W \mid v_1 \in V_1, v_2 \in V_2 \text{ mit } \|v_1\|_{V_1}, \|v_2\|_{V_2} \leq 1 \} \leq c$$

zeigt, also *iv.*). Die Implikation *i.*)  $\implies$  *ii.*) ist trivial. Es gelte also zunächst *ii.*). Da  $B_1(0) \subseteq W$  eine (offene) Umgebung von  $0 \in W$  ist, ist  $\Phi^{-1}(B_1(0)) \subseteq V_1 \times V_2$  eine Umgebung von  $(0, 0) \in V_1 \times V_2$ . Es gibt also  $r_1, r_2 > 0$  mit  $B_{r_1}(0) \times B_{r_2}(0) \subseteq \Phi^{-1}(B_1(0))$ , was eben

$$\|v_1\|_{V_1} < r_1 \text{ und } \|v_2\|_{V_2} < r_2 \implies \|\Phi(v_1, v_2)\|_W < 1$$

bedeutet. Seien nun  $v_1 \in V_1$  und  $v_2 \in V_2$  ungleich Null. Dann gilt

$$\frac{r_1}{2\|v_1\|_{V_1}}v_1 \in B_{r_1}(0) \quad \text{und} \quad \frac{r_2}{2\|v_2\|_{V_2}}v_2 \in B_{r_2}(0),$$

womit

$$\begin{aligned} 1 &> \left\| \Phi \left( \frac{r_1}{2\|v_1\|_{V_1}}v_1, \frac{r_2}{2\|v_2\|_{V_2}}v_2 \right) \right\|_W \\ &= \frac{r_1 r_2}{4\|v_1\|_{V_1}\|v_2\|_{V_2}} \|\Phi(v_1, v_2)\|_W \end{aligned}$$

und damit (8.4.13) mit  $c = \frac{4}{r_1 r_2}$  folgt. Gilt umgekehrt (8.4.13) und ist  $(v_{1,n}, v_{2,n})_{n \in \mathbb{N}}$  eine konvergente Folge mit Grenzwert  $(v_1, v_2) \in V_1 \times V_2$ , dann gilt

$$\begin{aligned} &\|\Phi(v_{1,n}, v_{2,n}) - \Phi(v_1, v_2)\|_W \\ &= \|\Phi(v_{1,n} - v_1, v_{2,n}) + \Phi(v_1, v_{2,n}) - \Phi(v_1, v_2)\|_W \\ &\leq \|\Phi(v_{1,n} - v_1, v_{2,n})\|_W + \|\Phi(v_1, v_{2,n} - v_2)\|_W \\ &\leq c\|v_{1,n} - v_1\|_{V_1}\|v_{2,n}\|_{V_2} + c\|v_1\|_{V_1}\|v_{2,n} - v_2\|_{V_2} \rightarrow 0. \end{aligned}$$

Im ersten Summanden konvergiert  $\|v_{2,n}\|_{V_2}$  gegen  $\|v_2\|_{V_2}$ , womit die Konvergenz  $\|v_{1,n} - v_1\|_{V_1} \rightarrow 0$  zeigt, dass der erste Summand gegen Null konvergiert. Beim zweiten Summanden zeigt die Konvergenz  $\|v_{2,n} - v_2\|_{V_2} \rightarrow 0$  direkt, dass auch dieser gegen Null konvergiert. Damit folgt aber  $\Phi(v_{1,n}, v_{2,n}) \rightarrow \Phi(v_1, v_2)$ , was die (Folgen-) Stetigkeit von  $\Phi$  bei  $(v_1, v_2)$  zeigt.  $\square$

Man bemerkt, dass der Beweis dem von Satz 7.3.7 sehr ähnelt. Lediglich die Lipschitz-Stetigkeit ist nicht länger gewährleistet. Es ist nun schnell zu sehen, dass für eine stetige bilineare Abbildung  $\Phi \in L(V_1, V_2; W)$  durch (8.4.14) eine Norm definiert wird, die auch als kleinste Konstante  $c \geq 0$  in (8.4.13) gesehen werden kann, siehe auch Übung ???. Auch hier gibt es entsprechende Versionen für multilineare Abbildungen mit mehr als zwei Argumenten, siehe Übung ???.

Nach diesen Vorüberlegungen können wir also Proposition 8.4.2 mit den Stetigkeitsanforderungen zusammenbringen.

**Proposition 8.4.4** *Seien  $V_1, V_2$  und  $W$  normierte Vektorräume. Sei  $\Phi \in \text{Hom}(V_1, \text{Hom}(V_2, W))$  mit zugehörigem  $\tilde{\Phi} \in \text{Hom}(V_1, V_2; W)$ , wie in (8.4.10). Dann gilt  $\Phi \in L(V_1, L(V_2, W))$  genau dann, wenn  $\tilde{\Phi} \in L(V_1, V_2; W)$ . In diesem Fall gilt*

$$\|\Phi\| = \|\tilde{\Phi}\| \quad (8.4.15)$$

BEWEIS: Sei zunächst  $\Phi \in L(V_1, L(V_2, W))$ . Für die Operatornorm  $\|\cdot\|'$  in  $L(V_2, W)$  gilt dann

$$\|\Phi(v_1)\|' \leq \|\Phi\| \|v_1\|_{V_1}$$

mit der Operatornorm von  $L(V_1, L(V_2, W))$ . Weiter gilt für alle  $v_2 \in V_2$

$$\|\Phi(v_1)v_2\|_W \leq \|\Phi(v_1)\|' \|v_2\|_{V_2}$$

womit insgesamt

$$\|\Phi(v_1)v_2\|_W \leq \|\Phi\| \|v_1\|_{V_1} \|v_2\|_{V_2}$$

folgt. Da  $\tilde{\Phi}(v_1, v_2) = \Phi(v_1)v_2$  nach (8.4.10), zeigt dies

$$\|\tilde{\Phi}(v_1, v_2)\|_W \leq \|\Phi\| \|v_1\|_{V_1} \|v_2\|_{V_2},$$

womit zum einen  $\tilde{\Phi} \in L(V_1, V_2; W)$  folgt, zum anderen  $\|\tilde{\Phi}\| \leq \|\Phi\|$  gezeigt ist, da die Norm von  $\tilde{\Phi}$  die kleinste Konstante  $c \geq 0$  in (8.4.13) ist. Sei umgekehrt  $\Psi \in L(V_1, V_2; W)$  mit  $\hat{\Psi} \in \text{Hom}(V_1, \text{Hom}(V_2, W))$

als Urbild unter (8.4.10) wie im Beweis von Proposition 8.4.2. Wir wissen  $\|\Psi(v_1, v_2)\|_W \leq \|\Psi\| \|v_1\|_{V_1} \|v_2\|_{V_2}$  für alle  $v_1 \in V_1$  und  $v_2 \in V_2$ . Damit gilt

$$\|\hat{\Psi}(v_1)v_2\|_W \leq \|\Psi\| \|v_1\|_{V_1} \|v_2\|_{V_2},$$

was zunächst  $\hat{\Psi}(v_1) \in L(V_2, W)$  mit Operatornorm  $\|\hat{\Psi}(v_1)\|' \leq \|\Psi\| \|v_1\|_{V_1}$  zeigt. Damit ist aber auch  $\hat{\Psi} \in L(V_1, L(V_2, W))$  mit

$$\|\hat{\Psi}\| \leq \|\Psi\|$$

gezeigt. Insgesamt haben wir daher die Äquivalenz sowie beide nötigen Ungleichungen für die Gleichheit in (8.4.15) gezeigt.  $\square$

**Korollar 8.4.5** *Seien  $V_1, V_2$  und  $W$  normierte Vektorräume, so dass  $V_1$  und  $V_2$  beide endlichdimensional sind. Dann ist jede bilineare Abbildung  $\Phi: V_1 \times V_2 \rightarrow W$  stetig.*

BEWEIS: Da  $V_2$  endlichdimensional ist, wissen wir  $L(V_2, W) = \text{Hom}(V_2, W)$ . Da  $V_1$  endlichdimensional ist, wissen wir  $L(V_1, \text{Hom}(V_2, W)) = \text{Hom}(V_1, \text{Hom}(V_2, W))$ , womit Proposition 8.4.4 und Proposition 8.4.2 zum Einsatz kommen.  $\square$

Ohne die Details auszuformulieren, bemerken wir hier, dass entsprechende Resultate induktiv auch für multilineare Abbildungen mit entsprechend mehr als nur zwei Argumenten gelten, siehe auch Übung ???. Nach diesen Vorüberlegungen kommen wir zurück zu zweimaliger Differenzierbarkeit und der zweiten Ableitung. Die zweite Ableitung einer zweimal differenzierbaren Abbildung  $f: X \rightarrow W$  mit einer nicht-leeren offenen Teilmenge  $X \subseteq V$  und zwei normierten Vektorräumen liefert also eine zweite Ableitung

$$D(Df): X \rightarrow L(V, L(V, W)) \quad (8.4.16)$$

welche wir an jedem Punkt  $x \in X$  als stetige bilineare Abbildung

$$D^2f(x) \in L(V, V; W) \quad (8.4.17)$$

auffassen können, wobei wir den Isomorphismus

$$D^2f(x)(u, v) = (D(Df)(x)v)u \quad (8.4.18)$$

aus Proposition 8.4.4 verwenden. Die überraschende Eigenschaft der zweiten Ableitung ist nun ihre Symmetrie: 5.4.5

**Satz 8.4.6 (Symmetrie von  $D^2f$ )** *Seien  $V$  und  $W$  normierte Vektorräume und  $X \subseteq V$  eine nicht-leere offene Teilmenge. Ist  $f: X \rightarrow W$  zweimal total differenzierbar, so ist die zweite Ableitung*

$$D^2f(x) \in L(V, V; W) \quad (8.4.19)$$

*von  $f$  an jedem Punkt symmetrisch, und*

$$D^2f(x)(v, u) = D^2f(x)(u, v) \quad (8.4.20)$$

*für alle  $v, u \in V$ .*

BEWEIS: Da wir recht allgemeine Voraussetzungen verwenden, ist der Beweis etwas technisch: Die wesentliche Idee ist, die Ableitung als

$$D(Df(x)v)u = \lim_{t \rightarrow 0} \frac{f(x + tv + tu) - f(x + tv) - f(x + tu) + f(x)}{t^2} \quad (*)$$

zu erhalten. Ist dies gezeigt, so folgt (8.4.20) sofort, da die rechte Seite offensichtlich unter der Vertauschung  $u \leftrightarrow v$  invariant ist. Es gilt also (\*) zu zeigen. Da  $Df: X \rightarrow L(V, W)$  eine auf  $X$  differenzierbare Funktion ist, finden wir einen Restterm

$$R: X \rightarrow L(V, W)$$

der zweiten Ableitung derart, dass für  $D(Df(x)) \in L(V, L(V, W))$  die Gleichung

$$Df(x+h) = Df(x) + D(Df)(x)h + R(x+h) \quad (**)$$

mit

$$\lim_{h \rightarrow 0} \frac{R(x+h)}{\|h\|} = 0 \quad (\odot)$$

gilt. Insbesondere ist  $R(x) = 0$ . Man beachte, dass (\*\*) eine Gleichung zwischen linearen Abbildungen in  $L(V, W)$  ist, insbesondere können wir (\*\*) noch auf Elemente aus  $V$  auswerten, um eine Gleichung zwischen Vektoren in  $W$  zu erhalten. Wir betrachten nun Punkte  $x+h \in X$  mit

$$h = tv + stu$$

und

$$h = stu$$

in (\*\*). Dies liefert

$$\begin{aligned} Df(x+tv+stu) - Df(x+stu) \\ = D(Df(x)tv) + R(x+tv+stu) - R(x+stu), \end{aligned}$$

da der Term mit der zweiten Ableitung in (\*\*) linear in  $h$  ist. Wir werten dies auf dem Vektor  $tu$  aus und erhalten

$$\begin{aligned} Df(x+tv+stu)tu - Df(x+stu)tu \\ = D(Df(x)tv)tu + R(x+tv+stu)tu - R(x+stu)tu. \end{aligned} \quad (*)$$

Die linke Seite können wir nun folgendermaßen interpretieren. Wir betrachten für festes  $t$  die Abbildung

$$F_t(s) = f(x+tv+stu) - f(x+stu),$$

definiert für kleine  $t$  und  $s \in [0, 1]$ . Da  $f$  differenzierbar ist, und die Gerade  $s \mapsto x+tv+stu$  ebenfalls differenzierbar ist, können wir  $F_t$  differenzieren und erhalten mit der Kettenregel

$$DF_t(s) = Df(x+tv+stu)tu - Df(x+stu)tu.$$

Man beachte, dass  $DF_t(s)$  nun ein Vektor in  $W$  ist, da  $F_t$  eine Kurve in  $W$  darstellt. Hier identifizieren wir  $DF_t(s)$  wie schon zuvor mit  $DF_t(s)1$ , siehe Bemerkung 8.1.9, i.). Damit wird (\*) also zu

$$DF_t(s) = t^2 D(Df(x)v)u + tR(x+tv+stu)u - tR(x+stu)u, \quad (\star\star)$$

wobei wir erneut die Linearität der beiden Ableitungen sowie die Linearität der Werte der Abbildung  $R$  ausnutzen. Als nächstes benötigen wir eine Abschätzung für  $F_t$  ähnlich zum Schrankensatz aus Satz 8.3.1. Wir betrachten eine differenzierbare vektorwertige Funktion  $g: (a, b) \rightarrow W$ , womit nach Satz 8.3.1

$$\|g(x) - g(x')\|_W \leq \sup_{\xi \in [x, x']} \|Dg(\xi)\| \|x - x'\|$$



gilt. Sieht man nun

$$g(s) = F_t(s) - DF_t(0)s,$$

so gilt  $Dg(s) = DF_t(s) - DF_t(0)$  sowie  $Dg(0) = 0$ . Verwenden wir nun den Schrankensatz, so folgt die Abschätzung

$$\|f_t(s) - DF_t(0)s - F_t(0)\|_W \leq \sup_{\sigma \in [0,1]} \|DF_t(\sigma) - DF_t(0)\|_s. \quad (\odot)$$

Nun haben wir alle Bausteine vorliegen. Wir schreiben zunächst mit  $(\star\star)$  und der Definition von  $F_t$ .

$$\begin{aligned} & \frac{F_t(1) - F_t(0) - DF_t(0)}{t^2} \\ &= \frac{f(x + tv + tu) - f(x + tu) - f(x + tv) + f(0)}{t^2} - D(Df(x)v)u - \frac{R(x + tv)u}{t}, \end{aligned}$$

da  $R(x) = 0$  in  $(\star\star)$  verwendet werden kann. Die rechte Seite können wir daher abschätzen

$$\begin{aligned} & \left\| \frac{f(x + tv + tu) - f(x + tv) - f(x + tu) + f(0)}{t^2} - D(Df(x)v)u \right\|_W \\ &= \left\| \frac{F_t(1) - F_t(0) - DF_t(0)}{t^2} + \frac{R(x + tv)u}{t} \right\|_W \\ &\leq \left\| \frac{F_t(1) - F_t(0) - DF_t(0)}{t^2} \right\| + \left\| \frac{R(x + tv)u}{t} \right\|_W \\ &\stackrel{(\odot)}{\leq} \sup_{\sigma \in [0,1]} \frac{1}{t^2} \|DF_t(\sigma) - DF_t(0)\|_W + \frac{\|R(x + tv)\|}{|t|} \|u\|_V \\ &\stackrel{(\star\star)}{\leq} \sup_{\sigma \in [0,1]} \frac{1}{t^2} \|tR(x + tv + \sigma tu)u - tR(x + \sigma tu)u - tR(x + tv)u\|_W \\ &\quad + \frac{\|R(x + tv)\|}{|t|} \|u\|_V \\ &\leq \sup_{\sigma \in [0,1]} \frac{\|R(x + tv + \sigma tu)\|}{|t|} \|u\|_V + \sup_{\sigma \in [0,1]} \frac{\|R(x + \sigma tu)\|}{|t|} \|u\|_V \\ &\quad + \sup_{\sigma \in [0,1]} \frac{\|R(x + tv)\|}{|t|} \|u\|_V + \frac{\|R(x + tv)\|}{|t|} \|u\|_V \\ &= \sup_{\sigma \in [0,1]} \frac{\|R(x + tv + \sigma tu)\|}{\|tv + \sigma tu\|} \|v + \sigma u\|_V \|u\|_V + \sup_{\sigma \in [0,1]} \frac{\|R(x + \sigma tu)\|}{\|t\sigma u\|_V} \|\sigma u\|_V \|u\|_V \\ &\quad + 2 \frac{\|R(x + tv)\|}{\|tv\|_V} \|v\|_V \|u\|_V. \end{aligned}$$

Wir behaupten, dass für  $t \rightarrow 0$  alle drei Terme separat gegen Null gehen. Beim dritten Term ist dies klar nach der definierenden Eigenschaft des Restterms  $R$  aus  $(\odot)$ . Beim ersten Term beachte man zunächst, dass  $\|v + \sigma u\|_V$  für  $\sigma \in [0, 1]$  beschränkt ist, etwa durch  $\|v\|_V + \|u\|_V$ . Die Konvergenz  $(\odot)$  zeigt dann, dass auch

$$\lim_{t \rightarrow 0} \sup_{\sigma \in [0,1]} \frac{\|R(x + tv + \sigma tu)\|}{\|tv + \sigma tu\|} = 0$$

gilt. Beim zweiten Term argumentiert man genauso. Insgesamt erhält man damit die gewünschte Konvergenz  $(*)$  und somit die Symmetrie (8.4.20).  $\square$

**Korollar 8.4.7** Sei  $X \subseteq V$  eine offene nicht-leere Teilmenge eines normierten Vektorraums  $V$  und  $f: X \rightarrow W$  eine  $k$ -mal differenzierbare Funktion mit Werten in einem weiteren normierten Vektorraum  $W$  mit  $k \in \mathbb{N} \cup \{+\infty\}$ . Ist  $\ell \in \mathbb{N}$  mit  $2 \leq \ell \leq k$ , so ist  $D^\ell f$  total symmetrisch.

BEWEIS: Dies ist eine einfache Induktion, da

$$D^{\ell+1}f = D(D^\ell)f = D^\ell(Df). \quad \square$$

Wir können dieses Resultat also insbesondere auf  $\mathcal{C}^k$ -Funktionen anwenden und erhalten die Symmetrie aller Mehrfachableitungen. In endlichen Dimensionen haben wir nach Satz 8.3.5 einfaches Kriterium für die Existenz der totalen Ableitungen anhand der Stetigkeit der partiellen Ableitungen erhalten.

**Korollar 8.4.8 (Satz von Schwarz)** Sei  $X \subseteq \mathbb{R}^n$  eine offene nicht-leere Teilmenge und  $W$  ein normierter Vektorraum. Für  $f \in \mathcal{C}^k(X, W)$  gilt dann

$$\frac{\partial}{\partial x^{i_1}} \cdots \frac{\partial}{\partial x^{i_\ell}} f = \frac{\partial}{\partial x^{i_{\sigma(1)}}} \cdots \frac{\partial}{\partial x^{i_{\sigma(\ell)}}} f \quad (8.4.21)$$

für alle  $i_1, \dots, i_\ell \in \{1, \dots, n\}$  und  $\sigma \in S_\ell$ . Insbesondere gilt für eine  $\mathcal{C}^2$ -Funktion  $f$

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}. \quad (8.4.22)$$

BEWEIS: Die partiellen Ableitungen erhält man wie immer durch Auswerten der totalen Ableitungen auf den Einheitsvektoren, womit die Symmetrie (8.4.22) beziehungsweise (8.4.21) aus Satz 8.4.6 folgt. Wichtig ist, dass wir eine  $\mathcal{C}^k$ -Funktion anhand der Stetigkeit der partiellen Ableitungen charakterisieren können.  $\square$

Gegenbeispiele zeigen, dass die bloße Existenz der zweiten Ableitungen allein noch nicht für (8.4.22) ausreicht: Die Stetigkeit der zweiten partiellen Ableitungen ist eine zwar etwas stärkere Voraussetzung als für Satz (8.4.6) nötig, aber dafür typischerweise leicht zu kontrollieren, siehe auch Übung ??.

**Bemerkung 8.4.9** Der Beweis von Satz 8.4.6 kann unter leicht stärkeren Voraussetzungen vereinfacht werden. Ist beispielsweise  $W = \mathbb{K}$  und damit  $f: X \rightarrow \mathbb{K}$  eine Funktion und ist die zweite Ableitung selbst sogar noch stetig, so kann man die technischen Abschätzungen im Beweis mit Hilfe des Hauptsatzes der Differential- und Integralrechnung etwas konzeptioneller zeigen. Durch komponentenweise Betrachtung erhält man damit auch sofort den Fall endlich-dimensionaler Bildräume  $W$ . Erweitert man den Begriff des Riemann-Integrals auf Banach-Räume als Werte, was keine große Schwierigkeit darstellt, so kann man bei stetiger zweiter Ableitung direkt auch hier mit dem Hauptsatz argumentieren.

## 8.5 Funktionenfolgen und $\mathcal{C}^k$ -Konvergenz

Wie schon in der eindimensionalen Situation in Abschnitt 5.5 wollen wir wieder Folgen von differenzierbaren Funktionen auf  $X \subseteq \mathbb{R}^n$  betrachten und deren Konvergenzeigenschaften studieren. Die Motivation ist die gleiche wie schon in Abschnitt 5.5. Man will komplizierte Funktionen durch einfachere annähern und verstehen, ob gute Eigenschaften der einfachen Funktionen dabei auf die Grenzfunktion übertragen werden.

Wir beginnen dabei mit einem Nachtrag zur rein topologischen Situation, um die Stetigkeit unter Konvergenz zu verstehen. Hier können wir analog zu Abschnitt 3.7 vorgehen:

**Definition 8.5.1 (Lokal gleichmäßige Konvergenz)** Sei  $(M, \mathcal{M})$  ein topologischer Raum. Sei weiter  $(f_n)_{n \in \mathbb{N}}$  eine Folge von Abbildungen  $f_n \in \text{Abb}(M, W)$  mit Werten in einem normierten Vektorraum.

i.) Die Folge  $(f_n)_{n \in \mathbb{N}}$  konvergiert gleichmäßig gegen eine Funktion  $f \in \text{Abb}(X, W)$ , wenn für alle  $\epsilon > 0$  ein  $N \in \mathbb{N}$  existiert, so dass

$$\|f_n(p) - f(p)\|_W < \epsilon \quad (8.5.1)$$

für alle  $p \in M$  gilt.

ii.) Die Folge  $(f_n)_{n \in \mathbb{N}}$  konvergiert lokal gleichmäßig gegen  $f \in \text{Abb}(M, W)$ , wenn es zu jedem Punkt  $p_0 \in X$  eine Umgebung  $U \subseteq M$  von  $p_0$  gibt, so dass  $(f_n|_U)_{n \in \mathbb{N}}$  gleichmäßig gegen  $f|_U$  konvergiert.

Wie zuvor können wir (8.5.1) als

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0 \quad (8.5.2)$$

verstehen, wenn wir die Supremumsnorm

$$\|f\|_\infty = \sup_{p \in M} \|f(p)\|_W \quad (8.5.3)$$

nun mit Hilfe der Norm von  $W$  definieren.

**Satz 8.5.2 (Lokal gleichmäßige Stetigkeit)** Seien  $M$  ein topologischer Raum und  $W$  ein normierter Vektorraum. Sei weiter  $(f_n)_{n \in \mathbb{N}}$  eine Folge von stetigen Abbildungen  $f_n \in \mathcal{C}(M, W)$ . Konvergiert  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen  $f \in \text{Abb}(M, W)$ , so ist  $f$  ebenfalls stetig.

BEWEIS: Wir zeigen, dass  $f$  bei  $p_0 \in M$  für einen beliebigen Punkt  $p_0$  stetig ist. Mit der richtigen Interpretation können wir wörtlich so verfahren wie in Satz 3.6.5. Da  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig konvergiert, finden wir zu  $\epsilon > 0$  und ein  $N \in \mathbb{N}$  eine Umgebung  $U \subseteq M$  von  $p_0$  mit

$$\|f_n(p) - f(p)\|_W < \frac{\epsilon}{3}$$

für alle  $p \in U$  und alle  $n \geq N$ . Insbesondere gilt auch

$$\|f_n(p_0) - f(p_0)\|_W < \frac{\epsilon}{3}.$$

Da  $f_n$  stetig ist, ist  $U' = f_N^{-1}(B_{\frac{\epsilon}{3}}(f(p_0)))$  eine (sogar offene) Umgebung von  $p_0$ , da ja  $B_{\frac{\epsilon}{3}}(f(p_0))$  eine sogar offene Umgebung des Bildpunktes ist. Es gilt also

$$\|f_N(p) - f_N(p_0)\|_W < \frac{\epsilon}{3}$$

für alle  $p \in U'$ . Als Schnitt zweier Umgebungen ist auch  $U'' = U \cap U'$  eine Umgebung von  $p_0$ . Für  $p \in U''$  gilt dann

$$\begin{aligned} \|f(p) - f(p_0)\|_W &\leq \|f(p) - f_N(p)\|_W + \|f_N(p) - f_N(p_0)\|_W + \|f_N(p_0) - f(p_0)\|_W \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon, \end{aligned}$$

was die Stetigkeit von  $f$  bei  $p_0$  zeigt. □

Die wesentliche, neue Schwierigkeit ist nun, dass der topologische Raum  $M$  wenige kompakte Teilmengen haben könnte: In Abschnitt 3.7 haben wir die lokal gleichmäßige Konvergenz mit Hilfe der lokalen Supremumsnormen über kompakten Teilmengen beschreiben können, da es zu jedem Punkt Umgebungen gibt, die zudem kompakt waren. Im Allgemeinen ist dies nicht zu erwarten sondern stellt eine zusätzliche Eigenschaft von  $M$  dar:

**Definition 8.5.3 (Lokal kompakter Raum)** Ein topologischer Raum heißt lokal kompakt, wenn jeder Punkt  $p \in M$  eine kompakte Umgebung besitzt.

Insbesondere ist jeder kompakte topologische Raum lokal kompakt, da  $M$  immer eine Umgebung jedes Punktes in  $M$  ist. Es gibt aber auch andere für uns wichtige Beispiele:

**Beispiel 8.5.4** Sei  $X \subseteq \mathbb{R}^n$  eine offene Teilmenge von  $\mathbb{R}^n$  mit der Relativtopologie. Dann ist  $X$  lokal kompakt. Ist nämlich  $x_0 \in X$ , so gibt es ein  $r > 0$  mit  $B_r(x_0) \subseteq X$ , da  $X$  offen ist. Dann ist  $B_{\frac{r}{2}}(x_0)^{\text{cl}} \subseteq X$  eine kompakte Umgebung von  $x_0$  in  $X$ .

Ist  $M$  lokal kompakt und  $K \subseteq M$ , so ist für jede stetige Funktion  $f \in \mathcal{C}(X, W)$

$$\|f\|_K = \sup_{p \in K} \|f(p)\|_W < \infty, \quad (8.5.4)$$

da auch die Verkettung  $p \mapsto \|f(p)\|_W$  stetig ist und folglich auf dem Kompaktum  $K$  nach Korollar 7.6.5, ii.) sogar ihr Maximum annimmt. Allgemein können wir (8.5.4) noch für lokale beschränkte Funktionen definieren:

**Definition 8.5.5 (Lokal beschränkte Funktionen)** Sei  $M$  ein topologischer Raum und  $W$  ein normierter Vektorraum

i.) Eine Funktion  $f \in \text{Abb}(M, W)$  heißt beschränkt, wenn

$$\|f\|_\infty = \sup_{p \in M} \|f(p)\|_W < \infty \quad (8.5.5)$$

gilt. Die Menge der beschränkten Abbildungen wird nun mit

$$\mathcal{B}(M, W) = \{f \in \text{Abb}(M, W) \mid f \text{ ist beschränkt}\} \quad (8.5.6)$$

bezeichnet.

ii.) Eine Funktion heißt lokal beschränkt, wenn es zu jedem Punkt  $p_0 \in M$  eine Umgebung  $U \subseteq M$  von  $p_0$  mit  $f|_U \in \mathcal{B}(U, W)$  gib. Die Menge der lokal beschränkten Funktionen wird mit

$$\mathcal{B}_{\text{loc}}(M, W) = \{f \in \text{Abb}(M, W) \mid f \text{ ist lokal beschränkt}\} \quad (8.5.7)$$

bezeichnet.

Es ist nun schnell zu sehen, dass sowohl  $\mathcal{B}(M, W)$  als auch  $\mathcal{B}_{\text{loc}}(M, W)$  Untervektorräume von  $\text{Abb}(M, W)$  sind, siehe Übung ???. Für allgemeine topologische Räume lässt sich nicht viel zur Beziehung von  $\mathcal{C}(M, W)$  und  $\mathcal{B}_{\text{loc}}(M, W)$  sagen. Ist  $M$  dagegen lokal kompakt, so erhalten wir die aus Abschnitt 3.7 vertrauten Resultate auch in dieser viel größeren Allgemeinheit:

**Proposition 8.5.6** Seien  $M$  ein lokal kompakter topologischer Raum und  $W$  ein normierter Vektorraum.

i.) Es gilt  $f \in \mathcal{B}_{\text{loc}}(M, W)$  genau dann, wenn zu jedem Kompaktum  $K \subseteq M$

$$\|f\|_K < \infty \quad (8.5.8)$$

gilt

ii.) Es gilt

$$\mathcal{C}(M, W) \subseteq \mathcal{B}_{\text{loc}}(M, W). \quad (8.5.9)$$

BEWEIS: Es gilt zunächst (8.5.8). Da es zu jedem Punkt  $p_0 \in M$  eine kompakte Umgebung  $K \subseteq M$  von  $p_0$  gibt, folgt  $\|f\|_K < \infty$ , womit  $f \in \mathcal{B}_{\text{loc}}(M, W)$  lokal beschränkt ist. Sei umgekehrt  $f \in \mathcal{B}_{\text{loc}}(M, W)$  und  $K$  kompakt. Zu jedem  $p \in K$  gibt es dann eine Umgebung  $U_p \subseteq M$  von  $p$  mit  $f|_{U_p} \in \mathcal{B}(U_p, W)$ , also

$$m_p = \sup_{p' \in U_p} \|f(p')\|_W < \infty.$$

da  $K$  kompakt ist, genügen endlich viele  $p_1, \dots, p_n \in K$ , um

$$K \subseteq U_{p_1} \cup \dots \cup U_{p_n}$$

zu erreichen. Daher ist  $\|f\|_W$  auf  $K$  durch das Maximum

$$m = \max\{m_{p_1}, \dots, m_{p_n}\}$$

beschränkt, womit  $\|f\|_K < \infty$  gezeigt ist. Mit (8.5.4) ist dann der zweite Teil klar.  $\square$

Wir definieren nun eine lokal gleichmäßige Cauchy-Folge wie schon in Proposition 3.7.8 in der nun allgemeinen Situation.

**Definition 8.5.7 (Lokal gleichmäßige Cauchy-Folge)** Seien  $M$  ein lokal kompakter topologischer Raum und  $W$  ein normierter Vektorraum. Eine Folge  $(f_n)_{n \in \mathbb{N}}$  von lokal beschränkten Funktionen  $f_n \in \mathcal{B}_{\text{loc}}(M, W)$  heißt lokal gleichmäßige Cauchy-Folge, wenn es zu jedem  $\epsilon > 0$  und zu jeder kompakten Teilmenge  $K \subseteq M$  ein  $N \in \mathbb{N}$  gibt, so dass

$$\|f_n - f_m\|_K < \epsilon \quad (8.5.10)$$

gilt.

Wir können nun das zentrale Resultat zur lokal gleichmäßigen Konvergenz formulieren:

**Satz 8.5.8 (Vollständigkeit von  $\mathcal{B}_{\text{loc}}(M, W)$  und  $\mathcal{C}(M, W)$ )** Seien  $M$  ein lokal kompakter topologischer Raum und  $W$  ein normierter Vektorraum. Sei weiter  $(f_n)_{n \in \mathbb{N}}$  eine Folge von lokal beschränkten Abbildungen  $f_n \in \mathcal{B}_{\text{loc}}(M, W)$ .

- i.) Konvergiert  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen eine Abbildung  $f \in \text{Abb}(M, W)$ , so gilt  $f \in \mathcal{B}_{\text{loc}}(M, W)$ .
- ii.) Konvergiert  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig, so ist  $(f_n)_{n \in \mathbb{N}}$  eine lokal gleichmäßige Cauchy-Folge.
- iii.) Ist  $W$  sogar ein Banach-Raum und  $(f_n)_{n \in \mathbb{N}}$  eine lokal gleichmäßige Cauchy-Folge, so konvergiert  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen eine Funktion  $f \in \mathcal{B}_{\text{loc}}(M, W)$ .
- iv.) Ist  $W$  sogar ein Banach-Raum und  $(f_n)_{n \in \mathbb{N}}$  eine lokal gleichmäßige Cauchy-Folge von stetigen Funktionen  $f \in \mathcal{C}(M, W)$ , so konvergiert  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen eine stetige Funktion  $f \in \mathcal{C}(M, W)$ .

BEWEIS: Für den ersten Teil betrachten wir einen Punkt  $p_0 \in M$  und eine kompakte Umgebung  $K \subseteq M$  von  $p_0$ , die nach Voraussetzung ja existiert. Dann gibt es zu  $\epsilon = 1$  ein  $N \in \mathbb{N}$  mit

$$\|f_n - f\|_K < \epsilon$$

für  $n \geq N$ . Sei also  $p \in K$ , dann gilt zum einen  $f_n(p) \rightarrow f(p)$  punktweise. Damit gilt aber auch

$$\begin{aligned} \|f(p)\|_W &\leq \|f(p) - f_N(p)\|_W + \|f_N(p)\|_W \\ &\stackrel{(a)}{=} \lim_{n \rightarrow \infty} \|f_n(p) - f_N(p)\|_W + \|f_N(p)\|_W \end{aligned}$$

$$\leq 1 + \|f_N\|_K,$$

da zum einen die Norm  $\|\cdot\|_W$  stetig ist und daher der Grenzwert  $f(p) = \lim_{n \rightarrow \infty} f_n(p)$  in (a) aus der Norm gezogen werden kann und zum anderen  $f_N$  auf  $K$  durch  $\|f_N\|_K$  beschränkt ist. Damit folgt aber  $\|f\|_K \leq 1 + \|f_N\|_K < \infty$  und  $f \in \mathcal{B}_{\text{loc}}(M, W)$  ist gezeigt. Wir zeigen nun, dass  $(f_n)_{n \in \mathbb{N}}$  auch eine Cauchy-Folge ist. Sei dazu  $\epsilon > 0$  vorgegeben. Wir finden ein  $N \in \mathbb{N}$  mit  $\|f_n - f\|_K < \frac{\epsilon}{2}$  für  $n \geq N$ . Damit gilt auch

$$\|f_n - f_m\|_K \leq \|f_n - f\|_K + \|f - f_m\|_K < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

was die gewünschte Cauchy-Eigenschaft zeigt. Für den dritten Teil nehmen wir an,  $(f_n)_{n \in \mathbb{N}}$  sei eine lokal gleichmäßige Cauchy-Folge. Dann gibt es zu  $\epsilon > 0$  und jedem Kompaktum  $K \subseteq M$  ein  $N \in \mathbb{N}$  mit

$$\|f_n - f_m\|_K < \epsilon. \quad (*)$$

Wählen wir als Kompaktum  $K$  einen einzelnen Punkt  $p_0 \in M$ , so erhalten wir eine Cauchy-Folge  $(f_n(p_0))_{n \in \mathbb{N}}$  im als Banach-Raum angenommenen  $W$ . Diese konvergiert daher, was die punktweise Konvergenz gegen eine Funktion  $f \in \text{Abb}(M, W)$ , also

$$f(p) = \lim_{n \rightarrow \infty} f_n(p)$$

für alle  $p \in M$  zeigt. Sei nun  $K$  beliebig und  $N$  mit  $(*)$  gewählt. Für  $p \in K$  gilt dann

$$\|f(p) - f_n(p)\|_W = \lim_{m \rightarrow \infty} \|f_m(p) - f_n(p)\| \leq \epsilon$$

sofern  $n \geq N$ , da im Limes ja nur  $m \geq N$  relevant werden. Dies zeigt, dass  $(f_n)_{n \in \mathbb{N}}$  lokal gleichmäßig gegen  $f$  konvergiert, so dass nach i.) auch  $f \in \mathcal{B}_{\text{loc}}(M, W)$  folgt. Ist nun sogar  $f_n \in \mathcal{C}(M, W)$  stetig für alle  $n \in \mathbb{N}$ , so folgt der letzte Teil aus iii.) und Satz 8.5.2.  $\square$

Dies ist nun das erste Mal, dass die Vollständigkeit des Zielraumes  $W$  von entscheidender Bedeutung ist. Hier benötigen wir also nicht nur einen normierten Vektorraum sondern wirklich einen Banach-Raum. Das Resultat, dass lokal gleichmäßige (Cauchy-)Folgen von stetigen Funktionen wieder einen stetigen Grenzwert besitzen, verallgemeinert zum einen die Situation aus Satz 3.7.10 sowie aus Proposition 3.7.8, iv.). Da uns bereits in dieser sehr speziellen Situation viele interessante Konsequenzen und Folgerungen bekannt sind, erwarten wir zum anderen von Satz 8.5.8 ebenfalls weitreichende Anwendungen.

Nach diesem Nachtrag zu stetigen Funktionen gilt es nun, auch Differenzierbarkeit zu berücksichtigen. Hier wollen wir Abbildungen auf einem offenen Definitionsbereich  $X \subseteq \mathbb{R}^n$  betrachten, die in diesem endlich-dimensionalen Fall  $X$  lokal kompakt sind, siehe nochmals Beispiel 8.5.4. Wäre  $X \subseteq V$  eine nicht-leere offene Teilmenge eines unendlich-dimensionalen Banach-Raums, wäre  $X$  tatsächlich nicht länger lokal kompakt, eine Tatsache die wir hier nicht beweisen wollen. Als Bildbereich verwenden wir nach wie vor einen beliebigen normierten Vektorraum  $W$ , wobei im Hinblick auf Satz 8.5.8 ein Banach-Raum  $W$  interessant sein wird. Mit der  $\mathcal{C}^k$ -Konvergenz aus der eindimensionalen Situation von Definition 5.5.5 als Motivation wollen wir einen Konvergenzbegriff für  $\mathcal{C}^k(X, W)$  etablieren, der die  $k$ -fach stetige Differenzierbarkeit erhält. Hier ist die naheliegende Definition die folgende:

**Definition 8.5.9 ( $\mathcal{C}^k$ -Konvergenz)** Sei  $X \subseteq \mathbb{R}^n$  eine nicht-leere offene Teilmenge und sei  $W$  ein normierter Vektorraum. Sei weiter  $k \in \mathbb{N}_0 \cup \{\infty\}$ .

- i.) Eine Folge  $(f_n)_{n \in \mathbb{N}}$  von  $\mathcal{C}^k$ -Funktionen  $f_i \in \mathcal{C}^k(X, W)$  konvergiert gegen eine  $\mathcal{C}^k$ -Funktion  $f \in \mathcal{C}^k(X, W)$  im  $\mathcal{C}^k$ -Sinne, falls für alle  $\ell \in \mathbb{N}_0$  mit  $0 \leq \ell \leq k$  die Folgen  $(D^\ell f_i)_{i \in \mathbb{N}}$  lokal gleichmäßig auf  $X$  gegen  $D^\ell f$  konvergieren.

ii.) Für  $f \in \mathcal{C}^k(X, W)$  und eine kompakte Teilmenge  $K \subseteq X$  definiert man

$$\|f\|_{K,\ell} = \sup_{x \in K} \|D^\ell f(x)\| \quad (8.5.11)$$

wobei  $\|\cdot\|$  hier die Operatornorm von  $D^\ell f$  bezeichnet.

iii.) Eine Folge  $(f_i)_{i \in \mathbb{N}}$  von  $\mathcal{C}^k$ -Funktionen  $f_i \in \mathcal{C}^k(X, W)$  heißt  $\mathcal{C}^k$ -Cauchy-Folge, falls zu jedem  $\epsilon > 0$ , zu jedem Kompaktum  $K \subseteq X$  und zu jedem  $\ell \in \mathbb{N}_0$  mit  $0 \leq \ell \leq k$  ein  $N \in \mathbb{N}$  existiert, so dass

$$\|f_i - f_j\|_{K,\ell} < \epsilon \quad (8.5.12)$$

für alle  $i, j \geq N$  gilt.

Die  $\mathcal{C}^k$ -Konvergenz nennt man entsprechend auch die lokal gleichmäßige Konvergenz mit allen Ableitungen. In (8.5.11) verwendet man die Operatornorm der höheren Ableitungen, also die Norm von  $L(\mathbb{R}^n, L(\mathbb{R}^n, \dots, L(\mathbb{R}^n, W) \dots))$  mit  $\ell$  Kopien von  $\mathbb{R}^n$ . Alternativ können wir  $D^\ell f$  auch als (nach Satz 8.4.6 beziehungsweise Korollar 8.4.7 symmetrische) stetige  $\ell$ -lineare Abbildung

$$D^\ell f(x) \in L(\mathbb{R}^n, \dots, \mathbb{R}^n; W) \quad (8.5.13)$$

mit Werten in  $W$  auffassen und die von  $L(\mathbb{R}^n, \dots, \mathbb{R}^n; W)$  kommende Norm verwenden.

Da wir explizit  $X \subseteq \mathbb{R}^n$  fordern, können wir die  $\ell$ -te totale Ableitung auch mit Hilfe der  $\ell$ -ten partiellen Ableitungen beschreiben. Hierzu verwenden wir die Äquivalenz von stetiger totaler Differenzierbarkeit und stetigen partieller Differenzierbarkeit nach Satz 8.3.5, entsprechend induktiv erweitert für alle höheren Ableitungen.

**Lemma 8.5.10** Seien  $V_1, \dots, V_\ell$  endlich-dimensionale Vektorräume mit Basen  $\{e_\alpha^{(i)}\}_{i=1, \dots, \dim V_i = n_i}$  für  $i = 1, \dots, \ell$  gegeben. Sei weiter  $W$  ein normierter Vektorraum. Dann ist die Operatornorm von  $L(V_1, L(V_2, \dots, L(V_\ell, W) \dots))$  äquivalent zur Norm

$$\|\Phi\|' = \max \left\{ \|(\dots ((\Phi e_{\alpha_1}^{(1)}) e_{\alpha_2}^{(2)}) \dots) e_{\alpha_\ell}^{(\ell)})\|_W \mid i = 1, \dots, \ell, \alpha_i = 1, \dots, n_i \right\}. \quad (8.5.14)$$

BEWEIS: Für  $\ell = 1$  ist dies gerade die Aussage von Proposition 8.3.4, ii.). Der allgemeine Fall folgt nun induktiv. Wir wissen aus Übung ??, dass äquivalente Normen auf dem Zielbereich einer linearen Abbildung äquivalente Operatornormen liefern. Nach Induktionsannahme ist  $\|\cdot\|'$  auf  $L(V_2, \dots, L(V_\ell, W) \dots)$  äquivalent zur Operatornorm. Dann zeigt eine erneute Anwendung von Proposition 8.3.4, ii.), dass  $\|\cdot\|'$  auf  $L(V_1, L(V_2, \dots, L(V_\ell, W) \dots))$  zur Operatornorm bezüglich  $\|\cdot\|'$  auf  $L(V_2, \dots, L(V_\ell, W) \dots)$  äquivalent ist, also insbesondere äquivalent zur ursprünglichen Operatornorm.  $\square$

Interpretiert man die iterierte lineare Abbildung  $\Phi$  wieder als  $\ell$ -lineare Abbildung mittels des kartesischen Isomorphismus, den wir für  $\ell = 2$  im Detail in Proposition 8.4.2 beziehungsweise Proposition 8.4.4 kennengelernt haben, so erhält man

$$\|\Phi\|' = \max \left\{ \|\tilde{\Phi}(e_{\alpha_1}^{(1)}, \dots, e_{\alpha_\ell}^{(\ell)})\|_W \mid \alpha_i = 1, \dots, n_i, i = 1, \dots, \ell \right\} \quad (8.5.15)$$

für die zugehörige  $\ell$ -lineare Abbildung  $\tilde{\Phi}$ . Wir verwenden dies nun, um die Operatornorm der  $\ell$ -ten totalen Ableitung durch entsprechende Normen der  $\ell$ -ten partiellen Ableitungen zu ersetzen, da diese in konkreten Rechnungen typischerweise einfacher zugänglich sind. Es bietet sich nun folgende Multiindex-Schreibweise an: Für  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$  setzen wir

$$|a| = \alpha_1 + \dots + \alpha_n \quad (8.5.16)$$

sowie

$$\alpha! = \alpha_1! \cdot \alpha_n! \quad (8.5.17)$$

Weiter schreiben wir

$$\frac{\partial^{(\alpha)} f}{\partial x^\alpha} = \frac{\partial^{(\alpha_1)}}{\partial (x^1)^{\alpha_1}} \cdots \frac{\partial^{(\alpha_n)}}{\partial (x^n)^{\alpha_n}} f \quad (8.5.18)$$

für die partiellen Ableitungen einer entsprechend oft differenzierbaren Funktion. Da bei stetigen Ableitungen nach Korollar 8.4.8 die Reihenfolge der Ableitungen in (8.5.18) unerheblich ist, stellt dies eine sinnvolle Schreibweise dar. Ist nun  $f \in \mathcal{C}^k(X, W)$  und  $\ell \in \mathbb{N}_0$  mit  $\ell \leq k$ , so gilt

$$(D^\ell f(x))(e_{i_1}, \dots, e_{i_\ell}) = \frac{\partial^\ell f}{\partial x^{i_1} \dots \partial x^{i_\ell}}(x), \quad (8.5.19)$$

da Auswerten der totalen Ableitung auf Basisvektoren der kanonischen Basis  $e_1, \dots, e_n \in \mathbb{R}^n$  die entsprechenden partiellen Ableitungen liefert. Damit folgt

$$\max \left\{ \|(D^\ell f)(x)(e_{i_1}, \dots, e_{i_\ell})\|_W \mid i_1, \dots, i_\ell = 1, \dots, n \right\} = \max_{|\alpha| \leq \ell} \left\| \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x) \right\|_W. \quad (8.5.20)$$

Im Hinblick auf (8.5.20) definieren wir nun die Größe

$$p_{K,\ell}(f) = \max_{\substack{x \in K \\ |\alpha| \leq \ell}} \left\| \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x) \right\|_W \quad (8.5.21)$$

für jedes Kompaktum  $K \subseteq X$ . Da die in (8.5.11) verwendete Operatornorm zur Norm (8.5.14) und damit zu (8.5.20) äquivalent ist, erhalten wir folgendes Korollar:

**Korollar 8.5.11** *Seien  $X \subseteq \mathbb{R}^n$  eine nicht-leere offene Teilmenge und  $W$  ein normierter Vektorraum. Sei weiter  $(f_i)_{i \in \mathbb{N}}$  eine Folge von  $\mathcal{C}^k$ -Funktionen  $f_i \in \mathcal{C}^k(X, W)$  mit einem  $k \in \mathbb{N}_0 \cup \{\infty\}$ .*

*i.) Die Folge  $(f_i)_{i \in \mathbb{N}}$  konvergiert genau dann gegen  $f \in \mathcal{C}^k(X, W)$  im  $\mathcal{C}^k$ -Sinne, wenn für alle Kompakta  $K \subseteq X$  und alle  $\ell \in \mathbb{N}_0$  mit  $\ell \leq k$*

$$\lim_{i \rightarrow \infty} p_{K,\ell}(f_i - f) = 0 \quad (8.5.22)$$

*gilt.*

*ii.) Die Folge  $(f_i)_{i \in \mathbb{N}}$  ist genau dann eine  $\mathcal{C}^k$ -Cauchy-Folge, wenn es zu jedem  $\epsilon > 0$  und zu jedem Kompaktum und zu jedem  $\ell \in \mathbb{N}_0$  mit  $\ell \leq k$  ein  $N \in \mathbb{N}$  existiert, so dass für alle  $i, j \geq N$*

$$p_{K,\ell}(f_i - f_j) < \epsilon \quad (8.5.23)$$

*gilt.*

Wir kommen nun zum Hauptresultat dieses Abschnitts für  $\mathcal{C}^k$ -Konvergenz:

**Satz 8.5.12 (Vollständigkeit von  $\mathcal{C}^k(X, W)$ )** *Seien  $X \subseteq \mathbb{R}^n$  eine nicht-leere offene Teilmenge und  $W$  ein Banach-Raum sowie  $k \in \mathbb{N}_0 \cup \{\infty\}$ . Eine Folge  $(f_i)_{i \in \mathbb{N}}$  von  $\mathcal{C}^k$  Funktionen  $f_i \in \mathcal{C}^k(X, W)$  konvergiert genau dann im  $\mathcal{C}^k$  Sinne, wenn  $(f_i)_{i \in \mathbb{N}}$  eine  $\mathcal{C}^k$ -Cauchy-Folge ist.*

**BEWEIS:** Den Fall  $k = 0$  haben wir in deutlich größerer Allgemeinheit bereits in Satz 8.5.8 gezeigt, da  $X \subseteq \mathbb{R}^n$  ja lokal kompakt ist. Wir betrachten zunächst  $k = 1$ . Da  $(f_i)_{i \in \mathbb{N}}$  eine  $\mathcal{C}^1$ -Cauchy-Folge ist, konvergiert  $f_i \rightarrow f$  im  $\mathcal{C}^0$ -Sinne, also lokal gleichmäßig, gegen eine stetige Funktion  $f \in \mathcal{C}^0(X, W)$  nach Satz 8.5.8. Weiter sind auch die partiellen Ableitungen  $(\frac{\partial f_i}{\partial x^\alpha})_{i \in \mathbb{N}}$  lokal gleichmäßige Cauchy-Folgen, womit sie ebenfalls im  $\mathcal{C}^0$ -Sinne gegen stetige Funktionen  $g_\alpha \in \mathcal{C}^0(X, W)$  für  $\alpha = 1, \dots, n$



konvergieren, also  $g_\alpha = \lim_{i \rightarrow \infty} \frac{\partial f_i}{\partial x^\alpha}$ . Dies erreichen wir ebenfalls nach Satz 8.5.8. Es bleibt zu zeigen, dass  $f$  sogar differenzierbar ist mit

$$g_\alpha = \frac{\partial f}{\partial x^\alpha} \quad (*)$$

für alle  $\alpha = 1, \dots, n$ . Wenn wir die partielle Differenzierbarkeit von  $f$  mit  $(*)$  gezeigt haben, hat  $f$  automatisch stetige partielle Ableitungen, da  $g_\alpha \in \mathcal{C}^0(X, W)$  ja stetig ist. Daher ist  $f \in \mathcal{C}^1(X, W)$  nach Satz 8.3.5. Es ist also nur  $(*)$  zu zeigen. Sei also  $1 \leq \alpha \leq n$  und  $x_0 \in X$ . Wir betrachten analog zum Beweis von Proposition 5.5.2 die Hilfsfunktionen

$$F_i(t) = \begin{cases} \frac{f_i(x_0 + te_\alpha) - f_i(x_0)}{t} & \text{für } t \neq 0 \\ \frac{\partial f_i}{\partial x^\alpha}(x_0) & \text{für } t = 0 \end{cases}$$

und

$$F(t) = \frac{f(x_0 + te_\alpha) - f(x_0)}{t} \quad \text{für } t \neq 0,$$

für  $t \in [-t_0, t_0] \subseteq \mathbb{R}$  mit  $t_0 > 0$  derart, dass  $x_0 + te_\alpha \in X$  für alle  $t \in [-t_0, t_0]$ . Da  $X$  offen ist, können wir solch ein  $t_0$  finden. Wir wählen nun ein  $\epsilon > 0$  aus und finden mit der Cauchy-Eigenschaft auf dem Kompaktum  $K = \{x + te_\alpha \mid t \in [-t_0, t_0]\} \subseteq X$  einen Index  $N \in \mathbb{N}$  derart, dass

$$\|f_i - f_j\|_K \leq \frac{\epsilon}{2} \quad (*)$$

und

$$\left\| \frac{\partial f_i}{\partial x^\alpha} - \frac{\partial f_j}{\partial x^\alpha} \right\|_K < \min \left\{ \frac{\epsilon}{2}, \frac{\epsilon}{4t_0} \right\}$$

für alle  $i, j \geq N$ . Für  $t, t' \in [-t_0, t_0]$  erhalten wir dann mit dem Schrankensatz für die Funktion  $f_i - f_j$

$$\begin{aligned} & \|f_i(x_0 + te_\alpha) - f_j(x_0 + te_\alpha) - f_i(x_0 + t'e_\alpha) + f_j(x_0 + t'e_\alpha)\|_W \\ & \leq \left\| \frac{\partial f_i}{\partial x^\alpha} - \frac{\partial f_j}{\partial x^\alpha} \right\|_K |t - t'| < \frac{\epsilon}{2}. \end{aligned} \quad (**)$$

Wir betrachten nun  $t \in [-t_0, t_0] \setminus \{0\}$  und erhalten für die Differenzenquotienten

$$\begin{aligned} \|F_i(t) - F_j(t)\|_W &= \frac{1}{|t|} \|f_i(x_0 + te_\alpha) - f_i(x_0) - f_j(x_0 + te_\alpha) + f_j(x_0)\| \\ &< \frac{\epsilon}{2} \end{aligned}$$

mit  $(**)$ . Für  $t = 0$  erhalten wir direkter

$$\|F_i(0) - F_j(0)\|_W = \left\| \frac{\partial f_i}{\partial x^\alpha}(x_0) - \frac{\partial f_j}{\partial x^\alpha}(x_0) \right\|_W < \frac{\epsilon}{2}.$$

Somit gilt

$$\|F_i - F_j\|_{[-t_0, t_0]} < \frac{\epsilon}{2}$$

auf dem ganzen Intervall  $[-t_0, t_0]$ . Dies zeigt, dass  $F_i$  auf  $[-t_0, t_0]$  eine gleichmäßige Cauchy-Folge von stetigen Funktionen  $f_i \in \mathcal{C}([-t_0, t_0], W)$  ist. Daher konvergiert  $F_i$  im  $\mathcal{C}^0$ -Sinne gegen eine stetige Grenzfunktion auf ganz  $[-t_0, t_0]$ . Da für  $t \neq 0$  punktweise

$$\begin{aligned} \lim_{i \rightarrow \infty} F_i(t) &= \lim_{i \rightarrow \infty} \frac{f_i(x_0 + te_\alpha) - f_i(x_0)}{t} \\ &= \frac{f(x_0 + te_\alpha) - f(x_0)}{t} \end{aligned}$$

$$= F(t)$$

gilt, ist die stetige Grenzfunktion  $\lim_{i \rightarrow \infty} F_i \in \mathcal{C}^0([-t_0, t_0], W)$  eine stetige Fortsetzung von  $F$  in den Punkt  $t = 0$ . Damit folgt aber aus der Stetigkeit dieser Fortsetzung

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{f(x_0 + te_\alpha) - f(x_0)}{t} &= \lim_{i \rightarrow \infty} F_i(0) \\ &= \lim_{i \rightarrow \infty} \frac{\partial f_i}{\partial x^\alpha}(x_0) \\ &= g_\alpha(x_0). \end{aligned}$$

Dies zeigt nun zum einen, dass  $f$  bei  $x_0$  partiell differenzierbar ist. Zum anderen folgt (\*) da  $x_0 \in X$  beliebig war. Dies schließt den Beweis von Fall  $k = 1$  ab. Der allgemeine Fall  $k \geq 2$  ist nun einfach. Wir wenden den Fall  $k = 1$  auf die  $(k - 1)$ -te Ableitung der Folge  $(f_i)_{i \in \mathbb{N}}$  an und erhalten dann die stetige Differenzierbarkeit dieser  $(k - 1)$ -ten Ableitungen, womit die  $k$ -fache stetige Differenzierbarkeit gezeigt ist.  $\square$

**Bemerkung 8.5.13** Im Falle  $W = \mathbb{K}$  können wir mit Hilfe des Hauptsatzes einen sehr viel einfacheren Beweis führen, siehe Übung ???. Erweitert man den Integralbegriff auf Riemann-Integrale von beispielsweise stetigen Funktionen mit Werten in Banach-Räumen, so wird dieser einfachere Beweis auch für die allgemeine Situation von Satz 8.5.12 zugänglich.

## 8.6 Übungen

## Kapitel 9

# Implizite Funktionen und Untermannigfaltigkeiten

In diesem Kapitel wollen wir differenzierbare Abbildungen weiter untersuchen, wobei nun neben den analytischen Eigenschaften auch geometrische Eigenschaften eine Rolle spielen sollen. Die wesentliche Idee dabei ist, die analytischen Eigenschaften einer Abbildung

$$f: X \longrightarrow W \tag{9.0.1}$$

mit den geometrischen Eigenschaften ihres Bildes

$$\operatorname{im} f = f(X) \subseteq W \tag{9.0.2}$$

und ihres Graphen

$$\operatorname{graph}(f) \subseteq X \times W \tag{9.0.3}$$

in Verbindung zu bringen. Die im Zuge dieser Fragestellungen erlangten Techniken werden wir dann ihrerseits dazu nutzen können, allgemeine Teilmengen vom Zielraum  $W$  zu betrachten, die sich zumindest lokal noch als Bild (9.0.2) einer hinreichend einfachen Abbildung  $f$  schreiben lassen. Wir beginnen dazu, die (lokale) Umkehrbarkeit von  $\mathcal{C}^k$ -Funktionen zu studieren, was zum ersten zentralen Resultat des Kapitels führt. Allgemeiner wollen wir neben (lokalen) Diffeomorphismen auch die Klassen der Submersionen und Inversionen kennen lernen, welche zum einen anhand ihrer Ableitung charakterisiert werden, andererseits aber einfache lokale Normalformen besitzen, die durch geeignete lokale Koordinatenwechsel erzielt werden. Eine zentrale Anwendung ist dann der Satz zu impliziten Funktionen, der in vielen Bereichen der Mathematik entscheidend ist. Nach diesen lokalen Eigenschaften von differenzierbaren Funktionen betrachten wir dann die globale Situation und definieren Untermannigfaltigkeiten in einem normierten Vektorraum. Niveaulächen einer skalaren Funktion werden hier die wesentlichen Beispiele sein, die zudem auch unmittelbare Anwendungen weit jenseits der Mathematik besitzen.

### 9.1 Lokale Umkehrbarkeit

Wir wollen wieder differenzierbare Abbildungen der Form

$$f: X \longrightarrow W \tag{9.1.1}$$

von einer offenen Teilmenge  $X \subseteq V$  eines Banach-Raums mit Werten in einem anderen Banach-Raum untersuchen. Das erste Ziel ist, ein infinitesimales Kriterium für die (lokale) Invertierbarkeit zu finden, also ein Kriterium für die Ableitung von  $f$ , welches zumindest die lokale Invertierbarkeit garantiert.

Als ersten Schritt betrachten wir die lineare Situation im Detail. Hierzu bieten sich folgende Begriffsbildungen an:

**Definition 9.1.1** ( $\text{GL}(V)$  und  $\text{Iso}(V, W)$ ) *Seien  $V$  und  $W$  Banach-Räume.*

- i.) Eine stetige lineare Abbildung  $\Phi: V \rightarrow W$  heißt Isomorphismus, falls  $\Phi$  eine stetige Umkehrabbildung  $\Phi^{-1}: W \rightarrow V$  besitzt.*
- ii.) Die Menge der Isomorphismen von  $V$  nach  $W$  wird mit*

$$\text{Iso}(V, W) = \{\Phi: V \rightarrow W \mid \Phi \text{ ist ein Isomorphismus}\} \quad (9.1.2)$$

*bezeichnet.*

- iii.) Die Menge der Selbst-Isomorphismen von  $V$  nach  $V$  wird mit*

$$\text{GL}(V) = \text{Aut}(V) = \text{Iso}(V, V) \quad (9.1.3)$$

*bezeichnet.*

## 9.2 Implizite Funktionen

## 9.3 Submersionen und Immersionen

## 9.4 Übungen

# Literaturverzeichnis

- [1] HALMOS, P. R.: *Naive Mengenlehre*. Vandenhoeck & Ruprecht, Göttingen, 4. edition, 1976. 4
- [2] KELLEY, J. L.: *General Topology. Graduate Texts in Mathematics* no. **27**. Springer-Verlag, New York, Heidelberg, Berlin, 1955. 299
- [3] LANG, S.: *Algebra*. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 3. edition, 1997. 6
- [4] OEIS FOUNDATION INC.: *Decimal expansion of  $e$* . Published electronically at <http://oeis.org/A001113>, 2023. Entry A001113 in The On-Line Encyclopedia of Integer Sequences. 157
- [5] OEIS FOUNDATION INC.: *Decimal expansion of Euler's constant (or the Euler-Mascheroni constant),  $\gamma$* . Published electronically at <http://oeis.org/A001620>, 2023. Entry A001620 in The On-Line Encyclopedia of Integer Sequences. 268
- [6] QUERENBURG, B. V.: *Mengentheoretische Topologie*. Springer-Verlag, Berlin, Heidelberg, New York, 3. edition, 2001. 299
- [7] WALDMANN, S.: *Topology. An Introduction*. Springer-Verlag, Heidelberg, New York, 2014. 91, 109, 299
- [8] WALDMANN, S.: *Lineare Algebra I. Die Grundlagen für Studierende der Mathematik und Physik*. Springer-Verlag, Berlin, 2. edition, 2021. 4, 6, 143, 206, 258
- [9] WALDMANN, S.: *Lineare Algebra II. Anwendungen und Konzepte für Studierende der Mathematik und Physik*. Springer-Verlag, Berlin, 2. edition, 2022. 325, 328



# Symbolverzeichnis





# Index

|                             |                     |                                    |          |
|-----------------------------|---------------------|------------------------------------|----------|
| Abelsche Gruppe .....       | 6                   | Assoziativgesetz .....             | 5        |
| Abgeschlossene Kugel .....  | 64                  | Asymptotische Entwicklung .....    | 224      |
| in metrischem Raum .....    | 280                 |                                    |          |
| Abgeschlossene Menge .....  | 64, 294             | Banach-Raum .....                  | 291      |
| in metrischem Raum .....    | 280                 | Unterraum .....                    | 292      |
| relativ .....               | 105                 | Banachscher Fixpunktsatz .....     | 292      |
| Ableitung .....             | 180                 | Bernoulli-Ungleichung .....        | 14       |
| totale .....                | 316                 | Berührungspunkt .....              | 95       |
| Zwischenwertsatz .....      | 195                 | in topologischem Raum .....        | 296      |
| Abschätzung .....           | 14                  | Beschränkte Funktionen .....       | 124      |
| Abschluss .....             | 97                  | Banach-Raum .....                  | 291      |
| in topologischem Raum ..... | 296                 | Beschränkte Funktionen             |          |
| Absolutbetrag .....         | <i>siehe</i> Betrag | Vollständigkeit .....              | 291      |
| als Norm .....              | 276                 | Beschränkte Teilmenge .....        | 18       |
| Abzählbarkeit .....         | 44                  | Betrag .....                       | 15       |
| von $\mathbb{Q}$ .....      | 44, 93              | Approximation durch Polynome ..... | 207      |
| Abzählung .....             | 44                  | komplexe Zahl .....                | 36       |
| Addition .....              | 5                   | Stetigkeit .....                   | 110      |
| Additionstheoreme           |                     | Bijektion .....                    | 46       |
| Exponentialfunktion .....   | 153                 | Binomialkoeffizient .....          | 10, 38   |
| Sinus und Kosinus .....     | 164                 | verallgemeinerter .....            | 222      |
| Tangens und Kotangens ..... | 171                 | Binomialsatz                       |          |
| Äquivalenz von Normen ..... | 309                 | verallgemeinerter .....            | 222      |
| Allquantor .....            | 46                  | Binomischer Lehrsatz .....         | 10       |
| Approximation               |                     | $\mathcal{C}^1$ -Abbildung .....   | 330      |
| höherer Ordnung .....       | 180                 | $\mathcal{C}^1$ -Funktion .....    | 332      |
| konstante .....             | 177                 | Cauchy-Folge .....                 | 60       |
| lineare .....               | 179                 | in metrischem Raum .....           | 290      |
| Archimedischer Körper ..... | 15                  | Konvergenz .....                   | 60       |
| Arcuskosinus .....          | 173                 | Cauchy-Produkt .....               | 88, 154  |
| Differenzierbarkeit .....   | 187                 | formales .....                     | 144      |
| Arcuskotangens .....        | 173                 | Cauchy-Reihe .....                 | 74       |
| Arcussinus .....            | 173                 | Cauchy-Schwarz-Ungleichung .....   | 206      |
| Differenzierbarkeit .....   | 187                 | für Riemann-Integrale .....        | 257      |
| Arcustangens .....          | 173                 | Cauchy-Vollständigkeit .....       | 62       |
| Differenzierbarkeit .....   | 187                 | Charakteristik .....               | 8, 37    |
| Taylor-Entwicklung .....    | 273                 | Charakteristische Funktion .....   | 111      |
| Argument .....              | 175                 | $\mathcal{C}^k$ -Funktion .....    | 199      |
| Stetigkeit .....            | 175                 |                                    |          |
| Arithmetische Mittel .....  | 162                 | Darboux-Integral .....             | 237, 240 |

- Darboux-integrierbar ..... 237, 240  
 Dedekindscher Schnitt ..... 19  
 Dekadischer Logarithmus ..... 158  
 $\delta$ -Funktional  
     Stetigkeit ..... 289  
 Differenzenquotient ..... 177  
 Differenzierbarkeit ..... 180  
     bei Häufungspunkt ..... 180  
     Produktregel ..... 185  
     Quotientenregel ..... 185  
     Rechenregeln ..... 184  
     Verkettung ..... 186  
     von links ..... 180  
     von rechts ..... 180  
 Dirichlet-Funktion ..... 111  
 Distanzfunktion ..... *siehe* Metrik  
 Distributivgesetz ..... 5  
 Divergenz  
     gegen  $\pm\infty$  ..... 50  
     von Folgen ..... 46  
 Dreiecksungleichung ..... 15, 278  
  
 e ..... *siehe* Euler-Zahl  
 Einheitskreis ..... 40, 173  
 Einheitskreislinie ..... 163  
 Einsteinsche Summenkonvention ..... 325  
 Euklidische Norm ..... 276  
 Euler-Mascheroni Konstante ..... 268  
 Euler-Zahl ..... 156  
     Transzendenz ..... 157  
 Existenzquantor ..... 46  
 Exponentialfunktion ..... 131, 153  
     Differenzierbarkeit ..... 181, 185  
     Glattheit ..... 200  
     Monotonie ..... 157  
     Stetigkeit ..... 110  
 Exponentialreihe ..... 153  
 Extremum  
     globales ..... 189  
     lokales ..... 189, 326  
  
 Fakultät ..... 10, 38  
 Fibonacci-Folge ..... 44  
 Folge ..... 43  
     beschränkt ..... 48  
     Divergenz ..... 46, 93  
     Grenzwert ..... 46  
     Häufungspunkt ..... 68  
     konstante ..... 43  
     konvergent ..... 53  
     monoton fallend ..... 62  
     monoton wachsend ..... 62  
     streng monoton fallend ..... 62  
     streng monoton wachsend ..... 62  
     Supremumsnorm ..... 48  
     Umordnung ..... 46  
     unbeschränkt ..... 48  
     von Funktionen ..... 138  
 Folgenkompaktheit ..... 118  
 Folgenkonvergenz ..... 46, 280  
     Teilfolge ..... 48  
     Umordnung ..... 50  
 Folgenraum  
      $c$  ..... 53  
      $c_0$  ..... 53  
      $\ell^\infty$  ..... 51  
 Fortsetzung ..... 130  
     stetige ..... 131  
 Fréchet-Differenzierbarkeit ..... *siehe* Totale  
     Differenzierbarkeit  
 Fundamentalsatz der Algebra ..... 37  
 Funktion ..... 41  
     Partielle Ordnung ..... 42  
     punktweise Rechenoperationen ..... 41  
 Funktionalnorm ..... 277  
 Funktionenfolge ..... 138  
  
 $\gamma$  ..... *siehe* Euler-Mascheroni-Konstante  
 Gamma-Funktion ..... 268  
 Gateaux-Ableitung ..... 324  
 Gateaux-Differenzierbarkeit ..... 324  
 Geometrische Reihe ..... 74  
 Geometrische Summe ..... 11  
 Geometrisches Mittel ..... 162  
 Gerichtete Menge ..... 88  
     Maximum ..... 98  
 Glatte Funktion ..... 199  
 Glattheit an einem Punkt ..... 227  
 Gleichmäßige Konvergenz ..... 138  
 Gleichmäßige Stetigkeit ..... 133  
     auf Kompakta ..... 135  
 Grenzfunktion ..... 138  
 Grenzwert ..... 46  
     als lineares Funktional ..... 55  
     Eindeutigkeit ..... 47, 280  
     in metrischem Raum ..... 280  
     Rechenregeln ..... 53  
     Vertauschung ..... 92  
     von Netz ..... 90

- Häufungspunkt  
  Abgeschlossenheit ..... 69  
  einer Folge ..... 68  
  einer Teilmenge ..... 95  
  in metrischem Raum ..... 281  
  uneigentlicher ..... 70  
Harmonische Reihe ..... 74, 267  
Hauptsatz ..... 251  
Hausdorff-Topologie ..... 299  
Hilberts Hotel ..... 93  
Hölder-Ungleichung ..... 205  
  für Riemann-Integrale ..... 257  
Homöomorphismus ..... 128  
  von Intervallen ..... 130  
Imaginäre Einheit ..... 33  
Imaginärteil ..... 35  
Infimum ..... 18  
Intervall ..... 20  
  abgeschlossenes ..... 20  
  Grenzen ..... 20  
  kompakt ..... 113  
  Länge ..... 20  
  offenes ..... 20, 65  
  zusammenhängend ..... 119  
Intervallschachtelung ..... 21  
Intervallschachtelungsprinzip ..... 21  
Inverses ..... 5  
Isolierter Punkt ..... 97  
Isometrie ..... 278  
  
Jacobi-Determinante ..... 328  
Jacobi-Matrix ..... 328  
Jensen-Ungleichung  
  für Riemann-Integrale ..... 259  
Jensensche Ungleichung ..... 206  
  
Kartesisches Produkt  
  kompakt ..... 116  
  Topologie ..... 297  
Kettenregel ..... 186, 320  
Körper ..... 5  
  angeordneter ..... 12  
  komplexe Zahlen .... *siehe* Komplexe Zahlen  
  reelle Zahlen ..... *siehe* Reelle Zahlen  
  vollständig ..... 21  
Körpermorphismus ..... 37  
Koinzidenzmenge ..... 151  
Kommutativer Ring ..... 6  
Kommutativgesetz ..... 5  
Kompaktheit ..... 113, 116, 300  
  abgeschlossenes Intervall ..... 113  
  stetiges Bild ..... 123, 301  
  topologische ..... 113, 300  
Komplexe Konjugation ..... 35  
  Stetigkeit ..... 110  
Komplexe Zahlen ..... 33  
  algebraisch abgeschlossen ..... 37  
  Vollständigkeit ..... 37  
Komplexe Zahlenebene ..... 35  
Konkave Funktion ..... 203  
Kontraktion ..... 284, 292  
Konvergenz  
  absolute ..... 79  
  gleichmäßige ..... 138  
  in metrischem Raum ..... 280  
  lokal gleichmäßige ..... 138  
  punktweise ..... 138  
  von  $\mathcal{C}^k$ -Funktionen ..... 210  
  von Folgen ..... 46  
  von Funktionswerten ..... 99, 100  
  von glatten Funktionen ..... 210  
  von Netzen ..... 90  
Konvergenzradius ..... 144  
Konvexe Funktion ..... 203  
Konvexität ..... 194  
Kosinus ..... 163  
  Additionstheoreme ..... 164  
  Differenzierbarkeit ..... 185  
  Glattheit ..... 200  
  Monotonie ..... 170  
  Periodizität ..... 168  
  Stetigkeit ..... 164  
Kotangens ..... 170  
  Additionstheoreme ..... 171  
  Glattheit ..... 200  
  Monotonie ..... 172  
  Periodizität ..... 171  
  Stetigkeit ..... 171  
Kreisscheibe  
  Fläche ..... 254  
  offene ..... 65  
Kreiszahl ..... 168  
Kurve ..... *siehe* Weg  
Kurvendiskussion ..... 202  
Lagrange-Restglied ..... 219  
Leibniz-Kriterium ..... 77  
Leibniz-Regel ..... 184, 322  
Limes ..... *siehe* Grenzwert

|                                     |          |  |                        |
|-------------------------------------|----------|--|------------------------|
| Limes inferior .....                | 70       | Neutrales Element .....                              | 5                      |
| Limes superior .....                | 70       | Norm .....   | 275                    |
| Lineare Näherung .....              | 179      | Äquivalenz .....                                     | 308                    |
| Lipschitz-Konstante .....           | 284      | auf direkter Summe .....                             | 277                    |
| Lipschitz-Stetigkeit .....          | 284      | auf stetigem Dualraum <i>siehe</i> Funktionalnorm    |                        |
| Logarithmus .....                   | 158      | auf Tensorprodukt .....                              | 277                    |
| Differenzierbarkeit .....           | 187      | Normierter Raum .....                                | 275                    |
| Glattheit .....                     | 200      | Unterraum .....                                      | 276                    |
| konkav .....                        | 204      | Normierter Vektorraum . <i>siehe</i> Normierter Raum |                        |
| Lokal beschränkte Funktionen .....  | 124      | Nullfolge .....                                      | 47, 53                 |
| Lokal gleichmäßige Konvergenz ..... | 138      |  |                        |
| Majorante .....                     | 80       | Oberintegral .....                                   | 237                    |
| Maximum                             |          | Obersumme .....                                      | 237                    |
| globales .....                      | 189      | Offene Kugel .....                                   | 64                     |
| lokales .....                       | 189      | in metrischem Raum .....                             | 280                    |
| Maximumsnorm .....                  | 276      | Offene Menge .....                                   | 64, 294                |
| Metrik .....                        | 278      | in metrischem Raum .....                             | 280                    |
| diskrete .....                      | 279, 312 | relativ .....  | 105                    |
| Einschränkung .....                 | 279      | Offene Überdeckung .....                             | 112                    |
| Metrischer Raum .....               | 278      | Operatornorm .....                                   | 277, 286–288           |
| abgeschlossene Kugel .....          | 280      | induzierte Topologie .....                           | 330                    |
| abgeschlossene Menge .....          | 280      | Ordnung .....  | 89                     |
| Cauchy-Folge .....                  | 290      | Ordnungsrelation .....                               | 12                     |
| Häufungspunkt .....                 | 281      |  |                        |
| Lipschitz-Stetigkeit .....          | 284      | $p$ -Norm .....                                      | 276                    |
| offene Kugel .....                  | 280      | Partialsumme .....                                   | 73                     |
| offene Menge .....                  | 280      | Nullfolge .....                                      | 75                     |
| Stetigkeit .....                    | 283      | Partielle Ableitung .....                            | 326                    |
| Topologie .....                     | 294      | stetige .....  | 332                    |
| Umgebung .....                      | 280      | Partielle Ordnung .....                              | 89                     |
| Vollständigkeit .....               | 290      | Partition .....                                      | <i>siehe</i> Zerlegung |
| Minimum                             |          | Pascalsches Dreieck .....                            | 38                     |
| globales .....                      | 189      | Permutation .....                                    | 39                     |
| lokales .....                       | 189      | $\pi$ .....  | <i>siehe</i> Kreiszahl |
| Minkowski-Ungleichung .....         | 205      | Polarisierungsidentität .....                        | 248                    |
| für Riemann-Integrale .....         | 258      | Polarkoordinaten .....                               | 175                    |
| Mittelwert .....                    | 260      | Polynom .....  | 37, 101                |
| Mittelwertsatz .....                | 192      | Differenzierbarkeit .....                            | 185                    |
| Momentangeschwindigkeit .....       | 323      | Glattheit .....                                      | 200                    |
| Monotone Funktion .....             | 121      | Nullstelle .....                                     | 126                    |
| Stetigkeit .....                    | 122      | Stetigkeit .....                                     | 109                    |
| via Ableitung .....                 | 193      | Positive Elemente .....                              | 12                     |
| Multiplikation .....                | 5        | Potenz   |                        |
| Natürlicher Logarithmus .....       | 158      | komplexe .....                                       | 156                    |
| Negative Elemente .....             | 12       | rationale .....                                      | 28                     |
| Netz .....                          | 90, 98   | reelle .....   | 30                     |
| Grenzwert .....                     | 90       | Potenzfunktion                                       |                        |
| Konvergenz .....                    | 90, 233  | Differenzierbarkeit .....                            | 188                    |
|                                     |          | Stetigkeit .....                                     | 109                    |
|                                     |          | Potenzgesetze .....                                  | 30, 153                |

|                                    |                                 |  |               |
|------------------------------------|---------------------------------|--|---------------|
| ganzzahliger Exponent.....         | 9                               | Funktionalnorm.....                        | 289           |
| Potenzmenge.....                   | 294                             | gleichmäßige Konvergenz.....               | 271           |
| Potenzreihe.....                   | 143                             | Mittelwertsatz.....                        | 260           |
| asymptotische Entwicklung.....     | 224                             | Skalarprodukt.....                         | 258           |
| formale.....                       | 143                             | Stetigkeit als Funktional.....             | 289           |
| Glattheit.....                     | 213                             | Riemann-integrierbare Funktionen.....      | 233, 240      |
| Integration.....                   | 273                             | Banach-Raum.....                           | 291           |
| Konvergenz.....                    | 144                             | beschränkt.....                            | 236           |
| Konvergenzradius.....              | 144                             | Vollständigkeit.....                       | 291           |
| lokal gleichmäßige Konvergenz..... | 146                             | Riemann-Summe.....                         | 232           |
| Stammfunktion.....                 | 273                             | Riemannsche $\zeta$ -Funktion.....         | 77            |
| Stetigkeit.....                    | 146                             | Glattheit.....                             | 212           |
| Produktmetrik.....                 | 313                             | Riemannscher Umordnungssatz.....           | 83            |
| Produktregel.....                  | 322                             | Sandwich-Prinzip.....                      | 49            |
| Produkttopologie.....              | 297, 301                        | Satz                                       |               |
| Produktzeichen.....                | 9                               | Bolzano-Weierstraß.....                    | 63            |
| Pull-back.....                     | 42                              | Heine-Borel.....                           | 118, 306      |
| Punkt vor Strich.....              | 7                               | l'Hospital.....                            | 197           |
| Punktweise Konvergenz.....         | 138                             | Rolle.....                                 | 190           |
| Quadratwurzel.....                 | 27                              | Tikhonov.....                              | 116, 301      |
| Raten.....                         | 251                             | Schrankensatz.....                         | 194, 328      |
| Rationale Funktion.....            | 101                             | Sekante.....                               | 177           |
| Differenzierbarkeit.....           | 185                             | Sinus.....                                 | 163           |
| Glattheit.....                     | 200                             | Additionstheoreme.....                     | 164           |
| Stetigkeit.....                    | 109                             | Differenzierbarkeit.....                   | 185           |
| Realteil.....                      | 35                              | Glattheit.....                             | 200           |
| Reelle Zahlen.....                 | 24                              | Monotonie.....                             | 170           |
| Regel von l'Hospital.....          | 197                             | Periodizität.....                          | 168           |
| Reihe.....                         | 73                              | Stetigkeit.....                            | 164           |
| absolute Konvergenz.....           | 79, 83                          | Sinus cardinalis.....                      | 196, 265      |
| alternierend.....                  | 77                              | Skalarprodukt.....                         | 258           |
| Dreiecksungleichung.....           | 79                              | Stammfunktion.....                         | 250, 251      |
| Integralkriterium.....             | 266                             | Standardtopologie von $\mathbb{R}^n$ ..... | 305           |
| Konvergenz.....                    | 73                              | Stetige Funktion                           |               |
| Majorantenkriterium.....           | 80                              | beschränkt.....                            | 124           |
| Produkt.....                       | 87                              | Bilder von Kompakta.....                   | 123           |
| Quotientenkriterium.....           | 81                              | lokal beschränkt.....                      | 124           |
| unbedingte Konvergenz.....         | 82, 83                          | Minimum und Maximum.....                   | 125, 301      |
| Wert.....                          | 73                              | Monotonie.....                             | 122           |
| Wurzelkriterium.....               | 80                              | Nullstellen.....                           | 126           |
| Reihenmultiplikation.....          | <i>siehe</i> Cauchy-Produkt     | Verkettung.....                            | 123           |
| Relativtopologie.....              | <i>siehe</i> Unterraumtopologie | Zwischenwertsatz.....                      | 126, 303      |
| Richtung.....                      | 88                              | Stetige Funktionen                         |               |
| kanonische.....                    | 89                              | Banach-Raum.....                           | 291           |
| triviale.....                      | 89                              | Vollständigkeit.....                       | 291           |
| Richtungsableitung.....            | 324                             | Stetige Homomorphismen.....                | 277           |
| stetige.....                       | 332                             | Stetiger Dualraum.....                     | 277           |
| Riemann-Integral.....              | 233, 240                        | Stetigkeit.....                            | 104, 107, 298 |
|                                    |                                 | Addition.....                              | 310           |

|                                     |          |  |        |
|-------------------------------------|----------|--|--------|
| der Diagonalen .....                | 310      | Überabzählbarkeit .....                              | 44     |
| der Umkehrabbildung .....           | 127      | von $\mathbb{C}$ .....                               | 45     |
| $\epsilon\delta$ -Stetigkeit .....  | 105      | von $\mathbb{R}$ .....                               | 44     |
| Folgenstetigkeit .....              | 105      | Überdeckungskompaktheit ... <i>siehe</i> Kompaktheit |        |
| Fortsetzung .....                   | 131      | Umgebung .....                                       | 64     |
| in metrischen Räumen .....          | 283      | in metrischem Raum .....                             | 280    |
| lineare Abbildung .....             | 287      | in topologischem Raum .....                          | 295    |
| linksseitig .....                   | 148      | kompakte .....                                       | 307    |
| lokal gleichmäßige Konvergenz ..... | 142      | konvexe .....  | 307    |
| Multiplikation .....                | 310      | relativ .....  | 105    |
| Netzstetigkeit .....                | 105      | wegzusammenhängende .....                            | 307    |
| rechtsseitig .....                  | 148      | Umkehrfunktion                                       |        |
| und Wegzusammenhang .....           | 303      | Differenzierbarkeit .....                            | 187    |
| und Zusammenhang .....              | 125, 303 | Monotonie .....                                      | 160    |
| von Produktabbildungen .....        | 311      | Uneigentliches Riemann-Integral .....                | 261    |
| Streng monoton .....                | 121      | absolute Konvergenz .....                            | 264    |
| Summenzeichen .....                 | 9        | Cauchy-Kriterium .....                               | 264    |
| Supremum .....                      | 18       | Majorantenkriterium .....                            | 265    |
| Supremumsnorm .....                 | 48, 276  | Ungleichung  |        |
| Tangens .....                       | 170      | arithmetische und geometrische Mittel ...            | 162    |
| Additionstheoreme .....             | 171      | Cauchy-Schwarz .....                                 | 206    |
| Differenzierbarkeit .....           | 185      | Hölder .....   | 205    |
| Glattheit .....                     | 200      | Jensen .....   | 206    |
| Monotonie .....                     | 172      | Minkowski .....                                      | 205    |
| Periodizität .....                  | 171      | Youngsche .....                                      | 205    |
| Stetigkeit .....                    | 171      | Unterintegral .....                                  | 237    |
| Tangente .....                      | 177, 180 | Unterraumtopologie .....                             | 297    |
| Tangentialvektor                    |          | Untersumme .....                                     | 237    |
| an Kurve .....                      | 322      | Verdichtungskriterium .....                          | 76     |
| Taylor-Entwicklung .....            | 216, 256 | Vollständigkeit .....                                | 21, 24 |
| Logarithmus .....                   | 221      | der komplexen Zahlen .....                           | 37     |
| Restglied .....                     | 216, 256 | metrischer Raum .....                                | 290    |
| Taylor-Polynom .....                | 216, 226 | Teilmenge .....                                      | 291    |
| Taylor-Reihe                        |          | via Cauchy-Folgen .....                              | 62     |
| formale .....                       | 216      | via Infima .....                                     | 26     |
| Teilfolge .....                     | 45       | von $\mathbb{C}$ .....                               | 37, 62 |
| Teilüberdeckung .....               | 113      | Weg .....  | 302    |
| Topologie .....                     | 109, 294 | differenzierbarer .....                              | 322    |
| diskrete .....                      | 294      | Tangentialvektor .....                               | 322    |
| indiskrete .....                    | 294      | Wegzusammenhang .....                                | 302    |
| kofinite .....                      | 294      | stetiges Bild .....                                  | 303    |
| von $\mathbb{R}^n$ .....            | 305      | Wendepunkt .....                                     | 202    |
| Zariski .....                       | 294      | Winkel .....   | 175    |
| Topologischer Raum .....            | 294      | Wurzel .....   | 27     |
| Totale Differenzierbarkeit .....    | 316      | Wurzelfunktion                                       |        |
| stetige .....                       | 332      | Approximation .....                                  | 136    |
| Translation .....                   | 132      | Glattheit .....                                      | 200    |
| Trennungszahl .....                 | 19       | Stetigkeit .....                                     | 109    |
| Treppenfunktion .....               | 270      |  |        |

|                             |          |
|-----------------------------|----------|
| Youngsche Ungleichung ..... | 205      |
| Zahlen .....                | 5        |
| algebraische .....          | 45       |
| imaginäre .....             | 35       |
| irrationale .....           | 45       |
| komplexe .....              | 33       |
| rationale .....             | 5, 45    |
| reelle .....                | 24       |
| transzendente .....         | 45, 157  |
| Zahlenstrahl .....          | 13       |
| Zerlegung                   |          |
| Feinheit .....              | 232      |
| markierte .....             | 232      |
| Verfeinerung .....          | 232      |
| von Intervall .....         | 232      |
| Zusammenhang .....          | 120, 302 |
| stetiges Bild .....         | 125, 303 |
| Zwischenwertsatz            |          |
| für Ableitung .....         | 195      |
| stetige Funktion .....      | 126, 303 |
| Überdeckung .....           | 112      |