

Analysis für Informatik *

Martin Brokate **

Inhaltsverzeichnis

1	Reelle Zahlen und Vektoren	2
2	Folgen und Stetigkeit	12
3	Reihen	22
4	Trigonometrische Funktionen	33
5	Folgen und Stetigkeit II	37
6	Differenzierbarkeit	44
7	Sätze der Differenzialrechnung, Extremwerte	54
8	Das Integral	60
9	Taylorentwicklung	75
10	Kurven	80
11	Partielle Ableitungen	89
12	Gewöhnliche Differentialgleichungen	109
13	Das Integral im Mehrdimensionalen	124
14	Der Fixpunktsatz von Banach	137

*Vorlesungsskript, WS 2012/13

**Zentrum Mathematik, TU München

1 Reelle Zahlen und Vektoren

Wir erinnern an die **natürlichen Zahlen**

$$\mathbb{N} = \{1, 2, 3, \dots\}, \quad \mathbb{N}_0 = \{0, 1, 2, 3, \dots\},$$

die **ganzen Zahlen**

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

und die **rationalen Zahlen**

$$\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0 \right\}.$$

Bereits im Altertum war bekannt, dass der Zahlenraum \mathbb{Q} nicht ausreicht; Euklid hat bewiesen, dass die Länge x der Diagonalen im Einheitsquadrat, $x = \sqrt{2}$, nicht rational ist. Nach Pythagoras gilt $x^2 = 1^2 + 1^2 = 2$, wir können x also erhalten als Nullstelle des Polynoms

$$p(x) = x^2 - 2.$$

Allgemein heißt eine Zahl x **algebraische Zahl**, falls es ein Polynom mit ganzzahligen Koeffizienten gibt, welches x als Nullstelle hat. Aber selbst mit den algebraischen Zahlen hat man noch nicht ausreichend Zahlen zur Verfügung. Joseph Liouville hat 1844 bewiesen, dass es Zahlen gibt, die nicht algebraisch sind, sie heißen **transzendente Zahlen**. Es stellte sich heraus, dass die Zahl e transzendent ist (Charles Hermite 1873), und dass die Zahl π transzendent ist (Ferdinand Lindemann 1882). Letzteres hatte unter anderem zur Folge, dass die Quadratur des Kreises (zu einem gegebenen Kreis soll ein flächengleiches Quadrat konstruiert werden, wobei nur Zirkel und Lineal verwendet werden dürfen) als unmöglich erkannt wurde.

Die Menge \mathbb{R} der **reellen Zahlen** enthält e und π und ist ausreichend groß für den Gegenstandsbereich der Analysis,

Beschreibung und Analyse von Situationen im Kontinuum.

“Kontinuierlich” steht hier im Gegensatz zu “diskret”. Unsere klassische Vorstellung von Raum und Zeit ist kontinuierlich.

Bei der Modellierung realer Phänomene spielen sowohl diskrete also auch kontinuierliche Beschreibungen eine Rolle, je nach Problem und Gegenstand ist mal das eine, mal das andere geeigneter. Oft ist ein Zusammenspiel kontinuierlicher und diskreter Aspekte der angemessene Ansatz. Letztere kommen immer irgendwie ins Spiel, wenn Computer bei der Problemlösung eine Rolle spielen, da Computer nur endlich viele Zustände haben und also “diskrete Maschinen” sind.

Beliebige reelle Zahlen kann man sich am einfachsten in ihrer **Dezimaldarstellung** vorstellen, also etwa

$$1.003, \quad \frac{1}{3} = 0.33333\dots, \quad \pi = 3.141592653589793\dots$$

Die allgemeine Form einer solchen Dezimaldarstellung lautet

$$\pm d_0.d_1d_2d_3d_4\dots$$

mit dem ganzzahligen Anteil $d_0 \in \mathbb{N}_0$ und den Nachkommastellen $d_k \in \{0, 1, \dots, 9\}$ für $k \geq 1$. Man kann sie auch schreiben als unendliche Reihe

$$\pm \sum_{k=0}^{\infty} d_k \cdot 10^{-k}.$$

Mit solchen Reihen werden wir uns später beschäftigen. Eine Dezimaldarstellung ist nicht immer eindeutig bestimmt, nämlich dann nicht, wenn sie abbricht, wie etwa im Beispiel

$$1.003 = 1.0030000 \dots = 1.002999999 \dots$$

Wir können die Folge d_1, d_2, \dots der Nachkommastellen als Abbildung

$$d : \mathbb{N} \rightarrow \{0, 1, 2, \dots, 9\}$$

auffassen. Jede solcher Abbildungen beschreibt also eine reelle Zahl zwischen 0 und 1, und jede reelle Zahl zwischen 0 und 1 kann auf diese Weise beschrieben werden. Das bedeutet aber nicht, dass wir jede reelle Zahl auch tatsächlich berechnen können. Eine reelle Zahl x im abgeschlossenen Intervall $[0, 1]$ heißt **berechenbar**, falls es ein Programm gibt, welches zu jedem $n \in \mathbb{N}$ die Ziffern d_1, d_2, \dots, d_n der Dezimaldarstellung von x berechnet.

Jede algebraische Zahl ist berechenbar: Ist x eine algebraische Zahl, so nehmen wir ein zugehöriges Polynom p mit $p(x) = 0$ und ein Programm, welches die Nullstellen von p mit beliebiger Genauigkeit berechnet. Das ist möglich, wir beschäftigen uns aber nicht mit der Frage, wie das geht.

Außer den algebraischen gibt es auch transzendente Zahlen, die berechenbar sind, beispielsweise π und e .

Wie viele Zahlen gibt es? Natürlich unendlich viele. Aber ist unendlich gleich unendlich? Zunächst gilt

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{B},$$

wobei \mathbb{B} für die Menge der berechenbaren reellen Zahlen steht. Alle diese Mengen sind **abzählbar**, das heißt, es gibt surjektive Abbildungen $f_{\mathbb{Z}} : \mathbb{N} \rightarrow \mathbb{Z}$, $f_{\mathbb{Q}} : \mathbb{N} \rightarrow \mathbb{Q}$ und $f_{\mathbb{B}} : \mathbb{N} \rightarrow \mathbb{B}$. (Um \mathbb{B} abzuzählen, zählen wir die Menge der Programme ab, welche die Zahlen aus \mathbb{B} berechnen. Das geht, da wir die Menge aller Programme zu einem festen endlichen Alphabet abzählen können – wir ordnen sie der Länge nach, und dann alphabetisch bei Programmen gleicher Länge.) Wir erinnern an den Begriff der Kardinalität: Zwei Mengen X und Y haben gleiche Kardinalität, falls es eine bijektive Abbildung $f : X \rightarrow Y$ gibt (egal, ob wir diese angeben können oder nicht). Das führt etwa zu dem scheinbaren Paradox, dass es “gleichviele” gerade natürliche Zahlen wie natürliche Zahlen überhaupt gibt, denn

$$f : \mathbb{N} \rightarrow 2\mathbb{N} = \{2n : n \in \mathbb{N}\}, \quad f(n) = 2n,$$

ist bijektiv. Die Mengen \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{B} haben gleiche Kardinalität, da wir alle Abzählungen bijektiv machen können. Es stellt sich aber heraus:

Die Menge \mathbb{R} der reellen Zahlen ist nicht abzählbar.

Dafür hat Georg Cantor zwei Beweise gefunden, berühmt geworden ist der zweite aus dem Jahr 1877:

Wir nehmen an, es gebe eine Abzählung der reellen Zahlen im Intervall $[0, 1]$ als Folge x_1, x_2, x_3, \dots . Wir betrachten die Dezimaldarstellungen dieser Zahlen

$$\begin{aligned} x_1 &= 0. d_{11} d_{12} d_{13} \dots \\ x_2 &= 0. d_{21} d_{22} d_{23} \dots \\ x_3 &= 0. d_{31} d_{32} d_{33} \dots \\ &\vdots \end{aligned}$$

und definieren eine reelle Zahl $x = 0.d_1d_2\dots$, indem wir setzen $d_k = 2$ falls $d_{kk} = 1$ und $d_k = 1$ andernfalls. Dann ist x von jeder der Zahlen x_k verschieden (da die k -te Stelle unterschiedlich ist), also kommt x in der Abzählung nicht vor, ein Widerspruch.

Man beachte, dass dieses Argument nichts mit ‘‘Berechenbarkeit’’ zu tun hat, der Widerspruch bezieht sich bereits auf die bloÙe Existenz der Folge x_1, x_2, x_3, \dots . Die theoretische Fundierung von Berechenbarkeit und Algorithmik erfolgte im Zusammenhang mit Entwicklungen in der mathematischen Logik und der Mengenlehre (Kurt Godel) durch Alonzo Church und Alan Turing zwischen 1930 und 1940.

Die Menge \mathbb{R} hat also eine groÙere Kardinalitat als die Mengen \mathbb{N} , \mathbb{Z} , \mathbb{Q} und \mathbb{B} , in diesem Sinn gibt es ‘‘viel mehr’’ nicht berechenbare als berechenbare Zahlen. Ein anderer Aspekt dieses Sachverhalts ist: Greift man zufallig und ‘‘gleichwahrscheinlich’’ eine reelle Zahl x aus $[0, 1]$ heraus, so ist die Wahrscheinlichkeit, dass x berechenbar ist, gleich Null.

Eigenschaften der reellen Zahlen. Die reellen Zahlen bilden hinsichtlich Addition und Multiplikation einen kommutativen Korper mit Einselement. Die entsprechenden Rechenregeln setzen wir als bekannt voraus. Weiterhin ist auf den reellen Zahlen eine **Ordnungsstruktur** definiert durch die beiden folgenden Eigenschaften.

(1) Fur jedes $x \in \mathbb{R}$ ist genau eine der folgenden drei Aussagen wahr:

$$x > 0, \quad x = 0, \quad -x > 0. \tag{1.1}$$

(2) Fur alle $x, y \in \mathbb{R}$ gilt:

$$x, y > 0 \quad \Rightarrow \quad x + y > 0 \text{ und } x \cdot y > 0. \tag{1.2}$$

Wir definieren auf $\mathbb{R} \times \mathbb{R}$ vier Relationen, indem wir setzen

$$\begin{aligned} x > y &\quad \text{genau dann, wenn} && x - y > 0, \\ x \geq y &\quad \text{genau dann, wenn} && x > y \text{ oder } x = y, \\ x < y &\quad \text{genau dann, wenn} && y - x > 0, \\ x \leq y &\quad \text{genau dann, wenn} && x < y \text{ oder } x = y. \end{aligned} \tag{1.3}$$

Aus (1.1) und (1.2) sowie den Rechenregeln fur Addition und Multiplikation erhalt man

die üblichen Regeln für Ungleichungen, beispielsweise

$$\begin{aligned}
 x < y \text{ und } y < z &\Rightarrow x < z && \text{(Transitivität)} \\
 x \leq y \text{ und } y \leq x &\Rightarrow x = y && \text{(Antisymmetrie)} \\
 x \neq 0 &\Rightarrow x^2 > 0 \\
 x > 0 &\Rightarrow \frac{1}{x} > 0 \\
 0 < x < y &\Rightarrow 0 < \frac{1}{y} < \frac{1}{x}
 \end{aligned} \tag{1.4}$$

und so weiter. Wir verwenden die Notation

$$[a, b] = \{x : x \in \mathbb{R}, a \leq x \leq b\} \quad \text{(abgeschlossenes Intervall)} \tag{1.5}$$

und analog (a, b) , $[a, b)$ und $(a, b]$ für offene und halboffene Intervalle; ebenso

$$[a, \infty) = \{x : x \in \mathbb{R}, a \leq x < \infty\} \tag{1.6}$$

und analog (a, ∞) , $(-\infty, b]$ und so weiter.

Die Vollständigkeit von \mathbb{R} .

Definition 1.1 (Obere Schranke)

Eine Zahl $x \in \mathbb{R}$ heißt obere Schranke einer Teilmenge M von \mathbb{R} , falls $y \leq x$ gilt für alle $y \in M$. Eine Teilmenge M von \mathbb{R} heißt nach oben beschränkt, wenn es eine obere Schranke von M gibt, andernfalls heißt M nach oben unbeschränkt.

Beispielsweise ist 8 eine obere Schranke von $M = \{-3, 7, \pi, 1/2\}$; mehr noch, jedes $x \geq 7$ ist obere Schranke von M . Die Zahl 7 spielt in diesem Beispiel eine besondere Rolle: Sie ist einerseits Maximum von M und andererseits die kleinste obere Schranke von M .

Definition 1.2 (Maximum)

Eine Zahl $x \in \mathbb{R}$ heißt Maximum einer Teilmenge M von \mathbb{R} , falls gilt: $x \in M$ und $x \geq y$ für jedes $y \in M$.

Es kann passieren, dass eine Menge nach oben beschränkt ist, aber kein Maximum im Sinne von Definition 1.2 hat, beispielsweise das offene Intervall $M = (0, 1)$. Hier gilt: Jedes $x \geq 1$ ist obere Schranke von M , 1 ist kleinste obere Schranke von M , aber 1 ist nicht Maximum von M , da 1 nicht zu M gehört. In der Tat, $(0, 1)$ hat kein Maximum.

In beiden genannten Beispielen besitzt M eine kleinste obere Schranke, diese ist aber nur im ersten Beispiel ein Maximum. In der Mathematik hat man daher den Begriff des Supremums eingeführt.

Definition 1.3 (Supremum)

Eine reelle Zahl s heißt Supremum einer Teilmenge M von \mathbb{R} , falls s obere Schranke ist von M und falls $s \leq x$ gilt für jede obere Schranke x von M . In anderen Worten, s ist kleinste obere Schranke von M .

Aus der Definition folgt unmittelbar, dass das Supremum von M eindeutig bestimmt ist (sind s_1 und s_2 Suprema von M , so gilt $s_1 \leq s_2$ und $s_2 \leq s_1$, also $s_1 = s_2$). Wir bezeichnen es mit

$$\sup M \quad \text{oder} \quad \sup(M). \quad (1.7)$$

Also beispielsweise: $\sup(0, 1) = 1 = \sup[0, 1]$.

Als Konvention verabredet man, dass $\sup \emptyset = -\infty$, und dass $\sup M = +\infty$ falls M nach oben unbeschränkt ist.

Wir sind nun bei der Eigenschaft angekommen, die \mathbb{R} von \mathbb{Q} und \mathbb{B} unterscheidet.

$$\boxed{\text{Jede nichtleere nach oben beschränkte Teilmenge von } \mathbb{R} \text{ besitzt ein Supremum.}} \quad (1.8)$$

Wir haben eingangs die reellen Zahlen als die Menge aller Dezimalbrüche eingeführt, modulo der Gleichheit zweier solcher. Dieser Zugang ist einerseits anschaulich. Will man nun die Körper- und die Ordnungseigenschaften sowie (1.8) damit beweisen, so führt das andererseits zu längeren Überlegungen. Alternativ dazu kann man (wie auch im Mathematikstudium weitgehend üblich) diese Eigenschaften als Axiome setzen, (1.8) wird dann als Supremumsaxiom bezeichnet.

Die durch das Supremumsaxiom sich ergebende Eigenschaft bezeichnet man als die **Vollständigkeit** von \mathbb{R} .

Der Unterschied von \mathbb{R} und \mathbb{Q} wird an folgendem Beispiel deutlich: Wir betrachten

$$M = \{x : x \in \mathbb{Q}, x^2 \leq 2\}$$

Diese Menge hat ein Supremum in \mathbb{R} , nämlich $\sqrt{2}$, welches gleichzeitig Maximum ist. Fassen wir aber M als Teilmenge von \mathbb{Q} auf, so hat M zwar obere Schranken in \mathbb{Q} (nämlich jedes $x \in \mathbb{Q}$ mit $x > \sqrt{2}$), aber keine kleinste obere Schranke **in** \mathbb{Q} und auch kein Maximum in \mathbb{Q} . Die rationalen Zahlen \mathbb{Q} haben also nicht die Eigenschaft (1.8), ebensowenig hat \mathbb{B} die Eigenschaft (1.8).

Die natürlichen Zahlen \mathbb{N} sind nach oben unbeschränkt im Sinne von Definition 1.1, das heißt, es gilt:

$$\text{Zu jedem } x \in \mathbb{R} \text{ gibt es ein } n \in \mathbb{N} \text{ mit } x < n. \quad (1.9)$$

Diese Eigenschaft (bzw. eine dazu äquivalente) nennt man die **Archimedizität** der reellen Zahlen.

Beweis von (1.9): Sei \mathbb{N} nach oben beschränkt, sei $s = \sup \mathbb{N}$ gemäß Supremumsaxiom. $s - 1$ ist keine obere Schranke von \mathbb{N} (da s die kleinste obere Schranke ist), also gibt es $n \in \mathbb{N}$ mit $s - 1 < n$. Dann aber ist $s < n + 1 \in \mathbb{N}$ und damit s keine obere Schranke von \mathbb{N} , ein Widerspruch.

Aus (1.9) folgt

$$\text{zu jedem } 0 < x \in \mathbb{R} \text{ gibt es ein } n \in \mathbb{N} \text{ mit } 0 < \frac{1}{n} < x, \quad (1.10)$$

indem wir (1.9) auf $\frac{1}{x}$ anwenden.

Man kann (1.10) dazu verwenden, um zu beweisen, dass \mathbb{Q} **dicht** ist in \mathbb{R} im folgenden Sinn:

$$\text{Sind } x, y \in \mathbb{R} \text{ mit } x < y, \text{ so gibt es ein } r \in \mathbb{Q} \text{ mit } x < r < y. \quad (1.11)$$

Anders ausgedrückt: Zwischen zwei reellen Zahlen (egal, wie dicht diese beieinander liegen) liegt immer eine rationale Zahl.

Untere Schranke, Minimum, Infimum. Diese Begriffe entsprechen den Begriffen obere Schranke, Maximum und Supremum, nur schauen wir uns die reelle Zahlengerade “von links” statt “von rechts” an. Formal bedeutet das, dass alle Ungleichungen umgedreht werden.

Eine Zahl $x \in \mathbb{R}$ heißt untere Schranke einer Teilmenge M von \mathbb{R} , falls $y \geq x$ gilt für alle $y \in M$. Eine Teilmenge M von \mathbb{R} heißt nach unten beschränkt, wenn es eine untere Schranke von M gibt, andernfalls heißt M nach unten unbeschränkt.

Eine Zahl $x \in \mathbb{R}$ heißt Minimum einer Teilmenge M von \mathbb{R} , falls gilt: $x \in M$ und $x \leq y$ für jedes $y \in M$.

Eine Zahl $s \in \mathbb{R}$ heißt Infimum einer Teilmenge M von \mathbb{R} , falls s untere Schranke ist von M und falls $s \geq x$ gilt für jede untere Schranke x von M . In anderen Worten, s ist größte untere Schranke von M .

Wir schreiben $\inf(M)$ für das Infimum von M , mit der Konvention $\inf(\emptyset) = +\infty$ und $\inf(M) = -\infty$ falls M nach unten unbeschränkt ist.

Rechenregeln für Supremum und Maximum. Unmittelbar aus den Definitionen folgt: Hat eine Teilmenge M von \mathbb{R} ein Maximum, so ist dieses gleichzeitig das Supremum von M . Die folgenden Rechenregeln für das Supremum gelten daher genauso für das Maximum (falls die entsprechenden Maxima existieren).

Satz 1.4 *Seien $X, Y \subset \mathbb{R}$ mit $\sup(X), \sup(Y) \in \mathbb{R}$. Dann gilt*

$$\sup(X + Y) = \sup(X) + \sup(Y). \quad (1.12)$$

Falls $\lambda \geq 0$ ist, so gilt

$$\sup(\lambda X) = \lambda \cdot \sup(X). \quad (1.13)$$

Sind $X, Y \subset [0, \infty)$, so gilt

$$\sup(X \cdot Y) = \sup(X) \cdot \sup(Y). \quad (1.14)$$

Ist $X \subset Y$, so gilt

$$\sup(X) \leq \sup(Y). \quad (1.15)$$

Hierbei ist

$$\begin{aligned} X + Y &= \{x + y : x \in X, y \in Y\}, \\ \lambda X &= \{\lambda x : x \in X\}, \\ X \cdot Y &= \{x \cdot y : x \in X, y \in Y\}. \end{aligned}$$

Wegen

$$\inf(X) = -\sup(-X) \quad (1.16)$$

gelten dieselben Regeln auch für das Infimum, wobei in (1.15) die Ungleichung umgedreht werden muss.

Die Beweise lassen wir hier weg, sie benützen direkt die Definitionen.

Die Betragsfunktion. Sie ist für $x \in \mathbb{R}$ definiert durch

$$|x| = \max\{-x, x\}. \quad (1.17)$$

Zunächst gilt

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

Weiter gilt für alle $x, y \in \mathbb{R}$

$$|x| \geq 0, \quad |x| = 0 \Leftrightarrow x = 0, \quad (1.18)$$

$$|xy| = |x| \cdot |y|, \quad (1.19)$$

$$|x + y| \leq |x| + |y|. \quad (1.20)$$

Die Ungleichung (1.20) heißt **Dreiecksungleichung**, in ihr gilt Gleichheit genau dann, wenn $xy \geq 0$, das heißt, wenn x und y das gleiche Vorzeichen haben. Aus (1.20) erhält man wegen

$$|x| = \left| \frac{x}{y} \cdot y \right| = \left| \frac{x}{y} \right| \cdot |y|$$

die Regel

$$\left| \frac{x}{y} \right| = \frac{|x|}{|y|}, \quad \text{für alle } x, y \in \mathbb{R} \text{ mit } y \neq 0. \quad (1.21)$$

Die **umgekehrte Dreiecksungleichung** besagt, dass

$$|x - y| \geq \left| |x| - |y| \right| \quad (1.22)$$

gilt für alle $x, y \in \mathbb{R}$.

Die Bernoulli-Ungleichung. Sie besagt, dass

$$(1 + x)^n \geq 1 + nx \quad (1.23)$$

für alle $n \in \mathbb{N}$ und alle $x \in \mathbb{R}$ mit $x \geq -1$. Den Beweis überlassen wir der Übung.

Die Ungleichung zwischen geometrischem und arithmetischem Mittel. Seien $a, b \geq 0$. Es gilt

$$0 \leq (a - b)^2 = a^2 - 2ab + b^2 = a^2 + 2ab + b^2 - 4ab,$$

also

$$4ab \leq (a + b)^2, \quad ab \leq \frac{(a + b)^2}{4}.$$

Es ergibt sich

$$\sqrt{ab} \leq \frac{a + b}{2}. \quad (1.24)$$

Auf der linken Seite steht das geometrische Mittel von a und b , auf der rechten Seite das arithmetische Mittel von a und b . An der ersten Zeile der Rechnung erkennt man, dass Gleichheit in (1.24) genau dann gilt, wenn $a = b$.

Vektoren und Normen. Aus der Linearen Algebra ist bekannt der Vektorraum \mathbb{R}^n , bestehend aus Vektoren der Form

$$x = (x_1, x_2, \dots, x_n)$$

mit $x_i \in \mathbb{R}$ für $1 \leq j \leq n$. (Bisher war x eine Zahl, jetzt ist x ein Vektor.) Die Länge eines Vektors ist definiert worden als

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2} = \left(\sum_{j=1}^n x_j^2 \right)^{\frac{1}{2}}, \quad (1.25)$$

das Skalarprodukt zweier Vektoren $x, y \in \mathbb{R}^n$ als

$$\langle x, y \rangle = \sum_{j=1}^n x_j y_j. \quad (1.26)$$

Es gilt immer $\|x\| \geq 0$, sowie $\|x\| = 0$ genau dann, wenn $x = 0$.

Die **Cauchy-Schwarz-Ungleichung** besagt, dass

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \quad (1.27)$$

gilt für beliebige Vektoren $x, y \in \mathbb{R}^n$.

Wir führen sie auf die Ungleichung (1.24) zurück. Ist $\|x\| = 0$ oder $\|y\| = 0$, so ist (1.27) offensichtlich erfüllt. Andernfalls setzen wir

$$\xi_j = \frac{|x_j|}{\|x\|}, \quad \eta_j = \frac{|y_j|}{\|y\|},$$

und rechnen unter Verwendung von (1.24)

$$\begin{aligned} \sum_{j=1}^n \xi_j \eta_j &= \sum_{j=1}^n \sqrt{\xi_j^2 \eta_j^2} \leq \sum_{j=1}^n \frac{\xi_j^2 + \eta_j^2}{2} = \sum_{j=1}^n \frac{1}{2} \left(\frac{x_j^2}{\|x\|^2} + \frac{y_j^2}{\|y\|^2} \right) \\ &= \frac{1}{2} \left(\frac{\sum_{j=1}^n x_j^2}{\|x\|^2} + \frac{\sum_{j=1}^n y_j^2}{\|y\|^2} \right) = \frac{1}{2}(1 + 1) = 1, \end{aligned}$$

also

$$|\langle x, y \rangle| = \left| \sum_{j=1}^n x_j y_j \right| \leq \sum_{j=1}^n |x_j| |y_j| = \|x\| \cdot \|y\| \cdot \sum_{j=1}^n \xi_j \eta_j \leq \|x\| \cdot \|y\|.$$

Die **Dreiecksungleichung** besagt, dass

$$\|x + y\| \leq \|x\| + \|y\| \quad (1.28)$$

gilt für beliebige Vektoren $x, y \in \mathbb{R}^n$. Der Name rührt daher, dass in dem von den Vektoren $0, x, x + y$ aufgespannten Dreieck die Länge der Seite von 0 nach $x + y$ kleiner oder gleich der Summe der Längen der beiden anderen Seiten (von 0 nach x und von x nach $x + y$) ist.

Wir führen die Dreiecksungleichung auf die Cauchy-Schwarz-Ungleichung zurück. Es gilt

$$\begin{aligned}\|x + y\|^2 &= \sum_{j=1}^n (x_j + y_j)^2 = \sum_{j=1}^n (x_j^2 + 2x_j y_j + y_j^2) = \sum_{j=1}^n x_j^2 + \sum_{j=1}^n 2x_j y_j + \sum_{j=1}^n y_j^2 \\ &= \|x\|^2 + 2 \langle x, y \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,\end{aligned}$$

und Wurzelziehen ergibt (1.28).

Komplexe Zahlen. Eine komplexe Zahl z ist ein Paar $z = (x, y)$ reeller Zahlen. Wir können sie als Punkt in der Ebene \mathbb{R}^2 auffassen, man spricht von der komplexen Zahlenebene und bezeichnet sie mit \mathbb{C} . x heißt der Realteil von z , y heißt der Imaginärteil von z , geschrieben

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

Die reellen Zahlen entsprechen der x -Achse, das heißt, jeder reellen Zahl x entspricht die komplexe Zahl $(x, 0)$, insbesondere entspricht der reellen 1 die komplexe Zahl $(1, 0)$. Die y -Achse heißt die imaginäre Achse, die komplexe Zahl $(0, 1)$ heißt imaginäre Einheit und wird mit

$$i$$

bezeichnet. Statt $z = (x, y)$ schreibt man dann

$$z = x + iy$$

in Abkürzung für $z = x(1, 0) + y(0, 1)$. Die Addition in \mathbb{C} ist genauso definiert wie die Vektoraddition im \mathbb{R}^2 , also

$$z_1 + z_2 = (x_1 + x_2, y_1 + y_2),$$

falls $z_1 = (x_1, y_1)$ und $z_2 = (x_2, y_2)$, oder anders geschrieben,

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2).$$

In \mathbb{C} ist auch eine Multiplikation definiert. Man setzt zunächst

$$i^2 = i \cdot i = -1.$$

Sind $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$, so soll gelten

$$\begin{aligned}z_1 \cdot z_2 &= (x_1 + iy_1) \cdot (x_2 + iy_2) = x_1 \cdot x_2 + iy_1 \cdot x_2 + x_1 \cdot iy_2 + i^2 y_1 \cdot y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).\end{aligned}\tag{1.29}$$

In Vektorschreibweise liest sich die komplexe Multiplikation also als

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1),$$

man kann sie sich aber besser merken in der Form (1.29). Eine geometrisch anschauliche Interpretation werden wir später kennenlernen, wenn wir Polarkoordinaten und die Exponentialfunktion in \mathbb{C} betrachten.

Man kann beweisen, dass \mathbb{C} ein kommutativer Körper mit Einselement 1 (die reelle 1) ist, und dass für $z = x + iy \neq 0$

$$\frac{1}{z} = \frac{1}{x^2 + y^2}(x - iy)$$

gilt. Die zu $z = x + iy$ konjugiert komplexe Zahl \bar{z} ist definiert durch

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

man erhält sie durch Spiegeln von z an der reellen Achse. Es gelten die Rechenregeln

$$\overline{\bar{z}} = z, \quad \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2,$$

für beliebige $z, z_1, z_2 \in \mathbb{C}$.

Der Betrag von z ist gleich der Länge des Vektors (x, y) ,

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

Es gilt für jedes $z \in \mathbb{C}$

$$|z|^2 = z\bar{z}.$$

2 Folgen und Stetigkeit

Folgen entstehen, indem man Objekte desselben Typs aneinanderreicht. So ist

a,c,d,g,k,a,b,a,a,a,...

eine Buchstabenfolge; die Punkte deuten an, dass die Folge nicht endet. (Falls sie es tut, spricht man von einer endlichen Folge; falls man extra betonen will, dass sie nicht endet, von einer unendlichen Folge.) Die Folgen

$$1, 2, 4, 8, 16, 32, \dots \quad (2.1)$$

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots \quad (2.2)$$

sind Zahlenfolgen. Man kann auch Folgen von Vektoren oder von Funktionen betrachten. Allgemein kann man Folgen definieren als Abbildungen von \mathbb{N} in irgendeine Menge M . Meistens schreibt man sie in der Form

$$x_1, x_2, x_3, x_4, x_5, \dots$$

Will man eine Folge konkret angeben, so kann man das beispielsweise tun durch eine explizite Vorschrift, etwa

$$x_n = 2^n$$

oder durch eine Rekursionsformel, etwa

$$x_{n+1} = 2x_n, \quad x_0 = 1.$$

In beiden Fällen erhalten wir die Folge (2.1), die Zählung der Indizes beginnt hier bei 0. Folgen tauchen immer dann auf, wenn die Lösung eines Problems durch ein Iterationsverfahren berechnet wird, welches aus einer vorhandenen Näherung x_n eine (hoffentlich) bessere Näherung x_{n+1} liefert. Iterationsverfahren sind zentrale Bausteine vieler numerischer Simulationen in Wirtschaft, Technik und Wissenschaft.

Wir beschäftigen uns in diesem Kapitel mit grundlegenden Eigenschaften von Zahlenfolgen und dem unmittelbar damit zusammenhängenden Begriff der Stetigkeit von Funktionen. In beiden Fällen ist

Grenzwert

der zentrale Begriff. Auf ihm basiert die gesamte Analysis im Kontinuierlichen, seine Präzisierung im 19. Jahrhundert war einer der Ausgangspunkte der modernen Mathematik.

Die Folge

$$x_n = \frac{1}{n}, \quad \text{also} \quad 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$$

strebt in der Anschauung gegen 0, die Folge

$$x_n = (-1)^n, \quad \text{also} \quad -1, 1, -1, 1, -1, 1, -1, \dots$$

hingegen springt immer hin und her.

Definition 2.1 (Grenzwert einer Folge)

Sei $(x_n)_{n \in \mathbb{N}}$ eine reelle Folge. Ein $a \in \mathbb{R}$ heißt Grenzwert (oder Limes) von $(x_n)_{n \in \mathbb{N}}$, falls es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt mit

$$|x_n - a| < \varepsilon \quad \text{für alle } n \in \mathbb{N} \text{ mit } n \geq N. \quad (2.3)$$

Falls $(x_n)_{n \in \mathbb{N}}$ einen Grenzwert a hat, so sagt man auch: “ (x_n) konvergiert gegen a ” oder “ (x_n) ist konvergent”.

Anders ausgedrückt: Gibt man eine Genauigkeit $\varepsilon > 0$ vor, egal wie klein, so liegen ab einem gewissen Index (in der Definition heißt er N) alle Folgenglieder näher als ε am Grenzwert a . Dieser Index N hängt von ε ab; je kleiner man ε vorgibt, desto größer wird er im allgemeinen sein. Bei der Folge

$$x_n = \frac{1}{n}$$

mit Grenzwert $a = 0$ erreicht man das, wenn man

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

setzt, dann gilt für alle $n \geq N$ nämlich

$$|x_n - 0| = x_n = \frac{1}{n} \leq \frac{1}{N} < \varepsilon,$$

und die in Definition 2.1 verlangte Bedingung ist erfüllt.

Die Folge

$$x_n = (-1)^n$$

hat keinen Grenzwert: Gibt man $\varepsilon = 1$ (oder kleiner) vor, so gilt für jedes $a \in \mathbb{R}$ entweder $|1 - a| \geq \varepsilon$ oder $|(-1) - a| \geq \varepsilon$, das heißt, die in Definition 2.1 verlangte Bedingung kann für kein a erfüllt werden.

Satz 2.2 Jede reelle Folge $(x_n)_{n \in \mathbb{N}}$ hat höchstens einen Grenzwert. □

Für “ a ist Grenzwert der Folge (x_n) ” schreibt man

$$a = \lim_{n \rightarrow \infty} x_n \quad \text{oder} \quad x_n \rightarrow a. \quad (2.4)$$

Ein weiteres Beispiel ist die Dezimaldarstellung einer Zahl. Ist etwa $a \in [0, 1]$ mit Dezimaldarstellung

$$a = 0. d_1 d_2 d_3 \dots \quad \text{und} \quad x_n = 0. d_1 \dots d_n,$$

so gilt $x_n \rightarrow a$, da $|x_n - a| \leq 10^{-n}$.

Die Definition der Konvergenz verlangt nicht, dass der Abstand zum Grenzwert laufend kleiner wird, wenn n größer wird. So ist beispielsweise die Folge

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots$$

gegen 0 konvergent, obwohl der Abstand zu 0 immer wieder (sogar unendlich oft) größer wird.

Liegt eine Folge oberhalb einer anderen, so gilt das auch für ihre Grenzwerte, falls diese existieren.

Lemma 2.3 Seien (x_n) und (y_n) Folgen mit $x_n \rightarrow a$ und $y_n \rightarrow b$. Gilt außerdem $x_n \leq y_n$ für alle n , so gilt auch $a \leq b$. \square

In der Situation des Lemmas kann es sein, dass $x_n < y_n$ für alle n gilt, aber die Grenzwerte gleich sind, beispielsweise bei

$$x_n = \frac{2}{n}, \quad y_n = \frac{1}{n}, \quad a = b = 0.$$

Satz 2.4 (Einschließung)

Seien (x_n) und (y_n) Folgen mit $x_n \rightarrow a$, $y_n \rightarrow a$ und $x_n \leq y_n$ für alle n . Ist (w_n) eine weitere Folge mit $x_n \leq w_n \leq y_n$ für alle n , so konvergiert (w_n) , und zwar ebenfalls gegen a . \square

Im Gegensatz zum vorangehenden Lemma 2.3 braucht bei der Einschließung nicht vorausgesetzt zu werden, dass die mittlere Folge konvergiert, falls (wie im Satz stattdessen vorausgesetzt wird) die beiden äußeren Folgen gegen den **gleichen** Grenzwert konvergieren.

Ein Beispiel: Wir betrachten die Folge

$$x_n = \frac{1}{n + (\log n)^2 + 3}.$$

Da wir wissen, dass $\log n \geq 0$ ist für $n \geq 1$, ist der Nenner immer $\geq n$ und daher

$$0 \leq x_n \leq \frac{1}{n}.$$

Da $1/n \rightarrow 0$, folgt aus dem Einschließungssatz, dass $x_n \rightarrow 0$. Mehr rechnen müssen wir nicht, und wir müssen auch nicht die Definition des Grenzwerts (mit ε und N) heranziehen.

Uneigentliche Konvergenz. Die Folge

$$1, 2, 4, 8, 16, \dots \tag{2.5}$$

ist nicht konvergent im Sinne von Definition 2.1, aber sie “strebt gegen unendlich”.

Definition 2.5 (Uneigentliche Konvergenz)

Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt *uneigentlich konvergent gegen $+\infty$* , falls es zu jedem $C > 0$ ein $N \in \mathbb{N}$ gibt mit

$$x_n \geq C \quad \text{für alle } n \in \mathbb{N} \text{ mit } n \geq N. \tag{2.6}$$

Sie heißt *uneigentlich konvergent gegen $-\infty$* , falls die Folge $(-x_n)_{n \in \mathbb{N}}$ *uneigentlich konvergent gegen $+\infty$* konvergiert.

Die Folge (2.5) hat diese Eigenschaft, da wir zu beliebig vorgegebenem C als Index N jede Zahl mit $2^N \geq C$ wählen können.

Eine Folge, die nicht konvergent ist im Sinne von Definition 2.1, heißt **divergent**. Gemäß dieser Terminologie sind uneigentlich konvergente Folgen divergent. Um solche Folgen

von anderen divergenten Folgen wie etwa $1, -1, 1, -1, \dots$ zu unterscheiden, nennt man uneigentlich konvergente Folgen auch **bestimmt divergent**.

Beschränkte Folgen. Ist (x_n) eine Folge mit $x_n \rightarrow a$, so müssen ab einem gewissen Index N alle Folgenglieder x_n im Intervall $[a - 1, a + 1]$ liegen, es können also höchstens endlich viele (nämlich nur die x_n mit $n < N$) außerhalb von $[a - 1, a + 1]$ liegen. Setzen wir

$$K = \max\{|a - 1|, |a + 1|, \max_{1 \leq n < N} |x_n|\},$$

so muss also

$$|x_n| \leq K \quad \text{für alle } n \in \mathbb{N} \quad (2.7)$$

gelten. Solche Folgen (das heißt, Folgen für die ein K mit der Eigenschaft (2.7) existiert) heißen **beschränkt**. Die vorangehende Überlegung zeigt:

Satz 2.6 *Jede konvergente Folge ist beschränkt.* □

Umgekehrt muss eine beschränkte Folge nicht konvergent sein, das zeigt das Beispiel $-1, 1, -1, 1, -1, \dots$

Monotone Folgen. Eine Folge (x_n) heißt **monoton wachsend**, falls $x_n \leq x_{n+1}$ gilt für alle n . Gilt sogar $x_n < x_{n+1}$, so heißt sie **streng monoton wachsend**. Entsprechend ist “(streng) monoton fallend” definiert.

Sei (x_n) eine monoton wachsende Folge. Ist diese außerdem beschränkt, so hat die Menge $M = \{x_n : n \in \mathbb{N}\}$ eine kleinste obere Schranke, das Supremum

$$\sup_{n \in \mathbb{N}} x_n =: s.$$

Für jedes $\varepsilon > 0$ muss es ein Folgenglied x_N geben mit $s - \varepsilon < x_N$ (andernfalls wäre $s - \varepsilon$ eine obere Schranke von M und damit s nicht die kleinste obere Schranke). Für alle weiteren Folgenglieder x_n mit $n \geq N$ gilt

$$s - \varepsilon < x_N \leq x_n \leq s,$$

da die Folge monoton wachsend ist und da s eine obere Schranke ist. Die in der Definition des Grenzwerts verlangte Bedingung ist damit für s erfüllt. Wir sehen also:

Satz 2.7 *Jede monoton wachsende und beschränkte Folge ist konvergent, und es gilt*

$$\lim_{n \rightarrow \infty} x_n = \sup_{n \in \mathbb{N}} x_n. \quad (2.8)$$

□

Analog gilt, dass jede beschränkte monoton fallende Folge gegen ihr Infimum konvergiert.

Grenzwerte und Rechenoperationen. Die Grenzwerte von zusammengesetzten Folgen können wir oft auf die Grenzwerte ihrer Bestandteile zurückführen.

Satz 2.8 *Gilt $x_n \rightarrow a$ und $y_n \rightarrow b$, so gilt auch*

$$x_n + y_n \rightarrow a + b, \quad x_n - y_n \rightarrow a - b, \quad x_n y_n \rightarrow ab, \quad \frac{x_n}{y_n} \rightarrow \frac{a}{b}, \quad (2.9)$$

letzteres falls $b \neq 0$.

Beispielsweise ergibt sich der Grenzwert von

$$x_n = \frac{n^2 - 2n}{n^2 + n + 1}$$

aus der Rechnung (Zähler und Nenner werden durch n^2 dividiert)

$$\frac{n^2 - 1}{n^2 + 2n + 1} = \frac{1 - \frac{1}{n^2}}{1 + \frac{2}{n} + \frac{1}{n^2}} \rightarrow \frac{1 - 0}{1 + 0 + 0} = 1.$$

Stetige Funktionen. Als einführendes Beispiel betrachten wir die durch

$$f(x) = \begin{cases} x, & x \leq 0, \\ x + 1, & x > 0, \end{cases} \quad (2.10)$$

definierte Funktion. Sie hat im Punkt $x = 0$ eine Sprungstelle. Wenn wir uns dem Nullpunkt von links nähern, bemerken wir diese nicht: Für die Folge $x_n = -1/n$ beispielsweise gilt

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

Anders ist es, wenn wir von rechts kommen, beispielsweise entlang $x_n = 1/n$: Es ist

$$\frac{1}{n} \rightarrow 0, \quad f\left(\frac{1}{n}\right) = 1 + \frac{1}{n} \rightarrow 1, \quad \text{aber } f(0) = 0 \neq 1.$$

Es folgt die allgemeine Definition.

Definition 2.9 (Stetigkeit)

Eine reellwertige Funktion f mit Definitionsgebiet $D \subset \mathbb{R}$ (kurz $f : D \rightarrow \mathbb{R}$) heißt stetig in einem Punkt $x \in D$, falls für jede Folge (x_n) in D mit $x_n \rightarrow x$ gilt

$$\lim_{n \rightarrow \infty} f(x_n) = f(x). \quad (2.11)$$

f heißt stetig in D , falls f in jedem Punkt D stetig ist.

Sagt man “ f ist stetig” ohne weiteren Zusatz, so ist damit in der Regel gemeint, dass die Funktion f auf ihrem gesamten Definitionsgebiet stetig ist.

Die in (2.10) betrachtete Funktion f ist stetig in allen Punkten $x \neq 0$, in 0 ist sie unstetig.

Ändert man f in diesem Beispiel so ab, dass man zwar die Funktionsvorschrift beibehält, aber den Definitionsbereich verkleinert zu $D = (-\infty, 0]$ statt $D = \mathbb{R}$, so ist $f : D \rightarrow \mathbb{R}$ nun stetig, da $f = 0$ auf D gilt. Wählt man stattdessen $D = [0, \infty)$, so ist f nach wie vor unstetig in 0.

Neben Sprungstellen können Oszillationen eine weitere Quelle von Unstetigkeiten sein. Betrachten wir als Beispiel folgende Funktion $f : [0, 1] \rightarrow \mathbb{R}$. Gegeben sind die Punkte

$$\left(\frac{1}{n}, (-1)^n\right)$$

und f entsteht, indem wir diese Punkte durch Strecken verbinden. Diese Funktion ist in 0 unstetig (egal wie wir $f(0)$ definieren), da die Folge $f(1/n) = (-1)^n$ keinen Grenzwert hat für $n \rightarrow \infty$. Anders wäre es, wenn wir stattdessen die Punkte

$$\left(\frac{1}{n}, (-1)^n \frac{1}{n}\right)$$

wählen, in diesem Fall ist f stetig in 0.

Kehren wir zu stetigen Funktionen zurück. Konstante Funktionen sowie die Identität $f(x) = x$ sind stetig (egal, wie wir $D \subset \mathbb{R}$ wählen). Um festzustellen, ob eine gegebene Funktion stetig ist, hilft oft der folgende Satz.

Satz 2.10 *Seien f, g reellwertige Funktionen mit Definitionsbereich D . Sind f und g in einem Punkt $x \in D$ stetig, so auch $f + g$, $f - g$, fg und f/g . (letzteres, sofern $g(x) \neq 0$).*

Der Beweis benutzt, dass bei Folgen die algebraischen Operationen mit der Grenzwertbildung vertauscht werden können (Satz 2.8), also beispielsweise

Sind f, g stetig in x und $x_n \rightarrow x$, so gilt $f(x_n) \rightarrow f(x)$ und $g(x_n) \rightarrow g(x)$, also auch $(f + g)(x_n) = f(x_n) + g(x_n) \rightarrow f(x) + g(x)$.

Als Konsequenz von Satz 2.10 erhalten wir, dass Polynome

$$p(x) = \sum_{k=0}^n a_k x^k, \quad a_k \in \mathbb{R} \text{ gegeben,}$$

und rationale Funktionen

$$r(x) = \frac{p(x)}{q(x)}, \quad p, q \text{ Polynome,}$$

stetig sind auf ihren Definitionsbereichen.

Die Komposition (Hintereinanderausführung) stetiger Funktionen ist ebenfalls stetig. Genauer:

Satz 2.11 *Seien $f : D_f \rightarrow \mathbb{R}$ und $g : D_g \rightarrow \mathbb{R}$ mit $f(D_f) \subset D_g$. Ist f stetig in x und g stetig in $f(x)$, so ist $g \circ f : D_f \rightarrow \mathbb{R}$ stetig in x .*

Aus $x_n \rightarrow x$ folgt $f(x_n) \rightarrow f(x)$ und weiter $g(f(x_n)) \rightarrow g(f(x))$.

Wollen wir zwei stetige Funktionen $f_1 : [a, b] \rightarrow \mathbb{R}$ und $f_2 : [b, c] \rightarrow \mathbb{R}$ zu einer einzigen Funktion $f : [a, c] \rightarrow \mathbb{R}$ zusammensetzen, so kommt es auf das Verhalten am Punkt b an.

Satz 2.12 *Seien $f_1 : [a, b] \rightarrow \mathbb{R}$ und $f_2 : [b, c] \rightarrow \mathbb{R}$ stetig. Gilt $f_1(b) = f_2(b)$, so ist die durch*

$$f(x) = \begin{cases} f_1(x), & x \in [a, b], \\ f_2(x), & x \in [b, c], \end{cases} \quad (2.12)$$

definierte Funktion $f : [a, c] \rightarrow \mathbb{R}$ stetig.

Gilt in der Situation des Satzes $f_1(b) \neq f_2(b)$, so ist die zusammengesetzte Funktion f unstetig, egal wie man sie im Punkt b definiert. Dieser Fall gibt Anlass, einen “einseitigen” Grenzwert zu definieren.

Definition 2.13 Sei $f : D \rightarrow \mathbb{R}$. Wir sagen, dass f im Punkt $a \in D$ den linksseitigen Grenzwert c hat, geschrieben

$$\lim_{x \rightarrow a^-} f(x) = c, \quad (2.13)$$

falls $f(x_n) \rightarrow c$ gilt für jede Folge (x_n) in D mit $x_n \rightarrow a$ und $x_n < a$ für alle $n \in \mathbb{N}$. f heißt linksseitig stetig in a , falls

$$\lim_{x \rightarrow a^-} f(x) = f(a). \quad (2.14)$$

Analog wird der rechtsseitige Grenzwert

$$\lim_{x \rightarrow a^+} f(x)$$

definiert.

Für die Heaviside-Funktion

$$f(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0, \end{cases}$$

gilt beispielsweise, dass

$$\lim_{x \rightarrow 0^-} f(x) = 0, \quad \lim_{x \rightarrow 0^+} f(x) = 1.$$

Da $f(0) = 0$, ist f in 0 linksseitig stetig, aber nicht rechtsseitig stetig.

Eine Funktion $f : D \rightarrow \mathbb{R}$ hat in a den Grenzwert c im Sinne eines beidseitigen Grenzwerts,

$$\lim_{x \rightarrow a} f(x) = c, \quad (2.15)$$

falls $f(x_n) \rightarrow c$ gilt für jede Folge (x_n) in D mit $x_n \rightarrow a$.

Fixpunktiteration. Sie stellt eine Methode zur Lösung von Gleichungssystemen (und damit einen wesentlichen Baustein von Simulationen) dar. Wir wollen lösen

$$x = f(x). \quad (2.16)$$

Eine Lösung x von (2.16) heißt ein **Fixpunkt** von f . Wir setzen die Iteration

$$x_{n+1} = f(x_n), \quad x_0 \text{ gegeben}, \quad (2.17)$$

an und hoffen, dass x_n gegen eine Lösung konvergiert.

Bei Gleichungssystemen ist x ein Vektor; wir betrachten jetzt nur den Fall $x \in \mathbb{R}$, also eine Gleichung mit einer Unbekannten. Die folgenden Überlegungen lassen sich aber ungeändert auch auf Systeme übertragen.

Damit die Iteration innerhalb einer gewissen Menge $D \subset \mathbb{R}$ wohldefiniert ist, muss $f(D) \subset D$ gelten. Ist f außerdem stetig in D , so können wir schließen: Falls die Iteration (2.17) überhaupt konvergiert, so ist der Grenzwert eine Lösung von (2.16).

Gilt nämlich $x_n \rightarrow x$, so auch $f(x_n) \rightarrow f(x)$ und $x_{n+1} \rightarrow x$. Da $x_{n+1} = f(x_n)$ und Grenzwerte eindeutig bestimmt sind, folgt $x = f(x)$.

Wir untersuchen jetzt nicht allgemein die Frage, wann (2.17) konvergent ist, sondern betrachten ein Beispiel, nämlich die Bestimmung der positiven Quadratwurzel \sqrt{b} einer reellen Zahl $b > 0$. Sie ist Lösung von

$$x^2 = b. \quad (2.18)$$

Dazu äquivalent ist

$$x^2 = \frac{1}{2}x^2 + \frac{1}{2}b,$$

oder (Division durch x)

$$x = \frac{1}{2}\left(x + \frac{b}{x}\right) =: f(x). \quad (2.19)$$

Diese Gleichung ist eine Fixpunktgleichung, die Funktion f ist stetig. Wir untersuchen die Konvergenz der zugehörigen Fixpunktiteration

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{b}{x_n}\right), \quad x_0 = b > 0. \quad (2.20)$$

Es gilt $x_n > 0$ für alle $n \in \mathbb{N}$ (Beweis mit Induktion), und

$$\begin{aligned} x_n^2 - b &= \frac{1}{4}\left(x_{n-1} + \frac{b}{x_{n-1}}\right)^2 - b = \frac{1}{4}\left(x_{n-1}^2 + 2b + \frac{b^2}{x_{n-1}^2}\right) - b \\ &= \frac{1}{4}\left(x_{n-1}^2 - 2b + \frac{b^2}{x_{n-1}^2}\right) = \frac{1}{4}\left(x_{n-1} - \frac{b}{x_{n-1}}\right)^2 \\ &\geq 0, \end{aligned}$$

also $x_n^2 \geq b$ für alle $n \geq 1$, und weiter

$$x_n - x_{n+1} = x_n - \frac{1}{2}\left(x_n + \frac{b}{x_n}\right) = \frac{1}{2x_n}(x_n^2 - b) \geq 0$$

für alle $n \geq 1$, also ist (x_n) ab dem zweiten Folgenglied x_1 monoton fallend. Nach Satz 2.7 konvergiert (x_n) gegen $x := \inf x_n$. Aus Lemma 2.3 folgt $x \geq 0$ und $x^2 \geq b > 0$, also $x > 0$. Da f stetig ist, gilt $x = f(x)$ für den Grenzwert, also $x^2 = b$ und damit

$$x = \sqrt{b}.$$

In der Tat, das so als Grenzwert erhaltene x ist die einzige positive Lösung von $x^2 = b$:

Ist $y \in \mathbb{R}$ eine weitere Zahl mit $y^2 = b$, so folgt wegen

$$0 = x^2 - y^2 = (x - y)(x + y)$$

entweder $x - y = 0$ oder $x + y = 0$, also $y = x$ oder $y = -x$.

Die Konvergenz der Fixpunktiteration ist in diesem Beispiel sehr schnell, so ist etwa für $b = 2$ nach dem vierten Schritt

$$x_4 = 1.414213562$$

eine Näherung für $\sqrt{2}$ mit einem Fehler von höchstens 10^{-9} . Im Gegensatz dazu ist die Konvergenz der Folge $1/n$ gegen 0 sehr langsam (Fehler 10^{-3} nach 1000 Schritten).

Konvergenz im Mehrdimensionalen. Sei \mathbb{R}^d der d -dimensionale Raum. Bei der Beschreibung von Folgen im \mathbb{R}^d benötigt man zwei Indizes, einen für die Numerierung des Vektors als Folgeelement, einen für die Komponenten des Vektors. Sei $(x^n)_{n \in \mathbb{N}}$ eine Folge von Vektoren im \mathbb{R}^d , jeder einzelne Vektor x^n habe die Komponenten

$$x^n = (x_1^n, x_2^n, \dots, x_d^n).$$

Definition 2.14 Eine Folge $(x^n)_{n \in \mathbb{N}}$ im \mathbb{R}^d heißt konvergent gegen ein $a \in \mathbb{R}^d$, falls gilt

$$\lim_{n \rightarrow \infty} \|x^n - a\| = 0. \quad (2.21)$$

In diesem Fall schreiben wir ebenfalls

$$a = \lim_{n \rightarrow \infty} x^n. \quad (2.22)$$

Es wird also verlangt, dass die Abstände von x^n zu a eine gegen 0 konvergente Folge bilden. Ein Beispiel im \mathbb{R}^2 ist

$$x^n = \left(\frac{1}{n}, \frac{2}{n} \right). \quad (2.23)$$

Hier ist

$$\|x^n - a\| = \|x_n\| = \sqrt{\frac{1}{n^2} + \frac{4}{n^2}} = \frac{\sqrt{5}}{n} \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

also konvergiert die Folge (x^n) gegen 0. In diesem Beispiel bewegt sich die Folge entlang einer Geraden auf den Nullpunkt zu. Das wird aber für Konvergenz nicht verlangt. Betrachten wir im \mathbb{R}^2 die Folge (x^n) definiert durch

$$\left(1, 0 \right), \left(0, \frac{1}{2} \right), \left(-\frac{1}{3}, 0 \right), \left(0, -\frac{1}{4} \right), \left(\frac{1}{5}, 0 \right) \dots \quad (2.24)$$

so gilt $\|x^n - 0\| = 1/n \rightarrow 0$, also $x^n \rightarrow 0$, die Konvergenz erfolgt spiralförmig.

Um Konvergenz festzustellen, muss man nicht unbedingt Abstände ausrechnen, oft genügt es, das folgende Kriterium anzuwenden.

Satz 2.15 Eine Folge $(x^n)_{n \in \mathbb{N}}$ im \mathbb{R}^d konvergiert genau dann gegen ein $a \in \mathbb{R}^d$, wenn alle Komponentenfolgen $(x_k^n)_{n \in \mathbb{N}}$ gegen a_k konvergieren, $1 \leq k \leq d$.

Man vergleiche mit den genannten Beispielen (2.23) und (2.24).

Die Definition der Stetigkeit einer Funktion lässt sich fast wörtlich aus dem Eindimensionalen übernehmen.

Definition 2.16 (Stetigkeit im Mehrdimensionalen)

Eine Funktion $f : D \rightarrow \mathbb{R}^m$ mit Definitionsbereich $D \subset \mathbb{R}^d$ heißt stetig in einem Punkt $x \in D$, falls für jede Folge (x^n) in D mit $x^n \rightarrow x$ gilt

$$\lim_{n \rightarrow \infty} f(x^n) = f(x). \quad (2.25)$$

f heißt stetig in D , falls f in jedem Punkt D stetig ist.

Die Projektionen $p_k : \mathbb{R}^d \rightarrow \mathbb{R}$, $p_k(x) = x_k$ sind stetig in jedem Punkt $a \in \mathbb{R}^d$, da für jede Folge (x^n) mit $x^n \rightarrow a$ gilt, dass

$$|p_k(x^n) - p_k(a)| = |x_k^n - a_k| \leq \|x^n - a\|,$$

und daher aus $x^n \rightarrow a$ folgt, dass $p_k(x^n) \rightarrow p_k(a)$. Da Summen und Produkte stetiger Funktionen wieder stetig sind, sind Polynome in mehreren Variablen, beispielsweise $f : \mathbb{R}^3 \rightarrow \mathbb{R}$,

$$f(x) = 2x_1x_2^2 - 4x_2x_3 - 1,$$

stetig.

Da Stetigkeit unmittelbar auf Folgenkonvergenz zurückgeführt wird, ist es nicht überraschend, dass das Analogon zu Satz 2.15 auch für Stetigkeit gilt. Jede Funktion $f : D \rightarrow \mathbb{R}^m$ setzt sich zusammen aus Komponentenfunktionen $f_i : D \rightarrow \mathbb{R}$,

$$f(x) = (f_1(x), \dots, f_m(x)).$$

Satz 2.17 Eine Funktion $f : D \rightarrow \mathbb{R}^m$ mit Definitionsbereich $D \subset \mathbb{R}^d$ ist genau dann stetig in einem Punkt $x \in D$, falls alle Komponentenfunktionen $f_i : D \rightarrow \mathbb{R}$, $1 \leq i \leq m$, in x stetig sind.

Beispielsweise ist $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, definiert durch

$$f(x) = (2x_1x_2^2 - 4x_2x_3 - 1, x_3^2 - x_1 + 4),$$

stetig.

Das Verhalten unstetiger Funktionen im Mehrdimensionalen kann recht komplex sein. Als Beispiel betrachten wir eine Situation in der Bildverarbeitung. Ein quadratisches 2D-Bild in SW wird beschrieben durch eine Funktion $f : Q \rightarrow [0, 1]$ mit $Q = [0, 1] \times [0, 1]$, wobei der Grauwert jedes Punktes $x \in Q$ durch den Funktionswert $f(x)$ gegeben ist. Eine Kante im Bild entspricht einem Geradenstück, entlang derer f unstetig ist. Stoßen in einem Punkt a mehrere Kanten zusammen, so kann f bei Annäherung an a unterschiedliche Grenzwerte annehmen, je nachdem aus welchem Flächenstück man kommt. Ebenso kann oszillatorisches Verhalten in unterschiedlicher Ausprägung auftreten.

3 Reihen

Wir beginnen mit einem Beispiel. Aus einer Folge, etwa

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots$$

bilden wir durch Summation eine neue Folge

$$1, \quad 1 + \frac{1}{2} = \frac{3}{2}, \quad 1 + \frac{1}{2} + \frac{1}{4} = \frac{7}{4}, \quad \dots$$

also

$$1, \frac{3}{2}, \frac{7}{4}, \frac{15}{8}, \frac{31}{16}, \dots$$

Im allgemeinen Fall gehen wir aus von einer Folge reeller oder komplexer Zahlen

$$a_0, a_1, a_2 \dots$$

und bilden aus den **Partialsommen**

$$s_n = a_0 + \dots + a_n = \sum_{k=0}^n a_k \quad (3.1)$$

eine Folge (s_n) , diese heißt **Reihe** oder auch unendliche Reihe mit den Gliedern a_k . Die Reihe heißt **konvergent**, falls (s_n) konvergent ist, andernfalls heißt sie **divergent**. Ihr Grenzwert $s = \lim_{n \rightarrow \infty} s_n$ heißt der **Wert** oder die **Summe** der Reihe, geschrieben als

$$\sum_{k=0}^{\infty} a_k. \quad (3.2)$$

Im Falle der Konvergenz ist (3.2) also eine reelle bzw. komplexe Zahl. Oft schreibt man (3.2) als Bezeichnung für die Reihe (d.h. für die Folge der Partialsommen) selbst.

Die geometrische Reihe. Für $z \in \mathbb{C}$ betrachten wir die Reihe

$$\sum_{k=0}^{\infty} z^k, \quad (3.3)$$

dabei ist $z^0 = 1$ gesetzt. Für die Partialsommen gilt

$$(1 - z)s_n = (1 - z) \sum_{k=0}^n z^k = (1 - z) + (z - z^2) + \dots + (z^n - z^{n+1}) = 1 - z^{n+1},$$

also

$$s_n = \frac{1 - z^{n+1}}{1 - z}. \quad (3.4)$$

Für $|z| < 1$ gilt $|z|^{n+1} \rightarrow 0$, also auch $z^{n+1} \rightarrow 0$. Die Reihe (3.3) konvergiert für $|z| < 1$, und ihre Summe ist

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1 - z}. \quad (3.5)$$

Beispiele für geometrische Reihen sind

$$\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = 2, \quad \sum_{k=0}^{\infty} \left(\frac{1}{100}\right)^k = \frac{100}{99}, \quad \sum_{k=0}^{\infty} \left(\frac{99}{100}\right)^k = 100.$$

Die Teleskopreihe. Wir betrachten

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)}. \quad (3.6)$$

(Hier ist der Index des ersten Folgenglieds 1 statt 0.) Wegen

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

gilt

$$s_n = \sum_{k=1}^n \frac{1}{k(k+1)} = 1 - \frac{1}{n+1} \rightarrow 1 \quad \text{für } n \rightarrow \infty,$$

also

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1. \quad (3.7)$$

Weglassen endlich vieler Glieder. Für Frage, ob eine Reihe konvergiert, ist genau wie bei einer Folge das “Verhalten am Anfang” gleichgültig: Für beliebiges $m \geq n_0$ gilt, dass

$$\sum_{k=n_0}^{\infty} a_k$$

konvergiert genau dann, wenn

$$\sum_{k=m}^{\infty} a_k$$

konvergiert. Im Gegensatz zur Situation bei Folgen ändert sich aber der Grenzwert (die Summe der Reihe), wenn einzelne Glieder weggelassen oder hinzugefügt werden, beispielsweise ist

$$\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = 2, \quad \sum_{k=3}^{\infty} \left(\frac{1}{2}\right)^k = \frac{1}{4}.$$

Die harmonische Reihe. Die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k} \quad (3.8)$$

heißt die harmonische Reihe. Für $N = 2^j$ gilt

$$s_N = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \dots + \frac{1}{8}\right) + \dots \\ \dots + \left(\frac{1}{2^{j-1} + 1} + \dots + \frac{1}{2^j}\right)$$

Da jeder der eingeklammerten Ausdrücke größer ist als $1/2$, folgt

$$s_N \geq 1 + \frac{j}{2}, \quad \text{falls } N = 2^j. \quad (3.9)$$

Die harmonische Reihe ist also divergent. Da alle Reihenglieder positiv sind, sind die Partialsummen monoton wachsend und wegen (3.9) uneigentlich konvergent gegen $+\infty$. Wir schreiben daher

$$\sum_{k=1}^{\infty} \frac{1}{k} = +\infty. \quad (3.10)$$

Summe und skalares Vielfaches von Reihen. Sind die Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ konvergent, so ist auch deren Summe konvergent, und es gilt

$$\sum_{k=0}^{\infty} (a_k + b_k) = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k. \quad (3.11)$$

Entsprechend gilt auch für ein skalares Vielfaches

$$\sum_{k=0}^{\infty} ca_k = c \sum_{k=0}^{\infty} a_k. \quad (3.12)$$

(3.11) und (3.12) ergeben sich, da für die Folgen der Partialsummen gilt, dass Grenzwertbildung mit Summe und skalarem Vielfachen vertauschbar sind (siehe Abschnitt über Folgen).

Wir betrachten als Beispiel die Umwandlung einer rationalen Zahl aus der periodischen Dezimaldarstellung in einen Bruch, etwa

$$\begin{aligned} 0.01777\cdots &= 1 \cdot 10^{-2} + 7 \cdot 10^{-3} + 7 \cdot 10^{-4} + \dots = 10^{-2} + \sum_{k=0}^{\infty} 7 \cdot 10^{-3} \cdot \left(\frac{1}{10}\right)^k \\ &= \frac{1}{100} + \frac{7}{1000} \cdot \frac{1}{1 - \frac{1}{10}} = \frac{4}{225}. \end{aligned}$$

Ist eine Reihe konvergent, also $s_n \rightarrow s$ für eine reelle bzw. komplexe Zahl s , so gilt für die Glieder

$$a_n = s_n - s_{n-1} = (s_n - s) + (s - s_{n-1}) \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Satz 3.1 (Notwendige Bedingung für Konvergenz)

Ist eine Reihe $\sum_{k=0}^{\infty} a_k$ konvergent, so muss gelten

$$\lim_{k \rightarrow \infty} a_k = 0. \quad (3.13)$$

□

Umgekehrt folgt aus der Konvergenz $a_k \rightarrow 0$ aber **nicht**, dass die zugehörige Reihe konvergiert, wie das Beispiel der harmonischen Reihe zeigt.

Satz 3.2 Ist $\sum_{k=0}^{\infty} a_k$ eine Reihe mit $a_k \geq 0$ für alle $k \geq 0$, so ist sie konvergent genau dann, wenn die Folge (s_n) ihrer Partialsummen beschränkt ist.

Die Folge der Partialsummen ist monoton wachsend. Monoton wachsende Folgen sind konvergent genau dann, wenn sie beschränkt sind.

Kriterien für Konvergenz. Ob eine Reihe konvergiert oder divergiert, kann man oft feststellen, indem man sie mit einer Reihe vergleicht, deren Konvergenz oder Divergenz man bereits kennt.

Definition 3.3 (Majorante)

Sei $\sum_{k=0}^{\infty} a_k$ eine Reihe in \mathbb{C} . Eine reelle Reihe $\sum_{k=0}^{\infty} b_k$ mit $|a_k| \leq b_k$ für alle $k \geq 0$ heißt Majorante der Reihe $\sum_{k=0}^{\infty} a_k$.

Satz 3.4 (Majorantenkriterium)

Sei $\sum_{k=0}^{\infty} a_k$ eine Reihe in \mathbb{C} , welche eine konvergente Majorante $\sum_{k=0}^{\infty} b_k$ besitzt. Dann ist $\sum_{k=0}^{\infty} a_k$ konvergent, und es gilt

$$\left| \sum_{k=0}^{\infty} a_k \right| \leq \sum_{k=0}^{\infty} |a_k| \leq \sum_{k=0}^{\infty} b_k. \quad (3.14)$$

Für die Partialsummen $s_n = \sum_{k=0}^n a_k$ gilt, falls $m \geq n$,

$$|s_m - s_n| = \left| \sum_{k=n+1}^m a_k \right| \leq \sum_{k=n+1}^m |a_k| \leq \sum_{k=n+1}^m b_k.$$

Wir bilden das Supremum bezüglich m und erhalten, da $\sum_k b_k$ als konvergent vorausgesetzt ist,

$$\sup_{m \geq n} |s_m - s_n| \leq \sum_{k=n+1}^{\infty} b_k \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

Warum hieraus die Konvergenz von (s_n) folgt, behandeln wir nicht; man führt das auf die Vollständigkeit von \mathbb{R} zurück. Weiter gilt

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

Mit Grenzübergang $n \rightarrow \infty$ erhalten wir (3.14).

Als Beispiel betrachten wir

$$\sum_{k=1}^{\infty} \frac{1}{k^2}. \quad (3.15)$$

Die Teleskopreihe

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

ist eine konvergente Majorante von

$$\sum_{k=2}^{\infty} \frac{1}{k^2}, \quad (3.16)$$

also ist (3.16) konvergent und damit auch (3.15).

Folgerung 3.5 (Divergente Minorante)

Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ Reihen in \mathbb{R} mit $0 \leq a_k \leq b_k$ für alle $k \geq 0$. Ist $\sum_{k=0}^{\infty} a_k$ divergent, so ist auch $\sum_{k=0}^{\infty} b_k$ divergent.

Wäre $\sum_{k=0}^{\infty} b_k$ konvergent, so wäre nach dem Majorantenkriterium ebenso auch $\sum_{k=0}^{\infty} a_k$ konvergent.

Die Reihe $\sum_{k=0}^{\infty} a_k$ heißt in diesem Fall eine divergente Minorante von $\sum_{k=0}^{\infty} b_k$.

Aus dem Vergleich mit der geometrischen Reihe erhalten wir das Quotientenkriterium.

Satz 3.6 (Quotientenkriterium)

Sei $\sum_{k=0}^{\infty} a_k$ eine Reihe in \mathbb{C} . Es gebe ein $q \in \mathbb{R}$ mit $q < 1$ und ein $n_0 \geq 0$, so dass

$$\frac{|a_{k+1}|}{|a_k|} \leq q, \quad \text{für alle } k \geq n_0. \quad (3.17)$$

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

Das folgt aus dem Majorantenkriterium, da $|a_k| \leq |a_{n_0}|q^{k-n_0}$ für $k \geq n_0$ und daher die geometrische Reihe

$$\sum_{k=n_0}^{\infty} |a_{n_0}|q^{k-n_0}$$

eine konvergente Majorante von $\sum_{k=n_0}^{\infty} a_k$ ist.

Vorsicht: Die genaue Formulierung (“mit q ”) ist wesentlich. Um die Konvergenz zu garantieren, genügt es **nicht**, dass $|a_{k+1}|/|a_k| < 1$ gilt. Beispielsweise gilt für die harmonische Reihe $a_k = 1/k$, dass

$$\frac{|a_{k+1}|}{|a_k|} = \frac{\frac{1}{k+1}}{\frac{1}{k}} = \frac{k}{k+1} < 1,$$

aber sie ist divergent, wie wir bereits wissen.

Wir betrachten als Beispiel die **Exponentialreihe**

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}, \quad (3.18)$$

wobei $z \in \mathbb{C}$ eine feste komplexe Zahl ist. Mit $a_k = z^k/k!$ gilt

$$\frac{|a_{k+1}|}{|a_k|} = \frac{|z|^{k+1}}{(k+1)!} \cdot \frac{k!}{|z|^k} = \frac{|z|}{k+1},$$

also

$$\frac{|a_{k+1}|}{|a_k|} \leq \frac{1}{2}, \quad \text{für alle } k \geq 2|z| - 1,$$

also ist (3.18) konvergent. Sie definiert die **Exponentialfunktion**

$$\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} \quad (3.19)$$

als Funktion

$$\exp : \mathbb{C} \rightarrow \mathbb{C}.$$

Aus der Definition folgt unmittelbar, dass $\exp(z) \in \mathbb{R}$, falls $z \in \mathbb{R}$. Schränken wir \exp auf \mathbb{R} ein, so erhalten wir die reelle Exponentialfunktion. Aus (3.19) folgt

$$\exp(0) = 1.$$

Wir definieren die Zahl e durch

$$e = \exp(1). \quad (3.20)$$

Es ist

$$e = 2.718281828459045235 \dots$$

Alternierende Reihen. Eine Reihe $\sum_{k=n_0}^{\infty} a_k$ heißt **alternierend**, wenn die Folgenglieder abwechselnd positives und negatives Vorzeichen haben. Wir schreiben sie in der Form

$$a_0 - a_1 + a_2 - a_3 + \dots$$

mit $a_k \geq 0$.

Satz 3.7 (Leibnizkriterium)

Sei $(a_n)_{n \geq 0}$ eine monoton fallende Folge in \mathbb{R} mit $a_n \rightarrow 0$. Dann konvergiert die Reihe

$$\sum_{k=0}^{\infty} (-1)^k a_k, \quad (3.21)$$

und es gilt für alle $n \geq 0$

$$\left| \sum_{k=0}^{\infty} (-1)^k a_k - s_n \right| = \left| \sum_{k=0}^{\infty} (-1)^k a_k - \sum_{k=0}^n (-1)^k a_k \right| \leq a_{n+1}. \quad (3.22)$$

Wir beweisen den Satz. Für $k \geq 1$ ist $a_{2k} - a_{2k-1} \leq 0$ und $a_{2k} - a_{2k+1} \geq 0$, also

$$s_{2k} = s_{2k-2} - a_{2k-1} + a_{2k} \leq s_{2k-2},$$

$$s_{2k+1} = s_{2k-1} + a_{2k} - a_{2k+1} \geq s_{2k-1},$$

also ist $(s_{2k})_{k \in \mathbb{N}}$ monoton fallend, $(s_{2k-1})_{k \in \mathbb{N}}$ monoton wachsend. Es gilt

$$s_1 \leq s_{2k-1} = s_{2k} - a_{2k} \leq s_{2k} \leq s_0 \quad (3.23)$$

für alle k , und weiter für alle $m \geq k$

$$s_1 \leq s_{2k-1} \leq s_{2m-1} \leq s_{2m} \leq s_{2k} \leq s_0. \quad (3.24)$$

Also sind beide Teilfolgen (s_{2k}) und (s_{2k-1}) beschränkt, also konvergent nach Satz 2.7, und

$$s_{2k-1} \leq \underbrace{\lim_{m \rightarrow \infty} s_{2m-1}}_{=:b} \leq s_{2k}, \quad s_{2k-1} \leq \underbrace{\lim_{m \rightarrow \infty} s_{2m}}_{=:c} \leq s_{2k} \quad (3.25)$$

gilt für alle $k \in \mathbb{N}$. Aus (3.24) folgt

$$0 \leq |c - b| \leq a_{2k},$$

und wegen $a_{2k} \rightarrow 0$ gilt $c = b$. Da für alle n entweder $s_n \leq b \leq s_{n+1}$ oder $s_n \geq b \geq s_{n+1}$ gilt, folgt

$$0 \leq |b - s_n| \leq |s_n - s_{n+1}| = a_{n+1} \rightarrow 0,$$

also $|b - s_n| \rightarrow 0$ nach dem Einschließungssatz für Folgen. \square

Beispiel: Die alternierende harmonische Reihe

$$\sum_{k=1}^{\infty} (-1)^{k-1} \frac{1}{k} \quad (3.26)$$

ist nach dem Leibniz-Kriterium konvergent. (Ihr Grenzwert ist, wie sich später herausstellen wird, die Zahl $\ln 2$.) Wir betrachten (3.26) etwas näher. Die aus den Glieder mit geraden Indizes gebildete Reihe

$$\sum_{k=1}^{\infty} -\frac{1}{2k} \quad (3.27)$$

ist divergent (andernfalls wäre die harmonische Reihe konvergent), die aus den Glieder mit ungeraden Indizes gebildete Reihe

$$\sum_{k=1}^{\infty} \frac{1}{2k-1} \quad (3.28)$$

ist ebenfalls divergent (die Reihe $\sum_{k=1}^{\infty} 1/(2k)$ ist eine divergente Minorante). Geben wir nun eine beliebige Zahl $a \in \mathbb{R}$ vor, so können wir durch Umordnen erreichen, dass die umgeordnete Reihe gegen a konvergiert, z.B. für $a = 1$ können wir betrachten

$$1 - \frac{1}{2} + \frac{1}{3} + \frac{1}{5} - \frac{1}{4} + \frac{1}{7} \dots$$

(wir nehmen negative Glieder so lange, bis die Partialsumme kleiner als 1 ist, dann positive Glieder so lange, bis die Partialsumme größer als 1 ist, usw.). Ebenso kann man durch Umordnen erreichen, dass die umgeordnete Reihe divergiert.

Im Gegensatz zu einer Summe endlich vieler Zahlen kann also bei einer unendlichen Summe die Konvergenz verlorengehen oder der Wert der Summe sich ändern, wenn man eine Reihe umordnet.

Bei absolut konvergenten Reihen hingegen kann dieses Phänomen nicht auftreten.

Definition 3.8 (Absolute Konvergenz)

Eine Reihe $\sum_{k=0}^{\infty} a_k$ mit $a_k \in \mathbb{C}$ heißt absolut konvergent, falls

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

konvergent ist.

Wir können die Umordnung auf folgende Weise formal fassen. Ist $\tau : \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung, so wird durch

$$a_{\tau(0)}, a_{\tau(1)}, a_{\tau(2)}, \dots$$

eine Umordnung der Folge a_0, a_1, a_2, \dots definiert, sowie durch

$$\sum_{j=0}^{\infty} a_{\tau(j)}$$

eine **Umordnung der Reihe** $\sum_{k=0}^{\infty} a_k$.

Satz 3.9 Sei $\sum_{k=0}^{\infty} a_k$ eine Reihe in \mathbb{C} , sei $\tau : \mathbb{N} \rightarrow \mathbb{N}$ bijektiv. Ist $\sum_{k=0}^{\infty} a_k$ absolut konvergent, so ist auch $\sum_{j=0}^{\infty} a_{\tau(j)}$ absolut konvergent, und es gilt

$$\sum_{j=0}^{\infty} a_{\tau(j)} = \sum_{k=0}^{\infty} a_k. \quad (3.29)$$

Diesen Beweis lassen wir aus.

Sind p, q Polynome,

$$p(x) = \sum_{k=0}^n a_k x^k, \quad q(x) = \sum_{k=0}^m b_k x^k, \quad (3.30)$$

so hat deren Produkt die Form

$$(pq)(x) = \sum_{k=0}^{n+m} c_k x^k, \quad c_k = \sum_{j=0}^k a_{k-j} b_j. \quad (3.31)$$

Der Wunsch, diese Formel auf Potenzreihen der Form

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

zu übertragen, führt auf das sogenannte Cauchy-Produkt von Reihen.

Satz 3.10 Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergente Reihen in \mathbb{C} , sei für $k \in \mathbb{N}$

$$c_k = \sum_{j=0}^k a_{k-j} b_j. \quad (3.32)$$

Dann ist auch die Reihe $\sum_{k=0}^{\infty} c_k$ absolut konvergent, und es gilt

$$\sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right). \quad (3.33)$$

Auch diesen Beweis lassen wir weg.

Wenden wir den vorangehenden Satz auf die Exponentialreihe an (das Argument, das die Konvergenz zeigte, zeigt auch die absolute Konvergenz), so erhalten wir eine für das Rechnen mit der Exponentialfunktion fundamentale Gleichung. Wir erinnern an die binomische Formel

$$(x + y)^n = \sum_{j=0}^n \binom{n}{j} x^{n-j} y^j, \quad (3.34)$$

gültig für alle $x, y \in \mathbb{C}$ und alle $n \in \mathbb{N}$.

Satz 3.11 *Es gilt*

$$\exp(z + w) = \exp(z) \exp(w) \quad (3.35)$$

für alle $z, w \in \mathbb{C}$.

Mit

$$a_k = \frac{z^k}{k!}, \quad b_k = \frac{w^k}{k!}, \quad c_k = \sum_{j=0}^k a_{k-j} b_j$$

gilt wegen Satz 3.10 und der binomischen Formel (3.34)

$$\begin{aligned} \exp(z + w) &= \sum_{k=0}^{\infty} \frac{(z + w)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} z^{k-j} w^j \\ &= \sum_{k=0}^{\infty} \sum_{j=0}^k \frac{z^{k-j}}{(k-j)!} \frac{w^j}{j!} = \sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} a_k \right) \left(\sum_{k=0}^{\infty} b_k \right) \\ &= \left(\sum_{k=0}^{\infty} \frac{z^k}{k!} \right) \cdot \left(\sum_{k=0}^{\infty} \frac{w^k}{k!} \right) = \exp(z) \exp(w). \end{aligned}$$

Wir stellen einige Eigenschaften der Exponentialfunktion zusammen.

Folgerung 3.12 *Es gelten*

$$\exp(-z) = \frac{1}{\exp(z)}, \quad \text{für alle } z \in \mathbb{C}, \quad (3.36)$$

$$\exp(z) \neq 0, \quad \text{für alle } z \in \mathbb{C}, \quad (3.37)$$

$$\exp(x) > 0, \quad \text{für alle } x \in \mathbb{R}, \quad (3.38)$$

$$\exp(n) = e^n, \quad \text{für alle } n \in \mathbb{N}, \quad (3.39)$$

$$\overline{\exp(z)} = \exp(\bar{z}), \quad \text{für alle } z \in \mathbb{C}. \quad (3.40)$$

Die reelle Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$ ist monoton wachsend, und für die komplexe Funktion gilt

$$|\exp(z)| \leq \exp(|z|), \quad \text{für alle } z \in \mathbb{C}. \quad (3.41)$$

Aus

$$1 = \exp(0) = \exp(z - z) = \exp(z) \exp(-z) \quad (3.42)$$

folgen (3.36) und (3.37). Aus der Definition der Exponentialfunktion folgt

$$\exp(x) \geq 1 > 0$$

für alle $x \in \mathbb{R}$, $x \geq 0$, also wegen (3.42) auch $\exp(x) > 0$ für $x < 0$. Aus

$$\exp(n) = \exp\left(\sum_{k=1}^n 1\right) = \prod_{k=1}^n \exp(1)$$

folgt (3.39). Weiter ergibt sich (3.40) aus

$$\exp(\bar{z}) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{\bar{z}^k}{k!} = \lim_{n \rightarrow \infty} \overline{\sum_{k=0}^n \frac{z^k}{k!}} = \overline{\exp(z)}.$$

Sind $0 \leq x \leq y \in \mathbb{R}$, so können wir schließen

$$\begin{aligned} 0 \leq x \leq y &\Rightarrow \frac{x^k}{k!} \leq \frac{y^k}{k!} \quad \text{für alle } k \\ &\Rightarrow \exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \leq \sum_{k=0}^{\infty} \frac{y^k}{k!} = \exp(y). \end{aligned} \quad (3.43)$$

Ist $x \leq y \leq 0$, so ist $0 \leq -y \leq -x$, und mit (3.43) folgt

$$\frac{1}{\exp(y)} = \exp(-y) \leq \exp(-x) = \frac{1}{\exp(x)},$$

also ebenfalls $\exp(x) \leq \exp(y)$. Die Ungleichung (3.41) erhalten wir aus dem Majorantenkriterium, da die Exponentialreihe für $\exp(|z|)$ eine konvergente Majorante der Exponentialreihe für $\exp(z)$ ist und nach (3.14) gilt, dass

$$|\exp(z)| = \left| \sum_{k=0}^{\infty} \frac{z^k}{k!} \right| \leq \sum_{k=0}^{\infty} \frac{|z|^k}{k!} = \exp(|z|).$$

Wir stellen die Frage, ob die Exponentialfunktion stetig ist. Für $z \in \mathbb{C}$ mit $|z| \leq 1$ gilt

$$\begin{aligned} 0 \leq |\exp(z) - 1| &= \left| \sum_{k=1}^{\infty} \frac{z^k}{k!} \right| \leq \sum_{k=1}^{\infty} \frac{|z|^k}{k!} \leq |z| \sum_{k=1}^{\infty} \frac{|z|^{k-1}}{(k-1)!} = |z| \exp(|z|) \\ &\leq |z| \exp(1). \end{aligned}$$

Aus dem Einschließungssatz folgt, dass $|\exp(z_n) - 1| \rightarrow 0$ gilt für jede Folge (z_n) mit $|z_n| \rightarrow 0$ und damit

$$\lim_{n \rightarrow \infty} \exp(z_n) = 1 = \exp(0)$$

für jede solche Folge. Daraus erkennen wir, dass die Exponentialfunktion im Nullpunkt stetig ist. Sei nun $z \in \mathbb{C}$ beliebig und (z_n) eine beliebige Folge in \mathbb{C} mit $z_n \rightarrow z$. Dann gilt $|z_n - z| \rightarrow 0$ und wegen

$$\begin{aligned} |\exp(z_n) - \exp(z)| &= |\exp(z_n - z + z) - \exp(z)| = |\exp(z_n - z) \exp(z) - \exp(z)| \\ &= |(\exp(z_n - z) - 1) \cdot \exp(z)| \rightarrow 0 \end{aligned}$$

folgt, dass \exp auch in z stetig ist.

Satz 3.13 *Die Exponentialfunktion ist stetig auf \mathbb{C} .*

Da die reelle Exponentialfunktion die Einschränkung von \exp auf \mathbb{R} ist, ist sie ebenfalls stetig.

Ab sofort werden wir statt $\exp(z)$ auch e^z schreiben. Das macht Sinn, da die Gleichung $\exp(z + w) = \exp(z) \exp(w)$ zu

$$e^{z+w} = e^z e^w$$

wird und $\exp(n) = e^n$ gilt für $n \in \mathbb{N}$ gemäß (3.39).

4 Trigonometrische Funktionen

Geometrische Sachverhalte in der Ebene und im Raum kann man mathematisch beschreiben mit Konzepten der Linearen Algebra (lineare Gleichungssysteme, Unterräume, Skalarprodukt, Determinante, ...) und der Analysis. Hier geht es um die Analysis.

Trigonometrische Funktionen sind ursprünglich für die Geometrie von Dreiecken entwickelt worden. Unabhängig davon sind sie ein wesentliches Werkzeug bei der Beschreibung von Schwingungsvorgängen (Stichwort Fourieranalyse) und bei vielen anderen Fragestellungen der Analysis.

Wir betrachten einen Punkt z auf dem Einheitskreis (Mittelpunkt 0, Radius 1) in der Ebene. Dreht man die horizontale Achse gegen den Uhrzeigersinn solange, bis man die Verbindungslinie von 0 nach z erreicht, so wird ein Winkel überstrichen; diesen bezeichnen wir mit x . Die trigonometrischen Funktionen Cosinus und Sinus entstehen dadurch, dass wir die kartesischen Koordinaten von z als Funktionen des Winkels x auffassen,

$$z = (\cos x, \sin x). \quad (4.1)$$

Offenbar ist

$$\sin(0) = 0, \quad \cos(0) = 1. \quad (4.2)$$

Den Winkel kann man messen in Grad (rechter Winkel 90 Grad, Halbkreis 180 Grad, Vollkreis 360 Grad) oder im **Bogenmaß**. Dieses gibt an die Länge des Kreisbogens auf dem Einheitskreis, der zum Winkel x gehört, also 2π für den Vollkreis, π für den Halbkreis und $\pi/2$ für den rechten Winkel. Die Verwendung von Grad als Einheit stammt aus dem Altertum; das Jahr hat näherungsweise 360 Tage, als Zahl ist 360 durch viele Zahlen teilbar, z.B. durch alle natürlichen Zahlen von 1 bis 12 mit Ausnahme von 7 und 11. In der Neuzeit hat sich zunehmend das Bogenmaß durchgesetzt, es ist eine mathematisch fundamentalere Größe und für das Rechnen mit trigonometrischen Funktionen und der Exponentialfunktion besser geeignet. Wir verwenden das Bogenmaß.

Auf die beschriebene Weise werden Winkel x mit $0 \leq x \leq 2\pi$ definiert. Winkel $x > 2\pi$ entsprechen einem Durchlaufen von mehr als einem Vollkreis, Winkel $x < 0$ entsprechen einem Durchlaufen in mathematisch negativem Sinn (in Uhrzeigerichtung). Da ein Vollkreisumlauf zum selben Punkt zurückkehrt, gilt

$$\sin(x + 2\pi) = \sin x, \quad \cos(x + 2\pi) = \cos x \quad (4.3)$$

für alle $x \in \mathbb{R}$. Sinus und Cosinus sind damit **periodisch** mit der **Periode** 2π . Weiterhin entnimmt man der Definition (4.1), dass

$$\sin(-x) = -\sin x, \quad \cos(-x) = \cos x \quad (4.4)$$

sowie

$$\sin\left(x + \frac{\pi}{2}\right) = \cos x, \quad \cos\left(x + \frac{\pi}{2}\right) = -\sin x \quad (4.5)$$

gilt für alle $x \in \mathbb{R}$.

Aus dem Satz des Pythagoras folgt

$$\sin^2 x + \cos^2 x = 1. \quad (4.6)$$

Weitere trigonometrische Funktionen sind Tangens und Cotangens, definiert durch

$$\tan x = \frac{\sin x}{\cos x}, \quad \cot x = \frac{\cos x}{\sin x} = \frac{1}{\tan x}. \quad (4.7)$$

Wir betrachten ein Bild (siehe Vorlesung). Aus ihm geht durch Flächenvergleich hervor, dass

$$\frac{1}{2} \sin x \cos x \leq \frac{x}{2} \leq \frac{1}{2} \tan x$$

gilt für $0 \leq x \leq \pi/2$, und wegen (4.4) auch für $-\pi/2 \leq x \leq 0$. Wir multiplizieren mit 2, dividieren durch $\sin x$ und bilden die Kehrwerte. Es ergibt sich

$$\cos x \leq \frac{\sin x}{x} \leq \frac{1}{\cos x}.$$

Da (wie wir gleich noch sehen werden) Sinus und Cosinus stetige Funktionen sind, gilt $\cos x \rightarrow \cos 0 = 1$ für $x \rightarrow 0$ wegen (4.2), also folgt aus dem Einschließungssatz (der sich von Folgen unmittelbar auf Funktionen überträgt), dass

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1. \quad (4.8)$$

Zusammenhang mit der komplexen Exponentialfunktion. Zu gegebenem $x \in \mathbb{R}$ betrachten wir die komplexe Zahl e^{ix} . Es gilt $\overline{e^{ix}} = e^{-ix}$, also

$$|e^{ix}|^2 = e^{ix} \overline{e^{ix}} = e^{ix} e^{-ix} = e^{ix-ix} = e^0 = 1,$$

also $|e^{ix}| = 1$, d.h. für $x \in \mathbb{R}$ liegt e^{ix} auf dem Einheitskreis in der komplexen Ebene. Im Jahr 1748 hat Leonhard Euler die später nach ihm benannte Formel

$$e^{ix} = \cos x + i \sin x \quad (4.9)$$

publiziert. Sie besagt geometrisch, dass der Punkt auf dem Einheitskreis zur Bogenlänge x , also $z = (\cos x, \sin x)$, gerade durch $z = e^{ix}$ gegeben ist.

Man kann die Eulerformel beweisen, indem man den Kreisbogen von 1 nach e^{ix} durch Streckenzüge S_n mit $n+1$ auf dem Einheitskreis liegenden Teilpunkten approximiert, deren Länge l_n unter Verwendung der Rechenregeln für die Exponentialfunktion berechnet und zeigt, dass $l_n \rightarrow x$ für $n \rightarrow \infty$.

Wir betrachten spezielle Werte von x . Es gilt

$$e^{i\frac{\pi}{2}} = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = i, \quad e^{i\pi} = e^{i\frac{\pi}{2}} \cdot e^{i\frac{\pi}{2}} = -1. \quad (4.10)$$

Eine Umformulierung der Eulerformel (4.9) ist

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}), \quad \sin x = \frac{1}{2i}(e^{ix} - e^{-ix}), \quad (4.11)$$

Da die Exponentialfunktion stetig ist, wie wir aus dem vorigen Kapitel wissen, erhalten wir aus (4.11) unmittelbar:

Satz 4.1 *Sinus, Cosinus, Tangens und Cotangens sind stetig auf \mathbb{R} mit Ausnahme (bei Tangens und Cotangens) der Polstellen.*

Mit Hilfe der Eulerformel lassen sich eine Reihe von trigonometrischen Formeln auf einfache Weise erhalten. Beispielsweise führt die Identität

$$e^{i(x+y)} = e^{ix} e^{iy}$$

auf

$$\cos(x+y) + i \sin(x+y) = (\cos x + i \sin x)(\cos y + i \sin y).$$

Ausmultiplizieren und Vergleich der Real- und Imaginärteile liefert die Summenformeln

$$\cos(x+y) = \cos(x) \cos(y) - \sin(x) \sin(y), \quad (4.12)$$

und

$$\sin(x+y) = \sin(x) \cos(y) + \sin(y) \cos(x). \quad (4.13)$$

Reihendarstellung von Sinus und Cosinus. In die Formel

$$\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$$

setzen wir die Exponentialreihe ein und erhalten

$$\begin{aligned} \sin x &= \frac{1}{2i}(e^{ix} - e^{-ix}) = \frac{1}{2i} \left(\sum_{k=0}^{\infty} \frac{(ix)^k}{k!} - \sum_{k=0}^{\infty} \frac{(-ix)^k}{k!} \right) = \frac{1}{2i} \sum_{k=0}^{\infty} i^k (1 - (-1)^k) \frac{x^k}{k!} \\ &= \sum_{k \text{ ungerade}} i^{k-1} \frac{x^k}{k!} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}. \end{aligned}$$

Es ergibt sich eine für alle $x \in \mathbb{R}$ konvergente Reihendarstellung des Sinus,

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (4.14)$$

Eine analoge Rechnung für den Cosinus führt von

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$$

auf die für alle $x \in \mathbb{R}$ konvergente Reihe

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad (4.15)$$

Aus den Reihendarstellungen können wir Grenzwerte erhalten. So ist

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \dots = 1 - x \cdot \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+3)!}.$$

Die am Ende stehende Reihe wird durch die Reihe für $e^{|x|}$ majorisiert, also folgt

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1,$$

was wir schon wissen. Für den Cosinus erhalten wir aus (4.15)

$$\frac{\cos x - 1}{x} = -\frac{x}{2!} + \frac{x^3}{4!} - \frac{x^5}{6!} + \cdots = x \cdot \sum_{k=0}^{\infty} (-1)^{k+1} \frac{x^{2k}}{(2k+2)!}.$$

Auch hier wird die am Ende stehende Reihe durch die Reihe für $e^{|x|}$ majorisiert, und wir erhalten den Grenzwert

$$\lim_{x \rightarrow 0} \frac{\cos x - 1}{x} = 0. \quad (4.16)$$

Bei den beiden vorangehenden Grenzwertbetrachtungen wurde ein Argument verwendet, welches oft hilfreich ist bei der Bestimmung von Grenzwerten. Wir präsentieren es in allgemeiner Form.

Lemma 4.2 Seien $f, g : D \rightarrow \mathbb{R}$ mit $D \subset \mathbb{R}^d$, sei $a \in D$. Wir nehmen an, dass g beschränkt ist auf D und dass

$$\lim_{x \rightarrow a} f(x) = 0.$$

Dann gilt auch

$$\lim_{x \rightarrow a} f(x)g(x) = 0. \quad (4.17)$$

Ist nämlich $|g(x)| \leq C$ für alle $x \in D$ und (x^n) eine beliebige Folge in D mit $x^n \rightarrow a$, so gilt $|f(x^n)g(x^n)| \leq C|f(x^n)| \rightarrow 0$ für $n \rightarrow \infty$.

Bei der Anwendung von Lemma 4.2 nimmt man oft für D nicht den maximalen gemeinsamen Definitionsbereich von f und g , sondern nimmt einen verkleinerten Definitionsbereich, auf dem g beschränkt ist (etwa ein Intervall bzw. eine Kugel mit a als Mittelpunkt).

Polarkoordinaten in der komplexen Ebene. Für jede komplexe Zahl $z \neq 0$ gilt

$$z = |z| \cdot \frac{z}{|z|}.$$

Da

$$\left| \frac{z}{|z|} \right| = \frac{|z|}{|z|} = 1,$$

gilt $z/|z| = e^{i\varphi}$, wobei φ der Winkel ist, den $z/|z|$ mit der reellen Achse bildet. Wir erhalten die Darstellung von z in Polarkoordinaten

$$z = re^{i\varphi}, \quad r = |z|. \quad (4.18)$$

Die Multiplikation in \mathbb{C} wird in Polarkoordinaten anschaulich: Sind

$$z_1 = r_1 e^{i\varphi_1}, \quad z_2 = r_2 e^{i\varphi_2},$$

so gilt

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

Bei der Multiplikation werden die Längen multipliziert und die Winkel addiert.

5 Folgen und Stetigkeit II

Der Zwischenwertsatz. Wir betrachten eine Funktion $f : [a, b] \rightarrow \mathbb{R}$. Sie hat Werte $f(a)$ und $f(b)$ in den Randpunkten, es gelte $f(a) < f(b)$. Es stellt sich die Frage, ob Werte y zwischen $f(a)$ und $f(b)$ ebenfalls als Funktionswerte $y = f(x)$ vorkommen, das heißt, ob die Gleichung

$$f(x) = y \tag{5.1}$$

eine Lösung $x \in [a, b]$ hat. Falls f unstetig ist, braucht das nicht der Fall zu sein, beispielsweise wenn f nur die Werte $f(a)$ und $f(b)$ annimmt. Ist f aber stetig auf $[a, b]$, so besagt der Zwischenwertsatz, dass jeder Wert zwischen $f(a)$ und $f(b)$ auch tatsächlich angenommen wird.

Satz 5.1 (Zwischenwertsatz) *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, sei y eine reelle Zahl zwischen $f(a)$ und $f(b)$, also entweder $f(a) < y < f(b)$ oder $f(a) > y > f(b)$. Dann gibt es ein $x \in (a, b)$ mit $f(x) = y$.*

Wir betrachten den Fall $f(a) < y < f(b)$. Ein $x \in [a, b]$ mit $f(x) = y$ erhalten wir mit

$$M = \{t : t \in [a, b], f(t) \leq y\}, \quad x = \sup(M),$$

und zwar aus folgenden Gründen: Ist (t_n) eine Folge in M mit $t_n \leq x$ und $t_n \rightarrow x$, so folgt $f(t_n) \leq y$ und $f(t_n) \rightarrow f(x)$ mit Stetigkeit, also $f(x) \leq y$. Es folgt weiter $f(x) < f(b)$, also $x < b$. Da $f(x + 1/n) > y$ für hinreichend großes n , gilt $f(x) \geq y$ mit $n \rightarrow \infty$ wegen Stetigkeit. Also ist $f(x) = y$.

Häufungspunkte. Die Folge der Primzahlen

$$2, 3, 5, 7, 11, \dots$$

ist eine **Teilfolge** der Folge $1, 2, 3, 4, 5, \dots$ der natürlichen Zahlen. Sie entsteht, indem wir 1 sowie alle die Zahlen weglassen, die mehr als zwei Teiler haben, oder andersherum gesagt, indem wir die Zahlen der Reihe nach auszeichnen, die genau zwei Teiler haben.

Allgemein kann man eine Teilfolge einer gegebenen Folge $(x_n)_{n \in \mathbb{N}}$ dadurch festlegen, dass man die Indizes der zur Teilfolge gehörenden Glieder angibt,

$$n_1, n_2, n_3, \dots \quad n_k \in \mathbb{N},$$

mit $n_k < n_{k+1}$ für alle $k \in \mathbb{N}$. Die so festgelegte Teilfolge bezeichnet man mit

$$(x_{n_k})_{k \in \mathbb{N}}.$$

Die Folge

$$x_n = (-1)^n$$

ist uns bereits mehrfach begegnet. Sie ist nicht konvergent. Betrachtet man aber die Teilfolge (x_{n_k}) zu den geradzahligem Indizes $n_k = 2k$, so entsteht die konvergente Folge $1, 1, 1, \dots$. Aus der divergenten Folge

$$1, 2, \frac{1}{2}, 3, \frac{1}{3}, 4, \frac{1}{4}, \dots$$

erhalten wir eine konvergente Teilfolge, indem wir alle Zahlen streichen, die größer als 1 sind; deren Grenzwert ist 0. Ein solcher Grenzwert heißt Häufungspunkt der ursprünglichen Folge.

Definition 5.2 (Häufungspunkt)

Ein $x_* \in \mathbb{R}$ heißt Häufungspunkt einer Folge $(x_n)_{n \in \mathbb{N}}$ in \mathbb{R} , falls es eine Teilfolge von (x_n) gibt, welche gegen x_* konvergiert.

Ersetzen wir \mathbb{R} durch \mathbb{R}^d in 5.2, so erhalten wir die Definition des Häufungspunkts einer Folge im Mehrdimensionalen, und damit als Spezialfall ($d = 2$) auch im Komplexen.

Wir betrachten als Beispiel im \mathbb{R}^2 die Folge, die entsteht, wenn wir die vier Ecken des Quadrats $[0, 1] \times [0, 1]$ immer wieder durchlaufen, also

$$(0, 0), (0, 1), (1, 1), (1, 0), (0, 0), (0, 1), \dots$$

Alle vier Ecken sind Häufungspunkte dieser Folge.

Nicht jede Folge hat eine konvergente Teilfolge, beispielsweise nicht die Folge der natürlichen Zahlen. (Alle Teilfolgen von \mathbb{N} sind unbeschränkt, also divergent.)

Satz 5.3 (Bolzano-Weierstraß)

Jede beschränkte Folge in \mathbb{R} hat eine konvergente Teilfolge und damit auch einen Häufungspunkt.

Sei $(x_n)_{n \in \mathbb{N}}$ beschränkt, also in einem Intervall $I_1 = [a_1, b_1]$ enthalten, wir setzen $n_1 = 1$. Wir teilen I_1 in zwei Hälften und wählen als $I_2 = [a_2, b_2]$ diejenige Hälfte, welche unendlich viele Punkte der Restfolge $(x_n)_{n > n_1}$ enthält. Wir wählen $n_2 > n_1$ mit $x_{n_2} \in I_2$. Auf dieselbe Weise konstruieren wir $x_{n_{k+1}}$ und I_{k+1} per Induktion zu x_{n_k} und I_k . Es gilt für alle $k \in \mathbb{N}$

$$a_k \leq x_{n_k} \leq b_k, \quad b_k - a_k = 2^{-k}(b_0 - a_0),$$

die Folge $(a_k)_{k \in \mathbb{N}}$ ist monoton wachsend, die Folge $(b_k)_{k \in \mathbb{N}}$ ist monoton fallend.

Wir setzen

$$a_* = \sup_{k \in \mathbb{N}} a_k, \quad b_* = \inf_{k \in \mathbb{N}} b_k.$$

Da $a_k \leq a_* \leq b_* \leq b_k$ und $b_k - a_k \rightarrow 0$, folgt mit dem Einschließungssatz sowohl $a_* = b_*$ als auch $x_{n_k} \rightarrow a_*$.

Eine Folge (x^n) im \mathbb{R}^d heißt **beschränkt**, falls es eine Zahl M gibt mit $\|x^n\| \leq M$ für alle n . Es wird also verlangt, dass es eine d -dimensionale Kugel um den Nullpunkt gibt (M ist ihr Radius), welche alle Punkte der Folge enthält, oder anders gesagt, dass die aus den Längen $\|x^n\|$ gebildete reelle Folge beschränkt ist.

Folgerung 5.4 Jede beschränkte Folge im \mathbb{R}^d hat eine konvergente Teilfolge und damit auch einen Häufungspunkt.

Sei (x^n) beschränkte Folge im \mathbb{R}^d . Die aus den ersten Komponenten (x_1^n) gebildete Folge ist im \mathbb{R} beschränkt, hat also nach Satz 5.3 eine konvergente Teilfolge $(x_1^{n_j})$ mit Laufindex j . Die mit denselben Indizes aus den zweiten Komponenten gebildete Folge $(x_2^{n_j})$ ist ebenfalls beschränkt und hat daher ebenfalls eine konvergente Teilfolge $(x_2^{n_{j_k}})$ mit Laufindex k . Die entsprechende Teilfolge der ersten Komponenten $(x_1^{n_{j_k}})$ ist konvergent, da bereits $(x_1^{n_j})$ konvergent war. In derselben Weise gehen wir auch in den weiteren Komponenten zu konvergenten Teilfolgen über und erhalten schließlich eine Teilfolge der ursprünglichen Vektorfolge (x^n) , die in jeder Komponente konvergiert und damit auch im Sinne der Konvergenz im \mathbb{R}^d konvergiert.

Existenz von Maxima und Minima. Sei $f : D \rightarrow \mathbb{R}$ mit $D \subset \mathbb{R}$ gegeben. Ein $x \in D$ heißt **Maximum** (bzw. **Minimum**) von f auf D , falls $f(x) \geq f(t)$ (bzw. $f(x) \leq f(t)$) gilt für alle $t \in D$.

Nicht jede Funktion hat ein Maximum oder ein Minimum. So hat $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x$, weder Maximum noch Minimum auf \mathbb{R} ; f ist sowohl nach oben als auch nach unten unbeschränkt. Die durch

$$f(x) = \frac{x}{x+1}, \quad \text{falls } x \geq 0,$$

und $f(x) = -f(-x)$ für $x < 0$ definierte Funktion ist nach oben durch 1 und nach unten durch -1 beschränkt, hat aber auf \mathbb{R} ebenfalls weder Maximum noch Minimum. Auch die durch $f(0) = f(1) = 0$ und $f(x) = x - 1/2$ für $0 < x < 1$ definierte Funktion $f : [0, 1] \rightarrow \mathbb{R}$ hat weder Maximum noch Minimum.

Satz 5.5 *Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ hat ein Maximum und ein Minimum auf $[a, b]$.*

Zunächst sehen wir, dass die Bildmenge $f([a, b])$ nach oben beschränkt ist. Andernfalls gäbe es eine Folge (x_n) in $[a, b]$ mit $f(x_n) \rightarrow +\infty$. Diese Folge hätte aber nach Satz 5.3 eine gegen ein $x \in [a, b]$ konvergente Teilfolge x_{n_k} , was $f(x_{n_k}) \rightarrow f(x)$ zur Folge hätte, ein Widerspruch. Sei nun $s = \sup f([a, b])$, sei (t_n) Folge in $[a, b]$ mit $f(t_n) \rightarrow s$. Ist (t_{n_j}) eine konvergente Teilfolge gemäß Satz 5.3, so setzen wir $x_* = \lim_{j \rightarrow \infty} t_{n_j}$. Da f stetig ist, gilt $f(t_{n_j}) \rightarrow f(x_*)$, also $s = f(x_*)$, und x_* ist das gesuchte Maximum. Analog erhält man die Existenz des Minimums.

Für die Gültigkeit von Satz 5.5 sind drei Dinge wesentlich: f ist stetig, $I = [a, b]$ ist beschränkt, die Randpunkte a und b von I gehören zu I . Ein analoger Satz gilt auch im Mehrdimensionalen. Wir betrachten $f : D \rightarrow \mathbb{R}$ mit $D \subset \mathbb{R}^d$. Die Menge D heißt **beschränkt**, falls es ein $M > 0$ gibt mit

$$\|x\| \leq M, \quad \text{für alle } x \in D.$$

Wir müssen noch überlegen, was es heißt, dass “der Rand zu D dazugehört”.

Eine Menge $D \subset \mathbb{R}^d$ heißt **abgeschlossen in \mathbb{R}^d** , falls gilt:

Ist (x^n) Folge in D , welche gegen ein $x \in \mathbb{R}^d$ konvergiert, so muss $x \in D$ gelten.

Anders gesagt: Die Grenzwerte konvergenter Folgen in D liegen ebenfalls in D , oder: D ist abgeschlossen hinsichtlich Grenzwertbildung.

Die Einheitskreisscheibe $D = \{x : \|x\| \leq 1\}$ ist abgeschlossen im \mathbb{R}^d : Ist (x^n) eine konvergente Folge mit $\|x^n\| \leq 1$ für alle n , so gilt $\|x\| \leq 1$ für den Grenzwert, der somit auch in D liegt. Lassen wir aber irgendeinen "Randpunkt" x_* der Kreisscheibe (also einen Punkt x_* mit $\|x_*\| = 1$) weg, so konvergiert die Folge

$$x_n = \left(1 - \frac{1}{n}\right)x_*$$

gegen x_* , und es gilt $\|x_n\| = 1 - 1/n < 1$; die Menge $D \setminus \{x_*\}$ ist also nicht abgeschlossen, da sie zwar die Folge (x^n) , aber nicht deren Grenzwert x_* enthält.

Eine Teilmenge D von \mathbb{R}^d heißt **kompakt**, falls jede Folge in D eine konvergente Teilfolge hat, deren Grenzwert ebenfalls in D liegt.

Satz 5.6 *Eine Teilmenge D von \mathbb{R}^d ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Ist D unbeschränkt, so gibt es eine Folge (x^n) in D mit $\|x^n\| \geq n$; jede Teilfolge einer solchen Folge ist unbeschränkt und kann daher nicht konvergieren. Ist D nicht abgeschlossen, so gibt es eine Folge (x^n) in D , deren Grenzwert x nicht in D liegt; da jede Teilfolge ebenfalls gegen x konvergiert und Grenzwerte eindeutig sind, kann D nicht kompakt sein. Ist andererseits D beschränkt, so hat jede Folge (x^n) in D eine konvergente Teilfolge nach Folgerung 5.4; deren Limes muss in D liegen, wenn D außerdem abgeschlossen ist.

Wir können die Argumentation im Anschluss an Satz 5.5 unmittelbar aufs Mehrdimensionale übertragen und erhalten damit das folgende Ergebnis.

Satz 5.7 *Jede auf einer beschränkten und abgeschlossenen Teilmenge D von \mathbb{R}^d stetige Funktion $f : D \rightarrow \mathbb{R}$ hat auf D ein Maximum und ein Minimum.*

Umkehrfunktionen. Wir kennen bereits den allgemeinen Begriff einer bijektiven Abbildung. Ist $f : M \rightarrow N$ eine Abbildung von einer Menge M in eine Menge N , so heißt f bijektiv, falls es zu jedem $y \in N$ genau ein $x \in M$ gibt mit $f(x) = y$. Die so erhaltene Zuordnung $y \mapsto x$ definiert die Umkehrabbildung $f^{-1} : N \rightarrow M$.

In der Analysis sind M und N Teilmengen von \mathbb{R} , oder allgemeiner von \mathbb{R}^d . So ist etwa die Funktion

$$f(x) = x^2, \quad f : [0, \infty) \rightarrow [0, \infty),$$

auf den genannten Mengen (nicht aber auf \mathbb{R}) bijektiv, die Umkehrfunktion ist

$$f^{-1}(y) = \sqrt{y}, \quad f^{-1} : [0, \infty) \rightarrow [0, \infty).$$

Genausogut kann man schreiben

$$f^{-1}(x) = \sqrt{x},$$

da man für die Argumente einer Funktion beliebige Symbole verwenden kann.

Ist f formelmäßig in einer Form $y = f(x)$ gegeben, so kann man versuchen, die Umkehrfunktion formelmäßig zu bestimmen, indem man nach x auflöst, also etwa

$$y = x^2, \quad \text{jetzt Wurzelziehen,} \quad \sqrt{y} = x.$$

In der Regel geht das von Hand nicht, bzw. wenn überhaupt, dann mit einem Computeralgebra-Paket. Andererseits kann man den Graph der Umkehrfunktion erhalten, indem man den Graphen der ursprünglichen Funktion an der Winkelhalbierenden spiegelt.

Löst man sich von dem Anspruch, eine formelmäßige Darstellung der Umkehrfunktion zu erhalten, so kann man aus dem folgenden Satz erfahren, dass es in vielen Fällen Umkehrfunktionen gibt.

Satz 5.8 *Sei I Intervall im \mathbb{R} , sei $f : I \rightarrow \mathbb{R}$ stetig und streng monoton wachsend. Dann ist $f : I \rightarrow f(I)$ bijektiv und die Umkehrfunktion $f^{-1} : f(I) \rightarrow I$ ebenfalls stetig und streng monoton wachsend.*

Den Beweis lassen wir weg.

Als erstes Beispiel betrachten wir die reelle Exponentialfunktion. Wir wissen bereits, dass

$$\exp : \mathbb{R} \rightarrow (0, \infty)$$

stetig und streng monoton wachsend ist. Gemäß Satz 5.8 hat die Exponentialfunktion eine stetige und streng monoton wachsende Umkehrfunktion, sie heißt der **Logarithmus**, genauer der **natürliche** Logarithmus, wir schreiben sie als

$$\ln : (0, \infty) \rightarrow \mathbb{R}. \tag{5.2}$$

Die Komposition von Funktion und Umkehrfunktion ergibt, wie immer bei Umkehrabbildungen, in beide Richtungen die Identität,

$$\ln(\exp(x)) = x, \quad \exp(\ln(x)) = x, \tag{5.3}$$

jeweils auf den Definitionsgebieten \mathbb{R} bzw. $(0, \infty)$.

Die Rechenregeln für den Logarithmus kann man aus den Rechenregeln für die Exponentialfunktion erhalten. Wegen $e^0 = 1$ und $e^1 = e$ gilt

$$\ln(1) = 0, \quad \ln(e) = 1. \tag{5.4}$$

Aus

$$\exp(\ln x + \ln y) = \exp(\ln x) \cdot \exp(\ln y) = xy$$

folgt, indem wir auf beide Seiten den Logarithmus draufsetzen,

$$\ln x + \ln y = \ln(xy). \tag{5.5}$$

Wir können also zwei Zahlen multiplizieren, indem wir deren Logarithmen addieren und anschließend die Exponentialfunktion anwenden. Das war im 17. Jahrhundert einer der

Gründe für die Erfindung und rasche Verbreitung der Logarithmusfunktion, in Form von Tabellen.

Analog folgt

$$\ln(x^k) = k \ln x, \quad \text{für } x > 0 \text{ und } k \in \mathbb{Z}. \quad (5.6)$$

Auf der Basis von \ln und \exp kann man die **allgemeine Potenzfunktion** definieren. Wir setzen

$$x^a = \exp(a \ln x), \quad \text{für } x > 0 \text{ und } a \in \mathbb{R}. \quad (5.7)$$

Auch für sie gelten die üblichen Rechenregeln für Potenzen. So ist etwa

$$x^{a+b} = \exp((a+b) \ln x) = \exp(a \ln x) \cdot \exp(b \ln x) = x^a \cdot x^b.$$

Die übliche Schreibweise für Wurzeln ordnet sich hier ein als

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

Der Ausdruck x^a suggeriert, dass der Exponent a fest ist, während x variiert, wie etwa bei x^2 und \sqrt{x} . Anders ist es bei

$$e^x = \exp(x),$$

hier variiert der Exponent, aber die **Basis** e ist fest. Wählen wir statt e eine andere Zahl $b > 1$ als Basis, so erhalten wir

$$b^x = \exp(x \ln b). \quad (5.8)$$

Die so definierte Funktion $f(x) = b^x$ ist wegen $\ln b > 0$ streng monoton wachsend und stetig auf \mathbb{R} , ihre Umkehrfunktion heißt **Logarithmus zur Basis b** , geschrieben als

$$\log_b(b^x) = x. \quad (5.9)$$

Die Rechenregeln (5.5) und (5.6) gelten analog. Solange man Logarithmen wie ursprünglich bezweckt zum Rechnen mit Papier und Bleistift verwendete, war $b = 10$ die gebräuchliche Basis. Für Situationen, bei denen die Dualdarstellung einer Zahl eine Rolle spielt, ist die Wahl $b = 2$ nützlich. Für alle anderen Anwendungen steht wegen der besonderen Rolle der Exponentialfunktion die Basis e mit dem zugehörigen natürlichen Logarithmus \ln im Vordergrund.

Umkehrfunktionen der trigonometrischen Funktionen. Weitere Beispiele für Umkehrfunktionen liefern die trigonometrischen Funktionen. Der Tangens

$$\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \mathbb{R} \quad (5.10)$$

ist stetig und streng monoton wachsend, seine Bildmenge ist \mathbb{R} . Wir erhalten gemäß Satz 5.8 eine stetige und streng monoton wachsende Umkehrfunktion, sie heißt Arcustangens,

$$\arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right). \quad (5.11)$$

Da der Sinus stetig und streng monoton wachsend ist als Funktion

$$\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow \mathbb{R} \quad (5.12)$$

mit Bildmenge $[-1, 1]$, gibt es eine stetige und streng monoton wachsende Umkehrfunktion, den Arcussinus

$$\arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \quad (5.13)$$

Analoges gilt für den Arcuscosinus als Umkehrfunktion des Cosinus,

$$\arccos : [-1, 1] \rightarrow [0, \pi]. \quad (5.14)$$

6 Differenzierbarkeit

Der Begriff der Ableitung. Seien $f : \mathbb{R} \rightarrow \mathbb{R}$ und ein $a \in \mathbb{R}$ gegeben. Wir wollen f in der Nähe von a durch eine Gerade approximieren. Einen naheliegenden Kandidaten für eine gute Approximation stellt die Tangente an den Graphen von f durch den Punkt $(a, f(a))$ dar. Diese erhalten wir geometrisch-anschaulich aus der Sekante an f durch die Punkte $(a, f(a))$ und $(a+h, f(a+h))$, analytisch beschrieben durch die Geradengleichung

$$g_h(x) = f(a) + \frac{f(a+h) - f(a)}{h}(x-a), \quad (6.1)$$

indem wir den Grenzübergang $h \rightarrow 0$ vornehmen. Genauer gesagt: Wir betrachten den sogenannten **Differenzenquotienten**

$$d_a(h) = \frac{f(a+h) - f(a)}{h}, \quad (6.2)$$

welcher die Steigung der Sekante g_h angibt, und wollen den Differenzialquotienten (welcher die Steigung der Tangente angibt) als den Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \quad (6.3)$$

erhalten.

Für die allgemeine Definition betrachten wir $f : D \rightarrow \mathbb{R}$ mit $D \subset \mathbb{R}$, verlangen aber, dass f "in der Nähe von a überall definiert ist". Wir nennen a einen **inneren Punkt** von D , falls D ein Intervall der Form $(a - \varepsilon, a + \varepsilon) \subset D$ für ein festes $\varepsilon > 0$ enthält; es ist gleichgültig, wie klein ε ist. Dadurch erreicht man, dass die Grenzwertbildung in (6.3) beliebige gegen 0 konvergente Folgen (h_n) berücksichtigt.

Definition 6.1 (Ableitung, differenzierbare Funktion)

Sei $D \subset \mathbb{R}$, $f : D \rightarrow \mathbb{R}$, a ein innerer Punkt von D . Wir sagen, dass f differenzierbar in a ist, falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \quad (6.4)$$

existiert. Der Grenzwert heißt dann die Ableitung (oder Steigung) von f in a , geschrieben als

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}. \quad (6.5)$$

f heißt in D differenzierbar, falls f in jedem Punkt $a \in D$ differenzierbar ist.

Für konstante Funktionen $f(x) = c$, $c \in \mathbb{R}$ fest, gilt für alle $a \in \mathbb{R}$

$$\frac{f(a+h) - f(a)}{h} = 0, \quad \text{also} \quad f'(a) = 0.$$

Für affin-lineare Funktionen, also Geraden der Form $f(x) = mx + b$, gilt

$$\frac{f(a+h) - f(a)}{h} = \frac{m(a+h) + b - ma - b}{h} = \frac{mh}{h} = m,$$

also auch $f'(a) = m$, die Gerade hat in jedem Punkt die gleiche Steigung m .

Für die Standardparabel $f(x) = x^2$ gilt in jedem Punkt a

$$\frac{f(a+h) - f(a)}{h} = \frac{(a+h)^2 - a^2}{h} = \frac{2ah + h^2}{h} = 2a + h, \quad \text{also } f'(a) = 2a.$$

Hier ändert sich die Steigung von Punkt zu Punkt.

Wir können die Abbildung $a \mapsto f'(a)$ als Funktion von a auffassen; schreiben wir x statt a für das Argument, so hat $f(x) = x^2$ die Ableitung $f'(x) = 2x$. Man muss begrifflich zwischen "Ableitung in einem Punkt" und "Ableitungsfunktion" unterscheiden; im Sprachgebrauch lässt man aber normalerweise sein und spricht in beiden Fällen von der Ableitung.

Der Graph der Funktion f' liefert für jedes Argument x die Steigung der Tangente an f in x .

Für die Kehrwertbildung $f(x) = 1/x$ gilt an Stellen $x \neq 0$

$$\frac{f(x+h) - f(x)}{h} = \frac{\frac{1}{x+h} - \frac{1}{x}}{h} = -\frac{1}{(x+h)x},$$

also

$$f'(x) = -\frac{1}{x^2}. \quad (6.6)$$

Für den Sinus $f(x) = \sin x$ gilt

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{\sin(x+h) - \sin(x)}{h} = \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \sin(x) \frac{\cos h - 1}{h} + \cos(x) \frac{\sin h}{h}. \end{aligned}$$

Wir wissen bereits, dass

$$\lim_{h \rightarrow 0} \frac{\cos h - 1}{h} = 0, \quad \lim_{h \rightarrow 0} \frac{\sin h}{h} = 1,$$

also folgt

$$(\sin)'(x) = \cos x. \quad (6.7)$$

Analog erhält man aus dem Additionstheorem für den Cosinus, dass gilt

$$(\cos)'(x) = -\sin x. \quad (6.8)$$

Zur Berechnung der Ableitung der Exponentialfunktion gehen wir auf die Reihendarstellung zurück. Es gilt für beliebiges $z \in \mathbb{C}$

$$\exp(z) = 1 + z + \sum_{k=2}^{\infty} \frac{z^k}{k!},$$

also

$$\frac{\exp(z) - 1}{z} - 1 = \sum_{k=2}^{\infty} \frac{z^{k-1}}{k!} = \sum_{k=0}^{\infty} \frac{z^{k+1}}{(k+2)!} = z \cdot \sum_{k=0}^{\infty} \frac{z^k}{(k+2)!}.$$

Da die zuletzt stehende Reihe durch die Exponentialreihe für $e^{|z|}$ majorisiert wird, gilt

$$\lim_{z \rightarrow 0} \frac{\exp(z) - 1}{z} = 1 \quad (6.9)$$

sogar als Grenzwert im Komplexen; wir benötigen nur den Grenzwert im Reellen. Für $f(x) = \exp(x)$ gilt

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{\exp(x+h) - \exp(x)}{h} = \frac{\exp(x)\exp(h) - \exp(x)}{h} \\ &= \exp(x) \frac{\exp(h) - 1}{h}, \end{aligned}$$

also wegen (6.9)

$$(\exp)'(x) = \exp(x). \quad (6.10)$$

Die Betragsfunktion $f(x) = |x|$ ist differenzierbar an allen Stellen $x \neq 0$, ihre Ableitung ist

$$f'(x) = \begin{cases} 1, & x > 0, \\ -1, & x < 0. \end{cases}$$

In $x = 0$ ist die Betragsfunktion nicht differenzierbar. Für Folgen $h_n \rightarrow 0$ gilt nämlich

$$\frac{f(0+h_n) - f(0)}{h_n} = \frac{|h_n| - |0|}{h_n} = \begin{cases} 1, & h_n > 0, \\ -1, & h_n < 0. \end{cases}$$

Je nach Folge erhält man entweder 1 oder -1 als Grenzwert, oder überhaupt keinen Grenzwert. Für solche Situationen hat man, analog zur Situation bei der Stetigkeit, den Begriff der rechtsseitigen und der linksseitigen Ableitung eingeführt. Die rechtsseitige Ableitung ist definiert durch

$$f'_+(x) = \lim_{\substack{h \rightarrow 0 \\ h > 0}} \frac{f(x+h) - f(x)}{h}, \quad (6.11)$$

die linksseitige Ableitung durch

$$f'_-(x) = \lim_{\substack{h \rightarrow 0 \\ h < 0}} \frac{f(x+h) - f(x)}{h}. \quad (6.12)$$

Für die Betragsfunktion $f(x) = |x|$ gilt $f'_+(0) = 1$ und $f'_-(0) = -1$.

Bemerkung zur Notation von Ableitungen. Für die Ableitung gibt es unterschiedliche Notationen. Sie sind teilweise historisch bedingt, teilweise entstehen sie, wenn man nicht klar unterscheidet zwischen einer Funktion f und dem Wert $f(x)$ einer Funktion an einer Stelle x , also etwa “wir betrachten die Funktion $f(x)$ ” oder “wir betrachten die Funktion $f = f(x)$ ”. Wir schreiben in der Regel f' für die Ableitungsfunktion, $f'(x)$ für den Wert der Ableitungsfunktion an der Stelle x , also $f'(3)$ für den Wert von f' an der Stelle 3. Eine andere Notation für $f'(3)$ ist

$$f'(x) \Big|_{x=3}.$$

Hier eine Liste weiterer möglicher Bezeichnungen:

$$Df, Df(x), \frac{df}{dx}, \frac{df}{dx}(x), \frac{df}{dx}(x) \Big|_{x=3}, \dots$$

Die Information, auf die es ankommt, ist: 1. um welche Funktion es geht, 2. dass man ableiten will, 3. an welcher Stelle die Ableitung gebildet werden soll.

Ableitung als Approximation. Der wesentliche Grund, warum Ableitungen eine so große Rolle spielen, liegt weniger in der geometrischen Interpretation der Tangentensteigung als Grenzwert der Sekantensteigungen, sondern hauptsächlich in der (zugegebenermaßen damit zusammenhängenden) Tatsache, dass mit ihrer Hilfe sehr viele allgemeine nichtlineare Situationen durch eine lineare Approximation gut beschrieben werden können, und dass lineare Zusammenhänge bei Problemen aller Größenordnungen algorithmisch vergleichsweise schnell und effizient in den Griff zu kriegen sind.

Ist f in einem Punkt x differenzierbar und $x + h$ ein weiterer Punkt, so ist die Differenz zwischen Funktionswert $f(x + h)$ und Tangentenwert $f(x) + f'(x)h$ gleich

$$r(h) = f(x + h) - f(x) - f'(x)h. \quad (6.13)$$

Beispielsweise gilt für $f(x) = x^2$, also $f'(x) = 2x$, dass $f(1) = 1$ und $f'(1) = 2$ sowie

$$r(h) = (1 + h)^2 - 1 - 2h = h^2.$$

Im Punkt 1.01 unterscheiden sich Funktionswert und Tangentenwert lediglich um

$$r(0.01) = 0.0001,$$

also erst in der vierten Stelle. Dieses Verhalten ist typisch für differenzierbare Funktionen, denn es gilt

$$\frac{r(h)}{h} = \frac{f(x + h) - f(x)}{h} - f'(x) \rightarrow 0 \quad \text{für } h \rightarrow 0 \quad (6.14)$$

gemäß der Definition der Ableitung, das heißt, $r(h)$ geht **schneller** gegen 0 als h . Im Beispiel $f(x) = x^2$ geht $r(h)$ so schnell wie h^2 gegen 0; auch das ist typisch, wie wir später beim Thema Taylorreihen sehen werden.

Aus (6.13) sehen wir weiterhin, dass gilt

$$f(x + h) - f(x) = f'(x)h + r(h) = h \left(f'(x) + \frac{r(h)}{h} \right).$$

Hieraus schließen wir unmittelbar mit Lemma 4.2, da der geklammerte Ausdruck auf der rechten Seite beschränkt ist als Funktion von h , dass

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

Satz 6.2 *Ist f differenzierbar in einem Punkt x , so ist f stetig im Punkt x .*

Die Umkehrung gilt nicht: Die Betragsfunktion $f(x) = |x|$ ist stetig in 0, aber nicht differenzierbar in 0.

Ableitungsregeln. Summen und skalare Vielfache differenzierbarer Funktionen sind differenzierbar, und für solche Funktionen f, g gilt

$$(f + g)'(x) = f'(x) + g'(x), \quad (cf)'(x) = cf'(x). \quad (6.15)$$

Das folgt unmittelbar daraus, dass Summen und skalare Vielfache mit der Grenzwertbildung vertauscht werden können.

Für das Produkt differenzierbarer Funktionen f, g gilt die **Produktregel**

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x). \quad (6.16)$$

Wir betrachten

$$\frac{(fg)(x+h) - (fg)(x)}{h} = f(x+h) \frac{g(x+h) - g(x)}{h} + g(x) \frac{f(x+h) - f(x)}{h}.$$

Durch Grenzübergang $h \rightarrow 0$ folgt nun (6.16), da $\lim_{h \rightarrow 0} f(x+h) = f(x)$ nach Satz 6.2.

Wir wenden die Produktregel an auf das Monom

$$f(x) = x^n, \quad n \in \mathbb{N}.$$

Das Ergebnis ist

$$f'(x) = nx^{n-1}.$$

Für $n = 1$ gilt $f'(x) = 1$. Für den Induktionsschritt $n - 1 \rightarrow n$ wenden wir die Produktregel an auf $x^n = x \cdot x^{n-1}$ und erhalten

$$f'(x) = 1 \cdot x^{n-1} + x \cdot (n-1)x^{n-2} = nx^{n-1}.$$

Mit dem Ergebnis für Monome können wir beliebige Polynome differenzieren, so führt $f(x) = 2x^3 - 4x$ beispielsweise auf $f'(x) = 6x^2 - 4$.

Für den Quotient differenzierbarer Funktionen f, g gilt die **Quotientenregel**

$$\left(\frac{f}{g}\right)'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2} \quad (6.17)$$

an Stellen x mit $g(x) \neq 0$.

Man betrachtet zunächst den Spezialfall $f = 1$ und bestimmt die Ableitung von $1/g$ in x als $-g'(x)/g(x)^2$ durch eine Grenzwertbetrachtung wie bei der Produktregel. Für beliebiges f wendet man nun die Produktregel auf $f/g = f \cdot (1/g)$ an.

Als Beispiel betrachten wir

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

Die Quotientenregel ergibt

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

Wir behandeln als nächstes die Frage, wie die Ableitung einer invertierbaren Funktion f mit der Ableitung ihrer Inversen f^{-1} zusammenhängt. Für eine Gerade

$$f(x) = mx + b$$

können wir die Inverse direkt berechnen mit dem Kalkül

$$y = mx + b, \quad mx = y - b, \quad x = \frac{1}{m}(y - b),$$

also

$$f^{-1}(y) = \frac{1}{m}(y - b).$$

Für die Ableitungen gilt

$$f'(x) = m, \quad (f^{-1})'(y) = \frac{1}{m},$$

unabhängig von den Stellen x und y , da Geraden konstante Steigung haben. Im allgemeinen Fall gilt ebenfalls, dass die Ableitung der Inversen gleich dem Kehrwert der Ableitung der ursprünglichen Funktion ist, man muss die Ableitungen aber an den richtigen Stellen betrachten, nämlich an Punkten x, y mit $y = f(x)$ bzw. $x = f^{-1}(y)$.

Satz 6.3 Seien $[a, b] \subset \mathbb{R}$, $f : [a, b] \rightarrow \mathbb{R}$ stetig und streng monoton, sei f differenzierbar in $x \in [a, b]$ mit $f'(x) \neq 0$. Dann ist die Umkehrfunktion $f^{-1} : f([a, b]) \rightarrow \mathbb{R}$ in $y = f(x)$ differenzierbar, und es gilt

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}. \quad (6.18)$$

Eine zu (6.18) äquivalente Formulierung wäre

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

Die Formel (6.18) erhält man aus der Definition der Ableitung als Grenzwert der Differenzenquotienten; wegen der uns bereits bekannten Stetigkeit der Umkehrfunktion (Satz 5.8) werden aus Folgen $x_n \rightarrow x$ solche mit $y_n \rightarrow y = f(x)$ und umgekehrt, wobei $y_n = f(x_n)$. Die Einzelheiten führen wir hier nicht aus.

Wir können somit die Ableitung des Logarithmus aus der Ableitung $(\exp)'(x) = \exp(x)$ der Exponentialfunktion bestimmen,

$$(\ln)'(y) = \frac{1}{(\exp)'(\ln y)} = \frac{1}{\exp(\ln y)} = \frac{1}{y}. \quad (6.19)$$

Als Anwendung betrachten wir die Frage, wie sich Kapital durch Verzinsung vermehrt. Wir beginnen mit einem Kapital vom Wert c (in irgendeiner Geldeinheit). Der Jahreszinssatz sei $x \in (0, 1)$, $x = 0.04$ beispielsweise entspricht 4 Prozent Zins. Der Verzinsungsschritt am Ende des Jahres entspricht der Multiplikation mit $1 + x$. Nach k Jahren wäre das Kapital also auf den Betrag $c(1 + x)^k$ angewachsen. Wir betrachten als nächstes eine Verzinsung in kürzeren Zeitabschnitten Δt , beispielsweise eine monatliche Verzinsung. Bei festem Jahreszins x und Ausgangskapital c hat das Kapital am Jahresende den Wert

$$c\left(1 + \frac{x}{12}\right)^{12}.$$

Bei täglicher Verzinsung wären es

$$c\left(1 + \frac{x}{365}\right)^{365}.$$

Mit "kontinuierlicher Verzinsung" bezeichnet man die Grenzwertbildung $\Delta t \rightarrow 0$, sie führt am Jahresende auf den Wert

$$c \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Wir berechnen diesen Grenzwert. Es ist

$$\ln \left[\left(1 + \frac{x}{n}\right)^n \right] = n \ln \left(1 + \frac{x}{n}\right) = x \cdot \frac{\ln \left(1 + \frac{x}{n}\right) - \ln 1}{\frac{x}{n}},$$

Der Bruch auf der rechten Seite konvergiert gegen $(\ln)'(1)$ für $n \rightarrow \infty$, also gilt

$$\lim_{n \rightarrow \infty} \ln \left[\left(1 + \frac{x}{n}\right)^n \right] = x \cdot (\ln)'(1) = x \cdot 1 = x.$$

Wir wenden auf beide Seiten die Exponentialfunktion an und erhalten

$$e^x = \exp \left(\lim_{n \rightarrow \infty} \ln \left[\left(1 + \frac{x}{n}\right)^n \right] \right) = \lim_{n \rightarrow \infty} \exp \left(\ln \left[\left(1 + \frac{x}{n}\right)^n \right] \right) = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Wir halten das Ergebnis fest als

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x. \quad (6.20)$$

Das Anfangskapital c hat also nach einem Jahr bei kontinuierlicher Verzinsung zum Jahreszinssatz x den Wert ce^x .

Im Spezialfall $x = 1$ wird (6.20) zu

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n. \quad (6.21)$$

Die Ableitung einer zusammengesetzten Funktion $g \circ f$ ergibt sich aus den Ableitungen der einzelnen Funktionen f und g mit der Kettenregel.

Satz 6.4 (Kettenregel)

Seien $D, E \subset \mathbb{R}$, $f : D \rightarrow \mathbb{R}$, $f(D) \subset E$, $g : E \rightarrow \mathbb{R}$, a innerer Punkt von D und $f(a)$ innerer Punkt von E . Sind f in a und g in $f(a)$ differenzierbar, so ist auch $g \circ f$ in a differenzierbar, und es gilt

$$(g \circ f)'(a) = g'(f(a))f'(a). \quad (6.22)$$

Wir zerlegen den Differenzenquotienten für $g \circ f$ in

$$\frac{g(f(x)) - g(f(a))}{x - a} = \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \cdot \frac{f(x) - f(a)}{x - a}$$

und führen Grenzübergang $x \rightarrow a$ durch. Da f in a stetig ist, gilt mit $x \rightarrow a$ auch $f(x) \rightarrow f(a)$, und es ergibt sich (6.22). Für diejenigen x , für die $f(x) = f(a)$ gilt, ersetzt man den ersten Bruch auf der rechten Seite durch $g'(f(a))$, der Grenzübergang bleibt gültig.

Die Ableitung der Komposition entsteht also als das Produkt der Ableitungen der äußeren Funktion g und der inneren Funktion f , jeweils an den Stellen, an denen die einzelnen Funktionen ausgewertet werden ($f(a)$ bzw. a).

Als Beispiel betrachten wir

$$f(x) = \cos(\ln x).$$

Hier ist die Notation so, wie die Anwendung es nahelegt: f steht für die abzuleitende, also für die zusammengesetzte Funktion, die Ableitung soll im Punkt x gebildet werden. Nach Kettenregel gilt

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

Weitere Beispiele liefert die Potenzfunktion. Sei

$$f(x) = x^\alpha = e^{\alpha \ln x}, \quad \text{wobei } \alpha > 0.$$

Da $(\exp)' = \exp$, gilt

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

Sei andersherum

$$f(x) = \alpha^x = e^{x \ln \alpha}, \quad \text{wobei } \alpha > 0.$$

Es gilt

$$f'(x) = e^{x \ln \alpha} \cdot \ln \alpha = \alpha^x \ln \alpha.$$

Potenzreihen und deren Ableitung. Eine Reihe der Form

$$\sum_{k=0}^{\infty} a_k z^k, \quad z \in \mathbb{C}, \quad (6.23)$$

heißt **Potenzreihe** in \mathbb{C} mit Koeffizienten $a_k \in \mathbb{C}$. Wir kennen als Beispiele bereits die geometrische Reihe und die Exponentialreihe,

$$\sum_{k=0}^{\infty} z^k, \quad \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Die geometrische Reihe ist für $|z| < 1$ konvergent, die Exponentialreihe für alle $z \in \mathbb{C}$. Zur Konvergenz von Potenzreihen gilt:

Satz 6.5 *Für jede Potenzreihe der Form (6.23) gibt es ein $R \geq 0$, so dass die Potenzreihe konvergiert für alle $z \in \mathbb{C}$ mit $|z| < R$ und divergiert für alle $z \in \mathbb{C}$ mit $|z| > R$. Die Zahl R heißt **Konvergenzradius** der Potenzreihe, der Kreis $\{z : |z| = R\}$ heißt **Konvergenzkreis** der Potenzreihe. Die Fälle $R = 0$ und $R = +\infty$ sind möglich.*

Auf dem Konvergenzkreis ist – abhängig von z – sowohl Konvergenz als auch Divergenz möglich.

Den Beweis lassen wir weg; er beruht darauf, dass die Potenzreihe innerhalb ihres Konvergenzkreises durch eine konvergente geometrische Reihe majorisiert wird und im Äußeren $\{z : |z| > R\}$ eine divergente geometrische Reihe als Minorante hat.

Jede Potenzreihe definiert also im Innern ihres Konvergenzkreises eine Funktion

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \quad |z| < R. \quad (6.24)$$

Der Konvergenzradius R hängt ab von der Potenzreihe, das heißt, von der Folge (a_k) ihrer Koeffizienten. Ist (a_k) konvergent, so gilt die Formel von Hadamard,

$$R = \frac{1}{\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|}}. \quad (6.25)$$

Für divergente Folgen (a_k) bleibt (6.25) gültig, wenn wir den Limes durch den sogenannten Limes superior ersetzen (den behandeln wir hier nicht).

Für die Ableitung von Polynomen gilt

$$p(x) = \sum_{k=0}^n a_k x^k, \quad p'(x) = \sum_{k=1}^n k a_k x^{k-1}.$$

Betrachten wir nun eine reelle Potenzreihe

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

so erhalten wir durch gliedweises Differenzieren

$$\sum_{k=1}^{\infty} k a_k x^{k-1}. \quad (6.26)$$

Satz 6.6 (Ableitung einer Potenzreihe)

Die zu einer Potenzreihe gehörende Funktion

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \quad (6.27)$$

ist im Innern ihres Konvergenzkreises $\{x : |x| < R\}$ definiert und differenzierbar, ihre Ableitung ist gleich

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}. \quad (6.28)$$

Der Konvergenzradius für die Ableitung ist ebenfalls gleich R .

Auch diesen Satz beweist man durch geeignete Majoranten bzw. Minoranten in Form einer geometrischen Reihe.

Als Beispiele kann die Ableitungen der Exponentialfunktion, des Sinus und des Cosinus betrachten; durch Differenzieren der Potenzreihe erhalten wir ebenfalls die uns schon bekannten Formeln.

Ein weiteres Beispiel liefert die geometrische Reihe,

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

Differenzieren auf beiden Seiten ergibt

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}, \quad \text{falls } |x| < 1.$$

7 Sätze der Differenzialrechnung, Extremwerte

Wir betrachten eine Funktion $f : [a, b] \rightarrow \mathbb{R}$, welche in einem Punkt $x \in (a, b)$, also $a < x < b$, ein Maximum hat, also

$$f(x) = \max_{t \in [a, b]} f(t). \quad (7.1)$$

Für hinreichend kleines h gilt $x + h \in [a, b]$, also $f(x + h) \leq f(x)$, und weiter

$$\begin{aligned} \frac{f(x + h) - f(x)}{h} &\leq 0, \quad \text{falls } h > 0, \\ \frac{f(x + h) - f(x)}{h} &\geq 0, \quad \text{falls } h < 0. \end{aligned} \quad (7.2)$$

Falls f im Punkt x differenzierbar ist, ist in beiden Fällen ein Grenzübergang $h \rightarrow 0$ möglich und ergibt

$$f'(x) \leq 0 \quad \text{sowie} \quad f'(x) \geq 0$$

und damit $f'(x) = 0$. Ein analoges Argument liefert, dass in einem Minimum

$$f(x) = \min_{t \in [a, b]} f(t). \quad (7.3)$$

ebenfalls $f'(x) = 0$ gelten muss, falls $a < x < b$. Wir fassen zusammen.

Satz 7.1 Sei $f : [a, b] \rightarrow \mathbb{R}$, differenzierbar in (a, b) . Ist $x \in (a, b)$ ein Maximum oder ein Minimum von f auf $[a, b]$, so gilt $f'(x) = 0$.

Als Beispiel betrachten wir $f(x) = x^2$. Diese Funktion hat in 0 ein sogenanntes **globales Minimum**, das heißt,

$$f(0) = \min_{t \in \mathbb{R}} f(t),$$

welches folglich ein Minimum auch in jedem Intervall $[a, b]$ mit $a < 0 < b$ ist, und es gilt $f'(x) = 2x$, also $f'(0) = 0$.

Als weiteres Beispiel betrachten wir die Funktion $f(x) = x^3$. Da f stetig ist, nimmt f auf jedem Intervall $[a, b]$ das Minimum und das Maximum an. Da f streng monoton wachsend ist, sind Minimum und Maximum eindeutig bestimmt, und es gilt

$$f(a) = \min_{t \in [a, b]} f(t), \quad f(b) = \max_{t \in [a, b]} f(t),$$

das heißt, Minimum und Maximum liegen am Rand. Im Nullpunkt gilt zwar $f'(0) = 0$, aber im Intervall $[-1, 1]$ beispielsweise ist 0 weder Minimum noch Maximum von f . In allen anderen Punkten gilt $f'(x) = 3x^2 \neq 0$. Wir lernen daraus zweierlei: Satz 7.1 bezieht sich nur auf **Extrema** (= Minima oder Maxima) im **Innern** des Definitionsgebiets, und seine Umkehrung gilt nicht.

Wir behandeln nun einige Sätze, die sich aus Satz 7.1 ergeben.

Satz 7.2 (Satz von Rolle)

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar in (a, b) . Ist $f(a) = f(b)$, so gibt es ein $x \in (a, b)$ mit $f'(x) = 0$.

Falls f konstant ist, gilt ohnehin $f' = 0$. Ist f nicht konstant, so muss entweder das Maximum oder das Minimum von f (beide existieren, da f stetig) im Innern (a, b) von $[a, b]$ liegen, wir nennen es x . Nach Satz 7.1 folgt $f'(x) = 0$.

Satz 7.3 (Mittelwertsatz) Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar in (a, b) . Dann gibt es ein $x \in (a, b)$ mit

$$\frac{f(b) - f(a)}{b - a} = f'(x). \quad (7.4)$$

Dieser Satz bedeutet geometrisch, dass die Sekantensteigung auf der linken Seite von (7.4) auch als Tangentensteigung an einem geeigneten Zwischenpunkt $x \in (a, b)$ vorkommen muss.

Wir erhalten den Mittelwertsatz als Spezialfall eines weiter unten (Satz 7.6) behandelten Satzes.

Das folgende Resultat ermöglicht es, Differenzen von Funktionswerten von f abzuschätzen, solange wir untere und obere Schranken für die Ableitung f kennen, ohne dass wir sonst etwas über die Form von f wissen müssen.

Folgerung 7.4 Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar in (a, b) , sei

$$m \leq \inf_{t \in (a, b)} f'(t), \quad M \geq \sup_{t \in (a, b)} f'(t). \quad (7.5)$$

Dann gilt

$$m(y - x) \leq f(y) - f(x) \leq M(y - x) \quad (7.6)$$

für alle $x, y \in [a, b]$ mit $x \leq y$.

Wir erhalten die Aussage, indem wir Satz 7.3 an auf das Teilintervall $[x, y]$ von $[a, b]$ anwenden.

Folgerung 7.5 Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar in (a, b) , sei $f'(x) = 0$ für alle $x \in (a, b)$. Dann ist f konstant.

Das folgt sofort aus (7.6) mit $m = M = 0$.

Wir kommen zurück auf den Mittelwertsatz.

Satz 7.6 (Verallgemeinerter Mittelwertsatz)

Seien $f, g : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar in (a, b) , sei $g'(x) \neq 0$ für alle $x \in (a, b)$. Dann ist $g(a) \neq g(b)$, und es gibt ein $x \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x)}{g'(x)}. \quad (7.7)$$

Der Mittelwertsatz 7.3 ergibt sich aus Satz 7.6, indem wir $g(x) = x$ setzen.

Aus $g(a) = g(b)$ folgt $g'(t) = 0$ für mindestens ein $t \in (a, b)$ nach dem Satz von Rolle, im Widerspruch zur Voraussetzung. Es ist also $g(a) \neq g(b)$. Wir definieren $F : [a, b] \rightarrow \mathbb{R}$ durch

$$F(t) = f(t) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(t) - g(a)).$$

Es ist $F(a) = f(a) = F(b)$. Nach dem Satz von Rolle gibt es ein $x \in (a, b)$ mit

$$0 = F'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)}g'(x),$$

woraus die Behauptung folgt.

Der folgende Satz beruht auf dem verallgemeinerten Mittelwertsatz. Er behandelt die Frage, wie man Grenzwerte berechnen kann, bei denen Einsetzen des Arguments zur Situation "0/0" führt, wie etwa im Beispiel (dessen Ergebnis wir schon kennen)

$$\lim_{x \rightarrow 0} \frac{\sin x}{x}.$$

Satz 7.7 (Satz von de l'Hospital) Seien f, g stetig auf $[a, b]$ und differenzierbar in (a, b) , sei $x \in [a, b]$ mit $f(x) = g(x) = 0$, sei $g'(\xi) \neq 0$ für alle $\xi \in (a, b)$ mit $\xi \neq x$. Dann gilt: Falls der Grenzwert

$$\lim_{\xi \rightarrow x} \frac{f'(\xi)}{g'(\xi)} \tag{7.8}$$

existiert, so existiert auch der Grenzwert

$$\lim_{\xi \rightarrow x} \frac{f(\xi)}{g(\xi)}, \tag{7.9}$$

und

$$\lim_{\xi \rightarrow x} \frac{f'(\xi)}{g'(\xi)} = \lim_{\xi \rightarrow x} \frac{f(\xi)}{g(\xi)}. \tag{7.10}$$

Ist (x_n) eine Folge mit $x_n \rightarrow x$, so gibt es nach Satz 7.6 Zwischenstellen ξ_n zwischen x_n und x mit

$$\frac{f(x_n)}{g(x_n)} = \frac{f(x_n) - f(x)}{g(x_n) - g(x)} = \frac{f'(\xi_n)}{g'(\xi_n)}.$$

Grenzübergang $n \rightarrow \infty$ führt auf die Behauptung.

Als Beispiel erhalten wir unmittelbar

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

Komplizierter ist das folgende Beispiel. Wir wollen berechnen

$$\lim_{x \rightarrow 0} \left(\frac{1}{\sin x} - \frac{1}{x} \right). \quad (7.11)$$

Es ist

$$\frac{1}{\sin x} - \frac{1}{x} = \frac{x - \sin x}{x \sin x},$$

wir setzen

$$f(x) = x - \sin x, \quad g(x) = x \sin x,$$

und beachten, dass $f(0) = g(0) = 0$. Es gilt

$$\frac{f'(x)}{g'(x)} = \frac{1 - \cos x}{\sin x + x \cos x}, \quad f'(0) = g'(0) = 0,$$

und weiter (im Vorgriff auf die zweite Ableitung, siehe unten)

$$\frac{f''(x)}{g''(x)} = \frac{\sin x}{2 \cos x - x \sin x}, \quad f''(0) = 0, \quad g''(0) = 2.$$

Wir können Satz 7.7 zweimal anwenden und erhalten

$$0 = \lim_{x \rightarrow 0} \frac{f''(x)}{g''(x)} = \lim_{x \rightarrow 0} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow 0} \frac{f(x)}{g(x)}.$$

Ableitungen und Monotonie. Wenn wir wissen wollen, ob eine Funktion monoton oder streng monoton ist, können wir das unter Umständen aus der Ableitung erkennen.

Satz 7.8 Sei f stetig auf $[a, b]$ und differenzierbar in (a, b) . Dann gilt

$$f'(x) \geq 0 \text{ für alle } x \in (a, b) \quad \Leftrightarrow \quad f \text{ ist monoton wachsend.}$$

$$f'(x) > 0 \text{ für alle } x \in (a, b) \quad \Rightarrow \quad f \text{ ist streng monoton wachsend.}$$

$$f'(x) \leq 0 \text{ für alle } x \in (a, b) \quad \Leftrightarrow \quad f \text{ ist monoton fallend.}$$

$$f'(x) < 0 \text{ für alle } x \in (a, b) \quad \Rightarrow \quad f \text{ ist streng monoton fallend.}$$

Alle Behauptungen folgen aus dem Mittelwertsatz, etwa die erste: Ist $f'(x) \geq 0$ für alle x und ist $s < t$, so gibt es nach dem Mittelwertsatz ein $x \in (s, t)$ mit

$$f(s) - f(t) = f'(x)(s - t) \leq 0.$$

Gibt es andererseits ein $x \in (a, b)$ mit $f'(x) < 0$, so ist

$$\frac{f(x+h) - f(x)}{h} < 0$$

für hinreichend kleines $h > 0$ nach Definition der Ableitung und damit $f(x+h) < f(x)$, also ist f nicht monoton wachsend.

Als Anwendung kommen wir nochmals auf die Umkehrfunktionen von trigonometrischen Funktionen zu sprechen. Da

$$\sin\left(-\frac{\pi}{2}\right) = -1, \quad \sin\left(\frac{\pi}{2}\right) = 1, \quad \sin'(x) = \cos(x) > 0 \quad \text{für alle } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

ist $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow [-1, 1]$ stetig und streng monoton wachsend. Die Umkehrfunktion

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

ist also differenzierbar, und ihre Ableitung errechnet sich gemäß der allgemeinen Formel im vorigen Kapitel als

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

Analog kann man die Ableitung des Arcus-Cosinus berechnen. Für den Tangens gilt

$$(\tan)'(x) = \frac{1}{\cos^2 x} > 0, \quad \text{für alle } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

also ist $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \mathbb{R}$ umkehrbar und die Umkehrfunktion Arcus-Tangens

$$\arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

hat wegen

$$\frac{1}{\cos^2 y} = \frac{\sin^2 y + \cos^2 y}{\cos^2 y} = \tan^2 y + 1$$

die Ableitung

$$(\arctan)'(x) = \cos^2(\arctan x) = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}.$$

Höhere Ableitungen. Ist die Ableitungsfunktion f' einer differenzierbaren Funktion ebenfalls differenzierbar, so können wir die Ableitung f'' von f' bilden, sie heißt die zweite Ableitung von f . Analog können wir höhere Ableitungen bilden, f''' , f'''' usw., üblicherweise zählt man dann nicht Striche ab, sondern schreibt $f^{(3)}$ und $f^{(4)}$ oder so ähnlich. Allgemein spricht man von der n -ten Ableitung, geschrieben $f^{(n)}$.

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

- n -mal differenzierbar in einem Punkt $x \in (a, b)$, falls alle Ableitungen f' , f'' , \dots , $f^{(n)}$ im Punkt x existieren,
- n -mal differenzierbar in (a, b) , falls alle Ableitungen bis zur n -ten Ordnung in jedem Punkt $x \in (a, b)$ existieren,
- n -mal stetig differenzierbar in (a, b) , falls alle Ableitungen bis zur n -ten Ordnung außerdem stetig sind in (a, b) ,
- beliebig oft (oder unendlich oft) differenzierbar in (a, b) , falls die Ableitungen für jede Ordnung $n \in \mathbb{N}$ in (a, b) existieren.

Polynome sind beliebig oft differenzierbar, für $f(x) = 2x^3 - 6x$ beispielsweise gilt

$$f'(x) = 6x^2 - 6, \quad f''(x) = 12x, \quad f'''(x) = 12, \quad f^{(4)}(x) = 0.$$

Da sich beim Ableiten von Polynomen der Grad jeweils um 1 erniedrigt, erhalten wir nach $(n + 1)$ -maligem Ableiten eines Polynoms vom Grad n die Nullfunktion.

Für $f(x) = e^x$ gilt $f'(x) = e^x = f(x)$, also auch $f^{(n)}(x) = e^x$ für jedes $n \in \mathbb{N}$.

Für $f(x) = \sin x$ gilt

$$f'(x) = \cos x, \quad f''(x) = -\sin x, \quad f^{(3)}(x) = -\cos x, \quad f^{(4)}(x) = \sin x,$$

also auch $f^{(5)} = f'$ usw.

Für die Extremwertbestimmung ist die zweite Ableitung von Bedeutung. Sei etwa $f : [a, b] \rightarrow \mathbb{R}$ differenzierbar. Ist $x \in (a, b)$ mit $f'(x) = 0$, so ist x ein Kandidat für ein lokales Extremum. Ist außerdem $f''(x) > 0$ und f'' stetig im Punkt x , so ist $f''(t) > 0$ für alle t in einem hinreichend kleinen Intervall $(x - \varepsilon, x + \varepsilon)$ von x und daher f' streng monoton wachsend in diesem Intervall. Wegen $f'(x) = 0$ muss also $f'(t) > 0$ gelten in $(x, x + \varepsilon)$, sowie $f'(t) < 0$ in $(x - \varepsilon, x)$. Daher ist f streng monoton wachsend in $(x, x + \varepsilon)$ und streng monoton fallend in $(x - \varepsilon, x)$, und damit x ein lokales Minimum von f .

In umgekehrter Richtung kann man mit dem Mittelwertsatz argumentieren und erhält, dass $f''(x) \geq 0$ gelten muss, falls x ein lokales Minimum ist.

Analoges (mit umgekehrten Ungleichungen) gilt für ein lokales Maximum.

Satz 7.9 Sei $f : (a, b) \rightarrow \mathbb{R}$ zweimal stetig differenzierbar, sei $x \in (a, b)$. Dann gelten die Aussagen

$$\begin{array}{lll} f'(x) = 0, \quad f''(x) > 0 & \Rightarrow & x \text{ ist lokales Minimum} \\ x \text{ ist lokales Minimum} & \Rightarrow & f'(x) = 0, \quad f''(x) \geq 0 \\ f'(x) = 0, \quad f''(x) < 0 & \Rightarrow & x \text{ ist lokales Maximum} \\ x \text{ ist lokales Maximum} & \Rightarrow & f'(x) = 0, \quad f''(x) \leq 0 \end{array}$$

Falls $f''(x) = 0$ gilt für einen Kandidaten mit $f'(x) = 0$, so können alle Fälle auftreten. Für die Beispielfunktionen

$$f(x) = x^3, \quad f(x) = x^4, \quad f(x) = -x^4,$$

gilt in jedem Fall $f'(0) = f''(0) = 0$, aber im ersten Fall ist 0 weder ein lokales Minimum noch ein lokales Maximum, im zweiten Fall ist 0 ein lokales Maximum, im dritten Fall ist 0 ein lokales Minimum.

8 Das Integral

Definition des Integrals. Wir wollen das Integral

$$\int_a^b f(x) dx \quad (8.1)$$

einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ definieren. Wir orientieren uns dabei an der Vorstellung, dass der Wert des Integrals gerade den Inhalt der Fläche zwischen der x -Achse und dem Graphen der Funktion liefern soll, falls $f \geq 0$ gilt.

Ist f eine Konstante, etwa $f(x) = c$ für alle x , so definieren wir

$$\int_a^b f(x) dx = c(b - a). \quad (8.2)$$

Ist $c > 0$, so ist das gerade der Flächeninhalt des Rechtecks zwischen der x -Achse und dem Graphen von f ; ist $c < 0$, so liegt dieses Rechteck unterhalb der x -Achse, und (8.2) liefert den Flächeninhalt mit negativem Vorzeichen.

Wir betrachten nun den Fall, dass f stückweise konstant ist. Sei das Intervall $[a, b]$ zerlegt in der Form

$$Z = \{x_0, \dots, x_n\}, \quad \text{wobei} \quad a = x_0 < x_1 < \dots < x_n = b,$$

und sei $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion mit

$$f(x) = c_i, \quad \text{falls } x \in (x_{i-1}, x_i),$$

wobei $c_i \in \mathbb{R}$ gegeben sind. (Welchen Wert f in den Teilpunkten x_i annimmt, ist uns gleichgültig.) Eine solche Funktion f heißt **Treppenfunktion zur Zerlegung Z** . Wir definieren

$$\int_a^b f(x) dx = \sum_{i=1}^n c_i(x_i - x_{i-1}). \quad (8.3)$$

Beispiel:

$$f : [0, 6] \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 1, & 0 \leq x < 2, \\ 3, & 2 \leq x < 3, \\ 2, & 3 \leq x \leq 6. \end{cases}$$

Hier können wir $a = x_0 = 0$, $x_1 = 2$, $x_2 = 3$, $x_3 = b = 6$ setzen, es ist

$$\int_0^6 f(x) dx = 1 \cdot (2 - 0) + 3 \cdot (3 - 2) + 2 \cdot (6 - 3) = 11.$$

Falls $f \geq 0$ gilt, kann die Formel (8.3) so interpretiert werden, dass die Gesamtfläche sich als Summe der einzelnen Rechteckflächen mit der Grundseite $x_i - x_{i-1}$ und der Höhe c_i ergibt.

Wir betrachten nun eine beliebige beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$. Wir fixieren eine Zerlegung Z . Seien $\varphi, \psi : [a, b] \rightarrow \mathbb{R}$ Treppenfunktionen zur Zerlegung Z mit

$$\varphi(x) \leq f(x) \leq \psi(x), \quad \text{für alle } x \in [a, b],$$

so folgt aus (8.3), angewandt auf φ und ψ , dass

$$\int_a^b \varphi(x) dx \leq \int_a^b \psi(x) dx, \quad (8.4)$$

Das Integral von f soll einen Wert zwischen diesen beiden Zahlen haben. Wir setzen nun

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x), \quad M_i = \sup_{x \in [x_{i-1}, x_i]} f(x), \quad (8.5)$$

und definieren die **Untersumme** und **Obersumme** von f zur Zerlegung Z durch

$$\begin{aligned} U_Z(f) &= \int_a^b \varphi(x) dx, \quad \text{wobei } \varphi(x) = m_i \text{ für } x \in (x_{i-1}, x_i), \\ O_Z(f) &= \int_a^b \psi(x) dx, \quad \text{wobei } \psi(x) = M_i \text{ für } x \in (x_{i-1}, x_i). \end{aligned} \quad (8.6)$$

Nach (8.4) gilt

$$U_Z(f) \leq O_Z(f). \quad (8.7)$$

Das Integral von f soll einen Wert zwischen diesen beiden Zahlen haben.

Ist nun \tilde{Z} eine **feinere** Zerlegung als Z , das heißt eine solche, die aus Z durch Hinzufügen weiterer Teilpunkte entsteht, so gilt

$$U_Z(f) \leq U_{\tilde{Z}}(f) \leq O_{\tilde{Z}}(f) \leq O_Z(f). \quad (8.8)$$

Man erhält weiter, dass

$$U_{Z_1}(f) \leq O_{Z_2}(f) \quad (8.9)$$

gilt für zwei Zerlegungen Z_1 und Z_2 , indem man eine dritte Zerlegung \tilde{Z} betrachtet, die sowohl für Z_1 als auch für Z_2 eine Verfeinerung ist, und (8.8) anwendet. Wir definieren nun das **Unterintegral** und das **Oberintegral** von f durch

$$\begin{aligned} U(f) &= \sup\{U_Z(f) : Z \text{ ist Zerlegung von } [a, b],\} \\ O(f) &= \inf\{O_Z(f) : Z \text{ ist Zerlegung von } [a, b].\} \end{aligned} \quad (8.10)$$

Aus (8.9) und der Definition von Supremum und Infimum folgt, dass

$$U(f) \leq O(f). \quad (8.11)$$

Da das Integral von f für jede Zerlegung Z mindestens so groß wie $U_Z(f)$ und höchstens so groß wie $O_Z(f)$ sein soll, soll es auch zwischen den beiden Zahlen $U(f)$ und $O(f)$ liegen. Ist nun aber

$$U(f) = O(f), \quad (8.12)$$

so gibt es dafür nur eine Möglichkeit, nämlich

$$\int_a^b f(x) dx = U(f) = O(f). \quad (8.13)$$

Definition 8.1 (Integral)

Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Falls (8.12) gilt, also Unterintegral und Oberintegral von f übereinstimmen, so heißt f **integrierbar**, und das Integral von f ist definiert durch (8.13). \square

Dieser Integralbegriff wird normalerweise als **Riemann-Integral** bezeichnet.

Satz 8.2 Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist integrierbar.

Der Beweis ist nicht schwierig, er verwendet aber die Eigenschaft der sogenannten gleichmäßigen Stetigkeit, die wir nicht behandelt haben.

Nicht jede beschränkte Funktion ist integrierbar. Die Funktion $f : [0, 1] \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases}$$

hat das Unterintegral 0 und das Oberintegral 1, (8.13) gilt also nicht.

In der Definition 8.1 des Integrals kommt der Begriff des Grenzwerts nicht explizit vor, er verbirgt sich vielmehr hinter dem Übergang von Untersummen zum Unterintegral (bzw. von Obersummen zum Oberintegral) durch Bildung des Supremums bzw. Infimums.

Die Definition des Integrals wird einfacher, wenn wir das Integral direkt auf einen Grenzwert zurückführen können. Dazu betrachten wir für $n \in \mathbb{N}$ die Zerlegung Z_n definiert durch

$$x_k = a + kh, \quad h = 2^{-n}(b - a), \quad 0 \leq k \leq 2^n. \quad (8.14)$$

Wir setzen

$$u_n = U_{Z_n}(f), \quad o_n = O_{Z_n}(f),$$

dann gilt, falls f integrierbar ist,

$$u_n \leq \int_a^b f(x) dx \leq o_n,$$

Falls nun $\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} o_n$ gilt, folgt

$$\lim_{n \rightarrow \infty} u_n = \int_a^b f(x) dx = \lim_{n \rightarrow \infty} o_n. \quad (8.15)$$

Für alle stetigen Funktion ist das richtig.

Eigenschaften des Integrals. Wir benennen einige Eigenschaften des Integrals, deren Gültigkeit sich aus den Definitionen herleiten lässt, wir tun das nicht.

Das Integral ist linear in folgendem Sinn. Sind $f, g : [a, b] \rightarrow \mathbb{R}$ integrierbar und sind $\lambda, \mu \in \mathbb{R}$, so ist auch $\lambda f + \mu g$ integrierbar, und

$$\int_a^b \lambda f(x) + \mu g(x) dx = \lambda \int_a^b f(x) dx + \mu \int_a^b g(x) dx. \quad (8.16)$$

Das Integral ist monoton in folgendem Sinn. Gilt $f(x) \leq g(x)$ für alle $x \in [a, b]$, so gilt auch

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx. \quad (8.17)$$

Ist $f : [a, b] \rightarrow \mathbb{R}$ integrierbar, so ist auch der Positivteil von f ,

$$f^+(x) = \max\{f(x), 0\}, \quad (8.18)$$

integrierbar. Definieren wir den Negativteil von f durch

$$f^-(x) = (-f)^+(x) = -\min\{f(x), 0\},$$

so ist mit f auch f^- integrierbar, und wegen

$$|f| = f^+ + f^-, \quad f^+ \geq 0, \quad f^- \geq 0,$$

ist weiterhin $|f|$ integrierbar, und wegen $f \leq |f|$, $-f \leq |f|$ gilt

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx. \quad (8.19)$$

Hieraus und mit (8.17) erhalten wir sofort die Abschätzung

$$\left| \int_a^b f(x) dx \right| \leq M(b-a), \quad \text{falls } |f(x)| \leq M \text{ für alle } x \in [a, b]. \quad (8.20)$$

Ist $f : [a, b] \rightarrow \mathbb{R}$ integrierbar, und ändern wir den Funktionswert von f an einer einzigen Stelle $t \in [a, b]$, betrachten wir also

$$\tilde{f}(x) = \begin{cases} f(x), & x \in [a, b], x \neq t, \\ y, & x = t, \end{cases} \quad (8.21)$$

wobei $y \in \mathbb{R}$ eine beliebige Zahl ist, so ist auch \tilde{f} integrierbar, und

$$\int_a^b \tilde{f}(x) dx = \int_a^b f(x) dx. \quad (8.22)$$

Indem wir (8.22) wiederholt einsetzen, ergibt sich, dass das Integral einer Funktion sich nicht ändert, wenn wir diese an endlich vielen Stellen abändern.

Satz 8.3

Sei $a < b < c$, sei $f : [a, b] \rightarrow \mathbb{R}$. Dann ist f auf $[a, c]$ integrierbar genau dann, wenn f auf $[a, b]$ und auf $[b, c]$ integrierbar ist, und in diesem Fall gilt

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx. \quad (8.23)$$

□

Formel (8.23) können wir sowohl zur Zerlegung eines Integrals in zwei Integrale als auch zur Zusammenfügung zweier Integrale in ein Integral verwenden. Zusammen mit Satz 8.2 ergibt sich beispielsweise, dass stückweise stetige Funktionen integrierbar sind.

Um sich Fallunterscheidungen zu ersparen, definiert man

$$\int_a^b f(x) dx = - \int_b^a f(x) dx, \quad \text{falls } a \geq b. \quad (8.24)$$

Dann gilt die Formel

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx \quad (8.25)$$

für alle $a, b, c \in \mathbb{R}$ (nicht nur im Fall $a < b < c$), und es folgt

$$\int_a^a f(x) dx = 0. \quad (8.26)$$

Es stellt sich nun die Frage nach der **Berechnung** von Integralen. Das kann auf verschiedene Weise geschehen. Wir beschäftigen uns hier mit der Berechnung aufgrund von Formeln, die darauf beruhen, dass **Integration als Umkehrung der Differentiation** aufgefasst werden kann. Dieser Zusammenhang wird durch den Hauptsatz der Differential- und Integralrechnung hergestellt.

Der Hauptsatz der Differential- und Integralrechnung. Als Vorbereitung behandeln wir den **Mittelwertsatz der Integralrechnung**.

Satz 8.4

Seien $f : [a, b] \rightarrow \mathbb{R}$ stetig, $p : [a, b] \rightarrow \mathbb{R}$ integrierbar mit $p \geq 0$. Dann gibt es ein $\xi \in [a, b]$ mit

$$\int_a^b f(x)p(x) dx = f(\xi) \int_a^b p(x) dx. \quad (8.27)$$

Mit

$$m = \min_{x \in [a, b]} f(x), \quad M = \max_{x \in [a, b]} f(x),$$

gilt wegen $p \geq 0$

$$m \int_a^b p(x) dx \leq \int_a^b f(x)p(x) dx \leq M \int_a^b p(x) dx.$$

Wir wählen ein $y \in [m, M]$ mit

$$y \int_a^b p(x) dx = \int_a^b f(x)p(x) dx.$$

Da $m \leq y \leq M$, können wir nach dem Zwischenwertsatz ein $\xi \in [a, b]$ finden mit $f(\xi) = y$, dann folgt (8.27).

Folgerung 8.5 Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, so gibt es ein $\xi \in [a, b]$ mit

$$\int_a^b f(x) dx = (b - a)f(\xi). \quad (8.28)$$

Wir wenden Satz 8.4 an mit $p = 1$.

Satz 8.6 Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Wir definieren $F : [a, b] \rightarrow \mathbb{R}$ durch

$$F(x) = \int_a^x f(t) dt. \quad (8.29)$$

Dann ist F differenzierbar in jedem $x \in (a, b)$, und

$$F'(x) = f(x). \quad (8.30)$$

Sei $h \in \mathbb{R}$ mit $h > 0$ und $x + h \in [a, b]$. Dann ist

$$\begin{aligned} \frac{F(x+h) - F(x)}{h} &= \frac{1}{h} \left(\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) = \frac{1}{h} \int_x^{x+h} f(t) dt \\ &= \frac{1}{h} h f(\xi) = f(\xi), \end{aligned}$$

wobei ξ zwischen x und $x + h$ liegt, gemäß Folgerung 8.5. Mit $h \rightarrow 0$ gilt $\xi \rightarrow x$, und da f stetig ist, folgt

$$\frac{F(x+h) - F(x)}{h} \rightarrow f(x)$$

für den einseitigen Grenzwert $h \rightarrow 0$, $h > 0$. Analog schließt man für $h < 0$. Also ist F differenzierbar in x , und (8.30) gilt.

Definition 8.7 (Stammfunktion)

Seien $f, F : (a, b) \rightarrow \mathbb{R}$, sei F stetig auf $[a, b]$ und differenzierbar auf (a, b) . Gilt $F'(x) = f(x)$ für alle $x \in (a, b)$, so heißt F eine **Stammfunktion** von f auf (a, b) . \square

Stammfunktionen sind bis auf eine Konstante eindeutig bestimmt. Ist F eine Stammfunktion von f , so ist auch $F + c$, c Konstante, eine Stammfunktion, da

$$(F + c)' = F' + c' = F' = f$$

gilt. Sind umgekehrt F und G Stammfunktionen, so gilt

$$(F - G)' = F' - G' = f - f = 0,$$

also muss $F - G$ eine konstante Funktion sein, wie wir als Konsequenz des Mittelwertsatzes der Differentialrechnung festgestellt haben. In diesem Zusammenhang bezeichnet man die Konstante c auch als **Integrationskonstante**.

Satz 8.8 Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, sei $F : [a, b] \rightarrow \mathbb{R}$ Stammfunktion von f . Dann gilt

$$\int_a^b f(x) dx = F(b) - F(a). \quad (8.31)$$

Nach Satz 8.6 wird durch

$$F_a(x) = \int_a^x f(t) dt$$

eine Stammfunktion von f definiert, und es gilt

$$\int_a^b f(t) dt = F_a(b) = F_a(b) - F_a(a).$$

Da zwei Stammfunktionen sich nur um eine Konstante unterscheiden, muss gelten

$$F(b) - F_a(b) = F(a) - F_a(a),$$

also gilt (8.31).

Man schreibt auch

$$\int_a^b f(x) dx = F(b) - F(a) = F \Big|_a^b = F(x) \Big|_{x=a}^{x=b}.$$

Aus jeder Formel für die Differentiation erhalten wir eine Formel für die Integration. Beispiel: Es ist

$$(\ln)'(x) = \frac{1}{x}, \quad x > 0, \quad (8.32)$$

also

$$\int_1^2 \frac{1}{x} dx = \ln 2 - \ln 1 = \ln 2.$$

Ebenso erhalten wir, wie wir gleich sehen werden, aus jeder Rechenregel für die Differentiation eine Rechenregel für die Integration.

Bemerkungen zum Sprachgebrauch. Statt “Stammfunktion” sagt man auch “unbestimmtes Integral”, was die Unbestimmtheit der Integrationsgrenzen und der Integrationskonstanten betont. Das Integral

$$\int_a^b f(x) dx$$

wird dann auch als “bestimmtes Integral” bezeichnet. Für eine Stammfunktion von f gibt es die traditionelle Notation

$$\int f(x) dx \quad (8.33)$$

mit dem Integralzeichen, aber ohne untere und obere Grenzen. Beispielsweise schreibt man dann

$$\int \frac{1}{x} dx = \ln x. \quad (8.34)$$

Die mathematische Bedeutung dieser Formel ist: “Die durch $F(x) = \ln x$ definierte Funktion ist eine Stammfunktion der durch $f(x) = 1/x$ definierten Funktion.” Für sich genommen ist (8.34) verwirrend, denn der Buchstabe x wird auf beiden Seiten in völlig unterschiedlicher Bedeutung verwendet. Außerdem wird nicht zwischen Funktion und Funktionswert unterschieden. Aber wenn man erstmal weiß, was (8.34) bedeuten soll, kann man es als kompakte Notation benutzen (kein Text, nur Formel).

Partielle Integration. Die Regel für die partielle Integration entsteht aus der Produktregel. Setzen wir

$$F(x) = f(x)g(x),$$

und sind f, g differenzierbar, so gilt

$$F'(x) = f'(x)g(x) + f(x)g'(x).$$

Aus dem Hauptsatz folgt nun, falls die Ableitungen f' und g' stetig sind,

$$\int_a^b f'(x)g(x) + f(x)g'(x) dx = F(x) \Big|_{x=a}^{x=b}.$$

Wir erhalten also die **Regel der partiellen Integration**

$$\int_a^b f(x)g'(x) dx = f(x)g(x) \Big|_{x=a}^{x=b} - \int_a^b f'(x)g(x) dx. \quad (8.35)$$

Wir betrachten einige Beispiele.

1. Es ist

$$\int_0^2 xe^x dx = xe^x \Big|_{x=0}^{x=2} - \int_0^2 e^x dx = (x-1)e^x \Big|_{x=0}^{x=2} = e^2 - 1,$$

und wir erhalten

$$F(x) = (x-1)e^x$$

als Stammfunktion von $u(x) = xe^x$.

2. Für $0 < a < b$ gilt

$$\int_a^b \ln x dx = \int_a^b 1 \cdot \ln x dx = x \ln x \Big|_{x=a}^{x=b} - \int_a^b 1 dx = (x \ln x - x) \Big|_{x=a}^{x=b},$$

und wir erhalten

$$F(x) = x \ln(x) - x$$

als Stammfunktion des Logarithmus.

3. Mit partieller Integration können wir eine Rekursionsformel finden für

$$I_m = \int_a^b (\sin x)^m dx. \quad (8.36)$$

Es ist

$$I_0 = b - a, \quad I_1 = \cos a - \cos b.$$

Für $m \geq 2$ gilt

$$\begin{aligned}
 I_m &= \int_a^b (\sin x)^m dx = - \int_a^b (\sin x)^{m-1} (\cos)'(x) dx \\
 &= -(\sin x)^{m-1} \cos x \Big|_{x=a}^{x=b} + (m-1) \int_a^b (\sin x)^{m-2} \cos^2 x dx \\
 &= -(\sin x)^{m-1} \cos x \Big|_{x=a}^{x=b} + (m-1) \int_a^b (\sin x)^{m-2} (1 - \sin^2 x) dx \\
 &= -(\sin x)^{m-1} \cos x \Big|_{x=a}^{x=b} + (1-m)I_m + (m-1)I_{m-2}. \tag{8.37}
 \end{aligned}$$

Lösen wir diese Gleichung nach I_m auf, so erhalten wir die Rekursionsformel

$$I_m = \frac{m-1}{m} I_{m-2} - \frac{1}{m} (\sin x)^{m-1} \cos x \Big|_{x=a}^{x=b},$$

mit der wir die Berechnung von I_m sukzessive auf I_0 bzw. I_1 zurückführen können.

Substitution. Sei $g : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar. Auf dem Bildintervall $I = g([a, b])$ sei eine stetige Funktion $f : I \rightarrow \mathbb{R}$ definiert. Ist $F : I \rightarrow \mathbb{R}$ eine Stammfunktion von f auf I , so erhalten wir aus der Kettenregel

$$(F \circ g)'(t) = F'(g(t))g'(t) = f(g(t))g'(t), \quad t \in [a, b].$$

Zweimalige Anwendung des Hauptsatzes liefert nun

$$\int_a^b f(g(t))g'(t) dt = (F \circ g)(b) - (F \circ g)(a) = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(x) dx.$$

Die Formel

$$\int_a^b f(g(t))g'(t) dt = \int_{g(a)}^{g(b)} f(x) dx \tag{8.38}$$

heißt **Substitutionsregel**.

Wir betrachten einige Beispiele.

1.

$$\int_a^b f(t+c) dt = \int_{a+c}^{b+c} f(x) dx,$$

falls $f : [a+c, b+c] \rightarrow \mathbb{R}$ stetig ist, $c \in \mathbb{R}$. (Substitution $g(t) = t+c$, es ist $g'(t) = 1$.)

2.

$$\int_a^b f(ct) dt = \frac{1}{c} \int_{ac}^{bc} f(x) dx,$$

falls $f : [ac, bc] \rightarrow \mathbb{R}$ stetig ist, $c \neq 0$. (Substitution $g(t) = ct$, es ist $g'(t) = c$.)

3. Ist $g : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar mit $g(t) > 0$ für alle $t \in [a, b]$, so liefert die Substitutionsregel mit $f(x) = 1/x$

$$\int_a^b \frac{g'(t)}{g(t)} dt = \int_{g(a)}^{g(b)} \frac{1}{x} dx = \ln(g(t)) \Big|_{t=a}^{t=b}.$$

Für $g(t) = \cos t$ ergibt sich, falls $[a, b] \subset (-\pi/2, \pi/2)$,

$$\int_a^b \tan t \, dt = \int_a^b \frac{\sin t}{\cos t} \, dt = -\ln(\cos t) \Big|_{t=a}^{t=b}.$$

4. Für die Substitution

$$g(t) = 2 \arctan t, \quad g'(t) = \frac{2}{1+t^2},$$

gilt vermittels der Identität

$$\sin x = \frac{2 \tan \frac{x}{2}}{1 + \tan^2 \frac{x}{2}}$$

zunächst, dass

$$\sin(g(t)) = \frac{2t}{1+t^2}.$$

Hieraus ergibt sich zum Beispiel

$$\begin{aligned} \int_{g(a)}^{g(b)} \frac{1}{\sin x} \, dx &= \int_a^b \frac{1}{\sin(g(t))} g'(t) \, dt = \int_a^b \frac{1+t^2}{2t} \frac{2}{1+t^2} \, dt = \int_a^b \frac{1}{t} \, dt \\ &= \ln(b) - \ln(a), \end{aligned}$$

falls etwa $0 < a < b$, und damit für $d = 2 \arctan b$, $c = 2 \arctan a$

$$\int_c^d \frac{1}{\sin x} \, dx = \ln \left(\tan \frac{x}{2} \right) \Big|_{x=c}^{x=d}, \quad 0 < c < d < \pi.$$

Systematische Bestimmung von Stammfunktionen. Jede stetige Funktion hat eine Stammfunktion, aber kann man sie auch formelmäßig hinschreiben? Das Auffinden von Formeln für Stammfunktionen durch geschicktes Anwenden der Integrationsregeln war im 18. und 19. Jahrhundert ein wichtiges Thema in der Mathematik. Dabei stellte es sich heraus, dass nicht jede unter Verwendung “elementarer” Funktionen hinschreibbare Funktion auch eine ebensolche Stammfunktion hat, beispielsweise

$$f(x) = \frac{\sin x}{x}.$$

Für eine geeignet definierte Klasse elementarer Funktionen konnte vor etwa 40 Jahren herausgefunden werden, dass man erstens entscheiden kann, ob es eine elementare Stammfunktion gibt, und diese zweitens auch algorithmisch berechnen kann. In Computeralgebra-Paketen sind solche Algorithmen enthalten. Wir verweisen auf das Buch von Bornemann für weitere Ausführungen.

Uneigentliche Integrale. Unter uneigentliche Integrale versteht man Integrale, die einen endlichen Wert haben, bei denen aber entweder eine Integrationsgrenze im Unendlichen liegt, oder die Funktion an der Integrationsgrenze eine Singularität hat. Wir behandeln zunächst den ersten Fall. Sei $f : [a, \infty) \rightarrow \mathbb{R}$ eine Funktion, die auf jedem endlichen Intervall $[a, M]$ integrierbar ist. Wir definieren

$$\int_a^\infty f(x) \, dx = \lim_{M \rightarrow \infty} \int_a^M f(x) \, dx, \quad (8.39)$$

falls der Grenzwert auf der rechten Seite existiert. Analog wird

$$\int_{-\infty}^a f(x) dx = \lim_{M \rightarrow -\infty} \int_M^a f(x) dx \quad (8.40)$$

definiert. Als Beispiel betrachten wir

$$\int_1^{\infty} \frac{1}{x^\alpha} dx, \quad \alpha > 1.$$

Es ist

$$\int_1^M \frac{1}{x^\alpha} dx = \frac{1}{1-\alpha} x^{1-\alpha} \Big|_{x=1}^{x=M} = \frac{M^{1-\alpha} - 1}{1-\alpha},$$

also

$$\int_1^{\infty} \frac{1}{x^\alpha} dx = \lim_{M \rightarrow \infty} \frac{M^{1-\alpha} - 1}{1-\alpha} = \frac{1}{\alpha - 1}.$$

Für den Grenzfall $\alpha = 1$ gilt

$$\int_1^M \frac{1}{x} dx = \ln(M) \rightarrow \infty, \quad \text{falls } M \rightarrow \infty,$$

also existiert $\int_1^{\infty} \frac{1}{x} dx$ nicht. Ein weiteres Beispiel ist

$$\int_1^{\infty} \frac{\ln x}{x^2} dx. \quad (8.41)$$

Mit partieller Integration erhalten wir

$$\begin{aligned} \int_1^M \frac{\ln x}{x^2} dx &= -\frac{\ln x}{x} \Big|_{x=1}^{x=M} - \int_1^M \frac{1}{x} \cdot \left(-\frac{1}{x}\right) dx \\ &= -\frac{\ln x}{x} \Big|_{x=1}^{x=M} - \frac{1}{x} \Big|_{x=1}^{x=M} = -\frac{\ln M + 1}{M} + 1 \\ &\rightarrow 1, \quad \text{für } M \rightarrow \infty, \end{aligned}$$

also

$$\int_1^{\infty} \frac{\ln x}{x^2} dx = 1. \quad (8.42)$$

Wir betrachten nun den Fall, dass die zu integrierende Funktion an einer Integrationsgrenze eine Singularität hat. Sei eine Funktion $f : (a, b] \rightarrow \mathbb{R}$ gegeben, welche integrierbar ist auf jedem Intervall $[a + \varepsilon, b]$ für $\varepsilon > 0$, $\varepsilon < b - a$. Wir definieren

$$\int_a^b f(x) dx = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \int_{a+\varepsilon}^b f(x) dx, \quad (8.43)$$

falls der Grenzwert auf der rechten Seite existiert. Ein solches Integral heißt ebenfalls uneigentliches Integral.

Als Beispiel betrachten wir

$$\int_0^1 \frac{1}{x^\alpha} dx, \quad \alpha < 1.$$

Es ist

$$\int_{\varepsilon}^1 \frac{1}{x^{\alpha}} dx = \frac{1}{1-\alpha} x^{1-\alpha} \Big|_{x=\varepsilon}^{x=1} = \frac{1-\varepsilon^{1-\alpha}}{1-\alpha},$$

also

$$\int_0^1 \frac{1}{x^{\alpha}} dx = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{1-\varepsilon^{1-\alpha}}{1-\alpha} = \frac{1}{1-\alpha}.$$

Für den Grenzfall $\alpha = 1$ gilt

$$\int_{\varepsilon}^1 \frac{1}{x} dx = -\ln(\varepsilon) \rightarrow \infty, \quad \text{falls } \varepsilon \rightarrow 0,$$

also existiert $\int_0^1 \frac{1}{x} dx$ nicht. Die eben beschriebenen Situationen können an beiden Integrationsgrenzen auftreten. Ist etwa $f : \mathbb{R} \rightarrow \mathbb{R}$ integrierbar für alle Intervalle $[a, b]$, so definieren wir

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^c f(x) dx + \int_c^{\infty} f(x) dx, \quad (8.44)$$

falls für ein $c \in \mathbb{R}$ beide Integrale auf der rechten Seite existieren. (Sie existieren dann für alle $c \in \mathbb{R}$, und die Summe hängt nicht von der Wahl von c ab.)

Beispiel:

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx &= \int_{-\infty}^0 \frac{1}{1+x^2} dx + \int_0^{\infty} \frac{1}{1+x^2} dx \\ &= \lim_{M \rightarrow -\infty} (-\arctan M) + \lim_{M \rightarrow \infty} \arctan M \\ &= \frac{\pi}{2} + \frac{\pi}{2} = \pi. \end{aligned} \quad (8.45)$$

Vorsicht! Die Existenz von

$$\lim_{M \rightarrow \infty} \int_{-M}^M f(x) dx$$

ist nicht hinreichend für die Existenz von $\int_{-\infty}^{\infty} f(x) dx$ im Sinne von (8.44), so ist etwa

$$\int_{-M}^M x dx = 0$$

für alle M , aber

$$\int_0^M x dx = \frac{1}{2} M^2 \rightarrow \infty \quad \text{für } M \rightarrow \infty.$$

Integration von Potenzreihen. Wir betrachten eine durch eine Potenzreihe definierte Funktion (die Koeffizienten heißen jetzt c_k statt a_k)

$$f(x) = \sum_{k=0}^{\infty} c_k x^k. \quad (8.46)$$

Satz 8.9 Sei $R > 0$ der Konvergenzradius der Potenzreihe in (8.46). Dann ist f auf jedem Intervall $[a, b] \subset (-R, R)$ integrierbar, und das Integral ist gegeben durch

$$\int_a^b f(x) dx = \sum_{k=0}^{\infty} \int_a^b c_k x^k dx = \sum_{k=0}^{\infty} \frac{c_k}{k+1} (b^{k+1} - a^{k+1}). \quad (8.47)$$

Weiterhin ist

$$F(x) = \sum_{k=0}^{\infty} \frac{c_k}{k+1} x^{k+1} \quad (8.48)$$

eine Stammfunktion von f ; diese Potenzreihe hat ebenfalls den Konvergenzradius R .

Auch diesen Satz beweisen wir nicht.

Wir erhalten also die Stammfunktion einer Potenzreihe durch **gliedweises Integrieren**.

Aus der Formel

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

können wir für $|x| < 1$ aus der Potenzreihendarstellung

$$\frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k \quad (8.49)$$

die Potenzreihendarstellung

$$\ln(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} \quad (8.50)$$

erhalten. (Beide Seiten in (8.50) sind Stammfunktionen von (8.49), und sie sind gleich, da sie für $x = 0$ beide den Wert 0 haben.)

Grenzübergänge unter dem Integralzeichen. Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge von Funktionen $f_n : [a, b] \rightarrow \mathbb{R}$, sei $f : [a, b] \rightarrow \mathbb{R}$ eine weitere Funktion. Wir sagen, dass die Folge $(f_n)_{n \in \mathbb{N}}$ **punktweise gegen f konvergiert**, falls

$$\lim_{n \rightarrow \infty} f_n(x) = f(x), \quad \text{für alle } x \in [a, b]. \quad (8.51)$$

Ist beispielsweise $f_n : [0, 1] \rightarrow \mathbb{R}$ definiert durch

$$f_n(x) = x + \frac{1}{n},$$

so gilt

$$\lim_{n \rightarrow \infty} f_n(x) = x,$$

also konvergiert (f_n) punktweise gegen die durch $f(x) = x$ definierte Funktion f . Betrachten wir als weiteres Beispiel

$$f_n(x) = x^n, \quad f_n : [0, 1] \rightarrow \mathbb{R}, \quad (8.52)$$

so ergibt sich, dass f_n punktweise konvergiert gegen f ,

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

In diesem Fall ist die Grenzfunktion f in $x = 1$ unstetig, obwohl alle Funktionen f_n stetig sind. Betrachten wir dagegen die Integrale, so gilt

$$\int_0^1 f_n(x) dx = \frac{1}{n+1} x^{n+1} \Big|_{x=0}^{x=1} = \frac{1}{n+1} \rightarrow 0 = \int_0^1 f(x) dx,$$

das heißt, in diesem Beispiel konvergieren die Integrale von f_n gegen das Integral der Grenzfunktion f . Das muss aber nicht so sein, wie das folgende Beispiel zeigt. Wir betrachten $f_n : [0, 1] \rightarrow \mathbb{R}$,

$$f_n(x) = \begin{cases} n^2 x, & 0 \leq x \leq \frac{1}{n}, \\ 2n - n^2 x, & \frac{1}{n} < x \leq \frac{2}{n}, \\ 0, & x > \frac{2}{n}. \end{cases} \quad (8.53)$$

Dann gilt

$$\lim_{n \rightarrow \infty} f_n(x) = 0, \quad \text{für alle } x \in [0, 1],$$

das heißt, f_n konvergiert punktweise gegen $f = 0$, aber

$$\int_0^1 f_n(x) dx = 1,$$

während das Integral der Grenzfunktion $f = 0$ natürlich gleich Null ist. Aus der punktweisen Konvergenz stetiger Funktionen folgt also im allgemeinen weder die Stetigkeit der Grenzfunktion noch die Konvergenz der Integrale.

Ein Konvergenzbegriff, der beides gewährleistet, ist die **gleichmäßige Konvergenz**. Zunächst definieren wir für eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ deren **Supremumsnorm** auf dem Intervall $[a, b]$ durch

$$\|f\|_\infty = \sup_{x \in [a, b]} |f(x)|. \quad (8.54)$$

Die Supremumsnorm hängt vom Definitionsgebiet ab, für

$$f(x) = x^3$$

ist beispielsweise

$$\|f\|_\infty = 1 \quad \text{auf } [0, 1], \quad \|f\|_\infty = 8 \quad \text{auf } [-2, 1].$$

Wir sagen nun, dass die Folge $(f_n)_{n \in \mathbb{N}}$ von Funktionen $f_n : [a, b] \rightarrow \mathbb{R}$ **gleichmäßig gegen f konvergiert**, falls

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0, \quad (8.55)$$

das heißt

$$\lim_{n \rightarrow \infty} \sup_{x \in [a, b]} |f_n(x) - f(x)| = 0. \quad (8.56)$$

Ist beispielsweise

$$f_n(x) = x + \frac{1}{n}, \quad f_n : [0, 1] \rightarrow \mathbb{R},$$

und setzen wir $f(x) = x$, so gilt

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = \lim_{n \rightarrow \infty} \frac{1}{n} = 0,$$

also konvergiert f_n gleichmäßig gegen f .

Falls f_n gleichmäßig gegen f konvergiert, gilt

$$\begin{aligned} \left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| &= \left| \int_a^b f_n(x) - f(x) dx \right| \leq \int_a^b |f_n(x) - f(x)| dx \\ &\leq (b - a) \|f_n - f\|_\infty \\ &\rightarrow 0. \end{aligned}$$

Satz 8.10 *Sind $f_n : [a, b] \rightarrow \mathbb{R}$ stetig, und konvergiert f_n gleichmäßig gegen eine Funktion $f : [a, b] \rightarrow \mathbb{R}$, so ist auch f stetig, und es gilt*

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx. \quad (8.57)$$

Dass die Integrale konvergieren, also (8.57) gilt, zeigt die Rechnung unmittelbar vor dem Satz. Dass die Grenzfunktion f stetig ist, beweisen wir nicht.

9 Taylorentwicklung

Wir kehren zur Differentialrechnung zurück, und zwar zu der Frage, wie gut die Werte $f(x)$ einer Funktion f durch ihre Ableitungen im Punkt a approximiert werden, wenn x in der Nähe von a liegt.

Sei $f : I \rightarrow \mathbb{R}$ eine Funktion, I ein offenes Intervall, und $a \in I$ gegeben. Die Tangente an den Graphen von f im Punkt $(a, f(a))$ ist gegeben durch

$$g(x) = f(a) + f'(a)(x - a). \quad (9.1)$$

Wir betrachten die Differenz (das Restglied)

$$r(x) = f(x) - g(x). \quad (9.2)$$

Wir hatten bereits gesehen, dass

$$\frac{r(x)}{x - a} = \frac{f(x) - f(a)}{x - a} - f'(a) \rightarrow 0, \quad \text{für } x \rightarrow a, \quad (9.3)$$

also dass das Restglied für $x \rightarrow a$ schneller gegen 0 geht als die Funktion $x - a$, eine Eigenschaft, die die Tangente vor allen anderen Geraden durch $(a, f(a))$ auszeichnet.

Mit dem Hauptsatz der Differential- und Integralrechnung lässt sich diese Approximation genauer fassen. Aus ihm und mit partieller Integration erhalten wir

$$\begin{aligned} f(x) - f(a) &= \int_a^x f'(t) dt = -(x - t)f'(t) \Big|_{t=a}^{t=x} + \int_a^x (x - t)f''(t) dt \\ &= (x - a)f'(a) + \int_a^x (x - t)f''(t) dt, \end{aligned} \quad (9.4)$$

also (wir bezeichnen das Restglied jetzt mit R_2)

$$R_2(x) = \int_a^x (x - t)f''(t) dt. \quad (9.5)$$

Aus dem Mittelwertsatz der Integralrechnung (Satz 8.4) folgt, dass für ein geeignetes ξ mit $a < \xi < x$ gilt

$$R_2(x) = f''(\xi) \int_a^x (x - t) dt = \frac{(x - a)^2}{2} f''(\xi). \quad (9.6)$$

Falls die zweite Ableitung von f in der Nähe von a beschränkt ist (was der Fall ist, wenn sie stetig ist), etwa durch die Konstante C_2 , so folgt weiter

$$|R_2(x)| \leq C_2 \frac{(x - a)^2}{2}, \quad (9.7)$$

das heißt, das Restglied R_2 verhält sich in der Nähe von a wie eine quadratische Funktion, die in a die Steigung 0 hat.

Dieses quadratische Verhalten von R_2 lässt sich ebenfalls näher spezifizieren. Es ist

$$\begin{aligned} R_2(x) &= \int_a^x (x-t)f''(t) dt = -\frac{(x-t)^2}{2} f''(t) \Big|_{t=a}^{t=x} + \int_a^x \frac{(x-t)^2}{2} f'''(t) dt \\ &= \frac{(x-a)^2}{2} f''(a) + \int_a^x \frac{(x-t)^2}{2} f'''(t) dt. \end{aligned} \quad (9.8)$$

Wir setzen

$$R_3(x) = \int_a^x \frac{(x-t)^2}{2} f'''(t) dt. \quad (9.9)$$

Setzen wir (9.4) – (9.8) zusammen, so erhalten wir

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2} f''(a) + R_3(x), \quad (9.10)$$

das heißt, f verhält sich wie die quadratische Funktion

$$q(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2} f''(a) \quad (9.11)$$

“bis auf” das Restglied R_3 , für welches wir auf die gleiche Weise wie in (9.6) und (9.7) die Darstellung

$$R_3(x) = \frac{(x-a)^3}{6} f'''(\xi) \quad (9.12)$$

mit einem geeigneten $\xi \in (a, x)$, und die Abschätzung

$$|R_3(x)| \leq C_3 \frac{|x-a|^3}{6} \quad (9.13)$$

erhalten. Dieser Prozess lässt sich fortsetzen, falls f hinreichend oft differenzierbar ist.

Satz 9.1 (Taylorentwicklung)

Sei $I \subset \mathbb{R}$ offenes Intervall, sei f $m+1$ -mal stetig differenzierbar, sei $a \in I$. Dann gilt für alle $x \in I$

$$f(x) = \sum_{k=0}^m \frac{f^{(k)}(a)}{k!} (x-a)^k + R_{m+1}(x) \quad (9.14)$$

mit

$$R_{m+1}(x) = \frac{1}{m!} \int_a^x (x-t)^m f^{(m+1)}(t) dt. \quad (9.15)$$

□

Weiterhin gilt wie oben

$$R_{m+1}(x) = \frac{(x-a)^{m+1}}{(m+1)!} f^{(m+1)}(\xi), \quad (9.16)$$

$$|R_{m+1}(x)| \leq C_{m+1} \frac{|x-a|^{m+1}}{(m+1)!}, \quad (9.17)$$

wobei ξ zwischen a und x liegt bzw. C_{m+1} nicht von x abhängt.

Satz 9.1 bedeutet, dass wir die Funktion f durch ein Polynom vom Grade m approximieren können, nämlich das **Taylorpolynom**

$$T_m(x) = \sum_{k=0}^m \frac{f^{(k)}(a)}{k!} (x-a)^k, \quad (9.18)$$

so dass der Approximationsfehler R_{m+1} in der Nähe von a “von der Ordnung” $m+1$ ist. Um das Taylorpolynom aufstellen zu können, müssen wir die Ableitungen von f im **Entwicklungspunkt** a kennen.

Als Beispiel betrachten wir die Exponentialfunktion im Entwicklungspunkt $a = 0$. Da $(\exp)' = \exp$ und daher $\exp^{(k)} = \exp$ für alle k gilt, ist wegen $\exp(0) = 1$

$$T_m(x) = \sum_{k=0}^m \frac{1}{k!} x^k. \quad (9.19)$$

Wollen wir die Exponentialfunktion durch das Polynom (9.19) für ein $x \in (0, 1)$ approximieren, so können wir den Fehler gemäß (9.16) abschätzen durch

$$|R_{m+1}(x)| \leq \frac{x^{m+1}}{(m+1)!} e \leq 3 \frac{x^{m+1}}{(m+1)!}, \quad (9.20)$$

Für $x = 0.5$ und $m = 3$ ist der Fehler kleiner als 0.008, für $m = 8$ bereits kleiner als $2 \cdot 10^{-7}$. Für $x = 0.01$ ist schon für $m = 3$ der Fehler kleiner als $1.9 \cdot 10^{-9}$. Für größere x wird der Fehler sehr schnell größer, so liefert für $x = 10$ die Abschätzung (9.16) bei $m = 8$ eine Schranke von $4.5 \cdot 10^{11}$. In der Tat liefert die Taylorentwicklung nur in der Nähe des Entwicklungspunkts brauchbare Näherungen.

Wir nehmen nun an, die Funktion f habe eine Potenzreihenentwicklung im Punkt a der Form

$$f(x) = \sum_{k=0}^{\infty} c_k (x-a)^k, \quad c_k \in \mathbb{R}. \quad (9.21)$$

Bisher hatten wir nur den Fall $a = 0$ betrachtet. Den Fall $a \neq 0$ können wir darauf zurückführen: Setzen wir

$$\tilde{f}(x) = f(x+a),$$

so gilt

$$\tilde{f}(x) = \sum_{k=0}^{\infty} c_k x^k.$$

Die für \tilde{f} gültigen Aussagen zum Konvergenzradius, Differenzieren und Integrieren übertragen sich entsprechend.

Gliedweises Differenzieren der rechten Seite und Einsetzen von $x = a$ in (9.21) führt auf

$$f^{(j)}(a) = j! c_j, \quad (9.22)$$

also wird (9.21) zu

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k, \quad (9.23)$$

das heißt, das Taylorpolynom vom Grad m wird aus den ersten $m + 1$ Termen der Potenzreihe von f im Entwicklungspunkt a gebildet. Das trifft beispielsweise für die Exponentialfunktion zu, die Darstellung im Entwicklungspunkt $a = 0$

$$\exp(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad (9.24)$$

ist für alle $x \in \mathbb{R}$ gültig. In anderen Fällen ist die Gültigkeit der Potenzreihenentwicklung auf ein (mehr oder weniger großes) Konvergenzintervall um den Entwicklungspunkt eingeschränkt.

Umgekehrt können wir durch (9.22) die Koeffizienten der Potenzreihe (9.21) bestimmen, wenn wir die Ableitungen von f kennen; eine andere Frage ist aber, ob dann in (9.21) die Gleichheit gilt. Das muss nicht unbedingt so sein, beispielsweise hat die durch

$$f(x) = \exp\left(-\frac{1}{x^2}\right), \quad x \neq 0,$$

und $f(0) = 0$ definierte Funktion in $a = 0$ keine Darstellung der Form (9.21). Es sind nämlich alle Ableitungen von f im Nullpunkt gleich Null, und damit auch alle Taylorpolynome in $a = 0$. Die so gebildete Potenzreihe ist also die der Nullfunktion und liefert damit keine Darstellung von f .

Ist f in a beliebig oft differenzierbar, so heißt die Reihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k \quad (9.25)$$

die **Taylorreihe** von f in a . Wir fassen das in den vergangenen Absätzen Gesagte zusammen:

- Ist f in einem Konvergenzintervall $(a - R, a + R)$ um a als Potenzreihe

$$f(x) = \sum_{k=0}^{\infty} c_k (x - a)^k, \quad c_k \in \mathbb{R}, \quad (9.26)$$

darstellbar, so ist diese gleich der Taylorreihe von f in a , mit $k!c_k = f^{(k)}(a)$.

- Ist f in a beliebig oft differenzierbar, so ist die Taylorreihe (9.25) für $x \in \mathbb{R}$ definiert. Aufgrund ihrer Form ist sie eine Potenzreihe um den Entwicklungspunkt a und hat daher ein wohlbestimmtes Konvergenzintervall. Für $x \neq a$ kann es sein, dass

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

es kann aber auch sein, dass

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

Eine andere Schreibweise der Taylorformel erhalten wir, wenn wir $x = a + h$ setzen,

$$f(a + h) = \sum_{k=0}^m \frac{f^{(k)}(a)}{k!} h^k + \tilde{R}_{m+1}(h) \quad (9.27)$$

mit

$$\tilde{R}_{m+1}(h) = \frac{h^{m+1}}{(m+1)!} f^{(m+1)}(\xi), \quad |\tilde{R}_{m+1}(h)| \leq C_{m+1} \frac{|h|^{m+1}}{(m+1)!}. \quad (9.28)$$

Man sagt, dass eine Funktion φ **von der Ordnung** h^m ist, falls

$$|\varphi(h)| \leq C|h|^m \quad (9.29)$$

gilt mit einer (von h unabhängigen, aber sonst nicht eingeschränkten) Konstante C . Man schreibt

$$\varphi(h) = O(h^m). \quad (\text{Gesprochen: "Groß-O von h hoch m"}). \quad (9.30)$$

In diesem Sinne gilt $\tilde{R}_{m+1}(h) = O(h^{m+1})$. Man verwendet auch die Notation

$$\varphi(h) = o(h^m). \quad (\text{Gesprochen: "Klein-O von h hoch m"}). \quad (9.31)$$

Das bedeutet

$$\lim_{h \rightarrow 0} \frac{\varphi(h)}{h^m} = 0. \quad (9.32)$$

Für die Differenz

$$\tilde{r}(h) = f(a + h) - (f(a) + f'(a)h)$$

zwischen Funktionswert und Wert der Tangente besagt die Definition der Ableitung gerade, dass

$$\tilde{r}(h) = o(h).$$

Durch die Restgliedbetrachtung in (9.7), welche gültig ist, wenn die Ableitung f' in der Nähe von a stetig ist, haben wir mehr erhalten, nämlich

$$\tilde{r}(h) = O(h^2).$$

10 Kurven

Eine Kurve in der Ebene oder im Raum stellen wir uns vor als eine gerade oder gekrümmte Linie, welche zwei verschiedene Punkte verbindet, oder welche geschlossen ist wie etwa der Kreis. Mathematisch beschrieben wird sie durch eine Funktion k , die auf einem Intervall I definiert ist und jedem Punkt $t \in I$ einen Kurvenpunkt $k(t)$ zuordnet.

Definition 10.1 (Kurve)

Sei $I = [a, b]$ ein Intervall. Eine stetige Funktion $k : I \rightarrow \mathbb{R}^d$ heißt **Kurve** im \mathbb{R}^d mit dem **Anfangspunkt** $k(a)$ und dem **Endpunkt** $k(b)$. Ist $k(a) = k(b)$, so heißt die Kurve **geschlossen**. \square

Der Einheitskreis in der Ebene wird beschrieben durch die geschlossene Kurve

$$k : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad k(t) = (\cos t, \sin t). \quad (10.1)$$

Die Kurve, die wir auf dem Papier sehen, entspricht also der Bildmenge $k(I)$. Wir können uns t als Zeitvariable vorstellen, die Funktion k beschreibt dann das "Durchlaufen" der Bildmenge $k(I)$ im Zeitintervall $[a, b]$. Unabhängig von der Vorstellung der Zeit kann man t einfach als einen Parameter auffassen.

Es kann sein, dass derselbe Punkt mehrfach durchlaufen wird, beispielsweise wenn die Kurve sich kreuzt. Ein solcher Kreuzungspunkt heißt **Doppelpunkt** der Kurve. Ist $x \in \mathbb{R}^d$ ein Doppelpunkt einer Kurve $k : I \rightarrow \mathbb{R}^d$, so gibt es also Parameterwerte $s, t \in K$, $s \neq t$, mit $x = k(s) = k(t)$.

Es können auch ganze Abschnitte mehrmals durchlaufen werden, beispielsweise beschreibt

$$k : [0, 4\pi] \rightarrow \mathbb{R}^2, \quad k(t) = (\cos t, \sin t) \quad (10.2)$$

die Kurve, welche den Einheitskreis zweimal durchläuft. Weitere Beispiele sind:

1. Die Verbindungsstrecke zweier Punkte $p, q \in \mathbb{R}^d$

$$k(t) = p + t(q - p), \quad t \in [0, 1]. \quad (10.3)$$

2. Die Neil'sche Parabel im \mathbb{R}^2

$$k(t) = (t^2, t^3), \quad t \geq 0. \quad (10.4)$$

3. Die logarithmische Spirale im \mathbb{R}^2

$$k(t) = (e^{-t} \cos t, e^{-t} \sin t). \quad (10.5)$$

4. Die Schraubenlinie im \mathbb{R}^3

$$k(t) = (r \cos t, r \sin t, ct), \quad (10.6)$$

wobei $r > 0$, $c \in \mathbb{R}$ gegeben sind.

5. Die Zykloide

$$k(t) = (rt - d \sin t, r - d \cos t). \quad (10.7)$$

Man kann sie sich im Fall $d = r$ vorstellen als diejenige Kurve, die ein auf einem Kreis vom Radius r fixierter Punkt beschreibt, wenn dieser Kreis mit konstanter Geschwindigkeit auf der horizontalen Achse entlangrollt.

6. Der Graph einer stetigen Funktion $f : [a, b] \rightarrow \mathbb{R}$, er wird beschrieben durch

$$k(t) = (t, f(t)), \quad t \in [a, b]. \quad (10.8)$$

Eine Kurve $k : I \rightarrow \mathbb{R}^d$ heißt **differenzierbar** in $t \in I$, falls alle Komponentenfunktionen $k_1, \dots, k_d : I \rightarrow \mathbb{R}$ in t differenzierbar sind. Der Vektor

$$k'(t) = (k'_1(t), \dots, k'_d(t)) \quad (10.9)$$

heißt die **Ableitung** von k in t . Im Beispiel der Verbindungsstrecke von p nach q ,

$$k(t) = p + t(q - p), \quad k_i(t) = p_i + t(q_i - p_i),$$

gilt $k'_i(t) = q_i - p_i$, also $k'(t) = q - p$, der Vektor $k'(t)$ ist in jedem Punkt derselbe und zeigt in die Richtung der Strecke. Im allgemeinen Fall zeigt der Vektor $k'(t)$ in die Richtung der Tangente an die Bildmenge $k(I)$ im Punkt $k(t)$, entsprechend der Durchlaufrichtung der Kurve, man nennt $k'(t)$ daher auch den **Tangentialvektor** oder **Tangentenvektor** an die Kurve im Punkt $k(t)$.

Ist k in jedem inneren Punkt von I differenzierbar, so wird durch (10.9) eine Funktion k' definiert, sie heißt ebenfalls die Ableitung von k . Beispielsweise hat der Einheitskreis

$$k(t) = (\cos t, \sin t)$$

die Ableitung

$$k'(t) = (-\sin t, \cos t).$$

Eine Kurve $k : I \rightarrow \mathbb{R}^d$ heißt **C¹-Kurve**, falls alle Komponentenfunktionen k_i auf I stetig und im Innern von I stetig differenzierbar sind. Sie heißt **stückweise C¹-Kurve**, falls das Definitionsgebiet I in Teilintervalle $[a_i, b_i]$ zerfällt, so dass k auf jedem Teilintervall eine C^1 -Kurve ist. Alle oben angeführten Beispiele sind C^1 -Kurven. Die durch

$$k : [0, 3] \rightarrow \mathbb{R}^2, \quad k(t) = \begin{cases} (t, 0), & 0 \leq t \leq 1, \\ (2 - t, t - 1), & 1 \leq t \leq 2, \\ (0, 3 - t), & 2 \leq t \leq 3, \end{cases}$$

definierte Kurve ist eine stückweise C^1 -Kurve, welche nacheinander die Seiten des Dreiecks zu den Eckpunkten $(0, 0)$, $(1, 0)$, $(0, 1)$ durchläuft.

Schnittwinkel. Wir nennen die Kurve **regulär**, wenn

$$k'(t) \neq 0 \quad (10.10)$$

gilt für alle $t \in (a, b)$. Seien k, h zwei reguläre Kurven, die sich in einem Punkt $x = k(t) = h(\tau)$ schneiden, es gelte $k'(t) \neq 0$ und $h'(\tau) \neq 0$. Wir definieren den **Schnittwinkel** von

k und h in x als den Winkel φ , der von den Tangentenvektoren $k'(t)$ und $h'(\tau)$ gebildet wird. Wie wir aus der Linearen Algebra wissen, gilt

$$\cos \varphi = \frac{\langle k'(t), h'(\tau) \rangle}{\|k'(t)\| \|h'(\tau)\|}. \quad (10.11)$$

Länge einer Kurve. Man kann die Länge einer Kurve durch Integration berechnen. Sei $k : I \rightarrow \mathbb{R}^d$ eine Kurve, $I = [a, b]$. Wir betrachten einen Streckenzug k_h , der einzelne Kurvenpunkte $k(t_j)$ nacheinander verbindet, wobei

$$a = t_0 < t_1 < \dots < t_N = b, \quad t_j = a + jh, \quad h = \frac{b-a}{N}. \quad (10.12)$$

Die Länge L_h des gesamten Streckenzugs ist gleich der Summe der Länge der Verbindungsstrecken zwischen $k(t_{j+1})$ und $k(t_j)$,

$$L_h = \sum_{j=0}^{N-1} \|k(t_{j+1}) - k(t_j)\|. \quad (10.13)$$

Wir wollen L_h mit der Zahl

$$\int_a^b \|k'(t)\| dt = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \|k'(t)\| dt \quad (10.14)$$

vergleichen. Für die einzelnen Komponenten der Kurve gilt nach dem Mittelwertsatz

$$k_i(t_{j+1}) - k_i(t_j) = k'_i(s_{ij})(t_{j+1} - t_j), \quad (10.15)$$

wobei $s_{ij} \in [t_j, t_{j+1}]$ geeignet gewählt ist. Es folgt

$$\|k(t_{j+1}) - k(t_j)\| = \|d_j\|(t_{j+1} - t_j), \quad \text{wobei } d_j = (k'_1(s_{1j}), \dots, k'_d(s_{dj})),$$

Es folgt weiter

$$\int_{t_j}^{t_{j+1}} \|k'(t)\| dt - \|k(t_{j+1}) - k(t_j)\| = \int_{t_j}^{t_{j+1}} \varepsilon_h(t) dt, \quad (10.16)$$

wobei

$$\varepsilon_h(t) = \|k'(t)\| - \|d_j\|, \quad t \in [t_j, t_{j+1}]. \quad (10.17)$$

Wenn h klein wird und k' eine stetige Funktion ist, werden auch die Differenzen

$$k'_i(t) - k'_i(s_{ij})$$

klein. Man kann beweisen, dass

$$\lim_{h \rightarrow 0} \int_a^b \varepsilon_h(t) dt = 0. \quad (10.18)$$

Nun ist aber nach (10.13) – (10.16)

$$\int_a^b \|k'(t)\| dt = L_h + \int_a^b \varepsilon_h(t) dt,$$

also folgt wegen (10.18)

$$\int_a^b \|k'(t)\| dt = \lim_{h \rightarrow 0} L_h.$$

Definition 10.2 (Länge einer Kurve)

Sei $k : I \rightarrow \mathbb{R}^d$ eine stückweise C^1 -Kurve. Wir definieren die **Länge** von k durch

$$L(k) = \int_a^b \|k'(t)\| dt. \quad (10.19)$$

Statt Länge sagt man auch **Bogenlänge**. □

Der Umfang eines Kreises ist gerade die Länge im Sinne von Definition 10.2. Hier ist (Mittelpunkt 0, Radius r)

$$k : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad k(t) = (r \cos t, r \sin t),$$

also

$$L(k) = \int_0^{2\pi} \|(-r \sin t, r \cos t)\| dt = \int_0^{2\pi} \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} dt = \int_0^{2\pi} r dt = 2\pi r.$$

Für die Zykloide gilt im Falle $d = r$

$$k(t) = (r(t - \sin t), r(1 - \cos t)),$$

also für einen einzelnen Bogen, $t \in I = [0, 2\pi]$,

$$\begin{aligned} L(k) &= \int_0^{2\pi} \|(r(1 - \cos t), r \sin t)\| dt = \int_0^{2\pi} r \sqrt{(1 - \cos t)^2 + \sin^2 t} dt = r \int_0^{2\pi} 2 \sin \frac{t}{2} dt \\ &= 8r. \end{aligned}$$

Für den Graph einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ gilt

$$k(t) = (t, f(t)), \quad k'(t) = (1, f'(t)),$$

also erhalten wir die Länge des Graphen einer Funktion aus

$$L(k) = \int_a^b \|k'(t)\| dt = \int_a^b \sqrt{1 + f'(t)^2} dt = \int_a^b \sqrt{1 + f'(x)^2} dx, \quad (10.20)$$

wobei wir die Gleichheit der letzten beiden Integrale nur deshalb hingeschrieben haben, um noch einmal darauf aufmerksam zu machen, dass die Wahl des Buchstabens für die Integrationsvariable völlig unerheblich ist.

Verschiedene Parameterdarstellungen. Wie schon bemerkt, enthält die Darstellung einer Kurve durch eine Funktion $k : I \rightarrow \mathbb{R}^d$ nicht nur die Information über die Form (das Bild $k(I)$), sondern auch über die Art des Durchlaufens, so können wir die Länge des Tangentenvektors $\|k'(t)\|$ als Geschwindigkeit interpretieren, falls t die Zeit bedeutet. Man kann die gleiche Kurvenform durch verschiedene Funktionen k darstellen, so liefern etwa

$$\begin{aligned} k &: [0, 1] \rightarrow \mathbb{R}^2, \quad k(t) = (\cos(2\pi t), \sin(2\pi t)), \\ k &: [1, e] \rightarrow \mathbb{R}^2, \quad k(t) = (\cos(2\pi \ln t), \sin(2\pi \ln t)), \end{aligned}$$

ebenfalls Darstellungen des Einheitskreises. Die allgemeine Situation sieht folgendermaßen aus. Sei $h : [c, d] \rightarrow [a, b]$ eine stetige und streng monoton wachsende Funktion mit $h(c) = a$ und $h(d) = b$. Zu einer gegebenen Kurve $k : [a, b] \rightarrow \mathbb{R}^d$ können wir eine neue Kurve $\tilde{k} : [c, d] \rightarrow \mathbb{R}^d$ definieren, indem wir setzen

$$\tilde{k} = k \circ h, \quad \tilde{k}(\tau) = k(h(\tau)).$$

Wir sagen, dass \tilde{k} durch **Umparametrisieren** aus k entsteht. Man wird erwarten, dass die Länge sich dabei nicht ändert, also

$$L(\tilde{k}) = L(k). \quad (10.21)$$

In der Tat, es ist

$$\begin{aligned} L(\tilde{k}) &= \int_c^d \|\tilde{k}'(\tau)\| \, d\tau = \int_c^d \|k'(h(\tau))h'(\tau)\| \, d\tau = \int_c^d h'(\tau)\|k'(h(\tau))\| \, d\tau \\ &= \int_a^b \|k'(t)\| \, dt = L(k), \end{aligned}$$

wobei die vorletzte Gleichheit aus der Substitutionsregel der Integration folgt.

Eine bestimmte (Um-)Parametrisierung spielt eine besondere Rolle, nämlich die nach der Bogenlänge. Ist $k : [a, b] \rightarrow \mathbb{R}^d$ eine C^1 -Kurve, so ist die Länge des Kurvenstücks zwischen $k(a)$ und $k(t)$ gegeben durch

$$s(t) = \int_a^t \|k'(\tau)\| \, d\tau. \quad (10.22)$$

Wir nennen die Kurve **regulär**, wenn

$$k'(t) \neq 0 \quad (10.23)$$

gilt für alle $t \in [a, b]$. Da aus (10.22) folgt

$$s'(t) = \|k'(t)\|, \quad (10.24)$$

ist für eine reguläre Kurve die in (10.22) definierte Funktion $s : [a, b] \rightarrow [0, L(k)]$ streng monoton wachsend, dasselbe gilt für ihre Umkehrfunktion $s^{-1} : [0, L(k)] \rightarrow [a, b]$. Die durch

$$\tilde{k} : [0, L(k)] \rightarrow \mathbb{R}^n, \quad \tilde{k} = k \circ s^{-1}, \quad (10.25)$$

definierte Umparametrisierung von k heißt **Parametrisierung nach der Bogenlänge**. Sie hat die besondere Eigenschaft, dass alle Tangentenvektoren die Länge 1 haben, da

$$\tilde{k}'(\tau) = k'(s^{-1}(\tau)) \cdot (s^{-1})'(\tau) = k'(s^{-1}(\tau)) \cdot \frac{1}{s'(s^{-1}(\tau))} = \frac{k'(s^{-1}(\tau))}{\|k'(s^{-1}(\tau))\|}.$$

Zum Sprachgebrauch. Im Vorangegangenen bedeutet ‘‘Kurve’’ die Funktion $k : I \rightarrow \mathbb{R}^d$, sie enthält Information sowohl über die Bildmenge $k(I)$ als auch über das Art, wie die Bildmenge durchlaufen wird. Im allgemeinen Sprachgebrauch wird oft die Bildmenge $k(I)$ als Kurve bezeichnet und die Funktion $k : I \rightarrow \mathbb{R}^d$ als Parametrisierung der Kurve. Dieser Sprachgebrauch ist vor allem dann problematisch, wenn die Bildmenge $k(I)$ mehrfach durchlaufen wird (das sieht man dem Bild eben nicht an).

Krümmung einer ebenen Kurve. Eine Kurve ist gekrümmt, wenn ihr Tangentenvektor $k'(t)$ nicht immer in dieselbe Richtung zeigt. Die Richtung von $k'(t)$ lässt sich beispielsweise am Winkel $\varphi(t)$ in der Polarkoordinaten-Darstellung ablesen. Sei

$$k(t) = (x(t), y(t)), \quad k'(t) = (x'(t), y'(t)),$$

dann ist in Polarkoordinaten

$$x'(t) = r(t) \cos \varphi(t), \quad y'(t) = r(t) \sin \varphi(t), \quad r(t) = \sqrt{x'(t)^2 + y'(t)^2}, \quad (10.26)$$

also

$$\cos \varphi(t) = \frac{x'(t)}{r(t)}, \quad \sin \varphi(t) = \frac{y'(t)}{r(t)}.$$

Differenzieren und weitere Umformungen führen auf die Formel

$$\varphi'(t) = \frac{x'(t)y''(t) - y'(t)x''(t)}{r(t)^2}. \quad (10.27)$$

Die **Krümmung** $\kappa(t)$ der Kurve k im Kurvenpunkt $k(t)$ wird nun definiert durch die Änderung des Winkels, bezogen auf die Änderung der Bogenlänge,

$$\kappa(t) = \frac{\varphi'(t)}{s'(t)}. \quad (10.28)$$

Da $s'(t) = \|k'(t)\| = r(t)$, ergibt sich aus (10.27)

$$\kappa(t) = \frac{x'(t)y''(t) - y'(t)x''(t)}{(x'(t)^2 + y'(t)^2)^{\frac{3}{2}}}. \quad (10.29)$$

Ist k bereits nach der Bogenlänge parametrisiert, so ist $r(t) = s'(t) = 1$, und

$$\kappa(t) = x'(t)y''(t) - y'(t)x''(t). \quad (10.30)$$

Betrachten wir wieder die Kurve, die durch den Graphen einer Funktion f definiert ist,

$$k(t) = (t, f(t)),$$

so ist $x(t) = t$ und $y(t) = f(t)$, also $x' = 1$ und $x'' = 0$, und (10.29) wird zu

$$\kappa(t) = \frac{f''(t)}{\sqrt{(1 + f'(t)^2)^3}},$$

oder, wenn wir x und t identifizieren,

$$\kappa(x) = \frac{f''(x)}{\sqrt{(1 + f'(x)^2)^3}}. \quad (10.31)$$

Aus (10.28) erkennen wir wegen $s' > 0$, dass die Krümmung positiv ist, wenn φ wächst, wenn also die Krümmung gegen die Richtung des Uhrzeigersinns (nach links in Durchlaufrichtung) gerichtet ist.

Für den Kreis mit Mittelpunkt 0 und Radius r gilt

$$k(t) = (r \cos t, r \sin t), \quad k'(t) = (-r \sin t, r \cos t), \quad k''(t) = (-r \cos t, -r \sin t),$$

also

$$\kappa(t) = \frac{r^2}{r^3} = \frac{1}{r}.$$

Die Krümmung ist also in jedem Punkt des Kreises dieselbe, und sie ist umgekehrt proportional zum Radius. Sie ist positiv, falls der Kreis im mathematisch positiven Sinn (gegen den Uhrzeiger) durchlaufen wird, andernfalls negativ.

Krümmung von Kurven im Raum. Wir betrachten nun Kurven $k : I \rightarrow \mathbb{R}^3$, $I = [a, b]$. Einleitend bemerken wir, dass die Produktregel der Differentiation auch für das Skalarprodukt und das Vektorprodukt zweier Kurven gelten. Sind $k, l : I \rightarrow \mathbb{R}^3$ zwei Kurven, so gilt

$$\frac{d}{dt} \langle k(t), l(t) \rangle = \langle k'(t), l(t) \rangle + \langle k(t), l'(t) \rangle, \quad (10.32)$$

$$\frac{d}{dt} (k(t) \times l(t)) = k'(t) \times l(t) + k(t) \times l'(t), \quad (10.33)$$

wie man aus den Definitionen direkt nachprüfen kann. Sei nun $k : I \rightarrow \mathbb{R}^3$ eine reguläre C^1 -Kurve. Als fortlaufendes Beispiel betrachten wir in diesem Unterabschnitt die Schraubenlinie,

$$k(t) = (r \cos t, r \sin t, ct), \quad r, c > 0. \quad (10.34)$$

Wir definieren den **Tangenteneinheitsvektor**

$$T(t) = \frac{k'(t)}{\|k'(t)\|}. \quad (10.35)$$

Für die Schraubenlinie ergibt sich

$$T(t) = \frac{1}{\sqrt{r^2 + c^2}} (-r \sin t, r \cos t, c). \quad (10.36)$$

Wir stellen nun fest, dass der Vektor $T'(t)$ immer senkrecht auf $T(t)$ steht. Dies ergibt sich, indem wir die Gleichung

$$1 = \|T(t)\|^2 = \langle T(t), T(t) \rangle$$

differenzieren, was auf

$$0 = \langle T'(t), T(t) \rangle + \langle T(t), T'(t) \rangle = 2 \langle T'(t), T(t) \rangle \quad (10.37)$$

führt. Wir definieren den **Hauptnormalenvektor** durch

$$N(t) = \frac{T'(t)}{\|T'(t)\|}. \quad (10.38)$$

Dieser steht wie gezeigt senkrecht auf $T(t)$. Für die Schraubenlinie ergibt sich

$$T'(t) = \frac{1}{\sqrt{r^2 + c^2}} (-r \cos t, -r \sin t, 0), \quad \|T'(t)\| = \frac{r}{\sqrt{r^2 + c^2}},$$

also

$$N(t) = (-\cos t, -\sin t, 0), \quad (10.39)$$

Schließlich definieren wir den **Binormalenvektor** durch

$$B(t) = T(t) \times N(t), \quad (10.40)$$

für die Schraubenlinie ist das

$$B(t) = \frac{1}{\sqrt{r^2 + c^2}}(c \sin t, -c \cos t, r).$$

Die Vektoren $T(t)$, $N(t)$, $B(t)$ stehen senkrecht aufeinander, sie bilden ein Orthonormalsystem, welches sich beim Durchlaufen der Kurve ständig ändert. Es wird das **begleitende Dreibein** der Kurve genannt. (Dass auch $\|B(t)\| = 1$ ist, folgt aus (10.40), da $\|T(t)\| = \|N(t)\| = 1$ und $T(t) \perp N(t)$ gelten.)

Es hat sich herausgestellt, dass die wesentliche Information über das Krümmungsverhalten der Kurve bereits in zwei skalaren Größen enthalten ist. Das ist zum einen die **Krümmung**

$$\kappa(t) = \frac{\|T'(t)\|}{\|k'(t)\|}. \quad (10.41)$$

Für die Schraubenlinie ergibt sich

$$\kappa(t) = \frac{r}{\sqrt{r^2 + c^2}} \cdot \frac{1}{\sqrt{r^2 + c^2}} = \frac{r}{r^2 + c^2}.$$

Gemäß Definition (10.41) ist die Krümmung immer positiv, im zweidimensionalen Fall ergibt sich gerade der Betrag der in (10.29) definierten Zahl. Die andere skalare Größe ist die Torsion. Wir definieren zunächst den **Torsionsvektor**

$$\frac{B'(t)}{\|k'(t)\|}.$$

Wie in (10.37) ergibt sich, dass $B'(t)$ und damit auch der Torsionsvektor senkrecht auf $B(t)$ steht. Weiterhin gilt

$$\begin{aligned} B'(t) &= T'(t) \times N(t) + T(t) \times N'(t) = \frac{1}{\|T'(t)\|} T'(t) \times T'(t) + T(t) \times N'(t) \\ &= T(t) \times N'(t), \end{aligned} \quad (10.42)$$

das heißt, $B'(t)$ steht auch auf $T(t)$ senkrecht. Daher muss $B'(t)$ parallel zum Hauptnormalenvektor $N(t)$ sein. Wir definieren nun die **Torsion** $\tau(t)$ als diejenige Zahl, welche

$$-\tau(t)N(t) = \frac{B'(t)}{\|k'(t)\|} \quad (10.43)$$

erfüllt. Die Torsion entspricht also der mit einem gewissen Vorzeichen versehenen Länge des Torsionsvektors. Für die Schraubenlinie gilt

$$\frac{B'(t)}{\|k'(t)\|} = \frac{1}{\sqrt{r^2 + c^2}}(c \cos t, c \sin t, 0) \cdot \frac{1}{\sqrt{r^2 + c^2}} = \frac{c}{r^2 + c^2}(\cos t, \sin t, 0),$$

also

$$\tau(t) = \frac{c}{r^2 + c^2}.$$

In diesem Fall ist die Torsion also konstant. Ist c klein relativ zu r , so ist sie klein, für $c = 0$ ist die Torsion gleich Null. In diesem Fall degeneriert die Schraubenlinie zum immer wieder durchlaufenden Kreis in der Ebene. Allgemein gilt: Verläuft eine Kurve in einer Ebene E , so gilt $\tau(t) = 0$. Dann liegt nämlich $N'(t)$ ebenfalls in E , und da $N'(t)$ senkrecht ist auf $N(t)$, muss $N'(t)$ parallel zu $T(t)$ sein, und es folgt $B'(t) = 0$ nach (10.42) und damit $\tau(t) = 0$ nach (10.43). Die Torsion misst also in einem gewissem Sinn die “Abweichung von der Zweidimensionalität” einer Kurve.

11 Partielle Ableitungen

Wir wollen Funktionen differenzieren, die im Mehrdimensionalen definiert sind.

Funktionen mit Definitionsbereich \mathbb{R}^d oder einer Teilmenge Ω von \mathbb{R}^d nennt man auch **Felder**. Ist \mathbb{R} der Bildbereich, so spricht man von einem **Skalarfeld** oder skalaren Feld. Ein Skalarfeld ist beispielsweise die Funktion, die jedem Punkt im Raum seine Temperatur zuordnet.

Ein weiteres Beispiel ist die Funktion, welche jedem Vektor seine Länge zuordnet,

$$f(x) = \|x\|, \quad f : \mathbb{R}^d \rightarrow \mathbb{R}.$$

Solche Funktionen haben wir in Kapitel 2 und 5 bereits betrachtet.

Wir definieren die offene bzw. abgeschlossene Kugel um einen Mittelpunkt x mit Radius r ,

$$B_r(x) = \{y : y \in \mathbb{R}^d, \|y - x\| < r\}, \quad K_r(x) = \{y : y \in \mathbb{R}^d, \|y - x\| \leq r\}, \quad (11.1)$$

sowie ihren Rand, die Sphäre

$$S_r(x) = \{y : y \in \mathbb{R}^d, \|y - x\| = r\}. \quad (11.2)$$

Wir betrachten ein Intervall $I = [a, b]$ im \mathbb{R} . Innere Punkte x von I haben die Eigenschaft, dass es ein (möglicherweise sehr kleines) Intervall $(x - \varepsilon, x + \varepsilon)$ gibt, welches vollständig in $[a, b]$ enthalten ist. Der Randpunkt a (analog b) hat die Eigenschaft, dass jedes Intervall der Form $(a - \varepsilon, a + \varepsilon)$ sowohl Punkte aus I als auch Punkte außerhalb von I enthält.

Definition 11.1 Sei $\Omega \subset \mathbb{R}^d$.

Ein $x \in \Omega$ heißt **innerer Punkt** von Ω , falls es eine offene Kugel mit Mittelpunkt x gibt, welche ganz in Ω liegt. Die Menge aller inneren Punkte von Ω heißt das **Innere** von Ω , geschrieben $\text{int}(\Omega)$. Die Menge Ω heißt **offen**, wenn $\text{int}(\Omega) = \Omega$, das heißt, jeder Punkt von Ω ist innerer Punkt von Ω .

Ein $x \in \mathbb{R}^d$ heißt **Randpunkt** von Ω , falls jede offene Kugel mit Mittelpunkt x sowohl einen Punkt aus Ω als auch einen Punkt enthält, der nicht in Ω liegt. Die Menge aller Randpunkte von Ω heißt der **Rand** von Ω , geschrieben $\partial\Omega$. \square

Man kann zeigen: Die Menge Ω ist abgeschlossen genau dann, wenn $\partial\Omega \subset \Omega$, das heißt, wenn jeder Randpunkt von Ω auch zu Ω gehört.

Es ist etwa

$$\begin{aligned} \text{int}(B_r(x)) &= \text{int}(K_r(x)) = B_r(x), & \partial B_r(x) &= \partial K_r(x) = S_r(x), \\ \text{int}(S_r(x)) &= \emptyset, & \partial S_r(x) &= S_r(x). \end{aligned}$$

$B_r(x)$ ist offen, $K_r(x)$ und $S_r(x)$ sind abgeschlossen. Bei Mengen mit einem ‘scharfen Rand’ ist das auch anschaulich klar, z.B. bei Teilmengen des \mathbb{R}^2 , die von einer ‘glatten’ Kurve begrenzt werden. Betrachten wir aber beispielsweise \mathbb{Q} als Teilmenge von \mathbb{R} , so ist

$$\text{int}(\mathbb{Q}) = \emptyset, \quad \partial\mathbb{Q} = \mathbb{R},$$

da in beliebig kleiner Entfernung von jeder rationalen Zahl sowohl rationale als auch irrationale Zahlen liegen.

Wir betrachten Beispiele von Funktionen $f : \Omega \rightarrow \mathbb{R}$ mit $\Omega = \mathbb{R}^d$.

$$f(x) = a^T x = \langle a, x \rangle = \sum_{i=1}^d a_i x_i, \quad a \in \mathbb{R}^d \text{ gegeben,} \quad (\text{lineare Funktion})$$

$$f(x) = x^T A x + b^T x + c = \sum_{i,j=1}^d a_{ij} x_i x_j + \sum_{i=1}^d b_i x_i + c, \quad (\text{quadratische Funktion})$$

$$f(x) = \|x\| = \sqrt{\sum_{i=1}^d x_i^2}.$$

Im Falle $d = 2$ oder $d = 3$ schreibt man oft (x, y) bzw. (x, y, z) für die Komponenten des Arguments, also etwa

$$f(x, y) = 3x + 2y - 4, \quad f(x, y, z) = x^2 y - x e^z.$$

Im Falle $d = 2$, also $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^2$, ist es möglich, sich den Graphen von f im Raum zu veranschaulichen, indem man die Menge

$$M = \{(x, y, f(x, y)) : (x, y) \in \Omega\}$$

darstellt. Ist etwa $\Omega = K_1(0)$ die Einheitskreisscheibe in der Ebene, so liefert

$$f(x, y) = x^2 + y^2$$

als Menge M ein nach oben geöffnetes Paraboloid, welches sich über der Einheitskreisscheibe erhebt.

Ableitungen. Wir behandeln nun die Frage, wie man Funktionen differenzieren kann, deren Definitionsgebiete im Mehrdimensionalen liegen. Ist beispielsweise $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = x^2 y^4,$$

so können wir y als fest gegebenen Parameter betrachten und die Funktion

$$g(x) = x^2 y^4$$

“ganz normal” differenzieren mit dem Ergebnis

$$g'(x) = 2xy^4. \quad (11.3)$$

Genaugut könnten wir x festhalten und die Funktion

$$h(y) = x^2 y^4$$

nach ihrem Argument (also y) differenzieren mit dem Ergebnis

$$h'(y) = x^2 \cdot 4y^3.$$

Die Definition der Ableitung von g in (11.3),

$$g'(x) = \lim_{t \rightarrow 0} \frac{g(x+t) - g(x)}{t} = \lim_{t \rightarrow 0} \frac{f(x+t, y) - f(x, y)}{t},$$

ist ein Spezialfall der folgenden Definition.

Definition 11.2 (Partielle Ableitung)

Sei $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$, sei x innerer Punkt von Ω . Die **partielle Ableitung von f nach der i -ten Komponente im Punkt x** ist definiert als der Grenzwert (sofern dieser existiert)

$$\lim_{t \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_d) - f(x_1, \dots, x_d)}{t}. \quad (11.4)$$

□

Die partielle Ableitung von f nach x_i an der Stelle x ist also eine Zahl. Für sie sind verschiedene Bezeichnungen üblich, etwa

$$\frac{\partial f}{\partial x_i}(x), \quad \partial_i f(x), \quad \partial_{x_i} f(x). \quad (11.5)$$

Verwendet man im Zweidimensionalen die Notation $f(x, y)$, so schreibt man für die partiellen Ableitungen entsprechend

$$\frac{\partial f}{\partial x}(x, y), \quad \frac{\partial f}{\partial y}(x, y), \quad \partial_x f(x, y), \quad \partial_y f(x, y), \quad (11.6)$$

entsprechend im Dreidimensionalen

$$\frac{\partial f}{\partial x}(x, y, z), \quad \frac{\partial f}{\partial y}(x, y, z), \quad \frac{\partial f}{\partial z}(x, y, z), \quad \text{usw.}$$

Beispiel wie oben:

$$f(x, y) = x^2 y^4, \quad \frac{\partial f}{\partial y}(x, y) = 4x^2 y^3, \quad \frac{\partial f}{\partial y}(3, -2) = 4 \cdot 3^2 \cdot (-2)^3 = -288.$$

Wir können uns die partielle Ableitung folgendermaßen vorstellen. Ist etwa $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, so entspricht die Funktion

$$g(t) = f(x + t, y), \quad g : \mathbb{R} \rightarrow \mathbb{R},$$

gerade dem Verhalten von f in Richtung der x -Koordinate, ausgehend vom Punkt (x, y) . Es ist dann

$$\frac{\partial f}{\partial x}(x, y) = g'(0).$$

Wir kehren zur allgemeinen d -dimensionalen Situation zurück. Sei $f : \Omega \rightarrow \mathbb{R}$, sei $\Omega \subset \mathbb{R}^d$ offen. Falls der Grenzwert (11.4) in jedem Punkt $x \in \Omega$ und für jedes i , $1 \leq i \leq d$, existiert, so heißt f **partiell differenzierbar** in Ω . Die partiellen Ableitungen liefern in diesem Fall Funktionen

$$\frac{\partial f}{\partial x_i} : \Omega \rightarrow \mathbb{R}, \quad \text{oder auch} \quad \partial_i f : \Omega \rightarrow \mathbb{R}, \quad (11.7)$$

sie sind also ebenfalls Skalarfelder. Sind alle diese Funktionen stetig in Ω , so heißt f **stetig differenzierbar in Ω** . (In diesem Fall lässt man das Wort "partiell" weg.)

Da die partielle Ableitung nichts anderes ist als die Ableitung nach einer Koordinatenrichtung (bei festgehaltenen anderen Richtungen), kommen die Rechenregeln aus dem Eindimensionalen auch für die Berechnung von partiellen Ableitungen zum Tragen.

Beispiele: Die lineare Funktion

$$f(x) = \langle a, x \rangle = \sum_{i=1}^d a_i x_i, \quad \frac{\partial f}{\partial x_i}(x) = \partial_i f(x) = a_i,$$

In diesem Fall ist die partielle Ableitung an jeder Stelle x dieselbe, als Funktion $\partial_i f : \mathbb{R}^d \rightarrow \mathbb{R}$ aufgefasst ist sie konstant.

Für die quadratische Funktion mit symmetrischer Matrix $A \in \mathbb{R}^{(d,d)}$

$$f(x) = x^T A x + b^T x + c = \sum_{i,j=1}^d a_{ij} x_i x_j + \sum_{i=1}^d b_i x_i + c$$

gilt (man muss 3 verschiedene Indizes unterscheiden)

$$\frac{\partial f}{\partial x_k}(x) = \partial_k f(x) = \sum_{i=1}^d a_{ik} x_i + \sum_{j=1}^d a_{kj} x_j + b_k = 2 \sum_{j=1}^d a_{kj} x_j + b_k = 2 \langle A_{k\cdot}, x \rangle + b_k,$$

wobei $A_{k\cdot}$ die k -te Zeile von A bezeichnet. Die partielle Ableitung $\partial_k f$ ist also eine affin lineare Funktion, im Falle $b_k = 0$ eine lineare Funktion.

Man kann die partiellen Ableitungen zu einem Vektor zusammenfassen, er heißt der **Gradient** von f , geschrieben

$$\nabla f = \text{grad } f = \begin{pmatrix} \partial_1 f \\ \vdots \\ \partial_d f \end{pmatrix}. \quad (11.8)$$

Beispiel

$$f(x_1, x_2) = x_1^2 x_2^4, \quad \nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 x_2^4 \\ 4x_1^2 x_2^3 \end{pmatrix}, \quad \nabla f(2, 1) = \begin{pmatrix} 4 \\ 16 \end{pmatrix}.$$

Ist $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $g : \mathbb{R} \rightarrow \mathbb{R}$, so erhalten wir die partiellen Ableitungen von $g \circ f$,

$$\partial_i (g \circ f)(x) = \frac{\partial (g \circ f)}{\partial x_i}(x) = g'(f(x)) \cdot \frac{\partial f}{\partial x_i}(x) \quad (11.9)$$

direkt aus der **Kettenregel** im Eindimensionalen, da wir in diesem Fall f nur als Funktion einer Variablen auffassen. Als Beispiel betrachten wir

$$r(x) = \|x\| = \sqrt{\sum_{i=1}^d x_i^2}, \quad r : \mathbb{R}^d \rightarrow \mathbb{R}. \quad (11.10)$$

Es ergibt sich

$$\partial_i r(x) = \frac{1}{2\sqrt{\sum_{j=1}^d x_j^2}} \cdot 2x_i = \frac{x_i}{r(x)}, \quad 1 \leq i \leq d, \quad (11.11)$$

und damit auch

$$\nabla r(x) = \text{grad } r(x) = \frac{x}{r(x)} = \frac{x}{\|x\|}. \quad (11.12)$$

Auf diese Weise erhalten wir die Ableitungen von **radialsymmetrischen** Funktionen, das sind solche Funktionen $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$, welche nur von der Länge des Argumentvektors abhängen, etwa

$$\varphi(x) = g(\|x\|), \quad \text{wobei } g : \mathbb{R} \rightarrow \mathbb{R}, \quad (11.13)$$

als

$$\partial_i \varphi(x) = g'(\|x\|) \frac{x_i}{\|x\|}, \quad \nabla \varphi(x) = g'(\|x\|) \frac{x}{\|x\|}. \quad (11.14)$$

Beispiel:

$$\varphi(x) = \ln(\|x\|), \quad \partial_i \varphi(x) = \frac{1}{\|x\|} \frac{x_i}{\|x\|} = \frac{x_i}{\|x\|^2}, \quad \partial_1 \varphi(2, 1) = \frac{2}{5}.$$

Wir betrachten nun die Veränderung eines Skalarfeldes f entlang einer Kurve k ,

$$g(t) = f(k(t)), \quad k : \mathbb{R} \rightarrow \mathbb{R}^d, \quad f : \mathbb{R}^d \rightarrow \mathbb{R}. \quad (11.15)$$

Die Ableitung von g wird mit der entsprechenden Variante der **Kettenregel** berechnet,

$$g'(t) = \langle \nabla f(k(t)), k'(t) \rangle = \sum_{i=1}^n \partial_i f(k(t)) \cdot k'_i(t), \quad (11.16)$$

wobei $k'(t) \in \mathbb{R}^d$ der Tangentenvektor an die Kurve k im Punkt $k(t)$ ist, siehe Kapitel 10. Dass (11.16) richtig ist, ergibt sich nicht unmittelbar aus der eindimensionalen Situation, sondern als Spezialfall der allgemeinen mehrdimensionalen Situation bei Vektorfeldern. Als Beispiel betrachten wir

$$f(x_1, x_2) = x_1^2 + x_2^3, \quad k(t) = t \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (11.17)$$

Die Kurve k ist hier die Gerade durch 0 in Richtung der Winkelhalbierenden. Es ist

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 3x_2^2 \end{pmatrix}, \quad k(t) = \begin{pmatrix} t \\ t \end{pmatrix}, \quad k'(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

also

$$g'(t) = \langle \nabla f(k(t)), k'(t) \rangle = \left\langle \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = 2t + 3t^2.$$

Dasselbe Ergebnis erhalten wir, wenn wir g formelmäßig darstellen, indem wir k in f einsetzen,

$$g(t) = f(k(t)) = f(t, t) = t^2 + t^3,$$

und dann differenzieren. Wenn f und k durch einfache Formeln gegeben sind, ist das der einfachere Weg.

Wir betrachten eine Situation wie im vorangegangenen Beispiel (11.17), aber mit einer allgemeinen Geraden

$$k(t) = x + tv, \quad x, v \in \mathbb{R}^d. \quad (11.18)$$

Die Ableitung der zusammengesetzten Funktion $g(t) = f(k(t))$, $f : \mathbb{R}^d \rightarrow \mathbb{R}$, im Punkt $t = 0$ ergibt sich gemäß (11.16) als

$$g'(0) = \langle \nabla f(x), v \rangle = \sum_{i=1}^d \partial_i f(x) \cdot v_i. \quad (11.19)$$

Andererseits ist

$$g'(0) = \lim_{t \rightarrow 0} \frac{g(t) - g(0)}{t} = \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t}. \quad (11.20)$$

Definition 11.3 (Richtungsableitung)

Sei $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$, sei x innerer Punkt von Ω , sei $v \in \mathbb{R}^d$. Die **Richtungsableitung von f im Punkt x in Richtung v** ist definiert als der Grenzwert (sofern dieser existiert)

$$\partial_v f(x) = \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} \quad (11.21)$$

□

Gemäß (11.19) erhalten wir die Richtungsableitung in eine beliebige Richtung aus den partiellen Ableitungen mit

$$\partial_v f(x) = \langle \nabla f(x), v \rangle. \quad (11.22)$$

Die Richtungsableitungen in Richtung der Koordinatenachsen sind nichts anderes als die partiellen Ableitungen,

$$\partial_{e_i} f(x) = \langle \nabla f(x), e_i \rangle = \frac{\partial f}{\partial x_i}(x). \quad (11.23)$$

Höhere partielle Ableitungen. Ist beispielsweise

$$f(x, y) = x^2 y^3, \quad f : \mathbb{R}^2 \rightarrow \mathbb{R},$$

so ist

$$\partial_x f(x, y) = 2xy^3, \quad \partial_x f : \mathbb{R}^2 \rightarrow \mathbb{R},$$

und nichts hindert uns daran, diese Funktion nochmals partiell abzuleiten, beispielsweise nach y ,

$$\partial_y \partial_x f(x, y) = 6xy^2, \quad \partial_y \partial_x f : \mathbb{R}^2 \rightarrow \mathbb{R}.$$

Statt $\partial_y \partial_x f$ schreibt man auch

$$\frac{\partial^2 f}{\partial y \partial x}.$$

Für eine beliebige Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ geht das in einem Punkt $(x, y) \in \mathbb{R}^2$ natürlich nur, wenn der entsprechende Grenzwert existiert, in diesem Fall also

$$\lim_{t \rightarrow 0} \frac{\partial_x f(x, y + t) - \partial_x f(x, y)}{t}.$$

Definition 11.4 (Zweite partielle Ableitungen)

Sei $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$, x innerer Punkt von Ω . Die zweiten partiellen Ableitungen $\partial_j \partial_i f$ von f sind im Punkt x definiert als der Grenzwert

$$\partial_j \partial_i f(x) = \lim_{t \rightarrow 0} \frac{\partial_i f(x + t e_j) - \partial_i f(x)}{t}, \quad (11.24)$$

wobei e_j der j -te Einheitsvektor ist. □

Ist Ω offen, und existieren in jedem Punkt alle zweiten partiellen Ableitungen, so erhalten wir Funktionen

$$\partial_j \partial_i f : \Omega \rightarrow \mathbb{R}, \quad 1 \leq i, j \leq d.$$

Sind diese Funktionen alle stetig, so heißt f **zweimal stetig differenzierbar** in Ω .

Die partiellen Ableitungen $\partial_i f$ heißen auch **erste partielle Ableitungen** von f . Entsprechend definiert man die $(k+1)$ -ten (dritten, vierten, ...) partiellen Ableitungen als partielle Ableitungen der k -ten partiellen Ableitungen.

Aus den ersten partiellen Ableitungen kann man wie besprochen einen Vektor bilden, den Gradienten. Aus den zweiten partiellen Ableitungen in einem Punkt x kann man eine $(d \times d)$ -Matrix bilden, sie heißt die **Hesse-Matrix** und ist definiert als

$$H_f(x) = \begin{pmatrix} \partial_1 \partial_1 f(x) & \cdots & \partial_d \partial_1 f(x) \\ \partial_1 \partial_2 f(x) & \cdots & \partial_d \partial_2 f(x) \\ \vdots & \ddots & \vdots \\ \partial_1 \partial_d f(x) & \cdots & \partial_d \partial_d f(x) \end{pmatrix} \quad (11.25)$$

Für

$$f(x_1, x_2) = x_1^2 x_2^3$$

gilt beispielsweise $\partial_1 f(x_1, x_2) = 2x_1 x_2^3$, $\partial_2 f(x_1, x_2) = 3x_1^2 x_2^2$, und

$$H_f(x) = \begin{pmatrix} 2x_2^3 & 6x_1 x_2^2 \\ 6x_1 x_2^2 & 6x_1^2 x_2 \end{pmatrix}, \quad H_f(2, 1) = \begin{pmatrix} 2 & 12 \\ 12 & 24 \end{pmatrix}.$$

Dass diese Matrix symmetrisch ist, ist kein Zufall, es gilt nämlich

Satz 11.5 *Ist $f : \Omega \rightarrow \mathbb{R}$ zweimal stetig differenzierbar, so ist die Hesse-Matrix von f symmetrisch, das heißt, es gilt*

$$\partial_j \partial_i f(x) = \partial_i \partial_j f(x), \quad (11.26)$$

für alle x und alle i, j . □

Niveaumengen. Skalarfelder kann man sich auch durch ihre Niveaumengen

$$N_c(f) = \{x : x \in \Omega, f(x) = c\} \quad (11.27)$$

veranschaulichen, sowohl für $d = 2$ als auch für $d = 3$. So liefert

$$f(x) = \sqrt{(x_1 - 1)^2 + x_2^2}$$

als Niveaumengen $N_c(f)$ gerade die Kreise in der Ebene mit Mittelpunkt $(1, 0)$ und Radius c . In diesem Fall handelt es sich bei den Niveaumengen um Kurven, die sich alternativ darstellen lassen durch

$$k(t) = (1 + c \cos t, c \sin t), \quad k : [0, 2\pi] \rightarrow \mathbb{R}^2.$$

Im Fall $d = 3$ stellen die Niveaumengen einer Funktion im regulären Fall eine gekrümmte Fläche dar, so ist $N_c(f)$ für

$$f(x) = \sqrt{(x_1 - 1)^2 + x_2^2 + x_3^2}$$

gerade die Sphäre um $(1, 0, 0)$ mit Radius c .

Aus dem Zusammenwirken der beiden Darstellungen der Niveaumenge erhalten wir eine geometrische Interpretation des Gradienten. Ist $k : I \rightarrow \mathbb{R}^d$ eine auf einem Intervall I definierte Kurve, die vollständig innerhalb einer Niveaumenge $N_c(f)$ verläuft, so gilt

$$f(k(t)) = c, \quad \text{für alle } t \in I.$$

Differenzieren wir beide Seiten nach t , so erhalten wir aus der Kettenregel (11.16)

$$\langle \nabla f(k(t)), k'(t) \rangle = 0, \quad (11.28)$$

das heißt, im Punkt $k(t)$ steht der Gradient von f senkrecht auf dem Tangentenvektor an die Kurve. Ist nun $x \in N_c(f)$ ein fester Punkt in der Niveaumenge, so gilt (11.28) für jede Kurve durch x . Der Gradient $\nabla f(x)$ steht also senkrecht auf die von allen möglichen Tangentenvektoren solcher Kurven gebildeten tangentialen Hyperebene der Niveaumenge und zeigt in die Richtung des steilsten Anstiegs der Funktion f .

Definition 11.6 (Tangentenvektor an eine Niveaumenge)

Sei $f : \Omega \rightarrow \mathbb{R}$ stetig differenzierbar, $\Omega \subset \mathbb{R}^d$ offen, sei $x \in \Omega$ mit $f(x) = c$ und $\nabla f(x) \neq 0$. Ein Vektor $v \in \mathbb{R}^d$ heißt Tangentialvektor an $N_c(f)$ in x , falls

$$\langle \nabla f(x), v \rangle = 0. \quad (11.29)$$

Die Menge

$$T(x) = \{v : v \text{ ist Tangentialvektor an } N_c(f) \text{ in } x\} \quad (11.30)$$

heißt der Tangentialraum an $N_c(f)$ im Punkt x . □

In der Situation von Definition 11.6 gilt, dass der Tangentialraum $T(x)$ ein $(d - 1)$ -dimensionaler Unterraum des \mathbb{R}^d ist. Der affine Unterraum $x + T(x)$ heißt auch Tangentialebene an $N_c(f)$ im Punkt x .

Maxima und Minima von Skalarfeldern. Im Eindimensionalen werden Extremwerte einer Funktion, also insbesondere Maxima und Minima, dadurch bestimmt, dass man die erste und die zweite Ableitung untersucht. Ein entsprechendes Vorgehen funktioniert auch im Mehrdimensionalen.

Wir betrachten die Aufgabe: Gegeben ist eine Funktion $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$, gesucht ist $x \in \mathbb{R}^d$ mit

$$f(x) = \min_{y \in \Omega} f(y). \quad (11.31)$$

Ein Beispiel ist etwa die Funktion

$$f(x_1, x_2) = x_1^2 + x_2^2.$$

Es gilt

$$0 = f(0) < f(x), \quad \text{für alle } x = (x_1, x_2) \neq 0,$$

also ist 0 ein (striktes) Minimum von f auf $\Omega = \mathbb{R}^2$.

Wir nehmen nun einmal an, $x \in \Omega$ sei ein Minimum von f auf Ω , Ω sei offen. Zu einem beliebigen Vektor $v \in \mathbb{R}^d$ definieren wir

$$g(t) = f(x + tv). \quad (11.32)$$

Da Ω offen ist, ist g definiert in einem offenen Intervall $I \subset \mathbb{R}$ mit $0 \in I$, und 0 ist Minimum von g auf I , da

$$g(0) = f(x) \leq f(x + tv) = g(t), \quad \text{für alle } t \in I.$$

Ist f zweimal stetig differenzierbar, so auch g , und es muss gelten

$$g'(0) = 0, \quad g''(0) \geq 0. \quad (11.33)$$

Aus der Kettenregel (11.16) folgt

$$g'(t) = \langle \nabla f(x + tv), v \rangle,$$

also

$$0 = g'(0) = \langle \nabla f(x), v \rangle. \quad (11.34)$$

Da v ein beliebiger Vektor war, muss (11.34) für alle $v \in \mathbb{R}^d$ gelten. Es folgt

$$\nabla f(x) = 0, \quad \text{also } \partial_i f(x) = 0, \quad 1 \leq i \leq d. \quad (11.35)$$

Ein solcher Punkt x mit $\nabla f(x) = 0$ heißt **kritischer Punkt** von f .

Wir differenzieren nun

$$g'(t) = \langle \nabla f(x + tv), v \rangle = \sum_{i=1}^d v_i \partial_i f(x + tv),$$

indem wir nochmals die Kettenregel (11.16) auf die einzelnen Summanden anwenden,

$$\begin{aligned} g''(t) &= \sum_{i=1}^d v_i \frac{d}{dt} \partial_i f(x + tv) = \sum_{i=1}^d v_i \left(\sum_{j=1}^d \partial_j \partial_i f(x + tv) v_j \right) \\ &= \sum_{i,j=1}^d v_i \left[\partial_j \partial_i f(x + tv) \right] v_j = v^T H_f(x + tv) v, \end{aligned}$$

also

$$0 \leq g''(0) = v^T H_f(x) v. \quad (11.36)$$

Da v ein beliebiger Vektor war, muss (11.36) für alle $v \in \mathbb{R}^d$ gelten, das bedeutet aber gerade, dass $H_f(x)$, die Hesse-Matrix von f im Punkt x , positiv semidefinit ist. Wir fassen das Ergebnis dieser Überlegungen im folgenden Satz zusammen.

Satz 11.7 Sei $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$ offen. Ist x ein lokales Minimum von f in Ω , so gilt

$$\nabla f(x) = 0, \quad (11.37)$$

und die Hesse-Matrix $H_f(x)$ ist positiv semidefinit. \square

Analog erhalten wir für das Maximum einer Funktion f , indem wir Satz 11.7 auf $-f$ anwenden:

Satz 11.8 Sei $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$ offen. Ist x ein lokales Maximum von f in Ω , so gilt

$$\nabla f(x) = 0, \quad (11.38)$$

und die Hesse-Matrix $H_f(x)$ ist negativ semidefinit. \square

Im Eindimensionalen folgt umgekehrt aus $f'(x) = 0$ und $f''(x) > 0$, dass x ein striktes lokales Minimum ist. Eine analoger Satz gilt auch im Mehrdimensionalen, wir werden ihn nicht beweisen.

Satz 11.9 Sei $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$ offen, $x \in \Omega$.

Ist $\nabla f(x) = 0$ und $H_f(x)$ positiv definit, so ist x ein striktes lokales Minimum.

Ist $\nabla f(x) = 0$ und $H_f(x)$ negativ definit, so ist x ein striktes lokales Maximum. \square

Ob die Hessematrix die in den Sätzen 11.7 – 11.9 verlangten Eigenschaften hat, lässt sich an ihren Eigenwerten erkennen, siehe Lineare Algebra.

Wir illustrieren die verschiedenen Situationen an möglichst einfachen Beispielen. In allen diesen Beispielen ist 0 der einzige kritische Punkt von f und damit der einzige Kandidat für ein Minimum oder Maximum.

$$f(x, y) = x^2 + 2y^2, \quad \text{dann ist} \quad \nabla f(x, y) = \begin{pmatrix} 2x \\ 4y \end{pmatrix}, \quad H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Die Hesse-Matrix in $x = y = 0$ (die in diesem Fall nicht von (x, y) abhängt) hat die Eigenwerte 2 und 4, ist also positiv definit. Der Nullpunkt ist also ein striktes lokales Minimum (in diesem Fall sogar ein globales Minimum). Sei nun

$$f(x, y) = x^2 - 2y^2, \quad \text{dann ist} \quad \nabla f(x, y) = \begin{pmatrix} 2x \\ -4y \end{pmatrix}, \quad H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}.$$

Die Hesse-Matrix in 0 ist indefinit, da sie einen positiven und einen negativen Eigenwert hat. Der Nullpunkt ist also weder ein lokales Maximum noch ein lokales Minimum, er ist vielmehr ein **Sattelpunkt**.

Ein lokales Maximum in 0 erhalten wir für

$$f(x, y) = -x^2 - 2y^2, \quad \text{dann ist} \quad \nabla f(x, y) = \begin{pmatrix} -2x \\ -4y \end{pmatrix}, \quad H_f(x, y) = \begin{pmatrix} -2 & 0 \\ 0 & -4 \end{pmatrix},$$

hier hat die Hesse-Matrix im Nullpunkt die Eigenwerte -2 und -4 , und ist daher negativ definit.

Ist die Hesse-Matrix (positiv oder negativ) semidefinit, aber nicht (positiv oder negativ) definit, so lässt sich durch sie nicht unterscheiden, ob ein Minimum bzw. Maximum vorliegt oder nicht. In jedem der Beispiele

$$f(x, y) = x^2 + y^4, \quad f(x, y) = x^2, \quad f(x, y) = x^2 + y^3, \quad (11.39)$$

gilt

$$\nabla f(0, 0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad (11.40)$$

im ersten Fall ist 0 ein striktes lokales Minimum, im zweiten Fall ist 0 ein lokales, aber nicht striktes Minimum (da $f(0, y) = 0$ für alle y), im dritten Fall liegt weder ein lokales Minimum noch ein lokales Maximum vor. Durch (11.40) lässt sich lediglich ausschließen, dass 0 ein lokales Maximum ist, denn die Hesse-Matrix hat einen positiven Eigenwert.

Es hat sich also ein zur Kurvendiskussion im Eindimensionalen völlig analoges **Rechen-schema** zur Bestimmung von Maxima und Minima einer Funktion $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$ offen, ergeben.

1. Bestimmung aller kritischen Punkte von f , also aller Punkte $x \in \Omega$ mit $\nabla f(x) = 0$.
2. Für jeden kritischen Punkt x : Berechnung der Hesse-Matrix $H_f(x)$.
3. Auswertung der Information aus der Hesse-Matrix an einem kritischen Punkt:
 - Sind alle Eigenwerte positiv, so liegt ein striktes lokales Minimum vor.
 - Sind alle Eigenwerte positiv oder Null, so liegt jedenfalls kein lokales Maximum vor.
 - Sind alle Eigenwerte negativ, so liegt ein striktes lokales Maximum vor.
 - Sind alle Eigenwerte negativ oder Null, so liegt jedenfalls kein lokales Minimum vor.
 - Gibt es positive und negative Eigenwerte, so liegt weder ein Maximum noch ein Minimum vor.

Der erste Schritt, die Bestimmung der kritischen Punkte von f , führt im allgemeinen auf ein nichtlineares Gleichungssystem von d Gleichungen für d Unbekannte,

$$\partial_i f(x_1, \dots, x_d) = 0, \quad 1 \leq i \leq d. \quad (11.41)$$

Dafür gibt es numerische Verfahren, etwa das **Newton-Verfahren** und seine Varianten, welche Folgen (x^n) von Vektoren konstruieren, die gegen eine Lösung von (11.41) konvergieren. Andere Verfahren zur Bestimmung eines Minimums von f konstruieren Folgen (x^n) von Vektoren mit

$$f(x^{n+1}) < f(x^n), \quad (11.42)$$

welche gegen ein Minimum konvergieren. Ein Beispiel für ein solches Verfahren ist das **Gradientenverfahren**, welches sich die Beobachtung zunutze macht, dass der negative Gradient in die Richtung des steilsten Abstiegs von f zeigt. Es hat die Form

$$x^{n+1} = x^n - \lambda_n \nabla f(x^n), \quad (11.43)$$

wobei λ_n eine geeignet zu definierende Schrittweite ist. Die Idee dieses Verfahrens ist einfach zu verstehen, andere Verfahren sind aber besser.

Taylorentwicklung für Skalarfelder. Um die Taylorentwicklung für ein Skalarfeld $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^d$, in einem Punkt $x+h$ in der Nähe eines Punktes $x \in \text{int}(\Omega)$ zu erhalten, betrachten wir wieder die (auf einem geeigneten Intervall I definierte) Funktion

$$g(t) = f(x + th), \quad g : I \rightarrow \mathbb{R}. \quad (11.44)$$

Für $f(x+h) = g(1)$ gilt, falls f hinreichend oft differenzierbar ist,

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(0) + \dots + \frac{1}{m!}g^{(m)}(0) + \frac{1}{(m+1)!}g^{(m+1)}(t), \quad (11.45)$$

für ein $t \in (0, 1)$. Die Ableitungen von g erhalten wir mit Hilfe der Kettenregel aus den Ableitungen von f , wie wir das weiter vorne in diesem Kapitel für die erste und zweite Ableitung ausgerechnet haben. Es ist

$$\begin{aligned} g(0) &= f(x) \\ g'(0) &= \langle \nabla f(x), h \rangle = \sum_{i=1}^d \partial_i f(x) h_i \\ g''(0) &= h^T H_f(x) h = \sum_{i,j=1}^d \partial_i \partial_j f(x) h_i h_j \\ g'''(0) &= \sum_{i,j,k=1}^d \partial_i \partial_j \partial_k f(x) h_i h_j h_k \end{aligned} \quad (11.46)$$

und so weiter. Es ergibt sich für die Taylorentwicklung erster Ordnung ($m = 1$ in (11.45))

$$f(x+h) = f(x) + h^T \nabla f(x) + O(\|h\|^2), \quad (11.47)$$

und für die Taylorentwicklung zweiter Ordnung ($m = 2$ in (11.45))

$$f(x+h) = f(x) + h^T \nabla f(x) + \frac{1}{2} h^T H_f(x) h + O(\|h\|^3), \quad (11.48)$$

und entsprechend für die höheren Ordnungen.

Als Beispiel betrachten wir

$$f(x_1, x_2) = x_1^3 e^{x_2}.$$

Es ist

$$\partial_1 f(x_1, x_2) = 3x_1^2 e^{x_2}, \quad \partial_2 f(x_1, x_2) = x_1^3 e^{x_2}, \quad (11.49)$$

$$\partial_1 \partial_1 f(x_1, x_2) = 6x_1 e^{x_2}, \quad \partial_1 \partial_2 f(x_1, x_2) = 3x_1^2 e^{x_2}, \quad \partial_2 \partial_2 f(x_1, x_2) = x_1^3 e^{x_2}. \quad (11.50)$$

Im Punkt $x = (1, 0)$ gilt also

$$\nabla f(1, 0) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad H_f(1, 0) = \begin{pmatrix} 6 & 3 \\ 3 & 1 \end{pmatrix},$$

also hat die Taylorentwicklung zweiter Ordnung die Form

$$\begin{aligned} f(x+h) &= f(x) + h^T \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \frac{1}{2} h^T \begin{pmatrix} 6 & 3 \\ 3 & 1 \end{pmatrix} h + O(\|h\|^3) \\ &= 1 + 3h_1 + h_2 + 3h_1^2 + 3h_1h_2 + \frac{1}{2}h_2^2 + O(\|h\|^3). \end{aligned}$$

Parameterabhängige Integrale. Wir betrachten als Beispiel die Funktion

$$F(x) = \int_0^1 \frac{1}{x+y} dy, \quad x > 0. \quad (11.51)$$

Falls wir wüssten, dass F stetig ist, könnten wir Grenzwerte von F durch Einsetzen berechnen, also beispielsweise

$$\lim_{x \rightarrow 1} F(x) = F(1) = \int_0^1 \frac{1}{1+y} dy = \ln(1+y) \Big|_{y=0}^{y=1} = \ln(2) - \ln(1) = \ln(2).$$

Zur Berechnung der Ableitung von F würden wir gerne das Differenzieren (nach x) und das Integrieren (nach y) vertauschen und rechnen

$$F'(x) = \int_0^1 \frac{\partial}{\partial x} \frac{1}{x+y} dy = \int_0^1 -\frac{1}{(x+y)^2} dy = \frac{1}{x+y} \Big|_{y=0}^{y=1} = \frac{1}{x+1} - \frac{1}{x} = -\frac{1}{x(x+1)}.$$

Beides funktioniert in diesem Beispiel, wie man erkennt, wenn man die explizite Formel für F ,

$$F(x) = \ln(x+y) \Big|_{y=0}^{y=1} = \ln(x+1) - \ln(x),$$

betrachtet. Eine solche explizite Formel steht aber nicht immer zur Verfügung.

Allgemein stellen sich die beiden folgenden Fragen. Sei

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

wobei $f : I \times [c, d] \rightarrow \mathbb{R}$ gegeben ist, $I \subset \mathbb{R}$ Intervall. Gilt dann, falls $x_n \rightarrow x$,

$$\lim_{n \rightarrow \infty} F(x_n) = \lim_{n \rightarrow \infty} \int_c^d f(x_n, y) dy = \int_c^d \lim_{n \rightarrow \infty} f(x_n, y) dy = \int_c^d f(x, y) dy = F(x), \quad (11.52)$$

und gilt, falls f partiell nach x differenzierbar ist,

$$F'(x) = \int_c^d \partial_x f(x, y) dy \quad ? \quad (11.53)$$

In beiden Fällen liegt das Problem in der Vertauschbarkeit zweier Grenzprozesse.

Das folgende Beispiel zeigt, dass (11.52) nicht immer gilt. Wir definieren $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ durch

$$f(x, y) = \begin{cases} 0, & x = 0, \\ \frac{y}{x^2}, & 0 \leq y < x, \\ \frac{2x-y}{x^2}, & x \leq y < 2x, \\ 0, & 2x \leq y. \end{cases} \quad (11.54)$$

Für jedes feste y ist die Funktion f stetig bezüglich x , das heißt,

$$\lim_{n \rightarrow \infty} f(x_n, y) = f(x, y), \quad \text{falls } x_n \rightarrow x,$$

gilt für alle $y \in [0, 1]$, aber es ist

$$F(x) = \int_0^1 f(x, y) dy = 1 \neq 0 = \int_0^1 f(0, y) dy = F(0), \quad \text{falls } 0 < x \leq \frac{1}{2},$$

also gilt (11.52) nicht für $x = 0$.

Wir betrachten

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

sei $(x_n)_{n \in \mathbb{N}}$ eine beliebige Folge im Intervall I mit $x_n \rightarrow x \in I$. Aus dem Mittelwertsatz der Differentialrechnung folgt, falls $\partial_x f$ existiert,

$$f(x_n, y) - f(x, y) = (x_n - x) \cdot \partial_x f(\xi_n, y),$$

für eine geeignete Zwischenstelle $\xi_n \in (x, x_n)$, die auch noch von y abhängt. Ist $\partial_x f$ außerdem beschränkt, etwa

$$|\partial_x f(\xi, y)| \leq M,$$

für alle ξ und alle y , so folgt

$$0 \leq |F(x_n) - F(x)| \leq \left| \int_c^d f(x_n, y) - f(x, y) dy \right| \leq \int_c^d M|x_n - x| dy = (d - c)M|x_n - x|,$$

also $F(x_n) \rightarrow F(x)$. In solchen Fällen gilt also

$$\lim_{n \rightarrow \infty} \int_c^d f(x_n, y) dy = \int_c^d f(x, y) dy \quad (11.55)$$

für jede Folge $x_n \rightarrow x$.

Damit (11.55) gilt, genügt es aber, dass f auf seinem zweidimensionalen Definitionsbereich $I \times [c, d]$ stetig ist (siehe Kapitel 2), das heißt,

$$\lim_{n \rightarrow \infty} f(x_n, y_n) = f(x, y) \quad (11.56)$$

gilt für jede Folge (x_n, y_n) , welche gegen (x, y) konvergiert. Die Funktion in (11.54) erfüllt (11.56) nicht, es ist

$$f(x, x) = \frac{1}{x}, \quad \text{aber } f(0, 0) = 0.$$

Satz 11.10 Sei $I \subset \mathbb{R}$ Intervall, $f : I \times [c, d] \rightarrow \mathbb{R}$ stetig. Dann wird durch

$$F(x) = \int_c^d f(x, y) dy \quad (11.57)$$

eine stetige Funktion $F : I \rightarrow \mathbb{R}$ definiert, das heißt, aus $x_n \rightarrow x$ folgt $F(x_n) \rightarrow F(x)$. \square

Satz 11.11 Die Aussage von Satz 11.10 gilt auch, wenn f die folgende Voraussetzungen erfüllt:

(i) f ist stetig bezüglich x , das heißt, es gilt

$$\lim_{n \rightarrow \infty} f(x_n, y) = f(x, y), \quad (11.58)$$

für alle $x \in I$, $y \in [c, d]$ und alle Folgen $x_n \rightarrow x$.

(ii) Es gibt eine integrierbare Funktion $h : [c, d] \rightarrow \mathbb{R}$, so dass

$$|f(x, y)| \leq h(y) \quad (11.59)$$

gilt für alle $x \in I$, $y \in [c, d]$.

Unter diesen Voraussetzungen gilt die Aussage von Satz 11.10 auch dann, wenn das Integral (11.57) ein uneigentliches Integral ist. \square

Diese beiden Sätze beweisen wir hier nicht.

Als Beispiel betrachten wir die **Laplace-Transformierte**

$$(\mathcal{L}g)(x) = \int_0^{\infty} g(y)e^{-xy} dy \quad (11.60)$$

einer Funktion $g : [0, \infty) \rightarrow \mathbb{R}$. Falls beispielsweise $|g|$ durch eine Konstante C beschränkt ist, so gilt auf jedem Intervall $I = [a, b]$ mit $a > 0$ für $x \in [a, b]$

$$|g(y)e^{-xy}| \leq Ce^{-ay} =: h(y)$$

für alle $y \geq 0$. Da h uneigentlich integrierbar ist auf $[0, \infty)$, ist also die Laplace-Transformierte $\mathcal{L}g$ für solche Funktionen g stetig auf jedem Intervall $[a, b]$ mit $a > 0$, und damit auf $(0, \infty)$.

Ein anderes Beispiel ist die **Gamma-Funktion**

$$\Gamma(x) = \int_0^{\infty} e^{-y}y^{x-1} dy, \quad (11.61)$$

welche definiert und stetig ist für $x > 0$.

Zur Klärung der Frage (11.53) nach der Ableitung $F'(x)$ von

$$F(x) = \int_c^d f(x, y) dy. \quad (11.62)$$

haben wir folgenden Satz.

Satz 11.12 Sei $I \subset \mathbb{R}$ Intervall, seien $f : I \times [c, d] \rightarrow \mathbb{R}$ und $\partial_x f : I \times [c, d] \rightarrow \mathbb{R}$ stetig. Dann ist $F : I \rightarrow \mathbb{R}$ differenzierbar, und

$$F'(x) = \int_c^d \partial_x f(x, y) dy. \quad (11.63)$$

\square

Auch in diesem Fall ist die Stetigkeit von $\partial_x f$ nur bezüglich x erforderlich, solange in Analogie zu (11.59) eine Abschätzung

$$|\partial_x f(x, y)| \leq h(y) \quad (11.64)$$

mit einer integrierbaren Funktion h gilt, und der Fall des uneigentlichen Integrals ist damit ebenfalls eingeschlossen. Auf diese Weise erhalten wir beispielsweise die Ableitung der Gammafunktion,

$$\Gamma'(x) = \int_0^\infty e^{-y} y^{x-1} \ln(y) dy, \quad (11.65)$$

sowie unter geeigneten Voraussetzungen an g die Ableitung der Laplace-Transformierten

$$(\mathcal{L}g)'(x) = \int_0^\infty -yg(y)e^{-xy} dy = -(\mathcal{L}k)(x), \quad \text{wobei } k(y) = yg(y). \quad (11.66)$$

Wir wollen nun Funktionen differenzieren, bei denen die unabhängige Variable x sowohl in den Integralgrenzen als auch im Integranden auftritt,

$$F(x) = \int_{c(x)}^{d(x)} f(x, y) dy. \quad (11.67)$$

Falls f, c, d stetig differenzierbare Funktionen sind, erhalten wir die Ableitung von F , indem wir Satz 11.12 mit der Kettenregel kombinieren. Wir setzen

$$\tilde{F}(x, a, b) = \int_a^b f(x, y) dy,$$

dann ist

$$F(x) = \tilde{F}(x, c(x), d(x)).$$

Aus der Kettenregel folgt

$$F' = \partial_x \tilde{F} + \partial_a \tilde{F} \cdot c' + \partial_b \tilde{F} \cdot d',$$

genauer

$$F'(x) = \partial_x \tilde{F}(x, c(x), d(x)) + \partial_a \tilde{F}(x, c(x), d(x)) \cdot c'(x) + \partial_b \tilde{F}(x, c(x), d(x)) \cdot d'(x).$$

Es ist nach (11.63) und dem Hauptsatz der Differential- und Integralrechnung

$$\partial_x \tilde{F}(x, a, b) = \int_a^b \partial_x f(x, y) dy, \quad \partial_a \tilde{F}(x, a, b) = -f(x, a), \quad \partial_b \tilde{F}(x, a, b) = f(x, b).$$

Insgesamt ergibt sich die **Formel von Leibniz**

$$F'(x) = \int_{c(x)}^{d(x)} \partial_x f(x, y) dy + f(x, d(x))d'(x) - f(x, c(x))c'(x). \quad (11.68)$$

Als Beispiel betrachten wir die Funktion

$$F(x) = \int_{x^2}^{3-x} \sin(xy) dy.$$

Es ist

$$F'(x) = \int_{x^2}^{3-x} y \cos(xy) dy + \sin(x(3-x)) \cdot (-1) - \sin(x^3) \cdot 2x.$$

Vektorfelder. Ein Vektorfeld wird beschrieben durch eine Funktion (oder Abbildung)

$$f : \Omega \rightarrow \mathbb{R}^m, \quad \Omega \subset \mathbb{R}^d. \quad (11.69)$$

Es besteht aus m Skalarfeldern, den Komponentenfunktionen

$$f_i : \Omega \rightarrow \mathbb{R}, \quad 1 \leq i \leq m. \quad (11.70)$$

Ist beispielsweise $g : \Omega \rightarrow \mathbb{R}$ ein Skalarfeld, so ist $\nabla g : \Omega \rightarrow \mathbb{R}^d$ ein Vektorfeld.

Ein Vektorfeld f können wir differenzieren, indem wir seine Komponentenfunktionen f_1, \dots, f_m jeweils für sich genommen differenzieren. An jedem Punkt $x \in \Omega$ (wir setzen Ω als offen voraus) erhalten wir partielle Ableitungen

$$\partial_j f_i(x), \quad 1 \leq i \leq m, 1 \leq j \leq d. \quad (11.71)$$

Wir können sie zu einer $m \times d$ -Matrix zusammenfassen,

$$J_f(x) = \begin{pmatrix} \partial_1 f_1(x) & \cdots & \partial_d f_1(x) \\ \partial_1 f_2(x) & \cdots & \partial_d f_2(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f_m(x) & \cdots & \partial_d f_m(x) \end{pmatrix} \in \mathbb{R}^{(m,d)}. \quad (11.72)$$

Diese Matrix heißt **Jacobi-Matrix** oder **Funktionalmatrix** des Vektorfelds f , ihre Zeilen werden von den (transponierten) Gradienten der Komponentenfunktionen f_i gebildet. Ist $m = d$, so ist sie quadratisch.

Als Beispiel betrachten wir $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$f(x) = (f_1(x_1, x_2), f_2(x_1, x_2)) = (x_1^2 + x_2, x_1 \sin(x_2)).$$

Es ist

$$J_f(x) = \begin{pmatrix} \partial_1 f_1(x_1, x_2) & \partial_2 f_1(x_1, x_2) \\ \partial_1 f_2(x_1, x_2) & \partial_2 f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} 2x_1 & 1 \\ \sin(x_2) & x_1 \cos(x_2) \end{pmatrix}.$$

Im Punkt $x = (2, \pi)$ gilt

$$J_f(2, \pi) = \begin{pmatrix} 4 & 1 \\ 0 & -2 \end{pmatrix}.$$

Koordinatentransformationen sind ebenfalls Vektorfelder. Affin lineare Koordinatentransformationen haben die Form

$$f(x) = Ax + b, \quad A \in \mathbb{R}^{(d,d)}, b \in \mathbb{R}^d, \quad f : \mathbb{R}^d \rightarrow \mathbb{R}^d. \quad (11.73)$$

Es ist

$$f_i(x) = \sum_{j=1}^d a_{ij} x_j, \quad \partial_j f_i(x) = a_{ij},$$

also

$$J_f(x) = A,$$

die Jacobi-Matrix hängt also nicht von x ab und ist gleich der ursprünglichen Matrix A . Ein Beispiel für eine nichtlineare Koordinatentransformation sind die **Polarkoordinaten**, beschrieben durch

$$f(r, \varphi) = (r \cos \varphi, r \sin \varphi), \quad f : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2. \quad (11.74)$$

Die zugehörige Funktionalmatrix ist

$$J_f(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}. \quad (11.75)$$

Ein weiteres Beispiel sind die **Zylinderkoordinaten**, beschrieben durch

$$f(r, \varphi, z) = (r \cos \varphi, r \sin \varphi, z), \quad f : (0, \infty) \times (0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3. \quad (11.76)$$

Die zugehörige Funktionalmatrix ist

$$J_f(r, \varphi, z) = \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (11.77)$$

Ebenfalls eine Rolle spielen die **Kugelkoordinaten**, beschrieben durch

$$f(r, \theta, \varphi) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta), \quad f : (0, \infty) \times (0, \pi) \times (0, 2\pi) \rightarrow \mathbb{R}^3. \quad (11.78)$$

Die zugehörige Funktionalmatrix ist

$$J_f(r, \theta, \varphi) = \begin{pmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}. \quad (11.79)$$

Approximationseigenschaften der Ableitung bei Vektorfeldern. Sie übertragen sich direkt vom skalaren Fall. Für die Komponenten f_i eines Vektorfelds $f : \Omega \rightarrow \mathbb{R}^m$ haben wir im vorigen Kapitel erhalten (wir nennen den Entwicklungspunkt nun x statt a)

$$f_i(x + h) = f_i(x) + \nabla f_i(x)^T h + o(\|h\|), \quad 1 \leq i \leq m. \quad (11.80)$$

In Matrix-Vektor-Notation können wir diese Gleichungen zusammenfassen zu einer Gleichung von Spaltenvektoren

$$f(x + h) = f(x) + J_f(x)h + o(\|h\|). \quad (11.81)$$

Die Differenz $f(x + h) - f(x)$ wird also in der Nähe von x durch die lineare Abbildung

$$T_x : \mathbb{R}^d \rightarrow \mathbb{R}^m, \quad T_x(h) = J_f(x)h, \quad (11.82)$$

approximiert. Ein Vektorfeld f , für das (11.81) oder äquivalent

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

gilt, heißt **total differenzierbar** oder **Fréchet-differenzierbar**. Die lineare Abbildung T_x nennt man das **totale Differential** oder die **Fréchet-Ableitung** von f im Punkt x ,

ihre Matrixdarstellung in den kanonischen Koordinaten ist gemäß (11.82) gerade durch die Funktionalmatrix gegeben.

Die **Kettenregel** für Vektorfelder ist am übersichtlichsten, wenn man sie für das totale Differential hinschreibt. Seien

$$f : \Omega \rightarrow \mathbb{R}^m, \quad g : \mathbb{R}^m \rightarrow \mathbb{R}^l$$

total differenzierbar, seien

$$T_x : \mathbb{R}^d \rightarrow \mathbb{R}^m, \quad S_{f(x)} : \mathbb{R}^m \rightarrow \mathbb{R}^l$$

die totalen Differentiale in den Punkten x bzw. $f(x)$. Dann ist (was wir nicht beweisen wollen) das totale Differential der zusammengesetzten Funktion

$$g \circ f : \mathbb{R}^d \rightarrow \mathbb{R}^l. \quad (11.83)$$

gegeben durch

$$S_{f(x)} \circ T_x : \mathbb{R}^d \rightarrow \mathbb{R}^l, \quad (11.84)$$

entsprechend erhalten wir die Funktionalmatrix von $g \circ f$ als Produkt der Funktionalmatrizen von f und g ,

$$J_{g \circ f}(x) = J_g(f(x)) \cdot J_f(x). \quad (11.85)$$

In (11.85) sind die Formeln zur Berechnung aller partieller Ableitungen aller Komponentenfunktionen von $g \circ f$ zusammengefasst.

Als Beispiel für die Kettenregel betrachten wir, wie sich partielle Ableitungen transformieren, wenn wir von kartesischen zu Polarkoordinaten übergehen. Sei etwa

$$u(x, y) = x^2 + 2y^3, \quad u : \mathbb{R}^2 \rightarrow \mathbb{R}. \quad (11.86)$$

Wir wollen die Funktion u und ihre Ableitungen “in Polarkoordinaten betrachten”, das heißt, wir haben zu untersuchen die Funktion

$$\tilde{u}(r, \varphi) = u(r \cos \varphi, r \sin \varphi). \quad (11.87)$$

Es ist also

$$\tilde{u} = u \circ f, \quad f(r, \varphi) = (r \cos \varphi, r \sin \varphi). \quad (11.88)$$

Wir wenden die Kettenregel (11.85) an mit $g = u$, $d = m = 2$, $l = 1$, dann ist

$$(\partial_r \tilde{u} \quad \partial_\varphi \tilde{u}) = (\partial_x u \quad \partial_y u) \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}. \quad (11.89)$$

Hierbei sind die Matrixelemente abgekürzt notiert, es sind gemeint die partiellen Ableitungen an den entsprechenden Punkten,

$$\partial_r \tilde{u}(r, \varphi), \quad \partial_x u(r \cos \varphi, r \sin \varphi), \quad (11.90)$$

und Entsprechendes für $\partial_\varphi \tilde{u}$, $\partial_y u$. Für die Beispielfunktion (11.86) gilt

$$\partial_x u(x, y) = 2x, \quad \partial_y u(x, y) = 6y^2.$$

Aus (11.89) ergibt sich also

$$\partial_r \tilde{u}(r, \varphi) = \begin{pmatrix} 2r \cos \varphi & 6r^2 \sin^2 \varphi \end{pmatrix} \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} = 2r \cos^2 \varphi + 6r^2 \sin^3 \varphi, \quad (11.91)$$

und analog

$$\partial_\varphi \tilde{u}(r, \varphi) = -2r^2 \sin \varphi \cos \varphi + 6r^3 \sin^2 \varphi \cos \varphi. \quad (11.92)$$

Dieses Beispiel illustriert das generelle Rechenschema bei der Kettenregel. In diesem speziellen Fall kann man (11.91) und (11.92) direkt erhalten, indem man

$$\tilde{u}(r, \varphi) = r^2 \cos^2 \varphi + 2r^3 \sin^3 \varphi$$

nach r und φ partiell differenziert.

Beim Rechnen unterscheidet man oft nicht zwischen u und \tilde{u} , man schreibt dann $u(r, \varphi) = u(r \cos \varphi, r \sin \varphi)$, weil man ja "bis auf eine Koordinatentransformation dieselbe Funktion meint". Die Transformationsformel (11.89) für die Ableitungen der Polarkoordinaten kann man dann auch aufgeschrieben finden in der Form

$$\begin{aligned} \frac{\partial u}{\partial r} &= \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial r} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial r} = \cos \varphi \frac{\partial u}{\partial x} + \sin \varphi \frac{\partial u}{\partial y} \\ \frac{\partial u}{\partial \varphi} &= \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial \varphi} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial \varphi} = -r \sin \varphi \frac{\partial u}{\partial x} + r \cos \varphi \frac{\partial u}{\partial y}. \end{aligned}$$

Man muss dann aber mehr aufpassen, dass man nichts durcheinanderbringt.

12 Gewöhnliche Differentialgleichungen

Eine Differentialgleichung ist eine Gleichung, in der Ableitungen von Funktionen auftreten. Das wohl bekannteste Beispiel ist die Anwendung des Newtonschen Kraftgesetzes

$$F = mb \quad (12.1)$$

(Kraft gleich Masse mal Beschleunigung) auf eine in einer festen Richtung schwingende Feder. Bezeichnet $y(t)$ die Position des Endpunkts der Feder zum Zeitpunkt t , so lautet das Hookesche Gesetz

$$F = -ky(t), \quad (12.2)$$

wobei k eine Materialkonstante (die Federkonstante) ist. Da die Beschleunigung zum Zeitpunkt t durch die zweite Ableitung $y''(t)$ gegeben ist, erhält das Kraftgesetz die Form der Differentialgleichung

$$my''(t) = -ky(t). \quad (12.3)$$

Gesucht ist eine **unbekannte Funktion**, nämlich die Position y als Funktion der Zeit t , nicht nur ein unbekannter Vektor wie etwa bei der Lösung eines linearen Gleichungssystems $Ax = b$. Die Variable t heißt **unabhängige Variable**, die Variable y **abhängige Variable**. Bringt man alle Ausdrücke auf die linke Seite, so wird (12.3) zu

$$my''(t) + ky(t) = 0. \quad (12.4)$$

Will man ausdrücken, über welche Differentialgleichung man redet, so schreibt man (12.4) normalerweise in der Form

$$my'' + ky = 0. \quad (12.5)$$

Unter der **Lösung** einer Differentialgleichung versteht man eine Funktion, die auf einem geeigneten Definitionsbereich I definiert und dort hinreichend oft (im Beispiel (12.5) zweimal) differenzierbar ist, und die in jedem Punkt von I (in (12.4) heißt er t) von I die Differentialgleichung erfüllt. Die Bestimmung eines geeigneten bzw. eines möglichst großen Definitionsbereiches ist Bestandteil des Problems. Beispielsweise hat die Differentialgleichung

$$y' = -y^2 \quad (12.6)$$

die Funktionen

$$y(x) = \frac{1}{x - c} \quad (12.7)$$

als Lösung, wobei $c \in \mathbb{R}$ eine beliebige Konstante ist. Der Definitionsbereich dieser Funktion zerfällt in zwei Teile,

$$I = (-\infty, c) \cup (c, \infty). \quad (12.8)$$

Jede Lösung von (12.7) hat als maximalen Definitionsbereich entweder $(-\infty, c)$ oder (c, ∞) .

In den obigen Differentialgleichungen ist die unabhängige Variable ein Skalar (t bzw. x). Solche Differentialgleichungen heißen **gewöhnliche Differentialgleichungen**. Bei der Modellierung realer Probleme sind in vielen Fällen die gesuchten Funktionen vom Ort $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ und von der Zeit t abhängig, so dass (je nach Modellbildung) bis zu 4 unabhängige Variable gleichzeitig auftreten können. Als Ableitungen treten dann Ableitungen nach den einzelnen Komponenten x_i des Orts bzw. nach der Zeit t separat auf.

Die gesuchten Funktionen erhalten oft auch mehrere Komponenten. Solche Differentialgleichungen heißen **partielle Differentialgleichungen**. Sie gehen oft auf die allgemeinen Bilanzgleichungen (Erhaltungsgleichungen) für Masse, Impuls und Energie in der Kontinuumsmechanik zurück. Beispiele sind

- die Diffusionsgleichung zur Beschreibung der Wärmeausbreitung bzw. der Diffusion von Substanzen in einem umgebenden Medium,
- die Navier-Stokes-Gleichung zur Beschreibung von Strömungen,
- die Maxwellgleichungen zur Beschreibung elektromagnetischer Felder,
- die Naviergleichungen aus der Mechanik elastischer Körper.
- die Drift-Diffusions-Gleichung für Dichte und Stromdichte von Elektronen (Halbleiterbauteile oberhalb 200 Nanometer)
- Varianten der Fokker-Planck- und der Boltzmann-Gleichung (Bereich 10 – 100 Nanometer)

Wir behandeln in diesem Kapitel gewöhnliche Differentialgleichungen. Am einfachsten ist die Situation, wenn die gesuchte Funktion (in Form einer ihrer Ableitungen) nur an einer Stelle der Differentialgleichung auftritt. Wir betrachten das Beispiel

$$y'' = -g, \quad (12.9)$$

welche die Bewegung eines fallenden Körpers der Masse 1 beschreibt. Hier ist y die Höhe des Körpers, die unabhängige Variable ist die Zeit t , und g ist die Erdbeschleunigung. Im Ansatz (12.9) wird die Luftreibung ignoriert. Wir nehmen nun an, dass g konstant ist (und vernachlässigen also die Tatsache, dass sie sowohl von der Höhe als auch von den geographischen Koordinaten abhängt – letztere spielen eine Rolle, da die Erde nur näherungsweise eine Kugel ist). In diesem Fall können wir (12.9) durch Integration direkt lösen. Die Ableitung y' muss nämlich eine Stammfunktion von $-g$ sein, also

$$y'(t) = -gt + c, \quad (12.10)$$

und weiter

$$y(t) = -\frac{g}{2}t^2 + ct + d, \quad (12.11)$$

mit den beiden Integrationskonstanten c und d . Diese Freiheitsgrade sind erst festgelegt, wenn man zwei weitere Bedingungen kennt. Sind etwa Höhe und Geschwindigkeit zum gleichen Zeitpunkt, etwa $t = 1$, gegeben,

$$y(1) = y_0, \quad y'(1) = v_0, \quad (12.12)$$

so nennt man sie **Anfangsbedingungen** und die Gleichungen (12.9) und (12.12) zusammengekommen ein **Anfangswertproblem**. In diesem Fall ergibt sich aus (12.10) und (12.11)

$$y_0 = -\frac{g}{2} + c + d, \quad v_0 = -g + c, \quad (12.13)$$

also das lineare Gleichungssystem

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} y_0 + \frac{g}{2} \\ v_0 + g \end{pmatrix} \quad (12.14)$$

zur Bestimmung von c und d . Sind die Höhen zu zwei verschiedenen Zeitpunkten gegeben, etwa $t = 1$ und $t = 2$, etwa

$$y(1) = y_1, \quad y(2) = y_2, \quad (12.15)$$

so nennt man diese **Randbedingungen** und die Gleichungen (12.9) und (12.15) zusammengefasst ein **Randwertproblem**. Aus (12.11) erhalten wir

$$y_1 = -\frac{g}{2} + c + d, \quad y_2 = -\frac{g}{2} \cdot 4 + 2c + d. \quad (12.16)$$

also ebenfalls ein lineares Gleichungssystem,

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} y_1 + \frac{g}{2} \\ y_2 + 2g \end{pmatrix}. \quad (12.17)$$

Sind c und d bestimmt, so können wir Höhe und Geschwindigkeit zu beliebigen Zeiten t durch Einsetzen in (12.10) und (12.11) bestimmen.

Hat die Differentialgleichung die Form

$$y' = f(x), \quad (12.18)$$

wobei f eine bekannte Funktion ist und die unabhängige Variable mit x bezeichnet wird, so sind die Lösungen gerade die Stammfunktionen von f . Ob sich diese Lösung mit einer expliziten Formel

$$y(x) = \dots$$

angeben lässt, hängt ganz von der Funktion f ab. Erst recht gilt dies für die **allgemeine Differentialgleichung erster Ordnung**

$$y' = f(x, y), \quad f : D \rightarrow \mathbb{R} \quad \text{gegeben}, \quad (12.19)$$

wobei $D \subset \mathbb{R}^2$ ein geeignetes Definitionsgebiet der **rechten Seite** f ist. Die Lösungen von (12.19) sind Funktionen y , welche auf einem geeigneten Definitionsgebiet $I \subset \mathbb{R}$ definiert sind und für die

$$y'(x) = f(x, y(x)), \quad x \in I, \quad (12.20)$$

gilt. (Damit (12.20) gilt, muss insbesondere auch $(x, y(x)) \in D$ gelten.) Wir können uns die Lösung als eine Funktion vorstellen, deren Steigung an jedem Punkt $(x, y(x)) \in \mathbb{R}^2$ ihres Graphen mit dem Funktionswert von f an diesem Punkt übereinstimmt.

Mindestens genauso wichtig wie die formelmäßige Lösung einer Differentialgleichung ist ihre **numerische Lösung**. Wir erläutern dies am Beispiel des **Eulerverfahrens**. Nehmen wir an, wir wollen (12.20) zum Anfangswert

$$y(x_0) = y_0 \quad (12.21)$$

mit gegebenem $x_0, y_0 \in \mathbb{R}$ lösen. Wir wählen ein $x_1 > x_0$ und definieren einen Näherungswert y_1 für den Wert $y(x_1)$ der wahren Lösung, indem wir setzen

$$y_1 = y_0 + h_1 f(x_0, y_0), \quad h_1 = x_1 - x_0. \quad (12.22)$$

Der Wert y_1 ist also der Funktionswert der Geraden durch (x_0, y_0) an der Stelle x_1 . Die Zahl h_1 heißt **Schrittweite**. Dieses Verfahren kann man fortsetzen, indem man (x_1, y_1) als neuen Anfangswert betrachtet. Es ergibt sich die Vorschrift

$$y_{k+1} = y_k + h_{k+1}f(x_k, y_k), \quad h_{k+1} = x_{k+1} - x_k, \quad k = 0, 1, 2, \dots \quad (12.23)$$

Ein Spezialfall ist die Wahl einer konstanten (“**äquidistanten**”) Schrittweite $h > 0$, also

$$x_k = x_0 + kh. \quad (12.24)$$

Verbindet man die Näherungswerte (x_k, y_k) durch Strecken, so erhält man eine Näherungsfunktion y_h mit

$$y_h(x) = y_k + (x - x_k)f(x_k, y_k), \quad x \in [x_k, x_{k+1}]. \quad (12.25)$$

Diese ist stückweise linear, man nennt das Eulerverfahren daher auch **Polygonzugverfahren** oder **Linienmethode**.

Das Eulerverfahren ist einfach zu verstehen und zu programmieren. Es gibt aber viele andere Verfahren, die bei vorgegebener Genauigkeit der Näherungslösung einen erheblich geringeren Rechenaufwand als das Eulerverfahren benötigen (da sie mit größeren Schrittweiten h_k als das Eulerverfahren arbeiten können).

Mit der Konstruktion und den Eigenschaften solcher **Diskretisierungsverfahren** befasst man sich in der Numerischen Mathematik und im Wissenschaftlichen Rechnen.

Trennung der Variablen. Wir nehmen an, die Differentialgleichung hat die Form

$$y' = f(x)g(y), \quad (12.26)$$

wobei f und g gegebene Funktionen sind. Zunächst gilt: Ist a eine Nullstelle von g , also $g(a) = 0$, so ist die konstante Funktion $y(x) = a$ eine Lösung von (12.26). Wir betrachten nun Bereiche, in denen g von Null verschieden ist. Zur Bestimmung einer Lösung $x \mapsto y(x)$ leiten wir eine Gleichung her, in der nur noch die Funktion y , aber nicht mehr deren Ableitung auftritt. Zu diesem Zweck nehmen wir an, dass wir eine Stammfunktion F von f und eine Stammfunktion G von $1/g$ finden können, also

$$F' = f, \quad G' = \frac{1}{g}.$$

Wir betrachten hilfsweise die Funktion (y soll eine Lösung sein)

$$z(x) = G(y(x)) - F(x). \quad (12.27)$$

Es gilt

$$z'(x) = G'(y(x))y'(x) - F'(x) = \frac{1}{g(y(x))}f(x)g(y(x)) - f(x) = 0,$$

also ist $z(x) = c$ konstant für ein $c \in \mathbb{R}$, und aus (12.27) folgt

$$G(y(x)) = F(x) + c. \quad (12.28)$$

Falls G eine Umkehrfunktion G^{-1} hat, gilt also

$$y(x) = G^{-1}(F(x) + c). \quad (12.29)$$

Diese Schritte lassen sich anhand des folgenden **Rechenschemas** einfacher merken und ausführen. Wir rechnen folgendermaßen:

$$\begin{aligned} \frac{dy}{dx} &= f(x)g(y) \\ \frac{dy}{g(y)} &= f(x) dx \\ \int \frac{dy}{g(y)} &= \int f(x) dx \\ G(y) &= F(x) + c && \text{(Bilden von Stammfunktionen)} \\ y &= G^{-1}(F(x) + c) && \text{(Auflösen der vorigen Zeile nach } y) \end{aligned}$$

Ein Beispiel: Wir wollen lösen

$$y' = y^2. \tag{12.30}$$

Wir rechnen

$$\begin{aligned} \frac{dy}{dx} &= y^2 \\ \frac{dy}{y^2} &= 1 dx \\ \int \frac{dy}{y^2} &= \int 1 dx \\ -\frac{1}{y} &= x + c \\ y &= -\frac{1}{x + c}. \end{aligned}$$

Die Funktion

$$y(x) = -\frac{1}{x + c} \tag{12.31}$$

ist für beliebiges $c \in \mathbb{R}$ eine Lösung von (12.30), wie man durch Differenzieren nachprüfen kann. Hat man zusätzlich einen bestimmten Anfangswert zu erfüllen, so wird c durch ihn festgelegt,

$$y(0) = 1 \quad \text{führt auf} \quad 1 = -\frac{1}{0 + c}, \quad c = -1.$$

Das ist ein Rechenschema. Man kann sich natürlich die Frage stellen: Was bedeuten in dieser Rechnung “ dy ” und “ dx ” für sich genommen? Die Antwort dieser Vorlesung ist: **Nichts**. Im Zeitraum vor 1900 hat man andere Antworten gesucht, Stichwort “unendlich kleine Größen”. Das Ergebnis war schließlich die Definition des Grenzwertbegriffs als Grundlage für die Differential- und Integralrechnung, wie wir das in den vorigen Kapiteln gesehen haben.

Es soll aber nicht verschwiegen werden, dass – in anderen mathematischen Situationen, etwa in der Theorie der Differentialformen und auch in der fortgeschrittenen Integrations- und Integralrechnung – die Ausdrücke “ dy ” und “ dx ” mit Bedeutung belegt werden.

Differentialgleichungen, die nicht die Form (12.26) haben, lassen sich unter Umständen durch eine geeignete **Substitution** darauf zurückführen. Ein Beispiel ist

$$y' = f(ax + by + d), \tag{12.32}$$

wobei $a, b, d \in \mathbb{R}$ Konstante sind. Wir führen als neue Variable ein

$$u = ax + by + d. \quad (12.33)$$

Gemeint sind Funktionen u und y , die vermittels

$$u(x) = ax + by(x) + d \quad (12.34)$$

ineinander übergeführt werden. Es ist dann

$$u' = a + by' = a + bf(u). \quad (12.35)$$

Wir lösen also

$$u' = a + bf(u) \quad (12.36)$$

durch Trennung der Variablen und berechnen anschließend die Lösung von (12.32) aus (12.34). Beispiel:

$$y' = (x - y)^2 \quad (12.37)$$

führt auf

$$u = x - y, \quad u' = 1 - u^2,$$

Trennung der Variablen ergibt

$$\int \frac{1}{1 - u^2} du = \int 1 dx = x + c.$$

Stammfunktion von $1/(1-u^2)$ ist die Umkehrfunktion \tanh^{-1} des **Tangens hyperbolicus**

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (12.38)$$

Es ist also

$$\tanh^{-1}(u) = x + c,$$

also

$$u(x) = \tanh(x + c),$$

und damit erhalten wir

$$y(x) = x - u(x) = x - \tanh(x + c) \quad (12.39)$$

als Lösung von (12.37).

Die lineare Differentialgleichung 1. Ordnung. Die **homogene** lineare Differentialgleichung 1. Ordnung hat die Form

$$y' + a(x)y = 0, \quad (12.40)$$

wobei a eine gegebene Funktion ist. Man kann sie ebenfalls durch Trennung der Variablen lösen, die allgemeine Lösung ergibt sich als

$$y(x) = ce^{-A(x)}, \quad c \in \mathbb{R}, \quad (12.41)$$

wobei A eine Stammfunktion von a ist,

$$A' = a. \quad (12.42)$$

Dass y aus (12.41) tatsächlich eine Lösung ist, ergibt sich aus

$$y'(x) = ce^{-A(x)} \cdot (-A'(x)) = -a(x)y(x).$$

Beispiele: Für

$$y' + 2y = 0$$

ist

$$a(x) = 2, \quad A(x) = 2x, \quad y(x) = ce^{-2x}.$$

Für

$$y' + 2xy = 0$$

ist

$$a(x) = 2x, \quad A(x) = x^2, \quad y(x) = ce^{-x^2}.$$

Die zu (12.40) **inhomogene** Gleichung lautet

$$y' + a(x)y = f(x), \quad (12.43)$$

wobei außer a auch f eine gegebene Funktion ist. Sei A wieder eine Stammfunktion von a . Ist y eine Lösung von (12.43), und setzen wir

$$z(x) = y(x)e^{A(x)}, \quad (12.44)$$

so gilt

$$z'(x) = y'(x)e^{A(x)} + y(x)e^{A(x)}A'(x) = (y'(x) + a(x)y(x))e^{A(x)} = f(x)e^{A(x)}.$$

Sei nun $x_0 \in \mathbb{R}$ fest gewählt. Integration führt auf

$$z(x) = z(x_0) + \int_{x_0}^x z'(s) ds = z(x_0) + \int_{x_0}^x e^{A(s)} f(s) ds,$$

und Multiplikation mit $e^{-A(x)}$ ergibt

$$y(x) = e^{-A(x)} \left(e^{A(x_0)} y(x_0) + \int_{x_0}^x e^{A(s)} f(s) ds \right). \quad (12.45)$$

Die allgemeine Lösung von (12.43) lässt sich schreiben als

$$\begin{aligned} y(x) &= e^{-A(x)} \left(c + \int_{x_0}^x e^{A(s)} f(s) ds \right) \\ &= ce^{-A(x)} + \int_{x_0}^x e^{A(s)-A(x)} f(s) ds. \end{aligned} \quad (12.46)$$

Setzt man als Anfangswert

$$y(x_0) = y_0 \quad (12.47)$$

fest, und wählt man außerdem A so, dass $A(x_0) = 0$, so ergibt sich aus (12.45)

$$y(x) = e^{-A(x)} \left(y_0 + \int_{x_0}^x e^{A(s)} f(s) ds \right). \quad (12.48)$$

Vergleicht man diese Formel mit der Lösungsformel (12.41) für die homogene Gleichung, so stellt man fest, dass die Konstante c in (12.41) durch eine nichtkonstante Funktion (die, die in (12.48) in der großen Klammer steht) ersetzt wurde. Dementsprechend wird Formel (12.48) oft mit dem Namen **Variation der Konstanten** belegt.

Als Beispiel betrachten wir

$$y' + 2y = \sin x.$$

Die allgemeine Lösung ist

$$y(x) = ce^{-2x} + \int_{x_0}^x e^{2(x-s)} \sin(s) ds,$$

und die Lösung zum Anfangswert $y(0) = 3$

$$y(x) = 3e^{-2x} + \int_0^x e^{2(x-s)} \sin(s) ds.$$

Die lineare Differentialgleichung 2. Ordnung. Wir betrachten

$$y'' + 2a(x)y' + b(x)y = 0, \quad (12.49)$$

wobei a, b gegebene Funktionen sind. Seien $y_1, y_2 : I \rightarrow \mathbb{R}$ zwei Lösungen von (12.49) auf einem Intervall I . Dann ist jede durch

$$y(x) = c_1 y_1(x) + c_2 y_2(x), \quad c_1, c_2 \in \mathbb{R}, \quad (12.50)$$

definierte Funktion ebenfalls eine Lösung von (12.49), da

$$\begin{aligned} y''(x) + 2a(x)y'(x) + b(x)y(x) &= \\ &= c_1 y_1''(x) + c_2 y_2''(x) + 2a(x)(c_1 y_1'(x) + c_2 y_2'(x)) + b(x)(c_1 y_1(x) + c_2 y_2(x)) \\ &= c_1 (y_1''(x) + 2a(x)y_1'(x) + b(x)y_1(x)) + c_2 (y_2''(x) + 2a(x)y_2'(x) + b(x)y_2(x)) \\ &= c_1 \cdot 0 + c_2 \cdot 0 = 0. \end{aligned}$$

Jede Linearkombination zweier (und damit auch endlich vieler) Lösungen ist also wieder eine Lösung. Diese Eigenschaft bezeichnet man als **Superpositionsprinzip**. Es gilt für **homogene** lineare Gleichungen (das heißt lineare Gleichungen, bei denen die rechte Seite gleich 0 ist).

Da es lineare Differentialgleichungen in ganz unterschiedlicher Form gibt, ist es zweckmäßig, sie von einem allgemeineren Standpunkt her zu betrachten. Ist $y : I \rightarrow \mathbb{R}$ eine Funktion (hier: zweimal differenzierbar), so definieren wir eine neue Funktion $L[y] : I \rightarrow \mathbb{R}$ durch die linke Seite von (12.49),

$$(L[y])(x) = y''(x) + 2a(x)y'(x) + b(x)y(x), \quad (12.51)$$

wobei a und b wie oben gegebene Funktionen sind. Hierdurch wird eine Abbildung L definiert, welche Funktionen y auf andere Funktionen $L[y]$ abbildet. L heißt **Differentialoperator**. Die Differentialgleichung (12.49) lässt sich nun kurz schreiben als

$$L[y] = 0. \quad (12.52)$$

Gleichung (12.52) ist eine Gleichung zwischen zwei Funktionen, gemeint ist also, dass die durch (12.51) definierte Funktion, die auf der linken Seite von (12.52) steht, gleich der Nullfunktion ist. Die Linearität der Differentialgleichung drückt sich nun in der Beziehung

$$L[c_1y_1 + c_2y_2] = c_1L[y_1] + c_2L[y_2], \quad (12.53)$$

aus, welche für alle zweimal differenzierbaren Funktionen y_1, y_2 und alle $c_1, c_2 \in \mathbb{R}$ gilt.

L ist eine lineare Abbildung $L : X \rightarrow Y$ im Sinne der Linearen Algebra, sobald man geeignete Vektorräume X und Y definiert hat. Das kann man tun, es handelt sich dabei um Vektorräume, deren Elemente Funktionen sind, solche Vektorräume nennt man **Funktionsräume**.

Sind y_1, y_2 Lösungen von (12.49), so ergibt sich das Superpositionsprinzip aus der Rechnung

$$L[c_1y_1 + c_2y_2] = c_1L[y_1] + c_2L[y_2] = c_1 \cdot 0 + c_2 \cdot 0 = 0.$$

Wir betrachten nun den Spezialfall der **konstanten Koeffizienten**

$$y'' + 2ay' + by = 0, \quad a, b \in \mathbb{R}. \quad (12.54)$$

Als Lösungsansatz wählen wir

$$y(x) = e^{\lambda x}. \quad (12.55)$$

Setzen wir (12.55) in (12.54) ein, so erhalten wir

$$(\lambda^2 + 2a\lambda + b)e^{\lambda x} = 0.$$

Diese Gleichung ist (unabhängig davon, welchen Wert x hat) genau dann erfüllt, wenn λ eine Lösung der quadratischen Gleichung

$$\lambda^2 + 2a\lambda + b = 0 \quad (12.56)$$

ist. Diese Gleichung hat die beiden Lösungen

$$\lambda_{1,2} = -a \pm \sqrt{a^2 - b}. \quad (12.57)$$

Wie bei quadratischen Gleichungen üblich, unterscheiden wir drei Fälle.

Fall 1, $a^2 > b$. Beide Lösungen von (12.57) sind reell, zugehörige Lösungen von (12.54) sind

$$y_1(x) = e^{\lambda_1 x}, \quad y_2(x) = e^{\lambda_2 x}, \quad (12.58)$$

die allgemeine Lösung ist

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}. \quad (12.59)$$

Das Verhalten der Lösungen für $x \rightarrow \infty$ hängt vom Vorzeichen von λ_1 und λ_2 ab. Gilt $a > 0$ und $b > 0$, so sind $\lambda_1 < 0$ und $\lambda_2 < 0$, und

$$\lim_{x \rightarrow \infty} y(x) = 0, \quad (12.60)$$

für jede Lösung (12.59). Falls $a < 0$, $b > 0$, so sind beide λ_i positiv, und alle Lösungen sind exponentiell wachsend. Bei anderen Konstellationen können λ_1 und λ_2 unterschiedliche Vorzeichen haben, etwa $\lambda_1 > 0 > \lambda_2$. In diesem Fall dominiert die exponentiell wachsende Komponente (außer wenn $c_1 = 0$) für $x \rightarrow \infty$. Ein Beispiel ist etwa die Gleichung

$$y'' + 2y' - 3y = 0.$$

Die Gleichung

$$\lambda^2 + 2\lambda - 3 = 0$$

hat die beiden Lösungen $\lambda_1 = 1$, $\lambda_2 = -3$, die allgemeine Lösung der Differentialgleichung ist also

$$y(x) = c_1 e^x + c_2 e^{-3x}, \quad c_1, c_2 \in \mathbb{R}.$$

Fall 2, $a^2 = b$. Es gilt dann

$$\lambda_1 = \lambda_2 = -a, \quad \text{sei } \lambda = \lambda_1. \quad (12.61)$$

Neben

$$y_1(x) = e^{\lambda x}$$

ist auch

$$y_2(x) = x e^{\lambda x}$$

eine Lösung, wie man durch Einsetzen in (12.54) nachprüfen kann. Die allgemeine Lösung ist

$$y(x) = (c_1 + c_2 x) e^{\lambda x}. \quad (12.62)$$

Für sie gilt

$$\lim_{x \rightarrow \infty} y(x) = 0 \quad (12.63)$$

genau dann, wenn $a > 0$ (bis auf den Ausnahmefall der trivialen Lösung $y \equiv 0$). Als Beispiel betrachten wir die Gleichung

$$y'' + 4y' + 4y = 0.$$

Die Gleichung

$$\lambda^2 + 4\lambda + 4 = 0$$

hat die doppelte Nullstelle $\lambda = -2$, die allgemeine Lösung der Differentialgleichung ist also

$$y(x) = (c_1 + c_2 x) e^{-2x}, \quad c_1, c_2 \in \mathbb{R}.$$

Fall 3, $a^2 < b$. In diesem Fall gibt es zwei konjugiert komplexe Lösungen

$$\lambda_{1,2} = -a \pm i\omega, \quad \omega = \sqrt{b - a^2}. \quad (12.64)$$

Die zugehörigen Lösungen gemäß dem Ansatz (12.55) sind

$$\tilde{y}_{1,2}(x) = e^{-ax \pm i\omega x} = e^{-ax} (\cos(\omega x) \pm i \sin(\omega x)). \quad (12.65)$$

Diese Lösungen sind allerdings komplexwertige Funktionen. Bilden wir daraus die Linearkombinationen

$$y_1 = \frac{1}{2}(\tilde{y}_2 + \tilde{y}_1), \quad y_2 = \frac{1}{2}(i\tilde{y}_2 - i\tilde{y}_1),$$

so erhalten wir die reellen Lösungen

$$y_1(x) = e^{-ax} \cos(\omega x), \quad y_2(x) = e^{-ax} \sin(\omega x),$$

und hieraus die allgemeine Lösung

$$y(x) = c_1 e^{-ax} \cos(\omega x) + c_2 e^{-ax} \sin(\omega x). \quad (12.66)$$

Der Fall 3 beschreibt Schwingungsvorgänge. Ist $a = 0$, so ergibt sich eine **periodische Lösung** (ungedämpfte Schwingung). Die zugehörige Differentialgleichung lautet

$$y'' + by = 0, \quad (12.67)$$

man nennt sie die Gleichung des **harmonischen Oszillators**. Die Situation $a > 0$ beschreibt eine gedämpfte Schwingung, deren Amplitude für $x \rightarrow \infty$ exponentiell abklingt.

Wir betrachten nun die zugehörige **inhomogene** Gleichung

$$y'' + 2ay' + by = f(x), \quad (12.68)$$

wobei zusätzlich zu den Koeffizienten $a, b \in \mathbb{R}$ eine Funktion f gegeben ist. Sie steht in der Regel für eine äußere Einwirkung (etwa eine Kraft) auf das durch die Differentialgleichung beschriebene System.

Es stellt sich heraus, dass das Superpositionsprinzip in derselben Form wie bei linearen Gleichungssystemen gilt, das heißt, die allgemeine Lösung von (12.68) hat die Form

$$y(x) = \tilde{y}(x) + c_1 y_1(x) + c_2 y_2(x), \quad (12.69)$$

wobei \tilde{y} eine feste Lösung der inhomogenen Gleichung (12.68) und $c_1 y_1 + c_2 y_2$ die wie oben bestimmte allgemeine Lösung der homogenen Gleichung ist. In der Schreibweise mit Differentialoperator erhält die inhomogene Gleichung nämlich die Form

$$L[y] = f. \quad (12.70)$$

Die Tatsache, dass jede Funktion der Form (12.69) eine Lösung von (12.68) ist, ergibt sich aus der Rechnung

$$L[y] = L[\tilde{y} + c_1 y_1 + c_2 y_2] = L[\tilde{y}] + L[c_1 y_1 + c_2 y_2] = f + 0 = f.$$

Es erhebt sich nun die Frage, wie man eine einzelne Lösung \tilde{y} der inhomogenen Gleichung bestimmen kann. Eine Möglichkeit stellt die Methode der **Variation der Konstanten** dar. Deren Idee ist der Ansatz

$$\tilde{y}(x) = c_1(x) y_1(x) + c_2(x) y_2(x), \quad (12.71)$$

wobei y_1 und y_2 Lösungen der homogenen Gleichung sind. Man will nun die Funktionen c_1 und c_2 so bestimmen, dass (12.68) gilt. Setzt man zu diesem Zweck (12.71) in (12.68) ein, so erhält man eine Gleichung für die beiden unbekannt Funktionen c_1 und c_2 . Man kann dann eine der beiden Funktionen beliebig vorgeben und die andere aus der Gleichung bestimmen. Das wollen wir aber hier nicht im Einzelnen behandeln.

Nichtexistenz und Nichteindeutigkeit. Nicht jedes Anfangswertproblem hat eine Lösung. Beispiel: $I = [0, 1]$,

$$y' = f(y) = \begin{cases} -1, & y > 0, \\ 1, & y \leq 0, \end{cases} \quad y(0) = 0. \quad (12.72)$$

Ist y Lösung in I , so ist y stetig in I . y kann außer in $t = 0$ keine weitere Nullstelle in $(0, 1]$ haben, da andernfalls (Satz von Rolle) $y'(\tau) = 0$ für ein $\tau \in (0, 1)$. Ist $y(t) > 0$ für alle $t > 0$, dann ist $y'(t) = -1$ für alle $t > 0$, also y streng monoton fallend im Widerspruch zu $y(0) = 0$. Für den Fall $y(t) < 0$ ergibt sich ein analoger Widerspruch.

Man beachte, dass (12.72) auf $I = [-1, 0]$ eine Lösung hat, nämlich $y(t) = t$. Auf $I = [0, 1]$ erhalten wir eine Lösung, nämlich $y = 0$, falls wir die Definition von f modifizieren zu

$$y' = f(y) = \begin{cases} -1, & y > 0, \\ 0, & y = 0, \\ 1, & y \leq 0. \end{cases} \quad (12.73)$$

Manche Anfangswertprobleme haben mehrere Lösungen. Beispiel:

$$y' = \sqrt{|y|}, \quad y(0) = 0. \quad (12.74)$$

Außer der Funktion $y = 0$ ist beispielsweise auch

$$y(t) = \begin{cases} \frac{1}{4}t^2, & t > 0, \\ 0, & t \leq 0, \end{cases} \quad (12.75)$$

eine Lösung. In diesem Fall ist die rechte Seite f stetig, aber im Punkt $t = 0$ nicht differenzierbar, und die Lösung (12.75) ist stetig differenzierbar, aber im Punkt $t = 0$ nicht zweimal differenzierbar.

Systeme von Differentialgleichungen. Wir betrachten ein System gewöhnlicher Differentialgleichungen erster Ordnung, in Vektornotation geschrieben als

$$y' = f(t, y), \quad (12.76)$$

sowie das zugehörige Anfangswertproblem

$$y' = f(t, y), \quad y(t_0) = y_0. \quad (12.77)$$

Die gesuchte Funktion y besteht aus d Komponenten, $y = (y_1, \dots, y_d)$, und (12.76) bedeutet komponentenweise

$$\begin{aligned} y'_1 &= f_1(t, y_1, \dots, y_d) \\ &\vdots \\ y'_d &= f_d(t, y_1, \dots, y_d). \end{aligned}$$

Die Definition einer Lösung ist völlig analog zum Fall einer einzelnen Gleichung (der Fall $d = 1$). Ist $f : D \rightarrow \mathbb{R}^d$ mit $D \subset \mathbb{R} \times \mathbb{R}^d$, so heißt $y : I \rightarrow \mathbb{R}^d$ eine Lösung von (12.76) auf einem Intervall I , falls y in I differenzierbar ist und für alle $t \in I$ gilt, dass $(t, y(t)) \in D$ und

$$y'(t) = f(t, y(t)) \quad (12.78)$$

Falls außerdem $(t_0, y_0) \in D$ gegeben ist und $y(t_0) = y_0$ gilt, so heißt y Lösung des Anfangswertproblems (12.77) in I .

Hängt f nicht von y ab, so ist (wie im Fall einer einzelnen Differentialgleichung)

$$y(t) = y_0 + \int_{t_0}^t f(s) ds$$

eine Lösung des Anfangswertproblems (12.77), falls f stetig ist. (Die Komponentenfunktionen y_i von y sind Stammfunktionen der Komponentenfunktionen f_i von f .)

Differentialgleichungen höherer Ordnung (d.h. solche, in denen auch Ableitungen höherer Ordnung der unbekannteten Funktion auftreten), können auf die Form (12.76) zurückgeführt werden. Ist etwa gesucht eine Funktion $z : I \rightarrow \mathbb{R}$ als Lösung von

$$z^{(d)} = g(t, z, z', \dots, z^{(d-1)}), \quad (12.79)$$

so führen wir die neuen Variablen

$$y_1 = z, \quad y_2 = z', \dots, y_d = z^{(d-1)}$$

ein und erhalten die Form (12.76), indem wir setzen

$$y'_1 = y_2, \quad y'_2 = y_3, \dots, y'_d = g(t, y_1, \dots, y_d).$$

Als Anfangswerte sind Werte für $z(t_0), \dots, z^{(d-1)}(t_0)$ vorzuschreiben.

Ein Beispiel dafür stellt der gedämpfte harmonische Oszillator

$$mz'' + bz' + kz = 0, \quad m, b, k \text{ Konstante,}$$

dar, den wir (mit anderen Bezeichnungen) bereits behandelt haben. Seine Umschreibung in ein System ergibt

$$y'_1 = y_2, \quad y'_2 = -\frac{k}{m}y_1 - \frac{b}{m}y_2.$$

Existenz und Eindeutigkeit. Wir betrachten das Anfangswertproblem

$$y' = f(t, y), \quad y(t_0) = y_0. \quad (12.80)$$

Satz 12.1 Sei $D \subset \mathbb{R} \times \mathbb{R}^d$ offen, sei $f : D \rightarrow \mathbb{R}^d$ stetig differenzierbar, sei $(t_0, y_0) \in D$. Dann gibt es ein $\delta > 0$, so dass das Anfangswertproblem (12.80) auf dem Intervall $I = (t_0 - \delta, t_0 + \delta)$ eine eindeutige Lösung $y : I \rightarrow \mathbb{R}^d$ hat.

Diesen Satz kann man auf verschiedene Arten beweisen. Eine Möglichkeit ist es, zu einer gegebenen Schrittweite $h > 0$ den Polygonzug y_h (eine stückweise lineare Funktion) zu betrachten, der vom Eulerverfahren geliefert wird. Man kann dann zeigen, dass für $h \rightarrow 0$ die Funktionen y_h gleichmäßig gegen eine Funktion y konvergieren, und dass diese Funktion y das Anfangswertproblem löst. Damit erhält man die Existenz, aber nicht die Eindeutigkeit der Lösung. Letztere erhält man aus einem zweiten Schritt, indem man beweist, dass die Abstände zweier Lösungen sich durch ein Vielfaches der Abstände am Anfangspunkt (t_0, y_0) abschätzen lassen (und folglich gleich 0 sind, wenn

beide Lösungen vom selben Punkt starten). Eine andere Möglichkeit (der Satz von Picard-Lindelöf) ist es, eine geeignete Fixpunktiteration anzusetzen und zu zeigen, dass die Lösungen des Anfangswertproblems mit den Fixpunkten der Iterationsvorschrift übereinstimmen, und dass die Iteration gegen einen Fixpunkt konvergiert, der wiederum eindeutig bestimmt ist.

Satz 12.1 ist ein **lokaler** Existenz- und Eindeigkeitssatz, er garantiert die Lösbarkeit auf einem (möglicherweise sehr kleinen) Existenzintervall. Ob und wie weit sich die Lösungen fortsetzen, ist unterschiedlich. Hier einige Beispiele: Das AWP

$$y' = y^2 + 1, \quad y(0) = 0,$$

hat die Lösung und das maximale Existenzintervall

$$y(t) = \tan t, \quad I = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Das AWP

$$y' = y^2, \quad y(0) = 1,$$

hat die Lösung und das maximale Existenzintervall

$$y(t) = \frac{1}{1-t}, \quad I = (-\infty, 1).$$

Das AWP

$$y' = y^2 - 1, \quad y(0) = 0,$$

hat die Lösung und das maximale Existenzintervall

$$y(t) = \tanh t, \quad I = \mathbb{R}.$$

In allen diesen drei Beispielen ist die rechte Seite stetig differenzierbar. In den beiden ersten Beispielen kann die Lösung nicht auf ganz \mathbb{R} fortgesetzt werden, da sie eine Polstelle hat.

Etwas anders gelagert ist die Situation für das AWP

$$y' = \frac{1}{y}, \quad y(0) = 1.$$

Hier ist $D = \{(t, y) : y \neq 0\} = \mathbb{R} \times (\mathbb{R} \setminus \{0\})$, die Lösung ist

$$y(t) = \sqrt{2t+1}, \quad \text{und } I = \left(-\frac{1}{2}, \infty\right)$$

ist das maximale Existenzintervall. Es gilt hier

$$\lim_{t \downarrow -\frac{1}{2}} (t, y(t)) = \left(-\frac{1}{2}, 0\right) \in \partial D.$$

Es sind also drei verschiedene Fälle für das Verhalten der Lösung für $t > t_0$ (analog $t < t_0$) erkennbar.

1. Die Lösung existiert für alle $t > t_0$.

2. Es gibt ein $b > t_0$ mit

$$\lim_{\substack{t \rightarrow b \\ t < b}} \|y(t)\| = \infty.$$

3. Es gibt ein $b > t_0$ mit

$$\lim_{\substack{t \rightarrow b \\ t < b}} \text{dist}((t, y(t)), \partial D) = 0.$$

13 Das Integral im Mehrdimensionalen

Wir beschäftigen uns zuerst mit dem zweidimensionalen Fall.

Definition des Integrals auf achsenparallelen Rechtecken in der Ebene. Sei

$$\Omega = [a, b] \times [c, d], \quad (13.1)$$

das heißt, Ω ist ein achsenparalleles Rechteck im \mathbb{R}^2 . Für eine Funktion $f : \Omega \rightarrow \mathbb{R}$ wollen wir das “zweidimensionale” Integral

$$\iint_{\Omega} f(x, y) \, dx \, dy \quad (13.2)$$

definieren. Seine Konstruktion erfolgt analog zum eindimensionalen Fall. Seien

$$\begin{aligned} Z_1 &= \{x_0, \dots, x_n\}, & \text{wobei } a &= x_0 < x_1 < \dots < x_n = b, \\ Z_2 &= \{y_0, \dots, y_m\}, & \text{wobei } c &= y_0 < y_1 < \dots < y_m = d, \end{aligned} \quad (13.3)$$

Zerlegungen von $[a, b]$ bzw. $[c, d]$. Hieraus ergibt sich eine Zerlegung Z von $[a, b] \times [c, d]$ in nm Rechtecke

$$Q_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j], \quad 1 \leq i \leq n, \quad 1 \leq j \leq m. \quad (13.4)$$

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

$$f(x, y) = c_{ij}, \quad \text{falls } x \in (x_{i-1}, x_i), \, y \in (y_{j-1}, y_j), \quad (13.5)$$

mit gegebenen $c_{ij} \in \mathbb{R}$ heißt **Treppenfunktion zur Zerlegung Z** . (Welchen Wert f auf den Rändern der Rechtecke Q_{ij} annimmt, ist uns gleichgültig.) Für diese Funktion f definieren wir

$$\iint_{Q_{ij}} f(x, y) \, dx \, dy = c_{ij}(x_i - x_{i-1})(y_j - y_{j-1}), \quad (13.6)$$

das ist gerade das Volumen des Quaders mit Grundfläche Q_{ij} und Höhe c_{ij} , und weiter

$$\iint_{\Omega} f(x, y) \, dx \, dy = \sum_{i=1}^n \sum_{j=1}^m \iint_{Q_{ij}} f(x, y) \, dx \, dy = \sum_{i=1}^n \sum_{j=1}^m c_{ij}(x_i - x_{i-1})(y_j - y_{j-1}). \quad (13.7)$$

Wir betrachten nun eine beliebige **beschränkte** Funktion $f : \Omega \rightarrow \mathbb{R}$, sei etwa

$$|f(x, y)| \leq C, \quad \text{für alle } (x, y) \in [a, b].$$

Wir fixieren eine Zerlegung Z der Form (13.3), (13.4). Sind $\varphi, \psi : \Omega \rightarrow \mathbb{R}$ Treppenfunktionen zur Zerlegung Z mit

$$\varphi(x, y) \leq f(x, y) \leq \psi(x, y), \quad \text{für alle } (x, y) \in \Omega, \quad (13.8)$$

so folgt aus (13.7), angewandt auf φ und ψ , dass

$$\iint_{\Omega} \varphi(x, y) \, dx \, dy \leq \iint_{\Omega} \psi(x, y) \, dx \, dy. \quad (13.9)$$

Das Integral von f soll einen Wert zwischen diesen beiden Zahlen haben. Wir setzen nun

$$m_{ij} = \inf_{(x,y) \in Q_{ij}} f(x,y), \quad M_{ij} = \sup_{(x,y) \in Q_{ij}} f(x,y), \quad (13.10)$$

und definieren die **Untersumme** und **Obersumme** von f zur Zerlegung Z durch

$$U_Z(f) = \iint_{\Omega} \varphi(x,y) dx dy, \quad \text{wobei } \varphi(x,y) = m_{ij} \text{ f\"ur } (x,y) \in Q_{ij}, \quad (13.11)$$

$$O_Z(f) = \iint_{\Omega} \psi(x,y) dx dy, \quad \text{wobei } \psi(x,y) = M_{ij} \text{ f\"ur } (x,y) \in Q_{ij}. \quad (13.12)$$

Nach (13.9) gilt

$$U_Z(f) \leq O_Z(f). \quad (13.13)$$

Ziel ist, dass das Integral von f einen Wert zwischen diesen beiden Zahlen haben soll.

Ist \tilde{Z} eine **feinere** Zerlegung als Z , das heißt eine solche, die aus Z durch Hinzufügen weiterer Teilpunkte entsteht, so gilt

$$U_Z(f) \leq U_{\tilde{Z}}(f) \leq O_{\tilde{Z}}(f) \leq O_Z(f). \quad (13.14)$$

Man erhält weiter, dass

$$U_Z(f) \leq O_{\hat{Z}}(f) \quad (13.15)$$

gilt für zwei Zerlegungen Z und \hat{Z} , indem man eine dritte Zerlegung \tilde{Z} betrachtet, die sowohl für Z als auch für \hat{Z} eine Verfeinerung ist. Wir definieren nun das **Unterintegral** und das **Oberintegral** von f durch

$$U(f) = \sup\{U_Z(f) : Z \text{ ist Zerlegung von } Q\}, \quad (13.16)$$

$$O(f) = \inf\{O_Z(f) : Z \text{ ist Zerlegung von } Q\}. \quad (13.17)$$

Aus der Definition von Supremum und Infimum folgt wegen (13.15), dass

$$U(f) \leq O(f). \quad (13.18)$$

Da das Integral von f für jede Zerlegung Z mindestens so groß wie $U_Z(f)$ und höchstens so groß wie $O_Z(f)$ sein soll, soll es auch zwischen den beiden Zahlen $U(f)$ und $O(f)$ liegen.

Ist nun aber

$$U(f) = O(f), \quad (13.19)$$

so gibt es dafür nur eine Möglichkeit, nämlich

$$\iint_{\Omega} f(x,y) dx dy = U(f) = O(f). \quad (13.20)$$

Definition 13.1 (Integral)

Sei $f : \Omega \rightarrow \mathbb{R}$ beschränkt. Falls (13.19) gilt, also Unterintegral und Oberintegral von f übereinstimmen, so heißt f **integrierbar**, (genauer: **Riemann-integrierbar**) auf Ω , und das Integral von f ist definiert durch (13.20). \square

Wie im Eindimensionalen ist das so definierte Integral **linear** und **monoton**, das heißt, für integrierbare Funktionen $f, g : \Omega \rightarrow \mathbb{R}$ gilt

$$\iint_{\Omega} \alpha f(x, y) + \beta g(x, y) dx dy = \alpha \iint_{\Omega} f(x, y) dx dy + \beta \iint_{\Omega} f(x, y) dx dy, \quad \alpha, \beta, \in \mathbb{R}, \quad (13.21)$$

sowie

$$\iint_{\Omega} f(x, y) dx dy \leq \iint_{\Omega} g(x, y) dx dy, \quad \text{falls } f \leq g \text{ in } \Omega, \quad (13.22)$$

und

$$\left| \iint_{\Omega} f(x, y) dx dy \right| \leq \iint_{\Omega} |f(x, y)| dx dy. \quad (13.23)$$

Berechnung von Integralen auf achsenparallelen Rechtecken. Wir können das zweidimensionale Integral auf $\Omega = [a, b] \times [c, d]$ auf zwei eindimensionale Integrale zurückführen. Sei $f : \Omega \rightarrow \mathbb{R}$ eine Treppenfunktion wie oben,

$$f(x, y) = c_{ij}, \quad \text{falls } x \in (x_{i-1}, x_i), y \in (y_{j-1}, y_j).$$

Für festgehaltenes y ist die durch $x \mapsto f(x, y)$ definierte Funktion eine Treppenfunktion im eindimensionalen Intervall $[a, b]$ zur Zerlegung Z_1 , und es gilt

$$\int_a^b f(x, y) dx = \sum_{i=1}^n c_{ij}(x_i - x_{i-1}), \quad \text{falls } y_{j-1} < y < y_j. \quad (13.24)$$

Die Funktion

$$G(y) = \int_a^b f(x, y) dx \quad (13.25)$$

ist wegen (13.24) eine Treppenfunktion in $[c, d]$ zur Zerlegung Z_2 , und es gilt

$$\int_c^d G(y) dy = \sum_{j=1}^m \left[\sum_{i=1}^n c_{ij}(x_i - x_{i-1}) \right] (y_j - y_{j-1}) = \iint_{\Omega} f(x, y) dx dy \quad (13.26)$$

nach Definition (13.7). Analog gilt für festgehaltenes x

$$\int_c^d f(x, y) dy = \sum_{j=1}^m c_{ij}(y_j - y_{j-1}), \quad \text{falls } x_{i-1} < x < x_i, \quad (13.27)$$

und für

$$F(x) = \int_c^d f(x, y) dy \quad (13.28)$$

folgt wie in (13.26)

$$\int_a^b F(x) dx = \iint_{\Omega} f(x, y) dx dy. \quad (13.29)$$

Insgesamt ergibt sich, dass für Treppenfunktionen gilt

$$\iint_{\Omega} f(x, y) dx dy = \int_a^b \int_c^d f(x, y) dy dx = \int_c^d \int_a^b f(x, y) dx dy. \quad (13.30)$$

Der Satz von Fubini besagt, dass diese Formel auch für beliebige Funktionen f gilt, falls $|f|$ integrierbar ist, also falls

$$\iint_{\Omega} |f(x, y)| dx dy < \infty. \quad (13.31)$$

Satz 13.2 (Fubini)

Sei $\Omega = [a, b] \times [c, d]$, $f : \Omega \rightarrow \mathbb{R}$, sei $|f|$ integrierbar auf Ω . Dann gilt

$$\iint_{\Omega} f(x, y) dx dy = \int_a^b \int_c^d f(x, y) dy dx = \int_c^d \int_a^b f(x, y) dx dy. \quad (13.32)$$

□

Außer der Integrierbarkeit von $|f|$ auf Ω sind keine weiteren Voraussetzungen erforderlich, es ist z.B. gleichgültig, ob f stetig ist oder nicht.

Als Beispiel betrachten wir

$$\iint_{\Omega} x^2 y dx dy, \quad \Omega = [0, 1] \times [1, 2].$$

Es gibt zwei Möglichkeiten, dieses Integral mit dem Satz von Fubini auszurechnen, je nachdem ob man zuerst bezüglich y oder zuerst bezüglich x integriert. Entweder:

$$\begin{aligned} \iint_{\Omega} x^2 y dx dy &= \int_0^1 \int_1^2 x^2 y dy dx = \int_0^1 \left(x^2 \cdot \frac{1}{2} y^2 \Big|_{y=1}^{y=2} \right) dx = \int_0^1 2x^2 - \frac{1}{2} x^2 dx \\ &= \int_0^1 \frac{3}{2} x^2 dx = \frac{1}{2} x^3 \Big|_{x=0}^{x=1} = \frac{1}{2}. \end{aligned}$$

Oder:

$$\begin{aligned} \iint_{\Omega} x^2 y dx dy &= \int_1^2 \int_0^1 x^2 y dx dy = \int_1^2 \left(\frac{1}{3} x^3 y \Big|_{x=0}^{x=1} \right) dy = \int_1^2 \frac{1}{3} y dy \\ &= \frac{1}{6} y^2 \Big|_{y=1}^{y=2} = \frac{1}{2}. \end{aligned}$$

In diesem Beispiel geht es sogar viel einfacher: Wir beginnen wie oben,

$$\iint_{\Omega} x^2 y dx dy = \int_0^1 \int_1^2 x^2 y dy dx,$$

und beobachten nun, dass die x -Abhängigkeit im inneren Integral in Form eines Faktors auftritt, also

$$\int_0^1 \int_1^2 x^2 y \, dy \, dx = \int_0^1 x^2 \int_1^2 y \, dy \, dx.$$

Wir stellen als nächstes fest, dass nunmehr das innere Integral nicht von x abhängt und somit als konstanter Faktor im äußeren Integral aufgefasst werden kann, also

$$\int_0^1 x^2 \int_1^2 y \, dy \, dx = \int_0^1 x^2 \, dx \cdot \int_1^2 y \, dy.$$

Die beiden Integrale auf der rechten Seite können direkt ausgewertet werden, insgesamt ist also

$$\iint_{\Omega} x^2 y \, dx \, dy = \int_0^1 x^2 \, dx \cdot \int_1^2 y \, dy = \frac{1}{3} \cdot \frac{3}{2} = \frac{1}{2}.$$

Allgemein tritt dieser Fall dann auf, wenn der Integrand sich schreiben lässt als

$$f(x, y) = f_1(x) \cdot f_2(y), \quad (13.33)$$

dann ist

$$\iint_{\Omega} f(x, y) \, dx \, dy = \int_a^b f_1(x) \, dx \cdot \int_c^d f_2(y) \, dy. \quad (13.34)$$

Integrale auf höherdimensionalen achsenparallelen Quadern. Für einen Quader

$$\Omega = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$$

im \mathbb{R}^3 definieren wir das Integral, in diesem Fall **Volumenintegral** genannt,

$$\iiint_{\Omega} f(x, y, z) \, dx \, dy \, dz \quad (13.35)$$

analog zum Zweidimensionalen, indem wir die Intervalle $[a_i, b_i]$ durch Zerlegungen Z_1, Z_2 und Z_3 unterteilen, Ω entsprechend in kleine Quader Q_{ijk} zerlegen, Treppenfunktionen (das heißt, auf den einzelnen Quadern Q_{ijk} konstante Funktionen) betrachten, Unter- und Obersummen definieren und das Integral durch die beschriebene Einschließung im Grenzwert erhalten. Es gilt ebenfalls der Satz von Fubini, so dass die Berechnung von (13.35) auf die Berechnung von drei eindimensionalen Integralen zurückgeführt werden kann, wobei die Reihenfolge der Integrationen beliebig gewählt werden kann,

$$\begin{aligned} \iiint_{\Omega} f(x, y, z) \, dx \, dy \, dz &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} f(x, y, z) \, dz \, dy \, dx \\ &= \int_{a_1}^{b_1} \int_{a_3}^{b_3} \int_{a_2}^{b_2} f(x, y, z) \, dy \, dz \, dx \\ &\quad \vdots \\ &= \int_{a_3}^{b_3} \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x, y, z) \, dx \, dy \, dz, \end{aligned} \quad (13.36)$$

das sind 6 verschiedene Möglichkeiten. Als Beispiel betrachten wir

$$\iiint_{\Omega} x^2 y z \, dx \, dy \, dz, \quad \Omega = [0, 1] \times [1, 2] \times [2, 3].$$

Wir rechnen nur eine Variante aus,

$$\begin{aligned} \iiint_{\Omega} x^2 y z \, dx \, dy \, dz &= \int_2^3 \int_1^2 \int_0^1 x^2 y z \, dx \, dy \, dz = \int_2^3 \int_1^2 \left(\frac{1}{3} x^3 y z \Big|_{x=0}^{x=1} \right) dy \, dz \\ &= \int_2^3 \int_1^2 \frac{1}{3} y z \, dy \, dz = \int_2^3 \frac{1}{6} y^2 z \Big|_{y=1}^{y=2} dz = \int_2^3 \frac{1}{2} z \, dz = \frac{5}{4}. \end{aligned}$$

Auch hier ist es aufgrund der Produktform des Integranden möglich, das dreidimensionale Integral direkt in ein Produkt eindimensionaler Integrale zu zerlegen,

$$\iiint_{\Omega} x^2 y z \, dx \, dy \, dz = \int_2^3 z \, dz \cdot \int_1^2 y \, dy \cdot \int_0^1 x^2 \, dx = \frac{5}{2} \cdot \frac{3}{2} \cdot \frac{1}{3} = \frac{5}{4}.$$

Die allgemeine Formel für diesen Fall ist

$$\iiint_{\Omega} f_1(x) f_2(y) f_3(z) \, dx \, dy \, dz = \int_{a_1}^{b_1} f_1(x) \, dx \cdot \int_{a_2}^{b_2} f_2(y) \, dy \cdot \int_{a_3}^{b_3} f_3(z) \, dz. \quad (13.37)$$

Integrale im \mathbb{R}^d für $d > 3$ für einen achsenparallelen Quader

$$\Omega = [a_1, b_1] \times \cdots \times [a_d, b_d]$$

werden analog definiert, es gilt ebenfalls der Satz von Fubini, also

$$\int_{\Omega} \cdots \int f(x_1, \dots, x_d) \, dx_1 \cdots dx_d = \int_{a_1}^{b_1} \cdots \int_{a_d}^{b_d} f(x_1, \dots, x_d) \, dx_d \cdots dx_1, \quad (13.38)$$

wobei die Reihenfolge der Integrale auf der rechten Seite beliebig abgeändert werden kann.

In der Mathematik hat sich als einheitliche Notation für Integrale im Mehrdimensionalen die Form

$$\int_{\Omega} f(x) \, dx \quad (13.39)$$

durchgesetzt, also genau wie im Eindimensionalen. Hier bedeutet Ω den Integrationsbereich (bisher haben wir nur achsenparallele Quader behandelt), integriert wird eine skalare Funktion $f : \Omega \rightarrow \mathbb{R}$, das Argument von f ist ein Vektor $x \in \mathbb{R}^3$. Da die Dimension d im Definitionsbereich von f implizit enthalten ist, ist es überflüssig, das Integralzeichen d -mal hinzuschreiben; die Zerlegung in n eindimensionale Integrale ist erst dann relevant, wenn man die Berechnung nach Fubini vornimmt.

Flächeninhalt von Teilmengen der Ebene. Ist $\Omega = [a, b] \times [c, d]$ ein achsenparalleles Rechteck, so ist sein Flächeninhalt gegeben durch

$$F(\Omega) = (b - a)(d - c). \quad (13.40)$$

Wir wollen nun den Flächeninhalt einer beliebigen beschränkten Teilmenge $\Omega \subset \mathbb{R}^2$ bestimmen. Wir wählen dazu einen Quader der Form $Q = [-M, M] \times [-M, M]$, wobei M so groß ist, dass $\Omega \subset Q$. Wir definieren den Flächeninhalt von Ω als

$$F(\Omega) = \iint_Q 1_\Omega(x, y) dx dy, \quad \text{wobei} \quad 1_\Omega(x, y) = \begin{cases} 1, & (x, y) \in \Omega, \\ 0, & (x, y) \notin \Omega. \end{cases} \quad (13.41)$$

Der Flächeninhalt ist also definiert, wenn die Funktion 1_Ω integrierbar ist. Wenn man den Begriff "integrierbar" etwas allgemeiner fasst als in den in dieser Vorlesung angegebenen Definitionen, so ist das für alle "normalerweise auftretenden" Mengen der Fall.

Die Bestimmung des Flächeninhalts wird durch (13.41) auf den durch das Integral definierten Grenzprozess zurückgeführt, man kann sich das wie folgt vorstellen. Sei Z eine Zerlegung von Q in kleine Rechtecke Q_{ij} , wie in der Definition des zweidimensionalen Bereichsintegrals. Die Menge Ω wird angenähert von innen durch

$$\Omega_Z = \bigcup_{Q_{ij} \subset \Omega} Q_{ij}, \quad \text{also} \quad \Omega_Z \subset \Omega, \quad (13.42)$$

und von außen durch

$$\Omega^Z = \bigcup_{Q_{ij} \cap \Omega \neq \emptyset} Q_{ij}, \quad \text{also} \quad \Omega \subset \Omega^Z. \quad (13.43)$$

Es ist dann $\varphi = 1_{\Omega_Z}$ eine Treppenfunktion, und

$$F(\Omega_Z) = \iint_Q \varphi(x, y) dx dy, \quad (13.44)$$

und analog $\psi = 1_{\Omega^Z}$ eine Treppenfunktion mit

$$F(\Omega^Z) = \iint_Q \psi(x, y) dx dy. \quad (13.45)$$

Die Einschachtelung in der Definition des Integrals

$$\iint_Q \varphi(x, y) dx dy \leq \iint_Q 1_\Omega(x, y) dx dy \leq \iint_Q \psi(x, y) dx dy$$

entspricht also der Einschachtelung des Flächeninhalts

$$F(\Omega_Z) \leq F(\Omega) \leq F(\Omega^Z).$$

Ist $\Omega \subset \mathbb{R}^2$ eine Menge, deren Rand aus Kurven besteht, so können wir $F(\Omega)$ mit dem Satz von Fubini berechnen. Sei

$$\Omega = \{(x, y) : a \leq x \leq b, c(x) \leq y \leq d(x)\}. \quad (13.46)$$

Die Fubini-Formel

$$\iint_Q 1_\Omega(x, y) dx dy = \int_{-M}^M \int_{-M}^M 1_\Omega(x, y) dy dx$$

wird zu

$$F(\Omega) = \int_a^b \int_{c(x)}^{d(x)} 1 \, dy \, dx = \int_a^b d(x) - c(x) \, dx, \quad (13.47)$$

da außerhalb der Integrationsgrenzen die Funktion 1_Ω den Wert Null hat. Analog gilt für

$$\Omega = \{(x, y) : c \leq y \leq d, a(y) \leq x \leq b(y)\}, \quad (13.48)$$

dass

$$F(\Omega) = \int_c^d \int_{a(y)}^{b(y)} 1 \, dx \, dy = \int_c^d b(y) - a(y) \, dy. \quad (13.49)$$

Sei nun Ω die disjunkte Vereinigung zweier Mengen,

$$\Omega = \Omega_1 \cup \Omega_2, \quad \Omega_1 \cap \Omega_2 = \emptyset. \quad (13.50)$$

Es ist dann

$$1_\Omega = 1_{\Omega_1} + 1_{\Omega_2},$$

und aus der Linearität des Integrals folgt

$$F(\Omega) = F(\Omega_1) + F(\Omega_2). \quad (13.51)$$

Ist Ω eine Menge, die sich aus endlich vielen Mengen der Form (13.46) oder (13.48) zusammensetzt, so erhalten wir aus (13.47), (13.49) und (13.51) ihren Flächeninhalt $F(\Omega)$ durch Addition der entsprechenden Formeln.

Aus der Monotonie des Integrals folgt sofort, dass

$$F(\Omega) \leq F(\tilde{\Omega}), \quad \text{falls } \Omega \subset \tilde{\Omega}, \quad (13.52)$$

da in diesem Fall $1_\Omega \leq 1_{\tilde{\Omega}}$.

Es taucht die Frage auf, ob der Flächeninhalt einer Menge $\Omega \subset \mathbb{R}^2$ davon abhängt, ob der Rand von Ω zu Ω gehört oder nicht. Ist etwa $\Omega = (0, 1) \times (0, 1)$ das offene Einheitsquadrat, so besteht $\partial\Omega$ aus den 4 Randseiten einschließlich der Ecken, und

$$\bar{\Omega} = [0, 1] \times [0, 1] = \Omega \cup \partial\Omega.$$

Gemäß (13.50) gilt, da die Vereinigung disjunkt ist,

$$F(\bar{\Omega}) = F(\Omega) + F(\partial\Omega), \quad (13.53)$$

so dass $\bar{\Omega}$ und Ω denselben Flächeninhalt haben, falls $F(\partial\Omega) = 0$. Wir sagen, dass eine Teilmenge N von \mathbb{R}^2 eine **Nullmenge** ist, falls

$$F(N) = 0. \quad (13.54)$$

Man kann zeigen, dass der Graph einer Kurve diese Eigenschaft hat, falls die Kurve eine endliche Länge besitzt. Das hat zur Folge, dass für alle Teilmengen der Ebene, deren Rand aus endlich vielen solcher Kurven besteht, es für die Flächenbestimmung gleichgültig ist, ob wir den Rand (oder Teile davon) mitbetrachten oder nicht.

Integrale auf allgemeinen Teilmengen. Wir behandeln zunächst den zweidimensionalen Fall. Ist $f : \Omega \rightarrow \mathbb{R}$ eine beschränkte Funktion auf einer beschränkten Teilmenge

$\Omega \subset \mathbb{R}^2$, so betten wir Ω wieder in einen Quader Q ein, setzen f außerhalb von Ω durch 0 fort und definieren

$$\iint_{\Omega} f(x, y) dx dy = \iint_Q f(x, y) 1_{\Omega}(x, y) dx dy, \quad (13.55)$$

falls die Funktion $f 1_{\Omega}$ integrierbar ist. Das ist beispielsweise dann der Fall, wenn f stetig ist und Ω einen Flächeninhalt hat gemäß (13.41). Auch in diesem Fall ist das Integral linear und monoton, und weiterhin analog zu (13.51)

$$\iint_{\Omega} f(x, y) dx dy = \iint_{\Omega_1} f(x, y) dx dy + \iint_{\Omega_2} f(x, y) dx dy, \quad (13.56)$$

falls Ω die disjunkte Vereinigung von Ω_1 und Ω_2 ist. Weiterhin gilt

$$\iint_{\Omega} f(x, y) dx dy = 0, \quad \text{falls } F(\Omega) = 0, \quad (13.57)$$

das heißt, falls Ω eine Nullmenge ist.

Hat Ω die Form

$$\Omega = \{(x, y) : a \leq x \leq b, c(x) \leq y \leq d(x)\}, \quad (13.58)$$

so ergibt sich wie oben mit Fubini

$$\iint_{\Omega} f(x, y) dx dy = \int_a^b \int_{c(x)}^{d(x)} f(x, y) dy dx. \quad (13.59)$$

Als Beispiel betrachten wir die Menge Ω deren Rand durch die Geraden $y = x$ und $y = 2$ sowie die Kurve $xy = 1$ gegeben ist, und bestimmen

$$\iint_{\Omega} x^2 y dx dy.$$

Wir können Ω darstellen als

$$\Omega = \Omega_1 \cup \Omega_2,$$

mit

$$\Omega_1 = \{(x, y) : \frac{1}{2} \leq x \leq 1, \frac{1}{x} \leq y \leq 2\}, \quad \Omega_2 = \{(x, y) : 1 \leq x \leq 2, x \leq y \leq 2\},$$

und es ergibt sich

$$\begin{aligned} \iint_{\Omega} x^2 y dx dy &= \int_{\frac{1}{2}}^1 \int_{\frac{1}{x}}^2 x^2 y dy dx + \int_1^2 \int_x^2 x^2 y dy dx \\ &= \int_{\frac{1}{2}}^1 \left(\frac{1}{2} x^2 y^2 \Big|_{y=\frac{1}{x}}^{y=2} \right) dx + \int_1^2 \left(\frac{1}{2} x^2 y^2 \Big|_{y=x}^{y=2} \right) dx \\ &= \int_{\frac{1}{2}}^1 \left(2x^2 - \frac{1}{2} \right) dx + \int_1^2 \left(2x^2 - \frac{1}{2} x^4 \right) dx \\ &= \frac{57}{30}. \end{aligned}$$

Wir können Ω auch darstellen als

$$\Omega = \{(x, y) : 1 \leq y \leq 2, \frac{1}{y} \leq x \leq y\},$$

dann ergibt sich

$$\iint_{\Omega} x^2 y \, dx \, dy = \int_1^2 \int_{\frac{1}{y}}^y x^2 y \, dx \, dy = \int_1^2 \left(\frac{1}{3} x^3 y \Big|_{x=\frac{1}{y}}^{x=y} \right) dy = \int_1^2 \left(\frac{1}{3} y^4 - \frac{1}{3y^2} \right) dy = \frac{57}{30}.$$

In höheren Dimensionen geht das genauso. Für eine beschränkte Menge $\Omega \subset \mathbb{R}^d$ definieren wir

$$\int \cdots \int_{\Omega} f(x_1, \dots, x_d) \, dx_1 \cdots dx_d = \int \cdots \int_Q f(x_1, \dots, x_d) 1_{\Omega}(x_1, \dots, x_d) \, dx_1 \cdots dx_d, \quad (13.60)$$

wobei Q ein Ω enthaltender achsenparalleler Quader ist. Die im Zweidimensionalen geltenden Rechenregeln gelten entsprechend auch in höheren Dimensionen.

Für $\Omega \subset \mathbb{R}^3$ ergibt

$$V(\Omega) = \iiint_{\Omega} 1 \, dx \, dy \, dz \quad (13.61)$$

den **Rauminhalt** oder das **Volumen** von Ω .

Substitution, Koordinatentransformationen. Es kann günstig sein, Integrale in anderen als den kartesischen Koordinaten zu berechnen. Es erhebt sich die Frage, wie sich Integrale transformieren, wenn die Koordinaten transformiert werden, beispielsweise auf Polar- oder Kugelkoordinaten. Wir brauchen dafür eine Integrationsformel, die der Kettenregel beim Differenzieren entspricht. Im Eindimensionalen ist das die **Substitutionsregel**

$$\int_{T(a)}^{T(b)} f(x) \, dx = \int_a^b f(T(\xi)) T'(\xi) \, d\xi. \quad (13.62)$$

Hierbei entspricht T einer Koordinatentransformation, und f ist die Funktion, die man integrieren will.

Wir behandeln nun die Verallgemeinerung der Substitutionsregel auf den mehrdimensionalen Fall. Sie heißt dann **Transformationsformel** und lautet in kompakter Notation

$$\int_{T(\Omega)} f(x) \, dx = \int_{\Omega} f(T(\xi)) |\det(J_T(\xi))| \, d\xi. \quad (13.63)$$

Hierbei ist $\Omega \subset \mathbb{R}^d$, $T : \Omega \rightarrow \mathbb{R}^d$, $f : T(\Omega) \rightarrow \mathbb{R}$, J_T bezeichnet die Funktionalmatrix von T , das Integral ist als d -dimensionales Integral aufzufassen. Die Formel (13.63) gilt jedenfalls dann, wenn T auf einer Ω umfassenden offenen Menge Q definiert, stetig differenzierbar und injektiv ist, und $\det(J_T(\xi)) \neq 0$ ist für alle $\xi \in Q$.

Wir behandeln die Transformationsformel in einer Reihe von speziellen Situationen. Sei

$$T : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad T(\xi) = A\xi + b, \quad A \in \mathbb{R}^{(2,2)}, \quad b \in \mathbb{R}^2, \quad (13.64)$$

eine **affine Transformation in der Ebene**. Es ist dann

$$J_T(\xi) = A,$$

unabhängig von ξ , und (13.63) wird zu

$$\int_{T(\Omega)} f(x) dx = \int_{\Omega} f(A\xi + b) |\det(A)| d\xi. \quad (13.65)$$

Als Beispiel bestimmen wir die Fläche einer Ellipse mit Mittelpunkt 0 und den Halbachsen a und b . Wir setzen

$$\Omega' = \{(x_1, x_2) : \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1\}.$$

Es ist

$$\Omega' = T(\Omega), \quad \Omega = \{\xi : \xi \in \mathbb{R}^2, \|\xi\| \leq 1\}, \quad T(\xi) = A\xi, \quad A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

Aus (13.65) folgt

$$F(\Omega') = \iint_{\Omega'} 1 dx_1 dx_2 = \iint_{\Omega} ab d\xi_1 d\xi_2 = abF(\Omega) = \pi ab,$$

da der Einheitskreis die Fläche π hat.

Als weiteres Beispiel bestimmen wir das Integral

$$\iint_{\Omega'} x_1 x_2 dx_1 dx_2, \quad (13.66)$$

wobei $\Omega' \subset \mathbb{R}^2$ das Parallelogramm mit den Ecken $(0, 0)$, $(1, 1)$, $(2, 0)$ und $(3, 1)$ ist. Wir führen das Integral über das Parallelogramm auf ein Integral über dem Einheitsquadrat

$$\Omega = [0, 1] \times [0, 1]$$

zurück, indem wir setzen

$$T(\xi) = A\xi, \quad A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}.$$

Mit

$$f(x_1, x_2) = x_1 x_2$$

erhalten wir

$$f(A\xi) = f\begin{pmatrix} 2\xi_1 + \xi_2 \\ \xi_2 \end{pmatrix} = (2\xi_1 + \xi_2)\xi_2, \quad T(\Omega) = \Omega', \quad |\det(A)| = 2,$$

also

$$\iint_{\Omega'} x_1 x_2 dx_1 dx_2 = \iint_{\Omega} (2\xi_1 + \xi_2)\xi_2 \cdot 2 d\xi_1 d\xi_2 = \int_0^1 \int_0^1 4\xi_1 \xi_2 + 2\xi_2^2 d\xi_1 d\xi_2 = \frac{5}{3}.$$

Wir betrachten die Transformation **von Polarkoordinaten auf kartesische Koordinaten**,

$$T(r, \varphi) = (r \cos \varphi, r \sin \varphi). \quad (13.67)$$

In diesem Fall ist

$$J_T(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}, \quad \det J_T(r, \varphi) = r, \quad (13.68)$$

also wird (13.63) zu

$$\iint_{T(\Omega)} f(x, y) dx dy = \iint_{\Omega} f(r \cos \varphi, r \sin \varphi) r d\varphi dr. \quad (13.69)$$

Als erstes Beispiel berechnen wir nochmals die Fläche der Einheitskreisscheibe. Für

$$\Omega = (0, 1) \times [0, 2\pi)$$

ist $T(\Omega)$ gerade die offene Einheitskreisscheibe (ohne den Nullpunkt, was für die Integration keine Rolle spielt), und

$$F(T(\Omega)) = \iint_{T(\Omega)} 1 dx dy = \iint_{\Omega} r dr d\varphi = \int_0^{2\pi} \int_0^1 r dr d\varphi = 2\pi \frac{1}{2} = \pi.$$

Wir betrachten nun den Übergang **von Kugelkoordinaten auf kartesische Koordinaten im Dreidimensionalen**. Die Transformationsvorschrift ist

$$T(r, \theta, \varphi) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta), \quad T : (0, \infty) \times (0, \pi) \times (0, 2\pi) \rightarrow \mathbb{R}^3. \quad (13.70)$$

Die zugehörige Funktionalmatrix ist

$$J_T(r, \theta, \varphi) = \begin{pmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}, \quad (13.71)$$

mit der Determinante

$$\det J_T(r, \theta, \varphi) = r^2 \sin \theta. \quad (13.72)$$

Die Transformationsformel (13.63) nimmt die Form an (man beachte $0 \leq \theta \leq \pi$, so dass $\sin \theta \geq 0$)

$$\begin{aligned} \iint\limits_{T(\Omega)} f(x, y, z) dx dy dz \\ = \iiint_{\Omega} f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta dr d\theta d\varphi. \end{aligned} \quad (13.73)$$

Als Beispiel betrachten wir das Volumen und den geometrischen Schwerpunkt des Kugeloktanten

$$\Omega' = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1, x \geq 0, y \geq 0, z \geq 0\}. \quad (13.74)$$

Es ist

$$\Omega' = T(\Omega), \quad \Omega = (0, 1) \times (0, \frac{\pi}{2}) \times (0, \frac{\pi}{2}).$$

Es folgt

$$\begin{aligned} V(\Omega') &= \iiint_{\Omega'} 1 \, dx \, dy \, dz = \iiint_{\Omega} r^2 \sin \theta \, dr \, d\theta \, d\varphi = \int_0^1 \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} r^2 \sin \theta \, dr \, d\theta \, d\varphi \\ &= \int_0^1 r^2 \, dr \cdot \int_0^{\frac{\pi}{2}} \sin \theta \, d\theta \cdot \int_0^{\frac{\pi}{2}} 1 \, d\varphi = \frac{1}{3} \cdot 1 \cdot \frac{\pi}{2} = \frac{\pi}{6}. \end{aligned}$$

Für die x -Koordinate des geometrischen Schwerpunkts gilt

$$x_S = \frac{1}{V(\Omega')} \iiint_{\Omega'} x \, dx \, dy \, dz, \quad (13.75)$$

also

$$\begin{aligned} x_S \cdot V(\Omega') &= \iiint_{\Omega'} x \, dx \, dy \, dz = \iiint_{\Omega} r \sin \theta \cos \varphi \cdot r^2 \sin \theta \, dr \, d\theta \, d\varphi \\ &= \int_0^1 \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} r \sin \theta \cos \varphi \cdot r^2 \sin \theta \, dr \, d\theta \, d\varphi = \int_0^1 r^3 \, dr \cdot \int_0^{\frac{\pi}{2}} \sin^2 \theta \, d\theta \cdot \int_0^{\frac{\pi}{2}} \cos \varphi \, d\varphi \\ &= \frac{1}{4} \cdot \frac{\pi}{4} \cdot 1 = \frac{\pi}{16}, \end{aligned}$$

also

$$x_S = \frac{3}{8}.$$

Aufgrund der Symmetrie der Konfiguration muss $x_S = y_S = z_S$ gelten.

14 Der Fixpunktsatz von Banach

Wir greifen nochmals das Thema ‘‘Fixpunktiteration’’ auf, welches wir bereits im 2. Kapitel angesprochen hatten.

Definition 14.1 (Fixpunkt)

Sei X Menge, $f : X \rightarrow X$. Ein $x \in X$ mit

$$x = f(x) \tag{14.1}$$

heißt *Fixpunkt* von f . Die *Iteration*

$$x_{k+1} = f(x_k), \quad x_0 \in X \text{ gegeben,} \tag{14.2}$$

heißt *Fixpunktiteration*. □

Fixpunktsätze machen Aussagen über Lösungen der Fixpunktgleichung $x = f(x)$, etwa über Existenz und Eindeutigkeit von Fixpunkten, unter gewissen Voraussetzungen an X und f . Fixpunktsätze stellen ein zentrales Werkzeug der Analysis dar.

Der Fixpunktsatz von Banach ist zugeschnitten auf **metrische Räume**, das sind Mengen, in denen zu je zwei Punkten x, y der Abstand $d(x, y)$ dieser beiden Punkte definiert ist. Ist X eine Menge, so verlangen wir von der Abstandsfunktion $d : X \times X \rightarrow \mathbb{R}$, dass für beliebige Punkte $x, y, z \in X$ gilt

$$\begin{aligned} d(x, y) &\geq 0, \quad d(x, y) = 0 \Leftrightarrow x = y, \\ d(x, y) &= d(y, x), \\ d(x, z) &\leq d(x, y) + d(x, z). \end{aligned} \tag{14.3}$$

Definition 14.2 (Metrischer Raum)

Sei X Menge, $d : X \times X \rightarrow \mathbb{R}$. Falls (14.3) gilt für alle $x, y, z \in X$, so heißt d *Metrik* und (X, d) *metrischer Raum*. □

Setzen wir $X = \mathbb{R}^d$ und $d(x, y) = \|x - y\|$, so sehen wir aufgrund der Eigenschaften der Norm, dass der d -dimensionale Raum \mathbb{R}^d ein metrischer Raum ist.

Setzen wir $X = C[a, b]$, definiert als die Menge aller stetigen Funktionen auf dem Intervall $[a, b]$, und

$$d(f, g) = \|f - g\|_\infty := \max_{t \in [a, b]} |f(t) - g(t)|,$$

so erhalten wir ebenfalls einen metrischen Raum, er hat unendliche Dimension im Sinne der Linearen Algebra.

Definition 14.3 (Lipschitz-Stetigkeit)

Seien $(X, d_1), (Y, d_2)$ metrische Räume. Ein $f : X \rightarrow Y$ heißt *Lipschitz-stetig* mit der *Lipschitz-Konstante* L , falls gilt

$$d_2(f(x_1), f(x_2)) \leq L d_1(x_1, x_2), \quad \text{für alle } x_1, x_2 \in X. \tag{14.4}$$

Falls $Y = X$, $d_1 = d_2$ und $L < 1$, so heißt f *Kontraktion*. □

Eine Abbildung f ist also eine Kontraktion, wenn durch ihre Anwendung die Abstände zwischen beliebigen Punkten um mindestens den Faktor $L < 1$ verkleinert werden.

Ist (X, d) ein metrischer Raum, so heißt $x \in X$ **Grenzwert** der Folge $(x_n)_{n \in \mathbb{N}}$, falls

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0$$

gilt.

Definition 14.4 (Vollständigkeit)

Sei (X, d) metrischer Raum. Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt *Cauchyfolge* in (X, d) , falls gilt

$$\lim_{k \rightarrow \infty} \sup_{m \geq 1} d(x_{k+m}, x_k) = 0. \quad (14.5)$$

Der metrische Raum (X, d) heißt *vollständig*, falls jede Cauchyfolge einen Grenzwert hat. \square

Man kann zeigen, dass die beiden Räume \mathbb{R}^d und $C[a, b]$ vollständig sind, wenn sie mit der oben angeführten Metrik versehen werden.

Satz 14.5 (Banachscher Fixpunktsatz, Kontraktionssatz)

Sei (X, d) vollständiger metrischer Raum, $f : X \rightarrow X$ Kontraktion. Dann gilt:

- (i) f hat genau einen Fixpunkt x .
- (ii) Für jedes $x_0 \in X$ konvergiert die durch die Fixpunktiteration $x_{k+1} = f(x_k)$ definierte Folge $(x_k)_{k \in \mathbb{N}}$ gegen x .
- (iii) Ist $L < 1$ eine Lipschitz-Konstante für f , so gilt die Fehlerabschätzung

$$d(x_k, x) \leq \frac{L}{1-L} d(x_{k-1}, x_k). \quad (14.6)$$

(Veranschaulichung der Fixpunktiteration für $X = \mathbb{R}$, siehe Bilder in der Vorlesung.)

Beweis: Sei $x_0 \in X$ beliebig, sei $L < 1$ Lipschitz-Konstante für f . Dann gilt für die durch die Fixpunktiteration definierte Folge $(x_k)_{k \in \mathbb{N}}$

$$d(x_k, x_{k+1}) = d(f(x_{k-1}), f(x_k)) \leq L d(x_{k-1}, x_k), \quad \text{für alle } k \in \mathbb{N},$$

also (mit Induktion)

$$d(x_k, x_{k+1}) \leq L^k d(x_0, x_1), \quad \text{für alle } k \in \mathbb{N},$$

also für alle $k, m \in \mathbb{N}$

$$\begin{aligned} d(x_k, x_{k+m}) &\leq d(x_k, x_{k+1}) + \dots + d(x_{k+m-1}, x_{k+m}) \\ &\leq L^k d(x_0, x_1) + \dots + L^{k+m-1} d(x_0, x_1) = L^k \left(\sum_{j=0}^{m-1} L^j \right) d(x_0, x_1) \\ &\leq \frac{L^k}{1-L} d(x_0, x_1) \quad \text{da } L < 1. \end{aligned}$$

Die Folge $(x_k)_{k \in \mathbb{N}}$ hat also die Eigenschaft (14.5) und konvergiert also wegen der Vollständigkeit von X gegen ein $x \in X$. Aus der Lipschitz-Stetigkeit von f folgt wegen

$$0 \leq d(f(x_k), f(x)) \leq Ld(x_k, x),$$

dass $f(x_k)$ gegen $f(x)$ konvergiert, und daher

$$x = \lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} f(x_{k-1}) = f(x).$$

Somit ist x tatsächlich ein Fixpunkt von f . Ist y ebenfalls Fixpunkt von f , so gilt

$$d(x, y) = d(f(x), f(y)) \leq Ld(x, y),$$

also $(1 - L)d(x, y) \leq 0$, also $d(x, y) = 0$ und damit $x = y$. Zum Beweis von (14.6) zeigen wir wie oben

$$d(x_k, x_{k+m}) \leq L \left(\sum_{j=0}^{m-1} L^j \right) d(x_{k-1}, x_k) \leq \frac{L}{1-L} d(x_{k-1}, x_k),$$

also

$$d(x_k, x) = \lim_{m \rightarrow \infty} d(x_k, x_{k+m}) \leq \frac{L}{1-L} d(x_{k-1}, x_k).$$

□

Man kann den Banachschen Fixpunktsatz benutzen, um zu zeigen, dass das in Kapitel 12 behandelte Anfangswertproblem

$$y' = f(t, y), \quad y(t_0) = y_0,$$

zumindest auf einem hinreichend kleinen Intervall um t_0 eine eindeutige Lösung hat, falls die rechte Seite f Lipschitz-stetig ist. Man kann ihn außerdem benutzen, um zu zeigen, dass viele der zur Lösung großer linearer Gleichungssysteme

$$Ax = b$$

verwendeten iterativen Verfahren tatsächlich gegen eine Lösung konvergieren, und zwar mit einer Konvergenzrate, die sich letztlich aus der Fehlerabschätzung (14.6) ergibt.