

Mathematik II für Studierende der Informatik
(Analysis und Lineare Algebra)

THOMAS ANDREAE

Sommersemester 2014

(Stand: 24. Juli 2014)

Vorwort

Das vorliegende Skript umfasst im Wesentlichen den Stoff der Vorlesung „Mathematik II für Studierende der Informatik (Analysis und Lineare Algebra)“ an der Universität Hamburg im Sommersemester 2014. Auf die Behandlung der Kapitel 8 und 9 wurde aus Zeitgründen verzichtet, was insofern vertretbar war, da das Allernötigste zur Linearen Algebra bereits im Wintersemester 2013/14 Stoff der Vorlesung „Mathematik I für Studierende der Informatik“ war.

Den ersten Abschnitten des Skripts liegt größtenteils das im Literaturverzeichnis aufgeführte Lehrbuch von K. Habetha zugrunde. Beginnend mit Abschnitt 1.4 habe ich häufig auch andere Literatur herangezogen: Von sämtlichen im Literaturverzeichnis aufgeführten Lehrbüchern habe ich in Bezug auf Stoffauswahl und Darstellung profitiert. Hervorzuheben ist neben den Büchern von E. Behrends, H. Heuser und B. L. van der Waerden vor allem das 2-bändige Werk von W. Luh, an das ich besonders häufig angeknüpft habe.

Wie schon das Skript „Mathematik I für Studierende der Informatik“ wurde auch das vorliegende Skript von Steven Köhler in \LaTeX gesetzt, wofür ich mich herzlich bedanke.

Hamburg, Juli 2014

Thomas Andreae

Inhaltsverzeichnis

1	Reelle Zahlen, Grenzwerte und Stetigkeit	5
1.1	Die Axiome der reellen Zahlen	5
1.2	Ungleichungen, Betrag und Abstand	6
1.3	Konvergenz	9
1.4	Funktionsgrenzwerte	21
1.5	Stetigkeit	24
1.6	Ergänzungen zu den Abschnitten 1.4 und 1.5	27
1.6.1	Einseitige Grenzwerte	27
1.6.2	Einseitige Stetigkeit	27
1.6.3	ε, δ -Definition der Stetigkeit	27
1.6.4	Stetigkeit und Grenzwerte	28
2	Differentialrechnung	29
2.1	Die Ableitung einer Funktion	29
2.2	Ableitungsregeln	32
2.3	Die Funktionen a^x , $\log_a x$, x^r und ihre Ableitungen	36
2.3.1	Die Exponentialfunktion a^x	38
2.3.2	Die Logarithmusfunktion $\log_a x$	39
2.3.3	Die Ableitung der Funktion $\ln x$	41
2.3.4	Die Ableitung der Funktion $\log_a x$	42
2.3.5	Die Ableitung der Funktion e^x	42
2.3.6	Die Ableitung der Funktion a^x	42
2.3.7	Die Ableitung von allgemeinen Potenzfunktionen	43
2.3.8	Logarithmisches Differenzieren	44
2.4	Die geometrische Bedeutung der ersten und zweiten Ableitung, Extremstellen und Kurvendiskussion	45
2.4.1	Beispiele zur Kurvendiskussion	51
2.4.2	Der Zwischenwertsatz	55
2.5	Trigonometrische Funktionen	56
2.5.1	Definition der Trigonometrischen Funktionen	56
2.5.2	Die Ableitungen der Trigonometrischen Funktionen	58
2.5.3	Die Umkehrfunktionen der Trigonometrischen Funktionen	60
2.6	Das Newtonsche Verfahren	62
2.7	Die Regeln von de l'Hospital	63
2.7.1	Der Typ $\frac{0}{0}$	64
2.7.2	Der Typ $\frac{\infty}{\infty}$	65
2.7.3	Der Typ $0 \cdot \infty$	67
2.7.4	Der Typ $\infty - \infty$	67
2.7.5	Die Typen 0^0 , 1^∞ und ∞^0	68
3	Integralrechnung	69
3.1	Berechnung der Fläche unter einer Kurve	69
3.2	Das bestimmte Integral	72
3.3	Der Fundamentalsatz der Differential- und Integralrechnung	77
3.4	Berechnung von Integralen	80
3.4.1	Partielle Integration	83

3.4.2	Integration durch Substitution	84
3.4.3	Weitere Beispiele zur Substitutionsregel	90
3.4.4	Integration rationaler Funktionen	96
3.5	Interpolation und Interpolationsfehler	104
3.5.1	Lagrange-Interpolation	104
3.5.2	Newton-Interpolation	106
3.5.3	Lineare Interpolation	107
3.6	Numerische Integration (Trapezregel)	108
3.7	Uneigentliche Integrale	112
4	Reihen	116
4.1	Konvergenzkriterien für Reihen	116
4.1.1	Das Majorantenkriterium	117
4.1.2	Das Quotientenkriterium	119
4.1.3	Das Wurzelkriterium	120
4.1.4	Das Leibnizsche Kriterium	120
4.2	Potenzreihen	121
4.3	Der Satz von Taylor	128
4.4	Taylorreihen	134
5	Differentialrechnung für Funktionen mehrerer Variablen	138
5.1	Partielle Ableitungen	138
5.2	Extremstellen, Gradient und Hessesche Matrix	141
5.3	Die geometrische Bedeutung der partiellen Ableitungen und des Gradienten	146
5.4	Extrema mit Nebenbedingungen und Lagrangesche Multiplikatorenregel	150
6	Integration von Funktionen mehrerer Variablen	154
6.1	Definition zweidimensionaler Integrale	154
6.2	Berechnung zweidimensionaler Integrale	157
7	Komplexe Zahlen	161
7.1	Definition der komplexen Zahlen	161
7.2	Veranschaulichung der komplexen Zahlen als Punkte der Gaußschen Zahlenebene	163
7.3	Polarkoordinaten	168
7.4	Quadratische Gleichungen	170
7.5	Der Fundamentalsatz der Algebra	172
7.6	Die n-ten Wurzeln einer komplexen Zahl	173
7.7	Die komplexe Exponentialfunktion	174
8	Vektorräume über einem Körper K	177
8.1	Vorbemerkungen	177
8.2	Der Begriff des Vektorraums über einem Körper	177
8.3	Einige Folgerungen aus den Axiomen	181
8.4	Untervektorräume	182
8.5	Lineare Unabhängigkeit	183
8.6	Die Dimension eines Vektorraums	190
8.7	Vektoren im \mathbb{R}^2 und \mathbb{R}^3	193
8.7.1	Addition für $V = \mathbb{R}^2$	193
8.7.2	Skalare Multiplikation für $V = \mathbb{R}^2$	194
8.7.3	Linear abhängige und linear unabhängige Vektoren für $V = \mathbb{R}^2$	194
8.7.4	Bestimmung der Unterräume des \mathbb{R}^2	195
8.7.5	Vektoren im \mathbb{R}^3	196
9	Lineare Abbildungen und Matrizen	198
9.1	Lineare Abbildungen	198

9.2	Der Zusammenhang zwischen linearen Abbildungen und Matrizen	205
9.3	Der Rang einer Matrix	211
9.4	Einige Beispiele	213
9.5	Eigenwerte und Eigenvektoren	215
10	Literatur	220

1 Reelle Zahlen, Grenzwerte und Stetigkeit

1.1 Die Axiome der reellen Zahlen

In diesem Abschnitt führen wir die reellen Zahlen mithilfe der axiomatischen Methode ein, d.h., wir gehen von einer Menge aus, die wir mit \mathbb{R} bezeichnen und deren Elemente wir *reelle Zahlen* nennen. Durch Axiome legen wir fest, welche Grundeigenschaften diese Menge haben soll. Aus diesen Grundeigenschaften leitet man dann alle weiteren Aussagen über reelle Zahlen her. Wir unterscheiden drei Axiomengruppen für die reellen Zahlen: die Körperaxiome, die Anordnungsaxiome und das Vollständigkeitsaxiom.

(I) Die Körperaxiome.

In \mathbb{R} gebe es zwei binäre Operationen, die mit $+$ und \cdot bezeichnet seien und für die Folgendes gelten soll:

(I.1) Assoziativgesetze: $a + (b + c) = (a + b) + c$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

(I.2) Kommutativgesetze: $a + b = b + a$, $a \cdot b = b \cdot a$

(I.3) Distributivgesetz: $a \cdot (b + c) = a \cdot b + a \cdot c$

(I.4) Existenz neutraler Elemente: Es gibt in \mathbb{R} verschiedene Elemente 0 und 1 derart, dass für jedes $a \in \mathbb{R}$ gilt: $a + 0 = a$ und $a \cdot 1 = a$.

(I.5) Existenz inverser Elemente: Zu jedem $a \in \mathbb{R}$ gibt es ein Element $-a$ aus \mathbb{R} mit $a + (-a) = 0$; ferner gibt es zu jedem $a \in \mathbb{R}$, $a \neq 0$, ein Element a^{-1} mit $a \cdot a^{-1} = 1$.

(II) Die Anordnungsaxiome.

In \mathbb{R} gebe es eine binäre Relation, die mit $<$ bezeichnet sei und für die Folgendes gelten soll:

(II.1) Für jedes Paar reeller Zahlen a , b ist genau eine der drei Möglichkeiten $a < b$, $a = b$ oder $b < a$ richtig.

(II.2) Transitivität: Aus $a < b$ und $b < c$ folgt $a < c$.

(II.3) Aus $a < b$ folgt $a + c < b + c$ für alle $c \in \mathbb{R}$.

(II.4) Aus $a < b$ und $0 < c$ folgt $a \cdot c < b \cdot c$.

Gilt entweder $a < b$ oder $a = b$, so schreibt man $a \leq b$.

Sind die Axiome der Gruppen (I) und (II) erfüllt, so spricht man von einem *angeordneten Körper*. Ebenso wie \mathbb{R} bilden auch die rationalen Zahlen \mathbb{Q} einen angeordneten Körper. Veranschaulicht man sich \mathbb{Q} auf der Zahlengeraden, so füllt \mathbb{Q} die Zahlengerade nicht lückenlos aus (siehe DM-Skript, Abschnitte 2.2 und 2.3). Von den reellen Zahlen verlangt man dagegen, dass man sie sich als Menge sämtlicher Punkte auf der Zahlengeraden veranschaulichen kann. Daher braucht man noch ein weiteres Axiom, das die Vorstellung von der „Lückenlosigkeit der Zahlengeraden“ widerspiegelt; dies leistet das Vollständigkeitsaxiom.

(III) Vollständigkeitsaxiom.

A und B seien nichtleere Teilmengen von \mathbb{R} , so dass $a < b$ für alle $a \in A$ und $b \in B$ gilt. Dann existiert (mindestens) ein $c \in \mathbb{R}$, so dass $a \leq c \leq b$ für alle $a \in A$ und $b \in B$ gilt.

Auf der Zahlengeraden bedeutet $a < b$, dass a links von b liegt. Das Vollständigkeitsaxiom besagt also, wenn man es sich auf der Zahlengeraden veranschaulicht: Liegt jedes Element von A links von jedem Element von B , so gibt es immer noch mindestens ein $c \in \mathbb{R}$, so dass $a \leq c \leq b$ für alle $a \in A$ und $b \in B$ gilt. Dabei ist es auch möglich, dass $c \in A$ oder $c \in B$ gilt: Ist beispielsweise $A = \{a \in \mathbb{R} : a < 1\}$, $B = \{b \in \mathbb{R} : 1 \leq b\}$ und wählt man $c = 1 \in B$, so gilt $a \leq c \leq b$ für alle $a \in A$ und $b \in B$. (Wegen $a < b$

für alle $a \in A$ und $b \in B$ kann natürlich niemals $c \in A$ und $c \in B$ gelten.)

Es sei darauf hingewiesen, dass das Vollständigkeitsaxiom nicht erfüllt ist, wenn wir anstelle von \mathbb{R} die rationalen Zahlen \mathbb{Q} betrachten. Für \mathbb{Q} gilt das Vollständigkeitsaxiom nicht. Wir skizzieren, weshalb dies so ist: Es sei $A = \{x \in \mathbb{Q} : x \geq 0 \text{ und } x^2 < 2\}$ und $B = \{x \in \mathbb{Q} : x \geq 0 \text{ und } 2 < x^2\}$. Man stellt unschwer fest, dass A und B nichtleere Teilmengen von \mathbb{Q} sind, für die $a < b$ für alle $a \in A$ und $b \in B$ gilt. Es gibt jedoch kein $c \in \mathbb{Q}$, so dass $a \leq c \leq b$ für alle $a \in A$ und $b \in B$ gilt, da für ein solches c (wie man sich unschwer überlegt) $c^2 = 2$ gelten müsste, im Widerspruch zu $c \in \mathbb{Q}$. Anders gesagt: Ein solches c könnte nur $\sqrt{2}$ sein, es gilt aber $\sqrt{2} \notin \mathbb{Q}$.

Aus den obigen Axiomen lassen sich sämtliche Rechenregeln für die vier Grundrechenarten und für die $<$ -Beziehung herleiten. Wir wollen dies nicht in jedem Einzelfall durchführen. Da die Rechenregeln für die vier Grundrechenarten alle aus der Schule geläufig sind, wollen wir hier auf eine Herleitung verzichten¹. Wir werden im Folgenden immer von so einfachen Regeln wie etwa $a(-b) = (-a)b = -ab$ ohne Weiteres Gebrauch machen. Da die Rechenregeln für die $<$ -Beziehung im Allgemeinen etwas weniger geläufig sind, werden wir im nächsten Abschnitt die wichtigsten Regeln für das Rechnen mit Ungleichungen aus den Axiomen herleiten.

Eine weitere Eigenschaft der reellen Zahlen, die man aus den obigen Axiomen herleiten kann, auf deren Beweis wir jedoch verzichten wollen, ist die *Archimedische Eigenschaft*:

$$\text{Zu jeder reellen Zahl } r \text{ gibt es eine natürliche Zahl } n, \text{ für die } r < n \text{ gilt.} \quad (1.1)$$

Veranschaulicht man sich \mathbb{R} wie üblich auf der Zahlengeraden, so ist (1.1) anschaulich klar. Es sei darauf hingewiesen, dass (1.1) oft auch „Archimedisches Axiom“ genannt wird. Dies erklärt sich dadurch, dass (1.1) in alternativen Axiomensystemen für die reellen Zahlen als Axiom auftritt.

1.2 Ungleichungen, Betrag und Abstand

Im folgenden Satz sind die neben den Anordnungsaxiomen wichtigsten Regeln für das Rechnen mit Ungleichungen zusammengestellt.

Satz 1.

- (a) $a < b, c < d \Rightarrow a + c < b + d$ („Ungleichungen kann man addieren.“)
- (b) $a < b, c < 0 \Rightarrow ac > bc$
- (c) $0 < 1$
- (d) $0 < a < b \Rightarrow 0 < \frac{1}{b} < \frac{1}{a}$
- (e) Für $a > 0, b > 0$ ist $a < b$ genau dann erfüllt, wenn $a^2 < b^2$.

Beweis.

- (a) Aus $a < b$ und $c < d$ folgt nach Axiom (II.3) $a + c < b + c$ und $b + c < b + d$. Nach (II.2) folgt dann die Behauptung.
- (b) Aus $c < 0$ folgt $0 < -c$ (nach (II.3) durch Addition von $-c$). Nach (II.4) folgt somit aus $a < b$, dass $a(-c) < b(-c)$ gilt. Addiert man auf beiden Seiten $ac + bc$, so folgt $bc < ac$.
- (c) Wegen (I.4) gilt $0 \neq 1$. Wäre $1 < 0$, so würde aus (b) $1 \cdot 1 > 0 \cdot 1$ folgen, d.h. $1 > 0$. Wir hätten also gleichzeitig $1 < 0$ und $1 > 0$ im Widerspruch zu (II.1). Es gilt also weder $1 < 0$ noch $1 = 0$. Nach

¹Für eine vollständige Herleitung aller dieser Regeln verweisen wir auf die zahlreichen Lehrbücher der Analysis; exemplarisch seien die Bücher von Behrends, Forster, Heuser und Königsberger genannt (siehe das Literaturverzeichnis am Ende dieses Skripts). Wer Weiterführendes über die reellen Zahlen wissen möchte (alternative Axiomensysteme, konstruktive Einführung der reellen Zahlen, Historisches), dem sei neben den genannten Lehrbüchern der Analysis das Buch von Ebbinghaus empfohlen.

(II.1) muss $0 < 1$ gelten.

- (d) Aus $0 < a$ und $0 < b$ folgt nach (II.4) $0 < ab$. Hieraus folgt $0 < \frac{1}{ab}$. (Denn wäre $\frac{1}{ab} < 0$, so könnte man mit ab multiplizieren und erhielte $1 < 0$ im Widerspruch zu (c).) Multipliziert man $0 < a < b$ mit $\frac{1}{ab}$, so folgt $0 < \frac{1}{b} < \frac{1}{a}$.
- (e) Ist $0 < a < b$, so folgt $aa < ab$ und $ab < bb$. Also gilt $a^2 < b^2$ nach (II.2). Umgekehrt gelte nun $a^2 < b^2$. Dann kann nicht $a = b$ gelten, da hieraus $a^2 = b^2$ folgen würde; es kann auch nicht $b < a$ gelten, da hieraus (wie oben) $bb < ab$ und $ab < aa$, also $b^2 < a^2$ folgen würde. Also gilt $a < b$. \square

Die obige Aussage (c), dass $0 < 1$ gilt, ist nur scheinbar selbstverständlich. Sie muss bewiesen werden, da in den Axiomen nur gesagt wird, dass $0 \neq 1$ gelten soll (Axiom (I.4)). Aussage (e) lässt sich auch so ausdrücken: Für alle $a, b > 0$ gilt $a < b \Leftrightarrow a^2 < b^2$.

Aus den obigen Regeln für das Rechnen mit der $<$ -Beziehung kann man leicht entsprechende Regeln für das Rechnen mit \leq gewinnen: Man braucht nur daran zu denken, dass $a \leq b$ nichts weiter bedeutet als $a < b$ oder $a = b$. Der Vollständigkeit halber führen wir einige dieser Regeln auf.

- (a) $a \leq b, b \leq c \Rightarrow a \leq c$
(b) $a \leq b, c \leq d \Rightarrow a + c \leq b + d$
(c) $a \leq b, c > 0 \Rightarrow ac \leq bc$
(d) $a \leq b, c < 0 \Rightarrow ac \geq bc$

Definition.

Der *absolute Betrag* $|a|$ einer reellen Zahl a wird definiert durch:

$$|a| := \begin{cases} a, & \text{falls } a \geq 0 \\ -a, & \text{falls } a < 0 \end{cases}$$

Statt „absoluter Betrag“ sagt man meistens kurz „Betrag“. Anschauliche Bedeutung: $|a|$ gibt den Abstand der Zahl a vom Nullpunkt an. Die wichtigsten Eigenschaften des absoluten Betrages sind im folgenden Satz zusammengestellt.

Satz 2.

Für alle $a, b \in \mathbb{R}$ gilt:

- (a) $|a| \geq 0$; $|a| = 0$ gilt genau dann, wenn $a = 0$.
(b) $|ab| = |a||b|$
(c) $\left| \frac{a}{b} \right| = \frac{|a|}{|b|}$ ($b \neq 0$)
(d) $|a + b| \leq |a| + |b|$ (Dreiecksungleichung)
(e) $|a - b| \geq |a| - |b|$

Beweis.

- (a) Für $a > 0$ ist $|a| = a > 0$, für $a < 0$ ist $|a| = -a > 0$ und für $a = 0$ ist $|a| = a = 0$.
(b) Für $a \geq 0, b \geq 0$ ist $ab \geq 0$ und daher $|ab| = ab = |a||b|$. Für $a \geq 0, b < 0$ ist $ab \leq 0$ und daher $|ab| = -ab = a(-b) = |a||b|$. Den Fall $a < 0, b \geq 0$ behandelt man analog. Für $a < 0, b < 0$ ist $ab > 0$ und daher $|ab| = ab = (-a)(-b) = |a||b|$.
(c) Wegen $a = \frac{a}{b}b$ folgt nach (b) $|a| = \left| \frac{a}{b}b \right| = \frac{|a|}{|b|}|b|$. Hieraus folgt $\frac{|a|}{|b|} = \frac{|a|}{|b|}$.
(d) Aus der Definition des Betrages folgt $a \leq |a|, b \leq |b|, -a \leq |a|$ und $-b \leq |b|$. Nach Regel (b) für \leq folgt hieraus $a + b \leq |a| + |b|$ und $-(a + b) \leq |a| + |b|$. Hieraus folgt für $a + b \geq 0$, dass $|a + b| = a + b \leq |a| + |b|$ gilt; für $a + b < 0$ folgt $|a + b| = -(a + b) \leq |a| + |b|$.
(e) Nach (d) gilt $|a| = |b + a - b| \leq |b| + |a - b|$. Hieraus folgt $|a| - |b| \leq |a - b|$. \square

Die Dreiecksungleichung lässt sich mittels vollständiger Induktion auf Summen mit n Summanden verallgemeinern:

$$\left| \sum_{i=1}^n a_i \right| \leq \sum_{i=1}^n |a_i|. \quad (1.2)$$

Beweis. Wir beweisen (1.2) durch vollständige Induktion.

(I) Induktionsanfang:

Für $n = 1$ besagt (1.2) $|a_1| \leq |a_1|$, was klarerweise richtig ist.

(II) Induktionsschritt:

Wir nehmen an, dass (1.2) für ein $n \in \mathbb{N}$ richtig ist (Induktionsannahme). Dann gilt:

$$\left| \sum_{i=1}^{n+1} a_i \right| = \left| \sum_{i=1}^n a_i + a_{n+1} \right| \stackrel{\text{Satz 2 (d)}}{\leq} \left| \sum_{i=1}^n a_i \right| + |a_{n+1}| \stackrel{\text{Induktionsannahme}}{\leq} \sum_{i=1}^n |a_i| + |a_{n+1}| = \sum_{i=1}^{n+1} |a_i| \quad \square$$

Beispiele zum Rechnen mit Ungleichungen und Beträgen.

Wir wollen die Regel (b) aus Satz 1 und die entsprechende Regel (d) für das \leq -Zeichen besonders hervorheben: *Nach diesen Regeln kehrt sich das $<$ -Zeichen bzw. das \leq -Zeichen um, wenn man beide Seiten einer Ungleichung mit einer negativen Zahl multipliziert.* Wenn beide Seiten einer Ungleichung mit einer positiven Zahl multipliziert werden, kehren sich die Relationszeichen $<$ und \leq hingegen nicht um (siehe Axiom (II.4) bzw. Regel (c) für das \leq -Zeichen). Dies führt häufig dazu, dass man in Rechnungen Fallunterscheidungen vornehmen muss.

Hierzu ein **Beispiel**.

Es sollen diejenigen $x \in \mathbb{R}$ bestimmt werden, für die Folgendes gilt:

$$\frac{3x+1}{x} < 4. \quad (*)$$

Lösung.

Für $x = 0$ gilt (*) nicht, da der Ausdruck $\frac{3x+1}{x}$ dann nicht definiert ist. („Der Nenner darf nicht gleich Null sein.“)

1. Fall: $x > 0$

In diesem Fall ist der Nenner positiv und wir erhalten:

$$\frac{3x+1}{x} < 4 \iff 3x+1 < 4x \iff 1 < x.$$

2. Fall: $x < 0$

In diesem Fall ist der Nenner negativ und wir erhalten:

$$\frac{3x+1}{x} < 4 \iff 3x+1 > 4x \iff 1 > x.$$

Zusammenfassung.

Ist $x > 0$, so gilt (*) genau für diejenigen x , die $1 < x$ erfüllen; ist $x < 0$, so gilt (*) immer, da alle negativen x die ermittelte Ungleichung $1 > x$ erfüllen. Bezeichnen wir die Menge aller $x \in \mathbb{R}$, die (*) erfüllen, mit L , so gilt also:

$$L = \{x \in \mathbb{R} : x > 1\} \cup \{x \in \mathbb{R} : x < 0\}.$$

Benutzt man die *Intervallschreibweise* (vgl. Abschnitt 1.4), so kann man die Lösungsmenge L auch wie folgt darstellen:

$$L = (-\infty, 0) \cup (1, \infty).$$

Auch wenn man mit Beträgen rechnet, hat man in ähnlicher Weise Fallunterscheidungen vorzunehmen. Das liegt daran, dass Ausdrücke wie $|x|$ oder beispielsweise $|2-x|$ Unterschiedliches bedeuten, je nachdem, ob $x \geq 0$ oder $x < 0$ bzw. $2-x \geq 0$ oder $2-x < 0$ gilt:

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

$$|2-x| := \begin{cases} 2-x & , \text{ falls } 2-x \geq 0 \\ x-2 & , \text{ falls } 2-x < 0 \end{cases}$$

Statt $2-x \geq 0$ und $2-x < 0$ hätte man hier natürlich auch die gleichwertigen Bedingungen $x \leq 2$ und $2 < x$ hinschreiben können. Das Rechnen mit Beträgen wird in den Übungen anhand von Beispielen geübt werden.

Definition.

Unter dem *Abstand* zweier reeller Zahlen a und b versteht man die Zahl $|a-b|$.

Einige **Beispiele** hierzu:

- (i) $a = 2, b = 8$: Dann gilt $|a-b| = |2-8| = |-6| = 6$.
- (ii) $a = -3, b = 4$: Dann gilt $|a-b| = |-3-4| = |-7| = 7$.
- (iii) $a = -3, b = -12$: Dann gilt $|a-b| = |-3-(-12)| = |9| = 9$.

Aufgabe: Berechnen Sie den Abstand der beiden Zahlen $\frac{45}{11}$ und $\frac{86}{21}$.

Für je zwei reelle Zahlen a, b gilt klarerweise $|a-b| = |b-a|$. (*Nachweis:* $|a-b| = |-(a-b)| = |b-a|$.) Aus der Dreiecksungleichung für den Betrag ergibt sich außerdem folgende Beziehung, die ebenfalls als *Dreiecksungleichung* (für den Abstand) bezeichnet wird:

$$|a-b| \leq |a-c| + |b-c| \quad (\text{für alle } a, b, c \in \mathbb{R}).$$

Beweis.

$$|a-b| = |a-c+c-b| \leq |a-c| + |c-b| = |a-c| + |b-c|$$

1.3 Konvergenz

Definition.

Unter einer *Folge* von Elementen aus einer Menge M versteht man eine Abbildung der natürlichen Zahlen \mathbb{N} in M . Eine Folge ist also eine Abbildung:

$$a : \mathbb{N} \rightarrow M.$$

Der n zugeordnete Funktionswert $a(n)$ wird üblicherweise mit a_n bezeichnet. Man nennt a_n ein *Glied* der Folge. Wir interessieren uns hier und in den folgenden Abschnitten in erster Linie für *Folgen reeller Zahlen*, d.h. Folgen vom Typ $a : \mathbb{N} \rightarrow \mathbb{R}$. Alle betrachteten Folgen sollen Folgen reeller Zahlen sein – es sei denn, es ist ausdrücklich etwas anderes gesagt. Man beschreibt Folgen im Allgemeinen durch die Angabe ihrer Glieder a_1, a_2, a_3, \dots

Übliche Schreibweisen zur Darstellung einer Folge:

$$(a_1, a_2, a_3, \dots), \quad (a_n)_{n \in \mathbb{N}} \quad \text{oder} \quad (a_n).$$

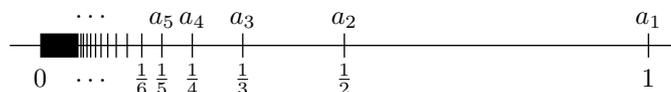
Varianten: Oft werden auch Abbildungen $a : \mathbb{N} \cup \{0\} \rightarrow M$ betrachtet. Auch hier spricht man von einer Folge; es handelt sich um eine Folge mit den Gliedern a_0, a_1, a_2, \dots . Weitere Varianten sind möglich, etwa Folgen vom Typ a_2, a_3, a_4, \dots

Reelle Zahlenfolgen können sich recht unterschiedlich verhalten, wie man an folgenden drei Beispielen sieht.

Beispiel 1. Die betrachtete Folge sei $1, \frac{1}{2}, \frac{1}{3}, \dots$

Ihre Glieder sind: $a_1 = 1, a_2 = \frac{1}{2}, a_3 = \frac{1}{3}, \dots$; oder allgemein: $a_n = \frac{1}{n}$.

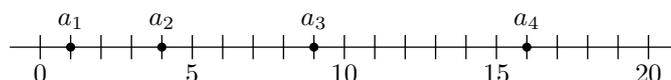
Anschauliche Darstellung dieser Folge:



Beispiel 2. Die betrachtete Folge sei $1^2, 2^2, 3^2, 4^2, \dots$

Ihre Glieder sind also $a_1 = 1^2, a_2 = 2^2, a_3 = 3^2, \dots$; allgemein: $a_n = n^2$.

Anschauliche Darstellung dieser Folge:



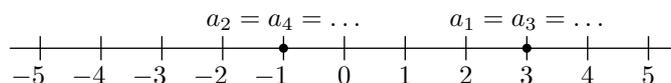
Beispiel 3. Die betrachtete Folge sei $3, -1, 3, -1, 3, \dots$

Die Folge oszilliert also zwischen 3 und -1 . Ihre Glieder sind $a_1 = 3, a_2 = -1, a_3 = 3, a_4 = -1, \dots$

Allgemein:

$$a_n = \begin{cases} 3 & , \text{ falls } n \text{ ungerade} \\ -1 & , \text{ falls } n \text{ gerade} \end{cases}$$

Anschauliche Darstellung dieser Folge:



Definition.

Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen *konvergiert gegen eine reelle Zahl* a , wenn es für jede reelle Zahl $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $|a_n - a| < \varepsilon$ für alle $n \geq N$ gilt.

Schreibweisen: $(a_n) \rightarrow a$ für $n \rightarrow \infty$; oder auch einfach nur $(a_n) \rightarrow a$; oder noch einfacher: $a_n \rightarrow a$.

Sprechweisen: (a_n) besitzt den *Grenzwert* a . Konvergiert eine Folge (a_n) gegen eine reelle Zahl a , so nennt man (a_n) *konvergent*; andernfalls heißt (a_n) *divergent*.

Bestimmte divergente Folgen nennt man „uneigentlich konvergent“ (genauer gesagt spricht man von „uneigentlich konvergent gegen ∞ “ und „uneigentlich konvergent gegen $-\infty$ “). Was das genau bedeuten soll, wird in der nachfolgenden Definition dargelegt.

Definition.

Eine Folge $(a_n)_{n \in \mathbb{N}}$ *konvergiert uneigentlich gegen* ∞ , falls es zu jeder reellen Zahl $r > 0$ ein $N \in \mathbb{N}$ gibt, so dass $a_n > r$ für alle $n \geq N$ gilt.

Analog definiert man: $(a_n)_{n \in \mathbb{N}}$ *konvergiert uneigentlich gegen* $-\infty$, falls es zu jeder reellen Zahl $r < 0$ ein $N \in \mathbb{N}$ gibt, so dass $a_n < r$ für alle $n \geq N$ gilt.

Schreibweisen: $(a_n) \rightarrow \infty$ für $n \rightarrow \infty$, $(a_n) \rightarrow -\infty$ für $n \rightarrow \infty$, sowie vereinfachte Schreibweisen (wie oben).

Sprechweise: (a_n) besitzt den *uneigentlichen Grenzwert* ∞ (bzw. $-\infty$).

Beispiele:

1. Für die Folge (a_n) mit $a_n = \frac{1}{n}$ aus Beispiel 1 ist zu vermuten, dass der Grenzwert 0 ist. Es ist $|a_n - 0| = \frac{1}{n}$; zum Beweis, dass $(a_n) \rightarrow 0$ gilt, müssen wir also zu jedem gegebenen $\varepsilon > 0$ ein $N \in \mathbb{N}$ angeben, so dass $\frac{1}{n} < \varepsilon$ für alle $n \geq N$ gilt. Wir wählen $N > \frac{1}{\varepsilon}$. Für alle $n \geq N$ gilt dann $n > \frac{1}{\varepsilon}$, woraus $\frac{1}{n} < \varepsilon$ folgt.
2. Die Folge (a_n) mit $a_n = \frac{n-1}{n}$ hat den Grenzwert 1. Denn: $|a_n - 1| = \left| \frac{n-1}{n} - 1 \right| = \left| \frac{-1}{n} \right| = \frac{1}{n}$. Hieraus folgt $(a_n) \rightarrow 1$ wie in 1.
3. Für die Folge (a_n) mit $a_n = \sqrt{n}$ gilt $(a_n) \rightarrow \infty$. Denn: Es sei $r > 0$ eine beliebige reelle Zahl. Wir müssen ein $N \in \mathbb{N}$ angeben, so dass $a_n > r$ für alle $n \geq N$ gilt. Wir wählen $N > r^2$. Für alle $n \geq N$ gilt dann $n > r^2$, woraus $a_n = \sqrt{n} > r$ folgt.

Sie kennen die Symbole \forall und \exists :

- \forall heißt *Allquantor* (Bedeutung: für alle).
- \exists heißt *Existenzquantor* (Bedeutung: es gibt).

Unter Benutzung dieser Symbole können wir die Definition der Konvergenz wie folgt schreiben: Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen *konvergiert gegen eine reelle Zahl* a , falls gilt:

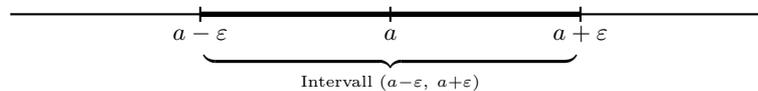
$$\forall \varepsilon > 0 : \exists N \in \mathbb{N}, \text{ so dass } |a_n - a| < \varepsilon \quad \forall n \geq N. \quad (*)$$

Dasselbe noch etwas kompakter aufgeschrieben:

$$\forall_{\varepsilon > 0} \exists_{N \in \mathbb{N}} \forall_{n \geq N} |a_n - a| < \varepsilon.$$

Der in der Definition der Konvergenz auftretende Term $|a_n - a|$ bezeichnet den Abstand zwischen a_n und a ; $|a_n - a| < \varepsilon$ bedeutet somit, dass der Abstand von a_n und a kleiner als ε ist. In Worten kann man (*) also auch so ausdrücken:

Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für sämtliche Folgenglieder $a_N, a_{N+1}, a_{N+2}, \dots$ gilt: Der Abstand von a ist für jedes dieser Folgenglieder kleiner als ε .



a_N und sämtliche nachfolgenden Folgenglieder (d.h. a_{N+1}, a_{N+2}, \dots)
liegen im Intervall $(a - \varepsilon, a + \varepsilon)$

Dasselbe, nur etwas weniger formal ausgedrückt: *Der Abstand $|a_n - a|$ der Folgenglieder zum Grenzwert a wird beliebig klein, wenn n gegen unendlich geht.*

Es gibt Folgen mit und ohne Grenzwert. Aber es gilt:

$$\text{Eine Folge } (a_n) \text{ kann höchstens einen Grenzwert besitzen.} \quad (1.3)$$

Beweis. Angenommen es gelte $a_n \rightarrow a$ und $a_n \rightarrow b$ mit $a \neq b$. Wir wählen $\varepsilon = \frac{1}{2}|a - b|$. Nach Definition der Konvergenz gibt es ein N_1 und ein N_2 , so dass $|a_n - a| < \varepsilon$ für alle $n \geq N_1$ und $|a_n - b| < \varepsilon$ für alle $n \geq N_2$ gilt. Für $n \geq \max(N_1, N_2)$ gilt also:

$$|a - b| = |a - a_n + a_n - b| \leq |a - a_n| + |a_n - b| < 2\varepsilon = |a - b|.$$

Dies ist ein Widerspruch. \square

Eine nützliche Eigenschaft für die Bestimmung von Grenzwerten liefert der folgende Satz.

Satz 3.

Für die Glieder a_n einer Folge (a_n) gelte $a_n > 0$. Dann gilt $a_n \rightarrow \infty$ genau dann, wenn $\frac{1}{a_n} \rightarrow 0$ gilt.

Beweis.

1. Es gelte $a_n \rightarrow \infty$. Um $\frac{1}{a_n} \rightarrow 0$ nachzuweisen, betrachten wir ein beliebiges $\varepsilon > 0$. Wir setzen $r = \frac{1}{\varepsilon}$. Wegen $a_n \rightarrow \infty$ gibt es ein N , so dass $a_n > r = \frac{1}{\varepsilon}$ für alle $n \geq N$ gilt. Es folgt $\frac{1}{a_n} < \varepsilon$ für alle $n \geq N$. Also gilt $\frac{1}{a_n} \rightarrow 0$.
2. Umgekehrt kann man aus $\frac{1}{a_n} \rightarrow 0$ auf ganz ähnliche Art auch $a_n \rightarrow \infty$ folgern. (Beweis zur Übung empfohlen.) \square

Eine einfache (aber nützliche) Ungleichung, die sich unmittelbar aus dem Binomischen Lehrsatz ergibt:

$$(1+b)^n \geq 1+nb \text{ für } n \in \mathbb{N} \text{ und } b \in \mathbb{R}, b \geq 0. \quad (1.4)$$

Denn: $(1+b)^n = 1 + \binom{n}{1}b + \binom{n}{2}b^2 + \dots + \binom{n}{n}b^n \geq 1+nb$, da $\binom{n}{1} = n$ und weil nach der Voraussetzung $b \geq 0$ die Summe $\binom{n}{2}b^2 + \dots + \binom{n}{n}b^n$ nicht negativ ist.

Weitere **Beispiele**:

4. Es sei $c \in \mathbb{R}$ eine (fest gewählte) reelle Zahl. Wir betrachten die Folge c^1, c^2, c^3, \dots . Für das n -te Folgenglied a_n dieser Folge gilt also $a_n = c^n$ ($n = 1, 2, \dots$). Hier sind einige Fälle zu unterscheiden:

1. Fall: $c > 1$

Dann ist $c = 1+b$ für ein $b > 0$. Wir zeigen $c^n \rightarrow \infty$. Hierzu sei $r > 0$. Wählen wir $N > \frac{r}{b}$, so folgt für alle $n \geq N$: $c^n = (1+b)^n \geq 1+nb > nb$ nach (1.4) und daher $c^n > nb \geq Nb > r$. Also gilt $c^n \rightarrow \infty$.

2. Fall: $0 < |c| < 1$

Wir zeigen $c^n \rightarrow 0$. Hierzu setzen wir $d := \frac{1}{|c|}$. Wegen $0 < |c| < 1$ gilt $d > 1$, also nach Fall 1: $d^n \rightarrow \infty$. Aus Satz 3 folgt $\frac{1}{d^n} \rightarrow 0$. Da $\frac{1}{d^n} = |c|^n = |c^n|$ gilt, haben wir erhalten: Für die Folge $(|c^n|)$ gilt $(|c^n|) \rightarrow 0$. Damit folgt, dass auch $c^n \rightarrow 0$ gilt².

3. Fall: $c = 0$ oder $c = 1$

Dann ist (c^n) eine Folge, für die alle Glieder gleich sind. Man spricht von einer *konstanten Folge*. Konstante Folgen sind natürlich immer konvergent: Der Grenzwert ist der gemeinsame Wert aller Folgenglieder.

4. Fall: $c \leq -1$

Hier gilt $c^n \leq -1$ für ungerade n bzw. $c^n \geq 1$ für gerade n . Der Abstand zweier aufeinanderfolgender Folgenglieder ist also immer mindestens 2; daher ist in diesem Fall (a_n) nicht konvergent. (Auch nicht uneigentlich, da das Vorzeichen wechselt.)

5. Wir betrachten für ein festes $c \in \mathbb{R}$ die Folge $(a_n) = (\frac{c^n}{n!})$.

Behauptung: $(a_n) \rightarrow 0$.

Beweis. Wir wählen ein $n_0 \in \mathbb{N}$ mit $n_0 \geq |c|$. Für $n > n_0$ gilt dann

$$|a_n - 0| = \frac{|c|^n}{n!} = \frac{|c|^{n_0}}{n_0!} \cdot \frac{|c|}{n_0+1} \cdot \frac{|c|}{n_0+2} \cdot \dots \cdot \frac{|c|}{n} \leq \frac{|c|^{n_0}}{n_0!} \cdot \frac{|c|}{n},$$

da $|c| < i$ für $i = n_0+1, \dots, n-1$ gilt. Es sei nun $\varepsilon > 0$ gegeben. Wählen wir $N \in \mathbb{N}$, so dass $N > n_0$ und $N > \frac{|c|^{n_0}}{n_0!} \cdot \frac{|c|}{\varepsilon}$ gilt, so folgt für alle $n \geq N$:

$$|a_n - 0| \leq \frac{|c|^{n_0}}{n_0!} \cdot \frac{|c|}{n} \leq \frac{|c|^{n_0}}{n_0!} \cdot \frac{|c|}{N} < \varepsilon. \quad \square$$

In den bisherigen Beispielen für konvergente Folgen hatten wir immer von Anfang an eine Vermutung, welches der Grenzwert der Folge ist. Das braucht nicht immer so zu sein. Betrachten wir etwa das folgende (wichtige) **Beispiel**.

²Eine Folge mit Grenzwert 0 heißt *Nullfolge*. Allgemein gilt: Eine Folge (a_n) ist genau dann eine Nullfolge, wenn die zugehörige Folge der Beträge $|a_n|$ eine Nullfolge ist. (Der einfache Beweis sei zur Übung empfohlen.)

6. $a_n = (1 + \frac{1}{n})^n$. Einige (gerundete) Werte von a_n :

n	a_n	$= (1 + \frac{1}{n})^n = (\frac{n+1}{n})^n$
1	$(\frac{2}{1})^1$	= 2
2	$(\frac{3}{2})^2$	= 2.25
3	$(\frac{4}{3})^3$	= 2.37037
4	$(\frac{5}{4})^4$	= 2.44141
5	$(\frac{6}{5})^5$	= 2.48832
6	$(\frac{7}{6})^6$	= 2.52163
10	$(\frac{11}{10})^{10}$	= 2.593742460
100	$(\frac{101}{100})^{100}$	= 2.704813815
10^3		2.716923842
10^4		2.718145918
10^5		2.718254646
10^6		2.718281828

Bisher haben wir den Nachweis der Konvergenz immer dadurch geführt, dass wir von einem vermuteten Grenzwert nachgewiesen haben, dass es sich tatsächlich um den Grenzwert handelt. Diese Vorgehensweise ist hier nicht möglich, da wir keinen Anhaltspunkt haben, welches der Grenzwert sein könnte. Wir können zwar vermuten, dass die obige Folge konvergent ist und dass der Grenzwert irgendwo zwischen 2.71 und 2.72 liegt – aber dies hilft uns für den Nachweis der Konvergenz nicht weiter. Wir werden im Folgenden zwei Methoden kennenlernen, mit deren Hilfe man in vielen Fällen über die Konvergenz einer Folge entscheiden kann, ohne eine Vermutung zu haben, welches ihr Grenzwert ist. Die erste Methode wird durch den Satz über monotone, beschränkte Folgen bereitgestellt (Satz 5), die zweite Methode durch das Cauchysche Konvergenzkriterium (Satz 6). Beide Methoden beruhen auf dem Vollständigkeitsaxiom.

Definition.

- (a) Es sei M eine Menge reeller Zahlen. Existiert ein $K \in \mathbb{R}$, so dass $x \leq K$ für alle $x \in M$ gilt, so heißt M *nach oben beschränkt* und K heißt eine *obere Schranke* von M . Entsprechend definiert man *nach unten beschränkt* und *untere Schranke*.
- (b) Ist M nach oben und nach unten beschränkt, so heißt M *beschränkt*. Mit anderen Worten: M ist genau dann beschränkt, wenn es ein $K \in \mathbb{R}$ gibt, so dass $|x| \leq K$ für alle $x \in M$ gilt.
- (c) Ist K eine obere Schranke von M und gilt für jede reelle Zahl $K' < K$, dass K' keine obere Schranke von M ist, so heißt K *obere Grenze*, *Supremum* oder *kleinste obere Schranke* von M . Entsprechend definiert man *untere Grenze* (*Infimum*, *größte untere Schranke*).

Satz 4.

Jede nichtleere, nach oben beschränkte Menge $M \subseteq \mathbb{R}$ besitzt ein Supremum; entsprechend besitzt jede nichtleere, nach unten beschränkte Menge $M \subseteq \mathbb{R}$ ein Infimum.

Für den Beweis von Satz 4 verweisen wir auf die Literatur³. Wir erwähnen nur, dass man für den Beweis in entscheidender Weise das Vollständigkeitsaxiom benutzt. Man nennt die erste Teilaussage von Satz 4 auch das *Supremumsprinzip*. Das Supremumsprinzip wird in manchen Axiomensystemen für die reellen Zahlen auch als ein Axiom benutzt (anstelle unseres Vollständigkeitsaxioms). Aus der Definition des Supremums als kleinste obere Schranke folgt sofort, dass eine Menge niemals zwei verschiedene Suprema haben kann; entsprechendes gilt für das Infimum einer Menge. Für das Supremum (Infimum) einer Menge M wollen wir abkürzend schreiben:

$$\sup M \quad \text{und} \quad \inf M.$$

³Siehe die im Literaturverzeichnis aufgeführten Lehrbücher der Analysis, oder beispielsweise auch K. Habetha: Höhere Mathematik für Ingenieure und Physiker, Band 1.

Wir weisen noch darauf hin, dass das Supremum (Infimum) einer Menge M keineswegs zu M gehören muss. Mit anderen Worten: Es kann durchaus $\sup M \notin M$ oder $\inf M \notin M$ gelten.

Hierzu zwei **Beispiele**:

1. $M = \left\{ \frac{1}{n} : n \in \mathbb{N} \right\}$

Dann gilt $\sup M = 1$ und $\inf M = 0$ und außerdem gilt $\sup M \in M$ sowie $\inf M \notin M$.

2. $M = \left\{ q \in \mathbb{Q} : q > 0 \text{ und } q^2 < 2 \right\}$

Dann gilt $\inf M = 0$ und $\sup M = \sqrt{2}$; außerdem gilt $\inf M \notin M$ sowie (wegen $\sqrt{2} \notin \mathbb{Q}$) $\sup M \notin M$.

Definition.

- (a) Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen heißt *monoton steigend* (oder auch *monoton wachsend*), falls $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$ gilt. Entsprechend definiert man *monoton fallend*. Eine Folge heißt *monoton*, falls sie monoton steigend oder monoton fallend ist.
- (b) Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen heißt *beschränkt*, falls die Menge ihrer Folgenglieder beschränkt ist (d.h., falls die Menge $M = \{a_n : n \in \mathbb{N}\}$ beschränkt ist).

Wir kommen nun wie angekündigt zu den beiden Sätzen, die uns erlauben, Konvergenz einer Folge nachzuweisen, ohne eine Vermutung zu haben, welches ihr Grenzwert ist.

Satz 5 (Satz über monotone, beschränkte Folgen).

Jede monotone und beschränkte Folge ist konvergent.

Beweis. Es sei (a_n) eine monotone, beschränkte Folge; (a_n) sei etwa monoton steigend. (Ist (a_n) monoton fallend, so führt man den Beweis entsprechend.) Da (a_n) beschränkt ist, besitzt die Menge M der Folgenglieder ein Supremum (Satz 4). Es sei $a = \sup M$. Zu jedem $\varepsilon > 0$ ist dann die Zahl $a - \varepsilon$ keine obere Schranke von M , d.h., es gibt eine natürliche Zahl N mit $a_N > a - \varepsilon$. Da (a_n) monoton steigend ist, folgt $a_n \geq a_N$ für alle $n \geq N$. Ferner gilt $a \geq a_n$, da $a = \sup M$. Fassen wir diese Ungleichungen zusammen, so erhalten wir $a \geq a_n > a - \varepsilon$ für alle $n \geq N$. Also gilt $|a_n - a| = a - a_n < \varepsilon$ für alle $n \geq N$. Damit ist $(a_n) \rightarrow a$ gezeigt. \square

Satz 6 (Cauchysches Konvergenzkriterium).

Eine Folge (a_n) ist genau dann konvergent, wenn zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass für alle $n, m \geq N$ gilt:

$$|a_n - a_m| < \varepsilon$$

Dass das Erfülltsein des Cauchyschen Konvergenzkriteriums *notwendig* für die Konvergenz von (a_n) ist, ist unschwer nachzuweisen. (Als Übungsaufgabe empfohlen!) Für den Beweis, dass das Kriterium auch hinreichend ist, verweisen wir auf die Literatur. Wir erwähnen hier nur, dass der Beweis auf dem Vollständigkeitsaxiom beruht; er liefert kein Verfahren zur Berechnung des Grenzwertes. Wir kehren nun zu Beispiel 6 zurück und wollen die Konvergenz der Folge $a_n = \left(1 + \frac{1}{n}\right)^n$ beweisen. Hierzu benötigen wir die folgende Ungleichung, die eine Verallgemeinerung der Ungleichung (1.4) ist.

Satz 7 (Bernoullische Ungleichung).

Für alle reellen Zahlen $b \geq -1$ und alle $n \in \mathbb{N}$ gilt:

$$\left(1 + b\right)^n \geq 1 + nb.$$

Für $n \geq 2$ und $b \neq 0$ gilt dies sogar mit strenger Ungleichung:

$$\left(1 + b\right)^n > 1 + nb.$$

Beweis. Der Beweis erfolgt durch vollständige Induktion.

(I) Induktionsanfang:

Für $n = 1$ haben wir $1 + b \geq 1 + b$, d.h., die Behauptung von Satz 7 ist richtig.

(II) Induktionsschritt:

Wir nehmen an, dass Satz 7 für ein $n \in \mathbb{N}$ richtig ist (Induktionsannahme). Wegen $b \geq -1$ gilt $1 + b \geq 0$. Es folgt:

$$(1 + b)^{n+1} = (1 + b)^n(1 + b) \stackrel{(*)}{\geq} (1 + nb)(1 + b) = 1 + (n + 1)b + nb^2.$$

(Die Ungleichung $(*)$ gilt nach Induktionsannahme und wegen $1 + b \geq 0$.)

Da $nb^2 \geq 0$ gilt und da für $b \neq 0$ sogar $nb^2 > 0$ gilt, folgt $(1 + b)^{n+1} \geq 1 + (n + 1)b$, wobei für $b \neq 0$ sogar die strenge Ungleichung gilt. \square

Wir zeigen nun:

$$\text{Die Folge } (a_n) \text{ mit } a_n = \left(1 + \frac{1}{n}\right)^n \text{ ist konvergent.} \quad (1.5)$$

Beweis. Wir zeigen 1. (a_n) ist monoton wachsend und 2. (a_n) ist beschränkt.

1. Wir zeigen $a_{n-1} < a_n$ für $n > 1$, d.h. $\left(1 + \frac{1}{n-1}\right)^{n-1} < \left(1 + \frac{1}{n}\right)^n$. Wir formen dies gleichwertig um:

$$\begin{aligned} \left(1 + \frac{1}{n-1}\right)^{n-1} &< \left(1 + \frac{1}{n}\right)^n \\ \iff \left(\frac{n}{n-1}\right)^{n-1} &< \left(\frac{n+1}{n}\right)^n \\ \iff \left(\frac{n}{n-1}\right)^n &< \left(\frac{n+1}{n}\right)^n \cdot \frac{n}{n-1} \\ \iff \left(\frac{n^2}{(n-1)(n+1)}\right)^n &< \frac{n}{n-1} \\ \iff \left(\frac{n^2}{n^2-1}\right)^n &< \frac{n}{n-1} \\ \iff \left(\frac{n^2-1}{n^2}\right)^n &> \frac{n-1}{n} \\ \iff \left(1 - \frac{1}{n^2}\right)^n &> 1 - \frac{1}{n} \end{aligned}$$

Die letzte Ungleichung ist richtig: Sie folgt nämlich aus der Bernoullischen Ungleichung, wenn man dort $b = -\frac{1}{n^2}$ einsetzt.

2. Wir zeigen $a_n < 3$ für alle $n \in \mathbb{N}$. Hierzu berechnen wir $a_n = \left(1 + \frac{1}{n}\right)^n$ nach dem binomischen Lehrsatz:

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

Für $1 \leq k \leq n$ gilt:

$$\begin{aligned} \binom{n}{k} \frac{1}{n^k} &= \frac{n(n-1)\dots(n-k+1)}{k!n^k} = \frac{1}{k!} \cdot \frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{n-k+1}{n} \leq \frac{1}{k!} = \frac{1}{1 \cdot 2 \cdot \dots \cdot k} \leq \frac{1}{2^{k-1}} \\ \implies \left(1 + \frac{1}{n}\right)^n &\leq 1 + \sum_{k=1}^n \frac{1}{2^{k-1}} = 1 + \sum_{k=0}^{n-1} \left(\frac{1}{2}\right)^k = 1 + \frac{1 - \left(\frac{1}{2}\right)^n}{1 - \frac{1}{2}} = 1 + 2 - \left(\frac{1}{2}\right)^{n-1} < 3 \end{aligned}$$

Dabei haben wir die geometrische Summenformel benutzt. Es gilt also $a_n < 3$ für alle $n \in \mathbb{N}$. Da ferner $a_1 = 2$ ist, folgt (wegen 1.) $2 \leq a_n < 3$ für alle $n \in \mathbb{N}$.

(a_n) ist monoton und beschränkt, also nach Satz 5 konvergent. \square

Den Grenzwert der Folge $(1 + \frac{1}{n})^n$ bezeichnet man mit e . Man nennt e die *Eulersche Zahl*. Man kann zeigen, dass e nicht in \mathbb{Q} liegt. Die ersten Stellen der Dezimalbruchdarstellung von e lauten:

$$e = 2.71828182845904\dots$$

Bei der Eulerschen Zahl e handelt es sich um eine reelle Zahl, die uns noch häufig begegnen wird.

Satz 8 (Rechenregeln für konvergente Folgen).

Es gelte $(a_n) \rightarrow a$ und $(b_n) \rightarrow b$ (für $a, b \in \mathbb{R}$). Dann folgt:

- (a) $(a_n + b_n) \rightarrow a + b$
- (b) $(a_n \cdot b_n) \rightarrow a \cdot b$
- (c) $(c \cdot a_n) \rightarrow c \cdot a$ für $c \in \mathbb{R}$
- (d) $\frac{a_n}{b_n} \rightarrow \frac{a}{b}$, falls $b_n \neq 0$ für alle n und $b \neq 0$.
- (e) Falls $a_n \leq b_n$ für alle n , so folgt $a \leq b$.

Beweis. Wir wollen hier nur (a) beweisen⁴. Es sei $\varepsilon > 0$ gegeben. Da $(a_n) \rightarrow a$ und $(b_n) \rightarrow b$ gilt, gibt es ein N_1 , so dass $|a_n - a| < \frac{\varepsilon}{2}$ für alle $n \geq N_1$ gilt und es gibt ein N_2 , so dass $|b_n - b| < \frac{\varepsilon}{2}$ für alle $n \geq N_2$ gilt. Es folgt für alle $n \geq N := \max(N_1, N_2)$:

$$\left| a_n + b_n - (a + b) \right| = \left| a_n - a + b_n - b \right| \leq \left| a_n - a \right| + \left| b_n - b \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \square$$

Schreibweise. Anstelle von $(a_n) \rightarrow a$ schreibt man auch:

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

Unter Verwendung dieser Schreibweise lassen sich die Rechenregeln aus Satz 8 wie folgt angeben:

- (a) $\lim (a_n + b_n) = \lim a_n + \lim b_n$
- (b) $\lim (a_n \cdot b_n) = \lim a_n \cdot \lim b_n$
- (c) $\lim (c \cdot a_n) = c \cdot \lim a_n$
- (d) $\lim \frac{a_n}{b_n} = \frac{\lim a_n}{\lim b_n}$, falls $b_n \neq 0$ für alle n und $\lim b_n \neq 0$.
- (e) Falls $a_n \leq b_n$ für alle n , so folgt $\lim a_n \leq \lim b_n$.

Definition von e in dieser Schreibweise:

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

Neben den Regeln (a) - (e) lassen sich weitere Regeln angeben, die den Fall behandeln, dass $(b_n) \rightarrow \infty$ oder $(b_n) \rightarrow -\infty$ gilt. Wir erwähnen hier nur:

(f) Gilt $\lim a_n = a$ für $a \in \mathbb{R}$, $a \neq 0$ und $\lim b_n = \infty$, so folgt:

$$\lim (a_n b_n) = \begin{cases} \infty & , \text{ falls } a > 0 \\ -\infty & , \text{ falls } a < 0 \end{cases}$$

Beispiele:

1. Gesucht ist $\lim \frac{2n-1}{3n^2+n-2}$. Anwendung von Regel (d) hilft nicht weiter, da Zähler und Nenner gegen ∞ streben. Klammert man im Zähler n und im Nenner n^2 aus, so erhält man:

$$\frac{2n-1}{3n^2+n-2} = \frac{n \left(2 - \frac{1}{n} \right)}{n^2 \left(3 + \frac{1}{n} - \frac{2}{n^2} \right)} = \frac{1}{n} \cdot \frac{2 - \frac{1}{n}}{3 + \frac{1}{n} - \frac{2}{n^2}}.$$

⁴Die Beweise von (b) - (e) sind dem Beweis von (a) sehr ähnlich. Man findet sie in fast allen gängigen Lehrbüchern der Analysis.

Anwendung von Regel (a) - (d) ergibt $\lim \left(\frac{1}{n} \cdot \frac{2 - \frac{1}{n}}{3 + \frac{1}{n} - \frac{2}{n^2}} \right) = 0 \cdot \frac{2}{3} = 0$.

2. Entsprechend: $\lim \left(\frac{2n^3 + n - 4}{-3n^2 + 1} \right) = \lim \left(n \cdot \frac{2 + \frac{1}{n^2} - \frac{4}{n^3}}{-3 + \frac{1}{n^2}} \right) = -\infty$ nach Regel (f), da $\lim n = \infty$ und

$$\lim \left(\frac{2 + \frac{1}{n^2} - \frac{4}{n^3}}{-3 + \frac{1}{n^2}} \right) = -\frac{2}{3}.$$

3. Entsprechend: $\lim \left(\frac{2n^3 + n - 4}{3n^3 - 5} \right) = \lim \left(\frac{2 + \frac{1}{n^2} - \frac{4}{n^3}}{3 - \frac{5}{n^3}} \right) = \frac{2}{3}$

Allgemein gilt: Sind $P_k(x) = a_k x^k + \dots + a_0$ und $Q_m(x) = b_m x^m + \dots + b_0$ Polynome mit $a_k, b_m \neq 0$, so gilt:

$$\lim_{n \rightarrow \infty} \left(\frac{P_k(n)}{Q_m(n)} \right) = \begin{cases} 0, & \text{falls } k < m \\ \frac{a_k}{b_m}, & \text{falls } k = m \\ +\infty \text{ oder } -\infty, & \text{falls } k > m \end{cases} \quad (1.6)$$

Im dritten Fall ist das Vorzeichen von ∞ gleich dem von $\frac{a_k}{b_m}$. Feststellung (1.6) ergibt sich wie in den obigen Beispielen aus den Regeln (a) - (d) und (f), wenn man für $\frac{P_k(n)}{Q_m(n)}$ im Zähler und im Nenner die jeweils höchsten Potenzen von n ausklammert:

$$\frac{P_k(n)}{Q_m(n)} = n^{k-m} \cdot \frac{a_k + a_{k-1}n^{-1} + \dots + a_0n^{-k}}{b_m + b_{m-1}n^{-1} + \dots + b_0n^{-m}}.$$

Es folgt:

$$\lim \left(\frac{P_k(n)}{Q_m(n)} \right) = \lim \left(n^{k-m} \right) \cdot \lim \left(\frac{a_k + a_{k-1}n^{-1} + \dots + a_0n^{-k}}{b_m + b_{m-1}n^{-1} + \dots + b_0n^{-m}} \right) = \lim \left(n^{k-m} \right) \cdot \frac{a_k}{b_m}.$$

Man braucht nur noch die Fälle $k < m$, $k = m$ und $k > m$ zu unterscheiden, um (1.6) zu erhalten.

Ein in der Praxis oft gebrauchtes Mittel, um Grenzwerte zu bestimmen, liefert der folgende Satz.

Satz 9 (Einschließungssatz).

Es seien (a_n) , (b_n) und (c_n) Folgen mit $a_n \leq b_n \leq c_n$ für alle $n \in \mathbb{N}$. Die Folgen (a_n) und (c_n) seien konvergent und es gelte:

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

Dann konvergiert auch (b_n) , nämlich ebenfalls gegen a ; d.h., es gilt:

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

Beweis. Da $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} c_n = a$ gilt, gibt es zu gegebenem $\varepsilon > 0$ Zahlen N_1 und N_2 , so dass $|a_n - a| < \varepsilon$ für alle $n \geq N_1$ und $|c_n - a| < \varepsilon$ für alle $n \geq N_2$. Also gilt für alle $n \geq N := \max(N_1, N_2)$: Für $b_n \geq a$ ist $|b_n - a| = b_n - a \leq c_n - a = |c_n - a| < \varepsilon$ und für $b_n < a$ ist $|b_n - a| = a - b_n \leq a - a_n = |a_n - a| < \varepsilon$. \square

Zusammenfassung.

Wir haben sechs Methoden bzw. Hilfsmittel zum Nachweis der Konvergenz oder zur Berechnung des Grenzwertes einer Folge (a_n) kennengelernt.

1. Direkter Nachweis.

Man gibt ein $\varepsilon > 0$ vor und zeigt für ein passend gewähltes N , dass $|a_n - a| < \varepsilon$ für alle $n \geq N$ gilt. Hierzu muss man allerdings eine Vermutung haben, welches der Grenzwert a ist.

2. Satz über monotone, beschränkte Folgen.

Man weist von der Folge (a_n) nach, dass (a_n) monoton und beschränkt ist; daraus folgt dann die Konvergenz von (a_n) . Und zwar gilt: Ist (a_n) monoton wachsend, so ist der Grenzwert $a =$

$\sup \{a_n : n \in \mathbb{N}\}$; ist (a_n) monoton fallend, so ist $a = \inf \{a_n : n \in \mathbb{N}\}$ der Grenzwert⁵. Damit hat man allerdings i.A. nur die Existenz eines Grenzwertes a nachgewiesen; berechnet hat man a noch nicht. Außerdem ist natürlich längst nicht jede Folge, die man untersuchen will, auch monoton; in solchen Fällen kann man den Satz nicht anwenden.

3. Cauchyscher Konvergenztest.

Man prüft, ob das Cauchysche Konvergenzkriterium erfüllt ist. Wenn ja, so ist (a_n) konvergent; andernfalls ist (a_n) divergent. Allerdings erhält man nur eine Entscheidung darüber, ob Konvergenz vorliegt oder nicht; im Fall, dass Konvergenz vorliegt, ist das (oftmals sehr schwierige) Problem, den Grenzwert tatsächlich zu berechnen, durch diesen Test noch nicht gelöst.

Die Methoden 2. und 3. liefern also reine Existenzaussagen. Wir werden von 2. häufig Gebrauch machen. Mithilfe der Methoden 4. und 5. führt man die Berechnung des Grenzwertes einer Folge auf bereits vorher berechnete Grenzwerte zurück. In der Praxis wird man diese Methoden am häufigsten anwenden.

4. Anwendung von Rechenregeln.

Unter „Rechenregeln“ sind hier die Regeln (a) - (e) aus Satz 8 sowie (f) und weitere ähnliche Regeln zu verstehen. Auch (1.6), Satz 3 sowie die Fußnote 2 lassen sich als „Rechenregeln“ ansehen.

5. Einschließungssatz.

Dieser wird häufig auch für den Spezialfall angewendet, dass (a_n) (oder (c_n)) eine konstante Folge ist, d.h., $a_n = a$ (oder $c_n = c$) für alle $n \in \mathbb{N}$.

6. Ungleichungen.

Wie aus den Beispielen klar geworden sein dürfte, spielen Ungleichungen beim Nachweis von Konvergenz und beim Berechnen von Grenzwerten eine besonders wichtige Rolle. Von grundlegender Wichtigkeit ist dabei die ständig verwendete Dreiecksungleichung $|a + b| \leq |a| + |b|$.

Weitere nützliche Ungleichungen:

1. *Bernoullische Ungleichung*: $(1 + b)^n \geq 1 + nb$ für $b \geq -1$;
2. Ungleichungen, die man aus dem *Binomischen Lehrsatz* gewinnt, wie beispielsweise die Ungleichung $(1 + b)^n > \binom{n}{2}b^2 = \frac{n(n-1)}{2}b^2$ für $b \geq 0$ und $n \geq 2$ (vgl. auch (1.4)).

Nach dieser Zusammenfassung betrachten wir abschließend noch einige Beispiele für **Reihen**. Wegen ihrer besonderen Wichtigkeit werden Reihen später noch wesentlich genauer behandelt werden. Aus einer gegebenen Folge a_1, a_2, a_3, \dots lässt sich eine neue Folge s_1, s_2, s_3, \dots bilden, indem man jeweils die ersten n Glieder der Folge (a_i) summiert:

$$\begin{aligned} s_1 &= a_1 \\ s_2 &= a_1 + a_2 \\ s_3 &= a_1 + a_2 + a_3 \\ s_4 &= a_1 + a_2 + a_3 + a_4 \\ &\vdots \\ s_n &= \sum_{i=1}^n a_i = a_1 + a_2 + \dots + a_n \end{aligned}$$

Die neu gewonnene Folge $(s_n)_{n \in \mathbb{N}}$ nennt man eine *Reihe*. Die Summe $s_n = \sum_{i=1}^n a_i$ heißt *n-te Partialsumme* der Reihe, die Summanden a_i heißen *Glieder der Reihe*.

Eine Reihe $(s_n)_{n \in \mathbb{N}}$ mit den Gliedern a_1, a_2, \dots bezeichnet man meistens mit dem Symbol $\sum_{i=1}^{\infty} a_i$. Durch

⁵Dies ergibt sich unmittelbar aus dem Beweis von Satz 5.

$\sum_{i=1}^{\infty} a_i$ wird also keineswegs eine „unendliche Summe“ dargestellt; es handelt sich hier vielmehr um eine Bezeichnung für die Reihe $(s_n)_{n \in \mathbb{N}}$.

Eine Reihe $\sum_{i=1}^{\infty} a_i$ heißt *konvergent*, wenn die Folge $(s_n)_{n \in \mathbb{N}}$ der Partialsummen konvergiert, d.h., wenn $(s_n) \rightarrow s$ für ein $s \in \mathbb{R}$ gilt. Hierfür schreibt man:

$$\sum_{i=1}^{\infty} a_i = s. \quad (1.7)$$

Das Zeichen $\sum_{i=1}^{\infty} a_i$ wird also in zwei unterschiedlichen Bedeutungen verwendet: Einerseits bezeichnet $\sum_{i=1}^{\infty} a_i$ die Reihe $(s_n)_{n \in \mathbb{N}}$ mit $s_n = \sum_{i=1}^n a_i$, andererseits bezeichnet $\sum_{i=1}^{\infty} a_i$ auch (wie in (1.7)) den Grenzwert s der Reihe $(s_n)_{n \in \mathbb{N}}$, falls dieser existiert. Gilt $(s_n) \rightarrow \infty$ bzw. $(s_n) \rightarrow -\infty$, so schreibt man $\sum_{i=1}^{\infty} a_i = \infty$ bzw. $\sum_{i=1}^{\infty} a_i = -\infty$.

Die Glieder a_i einer Reihe können natürlich auch mit dem Index 0 oder irgendeinem anderen Index beginnen.

Beispiele:

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

Die Glieder dieser Reihe sind $a_1 = 1, a_2 = \frac{1}{2}, a_3 = \frac{1}{3}, a_4 = \frac{1}{4}, \dots$

Die Partialsummen sind:

$$\begin{aligned} s_1 &= 1 \\ s_2 &= 1 + \frac{1}{2} \\ s_3 &= 1 + \frac{1}{2} + \frac{1}{3} \\ &\vdots \\ s_n &= \sum_{i=1}^n \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \end{aligned}$$

Diese Reihe heißt *harmonische Reihe*. Um zu entscheiden, ob die harmonische Reihe konvergent ist, fassen wir die Glieder a_1, a_2, \dots wie folgt zu Summen zusammen:

$$\begin{aligned} a_1 &= 1 \\ a_2 &= \frac{1}{2} \\ a_3 + a_4 &= \frac{1}{3} + \frac{1}{4} > \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \\ a_5 + a_6 + a_7 + a_8 &= \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{4}{8} = \frac{1}{2} \\ a_9 + a_{10} + a_{11} + a_{12} + a_{13} + a_{14} + a_{15} + a_{16} &> 8 \cdot \frac{1}{16} = \frac{1}{2} \end{aligned}$$

Allgemein gilt für die Glieder a_i mit $i = 2^{n-1} + 1, \dots, 2^n$: $a_i = \frac{1}{i} \geq \frac{1}{2^n}$. Daher folgt für die Summe dieser 2^{n-1} Glieder:

$$\sum_{i=2^{n-1}+1}^{2^n} a_i \geq \sum_{i=2^{n-1}+1}^{2^n} \frac{1}{2^n} = 2^{n-1} \cdot \frac{1}{2^n} = \frac{1}{2}.$$

Also folgt für die 2^n -te Partialsumme s_{2^n} :

$$s_{2^n} = a_1 + a_2 + (a_3 + a_4) + \dots + (a_{2^{n-1}+1} + \dots + a_{2^n}) \geq 1 + \underbrace{\frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{2}}_{n\text{-mal}} = 1 + \frac{n}{2}.$$

Da $1 + \frac{n}{2}$ gegen ∞ strebt, folgt, dass auch die Partialsummen der harmonischen Reihe gegen ∞ streben, d.h., *die harmonische Reihe ist divergent*⁶.

2. $\sum_{i=0}^{\infty} q^i$ (mit $q \in \mathbb{R}$).

Diese Reihe heißt *geometrische Reihe*.

Die Glieder dieser Reihe sind $a_0 = q^0$, $a_1 = q^1$, $a_2 = q^2$, $a_3 = q^3$, \dots

Die Partialsummen sind $s_0 = a_0 = q^0 = 1$, $s_1 = a_0 + a_1 = q^0 + q^1 = 1 + q$, $s_2 = a_0 + a_1 + a_2 = q^0 + q^1 + q^2 = 1 + q + q^2$, \dots

Allgemein: $s_n = \sum_{i=0}^n a_i = \sum_{i=0}^n q^i = \frac{1-q^{n+1}}{1-q}$ nach der geometrischen Summenformel (für $q \neq 1$).

Hieraus folgt (Wieso nämlich?):

$$\sum_{i=0}^{\infty} q^i = \frac{1}{1-q} \quad \text{für } |q| < 1. \quad (1.8)$$

Ist $|q| \geq 1$, so divergiert die geometrische Reihe, wie man sich zumindest für den Fall $q \geq 1$ leicht überlegt. (Eine Begründung, die sowohl für $q \geq 1$ als auch für den Fall $q \leq -1$ Gültigkeit besitzt, werden wir später in Abschnitt 4.1 kennenlernen.) Es gilt beispielsweise:

$$\begin{aligned} \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^i &= \frac{1}{1-\frac{1}{2}} = 2 \\ \sum_{i=0}^{\infty} \left(-\frac{1}{2}\right)^i &= \frac{1}{1+\frac{1}{2}} = \frac{2}{3} \\ \sum_{i=0}^{\infty} \left(\frac{14}{17}\right)^i &= \frac{1}{1-\frac{14}{17}} = \frac{17}{3} \\ \sum_{i=0}^{\infty} \left(\frac{3}{2}\right)^i &= \infty \end{aligned}$$

3. Oben haben wir gesehen, dass die harmonische Reihe $\sum_{i=1}^{\infty} \frac{1}{i}$ divergiert. Wir betrachten nun die Reihe $\sum_{i=1}^{\infty} \frac{1}{i^2}$ und zeigen, dass sie konvergiert. Wir benutzen dazu den Satz über monotone, beschränkte Folgen; wir haben also zu zeigen, dass die Partialsummen $s_n = \sum_{i=1}^n \frac{1}{i^2}$ eine monotone und beschränkte Folge bilden.

⁶Die harmonische Reihe strebt sehr langsam gegen ∞ . Es ist beispielsweise $\sum_{i=1}^{10000} \frac{1}{i}$ immer noch kleiner als 10.

- (a) Dass die Folge $(s_n)_{n \in \mathbb{N}}$ monoton wachsend ist, ergibt sich daraus, dass die Zahlen $\frac{1}{i^2}$ positiv sind.
- (b) Es ist $\frac{1}{i^2} < \frac{1}{i(i-1)} = \frac{1}{i-1} - \frac{1}{i}$ für $i \geq 2$. Also gilt:

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

Die Folge (s_n) ist also beschränkt.

Aus (a) und (b) folgt, dass die Reihe $\sum_{i=1}^{\infty} \frac{1}{i^2}$ konvergiert. Den Grenzwert können wir mit den zur Verfügung stehenden Mitteln allerdings nicht bestimmen.

1.4 Funktionsgrenzwerte

Wir betrachten in dieser Vorlesung vor allem reelle Funktionen, das sind Funktionen $f : D \rightarrow \mathbb{R}$, wobei der Definitionsbereich D eine Teilmenge der reellen Zahlen ist ($D \subseteq \mathbb{R}$). Ist nichts Gegenteiliges gesagt, so sollen die betrachteten Funktionen stets reell sein. Den Definitionsbereich einer Funktion f wollen wir im Folgenden mit $D(f)$ bezeichnen. $D(f)$ wird meistens ein Intervall sein; unter einem Intervall verstehen wir Folgendes:

Definition.

Es seien $a, b \in \mathbb{R}$. Die Mengen

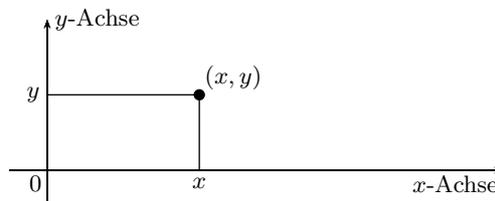
$$\begin{aligned} [a, b] &:= \{x \in \mathbb{R} : a \leq x \leq b\} \\ (a, b) &:= \{x \in \mathbb{R} : a < x < b\} \\ (a, b] &:= \{x \in \mathbb{R} : a < x \leq b\} \\ [a, b) &:= \{x \in \mathbb{R} : a \leq x < b\} \end{aligned}$$

heißen *endliche Intervalle*; $[a, b]$ heißt *abgeschlossenes Intervall*, (a, b) heißt *offenes Intervall*, $[a, b)$ und $(a, b]$ heißen *halboffene Intervalle*. Die Mengen

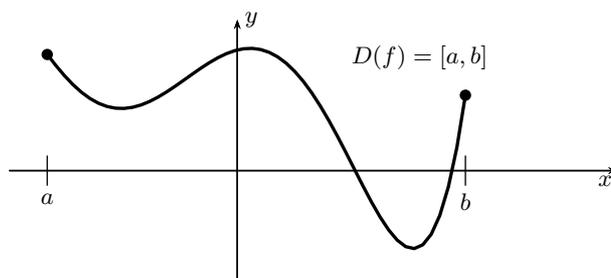
$$\begin{aligned} (-\infty, b] &:= \{x \in \mathbb{R} : x \leq b\} \\ (-\infty, b) &:= \{x \in \mathbb{R} : x < b\} \\ [a, \infty) &:= \{x \in \mathbb{R} : a \leq x\} \\ (a, \infty) &:= \{x \in \mathbb{R} : a < x\} \\ (-\infty, \infty) &:= \mathbb{R} \end{aligned}$$

heißen *unendliche Intervalle*.

Zur Veranschaulichung von reellen Funktionen verwendet man i.A. ein rechtwinkliges Koordinatensystem, d.h., ein Paar von Zahlengeraden, die ihren Nullpunkt gemeinsam haben und sich rechtwinklig schneiden („x-Achse“ und „y-Achse“). Jedem geordneten Paar (x, y) reeller Zahlen entspricht dann genau ein Punkt der Zeichenebene (und umgekehrt):



Das *Schaubild* (oder *graphische Darstellung* oder *Graph*) einer reellen Funktion f ist die Menge aller Punkte (x, y) mit $x \in D(f)$ und $y = f(x)$.



Im letzten Abschnitt haben wir Grenzwerte von Folgen (a_n) behandelt, d.h. Grenzwerte vom folgenden Typ:

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

Im Zusammenhang mit reellen Funktionen f betrachtet man darüber hinaus sehr oft Grenzwerte vom folgenden Typ:

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

Diese nennt man *Funktionsgrenzwerte*. Wir führen sie im Folgenden ein.

Gegeben sei eine reelle Funktion f mit Definitionsbereich $D(f)$ und ein $x_0 \in \mathbb{R}$. Wir betrachten Folgen $(x_n)_{n \in \mathbb{N}}$, für die $x_n \in D(f)$, $x_n \neq x_0$ und $\lim_{n \rightarrow \infty} x_n = x_0$ gilt. Zu jeder solchen Folge (x_n) bilden wir die zugehörige *Folge der Funktionswerte*

$$\left(f(x_n) \right)_{n \in \mathbb{N}}$$

und interessieren uns für das Konvergenzverhalten dieser Folge. Dies führt zur folgenden Definition.

Definition 3.

Es sei f eine reelle Funktion und $x_0 \in \mathbb{R}$. Wir betrachten Folgen $(x_n)_{n \in \mathbb{N}}$ für die

$$x_n \in D(f), \quad x_n \neq x_0, \quad \lim_{n \rightarrow \infty} x_n = x_0 \tag{1.9}$$

gilt und setzen voraus, dass mindestens eine Folge mit diesen Eigenschaften existiert. Es sei $a \in \mathbb{R}$. Wir sagen f hat *an der Stelle x_0 den Grenzwert a* und schreiben dafür $\lim_{x \rightarrow x_0} f(x) = a$, falls für jede Folge $(x_n)_{n \in \mathbb{N}}$, für die (1.9) erfüllt ist, gilt:

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

Wir sagen auch *der Grenzwert $\lim_{x \rightarrow x_0} f(x)$ existiert*, falls $\lim_{x \rightarrow x_0} f(x) = a$ für ein $a \in \mathbb{R}$ gilt.

Man beachte, dass in dieser Definition nicht $x_0 \in D(f)$ vorausgesetzt wird, d.h., x_0 kann zum Definitionsbereich gehören oder auch nicht. Es ist beispielsweise durchaus möglich, dass $D(f)$ ein offenes Intervall ist und dass x_0 ein Randpunkt von $D(f)$ ist; oder dass etwa $D(f) = \mathbb{R} \setminus \{0\}$ und $x_0 = 0$ gilt.

Bemerkungen und Ergänzungen.

- (1) Für die Frage, ob f an der Stelle x_0 den Grenzwert a hat, spielt es überhaupt keine Rolle, ob f in x_0 definiert ist und welchen Wert f in x_0 hat (falls f dort definiert ist).
- (2) Eine Erweiterung von Definition 3 erhält man, wenn man dort für a auch $-\infty$ oder ∞ zulässt. (Wenn $\lim_{x \rightarrow x_0} f(x) = \infty$ oder $\lim_{x \rightarrow x_0} f(x) = -\infty$ gilt, sagen wir jedoch nicht, dass der Grenzwert $\lim_{x \rightarrow x_0} f(x)$ existiert, sondern nur, dass er *im uneigentlichen Sinne* existiert.)
- (3) Eine andere Erweiterung: $\lim_{x \rightarrow \infty} f(x) = a$ bedeutet, dass für jede Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \in D(f)$ und $\lim_{n \rightarrow \infty} x_n = \infty$ gilt: $\lim_{n \rightarrow \infty} f(x_n) = a$. (Hierbei wird ähnlich wie in Definition 3 vorausgesetzt, dass es mindestens eine solche Folge gibt.) Analog: $\lim_{x \rightarrow -\infty} f(x) = a$.

(4) Anstelle von $\lim_{x \rightarrow x_0} f(x) = a$ schreibt man auch $f(x) \rightarrow a$ für $x \rightarrow x_0$; oder auch $f \rightarrow a$ für $x \rightarrow x_0$.

Beispiele:

1. Wir betrachten die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$. Es ist $D(f) = \mathbb{R}$.

(a) f hat beispielsweise an der Stelle $x_0 = 2$ den Grenzwert 4; denn es gilt für jede Folge (x_n) mit $x_n \in D(f)$, $x_n \neq 2$ und $\lim_{n \rightarrow \infty} x_n = 2$:

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_n^2 = \lim_{n \rightarrow \infty} x_n \cdot \lim_{n \rightarrow \infty} x_n = 2 \cdot 2 = 4.$$

(b) Allgemein hat f (für beliebiges $x_0 \in \mathbb{R}$) an der Stelle x_0 den Grenzwert x_0^2 ; denn es gilt für jede Folge (x_n) mit $x_n \in D(f)$, $x_n \neq x_0$ und $\lim_{n \rightarrow \infty} x_n = x_0$:

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_n^2 = \lim_{n \rightarrow \infty} x_n \cdot \lim_{n \rightarrow \infty} x_n = x_0 \cdot x_0 = x_0^2.$$

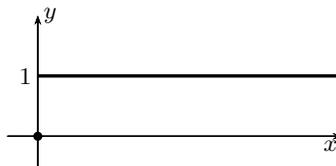
Es folgt also:

$$\lim_{x \rightarrow x_0} f(x) = x_0^2.$$

Andererseits gilt auch $f(x_0) = x_0^2$. Für die Funktion $f(x) = x^2$ stimmt also an jeder Stelle x_0 der Grenzwert $\lim_{x \rightarrow x_0} f(x)$ mit dem Funktionswert $f(x_0)$ überein. Dies ist nicht bei allen Funktionen so, wie das nächste Beispiel zeigt.

2. Wir betrachten die Funktion $f : [0, \infty) \rightarrow \mathbb{R}$ mit $f(x) = \begin{cases} 1, & \text{falls } x > 0 \\ 0, & \text{falls } x = 0 \end{cases}$.

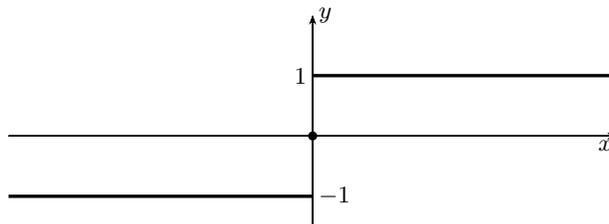
Graph von f :



Wir untersuchen, ob f an der Stelle $x_0 = 0$ einen Grenzwert hat. Hierzu sei $(x_n)_{n \in \mathbb{N}}$ eine Folge mit $x_n \in [0, \infty)$, $x_n \neq 0$ und $\lim_{n \rightarrow \infty} x_n = 0$. Es folgt $\lim_{n \rightarrow \infty} f(x_n) = 1$, da (wegen $x_n \neq 0$) $f(x_n) = 1$ gilt. Also hat f an der Stelle $x_0 = 0$ den Grenzwert 1. Es gilt:

$$\lim_{x \rightarrow 0} f(x) = 1 \neq 0 = f(0)$$

3. Wir betrachten die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = \begin{cases} 1, & \text{falls } x > 0 \\ 0, & \text{falls } x = 0 \\ -1, & \text{falls } x < 0 \end{cases}$



Wir untersuchen, ob f an der Stelle $x_0 = 0$ einen Grenzwert hat. Hierzu betrachten wir die Folge (x_n) mit $x_n = \frac{(-1)^n}{n}$. Für diese Folge gilt $x_n \in D(f)$, $x_n \neq 0$ und $\lim_{n \rightarrow \infty} x_n = 0$; für die zugehörige Folge $(f(x_n))$ der Funktionswerte gilt:

$$f(x_n) = \begin{cases} -1, & \text{falls } n \text{ ungerade} \\ 1, & \text{falls } n \text{ gerade} \end{cases}$$

Es folgt, dass $\lim_{n \rightarrow \infty} f(x_n)$ nicht existiert. Also hat f an der Stelle $x_0 = 0$ keinen Grenzwert.

4. Wir betrachten die Funktion $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ mit $f(x) = \frac{1}{x}$ und untersuchen, ob f an der Stelle x_0 einen Grenzwert hat.

1. Fall: $x_0 \neq 0$

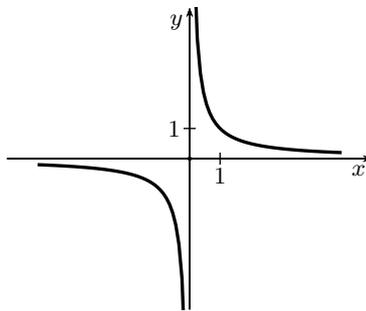
Es sei (x_n) eine Folge mit $x_n \in D(f)$ (d.h. $x_n \neq 0$), $x_n \neq x_0$ und $\lim_{n \rightarrow \infty} x_n = x_0$. Dann gilt (nach Satz 8 (d)):

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} \frac{1}{x_n} = \frac{1}{x_0}, \text{ also } \lim_{x \rightarrow x_0} f(x) = \frac{1}{x_0}.$$

2. Fall: $x_0 = 0$

Wir betrachten die Folgen $(\frac{1}{n})_{n \in \mathbb{N}}$ und $(-\frac{1}{n})_{n \in \mathbb{N}}$. Es gilt: $\lim_{n \rightarrow \infty} f(\frac{1}{n}) = \lim_{n \rightarrow \infty} n = \infty$ und $\lim_{n \rightarrow \infty} f(-\frac{1}{n}) = \lim_{n \rightarrow \infty} -n = -\infty$, d.h., die Funktion $f(x) = \frac{1}{x}$ hat an der Stelle $x_0 = 0$ keinen Grenzwert (auch nicht im uneigentlichen Sinne).

Graph der Funktion f mit $f(x) = \frac{1}{x}$:



5. Abschließend erwähnen wir noch ein weiteres Beispiel einer Funktion f mit $D(f) = \mathbb{R} \setminus \{0\}$, nämlich $f(x) = \frac{\sin x}{x}$. Diese Funktion besitzt an der Stelle $x_0 = 0$ einen Grenzwert; es gilt nämlich (wie wir später zeigen werden):

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

1.5 Stetigkeit

Definition 4.

- (a) Es sei f eine reelle Funktion und $x_0 \in D(f)$. Die Funktion f heißt *stetig an der Stelle x_0* , wenn für jede Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \in D(f)$ und $\lim_{n \rightarrow \infty} x_n = x_0$ gilt:

$$\lim_{n \rightarrow \infty} f(x_n) = f(x_0).$$

- (b) Die Funktion f heißt *stetig auf X* (für $X \subseteq D(f)$), falls f an jeder Stelle $x_0 \in X$ stetig ist.

Ist eine Funktion f an der Stelle $x_0 \in D(f)$ nicht stetig, so nennt man f *unstetig* an der Stelle x_0 ; x_0 nennt man dann auch eine *Unstetigkeitsstelle* von f . Dies bedeutet, dass es eine Folge (x_n) mit $x_n \in D(f)$ und $x_n \rightarrow x_0$ gibt, für welche die zugehörige Folge der Funktionswerte $(f(x_n))$ entweder überhaupt nicht oder gegen einen Wert $\neq f(x_0)$ konvergiert.

Es sei darauf hingewiesen, dass in Definition 4 $x_0 \in D(f)$ vorausgesetzt wird. Dies bedeutet, dass wir von Stetigkeit (bzw. Unstetigkeit) von f an der Stelle x_0 nur dann sprechen können, wenn x_0 zum Definitionsbereich von f gehört (im Gegensatz zu Definition 3, wo $x_0 \notin D(f)$ ja zugelassen war). Aus

Definition 3 und 4 erhält man den folgenden Zusammenhang zwischen Stetigkeit und der Existenz eines Grenzwertes an der Stelle x_0 .

Feststellung.

Es sei f eine reelle Funktion und $x_0 \in D(f)$; der Definitionsbereich $D(f)$ sei so beschaffen, dass es mindestens eine Folge $(x_n)_{n \in \mathbb{N}}$ gibt, für die $x_n \in D(f)$, $x_n \neq x_0$ und $\lim_{n \rightarrow \infty} x_n = x_0$ gilt. (In den folgenden Abschnitten und in allen in diesem Zusammenhang wichtigen Beispielen wird diese Anforderung an $D(f)$ immer ohne Weiteres erfüllt sein.) Dann gilt: f ist an der Stelle x_0 genau dann stetig, wenn Folgendes gilt:

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

Beispiele:

1. Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine konstante Funktion, d.h., $f(x) = c$ gilt für alle $x \in \mathbb{R}$. Dann ist f für jedes $x_0 \in \mathbb{R}$ stetig. Denn für jede Folge (x_n) mit $x_n \in D(f)$ und $\lim_{n \rightarrow \infty} x_n = x_0$ gilt:

$$\lim_{n \rightarrow \infty} f(x_n) = c = f(x_0).$$

2. Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $f(x) = x$. Man nennt diese Funktion die *Identität* oder *identische Funktion*. Die Identität ist stetig auf \mathbb{R} , da für jede Folge (x_n) mit $x_n \in D(f)$ ($D(f) = \mathbb{R}$) und $\lim_{n \rightarrow \infty} x_n = x_0$ ($x_0 \in \mathbb{R}$) gilt:

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_n = x_0 = f(x_0).$$

3. Nach Beispiel 1 des Abschnitts 1.4 ist die Funktion $f(x) = x^2$ auf ganz \mathbb{R} stetig.
 4. Nach Beispiel 4 des Abschnitts 1.4 ist die Funktion $f(x) = \frac{1}{x}$ auf ihrem gesamten Definitionsbereich stetig.
 5. Nach Beispiel 2 des Abschnitts 1.4 ist die dort betrachtete Funktion $f : [0, \infty) \rightarrow \mathbb{R}$ mit

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

an der Stelle $x_0 = 0$ unstetig.

6. Ebenfalls an der Stelle $x_0 = 0$ unstetig ist die Funktion f aus Beispiel 3 des Abschnitts 1.4.

Sind f und g reelle Funktionen mit den Definitionsbereichen $D(f)$ und $D(g)$, so versteht man unter $f + g$ die Funktion mit Definitionsbereich $D(f) \cap D(g)$, die definiert wird durch:

$$(f + g)(x) := f(x) + g(x).$$

Analog definiert man $f - g$, $f \cdot g$ und $\frac{f}{g}$, wobei im Falle von $\frac{f}{g}$ der Definitionsbereich die Menge $D(f) \cap \{x \in D(g) : g(x) \neq 0\}$ ist.

Satz 10.

f und g seien stetig an der Stelle x_0 . Dann gilt:

- (a) $f + g$, $f - g$ und $f \cdot g$ sind stetig an der Stelle x_0 .
 (b) Es gelte $g(x_0) \neq 0$. Dann ist auch $\frac{f}{g}$ stetig an der Stelle x_0 .

Beweis. Es sei (x_n) eine Folge mit $x_n \in D(f) \cap D(g)$ und $\lim_{n \rightarrow \infty} x_n = x_0$. Wegen der Stetigkeit von f und g an der Stelle x_0 folgt dann $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$ und $\lim_{n \rightarrow \infty} g(x_n) = g(x_0)$. Es folgt aus den Rechenregeln

für Grenzwerte (Satz 8):

$$\begin{aligned}\lim_{n \rightarrow \infty} (f + g)(x_n) &= \lim_{n \rightarrow \infty} (f(x_n) + g(x_n)) \\ &= \lim_{n \rightarrow \infty} f(x_n) + \lim_{n \rightarrow \infty} g(x_n) \\ &= f(x_0) + g(x_0) \\ &= (f + g)(x_0)\end{aligned}$$

Also ist $f + g$ an der Stelle x_0 stetig. Analog zeigt man die übrigen Behauptungen von Satz 10 (Beweis als Übungsaufgabe empfohlen). \square

Definition.

(a) Ein *Polynom* P vom Grad n ist eine Funktion $P : \mathbb{R} \rightarrow \mathbb{R}$:

$$P(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (\text{mit } a_n \neq 0).$$

(b) Als eine *rationale Funktion* bezeichnet man den Quotienten R zweier Polynome P und Q :

$$R(x) = \frac{P(x)}{Q(x)}.$$

Der Definitionsbereich von R ist die Menge $D(R) = \{x \in \mathbb{R} : Q(x) \neq 0\}$.

Als eine unmittelbare Folgerung aus Satz 10 sowie den Beispielen 1 und 2 erhält man den folgenden Satz.

Satz 11.

(a) Polynome sind auf \mathbb{R} stetig.

(b) Rationale Funktionen R sind auf ihrem Definitionsbereich $D(R)$ stetig.

Beweis.

(a) Nach den Beispielen 1 und 2 sind konstante Funktionen und die Identität stetig. Mehrmalige Anwendung von Satz 10 (a) ergibt die Behauptung.

(b) Nach Satz 10 (b) folgt aus der Stetigkeit von P und Q die Stetigkeit von R . \square

Will man untersuchen, ob f an der Stelle x_0 stetig ist, so braucht man nicht den gesamten Definitionsbereich $D(f)$ zu betrachten, sondern man kann sich auf ein Intervall $I = (a, b)$ mit $x_0 \in I$ beschränken, wobei man die Länge von I (d.h. $b - a$) beliebig klein wählen darf. Genauer gilt (wie aus Definition 4 (a) folgt):

$$\begin{aligned}f \text{ ist an } x_0 \text{ genau dann stetig}^7, & \text{ wenn es ein Intervall } I = (a, b) \text{ mit } x_0 \in I \text{ gibt,} \\ & \text{so dass die Einschränkung von } f \text{ auf } I \cap D(f) \text{ an } x_0 \text{ stetig ist.}\end{aligned} \tag{1.10}$$

Im Hinblick auf (1.10) sagt man auch, dass die Stetigkeit eine *lokale Eigenschaft* einer Funktion ist. Anders gesagt: Ob f an x_0 stetig ist, hängt nicht vom Gesamtverlauf von f ab, sondern nur vom Verhalten von f in einem beliebig kleinen Intervall $I = (a, b)$, das x_0 enthält.

In der Vorlesung (oder in den Übungen) zeigen wir ergänzend, dass die Nacheinanderausführung stetiger Funktionen wiederum stetig ist.

⁷Statt „an der Stelle x_0 stetig“ sagt man auch „an x_0 stetig“, „im Punkt x_0 stetig“ oder „in x_0 stetig“. Man kann auch „bei x_0 stetig“ sagen.

1.6 Ergänzungen zu den Abschnitten 1.4 und 1.5

1.6.1 Einseitige Grenzwerte

Definition 3'.

Es sei f eine reelle Funktion und $x_0 \in \mathbb{R}$. Wir betrachten Folgen $(x_n)_{n \in \mathbb{N}}$, für die

$$x_n \in D(f), \quad x_0 < x_n \quad \text{und} \quad \lim_{n \rightarrow \infty} x_n = x_0 \quad (1.9')$$

gilt und setzen voraus, dass mindestens eine Folge mit diesen Eigenschaften existiert. Es sei $a \in \mathbb{R}$. Wir sagen, f hat an der Stelle x_0 den *rechtsseitigen Grenzwert* a und schreiben dafür

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

falls für jede Folge $(x_n)_{n \in \mathbb{N}}$, für die (1.9') erfüllt ist, gilt: $\lim_{n \rightarrow \infty} f(x_n) = a$.

Analog definiert man *linksseitige Grenzwerte*. Schreibweise für linksseitige Grenzwerte:

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

1.6.2 Einseitige Stetigkeit

Definition 4'.

Es sei f eine reelle Funktion und $x_0 \in D(f)$. Die Funktion f heißt *rechtsseitig stetig* an der Stelle x_0 , wenn für jede Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \in D(f)$, $x_0 \leq x_n$ und $\lim_{n \rightarrow \infty} x_n = x_0$ gilt:

$$\lim_{n \rightarrow \infty} f(x_n) = f(x_0).$$

Analog definiert man *linksseitig stetig*.

Wir erwähnen (ohne Beweis), dass Folgendes gilt:

f ist stetig an der Stelle $x_0 \iff f$ ist sowohl rechts- als auch linksseitig stetig an der Stelle x_0

1.6.3 ε, δ -Definition der Stetigkeit

Eine andere Möglichkeit, den Begriff der Stetigkeit zu definieren, ist die folgende.

Definition.

Es sei f eine reelle Funktion und $x_0 \in D(f)$. Die Funktion f heißt *stetig* an der Stelle x_0 , wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass

$$|f(x) - f(x_0)| < \varepsilon$$

für alle $x \in D(f)$ gilt, die $|x - x_0| < \delta$ erfüllen.

Man kann zeigen, dass die ε, δ -Definition der Stetigkeit und die am Beginn von Abschnitt 1.5 gegebene Definition der Stetigkeit äquivalent sind. Beweise hierzu findet man in den Lehrbüchern der Analysis, etwa:

- H. Heuser, Lehrbuch der Analysis, Teil 1
- O. Forster, Analysis 1
- K. Königsberger, Analysis 1

Abschließend noch eine *besonders wichtige Bemerkung zur Stetigkeit*, in der festgehalten wird, welche Rolle die Stetigkeit einer Funktion beim *Ausrechnen von Grenzwerten* spielt.

1.6.4 Stetigkeit und Grenzwerte

Gegeben: eine reelle Funktion f und ein $x_0 \in D(f)$.

Wir setzen voraus, dass f an der Stelle x_0 stetig ist. Das bedeutet („Definition der Stetigkeit in Kurzfassung“):

$$\boxed{x_n \rightarrow x_0 \quad \Rightarrow \quad f(x_n) \rightarrow f(x_0).} \quad (\star)$$

Dasselbe in Worten: Wenn x_n gegen x_0 strebt, so strebt $f(x_n)$ gegen $f(x_0)$.

Noch einmal dasselbe mit dem Limeszeichen geschrieben:

$$\lim_{n \rightarrow \infty} x_n = x_0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} f(x_n) = f(x_0).$$

Setzt man in der letzten Gleichung für x_0 den Ausdruck $\lim_{n \rightarrow \infty} x_n$ ein, so erhält man eine Eigenschaft stetiger Funktionen, die in Rechnungen häufig benutzt wird:

$$\boxed{\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right).} \quad (\star\star)$$

Liegt eine stetige Funktion vor, so kann man in Rechnungen also das Limeszeichen und die Bezeichnung für die Funktion vertauschen.

Beispiel. Die Funktion \cos ist auf ganz \mathbb{R} stetig. Es gelte $x_n \rightarrow 0$. Es soll $\lim_{n \rightarrow \infty} \cos(x_n)$ bestimmt werden. Das gelingt mithilfe von $(\star\star)$:

$$\lim_{n \rightarrow \infty} \cos(x_n) \stackrel{(\star\star)}{=} \cos\left(\lim_{n \rightarrow \infty} x_n\right) = \cos(0) = 1.$$

2 Differentialrechnung

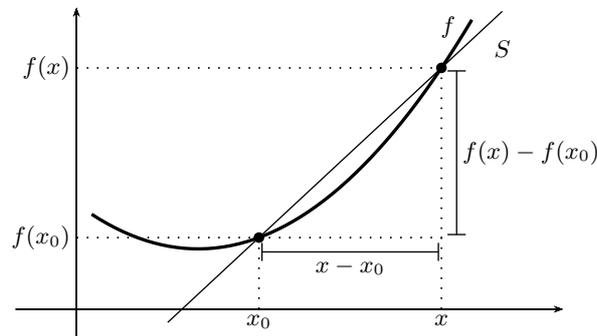
Wie findet man für eine reelle Funktion f die Tangente an den Graphen von f im Punkt $(x, f(x))$? Wie lässt sich überhaupt sinnvoll definieren, was unter der Tangente im Punkt $(x, f(x))$ zu verstehen ist? Wie berechnet man die momentane Geschwindigkeit eines sich bewegenden Körpers? Wie berechnet man seine momentane Beschleunigung? Was ist unter „momentaner Geschwindigkeit“ und „momentaner Beschleunigung“ überhaupt zu verstehen? Die Behandlung dieser (und vieler anderer) Fragen erfolgt mithilfe der Differentialrechnung, die auf NEWTON (1643-1727) und LEIBNIZ (1646-1716) zurückgeht.

2.1 Die Ableitung einer Funktion

Gegeben sei eine reelle Funktion $f : D \rightarrow \mathbb{R}$; $x_0 \in D$ sei fest gewählt. Zu jedem $x \in D$ mit $x \neq x_0$ können wir den *Differenzenquotienten* bilden:

$$\frac{f(x) - f(x_0)}{x - x_0}. \quad (2.1)$$

Veranschaulichen können wir uns den Differenzenquotienten als Steigung („Verhältnis der Änderung von f zur Änderung von x “) der Geraden durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$. Man nennt eine solche Gerade eine *Sekante* im Punkt $(x_0, f(x_0))$. Wir wollen sie hier mit S bezeichnen.



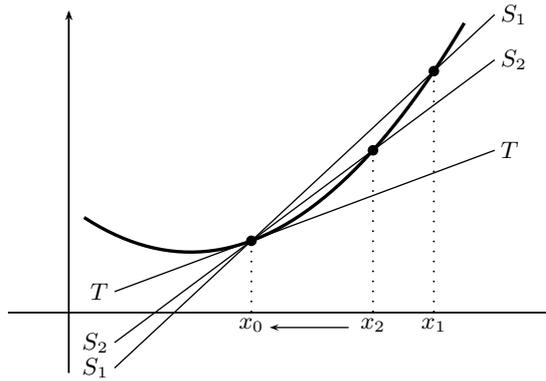
Weiter unten in der grundlegenden Definition 1 werden wir den folgenden Grenzwert betrachten:

$$\lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right). \quad (2.2)$$

Falls dieser Grenzwert (2.2) existiert, d.h., falls es ein $a \in \mathbb{R}$ gibt, für das

$$\lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right) = a$$

gilt, werden wir a als *Ableitung von f an der Stelle x_0* bezeichnen. Veranschaulichen können wir uns a als Steigung der Geraden T , an die sich die Sekante S annähert, wenn man x gegen x_0 streben lässt.



Man definiert: T ist die *Tangente* an f im Punkt $(x_0, f(x_0))$. Ihre Steigung bezeichnet man mit $f'(x_0)$. Also (falls dieser Grenzwert existiert):

$$f'(x_0) = a = \lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right). \quad (2.3)$$

In der folgenden grundlegenden Definition werden die obigen Betrachtungen zusammengefasst und erweitert; auf die der Motivation und Veranschaulichung dienenden geometrischen Betrachtungen wird dabei keinerlei Bezug genommen.

Definition 1.

(a) Eine reelle Funktion f heißt *differenzierbar an der Stelle* $x_0 \in D(f)$, wenn der Grenzwert

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

existiert, d.h., wenn es ein $a \in \mathbb{R}$ gibt mit $\lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right) = a$. Wir bezeichnen diesen Grenzwert a mit $f'(x_0)$ und nennen ihn die *Ableitung von f an der Stelle x_0* .

(b) f heißt *differenzierbar auf* $X \subseteq D(f)$, wenn f an jeder Stelle $x_0 \in X$ differenzierbar ist.

(c) Zu f lässt sich eine Funktion f' mit $D(f') = \{x_0 \in D(f) : f'(x_0) \text{ existiert}\}$ definieren, indem man jedem x_0 den Wert $f'(x_0)$ zuordnet. Diese Funktion f' nennt man die *Ableitung von f* .

Um nachzuweisen, dass f an der Stelle x_0 die Ableitung $a \in \mathbb{R}$ besitzt, muss man den Grenzwert $\lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right)$ berechnen. Man hat zu zeigen, dass für jede Folge (x_n) mit $x_n \in D(f)$, $x_n \neq x_0$ und $x_n \rightarrow x_0$ gilt (siehe Definition 3 aus Abschnitt 1.4):

$$\lim_{n \rightarrow \infty} \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) = a.$$

Beispiele:

1. Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine konstante Funktion, d.h., $f(x) = c$ gelte für alle $x \in \mathbb{R}$. Dann ist f für alle $x_0 \in \mathbb{R}$ differenzierbar und es gilt $f'(x_0) = 0$. Denn: Ist (x_n) eine Folge mit $x_n \rightarrow x_0$ und $x_n \neq x_0$, so folgt:

$$\lim_{n \rightarrow \infty} \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) = \lim_{n \rightarrow \infty} \left(\frac{c - c}{x_n - x_0} \right) = 0.$$

2. Die Funktion $f(x) = x$ ist auf ganz \mathbb{R} differenzierbar und es gilt $f'(x_0) = 1$ für alle $x_0 \in \mathbb{R}$. Denn: Ist (x_n) eine Folge mit $x_n \rightarrow x_0$ und $x_n \neq x_0$, so folgt:

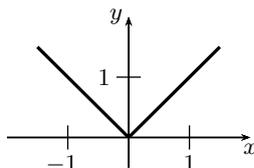
$$\lim_{n \rightarrow \infty} \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) = \lim_{n \rightarrow \infty} \left(\frac{x_n - x_0}{x_n - x_0} \right) = 1.$$

3. Die Funktion $f(x) = x^2$ ist auf ganz \mathbb{R} differenzierbar und es gilt $f'(x_0) = 2x_0$ für alle $x_0 \in \mathbb{R}$. Denn: Ist (x_n) eine Folge mit $x_n \rightarrow x_0$ und $x_n \neq x_0$, so folgt:

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) &= \lim_{n \rightarrow \infty} \left(\frac{x_n^2 - x_0^2}{x_n - x_0} \right) \\ &= \lim_{n \rightarrow \infty} \left(\frac{(x_n + x_0)(x_n - x_0)}{x_n - x_0} \right) \\ &= \lim_{n \rightarrow \infty} (x_n + x_0) \\ &= 2x_0 \end{aligned}$$

4. Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $f(x) = |x|$ („Betragfunktion“).

Graph von f :



Die Betragfunktion ist an der Stelle $x_0 = 0$ nicht differenzierbar, denn für die Folge $x_n = \frac{1}{n}$ gilt:

$$\lim_{n \rightarrow \infty} \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) = \lim_{n \rightarrow \infty} \left(\frac{\left| \frac{1}{n} \right| - |0|}{\frac{1}{n} - 0} \right) = \lim_{n \rightarrow \infty} \left(\frac{\frac{1}{n} - 0}{\frac{1}{n} - 0} \right) = 1.$$

Andererseits gilt für die Folge $x_n = -\frac{1}{n}$:

$$\lim_{n \rightarrow \infty} \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) = \lim_{n \rightarrow \infty} \left(\frac{\left| -\frac{1}{n} \right| - |0|}{-\frac{1}{n} - 0} \right) = \lim_{n \rightarrow \infty} \left(\frac{\frac{1}{n} - 0}{-\frac{1}{n} - 0} \right) = -1.$$

Also existiert $\lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right)$ nicht.

Satz 1.

Ist f auf einer Menge X differenzierbar, so ist f auf X stetig.

Beweis. Es sei $x_0 \in X$ und $(x_n)_{n \in \mathbb{N}}$ sei eine Folge mit $x_n \in D(f)$ mit $\lim_{n \rightarrow \infty} x_n = x_0$. Es ist $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$ zu zeigen. Es genügt, dies für Folgen (x_n) zu zeigen, für die $x_n \neq x_0$ für alle $n \in \mathbb{N}$ gilt, was wir im Folgenden annehmen. Die Behauptung $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$ ist gleichbedeutend mit:

$$\lim_{n \rightarrow \infty} (f(x_n) - f(x_0)) = 0.$$

Die Richtigkeit dieser Gleichung ergibt sich folgendermaßen aus der Differenzierbarkeit von f an x_0 :

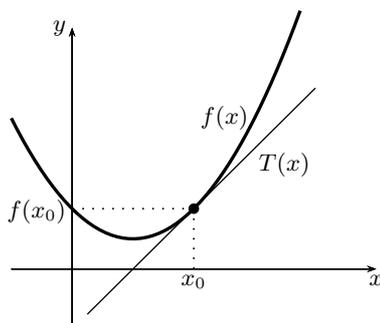
$$\begin{aligned} \lim_{n \rightarrow \infty} (f(x_n) - f(x_0)) &= \lim_{n \rightarrow \infty} \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} (x_n - x_0) \right) \\ &= \lim_{n \rightarrow \infty} \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) \cdot \lim_{n \rightarrow \infty} (x_n - x_0) \\ &= f'(x_0) \cdot 0 \\ &= 0 \quad \square \end{aligned}$$

Die Umkehrung von Satz 1 ist falsch: Aus der Stetigkeit einer Funktion folgt i.A. nicht ihre Differenzierbarkeit. Beispielsweise ist die Betragfunktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = |x|$, an der Stelle $x_0 = 0$ zwar stetig, aber nicht differenzierbar.

Bemerkung zur geometrischen Interpretation der Ableitung.

Ist f an der Stelle x_0 differenzierbar, so erhält man die folgende Gleichung der Tangente $T(x)$ im Punkt $(x_0, f(x_0))$:

$$T(x) = f(x_0) + f'(x_0) \cdot (x - x_0) \quad (2.4)$$



Höhere Ableitungen.

Diese werden rekursiv definiert, d.h. f'' ist die Ableitung von f' , f''' ist die Ableitung von f'' usw.; allgemein definiert man:

$$f^{(n+1)} := (f^{(n)})'$$

Die Funktion f selber bezeichnet man auch als 0-te Ableitung, d.h. $f^{(0)} = f$.

Alternative Schreibweisen.

1. Anstelle von $f'(x_0)$ schreibt man auch $\frac{df}{dx}(x_0)$; ebenfalls üblich sind $\frac{dy}{dx}(x_0)$, $Df(x_0)$ oder $y'(x_0)$.
2. Die Definition der Ableitung wird oft auch folgendermaßen geschrieben: Man setzt $x = x_0 + h$; dann bedeutet $h \rightarrow 0$ dasselbe wie $x \rightarrow x_0$. Es folgt:

$$f'(x_0) = \lim_{h \rightarrow 0} \left(\frac{f(x_0 + h) - f(x_0)}{h} \right).$$

Schreibt man statt x_0 einfach x , so ergibt sich folgende oft gebrauchte Schreibweise:

$$f'(x) = \lim_{h \rightarrow 0} \left(\frac{f(x + h) - f(x)}{h} \right).$$

2.2 Ableitungsregeln

Bei der Berechnung der Ableitung einer Funktion muss man nicht jedes Mal auf die Definition von f' zurückgreifen; stattdessen benutzt man Ableitungsregeln. Diese werden im Folgenden zusammengestellt.

Satz 2.

Die Funktionen f und g seien beide im Intervall $[a, b]$ definiert (d.h., es gilt sowohl $[a, b] \subseteq D(f)$ als auch $[a, b] \subseteq D(g)$); ferner seien beide Funktionen an der Stelle $x_0 \in [a, b]$ differenzierbar. Dann sind auch die Funktionen $f+g$, $f-g$, $c \cdot f$ und $f \cdot g$ an der Stelle x_0 differenzierbar ($c \in \mathbb{R}$); sofern $g(x_0) \neq 0$ gilt, ist auch $\frac{f}{g}$ an der Stelle x_0 differenzierbar. Es gelten die folgenden Regeln:

- (a) $(f+g)'(x_0) = f'(x_0) + g'(x_0)$ (Summenregel)
- (b) $(f-g)'(x_0) = f'(x_0) - g'(x_0)$
- (c) $(c \cdot f)'(x_0) = c \cdot f'(x_0)$
- (d) $(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0)$ (Produktregel)
- (e) $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0) \cdot g(x_0) - f(x_0) \cdot g'(x_0)}{(g(x_0))^2}$ (Quotientenregel)

Beweis. Wir führen den Beweis nur für $f+g$ und $f \cdot g$ durch. Die Beweise der übrigen Aussagen sind ähnlich bzw. lassen sich auf die Aussage für $f+g$ und $f \cdot g$ zurückführen.

Es sei (x_n) eine Folge mit $x_n \in D(f) \cap D(g)$, $x_n \neq x_0$ und $x_n \rightarrow x_0$.

(a)

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

(d) Da g an der Stelle x_0 differenzierbar ist, ist g an x_0 stetig (Satz 1). Also gilt $g(x_n) \rightarrow g(x_0)$.

Es folgt:

$$\begin{aligned} \frac{(f \cdot g)(x_n) - (f \cdot g)(x_0)}{x_n - x_0} &= \frac{f(x_n) \cdot g(x_n) - f(x_0) \cdot g(x_0)}{x_n - x_0} \\ &= \frac{f(x_n) - f(x_0)}{x_n - x_0} \cdot g(x_n) + f(x_0) \cdot \frac{g(x_n) - g(x_0)}{x_n - x_0} \\ &\rightarrow f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0) \end{aligned}$$

Satz 3.

Es sei $n \in \mathbb{N}$; $f: \mathbb{R} \rightarrow \mathbb{R}$ sei definiert durch $f(x) = x^n$. Dann ist f auf ganz \mathbb{R} differenzierbar und es gilt:

$$f'(x) = n \cdot x^{n-1}$$

Beweis. Wir verwenden vollständige Induktion:

- (I) Für $n = 1$ ist die Behauptung wegen Beispiel 2 in Abschnitt 2.1 richtig. (Man beachte $x^0 = 1$ für alle $x \in \mathbb{R}$.)
- (II) Für ein festes n sei die Behauptung richtig, d.h. $(x_n)' = nx^{n-1}$. Wir zeigen, dass die Behauptung

dann auch für $n + 1$ gilt:

$$\begin{aligned}
 (x^{n+1})' &= (x^n \cdot x)' \\
 &\stackrel{\text{Produktregel}}{=} (x^n)' \cdot x + x^n \cdot 1 \\
 &\stackrel{\text{Ind.annahme}}{=} nx^{n-1} \cdot x + x^n \\
 &= nx^n + x^n \\
 &= (n + 1) \cdot x^n \quad \square
 \end{aligned}$$

Als eine weitere Konsequenz der Regeln aus Satz 2 erhalten wir den folgenden Satz.

Satz 4.

- (a) Polynome sind auf ganz \mathbb{R} differenzierbar. Aus $P(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ folgt $P'(x) = a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1}$.
- (b) Rationale Funktionen $R(x) = \frac{P(x)}{Q(x)}$ sind an allen Stellen ihres Definitionsbereichs differenzierbar.

Beweis. Satz 4 folgt unmittelbar aus Satz 3 und Beispiel 1 (Abschnitt 2.1) durch wiederholte Anwendung der Regeln aus Satz 2. \square

Die Ableitung einer rationalen Funktion $R(x) = \frac{P(x)}{Q(x)}$ erfolgt mit der Quotientenregel.

Zwei weitere wichtige Ableitungsregeln sind die *Kettenregel* und die *Umkehrregel*.

Als Vorbereitung für die Kettenregel betrachten wir noch einmal den Begriff der Komposition zweier Funktionen.

Definition.

Es seien A, B, C und D Mengen und $f : A \rightarrow B$ und $g : C \rightarrow D$ seien Funktionen. Es gelte $f(A) \subseteq C$. Dann versteht man unter der *Komposition* $g \circ f$ diejenige Funktion $A \rightarrow D$, die wie folgt definiert wird:

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

Beispiel. Es seien $f : \mathbb{R} \rightarrow \mathbb{R}$ und $g : \mathbb{R}^+ \rightarrow \mathbb{R}$ gegeben durch $f(x) = x^2 + 1$ und $g(x) = \sqrt{x}$. (\mathbb{R}^+ bezeichne die Menge der nichtnegativen reellen Zahlen.) Da $f(\mathbb{R}) \subseteq \mathbb{R}^+$ gilt, lässt sich $g \circ f$ bilden; es gilt:

$$(g \circ f)(x) = \sqrt{x^2 + 1}.$$

Man nennt $g \circ f$ auch eine zusammengesetzte Funktion. f heißt *innere* und g heißt *äußere Funktion*. Man differenziert zusammengesetzte Funktionen nach der *Kettenregel*.

Satz 5 (Kettenregel).

Es seien $f : X \rightarrow \mathbb{R}$ und $g : Y \rightarrow \mathbb{R}$ Funktionen mit $f(X) \subseteq Y$. Die Funktion f sei differenzierbar an $x_0 \in X$ und g sei differenzierbar an $y_0 = f(x_0)$. Dann ist die zusammengesetzte Funktion $g \circ f$ differenzierbar an x_0 und es gilt:

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

Für den Beweis der Kettenregel verweisen wir auf O. Forster, Analysis I oder irgendein anderes gängiges Lehrbuch (siehe Literaturverzeichnis am Ende des Skripts).

Beispiele zur Anwendung der Kettenregel.

1. $h : \mathbb{R} \rightarrow \mathbb{R}$ sei definiert durch $h(x) = (3x^3 - 5x + 2)^7$.

Es gilt $h(x) = g(f(x))$ für die Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ und $g : \mathbb{R} \rightarrow \mathbb{R}$, die durch

$$f(x) = 3x^3 - 5x + 2 \quad \text{und} \quad g(y) = y^7$$

gegeben sind. Anwendung der Kettenregel ergibt

$$\begin{aligned} h'(x) &= (g \circ f)'(x) \\ &= g'(f(x)) \cdot f'(x) \\ &= 7 \cdot (3x^3 - 5x + 2)^6 \cdot (9x^2 - 5) \end{aligned}$$

2. Es soll $h(x) = \left(\frac{4x^3 + 1}{x + 1}\right)^{13}$ differenziert werden.

Unter Anwendung der Ketten- und Quotientenregel erhält man:

$$\begin{aligned} h'(x) &= 13 \left(\frac{4x^3 + 1}{x + 1}\right)^{12} \cdot \left(\frac{4x^3 + 1}{x + 1}\right)' \\ &= 13 \left(\frac{4x^3 + 1}{x + 1}\right)^{12} \cdot \frac{12x^2(x + 1) - (4x^3 + 1)}{(x + 1)^2} \\ &= 13 \left(\frac{4x^3 + 1}{x + 1}\right)^{12} \cdot \frac{8x^3 + 12x^2 - 1}{(x + 1)^2} \end{aligned}$$

Zur Formulierung der Umkehrregel benötigen wir die folgenden Begriffe ($X \subseteq D(f)$).

Definition.

(a) f heißt *monoton wachsend* auf X , falls für alle $x_1, x_2 \in X$ mit $x_1 < x_2$ gilt:

$$f(x_1) \leq f(x_2).$$

(b) f heißt *streng monoton wachsend* auf X , falls für alle $x_1, x_2 \in X$ mit $x_1 < x_2$ gilt:

$$f(x_1) < f(x_2).$$

(c) Analog definiert man *monoton fallend* und *streng monoton fallend*.

(d) f heißt *(streng) monoton*, falls f entweder (streng) monoton wachsend oder (streng) monoton fallend ist.

Es ist ebenfalls üblich, monoton *steigend* statt monoton wachsend zu sagen. Unmittelbar aus den Definitionen folgt: Ist f streng monoton auf X , so ist f injektiv und besitzt daher eine Umkehrfunktion.

Satz 6.

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ sei stetig und streng monoton auf I (d.h. insbesondere, dass die Umkehrfunktion f^{-1} existiert; f^{-1} bildet die Menge $f(I) = \{y \in \mathbb{R} : \text{Es gibt ein } x \in I \text{ mit } f(x) = y\}$ auf I ab).

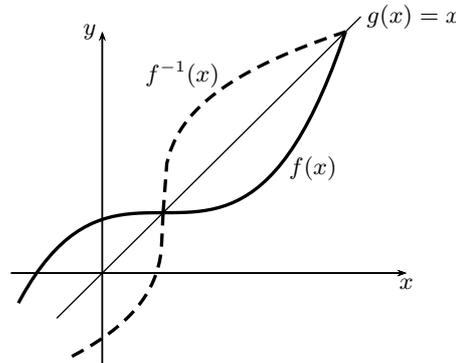
Dann gilt:

- (a) Das Bild $f(I) = \{y \in \mathbb{R} : \text{Es gibt ein } x \in I \text{ mit } f(x) = y\}$ ist ebenfalls ein Intervall.
- (b) f^{-1} ist stetig auf $f(I)$.
- (c) Ist f streng monoton wachsend (fallend), so ist auch f^{-1} streng monoton wachsend (fallend).
- (d) Ist f an $x_0 \in I$ differenzierbar und gilt $f'(x_0) \neq 0$, so ist f^{-1} differenzierbar an $y_0 = f(x_0)$ und es folgt:

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Die Aussage (d) nennt man *Umkehrregel*. Wir verzichten hier auf Beweise der Aussagen (a) - (d) und verweisen auf die Literatur.

Zur graphischen Darstellung der Umkehrfunktion f^{-1} : Man erhält den Graphen von f^{-1} , indem man den Graphen von f an der Geraden $g(x) = x$ spiegelt.



Beispiel zur Umkehrregel. Es sei $I = (0, \infty)$ und $f : I \rightarrow \mathbb{R}$ sei definiert durch $f(x) = x^2$.

Dann ist f stetig auf I (Satz 11, Abschnitt 1.5) und streng monoton (Satz 1, Abschnitt 1.2). Die Umkehrfunktion $f^{-1} : (0, \infty) \rightarrow (0, \infty)$ ist die Wurzelfunktion $\sqrt{y} = y^{\frac{1}{2}}$ (siehe auch den folgenden Abschnitt 2.3). Es sei $y = x^2$ bzw. $\sqrt{y} = x$. Nach der Umkehrregel gilt für die Ableitung von f^{-1} an der Stelle $y \neq 0$:

$$(\sqrt{y})' = \frac{1}{(x^2)'} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}.$$

Schreibt man hierin x statt y , so hat sich also die folgende Ableitungsregel für die Funktion $f(x) = \sqrt{x}$ ergeben:

$$(\sqrt{x})' = \frac{1}{2\sqrt{x}} \quad (\text{für } x > 0).$$

2.3 Die Funktionen a^x , $\log_a x$, x^r und ihre Ableitungen

Ist a eine reelle Zahl und n eine natürliche Zahl, so definiert man:

$$a^n = \underbrace{a \cdot a \cdot \dots \cdot a}_{n \text{ Faktoren}} \quad (2.5)$$

$$a^{-n} = \frac{1}{a^n} \quad (\text{falls } a \neq 0) \quad (2.6)$$

Ferner definiert man:

$$a^0 = 1. \tag{2.7}$$

Hierdurch sind also Potenzen a^x für ganzzahlige Exponenten x definiert. Es gelten folgende Rechenregeln, die unmittelbar aus den Definitionen (2.5) - (2.7) folgen ($x, y \in \mathbb{Z}$):

$$a^x \cdot a^y = a^{x+y} \tag{2.8}$$

$$a^x \cdot b^x = (a \cdot b)^x \tag{2.9}$$

$$(a^x)^y = a^{x \cdot y} \tag{2.10}$$

Als nächstes wollen wir Potenzen a^x für rationale Exponenten x definieren. Hierzu benötigen wir zunächst den folgenden grundlegenden Satz, der die Existenz der q -ten Wurzel ($q \in \mathbb{N}$) einer nichtnegativen reellen Zahl a sichert.

Satz 7.

Es sei $a \in \mathbb{R}$ mit $a \geq 0$ und $q \in \mathbb{N}$. Dann gibt es genau eine reelle Zahl $x \geq 0$, die die Gleichung $x^q = a$ erfüllt.

Beweis. Es sei $I = [0, \infty)$. Die Funktion $f : I \rightarrow \mathbb{R}$ mit $f(x) = x^q$ ist stetig (Satz 11, Abschnitt 1.5) und streng monoton steigend. Nach Satz 6 (a) ist $f(I)$ ein Intervall. Wegen $f(0) = 0$ und $\lim_{x \rightarrow \infty} f(x) = \infty$ folgt $f(I) = [0, \infty)$. Hieraus folgt Satz 7. \square

Definition.

Die eindeutig bestimmte Zahl x aus Satz 7 heißt die q -te Wurzel aus a , geschrieben $x = \sqrt[q]{a}$.

Man beachte, dass der Ausdruck $\sqrt[q]{a}$ nur definiert ist, wenn $a \geq 0$ ist. Definitionsgemäß gilt $\sqrt[q]{a} \geq 0$.

Nachdem wir durch Satz 7 die Existenz der q -ten Wurzel aus a ($a \geq 0$) gesichert haben, kommen wir nun wie oben angekündigt zur Definition von a^x für rationale Exponenten x .

Definition.

Es sei x eine rationale Zahl mit $x = \frac{p}{q}$, wobei $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ gilt. Dann setzen wir

$$a^x = a^{\frac{p}{q}} = (\sqrt[q]{a})^p \quad (\text{für } a > 0).$$

Wir haben bisher die Potenz a^x für $x \in \mathbb{Q}$ definiert ($a > 0$). Von Ausdrücken wie beispielsweise

$$2^{\frac{1}{2}}, \quad 2^{-5} \quad \text{oder} \quad 2^{-\frac{101}{99}}$$

haben wir also festgelegt, was sie bedeuten sollen. Was aber bedeuten beispielsweise die Ausdrücke

$$2^{\sqrt{2}}, \quad 2^{\sqrt[3]{5}}, \quad 2^\pi \quad \text{oder} \quad 2^{-\sqrt{2}} ?$$

Auch diese Ausdrücke können sinnvoll definiert werden: Für $a, x \in \mathbb{R}$ (mit $a > 0$) ist es möglich, a^x so zu definieren, dass die Regeln (2.8), (2.9) und (2.10) für das Rechnen mit Potenzen auch für beliebige reelle Exponenten gültig bleiben. Wie dies geschieht, wird im Folgenden erläutert.

Jede reelle Zahl x lässt sich eindeutig als ein unendlicher Dezimalbruch

$$x = r_0, r_1 r_2 r_3 \dots \tag{2.11}$$

darstellen, für den $r_0 \in \mathbb{Z}$ und $r_i \in \{0, 1, \dots, 9\}$ ($i \in \mathbb{N}$) gilt und der nicht die Periode 9 besitzt¹. Bricht

¹Beweise dieser Feststellung sowie weitere Informationen über die Dezimaldarstellung reeller Zahlen findet man in Lehrbüchern der Analysis, beispielsweise in E. Behrends, Analysis 1.

man diese Darstellung nach der n -ten Stelle ab, so erhält man eine rationale Zahl s_n :

$$\begin{aligned} s_0 &= r_0 \\ s_1 &= r_0, r_1 \\ s_2 &= r_0, r_1 r_2 \\ s_3 &= r_0, r_1 r_2 r_3 \\ &\vdots \\ s_n &= r_0, r_1 r_2 r_3 \dots r_n \end{aligned}$$

Die Folge (s_n) ist eine Folge von rationalen Zahlen, die gegen x konvergiert. Für gegebenes $a > 0$ können wir (wegen $s_n \in \mathbb{Q}$) die Potenzen

$$a^{s_0}, a^{s_1}, a^{s_2}, \dots$$

bilden. Von dieser Folge lässt sich zeigen, dass sie monoton und beschränkt ist, also nach Satz 5 (Abschnitt 1.3) einen Grenzwert besitzt. Diesen Grenzwert wählt man als a^x , d.h., man definiert a^x wie folgt.

Definition.

Für $a, x \in \mathbb{R}$, $a > 0$, definieren wir:

$$a^x = \lim_{n \rightarrow \infty} a^{s_n}. \tag{2.12}$$

(s_n) ist dabei die obige Folge rationaler Zahlen mit $\lim_{n \rightarrow \infty} s_n = x$. Ferner setzen wir für $x > 0$:

$$0^x = 0.$$

Wir erläutern die obige Definition am Beispiel 2^π . (π ist die bekannte Kreiszahl.) Es gilt $\pi \notin \mathbb{Q}$; die ersten Dezimalstellen: $\pi = 3,14159\dots$. Für s_n ($n = 0, 1, \dots, 5$) ergibt sich:

s_0	$=$	3	2^3	$=$	8
s_1	$=$	$3,1$	$2^{3,1}$	$=$	$8,57$
s_2	$=$	$3,14$	$2^{3,14}$	$=$	$8,815$
s_3	$=$	$3,141$	$2^{3,141}$	$=$	$8,82135$
s_4	$=$	$3,1415$	$2^{3,1415}$	$=$	$8,82441$
s_5	$=$	$3,14159$	$2^{3,14159}$	$=$	$8,82496$
		\downarrow			\downarrow
		π			2^π

Die Regeln (2.8), (2.9) und (2.10) für das Rechnen mit Potenzen bleiben auch für reelle Exponenten gültig. Dies halten wir noch einmal ausdrücklich in folgendem Satz fest. (Für den Beweis verweisen wir auf die Literatur.)

Satz 8.

Es seien a und b positive reelle Zahlen; x und y seien beliebige reelle Zahlen. Dann gilt:

$$a^x \cdot a^y = a^{x+y} \tag{2.13}$$

$$a^x \cdot b^x = (a \cdot b)^x \tag{2.14}$$

$$(a^x)^y = a^{x \cdot y} \tag{2.15}$$

2.3.1 Die Exponentialfunktion a^x

Wir kommen nun zur Einführung der Exponentialfunktion a^x .

Definition.

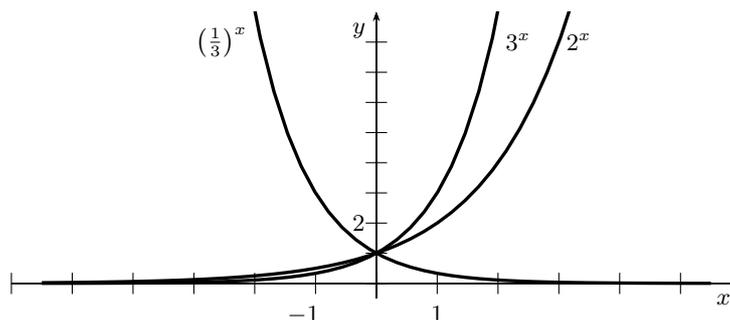
Es sei $a > 0$ eine reelle Zahl. Dann bezeichnet man die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, die durch

$$f(x) = a^x$$

gegeben ist, als *Exponentialfunktion* (zur Basis a).

Auch Funktionen der Form $c \cdot a^x$ werden gelegentlich als Exponentialfunktionen bezeichnet. Man verwechsle die Exponentialfunktionen nicht mit den Potenzfunktionen; Potenzfunktionen sind Funktionen der Form $f(x) = x^r$. Der Unterschied: Bei den Potenzfunktionen ist die Basis variabel, bei den Exponentialfunktionen der Exponent.

Skizze² der Exponentialfunktionen a^x für $a = 2$, $a = 3$ sowie $a = \frac{1}{3}$:

**Satz 9.**

- (a) Die Exponentialfunktionen a^x sind stetig auf ganz \mathbb{R} .
- (b) Ist $a > 1$, so ist die Exponentialfunktion a^x streng monoton steigend; für $a < 1$ ist die Exponentialfunktion a^x streng monoton fallend.

Für den Beweis verweisen wir auf die Standardlehrbücher der Analysis (siehe Literaturverzeichnis).

2.3.2 Die Logarithmusfunktion $\log_a x$

Da die Exponentialfunktion a^x für $a \neq 1$ streng monoton ist, lässt sich die Umkehrfunktion bilden. Diese bezeichnet man als *Logarithmus zur Basis a* .

Definition.

Ist $a \in \mathbb{R}$, $a > 0$ und $a \neq 1$, so bezeichnet man die Umkehrfunktion der Exponentialfunktion a^x als *Logarithmus zur Basis a* . Bezeichnung:

$$\log_a x.$$

Der Definitionsbereich von $\log_a x$ ist $(0, \infty)$. Dies ergibt sich aus der Tatsache, dass $a^x > 0$ für $a > 0$ und alle $x \in \mathbb{R}$ gilt, sowie aus den folgenden beiden Feststellungen (unter Benutzung von Satz 6(a), Abschnitt 2.2):

$$\text{Ist } a > 1, \text{ so gilt } \lim_{x \rightarrow -\infty} a^x = 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} a^x = \infty. \quad (2.16)$$

$$\text{Ist } a < 1, \text{ so gilt } \lim_{x \rightarrow -\infty} a^x = \infty \quad \text{und} \quad \lim_{x \rightarrow \infty} a^x = 0. \quad (2.17)$$

²Zur Übung: Skizzieren Sie zusätzlich a^x für $a = \frac{1}{4}, \frac{1}{2}, \frac{2}{3}, 1, \frac{3}{2}, 4$.

Als Umkehrfunktion der stetigen Funktion a^x ist die Logarithmusfunktion $\log_a x$ ebenfalls stetig (vgl. Satz 6 (b)). Eine besonders wichtige Rolle spielen die Exponentialfunktionen und die zugehörigen Logarithmusfunktionen zu den Basen e , 10 und 2 :

$$\log_e x, \quad \log_{10} x \quad \text{und} \quad \log_2 x.$$

Anstelle von $\log_e x$ schreibt man oft auch:

$$\ln x.$$

$\ln x$ heißt *natürlicher Logarithmus*.

Empfehlung: Fertigen Sie zunächst Skizzen der Funktionen e^x und 2^x an und skizzieren Sie anschließend die Funktionen $\ln x$ und $\log_2 x$.

Aus den Rechenregeln (2.13) und (2.15) für a^x (Satz 8) erhält man die folgenden Rechenregeln für $\log_a x$.

Satz 10.

Es seien $a, x, y, r \in \mathbb{R}$ mit $a > 0$, $x > 0$, $y > 0$ und $a \neq 1$; dann gilt:

$$\log_a (x \cdot y) = \log_a x + \log_a y, \quad (2.18)$$

$$\log_a (x^r) = r \cdot \log_a x. \quad (2.19)$$

Beweis. Es gilt:

$$\begin{aligned} a^{\log_a (x \cdot y)} &= x \cdot y \\ &= a^{\log_a x} \cdot a^{\log_a y} \\ &\stackrel{(2.13)}{=} a^{\log_a x + \log_a y} \end{aligned}$$

Da die Exponentialfunktion a^x streng monoton ist, folgt hieraus durch Vergleich der Exponenten die Aussage (2.18). Der Beweis von (2.19) sei dem Leser als Übungsaufgabe empfohlen. (*Hinweis:* Der Beweis von (2.19) geht ähnlich wie der Beweis von (2.18); anstelle von (2.13) benutzt man die Regel (2.15).) \square

Weitere Übungsaufgabe: Man zeige unter Benutzung der Sätze 8 und 10, dass die folgenden Regeln gelten:

$$\frac{a^x}{a^y} = a^{x-y}, \quad (2.20)$$

$$\log_a \left(\frac{x}{y} \right) = \log_a x - \log_a y. \quad (2.21)$$

Es gilt $x = a^{\log_a x}$. Setzt man dies in $\ln x$ ein, so erhält man folgende Beziehung zwischen dem Logarithmus zur Basis a und dem natürlichen Logarithmus:

$$\ln x = \ln (a^{\log_a x}) \stackrel{(2.19)}{=} \log_a x \cdot \ln a.$$

Stellt man diese Gleichung um, so ergibt sich:

$$\log_a x = \frac{\ln x}{\ln a}. \quad (2.22)$$

Fazit. Der Logarithmus zur Basis a unterscheidet sich vom natürlichen Logarithmus nur um einen konstanten Faktor, nämlich um den Faktor $\frac{1}{\ln a}$.

Wir kommen nun zur Differentiation der Funktionen a^x und $\log_a x$. Hierzu betrachten wir zunächst die beiden besonders wichtigen Funktionen e^x und $\ln x$.

2.3.3 Die Ableitung der Funktion $\ln x$

Aufgrund unserer Definition der Eulerschen Zahl e als Grenzwert der Folge $(1 + \frac{1}{n})^n$ (vgl. Abschnitt 1.3) gilt:

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

Dies können wir auch so aussprechen: Für die spezielle Nullfolge (x_n) mit $x_n = \frac{1}{n}$ gilt:

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

Allgemeiner gilt Folgendes (für den Beweis sei auf die bekannten Lehrbücher der Analysis verwiesen; siehe das Literaturverzeichnis am Ende dieses Skripts):

Ist (x_n) eine Nullfolge (d.h. $x_n \rightarrow 0$) und gilt $x_n \neq 0$ sowie $x_n > -1$, so folgt:

$$\lim_{n \rightarrow \infty} \left(1 + x_n\right)^{\frac{1}{x_n}} = e. \quad (2.23)$$

Wir benötigen (2.23) für den Beweis des folgenden Satzes.

Satz 11.

Es gilt für alle $x \in (0, \infty)$:

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

Beweis. Es sei $x_0 \in (0, \infty)$ und (x_n) sei eine Folge mit $x_n > 0$, $x_n \neq x_0$ und $x_n \rightarrow x_0$. Zum Beweis unseres Satzes haben wir zu zeigen:

$$\lim_{n \rightarrow \infty} \left(\frac{\ln x_n - \ln x_0}{x_n - x_0} \right) = \frac{1}{x_0}. \quad (*)$$

Das ergibt sich so:

$$\begin{aligned} \frac{\ln x_n - \ln x_0}{x_n - x_0} &\stackrel{(2.20)}{=} \frac{1}{x_n - x_0} \cdot \ln \left(\frac{x_n}{x_0} \right) \\ &= \frac{1}{x_n - x_0} \cdot \ln \left(1 + \frac{x_n - x_0}{x_0} \right) \\ &= \frac{1}{x_0} \cdot \frac{x_0}{x_n - x_0} \cdot \ln \left(1 + \frac{x_n - x_0}{x_0} \right) \end{aligned}$$

Schreiben wir zur Abkürzung $y_n = \frac{x_n - x_0}{x_0}$, so ergibt sich:

$$\begin{aligned} \frac{\ln x_n - \ln x_0}{x_n - x_0} &= \frac{1}{x_0} \cdot \frac{1}{y_n} \cdot \ln (1 + y_n) \\ &\stackrel{(2.19)}{=} \frac{1}{x_0} \cdot \ln \left(\left(1 + y_n\right)^{\frac{1}{y_n}} \right) \end{aligned}$$

Wegen (2.23) gilt $\lim_{n \rightarrow \infty} \left(1 + y_n\right)^{\frac{1}{y_n}} = e$, woraus wegen der Stetigkeit von \ln folgt:

$$\begin{aligned} \lim_{n \rightarrow \infty} \ln \left(\left(1 + y_n\right)^{\frac{1}{y_n}} \right) &= \ln \left(\lim_{n \rightarrow \infty} \left(\left(1 + y_n\right)^{\frac{1}{y_n}} \right) \right) \\ &= \ln e. \end{aligned}$$

Wegen $\ln e = 1$ haben wir also:

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\frac{\ln x_n - \ln x_0}{x_n - x_0} \right) &= \lim_{n \rightarrow \infty} \left(\frac{1}{x_0} \cdot \ln \left(\left(1 + y_n \right)^{\frac{1}{y_n}} \right) \right) \\ &= \frac{1}{x_0} \cdot \ln e \\ &= \frac{1}{x_0} \end{aligned}$$

womit (*) gezeigt ist. \square

Wir haben somit bewiesen, dass die Ableitung des natürlichen Logarithmus die Funktion $\frac{1}{x}$ ist (für $x \in (0, \infty)$). Die Ableitung der übrigen Logarithmusfunktionen $\log_a x$ (für beliebiges $a > 0$ mit $a \neq 1$) und der Exponentialfunktionen a^x ist nun einfach: Mit der Ableitung von $\ln x$ haben wir bereits die entscheidende Hürde genommen.

2.3.4 Die Ableitung der Funktion $\log_a x$

Nach (2.22) gilt $\log_a x = \frac{\ln x}{\ln a}$. Es folgt:

$$(\log_a x)' = \left(\frac{\ln x}{\ln a} \right)' = \frac{1}{\ln a} \cdot (\ln x)' = \frac{1}{\ln a} \cdot \frac{1}{x}. \quad (2.24)$$

2.3.5 Die Ableitung der Funktion e^x

Satz 12.

Für alle $x \in \mathbb{R}$ gilt:

$$(e^x)' = e^x.$$

Die Ableitung der Funktion $f(x) = e^x$ ist also die Funktion e^x selbst.

Beweis. Es sei $f(x) = \ln x = y$. Die Umkehrfunktion f^{-1} von f ist die Funktion e^y , d.h., es gilt $x = f^{-1}(y) = e^y$. Um Satz 12 zu beweisen, müssen wir also zeigen:

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

Aus Satz 11 wissen wir, dass $f'(x) = \frac{1}{x}$ gilt. Daher folgt nach der Umkehrregel (Satz 6, Abschnitt 2.2):

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

2.3.6 Die Ableitung der Funktion a^x

Aufgrund der Regel (2.19) gilt Folgendes:

$$a^x = e^{\ln(a^x)} \stackrel{(2.19)}{=} e^{x \cdot \ln a}. \quad (2.25)$$

Aus (2.25) folgt mit der Kettenregel:

$$\begin{aligned} (a^x)' &\stackrel{(2.25)}{=} (e^{x \cdot \ln a})' \\ &= e^{x \cdot \ln a} \cdot (x \cdot \ln a)' \\ &= e^{x \cdot \ln a} \cdot \ln a \\ &\stackrel{(2.25)}{=} a^x \cdot \ln a \end{aligned}$$

Insgesamt haben wir also erhalten:

$$(a^x)' = a^x \cdot \ln a. \quad (2.26)$$

2.3.7 Die Ableitung von allgemeinen Potenzfunktionen

Mit (2.24) und (2.26) haben wir die Ableitungen von $\log_a x$ und a^x gefunden. Wir bestimmen nun darüber hinaus noch die Ableitung der *allgemeinen Potenzfunktion*, d.h. der Funktion $f(x) = x^r$ für $r \in \mathbb{R}$. Ist r eine natürliche Zahl, so wissen wir, dass $f'(x) = r \cdot x^{r-1}$ für alle $x \in \mathbb{R}$ gilt. Wir zeigen nun, dass dieselbe Formel auch für beliebiges r gültig ist.

Satz 13.

Es sei r eine feste reelle Zahl. Dann gilt für alle $x \in (0, \infty)$:

$$(x^r)' = r \cdot x^{r-1}.$$

Beweis. Es gilt $x^r = e^{\ln(x^r)} \stackrel{(2.19)}{=} e^{r \cdot \ln x}$. Hieraus folgt mit der Kettenregel:

$$\begin{aligned} (x^r)' &= (e^{r \cdot \ln x})' \\ &= e^{r \cdot \ln x} \cdot (r \cdot \ln x)' \\ &= e^{r \cdot \ln x} \cdot r \cdot \frac{1}{x} \\ &= x^r \cdot r \cdot \frac{1}{x} \\ &= r \cdot x^{r-1} \quad \square \end{aligned}$$

Zusatz. Mithilfe der Quotientenregel kann man Folgendes leicht erkennen: Ist r eine negative ganze Zahl, so gilt die Formel $(x^r)' = r x^{r-1}$ nicht nur für $x \in (0, \infty)$, sondern ebenfalls für $x \in (-\infty, 0)$.

Beispiele:

1. Es sei $f(x) = 2^{-x^2}$ mit $D(f) = \mathbb{R}$.

Dann gilt nach der Kettenregel:

$$\begin{aligned} f'(x) &= 2^{-x^2} \cdot \ln 2 \cdot (-x^2)' \\ &= 2^{-x^2} \cdot \ln 2 \cdot (-2x) \\ &= -2x \cdot 2^{-x^2} \cdot \ln 2 \end{aligned}$$

2. Es sei $f(x) = x^r \cdot \log_2 x$ mit $D(f) = (0, \infty)$ und $r \in \mathbb{R}$.

Dann gilt nach der Produktregel:

$$\begin{aligned} f'(x) &= r x^{r-1} \cdot \log_2 x + x^r \cdot \frac{1}{\ln 2} \cdot \frac{1}{x} \\ &= x^{r-1} \left(r \log_2 x + \frac{1}{\ln 2} \right) \end{aligned}$$

3. Es sei $f(x) = \ln(\ln x)$ mit $D(f) = (1, \infty)$.

Dann gilt nach der Kettenregel:

$$\begin{aligned} f'(x) &= \frac{1}{\ln x} \cdot (\ln x)' \\ &= \frac{1}{x \ln x} \end{aligned}$$

4. Am Ende von Abschnitt 2.2 hatten wir mithilfe der Umkehrregel die Ableitung der Wurzelfunktion $f(x) = \sqrt{x}$ bestimmt:

$$(\sqrt{x})' = \frac{1}{2\sqrt{x}} \quad (\text{für } x > 0).$$

Bestätigen Sie dieses Ergebnis mithilfe von Satz 13.

2.3.8 Logarithmisches Differenzieren

Für eine Funktion $f : D \rightarrow \mathbb{R}$ der Form $f(x) = g(x)^{h(x)}$ soll die Ableitung berechnet werden, wobei angenommen wird, dass $g(x) > 0$ für alle $x \in D$ gilt und dass sowohl g als auch h auf D differenzierbar sind.

Beispiele für derartige Funktionen.

- $f(x) = a^x$ für $a > 0$
- $f(x) = x^x$ für $x > 0$
- $f(x) = x^{\sqrt{x+1}}$ für $x > 0$
- $f(x) = (x+1)^{x+2}$ für $x > -1$

Funktionen vom genannten Typ $f(x) = g(x)^{h(x)}$ lassen sich mit den bislang behandelten Methoden nicht ohne Weiteres ableiten. Man kann aber einen *Trick* anwenden, den wir bereits zuvor (z.B. bei der Ableitung von a^x und bei der Ableitung von x^r) verwendet haben; man formt $f(x) = g(x)^{h(x)}$ zunächst um:

$$f(x) = g(x)^{h(x)} = e^{\ln(g(x)^{h(x)})} = e^{h(x) \cdot \ln(g(x))}. \quad (*)$$

Nun kann man f ableiten:

$$\begin{aligned} f'(x) &\stackrel{(*)}{=} \left(e^{h(x) \cdot \ln(g(x))} \right)' \\ &= e^{h(x) \cdot \ln(g(x))} \cdot \left(h(x) \cdot \ln(g(x)) \right)' \\ &= e^{h(x) \cdot \ln(g(x))} \cdot \left(h'(x) \cdot \ln(g(x)) + h(x) \cdot \frac{g'(x)}{g(x)} \right) \\ &\stackrel{(*)}{=} f(x) \cdot \left(h'(x) \cdot \ln(g(x)) + h(x) \cdot \frac{g'(x)}{g(x)} \right) \end{aligned}$$

Beispiele:

1. Es sei $f(x) = x^x$ für $x > 0$.

Es folgt:

$$\begin{aligned} f'(x) &= (x^x)' \\ &= (e^{\ln(x^x)})' \\ &= (e^{x \cdot \ln x})' \\ &= e^{x \cdot \ln x} \cdot \left(1 \cdot \ln x + x \cdot \frac{1}{x} \right) \\ &= x^x \cdot (\ln x + 1) \end{aligned}$$

2. Es sei $f(x) = x^{\sqrt{x+1}}$ für $x > 0$.

Man erhält:

$$\begin{aligned} f(x) &= e^{\ln(x^{\sqrt{x+1}})} \\ &= e^{\sqrt{x+1} \cdot \ln x} \end{aligned}$$

Es folgt:

$$\begin{aligned} f'(x) &= \left(e^{\sqrt{x+1} \cdot \ln x} \right)' \\ &= e^{\sqrt{x+1} \cdot \ln x} \cdot \left(\frac{\ln x}{2\sqrt{x+1}} + \frac{\sqrt{x+1}}{x} \right) \\ &= x^{\sqrt{x+1}} \cdot \left(\frac{\ln x}{2\sqrt{x+1}} + \frac{\sqrt{x+1}}{x} \right) \end{aligned}$$

Für die in diesem Abschnitt beschriebene Methode zur Ableitung von Funktionen der Form $f(x) = g(x)^{h(x)}$ verwendet man die Bezeichnung *Logarithmisches Differenzieren*.

2.4 Die geometrische Bedeutung der ersten und zweiten Ableitung, Extremstellen und Kurvendiskussion

In diesem Abschnitt sei mit I immer ein Intervall bezeichnet, wobei die beiden uninteressanten Sonderfälle, dass I nur aus einem einzigen Punkt besteht bzw. $I = \emptyset$ gilt, ausgeschlossen seien. Ist nicht weiter über I vorausgesetzt, so kann I offen, abgeschlossen oder auch halboffen sein; ebenfalls kann I unendlich sein, etwa $I = [0, \infty)$ oder $I = (-\infty, \infty)$.

Wir wissen: Ist $f : I \rightarrow \mathbb{R}$ eine konstante Funktion, so ist die Ableitung f' überall gleich Null. Als Umkehrung hiervon gilt folgender Satz.

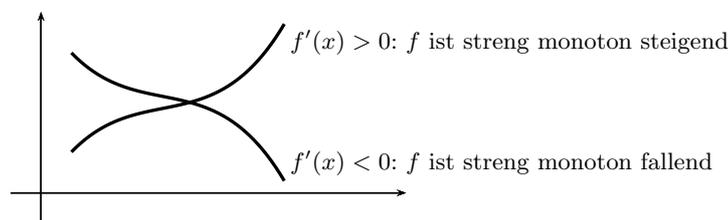
Satz 14.

Gilt $f'(x) = 0$ für alle $x \in I$, so ist f auf dem Intervall I konstant.

Dieser Satz ist anschaulich einleuchtend: Eine Funktion $f : I \rightarrow \mathbb{R}$, die nirgends steigt oder fällt, muss konstant sein.

Satz 15.

Für $a, b \in \mathbb{R}$ gelte $a < b$; die Funktion f sei stetig in $[a, b]$ und differenzierbar in (a, b) . Gilt $f'(x) > 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ streng monoton wachsend; gilt $f'(x) < 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ streng monoton fallend.



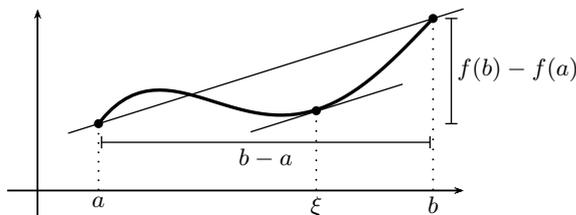
Sowohl Satz 14 als auch Satz 15 ist aufgrund der geometrischen Interpretation von $f'(x)$ als Steigung von f im Punkt $(x, f(x))$ einleuchtend. Daher wollen wir hier auf einen formalen Beweis verzichten. Wir erwähnen nur, dass formale Beweise dieser beiden Sätze vom sog. *Mittelwertsatz der Differentialrechnung*

Gebrauch machen. Da dieser Satz auch in anderen Zusammenhängen eine wichtige Rolle spielt, sei er an dieser Stelle aufgeführt.

Satz 16 (Mittelwertsatz der Differentialrechnung).

Für $a, b \in \mathbb{R}$ gelte $a < b$; die Funktion f sei stetig in $[a, b]$ und differenzierbar in (a, b) . Dann gibt es mindestens ein $\xi \in (a, b)$ mit

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



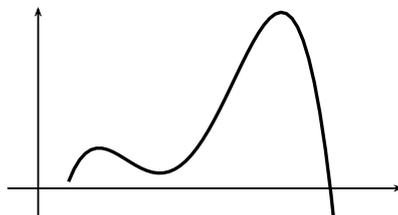
Der Mittelwertsatz ist ebenfalls anschaulich einleuchtend: $\frac{f(b)-f(a)}{b-a}$ gibt die Steigung der Geraden g durch die beiden Punkte $(a, f(a))$ und $(b, f(b))$ an. Der Mittelwertsatz besagt also, dass es (mindestens) eine Stelle $\xi \in (a, b)$ gibt, für die die Tangente im Punkt $(\xi, f(\xi))$ parallel zu g verläuft³.

Definition.

Die Funktion f sei auf dem Intervall (a, b) definiert. f hat an der Stelle $x_0 \in (a, b)$ ein *lokales Maximum*, falls es ein $\delta > 0$ gibt, für das gilt:

$$f(x_0) \geq f(x) \quad \text{für alle } x \in (x_0 - \delta, x_0 + \delta).$$

Analog: *lokales Minimum*.



Hat f an x_0 ein lokales Maximum oder ein lokales Minimum, so nennen wir x_0 ein (*lokales*) *Extremum* von f (oder auch Extremstelle von f).

Man spricht von einem *strengen lokalen Maximum*, falls Folgendes gilt:

$$f(x_0) > f(x) \quad \text{für alle } x \in (x_0 - \delta, x_0 + \delta) \quad \text{mit } x \neq x_0.$$

Analog: *strenges lokales Minimum* und *strenges (lokales) Extremum*.

Satz 17.

Die Funktion f sei in (a, b) definiert und an der Stelle $x_0 \in (a, b)$ differenzierbar. Hat f an x_0 eine Extremstelle, so gilt $f'(x_0) = 0$.

³Beweise der Sätze 14-16 findet man in fast allen Lehrbüchern der Analysis. Auch in den folgenden Abschnitten werden wir an einigen Stellen auf Beweise verzichten; in den Standardlehrbüchern der Analysis (wie etwa die Bücher von Behrends, Forster, Heuser oder Königsberger) findet man die fehlenden Beweise.

Beweis. Wir betrachten den Fall, dass f an x_0 ein lokales Maximum hat. (Der Fall, dass ein Minimum vorliegt, geht analog.) Es gibt also ein $\delta > 0$, so dass $f(x) \leq f(x_0)$ für alle $x \in (x_0 - \delta, x_0 + \delta)$. Es gilt, da f an der Stelle x_0 differenzierbar ist:

$$f'(x_0) = \lim_{n \rightarrow \infty} \left(\frac{f(x_0 + \frac{1}{n}) - f(x_0)}{\frac{1}{n}} \right) = \lim_{n \rightarrow \infty} \left(\frac{f(x_0 - \frac{1}{n}) - f(x_0)}{-\frac{1}{n}} \right). \quad (2.27)$$

Ist n hinreichend groß, so gilt $x_0 + \frac{1}{n} \in (x_0 - \delta, x_0 + \delta)$ und $x_0 - \frac{1}{n} \in (x_0 - \delta, x_0 + \delta)$ und somit $f(x_0 + \frac{1}{n}) \leq f(x_0)$ und $f(x_0 - \frac{1}{n}) \leq f(x_0)$.

Es folgt:

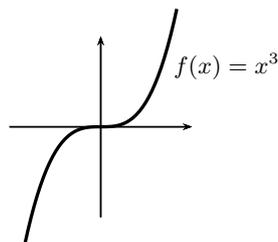
$$\frac{f(x_0 + \frac{1}{n}) - f(x_0)}{\frac{1}{n}} \leq 0 \quad \text{und} \quad \frac{f(x_0 - \frac{1}{n}) - f(x_0)}{-\frac{1}{n}} \geq 0.$$

Also gilt auch:

$$\lim_{n \rightarrow \infty} \left(\frac{f(x_0 + \frac{1}{n}) - f(x_0)}{\frac{1}{n}} \right) \leq 0 \quad \text{und} \quad \lim_{n \rightarrow \infty} \left(\frac{f(x_0 - \frac{1}{n}) - f(x_0)}{-\frac{1}{n}} \right) \geq 0.$$

Also (wegen (2.27)): $f'(x_0) = 0$. \square

Die Bedingung $f'(x_0) = 0$ ist notwendig, aber nicht hinreichend für das Vorliegen eines Extremums an der Stelle x_0 : Für die Funktion $f(x) = x^3$ gilt $f'(x) = 3x^2$, also $f'(0) = 0$. An der Stelle $x_0 = 0$ hat f aber keine Extremstelle.



Es sei x_0 eine Nullstelle der ersten Ableitung f' von f . Um zu entscheiden, ob ein Minimum, ein Maximum oder überhaupt kein Extremum vorliegt, kann man die folgenden Sätze 18 und 19 heranziehen.

Satz 18.

Die Funktion f sei in (a, b) zweimal differenzierbar. An der Stelle $x_0 \in (a, b)$ gelte $f'(x_0) = 0$.

- (a) Ist $f''(x_0) < 0$, so hat f an der Stelle x_0 ein strenges lokales Maximum.
- (b) Ist $f''(x_0) > 0$, so hat f an der Stelle x_0 ein strenges lokales Minimum.

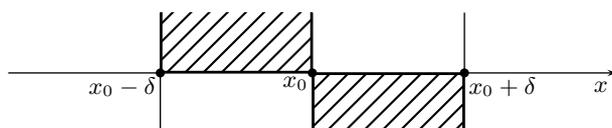
Beweis. Wir zeigen nur (a). (Der Beweis von (b) geht analog.)

Es gelte $f'(x_0) = 0$ und $f''(x_0) < 0$, also:

$$0 > f''(x_0) = \lim_{x \rightarrow x_0} \left(\frac{f'(x) - f'(x_0)}{x - x_0} \right) = \lim_{x \rightarrow x_0} \left(\frac{f'(x)}{x - x_0} \right).$$

Es folgt, dass es ein $\delta > 0$ gibt, so dass $\frac{f'(x)}{x - x_0} < 0$ für alle $x \in (x_0 - \delta, x_0 + \delta)$ mit $x \neq x_0$ gilt. (Andernfalls wäre der Grenzwert nicht kleiner als Null.) Hieraus folgt :

$$f'(x) > 0 \quad \text{für} \quad x \in (x_0 - \delta, x_0) \quad \text{und} \quad f'(x) < 0 \quad \text{für} \quad x \in (x_0, x_0 + \delta). \quad (2.28)$$



In der Zeichnung bedeutet (2.28), dass der Graph von f' im schraffierten Bereich verläuft. (Man sagt, dass f' an der Stelle x_0 einen *Vorzeichenwechsel von + nach -* hat.) Aus (2.28) und Satz 15 folgt: f ist im Intervall $[x_0 - \delta, x_0]$ streng monoton wachsend und im Intervall $[x_0, x_0 + \delta]$ streng monoton fallend. Also hat f an der Stelle x_0 ein strenges lokales Maximum. \square

Ist $f'(x_0) = f''(x_0) = 0$, so gibt Satz 18 keine Auskunft darüber, ob ein Minimum, ein Maximum oder keine Extremstelle vorliegt. Hier kann u.U. der folgende Satz 19 weiterhelfen, den wir ohne Beweis angeben.

Satz 19.

Die Funktion f sei in (a, b) n -mal differenzierbar ($n \geq 2$). An der Stelle $x_0 \in (a, b)$ gelte:

$$f'(x_0) = f''(x_0) = \dots = f^{(n-1)}(x_0) = 0 \quad \text{und} \quad f^{(n)}(x_0) \neq 0.$$

Ist n eine gerade Zahl, so hat f an der Stelle x_0 ein Extremum; und zwar:

- ein strenges lokales Maximum, falls $f^{(n)}(x_0) < 0$;
- ein strenges lokales Minimum, falls $f^{(n)}(x_0) > 0$.

Ist n ungerade, so liegt kein Extremum an der Stelle x_0 vor.

Man beachte, dass Satz 18 ein Spezialfall von Satz 19 ist: Für $n = 2$ liefert Satz 19 gerade die Aussage von Satz 18.

Beispiel. $f(x) = \frac{1}{4}x^3 - x^2$. Es gilt $f'(x) = \frac{3}{4}x^2 - 2x = x(\frac{3}{4}x - 2)$. Nullstellen von f' : $x_0 = 0$ und $x_1 = \frac{8}{3}$. Es gilt $f''(x) = \frac{3}{2}x - 2$. Wegen $f''(0) = -2$ und $f''(\frac{8}{3}) = 2$ liegt also an der Stelle $x_0 = 0$ ein Maximum und an der Stelle $x_1 = \frac{8}{3}$ ein Minimum vor.

Definition.

Ist $f : D \rightarrow \mathbb{R}$ eine Funktion, so hat f an der Stelle x_0 ein *absolutes Maximum*, falls $f(x) \leq f(x_0)$ für alle $x \in D$ gilt. Statt „absolutes Maximum“ sagt man auch *globales Maximum*.

Analog: *absolutes* (bzw. *globales*) *Minimum*.

Ist $D = [a, b]$ ein abgeschlossenes Intervall und soll das absolute Maximum oder Minimum von f auf D gefunden werden, so muss man sich auch um die Randpunkte a und b kümmern.

Beispiel. $D = [-1, 5]$, $f(x) = \frac{1}{4}x^3 - x^2$ (wie oben). Absolute Maxima bzw. Minima könnten an den Stellen $-1, 0, \frac{8}{3}$ und 5 vorliegen. Es gilt $f(-1) = -\frac{5}{4}$, $f(0) = 0$, $f(\frac{8}{3}) = -\frac{64}{27}$ und $f(5) = \frac{25}{4}$. Am Randpunkt $b = 5$ liegt also ein absolutes Maximum vor, während das absolute Minimum bei $\frac{8}{3}$ angenommen wird.

Ergänzung zu Satz 18. Es gelte $f'(x_0) = 0$. Will man, ohne von der zweiten Ableitung Gebrauch zu machen, entscheiden, ob ein Minimum oder Maximum vorliegt, so lässt sich dies häufig mithilfe des folgenden Satzes erreichen.

Satz 20.

Die Funktion f sei auf (a, b) differenzierbar und für $x_0 \in (a, b)$ gelte $f'(x_0) = 0$. Hat f' an der Stelle x_0 einen Vorzeichenwechsel von + nach - (d.h., für ein $\delta > 0$ gilt (2.28)), so hat f an der Stelle x_0 ein strenges lokales Maximum. Hat f' an der Stelle x_0 einen Vorzeichenwechsel von - nach + (dies wird analog definiert), so hat f an der Stelle x_0 ein strenges lokales Minimum.

Beweis. Man schließt wie am Schluss des Beweises von Satz 18, indem man Satz 15 anwendet. \square

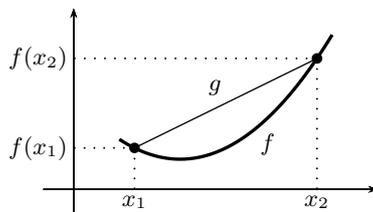
Es folgen einige Bemerkungen zur geometrischen Bedeutung der zweiten Ableitung. Zunächst einige Definitionen.

Definition.

Eine Funktion f heißt *streng konvex* auf dem Intervall I , falls für alle $x_1, x_2, x \in I$ mit $x_1 < x < x_2$ gilt:

$$f(x) < \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) + f(x_1). \quad (2.29)$$

Geometrisch bedeutet (2.29), dass zwischen x_1 und x_2 der Graph von f unterhalb der Verbindungsgeraden g der Punkte $(x_1, f(x_1))$ und $(x_2, f(x_2))$ verläuft. (Man beachte, dass der in (2.29) rechts stehende Ausdruck diese Gerade g beschreibt.)



Anschaulich kann man eine streng konvexe Funktion auch als „nach unten durchhängend“ beschreiben. Alternativ kann man auch sagen: Der Graph macht (in Richtung wachsender x -Werte) eine *Linkskurve*.

Analog definiert man, wann eine Funktion *konvex*, *streng konkav* bzw. *konkav* heißt; man braucht hierzu in (2.29) nur das Relationszeichen $<$ in \leq , $>$ bzw. \geq abzuändern.

Anschaulich kann man eine streng konkave Funktion auch als „nach oben gewölbt“ beschreiben oder man kann sagen: Der Graph macht eine *Rechtskurve* (in Richtung wachsender x -Werte).

Der folgende Satz stellt den Zusammenhang zur zweiten Ableitung f'' her.

Satz 21.

Es sei I ein Intervall und $f : I \rightarrow \mathbb{R}$ sei eine auf I zweimal differenzierbare Funktion.

- (a) Gilt $f''(x) > 0$ für alle $x \in I$, so ist f auf I streng konvex.
- (b) Gilt $f''(x) < 0$ für alle $x \in I$, so ist f auf I streng konkav.
- (c) Gilt $f''(x) \geq 0$ für alle $x \in I$, so ist f auf I konvex.
- (d) Gilt $f''(x) \leq 0$ für alle $x \in I$, so ist f auf I konkav.

Beweisskizze. Wir zeigen nur (a); (b), (c) und (d) gehen analog. Ist $f''(x) > 0$ für alle $x \in I$, so erhält man mittels Satz 15, dass f' auf I streng monoton steigend ist⁴. Es seien $x_1, x_2, x \in I$ mit $x_1 < x < x_2$. Nach dem Mittelwertsatz (Satz 16) gibt es $\xi_1 \in (x_1, x)$ und $\xi_2 \in (x, x_2)$, so dass gilt:

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) < f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}. \quad (2.30)$$

Die mittlere Ungleichung folgt aus der Tatsache, dass f' streng monoton steigend ist. Aus der Ungleichung zwischen den beiden äußeren Ausdrücken erhält man (2.29) durch eine nicht allzu schwierige Rechnung, auf deren Details hier nicht eingegangen werden soll. \square

Definition.

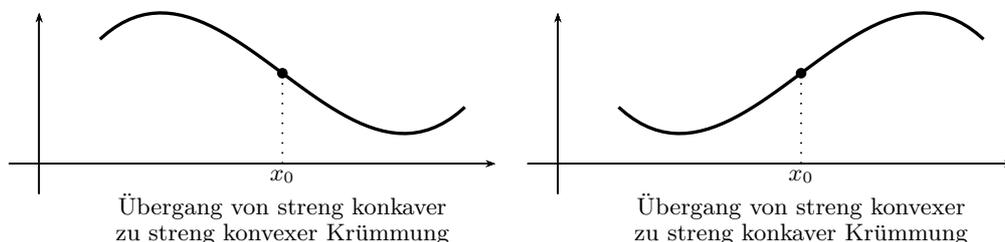
Eine Nullstelle x_0 der zweiten Ableitung von f heißt ein *Wendepunkt* von f , falls f'' in x_0 einen Vorzeichenwechsel hat, d.h. also, wenn für ein $\delta > 0$ entweder (2.31) oder (2.32) gilt:

$$f''(x) < 0 \text{ für alle } x \in (x_0 - \delta, x_0) \quad \text{und} \quad f''(x) > 0 \text{ für alle } x \in (x_0, x_0 + \delta); \quad (2.31)$$

$$f''(x) > 0 \text{ für alle } x \in (x_0 - \delta, x_0) \quad \text{und} \quad f''(x) < 0 \text{ für alle } x \in (x_0, x_0 + \delta). \quad (2.32)$$

⁴Unter Zuhilfenahme der Anschauung erkennt man hieran bereits, dass f streng konvex sein muss.

In einem Wendepunkt stößt also ein Bereich, in dem die Funktion streng konkav ist, mit einem Bereich, in dem die Funktion streng konvex ist, zusammen. Anders gesagt: An einem Wendepunkt geht die Funktion von streng konkaver zu streng konvexer Krümmung über (oder umgekehrt). Oder besonders anschaulich ausgedrückt: An einem Wendepunkt geht der Graph von einer Rechtskurve in eine Linkskurve über (oder umgekehrt).



Bei einer **Kurvendiskussion** geht es darum, den Graphen einer Funktion f zu skizzieren, um so eine Vorstellung vom Verlauf von f zu erhalten. Stichwortartige Zusammenfassung der Vorgehensweise, wobei wir voraussetzen, dass f zweimal differenzierbar und f'' stetig ist⁵:

1. Bestimmung des Definitionsbereichs, falls dieser nicht vorgegeben ist.
2. Bestimmung der Nullstellen von f , f' und f'' . (Dies ist oft nur näherungsweise möglich.) Daraus ergeben sich (neben den Schnittpunkten von f mit der x -Achse) mögliche Extremstellen sowie mögliche Wendepunkte.
3. Gegebenenfalls Berechnung von $\lim_{x \rightarrow \infty} f(x)$ und $\lim_{x \rightarrow -\infty} f(x)$ sowie Berechnung von einseitigen Grenzwerten für Randpunkte des Definitionsbereichs, die selbst nicht zum Definitionsbereich gehören.
4. Mithilfe der in 2. bestimmten Nullstellen grenzt man die Bereiche ab, in denen $f(x) > 0$ bzw. $f(x) < 0$ gilt. Ebenso für f' und f'' . Dadurch hat man die Bereiche ermittelt, in denen f positiv bzw. negativ, streng monoton wachsend bzw. streng monoton fallend sowie streng konvex bzw. streng konkav ist (Sätze 15 und 21).
5. Bestimmung der Maxima, Minima und Wendepunkte. (Dies ergibt sich in der Regel direkt aus 4.)
6. Berechnung von Asymptoten⁶ (falls vorhanden) und ggf. Schnittpunkten von f mit den Asymptoten.
7. Berechnung der Funktionswerte $f(x)$ für $x = 0$, für die Extrema und für die Wendepunkte (sowie möglicherweise für weitere x); Anfertigen einer Skizze.

Definition.

- (a) f sei eine Funktion, die (zumindest) auf einem Intervall $[c, \infty)$ definiert sei. Eine Gerade $g(x) = ax + b$ heißt *Asymptote von f für $x \rightarrow \infty$* , falls gilt:

$$\lim_{x \rightarrow \infty} (f(x) - g(x)) = 0. \quad (2.33)$$

- (b) Analog definiert man eine *Asymptote von f für $x \rightarrow -\infty$* .

- (c) Die Gerade $x = x_0$ heißt *Asymptote von f in x_0* , falls mindestens eine der folgenden vier Bedingungen erfüllt ist:

$$\begin{array}{ll} \text{(i)} & \lim_{x \rightarrow x_0+} f(x) = +\infty \\ \text{(ii)} & \lim_{x \rightarrow x_0+} f(x) = -\infty \\ \text{(iii)} & \lim_{x \rightarrow x_0-} f(x) = +\infty \\ \text{(iv)} & \lim_{x \rightarrow x_0-} f(x) = -\infty \end{array}$$

⁵Ist diese Voraussetzung an einigen Stellen nicht erfüllt, so ändert sich die Vorgehensweise nur wenig. Man hat aber beispielsweise zu berücksichtigen, dass f auch Sprungstellen und „Ecken“ haben kann.

⁶Siehe nächste Definition für den Begriff der Asymptote.

Der folgende Satz dient zur Berechnung von Asymptoten für $x \rightarrow \infty$ und $x \rightarrow -\infty$.

Satz 22.

f besitzt für $x \rightarrow \infty$ genau dann die Asymptote $g(x) = ax + b$, wenn gilt:

$$\lim_{x \rightarrow \infty} \left(\frac{f(x)}{x} \right) = a \quad \text{und} \quad \lim_{x \rightarrow \infty} (f(x) - ax) = b. \quad (2.34)$$

Beweis.

(i) Besitzt f die Asymptote $g(x) = ax + b$, so gilt nach (2.33):

$$\lim_{x \rightarrow \infty} (f(x) - ax - b) = 0.$$

Hieraus folgt klarerweise der zweite Teil von (2.34), aber auch der erste:

$$\begin{aligned} \lim_{x \rightarrow \infty} (f(x) - ax - b) = 0 &\Rightarrow \lim_{x \rightarrow \infty} \left(\frac{f(x) - ax - b}{x} \right) = 0 \\ &\Rightarrow \lim_{x \rightarrow \infty} \left(\frac{f(x)}{x} \right) = a \quad \left(\text{da } \lim_{x \rightarrow \infty} \left(\frac{b}{x} \right) = 0 \right) \end{aligned}$$

(ii) Gilt umgekehrt (2.34), so folgt aus $\lim_{x \rightarrow \infty} (f(x) - ax) = b$, dass f die Asymptote $g(x) = ax + b$ besitzt. \square

Bemerkungen.

- Die Bedingung (2.34) kann zur Berechnung der Koeffizienten a und b einer Asymptoten verwendet werden.
- Ein entsprechender Satz gilt auch für $x \rightarrow -\infty$.

2.4.1 Beispiele zur Kurvendiskussion

Es wird nach dem obigen Schema vorgegangen.

Beispiel 1. Es sei $f(x) = x^3 - 6x^2 + 9x$.

1. f ist auf ganz \mathbb{R} definiert.
2. $f'(x) = 3x^2 - 12x + 9$ und $f''(x) = 6x - 12$
Nullstellen von f : 0 und 3.
Nullstellen von f' : 1 und 3.
Nullstellen von f'' : 2.
3. Es ist $f(x) = x^3 - 6x^2 + 9x = x^3 \left(1 - \frac{6}{x} + \frac{9}{x^2} \right)$.

Hieran erkennt man, dass gilt:

- $\lim_{x \rightarrow \infty} f(x) = +\infty$;
 - $\lim_{x \rightarrow -\infty} f(x) = -\infty$.
4. Aus $f(1) = 4 > 0$ und 3. folgt:
 - $f(x) < 0$ für alle $x \in (-\infty, 0)$;
 - $f(x) > 0$ für alle $x \in (0, 3)$;
 - $f(x) > 0$ für alle $x \in (3, \infty)$.

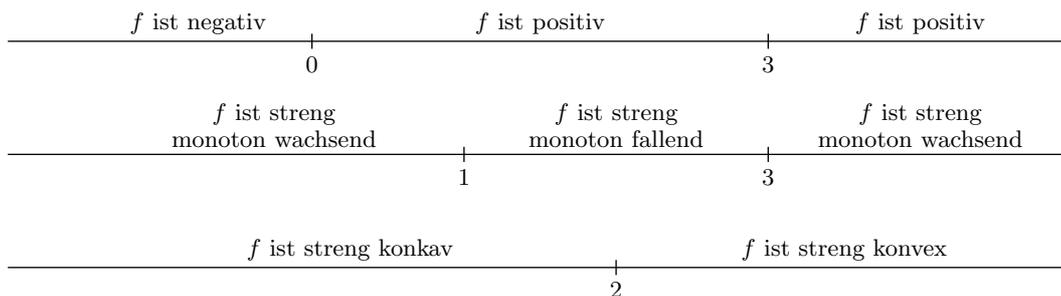
Auf ähnliche Art (etwa durch Einsetzen von 0, 2 und 4 in f') erkennt man, dass gilt:

- $f'(x) > 0$ für alle $x \in (-\infty, 1)$;
- $f'(x) < 0$ für alle $x \in (1, 3)$;
- $f'(x) > 0$ für alle $x \in (3, \infty)$.

Außerdem erhält man (etwa durch Einsetzen von 0 und 3 in f''), dass gilt:

- $f''(x) < 0$ für alle $x \in (-\infty, 2)$;
- $f''(x) > 0$ für alle $x \in (2, \infty)$.

Das Ergebnis lässt sich wie folgt darstellen („Abgrenzung von Bereichen, in denen f positiv bzw. negativ, streng monoton wachsend bzw. streng monoton fallend sowie streng konvex bzw. streng konkav ist“)



(Eine zusätzliche Erläuterung zu 4. wird in Abschnitt 2.4.2 gegeben. (Stichwort: *Zwischenwertsatz*))

5. Aus 4. ergibt sich Folgendes:

- f hat bei $x = 1$ ein Maximum.
- f hat bei $x = 3$ ein Minimum.
- f hat bei $x = 2$ einen Wendepunkt.

6. Es sind keine Asymptoten vorhanden.

Begründung: Es gilt $\frac{f(x)}{x} = x^2 - 6x + 9$ und somit:

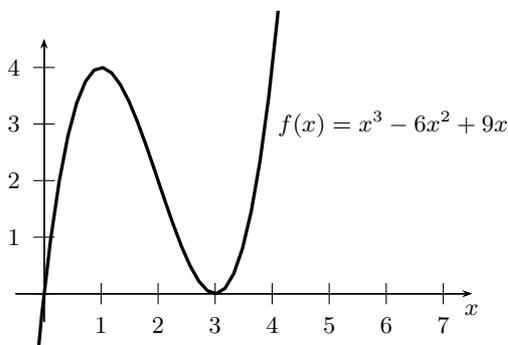
$$\lim_{x \rightarrow \infty} \frac{f(x)}{x} = \lim_{x \rightarrow -\infty} \frac{f(x)}{x} = \infty.$$

Also (vgl. Satz 22 sowie die anschließende Bemerkung) liegt weder eine Asymptote für $x \rightarrow \infty$ noch für $x \rightarrow -\infty$ vor. Außerdem gilt für alle $x_0 \in \mathbb{R}$:

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

Es liegt also auch keine Asymptote von f in x_0 vor.

7. *Berechnung einiger Funktionswerte und Skizze:*



x	-2	-1	0	1	2	3	4
$f(x)$	-50	-16	0	4	2	0	4

Beispiel 2. Es sei $f(x) = \frac{x^2 + 3}{x - 1}$.

- f ist auf $D(f) = \mathbb{R} \setminus \{1\}$ definiert.
- $f'(x) = \frac{x^2 - 2x - 3}{(x-1)^2}$ und $f''(x) = \frac{8}{(x+1)^3}$

Nullstellen von f : keine.

Nullstellen von f' : -1 und 3.

Nullstellen von f'' : keine.

- Für $x \rightarrow \infty$ ergibt sich:

$$\begin{aligned}
 \lim_{x \rightarrow \infty} (f(x)) &= \lim_{x \rightarrow \infty} \left(\frac{x^2 + 3}{x - 1} \right) \\
 &= \lim_{x \rightarrow \infty} \left(\frac{x^2 \left(1 + \frac{3}{x^2}\right)}{x \left(1 - \frac{1}{x}\right)} \right) \\
 &= \lim_{x \rightarrow \infty} \left(x \cdot \frac{1 + \frac{3}{x^2}}{1 - \frac{1}{x}} \right) \\
 &= \lim_{x \rightarrow \infty} (x) \cdot \lim_{x \rightarrow \infty} \left(\frac{1 + \frac{3}{x^2}}{1 - \frac{1}{x}} \right) \\
 &= \infty
 \end{aligned}$$

Für $x \rightarrow -\infty$ ergibt sich:

$$\begin{aligned}
 \lim_{x \rightarrow -\infty} (f(x)) &= \lim_{x \rightarrow -\infty} \left(\frac{x^2 + 3}{x - 1} \right) \\
 &= \lim_{x \rightarrow \infty} \left(\frac{(-x)^2 + 3}{-x - 1} \right) \\
 &= \lim_{x \rightarrow \infty} \left(\frac{x^2 + 3}{-x - 1} \right) \\
 &= \lim_{x \rightarrow \infty} \left(\frac{x^2 \left(1 + \frac{3}{x^2}\right)}{x \left(-1 - \frac{1}{x}\right)} \right) \\
 &= \lim_{x \rightarrow \infty} \left(x \cdot \frac{1 + \frac{3}{x^2}}{-1 - \frac{1}{x}} \right) \\
 &= \lim_{x \rightarrow \infty} (x) \cdot \lim_{x \rightarrow \infty} \left(\frac{1 + \frac{3}{x^2}}{-1 - \frac{1}{x}} \right) \\
 &= -\infty
 \end{aligned}$$

Des Weiteren gilt:

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

(Begründung: Für $x \rightarrow 1+$ strebt $x^2 + 3$ gegen eine positive Konstante (nämlich 4) und $\frac{1}{x-1}$ strebt gegen $+\infty$).

Mit ähnlicher Begründung erhält man:

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

4. Aus 3. folgt, dass gilt:

- $f(x) < 0$ für alle $x \in (-\infty, 1)$;
- $f(x) > 0$ für alle $x \in (1, \infty)$.

Durch Einsetzen von beispielsweise -2 , 0 , 2 und 4 in f' erhält man, dass gilt:

- $f'(x) > 0$ für alle $x \in (-\infty, -1)$;
- $f'(x) < 0$ für alle $x \in (-1, 1)$;
- $f'(x) < 0$ für alle $x \in (1, 3)$;
- $f'(x) > 0$ für alle $x \in (3, \infty)$.

Außerdem erhält man durch geeignetes Einsetzen in f'' , dass gilt:

- $f''(x) < 0$ für alle $x \in (-\infty, 1)$;
- $f''(x) > 0$ für alle $x \in (1, \infty)$.

Man erhält („Abgrenzung von Bereichen“):

f ist negativ			f ist positiv	
		1		
f ist streng monoton wachsend	f ist streng monoton fallend		f ist streng monoton fallend	f ist streng monoton wachsend
	-1	1	3	
f ist streng konkav			f ist streng konvex	
		1		

Wie im ersten Beispiel haben wir auch dieses Mal in Punkt 4 der Kurvendiskussion vom *Zwischenwertsatz* Gebrauch gemacht (vgl. Abschnitt 2.4.2).

5. Aus 4. ergibt sich Folgendes:

- f hat ein Maximum bei $x = -1$;
- f hat ein Minimum bei $x = 3$;
- f hat keine Wendepunkte.

6. Mithilfe von Satz 22 zeigen wir, dass f für $x \rightarrow \infty$ eine Asymptote $g(x) = ax + b$ besitzt und berechnen diese:

$$a = \lim_{x \rightarrow \infty} \left(\frac{f(x)}{x} \right) = \lim_{x \rightarrow \infty} \left(\frac{x^2 + 3}{x^2 - x} \right) = 1$$

$$b = \lim_{x \rightarrow \infty} (f(x) - ax) = \lim_{x \rightarrow \infty} \left(\frac{x^2 + 3}{x - 1} - x \right) = \lim_{x \rightarrow \infty} \left(\frac{x + 3}{x - 1} \right) = 1$$

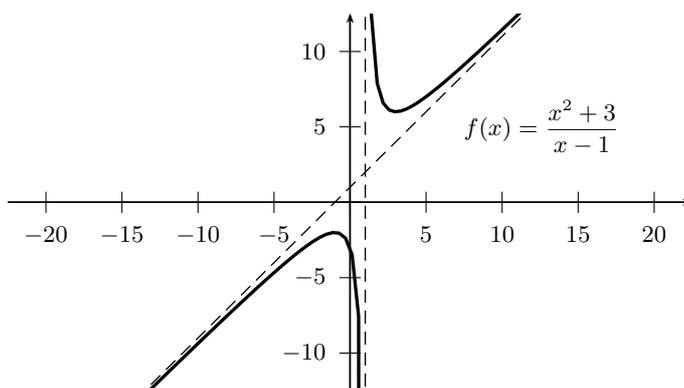
Also besitzt f für $x \rightarrow \infty$ die Asymptote $g(x) = x + 1$.

Eine ähnliche Überlegung ergibt, dass f für $x \rightarrow -\infty$ ebenfalls die Asymptote $g(x) = x + 1$ besitzt. Schnittpunkte von $f(x)$ mit der Asymptote gibt es keine, da

$$\frac{x^2 + 3}{x - 1} = x + 1 \iff x^2 + 3 = x^2 - 1$$

gilt und da diese Gleichung keine Lösung besitzt. Außerdem (vgl. 3.) ist die Gerade $x = 1$ eine Asymptote von f .

7. Berechnung einiger Funktionswerte und Skizze:



x	-2	-1	0	2	3
$f(x)$	$-\frac{7}{3}$	-2	-3	7	6

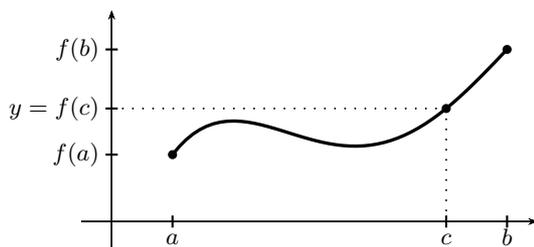
2.4.2 Der Zwischenwertsatz

In diesem Satz wird eine besonders nützliche Eigenschaft stetiger Funktionen festgehalten, die bereits in unseren Beispielen zur Kurvendiskussion (vgl. Abschnitt 2.4.1) eine Rolle gespielt hat.

Satz 23 (Zwischenwertsatz).

Für eine reelle Funktion f sei $[a, b]$ ein Intervall, auf dem f definiert und stetig ist. Es gelte $f(a) \neq f(b)$ und y sei ein beliebiger Wert zwischen $f(a)$ und $f(b)$, d.h. $f(a) < y < f(b)$ oder $f(b) < y < f(a)$. Dann gibt es ein c mit $a < c < b$, für das $f(c) = y$ gilt.

Illustration des Zwischenwertsatzes für den Fall $f(a) < f(b)$:



Da der Zwischenwertsatz anschaulich einleuchtend ist, verzichten wir auf einen Beweis.

Ein für die Anwendungen besonders wichtiger Spezialfall des Zwischenwertsatzes liegt für $y = 0$ vor. In diesem Fall besagt der Zwischenwertsatz Folgendes.

Satz 23' (Zwischenwertsatz für den Spezialfall $y = 0$).

Die reelle Funktion f sei auf $[a, b]$ definiert und stetig. Es gelte $f(a) < 0 < f(b)$ oder $f(b) < 0 < f(a)$. Dann gibt es ein c mit $a < c < b$, für das $f(c) = 0$ gilt, d.h., f besitzt eine Nullstelle in (a, b) .

Anwendungen hiervon kamen bereits in den Beispielen zur Kurvendiskussion (vgl. Abschnitt 2.4.1) vor. Wir erläutern dies anhand des ersten Beispiels: Dort galt $f(x) = x^3 - 6x^2 + 9x$. Unter Punkt 2 hatten wir festgestellt, dass 0 und 3 die einzigen Nullstellen von f sind. Unter Punkt 4 hatten wir ein $x \in (0, 3)$ gewählt (nämlich $x = 1$) und haben durch Einsetzen von $x = 1$ getestet, ob $f(x) > 0$ oder $f(x) < 0$ für alle $x \in (0, 3)$ gilt. Ergebnis: Wir hatten festgestellt, dass $f(1) > 0$ gilt und daraus gefolgert, dass $f(x) > 0$ für alle $x \in (0, 3)$ gelten muss. An dieser Stelle wurde der Zwischenwertsatz genutzt: Für kein $x \in (0, 3)$ kann $f(x) < 0$ gelten, da es andernfalls aufgrund des Zwischenwertsatzes ein zwischen x und 1 liegendes c geben müsste mit $f(c) = 0$, im Widerspruch zur Tatsache, dass im Intervall $(0, 3)$ keine Nullstellen von f vorkommen.

Um Fehlschlüsse zu vermeiden, ist es wichtig, vor Anwendung des Zwischenwertsatzes genau zu prüfen, ob die Voraussetzungen zur Anwendung erfüllt sind: Die betrachtete Funktion darf im Intervall $[a, b]$ „keine Sprünge machen“, d.h., sie darf dort weder Definitionslücken noch Unstetigkeitsstellen haben.

2.5 Trigonometrische Funktionen

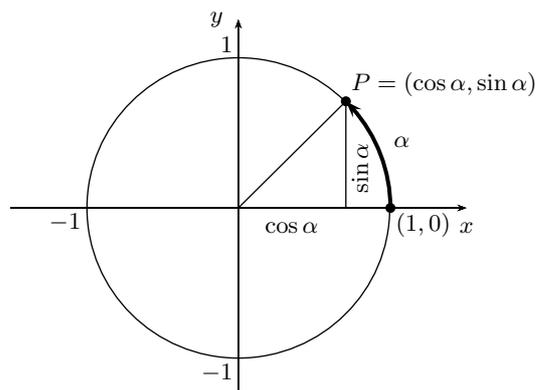
2.5.1 Definition der Trigonometrischen Funktionen

Wir betrachten die Menge der Punkte $(x, y) \in \mathbb{R}^2$, für die $x^2 + y^2 = 1$ gilt. Diese bilden einen Kreis mit Radius 1 und Mittelpunkt $(0, 0)$, den *Einheitskreis*. Die Zahl π definiert man als den halben Umfang dieses Kreises; der Umfang des Einheitskreises beträgt also 2π . Man kann zeigen, dass π eine irrationale Zahl ist, d.h. $\pi \notin \mathbb{Q}$. Die ersten Dezimalstellen von π :

$$\pi = 3,1415926535 \dots$$

Es sei P ein beliebiger Punkt auf dem Einheitskreis. Durchlaufen wir ausgehend vom Punkt $(1, 0)$ den Einheitskreis entgegen dem Uhrzeigersinn bis zum Punkt P , so sei mit α die Länge des durchlaufenen Bogens bezeichnet (siehe Zeichnung unten). Sind x und y die Koordinaten von P (d.h. $P = (x, y)$), so bezeichnet man x als $\cos \alpha$ und y als $\sin \alpha$, d.h., man definiert:

$$\cos \alpha = x \quad \text{und} \quad \sin \alpha = y \quad (\text{für } \alpha \in [0, 2\pi)).$$



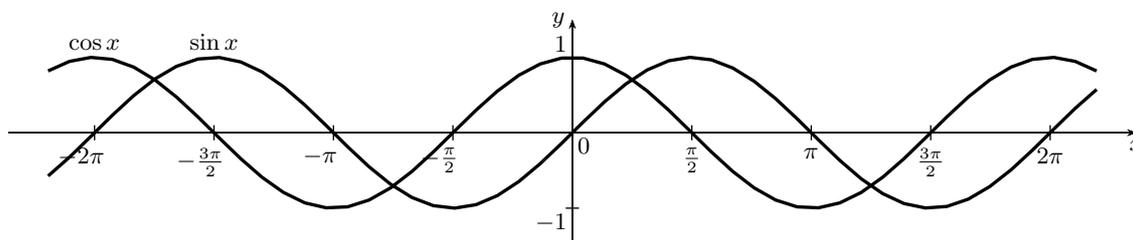
Hierdurch haben wir also zwei Funktionen $\cos : [0, 2\pi) \rightarrow \mathbb{R}$ und $\sin : [0, 2\pi) \rightarrow \mathbb{R}$ definiert. Diese beiden Funktionen erweitert man zu Funktionen, die auf ganz \mathbb{R} definiert sind, indem man sie periodisch fortsetzt. Man definiert für alle $n \in \mathbb{Z}$ und $\alpha \in [0, 2\pi)$:

$$\cos(2\pi n + \alpha) := \cos \alpha, \quad \sin(2\pi n + \alpha) := \sin \alpha.$$

Aus der Definition von sin und cos folgt:

$$\begin{array}{cccccc} \sin 0 = 0 & \sin \frac{\pi}{2} = 1 & \sin \pi = 0 & \sin \frac{3\pi}{2} = -1 & \sin 2\pi = 0 & \\ \cos 0 = 1 & \cos \frac{\pi}{2} = 0 & \cos \pi = -1 & \cos \frac{3\pi}{2} = 0 & \cos 2\pi = 1 & \end{array}$$

Skizze der Graphen von sin und cos:



Im Folgenden nennen wir die unabhängige Variable meist nicht mehr α , sondern (wie gewohnt) x ; wir schreiben also $\sin x$ und $\cos x$ anstelle von $\sin \alpha$ bzw. $\cos \alpha$. Aus der Definition der Funktionen sin und cos ergibt sich⁷ (Man mache sich dies am Einheitskreis klar!):

$$\sin^2 x + \cos^2 x = 1 \quad (2.35)$$

$$\sin(-x) = -\sin x \quad (2.36)$$

$$\cos(-x) = \cos x \quad (2.37)$$

Funktionen mit der Eigenschaft $f(-x) = -f(x)$ nennt man *ungerade Funktionen*; Funktionen mit der Eigenschaft $f(-x) = f(x)$ nennt man *gerade Funktionen*. In (2.36) und (2.37) wird also festgestellt, dass sin eine ungerade und cos eine gerade Funktion ist.

Anschaulich einleuchtend ist die folgende Feststellung, auf deren Beweis wir verzichten wollen.

Feststellung.

$$\text{Die Funktionen sin und cos sind auf ganz } \mathbb{R} \text{ stetig.} \quad (2.38)$$

Ebenfalls ohne Beweis geben wir die folgenden, oft benutzten Eigenschaften von sin und cos an, welche *Additionstheoreme* genannt werden:

$$\begin{array}{l} \sin(x_1 + x_2) = \sin x_1 \cos x_2 + \cos x_1 \sin x_2 \\ \cos(x_1 + x_2) = \cos x_1 \cos x_2 - \sin x_1 \sin x_2 \end{array} \quad (2.39)$$

Eine Folgerung aus den Additionstheoremen:

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

Ähnlich folgt:

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

Als Folgerung aus (2.39) erhält man die Formeln für den doppelten Winkel:

$$\sin(2x) = 2 \sin x \cos x \quad (2.40)$$

⁷Anstelle von $(\sin x)^2$ und $(\cos x)^2$ schreibt man häufig auch $\sin^2 x$ bzw. $\cos^2 x$.

$$\cos(2x) = \cos^2 x - \sin^2 x. \quad (2.41)$$

Beweis. Man setze $x_1 = x_2 = x$ in (2.39) ein. \square

Als *trigonometrische Funktionen* bezeichnet man die Funktionen \sin , \cos , \tan und \cot .

Nachdem wir oben \sin und \cos am Einheitskreis definiert haben, kommen wir nun zur Definition von \tan und \cot .

Definition.

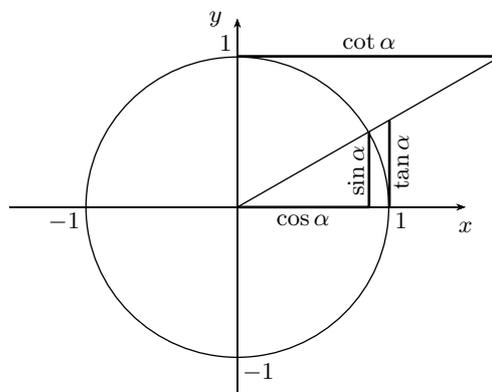
(1) Für alle x mit $\cos x \neq 0$ definieren wir:

$$\tan x := \frac{\sin x}{\cos x}.$$

(2) Für alle x mit $\sin x \neq 0$ definieren wir:

$$\cot x := \frac{\cos x}{\sin x}.$$

Die Funktionen \tan und \cot können ebenfalls am Einheitskreis veranschaulicht werden:

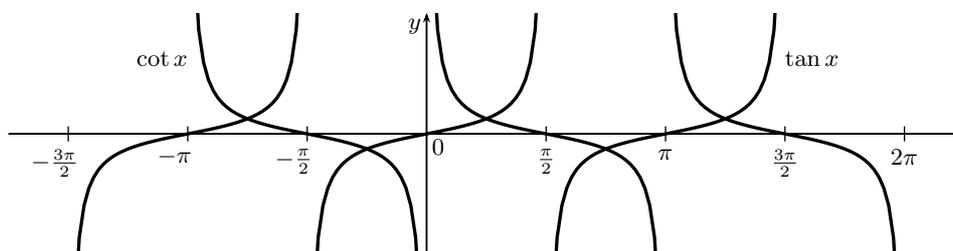


Die Funktion $\tan x$ ist also für alle $x \neq \pm \frac{\pi}{2}, \pm \frac{3\pi}{2}, \pm \frac{5\pi}{2}, \dots$ definiert; $\cot x$ ist für alle $x \neq 0, \pm\pi, \pm 2\pi, \dots$ definiert. Beide Funktionen sind periodisch mit der Periode π , d.h., es gilt:

$$\tan(x + \pi) = \tan x \quad \text{bzw.} \quad \cot(x + \pi) = \cot x. \quad (2.42)$$

Beweis. $\tan(x + \pi) = \frac{\sin(x + \pi)}{\cos(x + \pi)} = \frac{-\sin x}{-\cos x} = \tan x$. Der Beweis für $\cot x$ geht analog. \square

Die Schaubilder für \tan und \cot :



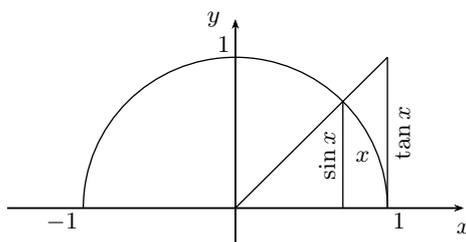
2.5.2 Die Ableitungen der Trigonometrischen Funktionen

Grundlegend für die Bestimmung der Ableitung von $\sin x$ ist die folgende Grenzwertbeziehung:

$$\lim_{x \rightarrow 0} \left(\frac{\sin x}{x} \right) = 1. \quad (2.43)$$

Beweisskizze. Für $0 < x < \frac{\pi}{2}$ gilt:

$$\sin x < x < \tan x = \frac{\sin x}{\cos x}.$$



Dividiert man durch $\sin x$ und geht anschließend zu den Kehrwerten über, so ergibt sich:

$$\begin{aligned} 1 &> \frac{\sin x}{x} > \cos x \\ \implies 1 &\geq \lim_{x \rightarrow 0} \left(\frac{\sin x}{x} \right) \geq \lim_{x \rightarrow 0} (\cos x) = \cos 0 = 1 \end{aligned}$$

(Wir haben benutzt, dass \cos eine stetige Funktion ist.) Also gilt⁸ (2.43). \square

Satz 1.

Die Funktionen $\sin x$, $\cos x$, $\tan x$ und $\cot x$ sind an jeder Stelle ihres Definitionsbereichs differenzierbar; es gilt:

$$\begin{aligned} (\sin x)' &= \cos x \\ (\cos x)' &= -\sin x \\ (\tan x)' &= \frac{1}{\cos^2 x} = 1 + \tan^2 x \\ (\cot x)' &= -\frac{1}{\sin^2 x} = -1 - \cot^2 x \end{aligned}$$

Beweis. Wir zeigen $(\sin x)' = \cos x$. Hierzu sei (x_n) eine Folge mit $x_n \neq x$ und $x_n \rightarrow x$. Wir haben nachzuweisen, dass die Folge der Differenzenquotienten $\frac{\sin x_n - \sin x}{x_n - x}$ gegen $\cos x$ konvergiert. Dies folgt so:

$$\begin{aligned} \frac{\sin x_n - \sin x}{x_n - x} &= \frac{\sin(x_n - x + x) - \sin x}{x_n - x} \\ (2.39) \quad &= \frac{\sin(x_n - x) \cdot \cos x + \cos(x_n - x) \cdot \sin x - \sin x}{x_n - x} \\ &= \frac{\sin(x_n - x)}{x_n - x} \cdot \cos x + \frac{\cos(x_n - x) - 1}{x_n - x} \cdot \sin x \\ &= \frac{\sin(x_n - x)}{x_n - x} \cdot \cos x + \frac{(\cos(x_n - x) - 1) \cdot (\cos(x_n - x) + 1)}{(x_n - x) \cdot (\cos(x_n - x) + 1)} \cdot \sin x \\ &= \frac{\sin(x_n - x)}{x_n - x} \cdot \cos x + \frac{\cos^2(x_n - x) - 1}{(x_n - x) \cdot (\cos(x_n - x) + 1)} \cdot \sin x \end{aligned}$$

⁸Wegen $\frac{\sin(-x)}{-x} = \frac{-\sin x}{-x} = \frac{\sin x}{x}$ muss man für den Beweis von (2.43) den Fall $x < 0$ nicht beachten. Auf den Beweis von $\sin x < x < \tan x$ wollen wir hier nicht weiter eingehen. Man mache sich jedoch am Einheitskreis klar, dass die Aussage $\sin x < x < \tan x$ anschaulich einleuchtend ist (vgl. Figur). Wer eine einfache, elementargeometrische Begründung der Ungleichung $\sin x < x < \tan x$ kennenlernen möchte, findet diese beispielsweise in W. Schäfer, K. Georgi, G. Trippler, Mathematik-Vorkurs, Teubner.

$$\begin{aligned}
&\stackrel{(2.35)}{=} \frac{\sin(x_n - x)}{x_n - x} \cdot \cos x - \frac{\sin^2(x_n - x)}{(x_n - x) \cdot (\cos(x_n - x) + 1)} \cdot \sin x \\
&= \frac{\sin(x_n - x)}{x_n - x} \cdot \left(\cos x - \frac{\sin(x_n - x) \cdot \sin x}{\cos(x_n - x) + 1} \right) \rightarrow \cos x \quad (\text{für } n \rightarrow \infty).
\end{aligned}$$

Der letzte Schritt gilt wegen $\frac{\sin(x_n - x)}{x_n - x} \rightarrow 1$ (wegen (2.43)) und $\frac{\sin(x_n - x) \sin x}{\cos(x_n - x) + 1} \rightarrow 0$ (wegen der Stetigkeit von $\sin x$ und $\cos x$). Damit ist $(\sin x)' = \cos x$ gezeigt.

Die übrigen Formeln folgen nun leicht:

$$\begin{aligned}
(\cos x)' &= \left(\sin \left(x + \frac{\pi}{2} \right) \right)' = \cos \left(x + \frac{\pi}{2} \right) = -\sin x \\
(\tan x)' &= \left(\frac{\sin x}{\cos x} \right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} \stackrel{(2.35)}{=} \frac{1}{\cos^2 x} = 1 + \tan^2 x \\
(\cot x)' &= \left(\frac{\cos x}{\sin x} \right)' = \frac{-\sin^2 x - \cos^2 x}{\sin^2 x} \stackrel{(2.35)}{=} -\frac{1}{\sin^2 x} = -1 - \cot^2 x \quad \square
\end{aligned}$$

2.5.3 Die Umkehrfunktionen der Trigonometrischen Funktionen

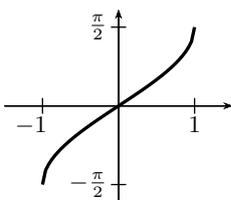
Die trigonometrischen Funktionen lassen sich auf bestimmten Intervallen umkehren. (Eine Umkehrung von \sin , \cos , \tan und \cot auf ihrem gesamten Definitionsbereich ist natürlich nicht möglich, da diese Funktionen nicht injektiv sind.)

- 1) Die Funktion $f(x) = \sin x$ ist im Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton wachsend. Es existiert also für dieses Intervall die Umkehrfunktion f^{-1} . Die Funktion f^{-1} nennt man *Arcussinus* (Bezeichnung: \arcsin). Da die Funktion $f(x) = \sin x$ das Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ auf das Intervall $[-1, 1]$ abbildet, gilt:

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

Da $\sin x$ auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton wachsend und stetig ist, ist auch $\arcsin x$ streng monoton wachsend und stetig (siehe Satz 6, Abschnitt 2.2).

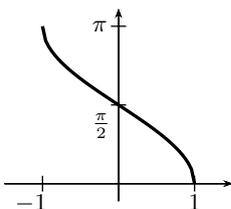
Skizze des Graphen von $\arcsin x$:



- 2) Die Funktion $f(x) = \cos x$ ist im Intervall $[0, \pi]$ stetig und streng monoton fallend. Analog zu $\arcsin x$ definiert man:

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

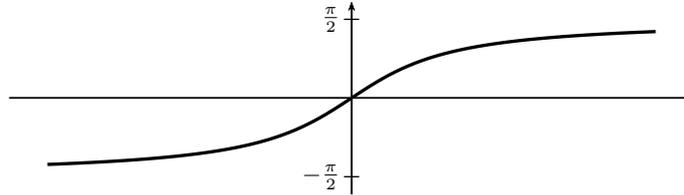
Skizze des Graphen von $\arccos x$:



- 3) Die Funktion $f(x) = \tan x$ ist im Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$ stetig und streng monoton wachsend. Analog zu $\arcsin x$ und $\arccos x$ definiert man:

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

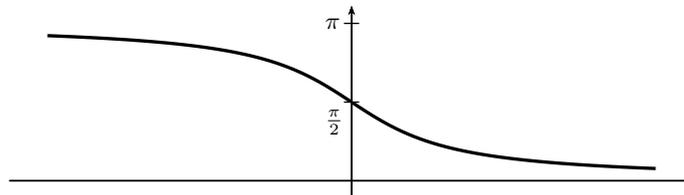
Skizze des Graphen von $\arctan x$:



- 4) Die Funktion $f(x) = \cot x$ ist im Intervall $(0, \pi)$ stetig und streng monoton fallend. Analog zu $\arctan x$ definiert man:

$$\operatorname{arccot} x : \mathbb{R} \rightarrow (0, \pi).$$

Skizze des Graphen von $\operatorname{arccot} x$:



Satz 2.

Die Funktionen \arcsin und \arccos sind auf $(-1, 1)$ differenzierbar, die Funktionen \arctan und arccot sind auf ganz \mathbb{R} differenzierbar. Es gilt:

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$$

$$(\arccos x)' = \frac{-1}{\sqrt{1-x^2}}$$

$$(\arctan x)' = \frac{1}{1+x^2}$$

$$(\operatorname{arccot} x)' = \frac{-1}{1+x^2}$$

Beweis. Wir benutzen die Umkehrregel (Satz 6(d), Abschnitt 2.2):

1. Für $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ gilt: Die Ableitung von $\sin x$ an der Stelle x ist ungleich Null; es gilt genauer $(\sin x)' = \cos x \neq 0$ für $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Also folgt für die Ableitung der Umkehrfunktion \arcsin an der Stelle $y = \sin x$ ($x \in (-\frac{\pi}{2}, \frac{\pi}{2})$):

$$(\arcsin y)' = \frac{1}{(\sin x)'} = \frac{1}{\cos x} \stackrel{(2.35)}{=} \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}}.$$

Dies beweist die Behauptung⁹. Analog zeigt man die Formel für $(\arccos x)'$.

⁹Wir haben $(\arcsin y)' = \frac{1}{\sqrt{1-y^2}}$ gezeigt. Da wir in dieser Formel die unabhängige Variable auch mit x (oder α , ξ , μ , ζ , z , usw.) bezeichnen können, gilt also - wie in Satz 2 behauptet - $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$.

2. Für $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ gilt: $(\tan x)' = 1 + \tan^2 x \neq 0$. Also gilt nach der Umkehrregel für $y = \tan x$:

$$(\arctan y)' = \frac{1}{(\tan x)'} = \frac{1}{1 + \tan^2 x} = \frac{1}{1 + y^2}.$$

Analog zeigt man die Formel für $(\operatorname{arccot} x)'$. \square

Man nennt die Funktionen \arcsin , \arccos , \arctan und arccot *zyklometrische Funktionen* oder auch *Arkusfunktionen*.

2.6 Das Newtonsche Verfahren

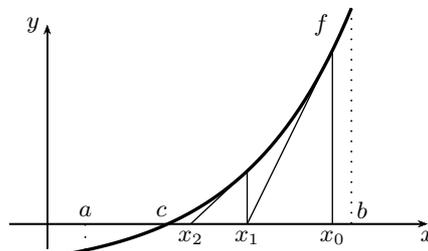
Dieses Verfahren dient zur näherungsweise Berechnung von Nullstellen einer Funktion f . Es zeichnet sich dadurch aus, dass es sehr gut konvergiert und daher meist schon nach wenigen Iterationsschritten gute Näherungswerte liefert. Wir formulieren zunächst Voraussetzungen für das Newtonsche Verfahren, danach beschreiben wir die geometrisch sehr anschauliche Vorgehensweise.

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion, für die Folgendes gilt:

- (1) $f'(x) \neq 0$ für alle $x \in [a, b]$.
- (2) f'' ist auf ganz $[a, b]$ vorhanden und stetig.
- (3) $f(a) \cdot f(b) < 0$.

Nach (1) gilt also $f'(x) > 0$ für alle $x \in [a, b]$ oder $f'(x) < 0$ für alle $x \in [a, b]$, d.h., (1) sorgt dafür, dass f *streng monoton* (steigend oder fallend) ist. Bedingung (3) sichert, dass (nach dem Zwischenwertsatz) überhaupt eine Nullstelle $c \in [a, b]$ vorhanden ist.

Dem Newtonschen Verfahren liegt eine einfache geometrische Idee zugrunde. Man geht von einer ersten Näherung x_0 für die Nullstelle c aus, legt im Punkt $(x_0, f(x_0))$ die Tangente an den Graphen von f und bestimmt den Schnittpunkt x_1 der Tangente mit der x -Achse (siehe Zeichnung).



Indem man $y = 0$ setzt, erhält man aus der Tangentengleichung $y = f(x_0) + f'(x_0)(x - x_0)$:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Man nimmt nun x_1 als verbesserte Näherung für die Nullstelle c . Aus x_1 berechnet man im nächsten Schritt:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Allgemein: Man erhält eine Folge von fortlaufend verbesserten Näherungswerten x_0, x_1, x_2, \dots für c durch:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Satz 28.

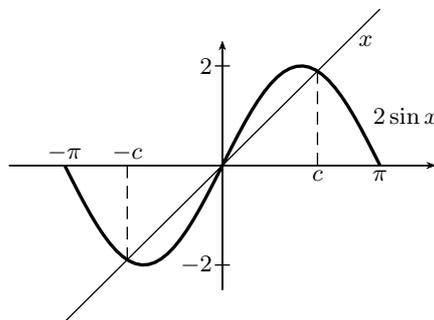
Die Funktion $f : [a, b] \rightarrow \mathbb{R}$ erfülle die Bedingungen (1) - (3). Dann existiert ein $\delta > 0$, für das gilt: Ist der Startwert x_0 so gewählt, dass $|x_0 - c| < \delta$ für die (eindeutig bestimmte) Nullstelle c von f in $[a, b]$ gilt, so wird durch

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

eine Folge x_0, x_1, \dots definiert, die gegen c konvergiert.

Wir wollen auf die Wiedergabe eines Beweises verzichten. Die Rolle der Zahl δ lässt sich auch wie folgt beschreiben: Wählt man den Startwert x_0 hinreichend nahe bei der Nullstelle c (nämlich nicht weiter entfernt als δ), so konvergiert die Folge x_0, x_1, \dots gegen c . In der Praxis wird man einen geeigneten Startwert x_0 einem möglichst guten Schaubild entnehmen.

Beispiel zum Newtonverfahren. Es sollen die Nullstellen von $f(x) = 2 \sin x - x$ berechnet werden. Da $f(x) = 0$ gleichbedeutend mit $2 \sin x = x$ ist, erkennt man, dass es genau eine positive Nullstelle c gibt und dass $-c$ die einzige weitere Nullstelle außer 0 ist (siehe Skizze). Es genügt also, c zu berechnen.



Es gilt $f(\frac{\pi}{2}) = 2 - \frac{\pi}{2} > 0$ und $f(\pi) = 0 - \pi < 0$, d.h., wenn wir $a = \frac{\pi}{2}$ und $b = \pi$ wählen, so liegt c im Intervall $[a, b]$. Für $x \in [a, b] = [\frac{\pi}{2}, \pi]$ gilt:

$$\begin{aligned} f'(x) &= 2 \cos x - 1 < 0 \\ f''(x) &= -2 \sin x. \end{aligned}$$

Es sind also die Bedingungen (1), (2) und (3) erfüllt. Als Startwert x_0 wählen wir (näherungsweise) die Zahl π :

$$x_0 = 3.141592654,$$

Man erhält mit diesem Startwert:

$$\begin{aligned} x_1 &= 2.094395102 \\ x_2 &= 1.913222955 \\ x_3 &= 1.895671752 \\ x_4 &= 1.895494285 \\ x_5 &= 1.895494267 = x_6 \end{aligned}$$

2.7 Die Regeln von de l'Hospital

Die Bestimmung eines Grenzwertes

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

kann schwierig sein, wenn Zähler und Nenner beide gegen 0 bzw. beide gegen ∞ streben (für $x \rightarrow x_0$). Die folgenden Sätze liefern eine schlagkräftige Methode zur Bestimmung derartiger Grenzwerte. Man fasst

diese Sätze unter dem Namen *Regeln von de l'Hospital*¹⁰ zusammen. Wir verzichten in diesem Abschnitt auf Beweise.

2.7.1 Der Typ $\frac{0}{0}$

Satz 24.

Es sei I ein Intervall und $x_0 \in I$. Die Funktionen f und g seien für alle $x \in I$ mit $x \neq x_0$ differenzierbar. Es gelte $g(x) \neq 0$ und $g'(x) \neq 0$ für alle $x \in I$, $x \neq x_0$. Ferner sei $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$.

Dann gilt

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

falls der rechte Grenzwert existiert bzw. gleich $+\infty$ oder $-\infty$ ist.

Beispiele:

1. Es soll $\lim_{x \rightarrow 0} \left(\frac{e^x - 1}{3x} \right)$ berechnet werden.

Für $x \rightarrow 0$ streben sowohl Zähler als auch Nenner gegen 0, d.h., die Berechnung des Grenzwertes ist nicht ohne Weiteres möglich. Anwendung von Satz 24 ergibt jedoch:

$$\lim_{x \rightarrow 0} \left(\frac{e^x - 1}{3x} \right) = \lim_{x \rightarrow 0} \left(\frac{e^x}{3} \right) = \frac{1}{3}.$$

2. Es soll $\lim_{x \rightarrow 0} \left(\frac{\ln(x+1)}{x} \right)$ berechnet werden.

Auch hier streben Zähler und Nenner gegen 0. Satz 24 ergibt:

$$\lim_{x \rightarrow 0} \left(\frac{\ln(x+1)}{x} \right) = \lim_{x \rightarrow 0} \left(\frac{\frac{1}{x+1}}{1} \right) = 1.$$

3. Es soll $\lim_{x \rightarrow 1} \left(\frac{x^2 - 1}{x - 1} \right)$ berechnet werden.

Erste Methode:

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

Zweite Methode (nach Satz 24):

$$\lim_{x \rightarrow 1} \left(\frac{x^2 - 1}{x - 1} \right) = \lim_{x \rightarrow 1} \left(\frac{2x}{1} \right) = 2.$$

4. Es soll $\lim_{x \rightarrow 1} \left(\frac{\sqrt{x} - 1}{\sqrt{x} - 1} \right)$ berechnet werden.

$$\lim_{x \rightarrow 1} \left(\frac{\sqrt{x} - 1}{\sqrt{x} - 1} \right) = \lim_{x \rightarrow 1} \left(\frac{\frac{1}{2\sqrt{x}}}{\frac{1}{2\sqrt{x-1}}} \right) = \lim_{x \rightarrow 1} \left(\frac{\sqrt{x-1}}{\sqrt{x}} \right) = \frac{0}{1} = 0.$$

¹⁰Nach Marquis de l'Hospital (1661-1704) benannt.

5. Es soll $\lim_{x \rightarrow 0} \left(\frac{e^x + e^{-x} - 2}{2x^2} \right)$ berechnet werden.

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

Dieses Beispiel zeigt, dass unter Umständen erst mehrfach wiederholte Anwendung von Satz 24 zum Erfolg führt.

6. Es soll $\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}}$ berechnet werden.

$$\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = \lim_{x \rightarrow 0} \left(e^{\left(\frac{1}{x} \ln(1+x)\right)} \right) \stackrel{e^x \text{ stetig}}{=} e^{\left(\lim_{x \rightarrow 0} \left(\frac{1}{x} \ln(1+x)\right)\right)} \stackrel{2.}{=} e^1 = e.$$

Durch Beispiel 6 haben wir die Beziehung (2.23) in Abschnitt 2.3 im Nachhinein bestätigt.

Ein Satz 24 entsprechender Satz gilt auch für $x \rightarrow \infty$ anstelle von $x \rightarrow x_0$.

Satz 24'.

Die Funktionen f und g seien für alle $x \in I = (a, \infty)$ differenzierbar. Es gelte $g(x) \neq 0$ und $g'(x) \neq 0$ für alle $x \in I$ sowie $\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) = 0$.

Dann gilt

$$\lim_{x \rightarrow \infty} \left(\frac{f(x)}{g(x)} \right) = \lim_{x \rightarrow \infty} \left(\frac{f'(x)}{g'(x)} \right),$$

falls der rechte Grenzwert existiert bzw. gleich $+\infty$ oder $-\infty$ ist.

2.7.2 Der Typ $\frac{\infty}{\infty}$

Satz 25.

Es sei I ein Intervall und $x_0 \in I$. Die Funktionen f und g seien für alle $x \in I$, $x \neq x_0$, differenzierbar. Es gelte $g(x) \neq 0$ und $g'(x) \neq 0$ für alle $x \in I$, $x \neq x_0$. Ferner gelte $\lim_{x \rightarrow x_0} |f(x)| = \lim_{x \rightarrow x_0} |g(x)| = \infty$.

Dann gilt

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

falls der rechte Grenzwert existiert bzw. gleich $+\infty$ oder $-\infty$ ist.

Zusatz. Ein Satz 25 entsprechender Satz gilt auch für $x \rightarrow \infty$ anstelle von $x \rightarrow x_0$.

Beispiele:

1. Es soll $\lim_{x \rightarrow 0} (x \cdot \ln x)$ berechnet werden.

Um Satz 25 anwenden zu können, schreiben wir $x \cdot \ln x$ zunächst um. Wir erhalten:

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

2. Es soll $\lim_{x \rightarrow \infty} \left(\frac{2^x}{x} \right)$ berechnet werden.

$$\lim_{x \rightarrow \infty} \left(\frac{2^x}{x} \right) = \lim_{x \rightarrow \infty} \left(\frac{2^x \cdot \ln 2}{1} \right) = \infty$$

3. Es soll $\lim_{x \rightarrow \infty} \left(x^{\frac{1}{x}}\right)$ berechnet werden.

$$\lim_{x \rightarrow \infty} \left(x^{\frac{1}{x}}\right) = \lim_{x \rightarrow \infty} \left(e^{\left(\frac{1}{x} \cdot \ln x\right)}\right) = e^{\left(\lim_{x \rightarrow \infty} \left(\frac{1}{x} \cdot \ln x\right)\right)}$$

Wegen

$$\lim_{x \rightarrow \infty} \left(\frac{\ln x}{x}\right) = \lim_{x \rightarrow \infty} \left(\frac{\frac{1}{x}}{1}\right) = 0$$

folgt also:

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

Insbesondere folgt $\lim_{n \rightarrow \infty} \left(\sqrt[n]{n}\right) = 1$.

Satz 24 nebst Zusatz (Satz 24') nennt man die *Regel von de l'Hospital für den Fall $\frac{0}{0}$* ;

Satz 25 nebst Zusatz nennt man die *Regel von de l'Hospital für den Fall $\frac{\infty}{\infty}$* .

Satz 26.

Für jedes $r \in \mathbb{R}$ gilt:

$$\lim_{x \rightarrow \infty} \left(\frac{e^x}{x^r}\right) = \infty.$$

Für $r \leq 0$ ist die Aussage von Satz 26 klarerweise richtig, da $x^r \rightarrow 0$ für $r < 0$ und $x^r = 1$ für $r = 0$ gilt. Für $r > 0$ kann Satz 26 leicht durch mehrfache Anwendung der Regel von de l'Hospital beweisen.

Wir schauen uns exemplarisch den Fall $r = 4$ an: Weshalb gilt $\lim_{x \rightarrow \infty} \left(\frac{e^x}{x^4}\right) = \infty$?

Man erhält das Ergebnis durch 4-maliges Anwenden der Regel von de l'Hospital für $\frac{\infty}{\infty}$:

$$\lim_{x \rightarrow \infty} \left(\frac{e^x}{x^4}\right) = \lim_{x \rightarrow \infty} \left(\frac{e^x}{4x^3}\right) = \lim_{x \rightarrow \infty} \left(\frac{e^x}{12x^2}\right) = \lim_{x \rightarrow \infty} \left(\frac{e^x}{24x}\right) = \lim_{x \rightarrow \infty} \left(\frac{e^x}{24}\right) = \infty.$$

Definition.

Es gelte $\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) = \infty$. Man sagt, $f(x)$ wächst für $x \rightarrow \infty$ schneller als $g(x)$, falls Folgendes gilt:

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

Im Fall $r > 0$ besagt Satz 26 also, dass die Exponentialfunktion e^x für $x \rightarrow \infty$ schneller wächst als jede Potenzfunktion x^r ($r > 0$).

Satz 27.

Jede Potenzfunktion x^r ($r > 0$) wächst schneller als $\ln x$ (für $x \rightarrow \infty$).

Beweis. Es ist $\lim_{x \rightarrow \infty} \left(\frac{x^r}{\ln x}\right) = \infty$ zu zeigen. Mit $r > 0$ und der Regel von de l'Hospital für $\frac{\infty}{\infty}$ folgt:

$$\lim_{x \rightarrow \infty} \left(\frac{x^r}{\ln x}\right) = \lim_{x \rightarrow \infty} \left(\frac{rx^{r-1}}{\frac{1}{x}}\right) = \lim_{x \rightarrow \infty} (rx^r) = r \cdot \lim_{x \rightarrow \infty} (x^r) = \infty. \quad \square$$

Eng verwandte Sprechweise. Es gelte $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$. Man sagt, $f(x)$ strebt für $x \rightarrow x_0$ schneller gegen Null als $g(x)$, falls Folgendes gilt:

$$\lim_{x \rightarrow x_0} \left(\frac{f(x)}{g(x)}\right) = 0.$$

Ähnlich definiert man: $f(x)$ strebt für $x \rightarrow \infty$ schneller gegen Null als $g(x)$.

Wir haben bisher Grenzwerte vom Typ $\frac{0}{0}$ und $\frac{\infty}{\infty}$ behandelt. In der Praxis treten aber noch andere (ähnliche) Fälle auf, in denen der Grenzwert nicht ohne Weiteres berechnet werden kann. Diese Typen lassen sich auf die bereits behandelten Typen $\frac{0}{0}$ und $\frac{\infty}{\infty}$ zurückführen und werden im Folgenden besprochen werden.

2.7.3 Der Typ $0 \cdot \infty$

Durch die symbolische Schreibweise $0 \cdot \infty$ wird nur angedeutet, um welchen Typ es sich handelt. Wir geben deshalb noch eine genaue Festlegung. Wir sprechen vom Typ $0 \cdot \infty$, wenn ein Grenzwert

$$\lim_{x \rightarrow x_0} (f(x) \cdot g(x)) \quad (2.44)$$

berechnet werden soll, für den $f(x) \rightarrow 0$ und $|g(x)| \rightarrow \infty$ für $x \rightarrow x_0$ gilt.

Lösung. Anstelle von $f(x) \cdot g(x)$ schreibt man:

$$\frac{f(x)}{\frac{1}{g(x)}} \quad \text{oder} \quad \frac{g(x)}{\frac{1}{f(x)}}.$$

Dadurch wird die Berechnung des Grenzwertes (2.44) auf den Typ $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ zurückgeführt. (Siehe Berechnung von $\lim_{x \rightarrow 0} (x \cdot \ln x)$.)

2.7.4 Der Typ $\infty - \infty$

Wir geben zunächst wieder eine präzise Festlegung, wann wir vom Typ $\infty - \infty$ sprechen wollen. Der Typ $\infty - \infty$ liegt vor, wenn ein Grenzwert

$$\lim_{x \rightarrow x_0} (f(x) - g(x)) \quad (2.45)$$

berechnet werden soll, für den $|f(x)| \rightarrow \infty$ und $|g(x)| \rightarrow \infty$ für $x \rightarrow x_0$ gilt.

Lösung. Man formt $f(x) - g(x)$ wie folgt um:

$$f(x) - g(x) = \frac{\frac{1}{g(x)} - \frac{1}{f(x)}}{\frac{1}{g(x) \cdot f(x)}}. \quad (2.46)$$

Dadurch wird die Berechnung des Grenzwertes (2.45) auf den Typ $\frac{0}{0}$ zurückgeführt.

Beispiel. Es soll $\lim_{x \rightarrow 0} \left(\frac{1}{\ln(x+1)} - \frac{1}{x} \right)$ berechnet werden.

$$\begin{aligned} \lim_{x \rightarrow 0} \left(\frac{1}{\ln(x+1)} - \frac{1}{x} \right) &\stackrel{(2.46)}{=} \lim_{x \rightarrow 0} \left(\frac{x - \ln(x+1)}{x \ln(x+1)} \right) \\ &= \lim_{x \rightarrow 0} \left(\frac{1 - \frac{1}{x+1}}{\ln(x+1) + \frac{x}{x+1}} \right) \\ &= \lim_{x \rightarrow 0} \left(\frac{\frac{1}{(x+1)^2}}{\frac{1}{x+1} + \frac{1}{(x+1)^2}} \right) \\ &= \frac{1}{2} \end{aligned}$$

2.7.5 Die Typen 0^0 , 1^∞ und ∞^0

Diese Typen entsprechen einem Grenzwert vom folgenden Typ:

$$\lim_{x \rightarrow x_0} \left(f(x)^{g(x)} \right) \quad (\text{mit } f(x) > 0), \quad (2.47)$$

wobei für $x \rightarrow x_0$ gilt:

- Typ 0^0 : $f(x) \rightarrow 0$ und $g(x) \rightarrow 0$
- Typ 1^∞ : $f(x) \rightarrow 1$ und $|g(x)| \rightarrow \infty$
- Typ ∞^0 : $f(x) \rightarrow \infty$ und $g(x) \rightarrow 0$

Lösung. Diese drei Typen löst man dadurch, dass man (2.47) wie folgt umformt:

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

Hierdurch wird die Berechnung von (2.47) auf die Berechnung von $\lim_{x \rightarrow x_0} (g(x) \cdot \ln f(x))$ zurückgeführt, d.h. auf den bereits behandelten Typ $0 \cdot \infty$.

Beispiele:

1. Es soll $\lim_{x \rightarrow 0} (x^x)$ berechnet werden. Es liegt hier also der Typ 0^0 vor. Zeigen Sie:

$$\lim_{x \rightarrow 0} (x^x) = 1.$$

2. Es soll $\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}}$ berechnet werden. Es liegt hier also der Typ 1^∞ vor.

Dieser Grenzwert wurde bereits in Beispiel 6 (Abschnitt 2.7.1) berechnet.

3. Es soll $\lim_{x \rightarrow \infty} (x^{\frac{1}{x}})$ berechnet werden. Es liegt hier also der Typ ∞^0 vor.

Dieser Grenzwert wurde bereits in Beispiel 3 (Abschnitt 2.7.2) berechnet.

Bemerkung. In (2.44), (2.45) und (2.47) ist auch $x_0 = +\infty$ und $x_0 = -\infty$ zugelassen. Ferner gelten alle Ergebnisse des Abschnitts 2.7 entsprechend sowohl für $x \rightarrow -\infty$ als auch für einseitige Grenzwerte $x \rightarrow x_0+$ und $x \rightarrow x_0-$.

Allgemeinere Version von Satz 26. Jede Exponentialfunktion a^x (mit $a > 1$) wächst für $x \rightarrow \infty$ schneller als jede Potenzfunktion x^r (mit $r > 0$) und auch schneller als jedes Polynom $P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ (mit $a_n > 0$).

3 Integralrechnung

In der Integralrechnung geht es um die folgenden beiden Grundfragen:

(a) **Das Flächeninhaltsproblem.**

Wie bestimmt man den Inhalt einer Fläche, die krummlinig berandet ist?

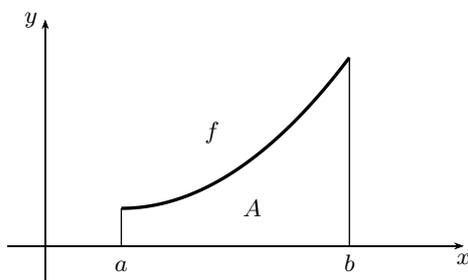
(b) **Das Stammfunktionsproblem.**

Wie findet man zu einer gegebenen Funktion f eine Stammfunktion, d.h. eine Funktion F , für die $F' = f$ gilt?

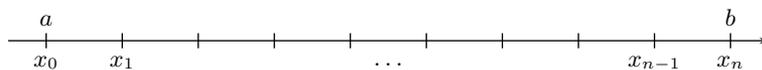
Wir werden sehen, dass es sich bei (a) und (b) im Wesentlichen um ein und dasselbe Problem handelt. Als Einstieg in die Integralrechnung wählen wir das Flächeninhaltsproblem, ähnlich wie wir als Ausgangspunkt für die Differentialrechnung das Tangentenproblem gewählt haben. Es sei jedoch darauf hingewiesen, dass die inner- und außermathematischen Anwendungen von Differential- und Integralrechnung weit über diese geometrischen Probleme hinausgehen; so findet beispielsweise die Integralrechnung auf vielfältige Weise Anwendung in den Natur- und Wirtschaftswissenschaften, in der Wahrscheinlichkeitstheorie und in der Statistik.

3.1 Berechnung der Fläche unter einer Kurve

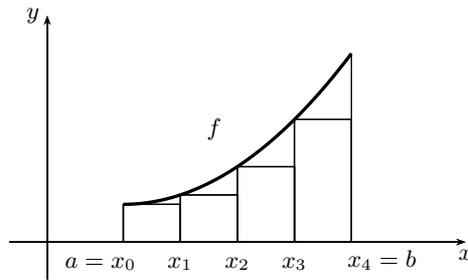
In diesem Abschnitt sei $f : [a, b] \rightarrow \mathbb{R}$ immer eine monoton steigende, nichtnegative Funktion. Mit A bezeichnen wir den Inhalt der Fläche, die durch den Graphen von f , die x -Achse und die beiden Geraden $x = a$ und $x = b$ eingeschlossen wird.



Zur Berechnung von A teilen wir das Intervall $[a, b]$ in n gleiche Teile, wobei wir die Teilungspunkte mit x_0, x_1, \dots, x_n bezeichnen wollen. Dabei gelte $a = x_0 < x_1 < x_2 < \dots < x_n = b$.



Der Abstand zweier Punkte x_{i-1} und x_i ist also immer derselbe (genauer: $x_i - x_{i-1} = \frac{b-a}{n}$). Wir nennen eine solche Zerlegung des Intervalls $[a, b]$ auch eine *äquidistante Zerlegung* von $[a, b]$. Wir zeichnen nun Strecken parallel zur y -Achse von den Punkten x_i bis zum Graphen von f und teilen dadurch die Fläche unterhalb der Kurve in Streifen auf. Sodann bilden wir zu diesen Streifen einbeschriebene Rechtecke, bei denen jeweils ein Eckpunkt auf dem Graphen von f liegt:



Den Flächeninhalt der entstehenden treppenförmigen Figur nennt man *Untersumme* bei Zerlegung von $[a, b]$ in n gleiche Teile. Oder kurz: n -te *Untersumme* (Bezeichnung: U_n). Es gilt:

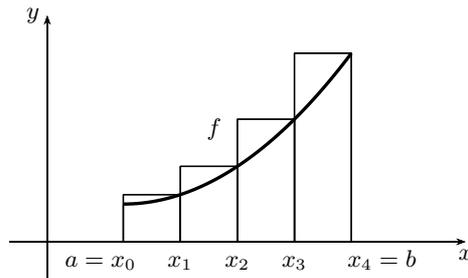
$$U_n = \sum_{i=1}^n f(x_{i-1})(x_i - x_{i-1}) = \frac{b-a}{n} \sum_{i=1}^n f(x_{i-1}). \quad (3.1)$$

(Man beachte: $x_i - x_{i-1} = \frac{b-a}{n}$ gilt für alle $i = 1, \dots, n$.)

Ganz entsprechend bildet man die n -te *Obersumme* O_n :

$$O_n = \sum_{i=1}^n f(x_i)(x_i - x_{i-1}) = \frac{b-a}{n} \sum_{i=1}^n f(x_i). \quad (3.2)$$

Darstellung der Obersumme:



Für die Differenz zwischen der n -ten Obersumme und der n -ten Untersumme gilt:

$$O_n - U_n = \frac{b-a}{n} \left(\sum_{i=1}^n f(x_i) - \sum_{i=1}^n f(x_{i-1}) \right) = \frac{b-a}{n} (f(x_n) - f(x_0)) = \frac{(b-a)(f(b) - f(a))}{n}. \quad (3.3)$$

Da $(b-a)(f(b) - f(a))$ konstant ist, folgt, dass die Differenz $O_n - U_n$ gegen Null strebt (für $n \rightarrow \infty$). Man kann ferner zeigen, dass die Grenzwerte $\lim_{n \rightarrow \infty} U_n$ und $\lim_{n \rightarrow \infty} O_n$ immer existieren. (Auf den Beweis dieser Tatsache wollen wir hier verzichten.) Es folgt:

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} (O_n - U_n) = \lim_{n \rightarrow \infty} O_n - \lim_{n \rightarrow \infty} U_n \\ &\implies \lim_{n \rightarrow \infty} O_n = \lim_{n \rightarrow \infty} U_n \end{aligned} \quad (3.4)$$

Die Folgen (O_n) und (U_n) konvergieren gegen einen gemeinsamen Grenzwert. Diesen gemeinsamen Grenzwert bezeichnen wir mit $\int_a^b f(x) dx$; wir definieren:

$$\int_a^b f(x) dx := \lim_{n \rightarrow \infty} O_n = \lim_{n \rightarrow \infty} U_n. \quad (3.5)$$

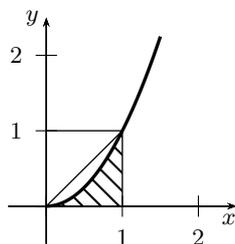
$\int_a^b f(x) dx$ heißt *bestimmtes Integral* von f über $[a, b]$.

Die Funktion $f(x)$ heißt *Integrand*. Die Bezeichnung der Variablen mit x ist willkürlich; man kann in (3.5) ebenso gut $\int_a^b f(t) dt$ oder $\int_a^b f(s) ds$ schreiben.

Wir kehren nun zurück zum Ausgangsproblem, den Inhalt A der Fläche zu berechnen, die durch den Graphen von f , die x -Achse und die beiden Geraden $x = a$ und $x = b$ eingeschlossen wird. Genaugenommen haben wir noch gar nicht definiert, was unter dem Inhalt A dieser Fläche verstanden werden soll. Unsere Anschauung legt uns aber nahe, dass für einen vernünftigt definierten Flächeninhalt A gelten muss: $U_n \leq A \leq O_n$ für alle $n \in \mathbb{N}$. Da $\lim_{n \rightarrow \infty} O_n = \lim_{n \rightarrow \infty} U_n = \int_a^b f(x) dx$ gilt, definieren wir den

Flächeninhalt A als den Wert von $\int_a^b f(x) dx$.

Beispiel. Wir wollen die Fläche berechnen, die von der Parabel $f(x) = x^2$, der x -Achse und der Geraden $x = 1$ eingeschlossen wird. Klar ist, dass der gesuchte Flächeninhalt zwischen 0 und $\frac{1}{2}$ liegen muss (siehe Skizze).



Wir berechnen $\lim_{n \rightarrow \infty} O_n$. Es gilt:

$$\begin{aligned} O_n &\stackrel{(3.2)}{=} \frac{b-a}{n} \sum_{i=1}^n f(x_i) = \frac{1}{n} \sum_{i=1}^n f\left(\frac{i}{n}\right) = \frac{1}{n} \sum_{i=1}^n \left(\frac{i}{n}\right)^2 = \frac{1}{n^3} \sum_{i=1}^n i^2 \\ &= \frac{1}{n^3} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{2n^2 + 3n + 1}{6n^2} \end{aligned}$$

Also gilt $\lim_{n \rightarrow \infty} O_n = \lim_{n \rightarrow \infty} \left(\frac{2n^2 + 3n + 1}{6n^2}\right) = \frac{1}{3}$. Der gesuchte Flächeninhalt hat den Wert $\frac{1}{3}$.

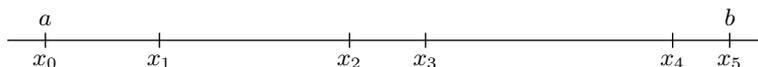
Anstelle der bisher betrachteten Unter- und Obersummen werden wir im nächsten Abschnitt sog. *Zwischensummen* behandeln. Als Vorbereitung wollen wir diesen Begriff hier für den Spezialfall einer monoton steigenden, nichtnegativen Funktion f betrachten.

Es sei also (wie bisher) $f : [a, b] \rightarrow \mathbb{R}$ eine monoton steigende, nichtnegative Funktion. Zwischensummen sind in zweierlei Hinsicht allgemeiner als die oben betrachteten Summen U_n und O_n :

1) Wir betrachten eine Zerlegung Z von $[a, b]$ der folgender Form:

$$Z : a = x_0 < x_1 < x_2 < \dots < x_n = b.$$

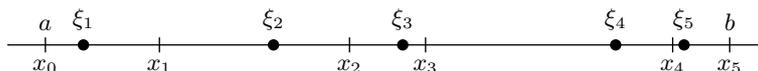
Diese Zerlegung braucht nicht äquidistant zu sein.



Als *Feinheitsmaß* $|Z|$ von Z bezeichnen wir den größten auftretenden Abstand $|x_i - x_{i-1}|$, d.h.:

$$|Z| = \max \left\{ |x_i - x_{i-1}| : i = 1, \dots, n \right\}. \quad (3.6)$$

2) Zu jedem Teilintervall $[x_{i-1}, x_i]$ der Zerlegung Z wählen wir einen beliebigen *Zwischenpunkt* $\xi_i \in [x_{i-1}, x_i]$.

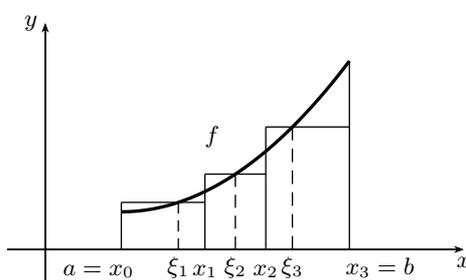


Als *Zwischensumme* S_Z bezeichnen wir die folgende Summe¹:

$$S_Z = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}). \quad (3.7)$$

Natürlich ist es auch erlaubt, dass $\xi_i = x_{i-1}$ oder $\xi_i = x_i$ gilt; das war bei den oben betrachteten Summen U_n und O_n ja gerade der Fall. (U_n und O_n sind also spezielle Zwischensummen.)

Man kann eine Zwischensumme S_Z ähnlich wie eine Untersumme U_n bzw. eine Obersumme O_n als einen Näherungswert für die Fläche unter dem Graphen von f auffassen. Wir illustrieren dies anhand der folgenden Zeichnung:



Wir wählen nun immer feinere Zerlegungen, d.h., wir lassen $|Z| \rightarrow 0$ streben. Bei jeweils beliebiger Wahl der Punkte ξ_i streben dann die zugehörigen Zwischensummen S_Z gegen das Integral $\int_a^b f(x) dx$, d.h. gegen den Inhalt der Fläche, die vom Graphen von f , der x -Achse und den beiden Geraden $x = a$ und $x = b$ eingeschlossen wird. (Auf den Beweis dieser Tatsache wollen wir hier verzichten.)

Als Ergebnis halten wir fest: Ist f eine monoton steigende, nichtnegative Funktion, so gilt bei jeweils beliebiger Wahl der Zwischenpunkte ξ_i :

$$|Z| \rightarrow 0 \quad \implies \quad S_Z \rightarrow \int_a^b f(x) dx. \quad (3.8)$$

3.2 Das bestimmte Integral

Im letzten Abschnitt haben wir den Spezialfall betrachtet, dass f eine monoton steigende, nichtnegative Funktion ist. Jetzt wollen wir den allgemeinen Fall behandeln. Von der Funktion $f : [a, b] \rightarrow \mathbb{R}$ wird jetzt nichts weiter vorausgesetzt, als dass f *beschränkt* ist, d.h., es gibt ein $S \in \mathbb{R}$ mit $|f(x)| \leq S$ für alle $x \in [a, b]$.

Wie im letzten Abschnitt definieren wir, was eine *Zerlegung*

$$Z : a = x_0 < x_1 < \dots < x_n = b$$

von $[a, b]$ ist, was das *Feinheitsmaß* $|Z|$ von Z ist und was man unter einer *Zwischensumme* S_Z versteht (für eine bestimmte Wahl von Zwischenpunkten $\xi_i \in [x_{i-1}, x_i]$). Wenn wir zum Ausdruck bringen wollen,

¹Anstelle von Zwischensumme sagt man auch *Riemannsche Summe*.

auf welche Funktion f und welche Wahl der Zwischenpunkte wir uns beziehen, so schreiben wir für eine Zwischensumme anstelle von S_Z auch

$$S_Z(f; \xi_1, \dots, \xi_n)$$

Definition.

Es sei $a < b$ und f sei eine in $[a, b]$ beschränkte Funktion. Wir betrachten Folgen Z_1, Z_2, \dots von Zerlegungen von $[a, b]$, für die $|Z_n| \rightarrow 0$ gilt. Streben für alle derartigen Folgen (Z_n) und für jede mögliche Wahl der Zwischenpunkte ξ_i die Zwischensummen S_{Z_n} immer gegen denselben Grenzwert A , so heißt f *integrierbar* auf $[a, b]$ und der Grenzwert A heißt *bestimmtes Integral* von f über $[a, b]$.

Den in dieser Definition A genannten Grenzwert bezeichnet man mit:

$$\int_a^b f(x) \, dx.$$

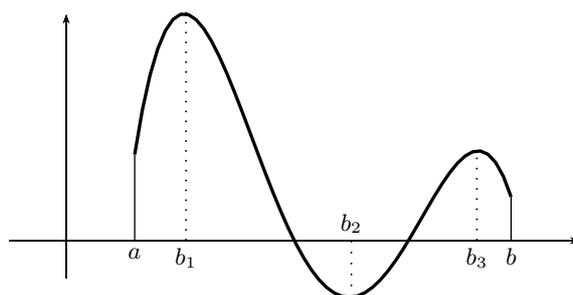
f heißt *Integrand*, a und b heißen *Integrationsgrenzen* und x heißt *Integrationsvariable*.

Es stellt sich die Frage, welche Funktionen integrierbar sind.

Definition.

Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt *stückweise monoton*, wenn sich das Intervall $[a, b]$ derart in endlich viele Teilintervalle zerlegen lässt, dass f auf jedem Teilintervall monoton ist.

Analog definiert man den Begriff *stückweise stetig*. (Nämlich wie?)



Die Funktion f , deren Graph oben skizziert ist, ist beispielsweise stückweise monoton, da f auf $[a, b_1]$ und $[b_2, b_3]$ monoton steigend und auf $[b_1, b_2]$ und $[b_3, b]$ monoton fallend ist.

Satz 1.

Ist f auf $[a, b]$ stückweise monoton, so ist f auf $[a, b]$ integrierbar.

Beweisidee (grobe Skizze). Man überlegt sich zunächst, dass jede monoton steigende, nichtnegative Funktion $f : [a, b] \rightarrow \mathbb{R}$ integrierbar ist. Dies haben wir (ohne auf alle Details des Beweises einzugehen) bereits in (3.8) festgestellt. Dieses Ergebnis lässt sich auf stückweise monotone Funktionen übertragen, indem man $[a, b]$ in Teilintervalle zerlegt, auf denen f monoton und entweder nichtnegativ oder nichtpositiv ist. \square

Satz 2.

Ist f auf $[a, b]$ stückweise stetig, so ist f auf $[a, b]$ integrierbar.

Auf den Beweis von Satz 2 wollen wir verzichten.

In den Sätzen 1 und 2 ist festgestellt worden, dass umfangreiche Klassen von Funktionen zu den integrierbaren Funktionen gehören. Insbesondere ist also jede stetige Funktion integrierbar.

Beispiel einer nichtintegrierbaren Funktion.

$$f : [0, 1] \rightarrow \mathbb{R} \text{ sei definiert durch } f(x) = \begin{cases} 1 & , \text{ falls } x \in \mathbb{Q} \\ 0 & , \text{ falls } x \notin \mathbb{Q} \end{cases}$$

Man nennt diese Funktion die *Dirichletsche Sprungfunktion*. Wir zeigen, dass die Dirichletsche Sprungfunktion nicht integrierbar ist.

Z_n sei die äquidistante Zerlegung von $[0, 1]$ in n Teilintervalle $[\frac{i-1}{n}, \frac{i}{n}]$ ($i = 1, \dots, n$). Zu jedem dieser Teilintervalle wähle man Punkte $\xi_{i,n}, \eta_{i,n} \in [\frac{i-1}{n}, \frac{i}{n}]$ mit

$$\xi_{i,n} \in \mathbb{Q} \quad \text{und} \quad \eta_{i,n} \notin \mathbb{Q},$$

beispielsweise $\xi_{i,n} = \frac{i}{n}$ und $\eta_{i,n} = \frac{\sqrt{2}+i-2}{n}$; f bezeichne die Dirichletsche Sprungfunktion. Dann gilt:

$$S_{Z_n}(f; \xi_{1,n}, \dots, \xi_{n,n}) = 1 \quad \text{und} \quad S_{Z_n}(f; \eta_{1,n}, \dots, \eta_{n,n}) = 0.$$

Für $n \rightarrow \infty$ gilt also:

$$S_{Z_n}(f; \xi_{1,n}, \dots, \xi_{n,n}) \rightarrow 1 \quad \text{und} \quad S_{Z_n}(f; \eta_{1,n}, \dots, \eta_{n,n}) \rightarrow 0.$$

Also ist f nicht integrierbar.

Wir haben bei unserer Definition von $\int_a^b f(x) dx$ bisher vorausgesetzt, dass die untere Integrationsgrenze kleiner ist als die obere. Darüber hinaus definiert man für $a < b$:

$$\int_b^a f(x) dx := - \int_a^b f(x) dx \quad \text{und} \quad \text{außerdem} \quad \int_a^a f(x) dx = 0. \quad (3.9)$$

In den folgenden beiden Sätzen werden grundlegende Eigenschaften des bestimmten Integrals aufgeführt, die man beim Rechnen immer wieder benutzt.

Satz 3.

$f : [a, b] \rightarrow \mathbb{R}$ und $g : [a, b] \rightarrow \mathbb{R}$ seien integrierbare Funktionen und es sei $\alpha \in \mathbb{R}$. Dann sind auch $f + g$ und αf auf $[a, b]$ integrierbar und es gilt:

$$\int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x) dx \quad (3.10)$$

$$\int_a^b \alpha f(x) dx = \alpha \int_a^b f(x) dx \quad (3.11)$$

Beweis. Es sei $S_Z = \sum_{i=1}^n (f + g)(\xi_i)(x_i - x_{i-1})$ eine beliebige Zwischensumme für $f + g$. Es gilt:

$$\begin{aligned} S_Z &= \sum_{i=1}^n (f + g)(\xi_i)(x_i - x_{i-1}) \\ &= \sum_{i=1}^n (f(\xi_i) + g(\xi_i))(x_i - x_{i-1}) \\ &= \sum_{i=1}^n (f(\xi_i)(x_i - x_{i-1}) + g(\xi_i)(x_i - x_{i-1})) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) + \sum_{i=1}^n g(\xi_i)(x_i - x_{i-1}) \\
&= S'_Z + S''_Z,
\end{aligned}$$

Da S'_Z eine Zwischensumme für f und S''_Z eine Zwischensumme für g ist, folgt mit $|Z| \rightarrow 0$:

$$S'_Z \rightarrow \int_a^b f(x) dx \quad \text{und} \quad S''_Z \rightarrow \int_a^b g(x) dx.$$

Wegen $S_Z = S'_Z + S''_Z$ gilt für $|Z| \rightarrow 0$ also $S_Z \rightarrow \int_a^b f(x) dx + \int_a^b g(x) dx$. Man beachte, dass S_Z eine beliebige Zwischensumme für $f + g$ ist und dass wir gezeigt haben: Aus $|Z| \rightarrow 0$ folgt immer $S_Z \rightarrow \int_a^b f(x) dx + \int_a^b g(x) dx$. Das bedeutet, dass $f + g$ integrierbar ist und dass (3.10) gilt. Analog zeigt man (3.11). \square

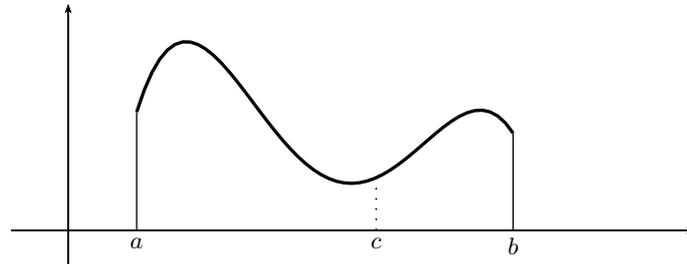
Satz 4.

Es sei $c \in [a, b]$. Ist f auf $[a, b]$ integrierbar, so ist f auch auf $[a, c]$ und $[c, b]$ integrierbar. Es gilt:

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (3.12)$$

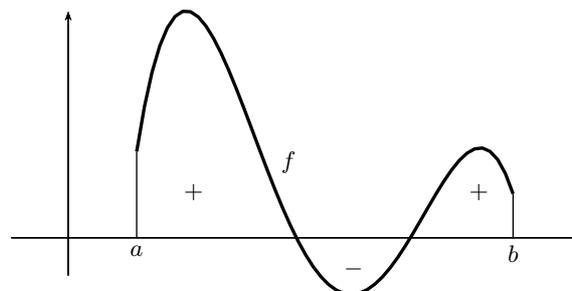
Umgekehrt gilt auch: Ist f auf $[a, c]$ und $[c, b]$ integrierbar, so ist f auf $[a, b]$ integrierbar und es gilt (3.12).

Auf den Beweis dieses Satzes wollen wir verzichten. Wir begnügen uns damit, uns die Aussage von Satz 4 anhand einer Skizze plausibel zu machen.



Es soll bei dieser Gelegenheit noch einmal ausdrücklich auf die *geometrische Bedeutung des bestimmten Integrals* hingewiesen werden, die sich aus der Definition des bestimmten Integrals ergibt.

Für $a \leq b$ gibt das Integral $\int_a^b f(x) dx$ den Flächeninhalt zwischen der x -Achse, dem Graphen von f und den Geraden $x = a$ und $x = b$ an, wobei Inhalte von Flächen oberhalb der x -Achse positiv, unterhalb der x -Achse negativ gerechnet werden.



Zusatz. Die beiden Formeln (3.10) und (3.11) bleiben auch richtig, wenn $b \leq a$ gilt. Ebenso bleibt die Formel (3.12) für jede beliebige Lage von a , b und c auf der Zahlengeraden richtig. $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$ gilt also beispielsweise auch für $a < b < c$ oder $b < c < a$.

Weitere beim Rechnen mit Integralen häufig benutzte Eigenschaften liefert der folgende Satz.

Satz 5.

Es sei $a \leq b$ und die Funktionen f und g seien auf $[a, b]$ integrierbar.

(a) Gilt $f(x) \leq g(x)$ für alle $x \in [a, b]$, so folgt $\int_a^b f(x) dx \leq \int_a^b g(x) dx$.

(b) Mit $f(x)$ ist auch $|f(x)|$ integrierbar und es gilt $\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$.

(c) $\left| \int_a^b (f(x) + g(x)) dx \right| \leq \int_a^b |f(x)| dx + \int_a^b |g(x)| dx$

Beweis.

(a) Dies folgt ähnlich wie Satz 3 aus der Definition der Integrierbarkeit.

(b) Auf den Beweis, dass $|f(x)|$ integrierbar ist, sei verzichtet. Die Formel folgt so: Es gilt $f(x) \leq |f(x)|$ und $-f(x) \leq |f(x)|$ für alle $x \in [a, b]$. Nach (a) und Satz 3 folgt:

$$\int_a^b f(x) dx \stackrel{(a)}{\leq} \int_a^b |f(x)| dx \quad \text{und} \quad -\int_a^b f(x) dx \stackrel{(3.11)}{=} \int_a^b -f(x) dx \stackrel{(a)}{\leq} \int_a^b |f(x)| dx.$$

Also gilt (b).

(c) Dies folgt aus (b) und Satz 3, indem die Dreiecksungleichung für den Betrag benutzt wird.

$$\begin{aligned} \left| \int_a^b (f(x) + g(x)) dx \right| &= \left| \int_a^b f(x) dx + \int_a^b g(x) dx \right| \\ &\leq \left| \int_a^b f(x) dx \right| + \left| \int_a^b g(x) dx \right| \\ &\stackrel{(b)}{\leq} \int_a^b |f(x)| dx + \int_a^b |g(x)| dx \quad \square \end{aligned}$$

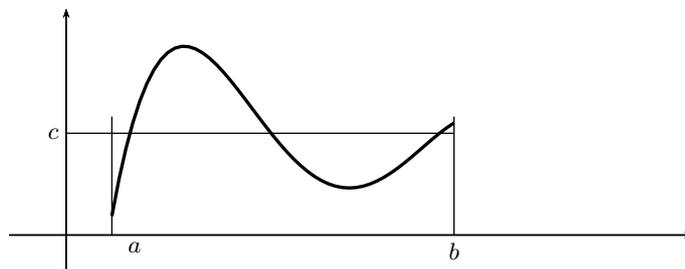
Bemerkung. Das in diesem Abschnitt eingeführte bestimmte Integral nennt man das *Riemannsches Integral*.

Wir schließen eine *Bemerkung zum Durchschnittswert einer Funktion auf einem Intervall* $[a, b]$ an.

Gegeben sei eine integrierbare Funktion $f : [a, b] \rightarrow \mathbb{R}$. In vielen Situationen möchte man den Durchschnittswert von f im Intervall $[a, b]$ bestimmen. Aber wie ist dieser Durchschnittswert c überhaupt definiert? Orientiert man sich an der geometrischen Anschauung, so ist unmittelbar einleuchtend, dass für c gelten sollte:

$$c(b - a) = \int_a^b f(x) dx. \tag{*}$$

Die Gleichung (*) bedeutet, wenn man sie sich anhand einer nichtnegativen Funktion klar macht: Das Rechteck mit den Seitenlängen c und $b - a$ hat dieselbe Fläche wie das Flächenstück, das durch den Graphen von f begrenzt wird:



Aus (*) erhält man als *Definition des Durchschnittswerts von f im Intervall $[a, b]$* :

$$c = \frac{1}{b-a} \int_a^b f(x) dx.$$

Anstelle von „Durchschnittswert“ sagt man auch „Mittelwert“. Wir schließen den Abschnitt mit dem folgenden Satz.

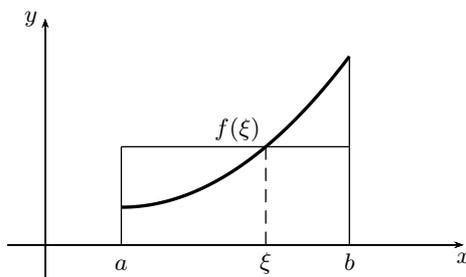
Satz 6 (Mittelwertsatz der Integralrechnung).

Die Funktion $f : [a, b] \rightarrow \mathbb{R}$ sei stetig. Dann gibt es ein $\xi \in [a, b]$ mit:

$$\int_a^b f(x) dx = f(\xi)(b-a). \quad (3.13)$$

Für den Beweis dieses Satzes verweisen wir auf die im Literaturverzeichnis aufgeführten Standardlehrbücher der Analysis.

Veranschaulichung für eine auf $[a, b]$ nichtnegative Funktion f : Die Fläche unter dem Graphen von f ist für ein geeignetes $\xi \in [a, b]$ gleich dem Inhalt des Rechtecks mit den Seitenlängen $b - a$ und $f(\xi)$.

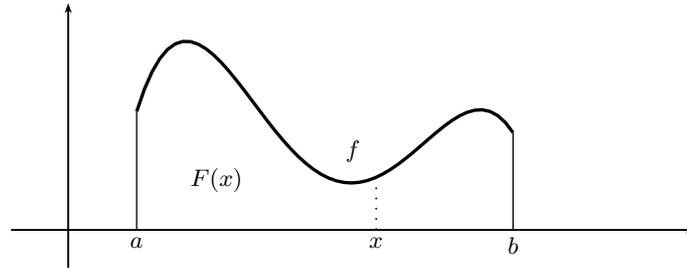


3.3 Der Fundamentalsatz der Differential- und Integralrechnung

Die Funktion f sei integrierbar auf $[a, b]$. Nach Satz 4 ist sie dann auch für jedes $x \in [a, b]$ auf dem Intervall $[a, x]$ integrierbar. Wir können also die folgende neue Funktion $F : [a, b] \rightarrow \mathbb{R}$ bilden:

$$F(x) := \int_a^x f(t) dt. \quad (3.14)$$

Diese Funktion $F(x)$ nennt man auch *Integral von f als Funktion der oberen Grenze*. Anschauliche Deutung (für nichtnegatives f): $F(x)$ gibt zwischen a und x den Flächeninhalt der Fläche unter dem Graphen von f an.



Klarerweise gilt $F(a) = 0$ und $F(b) = \int_a^b f(t) dt$.

Satz 7 (Fundamentalsatz der Differential- und Integralrechnung).

Die Funktion $f : [a, b] \rightarrow \mathbb{R}$ sei auf $[a, b]$ stetig. Dann ist die Funktion $F(x)$, die durch

$$F(x) := \int_a^x f(t) dt$$

definiert ist, differenzierbar auf $[a, b]$ und es gilt:

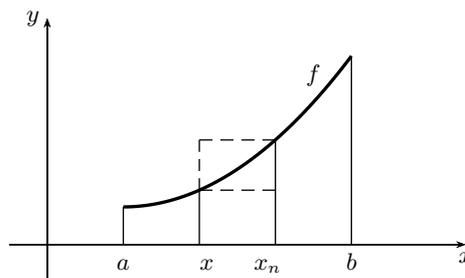
$$F'(x) = f(x). \tag{3.15}$$

Beweis. Zum Beweis dieses Satzes müssen wir für $x \in [a, b]$ und für eine beliebige Folge (x_n) mit $x_n \in [a, b]$, $x_n \neq x$ und $x_n \rightarrow x$ zeigen:

$$\frac{F(x_n) - F(x)}{x_n - x} \rightarrow f(x). \tag{3.16}$$

Wir wollen uns hier damit begnügen, die Beweisidee anhand des speziellen Falles vorzuführen, dass f nichtnegativ und monoton steigend ist und dass $x < x_n$ für alle n gilt.

Da $F(x_n) - F(x) = \int_a^{x_n} f(t) dt - \int_a^x f(t) dt \stackrel{(3.9)}{=} \int_a^{x_n} f(t) dt + \int_x^a f(t) dt = \int_x^{x_n} f(t) dt + \int_x^a f(t) dt \stackrel{(3.12)}{=} \int_x^{x_n} f(t) dt$ gilt², lässt sich der Zähler $F(x_n) - F(x)$ von (3.16) als Flächeninhalt veranschaulichen.



Da f monoton steigend ist und da $x < x_n$ gilt, folgt für diesen Flächeninhalt (siehe auch die Zeichnung):

$$f(x)(x_n - x) \leq F(x_n) - F(x) \leq f(x_n)(x_n - x).$$

Es folgt:

$$f(x) \leq \frac{F(x_n) - F(x)}{x_n - x} \leq f(x_n). \tag{3.17}$$

²Man beachte den Zusatz in Abschnitt 3.2.

Da wir f als stetig vorausgesetzt haben, folgt aus $x_n \rightarrow x$, dass $f(x_n) \rightarrow f(x)$ gilt. Also gilt wegen (3.17) $\frac{F(x_n) - F(x)}{x_n - x} \rightarrow f(x)$, was zu zeigen war. \square

Der Fundamentalsatz³ stellt die Verbindung zwischen Differential- und Integralrechnung her. Wir werden ihn in erster Linie dazu benutzen, bestimmte Integrale $\int_a^b f(x) dx$ zu berechnen.

Definition.

Gegeben sei eine Funktion $f : X \rightarrow \mathbb{R}$ (für $X \subseteq \mathbb{R}$). Ist $F : X \rightarrow \mathbb{R}$ eine auf X differenzierbare Funktion, für die $F'(x) = f(x)$ für alle $x \in X$ gilt, so nennt man F eine *Stammfunktion* von f .

Wir können den Fundamentalsatz also auch so formulieren:

Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, so ist die Funktion $F(x) = \int_a^x f(t) dt$ eine Stammfunktion von f .

Satz 8.

Sind F_1 und F_2 im Intervall I Stammfunktionen von f , so unterscheiden sich F_1 und F_2 nur um eine additive Konstante, d.h., es gibt eine Konstante $c \in \mathbb{R}$, so dass für alle $x \in I$ gilt:

$$F_1(x) = F_2(x) + c.$$

Beweis. Da F_1 und F_2 differenzierbar sind, ist auch $F_1 - F_2$ differenzierbar und es gilt für alle x :

$$(F_1 - F_2)'(x) = F_1'(x) - F_2'(x) = f(x) - f(x) = 0.$$

Die Ableitung der Funktion $F_1 - F_2$ ist also überall gleich Null. Folglich⁴ muss $F_1 - F_2$ eine konstante Funktion sein, d.h.:

$$(F_1 - F_2)(x) = F_1(x) - F_2(x) = c.$$

Es folgt Satz 8. \square

Als **wichtige Folgerung aus dem Fundamentalsatz** erhalten wir den folgenden Satz⁵.

Satz 9.

Die Funktion f sei stetig in $[a, b]$ und F sei irgendeine Stammfunktion zu f . Dann gilt:

$$\int_a^b f(x) dx = F(b) - F(a). \tag{3.18}$$

Beweis. Aus dem Fundamentalsatz wissen wir, dass die Funktion $\int_a^x f(t) dt$ eine Stammfunktion von f ist. Ist also $F(x)$ irgendeine Stammfunktion von f , so muss nach Satz 8 gelten (für ein $c \in \mathbb{R}$):

$$F(x) = \int_a^x f(t) dt + c. \tag{3.19}$$

Aus (3.19) ergibt sich:

$$F(b) - F(a) = \int_a^b f(t) dt + c - \left(\int_a^a f(t) dt + c \right) = \int_a^b f(t) dt. \quad \square$$

³Statt Fundamentalsatz sagt man auch *Hauptsatz der Differential- und Integralrechnung*.

⁴Siehe Satz 14, Abschnitt 2.4.

⁵Oft wird auch diese Folgerung als Fundamental- oder Hauptsatz der Differential- und Integralrechnung bezeichnet.

3.4 Berechnung von Integralen

Durch Satz 9 haben wir eine Methode gewonnen, um ein bestimmtes Integral

$$\int_a^b f(x) dx$$

zu berechnen (für stetiges f): Man beschaffe sich eine Stammfunktion F , berechne $F(a)$ und $F(b)$ und bilde $F(b) - F(a)$.

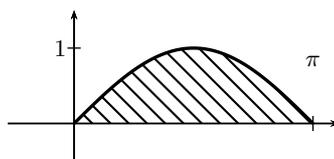
Wir werden auf das Problem geführt, zu einer gegebenen stetigen Funktion f eine Stammfunktion F zu ermitteln. Während das Ableiten einer Funktion i.A. unproblematisch ist, kann das Auffinden einer Stammfunktion größte Schwierigkeiten bereiten. Bevor wir uns mit etwas schwierigeren Fällen auseinandersetzen, wollen wir zunächst einfache Beispiele betrachten, in denen Satz 9 zur Berechnung von bestimmten Integralen eingesetzt wird.

Beispiele:

1. Es soll die Fläche berechnet werden, die von der Parabel $f(x) = x^2$, der x -Achse und der Geraden $x = 1$ eingeschlossen wird (siehe Beispiel in Abschnitt 3.1): Eine Stammfunktion von $f(x) = x^2$ ist $F(x) = \frac{1}{3}x^3$. Folglich gilt:

$$\int_0^1 x^2 dx = F(1) - F(0) = \frac{1}{3}.$$

2. Es soll die Fläche zwischen der Sinuskurve und der x -Achse zwischen 0 und π berechnet werden.



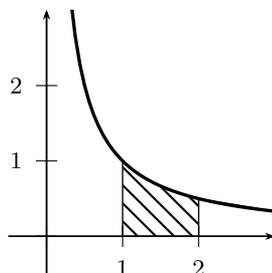
Eine Stammfunktion von $\sin x$ ist $-\cos x$. Folglich gilt:

$$\int_0^{\pi} \sin x dx = -\cos \pi - (-\cos 0) = 2.$$

3. Analog:

$$\int_0^{2\pi} \sin x dx = -\cos 2\pi - (-\cos 0) = 0.$$

4. Es soll die Fläche zwischen dem Graphen von $f(x) = \frac{1}{x}$, der x -Achse sowie den Geraden $x = 1$ und $x = 2$ berechnet werden.



Eine Stammfunktion von $f(x) = \frac{1}{x}$ ist $\ln x$. Folglich gilt:

$$\int_1^2 \frac{1}{x} dx = \ln 2 - \ln 1 \approx 0.693.$$

Zwei Schreibweisen.

1. Ist $F(x)$ eine Stammfunktion der stetigen Funktion $f(x)$ und ist $\int_a^b f(x) dx$ zu berechnen, so schreiben wir⁶:

$$\int_a^b f(x) dx = [F(x)]_a^b = F(b) - F(a).$$

Beispiel. $\int_{-4}^3 3x^3 dx = \left[\frac{3}{4}x^4 \right]_{-4}^3 = \frac{3}{4} \cdot 3^4 - \frac{3}{4} \cdot (-4)^4 = \frac{3}{4} \cdot (3^4 - 4^4) = -131.25.$

2. Ist F eine Stammfunktion von f , so wird F auch *unbestimmtes Integral* genannt und mit $\int f(x) dx$ bezeichnet. Üblich ist folgende Schreibweise: $\int f(x) dx = F(x)$. Dies ist keine Gleichung im üblichen Sinne, sondern nur eine kompakte Schreibweise für die Aussage: „ F ist eine Stammfunktion von f .“

Beispiele:

- $\int \sin x dx = -\cos x$
- $\int x^5 dx = \frac{1}{6}x^6$

Wie wir gesehen haben, ist es zur praktischen Berechnung eines Integrals $\int_a^b f(x) dx$ wichtig, eine Stammfunktion F von f zu kennen. Die folgende Tabelle gibt die Stammfunktionen einiger elementarer Funktionen an. (Durch Ableiten von F lässt sich deren Richtigkeit nachprüfen.)

Funktion f	Stammfunktion F
x^r (für $r \in \mathbb{R}, r \neq -1$)	$\frac{1}{r+1}x^{r+1}$
$\frac{1}{x}$	$\ln x $
$\frac{1}{1+x^2}$	$\arctan x$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$
e^x	e^x
$\sin x$	$-\cos x$
$\cos x$	$\sin x$
$\frac{1}{\cos^2 x} = 1 + \tan^2 x$	$\tan x$
$\frac{1}{\sin^2 x} = 1 + \cot^2 x$	$-\cot x$

⁶ $[F(x)]_a^b$ ist eine abkürzende Schreibweise für $F(b) - F(a)$. Statt $[F(x)]_a^b$ schreibt man häufig auch $F(x)\Big|_a^b$. Gebräuchlich sind auch die Schreibweisen $[F(x)]_{x=a}^b$ und $F(x)\Big|_{x=a}^b$ sowie $[F(x)]_{x=a}^{x=b}$ und $F(x)\Big|_{x=a}^{x=b}$.

Bemerkung zur zweiten Zeile der obigen Tabelle.

Die Funktion $f(x) = \ln|x|$ ist für alle $x \in \mathbb{R}$ mit $x \neq 0$ definiert. Prüfen wir nach, ob in der Tat $(\ln|x|)' = \frac{1}{x}$ für alle $x \neq 0$ gilt:

1. Fall: $x > 0$

Dann gilt $\ln|x| = \ln x$. Man erhält also (siehe Satz 11, Abschnitt 2.3.3):

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

2. Fall: $x < 0$

Dann gilt $\ln|x| = \ln(-x)$. Man erhält unter Verwendung der Kettenregel:

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

Damit haben wir die Richtigkeit der Formel $(\ln|x|)' = \frac{1}{x}$ in beiden Fällen bestätigt. *Empfehlung:* Fertigen Sie eine Skizze der für alle $x \in \mathbb{R} \setminus \{0\}$ definierten Funktion $f(x) = \ln|x|$ an.

Tabellen, in denen man Stammfunktionen nachschlagen kann, findet man beispielsweise in folgendem Buch:

- J. N. Bronstein, K. A. Semendjajew, G. Musiol, H. Mühlig: Taschenbuch der Mathematik, Frankfurt (Main), 2000.

In zahlreichen anderen Büchern finden sich ebenfalls Tabellen der wichtigen Grundintegrale.

Hat man ein bestimmtes Integral $\int_a^b f(x) dx$ zu berechnen, so steht man vor der Aufgabe, eine Stammfunktion $F(x)$ von $f(x)$ zu finden; ist diese Aufgabe gelöst, kann man das bestimmte Integral sofort ausrechnen, indem man die Grenzen a und b einsetzt:

$$\int_a^b f(x) dx = F(b) - F(a).$$

Anders gesagt: Die Aufgabe, das bestimmte Integral $\int_a^b f(x) dx$ zu berechnen, ist im Wesentlichen erledigt, wenn es einem gelingt, das *unbestimmte Integral*

$$\int f(x) dx$$

zu ermitteln. Wir beschäftigen uns daher im Folgenden meist mit der Berechnung unbestimmter Integrale.

Zwei häufig benutzte Methoden zur Integration sind *partielle Integration* und *Integration durch Substitution*. Bevor wir diese im Folgenden besprechen, halten wir zunächst drei grundlegende Rechenregeln für unbestimmte Integrale fest:

$$(I) \int (f(x) + g(x)) dx = \int f(x) dx + \int g(x) dx$$

$$(II) \int (f(x) - g(x)) dx = \int f(x) dx - \int g(x) dx$$

$$(III) \int c \cdot f(x) dx = c \int f(x) dx$$

Diese Regeln sind nichts weiter als *Umkehrungen der entsprechenden Ableitungsregeln*:

$$(I) \quad (f + g)' = f' + g'$$

$$(II) \quad (f - g)' = f' - g'$$

$$(III) \quad (c \cdot f)' = c \cdot f'$$

In Worten ausgedrückt besagt beispielsweise die Integrationsregel (I): Man erhält eine Stammfunktion von $f + g$, indem man eine Stammfunktion von f und eine Stammfunktion von g addiert.

Auch die Produktregel $(f \cdot g)' = f' \cdot g + f \cdot g'$ und die Kettenregel $[f(g(x))]' = f'(g(x)) \cdot g'(x)$ kann man durch Umkehrung in Integrationsregeln verwandeln, was zu den bereits erwähnten Regeln der partiellen Integration und zur Substitutionsregel führt.

3.4.1 Partielle Integration

Satz 10 (Partielle Integration).

Die Funktionen f und g seien differenzierbar mit (auf einem Intervall) stetigen Ableitungen f' und g' . Dann gilt:

$$\int f'(x)g(x) \, dx = f(x)g(x) - \int f(x)g'(x) \, dx. \quad (3.20)$$

Beweis. Es wird in (3.20) behauptet, dass $f(x)g(x) - \int f(x)g'(x) \, dx$ eine Stammfunktion von $f'(x)g(x)$ ist. Durch Differenzieren überzeugt man sich leicht von der Richtigkeit dieser Behauptung. Die Ableitung von

$$f(x)g(x) - \int f(x)g'(x) \, dx$$

ist nach der Produktregel nämlich gerade

$$f'(x)g(x) + f(x)g'(x) - f(x)g'(x) = f'(x)g(x). \quad \square$$

Beispiel. $\int x \cdot \cos x \, dx$ ist zu berechnen.

Der Term $\cos x$ wird durch Übergang zur Stammfunktion nicht komplizierter, der Term x wird durch Ableiten einfacher. Daher lässt sich partielle Integration erfolgreich einsetzen, wenn man $f'(x) = \cos x$ und $g(x) = x$ setzt. Dann ist $f(x) = \sin x$ und $g'(x) = 1$. Nach der Regel der partiellen Integration ergibt sich somit

$$\begin{aligned} \int x \cdot \cos x \, dx &= x \cdot \sin x - \int \sin x \, dx \\ &= x \cdot \sin x - (-\cos x) \\ &= x \cdot \sin x + \cos x \end{aligned}$$

Probe:

$$\begin{aligned} (x \cdot \sin x + \cos x)' &= 1 \cdot \sin x + x \cdot \cos x - \sin x \\ &= x \cdot \cos x \end{aligned}$$

Bei der partiellen Integration wird die Aufgabe, $f' \cdot g$ zu integrieren, im Wesentlichen durch die Aufgabe, $f \cdot g'$ zu integrieren, ersetzt, was (wie im obigen Beispiel) vorteilhaft sein kann, wenn $f \cdot g'$ ein weniger komplizierter Ausdruck als $f' \cdot g$ ist.

Weitere **Beispiele**:

1. Es soll $\int e^x x^2 dx$ berechnet werden.

$$\begin{aligned}\int e^x x^2 dx &\stackrel{(3.20)}{=} x^2 e^x - \int 2x e^x dx \\ &= x^2 e^x - 2 \int x e^x dx \\ &\stackrel{(3.20)}{=} x^2 e^x - 2 \left(x e^x - \int e^x dx \right) \\ &= x^2 e^x - 2x e^x + 2e^x \\ &= e^x (x^2 - 2x + 2)\end{aligned}$$

Probe:

$$\begin{aligned}\left(e^x (x^2 - 2x + 2) \right)' &= e^x (2x - 2) + e^x (x^2 - 2x + 2) \\ &= x^2 e^x\end{aligned}$$

2. Es soll $\int \ln x dx$ berechnet werden.

Hier hilft ein kleiner Trick: Statt $\ln x$ schreibt man $1 \cdot \ln x$ und wählt $f'(x) = 1$ und $g(x) = \ln x$. Es folgt:

$$\begin{aligned}\int \ln x dx &= \int 1 \cdot \ln x dx \\ &\stackrel{(3.20)}{=} x \ln x - \int x \cdot \frac{1}{x} dx \\ &= x \ln x - x \\ &= x(\ln x - 1)\end{aligned}$$

Machen Sie die Probe!

Wir erwähnen abschließend, dass die partielle Integration häufig auch *Produktintegration* genannt wird.

3.4.2 Integration durch Substitution

Wir beginnen mit einem besonders einfachen Fall. Es soll ein Integral der folgenden Art berechnet werden, wobei vorausgesetzt wird, dass f stetig und g differenzierbar mit stetiger Ableitung g' ist:

$$\int f(g(x)) \cdot g'(x) dx.$$

Der Integrand hat also eine besondere Bauart: Es werden zwei Funktionen nacheinander ausgeführt, was zum Term $f(g(x))$ führt. Außerdem wird dieser Term mit der Ableitung $g'(x)$ der inneren Funktion g multipliziert.

Liegt ein Integrand dieser speziellen Bauart vor und ist F eine Stammfunktion von f , so gilt die folgende Regel, die eine Umkehrung der Kettenregel ist:

$$\int f(g(x)) \cdot g'(x) dx = F(g(x)). \quad (3.21)$$

Probe: Anwendung der Kettenregel ergibt:

$$\left(F(g(x)) \right)' = f(g(x)) \cdot g'(x).$$

Beispiele:

1. Es soll $\int 3 \sin(3x + 1) dx$ berechnet werden.

Der Integrand hat offenbar die spezielle Bauart $f(g(x)) \cdot g'(x)$. Es folgt:

$$\int 3 \sin(3x + 1) dx = -\cos(3x + 1).$$

2. Es soll $\int \cos(3x + 1) dx$ berechnet werden.

Hier hat der Integrand nicht die verlangte spezielle Bauart. Durch eine leichte Umformung lässt sich jedoch die Bauart $f(g(x)) \cdot g'(x)$ herstellen und das Integral berechnen:

$$\begin{aligned} \int \cos(3x + 1) dx &= \frac{1}{3} \int 3 \cos(3x + 1) dx \\ &= \frac{1}{3} \sin(3x + 1) \end{aligned}$$

3. Es soll $\int \frac{2x}{x^2 + 5} dx$ berechnet werden.

Auch in diesem Beispiel hat der Integrand die spezielle Bauart $f(g(x)) \cdot g'(x)$, was man erkennt, wenn man $g(x) = x^2 + 5$ und $f(x) = \frac{1}{x}$ wählt. Es folgt:

$$\int \frac{2x}{x^2 + 5} dx = \ln|x^2 + 5|.$$

Machen Sie die Probe!

Im letzten Beispiel hatte der Integrand die Form $\frac{g'(x)}{g(x)}$, d.h. *im Zähler stand die Ableitung des Nenners*. Analog zu diesem Beispiel gilt:

$$\int \frac{g'(x)}{g(x)} dx = \ln|g(x)|. \quad (3.22)$$

Bestätigen Sie diese Integrationsregel, indem Sie die Probe machen!

Die Regel (3.22) ist ein häufig auftretender Spezialfall der Regel (3.21). Zum Einüben des Umgangs mit der Regel (3.22) berechne man die folgenden Integrale und mache jeweils die Probe:

(i) $\int \frac{1}{x - 5} dx$

(ii) $\int \frac{1}{2x - 5} dx$

(iii) $\int \frac{-\sin x}{\cos x} dx$

(iv) $\int \tan x dx$

Bei der zuvor besprochenen Regel (3.21) handelt es sich um eine Formel, die eng verwandt mit der Substitutionsregel ist. *Bei der nun zu besprechenden Substitutionsregel geht es um Folgendes:* Zu berechnen ist das Integral $\int f(x) dx$. In manchen Fällen gelingt das, indem man eine *Ersetzung* vornimmt (*substituieren* bedeutet ersetzen bzw. einsetzen): Man setzt $x = g(t)$, d.h., man ersetzt x durch eine Funktion $g(t)$, wobei man einen neuen Buchstaben für die Integrationsvariable wählt (hier t statt x). Das ist aber noch nicht alles. Auch das Symbol dx muss ersetzt werden. Wir werden sehen, dass $dx = g'(t) dt$ die richtige Ersetzung ist. Hat man die Funktion $g(t)$ geeignet gewählt (Darauf kommt es an!), so kann man das Integral berechnen, das nun wie folgt lautet:

$$\int f(g(t)) \cdot g'(t) dt.$$

Zum Schluss muss man dann im Ergebnis noch eine *Resubstitution* vornehmen, d.h., man muss die Variable t wieder ersetzen, nämlich durch $g^{-1}(x)$.

Im folgenden Satz wird all dies zusammengefasst und es werden die genauen Voraussetzungen genannt, unter denen die Vorgehensweise korrekt ist.

Satz 11 (Substitutionsregel).

Es sei f eine auf dem Intervall I stetige Funktion und es soll dort $\int f(x) dx$ berechnet werden. Für ein Intervall I' sei $g : I' \rightarrow I$ eine bijektive Funktion mit stetiger Ableitung. Dann lässt sich $\int f(x) dx$ berechnen, indem man

1. $\int f(g(t)) \cdot g'(t) dt$ ermittelt;
2. im Ergebnis $t = g^{-1}(x)$ substituiert („Resubstitution“).

Formelmäßig drückt man dies häufig auch durch die folgende Schreibweise aus:

$$\int f(x) dx = \left[\int f(g(t)) \cdot g'(t) dt \right]_{t=g^{-1}(x)} \quad (3.23)$$

Beweis. Es sei $\int f(g(t)) \cdot g'(t) dt = H(t)$.

Mit anderen Worten: $H(t)$ sei irgendeine Stammfunktion von $f(g(t)) \cdot g'(t)$. In (3.23) wird behauptet, dass $H(g^{-1}(x))$ eine Stammfunktion von $f(x)$ ist. Dies folgt so: Ist $F(x)$ eine Stammfunktion von $f(x)$, so ist nach der Kettenregel $F(g(t))$ eine Stammfunktion von $f(g(t)) \cdot g'(t)$. Da sowohl $H(t)$ als auch $F(g(t))$ eine Stammfunktion von $f(g(t)) \cdot g'(t)$ ist, gilt nach Satz 8:

$$H(t) = F(g(t)) + c.$$

Es folgt:

$$H(g^{-1}(x)) = F(x) + c.$$

$H(g^{-1}(x))$ ist also (wie behauptet) eine Stammfunktion von $f(x)$. \square

Wir erläutern den Satz an einem **Beispiel**. Es sei $f(x) = e^{\sqrt{x}}$. Die Funktion f ist auf dem Intervall $I = [0, \infty)$ definiert und es soll dort $\int f(x) dx = \int e^{\sqrt{x}} dx$ berechnet werden. Um die Substitutionsregel anwenden zu können, ist eine bijektive Funktion $g : I' \rightarrow I$ zu wählen. Wir wählen $g(t) = t^2$ mit $I' = [0, \infty)$ und setzen $x = g(t) = t^2$. Es gilt $g'(t) = 2t$ und $f(g(t)) = e^{\sqrt{t^2}} = e^t$ sowie $t = g^{-1}(x) = \sqrt{x}$. Nach der Substitutionsregel erhalten wir:

$$\int e^{\sqrt{x}} dx = \left[\int e^t \cdot 2t dt \right]_{t=\sqrt{x}}.$$

Es gilt:

$$\begin{aligned} \int e^t \cdot 2t dt &= 2 \int e^t \cdot t dt \\ &\stackrel{(*)}{=} 2 \left(e^t \cdot t - \int e^t dt \right) \\ &= 2 \left(e^t \cdot t - e^t \right) \\ &= 2e^t (t - 1) \end{aligned}$$

An der Stelle (*) wurde partielle Integration verwendet. Nach der Resubstitution erhält man:

$$\int e^{\sqrt{x}} dx = 2e^{\sqrt{x}} (\sqrt{x} - 1).$$

Machen Sie die *Probe* für $x > 0$, indem Sie die erhaltene Funktion $2e^{\sqrt{x}}(\sqrt{x} - 1)$ ableiten!

In Rechnungen lässt man auch häufig die eckigen Klammern $[\dots]_{t=g^{-1}(x)}$ weg. Man darf natürlich nicht vergessen, dass am Ende trotzdem die Resubstitution $t = g^{-1}(x)$ durchzuführen ist!

Wir haben bisher die Substitutionsregel eher aus dem Blickwinkel der *Theorie* besprochen. In der *Praxis* verwendet man die Substitutionsregel auf eine etwas andere Art, die inhaltlich gleichwertig ist, aber andere Schreibweisen benutzt. **Vorteil:** *Dies ermöglicht eine flexiblere Handhabung der Substitutionsregel.* In diesem Zusammenhang sei auf die folgende Notation für Ableitungen hingewiesen: Ist eine Funktion $x = g(t)$ gegeben, so schreibt man anstelle von $g'(t)$ auch $\frac{dx}{dt}$. (Man nennt dies die „Leibnizsche Schreibweise“; diese wird im nachfolgenden Beispiel noch ausführlich erklärt werden.) Ist nun $\int f(x) dx$ das Integral, das unter Verwendung der Substitutionsregel zu bestimmen ist, so geht man in zwei Schritten vor:

1. Man schreibt auf, welche Funktion gleich t sein soll. In unserem Beispiel schreibt man also $t = \sqrt{x}$. Allgemein würde man (mit den obigen Bezeichnungen) $t = g^{-1}(x)$ aufschreiben.
2. Im zweiten Schritt ermittelt man, wodurch dx zu ersetzen ist. Hierzu löst man zunächst die Gleichung $t = g^{-1}(x)$ nach x auf und erhält $x = g(t)$. Danach leitet man die Funktion $g(t)$ ab und erhält unter Benutzung der Leibnizschen Schreibweise:

$$\frac{dx}{dt} = g'(t).$$

Rechnet man mit den Symbolen dx und dt so, als seien sie Zahlen, so erhält man:

$$dx = g'(t) dt.$$

Man nennt diese Vorgehensweise *formales Rechnen mit dx und dt* . Die erhaltene Gleichung gibt an, wodurch dx zu ersetzen ist.

Kurze Zusammenfassung.

1. Man legt fest, was t sein soll: im Beispiel $t = \sqrt{x}$ bzw. allgemein $t = g^{-1}(x)$.
2. Man „rechnet“:

$$t = g^{-1}(x) \implies x = g(t) \implies \frac{dx}{dt} = g'(t) \implies dx = g'(t) dt.$$

Beispiel:

$$t = \sqrt{x} \implies x = t^2 \implies \frac{dx}{dt} = 2t \implies dx = 2t dt.$$

Häufig wird anstelle von 2) auch ein *abgekürztes Verfahren* verwendet, das auf der Umkehrregel basiert: Die Abkürzung besteht darin, dass man die Gleichung $t = g^{-1}(x)$ gar nicht erst nach x auflöst, sondern stattdessen gleich ableitet. Genauer: Man leitet die Funktion $t = g^{-1}(x)$ nach der Variablen x ab. Wir demonstrieren dies anhand unseres Beispiels:

$$t = \sqrt{x} \implies \frac{dt}{dx} = \frac{1}{2\sqrt{x}} = \frac{1}{2t} \implies dx = 2t dt.$$

Wir erläutern nun alles noch einmal ausführlich anhand eines weiteren Beispiels.

Beispiel. Das Integral $\int e^{\sqrt{3x+1}} dx$ soll mithilfe der Substitutionsregel so umgeformt werden, dass wir in der Lage sind, es zu lösen. Dies geschieht in zwei Schritten.

1. Schritt: Man setzt $t = \sqrt{3x+1}$.

In diesem Schritt ist nichts weiter geschehen, als dass man festgelegt hat, dass der relativ komplizierte (und deshalb störende) Ausdruck $\sqrt{3x+1}$ durch t ersetzt werden soll.

Im zweiten Schritt muss nun noch dx ersetzt werden, aber natürlich nicht irgendwie, sondern die Ersetzung von dx muss gemäß der Substitutionsregel erfolgen. In unserem Beispiel wird sich herausstellen, dass die richtige Ersetzung

$$dx = \frac{2}{3}t \, dt$$

ist. Es stellt sich die Frage, wie man darauf kommt. Oder, etwas anders gefragt:

Wie findet man die richtige Ersetzung von dx ?

Weiter unten im Schritt 2 werden wir rezeptartig angeben, wie man die richtige Ersetzung von dx findet. Der folgende Einschub bereitet dies vor.

Einschub.

Es liege eine Funktion in Form einer Gleichung vor, beispielsweise (für $x, y \in \mathbb{R}$):

$$y = x^2 + 1. \tag{3.24}$$

Um zum Ausdruck zu bringen, dass y abhängig von x ist, kann man auch $y(x)$ anstelle von y schreiben. In dieser Schreibweise ist $y = y(x)$ eine Funktion, die man ableiten kann:

$$y'(x) = 2x. \tag{3.25}$$

Wie Sie vermutlich wissen, sagt man: *Die Funktion y wird nach x abgeleitet.*

Für die Ableitung einer Funktion gibt es etliche Schreibweisen. NEWTON (1643 - 1727) benutzte beispielsweise die Schreibweise $\dot{y}(x)$, die heutzutage in der Physik noch vielfach verwendet wird. Auf LEIBNIZ (1646 - 1716) geht die Schreibweise $\frac{dy}{dx}$ (lies: dy nach dx) zurück, die besonders praktisch ist, beispielsweise im Zusammenhang mit der Substitutionsregel. Benutzt man die Leibnizsche Schreibweise, so geht (3.25) über in:

$$\frac{dy}{dx} = 2x.$$

Das Symbol $\frac{dy}{dx}$ ist also nur eine andere Schreibweise für die Ableitung $y'(x)$ und die Sprechweise „ dy nach dx “ spiegelt wider, dass die Funktion y nach x abgeleitet wird.

Alles in diesem Einschub Gesagte gilt natürlich ganz entsprechend, wenn man anstelle von y und x andere Buchstaben benutzt; liegt beispielsweise die Gleichung

$$x = t^3 + 2t - 5$$

vor, so ist $x = x(t)$ eine Funktion, die man nach t ableiten kann und man erhält:

$$\frac{dx}{dt} = 3t^2 + 2.$$

Zurück zur Aufgabe, das Integral $\int e^{\sqrt{3x+1}} dx$ zu bestimmen. Den 1. Schritt haben wir bereits oben beschrieben.

2. Schritt: Ersetzen von dx

Man löst zunächst die Gleichung $t = \sqrt{3x+1}$ nach x auf und erhält $x = \frac{1}{3}t^2 - \frac{1}{3}$. Ableiten nach t ergibt:

$$\frac{dx}{dt} = \frac{2}{3}t.$$

Man rechnet jetzt mit den Symbolen dx und dt so, als wären es Zahlen („formales Rechnen“) und erhält:

$$dx = \frac{2}{3}t \, dt.$$

Nun weiß man, wodurch dx zu ersetzen ist; man erhält:

$$\int e^{\sqrt{3x+1}} dx = \int e^t \frac{2}{3} t dt = \frac{2}{3} \int e^t t dt.$$

Das Integral $\int e^t t dt$ kann man leicht durch partielle Integration lösen:

$$\begin{aligned} \int e^t t dt &= e^t t - \int e^t dt \\ &= e^t t - e^t \\ &= e^t (t - 1) \end{aligned}$$

Es folgt:

$$\begin{aligned} \int e^{\sqrt{3x+1}} dx &= \frac{2}{3} \int e^t t dt \\ &= \frac{2}{3} e^t (t - 1) \\ &\stackrel{\text{Resubstitution}}{=} \frac{2}{3} e^{\sqrt{3x+1}} (\sqrt{3x+1} - 1) \end{aligned}$$

Empfehlung: Machen Sie die Probe!

Wie bereits erwähnt, verwendet man im 2. Schritt häufig auch eine abgekürzte Variante, die auf der Umkehrregel basiert.

Abgekürztes Verfahren.

1. *Schritt:* Man setzt $t = \sqrt{3x+1}$.

2. *Schritt:* Ersetzen von dx (*abgekürzte Variante*).

Man löst die Gleichung $t = \sqrt{3x+1}$ gar nicht erst nach x auf, sondern man leitet gleich ab (nämlich nach x); man erhält:

$$\frac{dt}{dx} = \frac{3}{2\sqrt{3x+1}}.$$

Durch formales Rechnen folgt:

$$dx = \frac{2}{3} \sqrt{3x+1} dt \stackrel{(*)}{=} \frac{2}{3} t dt.$$

Danach geht es weiter wie zuvor:

$$\int e^{\sqrt{3x+1}} dx = \frac{2}{3} \int e^t t dt \stackrel{(**)}{=} \dots$$

Wichtig: Damit es an der mit (***) gekennzeichneten Stelle wirklich weiter geht, muss man darauf achten, dass im Integranden die Variable x nicht mehr vorkommt, sondern nur noch t . In der obigen Rechnung wurde an der Stelle (*) dafür gesorgt. Auch beim nicht abgekürzten Verfahren muss auf Entsprechendes geachtet werden.

Abschließend stellen wir die besprochenen Regeln zur Berechnung von Stammfunktionen noch einmal in Kurzform zusammen.

Allgemeine Integrationsregeln.

- (I) $\int (f(x) + g(x)) dx = \int f(x) dx + \int g(x) dx$
(II) $\int (f(x) - g(x)) dx = \int f(x) dx - \int g(x) dx$
(III) $\int c \cdot f(x) dx = c \int f(x) dx$
(IV) $\int f'(x)g(x) dx = f(x)g(x) - \int f(x)g'(x) dx$ (Regel der partiellen Integration)
(V) $\int f(g(x))g'(x) dx = F(g(x))$ (für jede Stammfunktion F von f)
(VI) $\int \frac{g'(x)}{g(x)} dx = \ln |g(x)|$ (Spezialfall von (V))
(VII) $\int f(x) dx = \left[\int f(g(t))g'(t) dt \right]_{t=g^{-1}(x)}$ (Substitutionsregel)

3.4.3 Weitere Beispiele zur Substitutionsregel

Wir wollen die Anwendung der Substitutionsregel anhand zusätzlicher Beispiele studieren.

Beispiel 1. Zu berechnen sei $\int x^{\frac{3}{2}} e^{\sqrt{x}} dx$.

Wir setzen $t = \sqrt{x}$. Man erhält $x = t^2$, $\frac{dx}{dt} = 2t$ und folglich $dx = 2t dt$. Nimmt man in unserem Integranden die Ersetzungen $\sqrt{x} = t$ und $dx = 2t dt$ vor, so erhält man:

$$\int x^{\frac{3}{2}} e^{\sqrt{x}} dx = \int x^{\frac{3}{2}} e^t 2t dt.$$

Damit ist der Ersetzungsprozess aber noch nicht abgeschlossen, denn die Variable x darf ja nicht mehr vorkommen. Aus $t = \sqrt{x} = x^{\frac{1}{2}}$ erhält man $x^{\frac{3}{2}} = t^3$, also:

$$\int x^{\frac{3}{2}} e^{\sqrt{x}} dx = \int x^{\frac{3}{2}} e^t 2t dt = 2 \int t^4 e^t dt.$$

Das rechtsstehende Integral kann (mit etwas Rechenaufwand) ohne Weiteres ausgerechnet werden. (Nämlich wie?) Am Schluss hat man dann noch die Resubstitution vorzunehmen. (Die Details seien dem Leser überlassen.)

Zur Übung: Berechnen Sie $\int \frac{x+2}{\sqrt{x+1}} e^{\sqrt{x+1}} dx$.

Beispiel 2. Zu berechnen sei $\int \sin(\sqrt{x}) dx$.

Wir setzen $t = \sqrt{x}$. Es ergibt sich $x = t^2$, $\frac{dx}{dt} = 2t$ sowie $dx = 2t dt$. Also:

$$\int \sin(\sqrt{x}) dx = \int \sin t \cdot 2t dt = 2 \int t \cdot \sin t dt.$$

Das Integral $\int t \cdot \sin t dt$ erledigt man mit partieller Integration:

$$\begin{aligned} \int t \cdot \sin t dt &= t \cdot (-\cos t) - \int -\cos t dt \\ &= -t \cdot \cos t + \sin t \end{aligned}$$

Es folgt (Resubstitution):

$$\int \sin(\sqrt{x}) \, dx = 2 \left(\sin(\sqrt{x}) - \sqrt{x} \cdot \cos(\sqrt{x}) \right).$$

Beispiel 3. Zu berechnen sei $\int \sin(\ln x) \, dx$.

Wir setzen $t = \ln x$. Es ergibt sich $x = e^t$, $\frac{dx}{dt} = e^t$ und $dx = e^t \, dt$. Also:

$$\int \sin(\ln x) \, dx = \int \sin t \cdot e^t \, dt$$

Dieses Integral lässt sich mit zweifacher partieller Integration erledigen, nämlich so:

$$\begin{aligned} \int \sin t \cdot e^t \, dt &= \sin t \cdot e^t - \int \cos t \cdot e^t \, dt \\ &= \sin t \cdot e^t - \left(\cos t \cdot e^t + \int \sin t \cdot e^t \, dt \right) \end{aligned}$$

Es folgt:

$$\begin{aligned} \int \sin t \cdot e^t \, dt &= \frac{1}{2} (\sin t \cdot e^t - \cos t \cdot e^t) \\ &= \frac{1}{2} e^t (\sin t - \cos t) \end{aligned}$$

Der „Trick“ dieser zweifachen partiellen Integration besteht darin, dass auf der rechten Seite nach zweimaliger Anwendung der partiellen Integration das zu berechnende Integral $\int \sin t \cdot e^t \, dt$ wieder auftaucht (und zwar *mit negativem Vorzeichen*), wodurch man die Gleichung nach $\int \sin t \cdot e^t \, dt$ auflösen kann.

Es folgt nach Resubstitution ($t = \ln x$):

$$\int \sin(\ln x) \, dx = \frac{1}{2} x (\sin(\ln x) - \cos(\ln x)).$$

Machen Sie die Probe!

Bei Integranden, die aus $\sin x$ und $\cos x$ zusammengesetzt sind, kann man häufig mit Erfolg eine der Substitutionen $t = \tan x$ bzw. $t = \tan\left(\frac{x}{2}\right)$ anwenden. Wir illustrieren dies anhand des folgenden Beispiels.

Beispiel 4. Zu berechnen sei⁷ $\int \frac{1}{\cos^4 x} \, dx$.

Wegen $\frac{1}{\cos^2 x} = 1 + \tan^2 x$ gilt:

$$\int \frac{1}{\cos^4 x} \, dx = \int (1 + \tan^2 x)^2 \, dx.$$

Wir substituieren $t = \tan x$. Es ergibt sich $x = \arctan t$, $\frac{dx}{dt} = \frac{1}{1+t^2}$ und somit $dx = \frac{1}{1+t^2} \, dt$. Es folgt

$$\begin{aligned} \int (1 + \tan^2 x)^2 \, dx &= \left[\int (1 + t^2)^2 \frac{1}{1+t^2} \, dt \right]_{t=\tan x} \\ &= \left[\int (1 + t^2) \, dt \right]_{t=\tan x} \\ &= \left[t + \frac{1}{3} t^3 \right]_{t=\tan x} \\ &= \tan x + \frac{1}{3} \tan^3 x \end{aligned}$$

⁷Statt $\int \frac{1}{\cos^4 x} \, dx$ kann man auch $\int \frac{dx}{\cos^4 x}$ schreiben.

Also gilt:

$$\int \frac{1}{\cos^4 x} dx = \tan x + \frac{1}{3} \tan^3 x.$$

Beispiel 5. Zu berechnen sei $\int \arcsin x dx$.

Substitution: $t = \arcsin x$. Es folgt $x = \sin t$ sowie $dx = \cos t dt$. Also gilt:

$$\int \arcsin x dx = \int t \cdot \cos t dt.$$

Das Integral $\int t \cdot \cos t dt$ behandelt man mit partieller Integration und findet:

$$\begin{aligned} \int t \cdot \cos t dt &= t \cdot \sin t - \int \sin t dt \\ &= t \cdot \sin t + \cos t \\ &= t \cdot \sin t + \sqrt{1 - \sin^2 t} \end{aligned}$$

Es folgt durch Einsetzen von $t = \arcsin x$:

$$\int \arcsin x dx = x \cdot \arcsin x + \sqrt{1 - x^2}.$$

Ebenso wie in Beispiel 5 kann man allgemein in folgender Situation vorgehen: Für eine *injektive* Funktion f mit stetiger Ableitung sei $\int f(x) dx$ bekannt. Gesucht ist $\int f^{-1}(x) dx$.

Lösung. Man substituiert $t = f^{-1}(x)$; also $x = f(t)$, $\frac{dx}{dt} = f'(t)$ und $dx = f'(t) dt$. Es folgt:

$$\int f^{-1}(x) dx = \left[\int t f'(t) dt \right]_{t=f^{-1}(x)}$$

$\int t f'(t) dt$ löst man mit partieller Integration:

$$\int t \cdot f'(t) dt = t \cdot f(t) - \int f(t) dt.$$

Man beachte, dass wir oben $\int f(t) dt$ als bekannt angenommen haben; es sei etwa $\int f(t) dt = F(t)$. Insgesamt ergibt sich also (nach Einsetzen von $t = f^{-1}(x)$):

$$\int f^{-1}(x) dx = x \cdot f^{-1}(x) - F(f^{-1}(x)). \quad (3.26)$$

Besonders einfache Substitutionen sind *lineare Substitutionen*, d.h. Substitutionen vom Typ $t = ax + b$. Hierzu ein Beispiel.

Beispiel 6. Zu berechnen sei $\int \cos(3x + 1) dx$.

Wir hatten dieses Integral bereits in Abschnitt 3.4.2 durch Anwendung von Regel (3.21) berechnet. Alternativ wäre es auch mit der linearen Substitution $t = 3x + 1$ gegangen. Aus $t = 3x + 1$ erhält man $dx = \frac{1}{3} dt$. Es folgt:

$$\begin{aligned} \int \cos(3x + 1) dx &= \int \frac{1}{3} \cos t dt \\ &= \frac{1}{3} \sin t \\ &= \frac{1}{3} \sin(3x + 1) \end{aligned}$$

Wir haben gesehen, dass Substitution (eventuell mit anschließender partieller Integration) in gewissen Fällen weiterhelfen kann - häufig hilft dieses Verfahren aber leider auch nicht. Damit nicht der (falsche) Eindruck entsteht, Substitution mit anschließender partieller Integration sei ein „Allheilmittel“, betrachten wir noch das folgende Beispiel.

Beispiel 7. Zu berechnen sei $\int e^{x^2} dx$.

Wir versuchen die Substitution $t = x^2$ und erhalten $x = \sqrt{t}$, $\frac{dx}{dt} = \frac{1}{2\sqrt{t}}$ und $dx = \frac{1}{2\sqrt{t}} dt$. Es folgt:

$$\int e^{x^2} dx = \int e^t \cdot \frac{1}{2\sqrt{t}} dt.$$

Mit partieller Integration kommt man hier nicht weiter. (Es konnte nicht klappen, da man beweisen kann, dass das vorliegende Integral „nicht elementar integrierbar“ ist (vgl. auch Abschnitt 3.6)).

Als einen Spezialfall der Regel (V) haben wir die Regel

$$\int \frac{g'(x)}{g(x)} dx = \ln |g(x)|$$

erhalten (vgl. Übersicht am Ende von Abschnitt 3.4.2). Im nächsten Beispiel betrachten wir das folgende Integral:

$$\int g(x) \cdot g'(x) dx.$$

Wir werden sehen, dass auch hier ein Spezialfall von (V) vorliegt.

Beispiel 8. Integriert werden soll ein Produkt $g(x) \cdot g'(x)$.

Beispiele hierfür sind:

$$\sin x \cdot \cos x, \quad \frac{\ln x}{x} = \ln x \cdot \frac{1}{x} \quad \text{oder} \quad \frac{\arctan x}{x^2 + 1} = \arctan x \cdot \frac{1}{x^2 + 1}.$$

Lösung. Man wählt f als die identische Abbildung (d.h. $f(x) = x$) und wendet (V) an:

$$\int f(g(x))g'(x) dx = F(g(x)).$$

Da $F(x) = \frac{1}{2}x^2$ gilt, haben wir also folgende Formel erhalten (die man leicht durch Ableiten bestätigt):

$$\int g(x) \cdot g'(x) dx = \frac{1}{2}(g(x))^2. \tag{3.27}$$

Von der Regel der partiellen Integration und der Substitutionsregel gibt es auch *Varianten für bestimmte Integrale*, auf die hier noch kurz eingegangen werden soll.

Satz 10' (Partielle Integration für bestimmte Integrale).

f und g seien differenzierbar mit stetigen Ableitungen f' und g' (auf $[a, b]$). Dann gilt:

$$\int_a^b f'(x) \cdot g(x) dx = [f(x) \cdot g(x)]_a^b - \int_a^b f(x) \cdot g'(x) dx.$$

Beweis. Dies ergibt sich aus der Produktregel für Ableitungen. Es gilt:

$$\begin{aligned}(f \cdot g)'(x) &= f'(x) \cdot g(x) + f(x) \cdot g'(x) \\ \Rightarrow \int_a^b (f \cdot g)'(x) \, dx &= \int_a^b (f'(x) \cdot g(x) + f(x) \cdot g'(x)) \, dx \\ &= \int_a^b f'(x) \cdot g(x) \, dx + \int_a^b f(x) \cdot g'(x) \, dx\end{aligned}$$

Außerdem gilt:

$$\int_a^b (f \cdot g)'(x) \, dx = [f(x) \cdot g(x)]_a^b.$$

Also insgesamt:

$$[f(x) \cdot g(x)]_a^b = \int_a^b f'(x) \cdot g(x) \, dx + \int_a^b f(x) \cdot g'(x) \, dx.$$

Daraus folgt Satz 10'. \square

Beispiel 9. Zu berechnen sei $\int_0^\pi x \cdot \cos x \, dx$.

Hat man das bestimmte Integral $\int_0^\pi x \cdot \cos x \, dx$ zu berechnen, gibt es 2 Möglichkeiten.

- (i) Man bestimmt zunächst eine Stammfunktion mithilfe von Satz 10. Das hatten wir in Anschluss an Satz 10 bereits durchgeführt, wobei wir

$$\int x \cdot \cos x \, dx = x \cdot \sin x + \cos x$$

gefunden hatten. Anschließendes Einsetzen der Grenzen ergibt:

$$\begin{aligned}\int_0^\pi x \cdot \cos x \, dx &= [x \cdot \sin x + \cos x]_0^\pi \\ &= \pi \cdot \sin \pi + \cos \pi - (0 \cdot \sin 0 + \cos 0) \\ &= 0 - 1 - (0 + 1) \\ &= -2\end{aligned}$$

- (ii) Die zweite Möglichkeit ist, anstelle von Satz 10 den Satz 10' zu verwenden, was im Wesentlichen auf dieselbe Rechnung hinausläuft:

$$\begin{aligned}\int_0^\pi x \cdot \cos x \, dx &= [x \cdot \sin x]_0^\pi - \int_0^\pi \sin x \, dx \\ &= [x \cdot \sin x]_0^\pi - [-\cos x]_0^\pi \\ &= \pi \cdot \sin \pi - 0 \cdot \sin 0 + \cos \pi - \cos 0 \\ &= 0 - 0 - 1 - 1 \\ &= -2\end{aligned}$$

Satz 11' (Substitutionsregel für bestimmte Integrale).

Es sei $g : [a, b] \rightarrow \mathbb{R}$ differenzierbar und g' sei auf $[a, b]$ stetig; f sei auf dem Bild $g([a, b])$ von g definiert und stetig. Dann gilt:

$$\int_a^b f(g(t)) \cdot g'(t) dt = \int_{g(a)}^{g(b)} f(x) dx. \quad (3.28)$$

Beweis. Dies ergibt sich direkt aus der Kettenregel. Ist nämlich F auf $g([a, b])$ eine Stammfunktion von f , so ist $F(g(t))$ nach der Kettenregel eine Stammfunktion von $f(g(t)) \cdot g'(t)$. Es folgt:

$$\int_a^b f(g(t)) \cdot g'(t) dt = \left[F(g(t)) \right]_a^b = F(g(b)) - F(g(a))$$

und

$$\int_{g(a)}^{g(b)} f(x) dx = \left[F(x) \right]_{g(a)}^{g(b)} = F(g(b)) - F(g(a)). \quad \square$$

Man kann die Regel (3.28) „von links nach rechts“ sowie „von rechts nach links“ anwenden. Die Anwendung „von links nach rechts“ unterscheidet sich nur unwesentlich von der Verwendung der Regel (3.21) für unbestimmte Integrale, weshalb wir nicht näher darauf eingehen wollen. Wir beschreiben im Folgenden die Anwendung „von rechts nach links“ anhand eines Beispiels.

Beispiel 10. Zu berechnen sei $\int_2^3 e^{\sqrt{4x+1}} dx$.

Analog zu Beispiel 9 hat man auch hier 2 Möglichkeiten:

- (i) Man geht wie bisher vor und berechnet zunächst $\int e^{\sqrt{4x+1}} dx$ mit der Substitutionsregel für *unbestimmte Integrale* (Satz 11). Man erhält $t = \sqrt{4x+1}$, $x = \frac{1}{4}t^2 - \frac{1}{4}$, $\frac{dx}{dt} = \frac{t}{2}$ und $dx = \frac{1}{2}t dt$. Es folgt unter Verwendung von partieller Integration:

$$\begin{aligned} \int e^{\sqrt{4x+1}} dx &= \int e^t \cdot \frac{1}{2}t dt \\ &= \frac{1}{2} \int e^t \cdot t dt \\ &= \frac{1}{2} e^t (t - 1) \\ &\stackrel{(*)}{=} \frac{1}{2} e^{\sqrt{4x+1}} (\sqrt{4x+1} - 1) \end{aligned}$$

An der Stelle (*) wurde resubstituiert. (Man hat jetzt Gelegenheit, durch Ableiten eine Probe zu machen!) Einsetzen der Grenzen in die erhaltene Stammfunktion ergibt

$$\begin{aligned} \int_2^3 e^{\sqrt{4x+1}} dx &= \left[\frac{1}{2} e^{\sqrt{4x+1}} (\sqrt{4x+1} - 1) \right]_2^3 \\ &= \frac{1}{2} \left(e^{\sqrt{13}} (\sqrt{13} - 1) - 2e^3 \right) \\ &\approx 27.86 \end{aligned}$$

- (ii) Man verwendet Satz 11', wobei die Substitution $t = \sqrt{4x+1}$, $dx = \frac{1}{2}t \, dt$ dieselbe ist wie bei Möglichkeit (i). Es folgt nach Satz 11', indem man (3.28) von rechts nach links anwendet:

$$\int_2^3 e^{\sqrt{4x+1}} \, dx = \int_3^{\sqrt{13}} e^t \cdot \frac{1}{2}t \, dt$$

Die neuen Grenzen 3 und $\sqrt{13}$ ergeben sich durch Einsetzen der alten Grenzen 2 und 3 in die Substitutionsgleichung $t = \sqrt{4x+1}$:

- Für $x = 2$ erhält man $t = 3$.
- Für $x = 3$ erhält man $t = \sqrt{13}$.

Unter Verwendung von Satz 10' erhält man

$$\begin{aligned} \int_2^3 e^{\sqrt{4x+1}} \, dx &= \frac{1}{2} \int_3^{\sqrt{13}} e^t \cdot t \, dt \\ &= \frac{1}{2} \left(\left[e^t \cdot t \right]_3^{\sqrt{13}} - \int_3^{\sqrt{13}} e^t \, dt \right) \\ &= \frac{1}{2} \left(\left[e^t \cdot t \right]_3^{\sqrt{13}} - \left[e^t \right]_3^{\sqrt{13}} \right) \\ &= \frac{1}{2} \left(e^{\sqrt{13}} \cdot \sqrt{13} - 3e^3 - e^{\sqrt{13}} + e^3 \right) \\ &= \frac{1}{2} \left(e^{\sqrt{13}} \left(\sqrt{13} - 1 \right) - 2e^3 \right) \\ &\approx 27.86 \end{aligned}$$

Bei beiden Möglichkeiten hat man sehr ähnliche Rechnungen durchzuführen.

Einzigster Unterschied: Bei der Möglichkeit (ii) wird keine Resubstitution vorgenommen, sondern eine *Anpassung der Integrationsgrenzen*.

Ein *Vorteil von Möglichkeit (i)*: Nach der Resubstitution ergibt sich die Gelegenheit, durch Ableiten eine *Probe* zu machen.

3.4.4 Integration rationaler Funktionen

Wir beschreiben in diesem Abschnitt eine Methode, mit der man rationale Funktionen $R(x)$ systematisch integrieren kann. Zur Erinnerung: Eine rationale Funktion ist ein Quotient zweier Polynome:

$$R(x) = \frac{f(x)}{g(x)}. \quad (3.29)$$

(Beispielsweise $R(x) = \frac{x^3-2x+1}{x^2+x+1}$ oder $R(x) = \frac{1}{x^2-a^2}$).

Schauen wir in unsere Tabelle der Grundintegrale, so stellen wir fest, dass wir die beiden rationalen Funktionen $\frac{1}{x}$ und $\frac{1}{x^2+1}$ bereits integrieren können. Für diese Funktionen gelten die grundlegenden und besonders wichtigen Formeln:

$$\int \frac{1}{x} \, dx = \ln |x| \quad \text{und} \quad \int \frac{1}{x^2+1} \, dx = \arctan x.$$

Mithilfe der Formel $\int \frac{1}{x^2+1} dx = \arctan x$ kann man weitere Integrale berechnen, wie beispielsweise $\int \frac{1}{x^2+5} dx$ oder $\int \frac{1}{x^2+2x+7} dx$. Wir schauen uns im Folgenden an, wie man derartige Integrale auf das Grundintegral $\int \frac{1}{x^2+1} dx$ zurückführt.

(i) Zu berechnen: $\int \frac{1}{x^2+a^2} dx$ (für $a \neq 0$).

Es gilt:

$$\begin{aligned} \int \frac{1}{x^2+a^2} dx &= \int \frac{1}{a^2 \left(\frac{x^2}{a^2} + 1\right)} dx \\ &= \frac{1}{a^2} \int \frac{1}{\left(\frac{x}{a}\right)^2 + 1} dx \end{aligned}$$

Wir verwenden die Substitution $t = \frac{x}{a}$. Also gilt $\frac{dt}{dx} = \frac{1}{a}$ und $dx = a dt$. Es folgt:

$$\begin{aligned} \int \frac{1}{x^2+a^2} dx &= \left[\frac{1}{a^2} \int \frac{1}{t^2+1} a dt \right]_{t=\frac{x}{a}} \\ &= \left[\frac{1}{a} \int \frac{1}{t^2+1} dt \right]_{t=\frac{x}{a}} \\ &= \left[\frac{1}{a} \arctan t \right]_{t=\frac{x}{a}} \\ &= \frac{1}{a} \arctan \left(\frac{x}{a} \right) \end{aligned}$$

(ii) Zu berechnen $\int \frac{1}{x^2+ax+b} dx$.

Auch dieses Integral lässt sich auf bereits bekannte Integrale zurückführen. Hierzu formen wir den Nenner nach der Methode der quadratischen Ergänzung um:

$$\begin{aligned} x^2+ax+b &= x^2+ax+\frac{a^2}{4}+b-\frac{a^2}{4} \\ &= \left(x+\frac{a}{2}\right)^2+b-\frac{a^2}{4} \end{aligned}$$

Wir setzen $c = b - \frac{a^2}{4}$. Substitution: $t = x + \frac{a}{2}$, $\frac{dt}{dx} = 1$ und $dx = dt$. Also gilt:

$$\begin{aligned} \int \frac{1}{x^2+ax+b} dx &= \int \frac{1}{\left(x+\frac{a}{2}\right)^2+c} dx \\ &= \int \frac{1}{t^2+c} dt \end{aligned}$$

Wir unterscheiden die Fälle $c > 0$, $c = 0$ und $c < 0$. (Man beachte: $c > 0$ gilt genau dann, wenn das Nennerpolynom x^2+ax+b keine reelle Nullstelle hat. Entsprechend gilt: $c = 0 \Leftrightarrow x^2+ax+b$ hat genau eine reelle Nullstelle („doppelte Nullstelle“); $c < 0 \Leftrightarrow x^2+ax+b$ hat zwei verschiedene reelle Nullstellen.)

1. Fall: $c > 0$

Wir setzen $r = \sqrt{c}$. Dann ist $c = r^2$. Es ergibt sich nach (i):

$$\begin{aligned}\int \frac{1}{x^2 + ax + b} dx &= \int \frac{1}{t^2 + r^2} dt \\ &= \frac{1}{r} \cdot \arctan\left(\frac{t}{r}\right) \\ &= \frac{1}{\sqrt{c}} \cdot \arctan\left(\frac{t}{\sqrt{c}}\right)\end{aligned}$$

Wir haben nun noch $t = x + \frac{a}{2}$ einzusetzen (Resubstitution). Es folgt:

$$\int \frac{1}{x^2 + ax + b} dx = \frac{1}{\sqrt{c}} \arctan\left(\frac{x + \frac{a}{2}}{\sqrt{c}}\right)$$

2. Fall: $c = 0$

$$\int \frac{1}{t^2} dt = \int t^{-2} dt = -t^{-1} \implies \int \frac{1}{x^2 + ax + b} dx = \frac{-1}{x + \frac{a}{2}}.$$

3. Fall $c < 0$

In diesem Fall hat $t^2 + c$ zwei reelle Nullstellen. Wir werden ihn in Kürze erledigen.

Wir besprechen nun ein systematisches Verfahren zur Integration rationaler Funktionen. Gegeben sei eine rationale Funktion wie in (3.29), d.h. $R(x) = \frac{f(x)}{g(x)}$.

1. Schritt: Division mit Rest

Hat das Nennerpolynom $g(x)$ einen höheren Grad als $f(x)$, so entfällt der erste Schritt. Andernfalls führe man die Polynomdivision $f(x) : g(x)$ durch (Beispiel s.u.). Man erhält dann Polynome $Q(x)$, $r(x)$, für die gilt:

$$R(x) = \frac{f(x)}{g(x)} = Q(x) + \frac{r(x)}{g(x)}. \quad (3.30)$$

Der Grad von $r(x)$ ist dabei kleiner als der Grad von $g(x)$. Es folgt:

$$\int R(x) dx = \int Q(x) dx + \int \frac{r(x)}{g(x)} dx. \quad (3.31)$$

$\int Q(x) dx$ ist leicht zu bestimmen, da $Q(x)$ ein Polynom ist. In dem zweiten Integral steht im Zähler des Integranden ein Polynom (nämlich $r(x)$), dessen Grad kleiner ist als der Grad des Polynoms $g(x)$ im Nenner. Wir haben das ursprüngliche Integral also so umgeformt, dass wir uns nur noch um die Integration einer rationalen Funktion zu kümmern brauchen, bei der der Grad des Nenners größer ist als der Grad des Zähler.

Beispiel zur Polynomdivision.

$$R(x) = \frac{x^3 - 2x + 1}{x^2 + x + 1} = \frac{f(x)}{g(x)}$$

$$\begin{array}{r} (x^3 \quad \quad \quad - \quad 2x \quad + \quad 1) : (x^2 + x + 1) = x - 1 \quad \text{Rest: } -2x + 2 \\ -(x^3 \quad + \quad \quad x^2 \quad + \quad \quad x) \\ \hline \quad \quad \quad -x^2 \quad - \quad 3x \quad + \quad 1 \\ \quad \quad - \quad (-x^2 \quad - \quad \quad x \quad - \quad 1) \\ \hline \quad \quad \quad \quad \quad \quad -2x \quad + \quad 2\end{array}$$

Es folgt:

$$\begin{aligned} R(x) &= Q(x) + \frac{r(x)}{g(x)} \\ &= x - 1 + \frac{-2x + 2}{x^2 + x + 1} \\ \implies \int R(x) dx &= \int (x - 1) dx + \int \frac{-2x + 2}{x^2 + x + 1} dx \end{aligned}$$

Es gilt:

$$\int (x - 1) dx = \frac{1}{2}x^2 - x$$

Das Integral $\int \frac{-2x+2}{x^2+x+1}$ bleibt zu bestimmen.

2. Schritt: Faktorzerlegung des Nenners

Es sei $g(x)$ ein Polynom n -ten Grades (mit reellen Koeffizienten). Ein $a \in \mathbb{R}$ heißt *reelle Nullstelle* von $g(x)$, falls $g(a) = 0$ gilt. Mit a_1, \dots, a_r seien (sofern vorhanden) die reellen Nullstellen von $g(x)$ bezeichnet, α_i bezeichne die Vielfachheit der Nullstelle a_i ($i = 1, \dots, r$). Dann lässt sich $g(x)$ wie folgt als Produkt von Linearfaktoren und quadratischen Polynomen darstellen:

$$g(x) = c \cdot (x - a_1)^{\alpha_1} \cdot (x - a_2)^{\alpha_2} \cdot \dots \cdot (x - a_r)^{\alpha_r} \cdot (Q_1(x))^{\beta_1} \cdot \dots \cdot (Q_s(x))^{\beta_s}. \quad (3.32)$$

Hierin sind $Q_1(x), \dots, Q_s(x)$ verschiedene quadratische Polynome mit reellen Koeffizienten, aber **ohne reelle Nullstellen**; β_j ist die Vielfachheit, mit der $Q_j(x)$ als Faktor von $g(x)$ auftritt. (Dass eine solche Darstellung von $g(x)$ immer möglich ist, wird in einem späteren Kapitel besprochen werden, nachdem wir komplexe Zahlen und den *Fundamentalsatz der Algebra* kennengelernt haben.)

Beispiele:

- $g(x) = x^3 - x^2 - x + 1 = (x - 1)(x - 1)(x + 1) = (x - 1)^2(x + 1)$
- $g(x) = 2(x - 3)^2(x + 1)(x - 5)^4(x^2 + x + 1)^3(x^2 - x + 2)$

Wir werden im Folgenden der Einfachheit halber annehmen, dass

$$\begin{aligned} g(x) &= c \cdot (x - a_1)^{\alpha_1} \cdot (x - a_2)^{\alpha_2} \cdot \dots \cdot (x - a_r)^{\alpha_r} \quad \text{oder} \\ g(x) &= c \cdot (x - a_1)^{\alpha_1} \cdot (x - a_2)^{\alpha_2} \cdot \dots \cdot (x - a_r)^{\alpha_r} \cdot Q(x) \end{aligned} \quad (3.33)$$

gilt, d.h., es kommen nur lineare Faktoren oder nur genau ein quadratischer Faktor vor; dies sind die beiden wichtigsten Fälle. Wir betonen jedoch, dass die beschriebenen Methoden in ähnlicher Weise auch in dem Fall funktionieren, dass mehr als ein quadratischer Faktor vorkommt; Näheres findet man in den einschlägigen Lehrbüchern der Analysis, z.B. in H. Heuser, Analysis I.

Beispiel. $g(x) = x^3 - 1$.

Eine reelle Nullstelle von $g(x)$ ist 1. Also ist $x^3 - 1$ durch $x - 1$ teilbar. Durch Polynomdivision erhalten wir:

$$g(x) = x^3 - 1 = (x - 1)(x^2 + x + 1).$$

Der quadratische Faktor $x^2 + x + 1$ hat keine reellen Nullstellen; er kann im Reellen nicht weiter zerlegt werden.

3. Schritt: Partialbruchzerlegung

Für $R(x) = \frac{f(x)}{g(x)}$ gelte: Der Grad von $f(x)$ ist kleiner als der Grad von $g(x)$. Das Nennerpolynom liege

in der Darstellung (3.33) vor. Dann lässt sich $R(x)$ folgendermaßen darstellen (*Partialbruchzerlegung*):

$$\begin{aligned} \frac{f(x)}{g(x)} &= \frac{A_{11}}{(x-a_1)^{\alpha_1}} + \frac{A_{12}}{(x-a_1)^{\alpha_1-1}} + \dots + \frac{A_{1\alpha_1}}{(x-a_1)} \\ &+ \frac{A_{21}}{(x-a_2)^{\alpha_2}} + \frac{A_{22}}{(x-a_2)^{\alpha_2-1}} + \dots + \frac{A_{2\alpha_2}}{(x-a_2)} \\ &\vdots \\ &+ \frac{A_{r1}}{(x-a_r)^{\alpha_r}} + \frac{A_{r2}}{(x-a_r)^{\alpha_r-1}} + \dots + \frac{A_{r\alpha_r}}{(x-a_r)} \\ &+ \frac{px+q}{Q(x)} \end{aligned}$$

Der Partialbruch $\frac{px+q}{Q(x)}$ tritt nur dann auf, wenn in (3.33) ein quadratischer Faktor $Q(x)$ vorkommt⁸. (Dabei sind die Koeffizienten A_{ij} sowie p und q reelle Zahlen.)

Es gibt zahlreiche Methoden zur Berechnung der Koeffizienten A_{ij} , p und q der Partialbruchzerlegung. Wir stellen hier anhand von Beispielen die **Methode des Koeffizientenvergleichs** vor.

Beispiel 1. Zu berechnen sei $\int \frac{x+1}{x^2-5x+6} dx$.

Der 1. Schritt entfällt, da das Zählerpolynom einen kleineren Grad als das Nennerpolynom besitzt. Die Nullstellen des Nenners sind $a_1 = 2$ und $a_2 = 3$; den Integranden können wir also wie folgt aufschreiben:

$$\frac{x+1}{x^2-5x+6} = \frac{x+1}{(x-2)(x-3)} = \frac{A}{x-2} + \frac{B}{x-3}.$$

Der rechts stehende Ausdruck gibt an, wie die gesuchte Partialbruchzerlegung auszusehen hat. (A und B sind noch zu bestimmen; oben hatten wir eine etwas andere Bezeichnung benutzt: Was hier A und B heißt, wurde oben A_{11} und A_{21} genannt.)

Bei der Bestimmung von A und B nach der *Methode des Koeffizientenvergleichs* geht man wie folgt vor: Die rechte Seite der Gleichung

$$\frac{x+1}{x^2-5x+6} = \frac{A}{x-2} + \frac{B}{x-3}$$

„bringt man auf den Nenner x^2-5x+6 “ und erhält somit:

$$\frac{x+1}{x^2-5x+6} = \frac{A(x-3) + B(x-2)}{x^2-5x+6} = \frac{(A+B)x - 3A - 2B}{x^2-5x+6}.$$

Vergleicht man nun die Koeffizienten der Zählerpolynome, so erhält man die beiden Gleichungen:

$$\begin{aligned} A+B &= 1 \\ -3A-2B &= 1 \end{aligned}$$

Als Lösung dieser beiden Gleichungen erhält man:

$$A = -3 \quad \text{und} \quad B = 4.$$

Wir haben somit die gesuchte *Partialbruchzerlegung* erhalten:

$$\frac{x+1}{x^2-5x+6} = \frac{-3}{x-2} + \frac{4}{x-3}.$$

⁸Im allgemeinen Fall (3.32) sieht die Partialbruchzerlegung ganz ähnlich aus; dann kommen noch Partialbrüche der Form $\frac{p_{ij}x + q_{ij}}{(Q_i(x))^j}$ hinzu ($p_{ij}, q_{ij} \in \mathbb{R}$). Beweise, dass eine Partialbruchzerlegung wie angegeben tatsächlich immer existiert, kann man in den im Literaturverzeichnis aufgeführten Lehrbüchern der Analysis finden.

Nun können wir das gesuchte Integral berechnen (vgl. Abschnitt 3.4, Formel (3.22) bzw. (VI)):

$$\begin{aligned} \int \frac{x+1}{x^2-5x+6} dx &= \int \frac{-3}{x-2} dx + \int \frac{4}{x-3} dx \\ &= (-3) \ln|x-2| + 4 \ln|x-3| \end{aligned}$$

Im obigen Beispiel 1 konnte man das Integral nach Herstellung der Partialbruchzerlegung ohne Weiteres berechnen. Wir befassen uns nun im 4. Schritt unseres Verfahrens mit der Integration nach erfolgter Partialbruchzerlegung und werden sehen, dass die Integration immer ausführbar ist.

4. Schritt: Integration

Hat man eine rationale Funktion $R(x)$ in Partialbrüche zerlegt, so kann man diese einzeln integrieren. Die Partialbrüche haben eine der folgenden drei Formen:

$$(1) \quad \frac{c}{x-a} \quad \text{oder} \quad (2) \quad \frac{c}{(x-a)^j} \quad (\text{für } j \in \mathbb{N}, j \geq 2) \quad \text{oder} \quad (3) \quad \frac{px+q}{Q(x)}.$$

Die Typen (1) und (2) können wir sofort integrieren (siehe Abschnitt 3.4, Formel (3.22) bzw. (VI)):

$$\int \frac{c}{x-a} dx = c \ln|x-a| \quad (3.34)$$

$$\int c(x-a)^{-j} dx = c \frac{(x-a)^{-j+1}}{-j+1} \quad (3.35)$$

Zu (3): Es gelte $Q(x) = x^2 + ax + b$. Wir schreiben (quadratische Ergänzung)

$$Q(x) = x^2 + ax + b = \left(x + \frac{a}{2}\right)^2 + b - \frac{a^2}{4}$$

und setzen $c = b - \frac{a^2}{4}$. Man erinnere sich, dass $Q(x)$ keine reellen Nullstellen hat (nur solche quadratischen Faktoren kommen in der Zerlegung (3.32) vor). Es gilt also $c > 0$. Den speziellen Fall, dass $px + q = 1$ gilt, haben wir schon zu Beginn des Abschnitts 3.4.4 erledigt; dort haben wir unter (ii) gefunden, dass Folgendes gilt:

$$\int \frac{1}{x^2 + ax + b} dx = \frac{1}{\sqrt{c}} \arctan\left(\frac{x + \frac{a}{2}}{\sqrt{c}}\right). \quad (3.36)$$

Einen anderen speziellen Fall können wir ebenfalls sofort erledigen: Ist $px + q = 2x + a$, so steht im Zähler die Ableitung des Nenners und es gilt nach (3.22) bzw. (VI) im Abschnitt 3.4:

$$\int \frac{2x+a}{x^2+ax+b} dx = \ln|x^2+ax+b|. \quad (3.37)$$

Wir zeigen nun, wie man durch eine einfache Umformung den allgemeinen Fall

$$\int \frac{px+q}{x^2+ax+b} dx$$

auf die beiden Spezialfälle (3.36) und (3.37) zurückführt. Wir formen hierzu $px + q$ so um, dass die Ableitung $2x + a$ des Nenners auftritt:

$$\begin{aligned}
\int \frac{px+q}{x^2+ax+b} dx &= \int \frac{\frac{p}{2}\left(2x+\frac{2q}{p}\right)}{x^2+ax+b} dx \\
&= \frac{p}{2} \int \frac{2x+a+\frac{2q}{p}-a}{x^2+ax+b} dx \\
&= \frac{p}{2} \int \frac{2x+a}{x^2+ax+b} dx + \left(q-\frac{pa}{2}\right) \int \frac{1}{x^2+ax+b} dx \\
&= \frac{p}{2} \ln|x^2+ax+b| + \left(q-\frac{pa}{2}\right) \frac{1}{\sqrt{c}} \arctan\left(\frac{x+\frac{a}{2}}{\sqrt{c}}\right) \tag{3.38}
\end{aligned}$$

Wir betrachten nun zwei weitere Beispiele.

Beispiel 2. Zu bestimmen sei $\int \frac{3x+5}{x^3-x^2+x-1} dx$.

Eine Nullstelle des Nenners ist 1; Polynomdivision ergibt $x^3-x^2+x-1 = (x-1)(x^2+1)$.

Partialbruchzerlegung:

$$\frac{3x+5}{x^3-x^2+x-1} = \frac{A}{x-1} + \frac{px+q}{x^2+1}.$$

Wir bestimmen A , p und q mit der Methode des Koeffizientenvergleichs. Es ergibt sich:

$$\begin{aligned}
\frac{3x+5}{x^3-x^2+x-1} &= \frac{A(x^2+1) + (px+q)(x-1)}{(x-1)(x^2+1)} \\
&= \frac{(A+p)x^2 + (q-p)x + A-q}{(x-1)(x^2+1)}
\end{aligned}$$

Koeffizientenvergleich ergibt:

$$\begin{aligned}
A+p &= 0 \\
q-p &= 3 \\
A-q &= 5
\end{aligned}$$

Lösung dieses Gleichungssystems:

$$A = 4, \quad p = -4 \quad \text{und} \quad q = -1.$$

Es folgt:

$$\begin{aligned}
\int \frac{3x+5}{x^3-x^2+x-1} dx &= \int \frac{4}{x-1} dx + \int \frac{-4x-1}{x^2+1} dx \\
&= 4 \ln|x-1| - 2 \int \frac{2x}{x^2+1} dx - \int \frac{1}{x^2+1} dx \\
&= 4 \ln|x-1| - 2 \ln|x^2+1| - \arctan x
\end{aligned}$$

Beispiel 3. Zu bestimmen sei $\int \frac{x^3-x^2-6x+1}{x^2-6x+9} dx$.

Zunächst ist eine Polynomdivision auszuführen:

$$\begin{array}{r}
(x^3 - x^2 - 6x + 1) : (x^2 - 6x + 9) = x + 5 \quad \text{Rest: } 15x - 44 \\
-(x^3 - 6x^2 + 9x) \\
\hline
5x^2 - 15x + 1 \\
-(5x^2 - 30x + 45) \\
\hline
15x - 44
\end{array}$$

Es folgt:

$$\frac{x^3 - x^2 - 6x + 1}{x^2 - 6x + 9} = x + 5 + \frac{15x - 44}{x^2 - 6x + 9}.$$

Bestimmung der Partialbruchzerlegung der Funktion $\frac{15x-44}{x^2-6x+9}$:

$$\begin{aligned}\frac{15x - 44}{x^2 - 6x + 9} &= \frac{15x - 44}{(x - 3)^2} \\ &= \frac{A}{(x - 3)^2} + \frac{B}{x - 3} \\ &= \frac{A + B(x - 3)}{(x - 3)^2} \\ &= \frac{Bx + A - 3B}{(x - 3)^2}\end{aligned}$$

Es resultiert das folgende Gleichungssystem:

$$\begin{aligned}B &= 15 \\ A - 3B &= -44\end{aligned}$$

Lösung dieses Gleichungssystems: $A = 1$, $B = 15$.

Integration von $\frac{15x-44}{x^2-6x+9}$ ergibt:

$$\begin{aligned}\int \frac{15x - 44}{x^2 - 6x + 9} dx &= \int \frac{1}{(x - 3)^2} dx + \int \frac{15}{x - 3} dx \\ &= -\frac{1}{x - 3} + 15 \ln |x - 3|\end{aligned}$$

Insgesamt erhält man also:

$$\int \frac{x^3 - x^2 - 6x + 1}{x^2 - 6x + 9} dx = \frac{1}{2}x^2 + 5x - \frac{1}{x - 3} + 15 \ln |x - 3|.$$

Wir haben oben im 4. Schritt vorgeführt, wie man ein Integral

$$\int \frac{px + q}{x^2 + ax + b} dx$$

berechnet, wenn der Nenner $x^2 + ax + b$ keine reellen Nullstellen besitzt (vgl. (3.36) - (3.38)). Es ist nützlich, sich hierzu noch konkrete **Beispiele** anzuschauen.

Beispiel 4. Zu bestimmen sei $\int \frac{x + 3}{x^2 + 1} dx$.

$$\begin{aligned}\int \frac{x + 3}{x^2 + 1} dx &= \int \frac{x}{x^2 + 1} dx + \int \frac{3}{x^2 + 1} dx \\ &= \frac{1}{2} \int \frac{2x}{x^2 + 1} dx + 3 \int \frac{1}{x^2 + 1} dx \\ &= \frac{1}{2} \ln |x^2 + 1| + 3 \arctan x\end{aligned}$$

Beispiel 5. Zu bestimmen sei $\int \frac{x+3}{x^2+2x+5} dx$.

$$\begin{aligned} \int \frac{x+3}{x^2+2x+5} dx &= \frac{1}{2} \int \frac{2x+6}{x^2+2x+5} dx \\ &= \frac{1}{2} \int \frac{2x+2+4}{x^2+2x+5} dx \\ &= \frac{1}{2} \int \frac{2x+2}{x^2+2x+5} dx + 2 \int \frac{1}{x^2+2x+5} dx \\ &= \frac{1}{2} \ln|x^2+2x+5| + \arctan\left(\frac{x+1}{2}\right) \end{aligned}$$

Der letzte Schritt gilt, da:

$$\begin{aligned} \int \frac{1}{x^2+2x+5} dx &= \int \frac{1}{x^2+2x+1+4} dx \\ &= \int \frac{1}{(x+1)^2+4} dx \\ &= \frac{1}{4} \int \frac{1}{\left(\frac{x+1}{2}\right)^2+1} dx \\ &\stackrel{(*)}{=} \frac{1}{4} \int \frac{1}{t^2+1} \cdot 2 dt \\ &= \frac{1}{2} \arctan t \\ &= \frac{1}{2} \arctan\left(\frac{x+1}{2}\right) \end{aligned}$$

An der mit (*) bezeichneten Stelle wurde die Substitutionsregel benutzt; und zwar folgendermaßen:

$$t = \frac{x+1}{2} \implies \frac{dt}{dx} = \frac{1}{2} \implies dx = 2 dt.$$

3.5 Interpolation und Interpolationsfehler

Unter einer Interpolationsaufgabe versteht man die folgende Fragestellung.

Gegeben: $n+1$ verschiedene reelle Zahlen x_0, x_1, \dots, x_n (genannt *Stützstellen*) und dazugehörige, nicht notwendigerweise verschiedene reelle Zahlen y_0, y_1, \dots, y_n (genannt *Stützwerte*).

Gesucht: Ein Polynom P vom Grad $\leq n$ mit $P(x_i) = y_i$ für $i = 0, 1, \dots, n$.

3.5.1 Lagrange-Interpolation

Wir zeigen, dass ein solches Polynom immer vorhanden ist: Hierzu betrachten wir zunächst den Spezialfall, dass für ein festes $k \in \{0, 1, \dots, n\}$ $y_k = 1$ und $y_j = 0$ für alle $j \neq k$ gilt. Dann wird die Interpolationsaufgabe offenbar durch folgendes Polynom $L_k(x)$ gelöst (Man überprüfe dies durch Einsetzen!):

$$L_k(x) = \frac{(x-x_0) \cdot \dots \cdot (x-x_{k-1}) \cdot (x-x_{k+1}) \cdot \dots \cdot (x-x_n)}{(x_k-x_0) \cdot \dots \cdot (x_k-x_{k-1}) \cdot (x_k-x_{k+1}) \cdot \dots \cdot (x_k-x_n)}. \quad (3.39)$$

Man nennt $L_k(x)$ das *k-te Lagrangesche Polynom n-ten Grades*.

Mithilfe der Polynome $L_k(x)$, $k = 0, 1, \dots, n$ lässt sich jetzt auf ganz einfache Art auch die Lösung $L(x)$ der Interpolationsaufgabe für den allgemeinen Fall beliebiger Stützwerte y_0, \dots, y_n angeben. Offenbar leistet nämlich das folgende Polynom $L(x)$ das Gewünschte⁹:

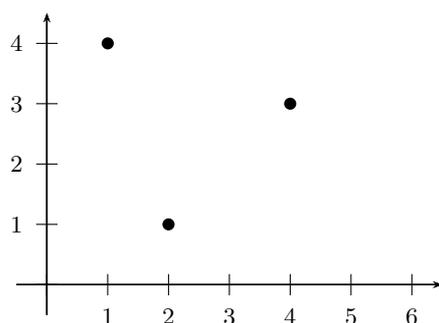
$$L(x) = y_0 L_0(x) + y_1 L_1(x) + \dots + y_n L_n(x). \quad (3.40)$$

Nach Konstruktion der Polynome $L_k(x)$ gilt ja gerade:

$$L_k(x_i) = \begin{cases} 1 & , \text{ falls } i = k \\ 0 & , \text{ falls } i \neq k \end{cases}$$

Beispiel. $x_0 = 1$, $x_1 = 2$, $x_2 = 4$, $y_0 = 4$, $y_1 = 1$ und $y_2 = 3$.

Gesucht ist ein Polynom $P(x)$ höchstens 2-ten Grades, für das $P(1) = 4$, $P(2) = 1$ und $P(4) = 3$ gilt. Gesucht ist also eine Gerade oder eine Parabel, die durch die Punkte $(1, 4)$, $(2, 1)$ und $(4, 3)$ geht:



Lösung:

$$L_0(x) = \frac{(x-2)(x-4)}{(1-2)(1-4)} = \frac{1}{3}(x-2)(x-4)$$

$$L_1(x) = \frac{(x-1)(x-4)}{(2-1)(2-4)} = -\frac{1}{2}(x-1)(x-4)$$

$$L_2(x) = \frac{(x-1)(x-2)}{(4-1)(4-2)} = \frac{1}{6}(x-1)(x-2)$$

$$L(x) = 4 \cdot L_0(x) + 1 \cdot L_1(x) + 3 \cdot L_2(x)$$

$$= \frac{4}{3}(x-2)(x-4) - \frac{1}{2}(x-1)(x-4) + \frac{1}{2}(x-1)(x-2)$$

$$= \frac{4}{3}x^2 - 7x + \frac{29}{3}$$

Frage: Kann es zwei verschiedene Polynome P und Q vom Grad $\leq n$ geben, die eine vorgelegte Interpolationsaufgabe lösen, d.h., für die $P(x_0) = Q(x_0) = y_0, \dots, P(x_n) = Q(x_n) = y_n$ gilt? Die Antwort lautet: Nein! Dies beschreibt der nächste Satz.

Satz 12.

Sind x_0, x_1, \dots, x_n $n+1$ Stützstellen und y_0, y_1, \dots, y_n die $n+1$ dazugehörigen Stützwerte, so gibt es genau ein Polynom $P(x)$ vom Grad $\leq n$, für das $P(x_i) = y_i$ ($i = 0, \dots, n$) gilt.

⁹Man nennt $L(x)$ *Interpolationspolynom in Lagrange-Form* oder auch einfach *Lagrangesches Interpolationspolynom*.

Beweis. Dass ein solches Polynom immer existiert, haben wir oben bereits gezeigt; $L(x)$ leistet ja das Gewünschte. Wir zeigen, dass es nur ein solches Polynom geben kann. Angenommen, $P(x)$ und $Q(x)$ seien zwei verschiedene Polynome vom Grad $\leq n$ mit $P(x_i) = Q(x_i) = y_i (i = 0, \dots, n)$. Es gelte etwa:

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0,$$

$$Q(x) = b_n x^n + b_{n-1} x^{n-1} + \dots + b_1 x + b_0.$$

Da $P(x)$ und $Q(x)$ verschieden sein sollen, muss es eine größte Zahl $m \leq n$ geben, für die $a_m \neq b_m$ gilt. Für das Polynom $D(x) = P(x) - Q(x)$ gilt dann

1. $D(x)$ hat den Grad $m < n+1$ und $D(x)$ ist nicht das Nullpolynom¹⁰ (da $P(x)$ und $Q(x)$ verschieden sind).
2. $D(x)$ hat $n+1$ verschiedenen Nullstellen $x_i (i = 0, 1, \dots, n)$ (da $D(x_i) = P(x_i) - Q(x_i) = y_i - y_i = 0$ gilt).

Dies ist aber ein Widerspruch, da andererseits allgemein gilt:

$$\text{Jedes Polynom vom Grad } m, \text{ das nicht das Nullpolynom ist, hat höchstens } m \text{ verschiedene Nullstellen.} \quad (3.41)$$

Damit ist Satz 12 bewiesen. \square

Bemerkung zu (3.41). (3.41) ergibt sich aus der Tatsache, dass ein Polynom $p(x)$ durch $(x - a)$ teilbar ist, wenn a eine Nullstelle von $p(x)$ ist („Man kann $(x - a)$ abspalten.“); teilt man nämlich $p(x)$ durch $(x - a)$, so sinkt der Grad von $p(x)$ um 1. Ist der Grad von $p(x)$ anfangs gleich m , so kann höchstens m mal ein Linearfaktor „abspalten“ werden, d.h., es kann höchstens m verschiedene Nullstellen geben.

Es sei $f(x)$ eine Funktion, von der wir an $n + 1$ Stützstellen x_0, \dots, x_n die Funktionswerte $f(x_0) = y_0, f(x_1) = y_1, \dots, f(x_n) = y_n$ kennen. Wir wollen $f(x)$ nun durch ein Polynom $P(x)$ ersetzen, das mit $f(x)$ an den Stützstellen übereinstimmt. An anderen Stellen x brauchen $f(x)$ und $P(x)$ natürlich nicht übereinstimmen, d.h., für diese Stellen nehmen wir einen gewissen Fehler in Kauf. Ist f $(n + 1)$ -mal differenzierbar, so ist es möglich, mithilfe der $(n + 1)$ -ten Ableitung $f^{(n+1)}$ eine Abschätzung dieses Fehlers anzugeben. Dies ist der Inhalt des nächsten Satzes, den wir ohne Beweis angeben¹¹.

Satz 13.

Es sei f auf $[a, b]$ $(n + 1)$ -mal differenzierbar und $f^{(n+1)}$ sei stetig. Für die $n + 1$ Stützstellen x_0, \dots, x_n gelte $f(x_i) = y_i$ und P sei das eindeutig bestimmte Polynom vom Grad $\leq n$, für das $P(x_i) = f(x_i) = y_i$ gilt ($i = 0, \dots, n$). Dann gilt für jedes $x \in [a, b]$:

$$\left| f(x) - P(x) \right| \leq M_{n+1} \frac{|x - x_0| \cdot |x - x_1| \cdot \dots \cdot |x - x_n|}{(n + 1)!} \quad (3.42)$$

$$\text{mit } M_{n+1} = \max \left\{ \left| f^{(n+1)}(\xi) \right| : \xi \in [a, b] \right\}.$$

3.5.2 Newton-Interpolation

In vielen Fällen ist es zweckmäßig, das eindeutig bestimmte Interpolationspolynom nicht wie oben in der Lagrangeschen Form (3.40) zu berechnen, sondern das folgende rekursive Verfahren, das auf Newton zurückgeht, zu verwenden. Man setzt das Interpolationspolynom vom Grad $\leq n$ dabei in der folgenden Form („Interpolationspolynom in Newton-Form“ oder auch einfach „Newtonsches Interpolationspolynom“) an:

$$N(x) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)(x - x_1) + \dots + \alpha_n(x - x_0)(x - x_1) \dots (x - x_{n-1}). \quad (3.43)$$

¹⁰Das Nullpolynom ist das Polynom, bei dem alle Koeffizienten gleich Null sind. $D(x)$ ist nicht das Nullpolynom, da der Koeffizient $a_m - b_m$ verschieden von Null ist.

¹¹Einen Beweis findet man in den gängigen Lehrbüchern der Analysis, etwa H. Heuser: Lehrbuch der Analysis 1.

Die Zahlen $\alpha_0, \dots, \alpha_n$ sind dabei zunächst noch unbestimmt. Die Forderung $N(x_i) = y_i$ für $i = 0, \dots, n$ führt zu folgendem Gleichungssystem, aus dem man die Zahlen $\alpha_0, \dots, \alpha_n$ (in dieser Reihenfolge) rekursiv berechnen kann:

$$\begin{aligned} y_0 &= \alpha_0 \\ y_1 &= \alpha_0 + \alpha_1(x_1 - x_0) \\ y_2 &= \alpha_0 + \alpha_1(x_2 - x_0) + \alpha_2(x_2 - x_0)(x_2 - x_1) \\ &\vdots \\ y_n &= \alpha_0 + \alpha_1(x_n - x_0) + \alpha_2(x_n - x_0)(x_n - x_1) + \dots + \alpha_n(x_n - x_0)(x_n - x_1) \dots (x_n - x_{n-1}) \end{aligned}$$

Beispiel. Wie oben sei $x_0 = 1, x_1 = 2, x_2 = 4, y_0 = 4, y_1 = 1$ und $y_2 = 3$. Die Gleichungen für α_0, α_1 und α_2 lauten:

$$\begin{aligned} 4 &= \alpha_0 \\ 1 &= \alpha_0 + \alpha_1 \\ 3 &= \alpha_0 + \alpha_1 \cdot 3 + \alpha_2 \cdot 3 \cdot 2 \end{aligned}$$

Es folgt $\alpha_0 = 4, \alpha_1 = -3$ und $\alpha_2 = \frac{4}{3}$; also gilt:

$$N(x) = 4 - 3(x - 1) + \frac{4}{3}(x - 1)(x - 2) = \frac{4}{3}x^2 - 7x + \frac{29}{3}.$$

Der Newtonsche Ansatz (3.43) hat gegenüber dem Lagrangeschen Ansatz (3.40) den Vorteil, dass man nicht wieder ganz von vorne anfangen muss, wenn eine weitere Stützstelle x_{n+1} und ein weiterer Stützwert y_{n+1} hinzugenommen werden.

Beispiel. Es seien $x_0 = 1, x_1 = 2, x_2 = 4, y_0 = 4, y_1 = 1$ und $y_2 = 3$ wie oben. Als weitere Stützstelle komme $x_3 = 3$ zusammen mit $y_3 = 0$ hinzu.

Gesucht: Ein Polynom Q höchstens 3-ten Grades, das $Q(x_i) = y_i$ ($i = 0, \dots, 3$) erfüllt. Die Gleichungen für α_0, α_1 und α_2 bleiben dieselben wie oben; wir können also die Werte $\alpha_0 = 4, \alpha_1 = -3$ und $\alpha_2 = \frac{4}{3}$ übernehmen und brauchen nur noch α_3 zu berechnen. Als Gleichung für α_3 ergibt sich:

$$0 = \alpha_0 + \alpha_1 \cdot 2 + \alpha_2 \cdot 2 \cdot 1 + \alpha_3 \cdot 2 \cdot 1 \cdot (-1).$$

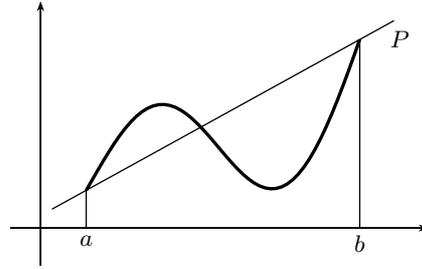
Es folgt durch Einsetzen von α_0, α_1 und α_2 : $\alpha_3 = \frac{1}{3}$. Also ist:

$$\begin{aligned} Q(x) &= \frac{4}{3}x^2 - 7x + \frac{29}{3} + \frac{1}{3}(x - 1)(x - 2)(x - 4) \\ &= \frac{1}{3}x^3 - x^2 - \frac{7}{3}x + 7 \end{aligned}$$

Wir behandeln nun einen Spezialfall, den wir im nächsten Abschnitt über Numerische Integration verwenden werden.

3.5.3 Lineare Interpolation

Es sei f auf $[a, b]$ definiert und 2-mal differenzierbar mit stetiger 2-ter Ableitung f'' . Wir ersetzen f auf dem Intervall $[a, b]$ durch die Gerade P , die durch die Punkte $(a, f(a))$ und $(b, f(b))$ geht. Was können wir über den Fehler aussagen, den wir dabei machen?



Antwort: Wir betrachten das (sehr einfache) Interpolationsproblem mit zwei Stützstellen $x_0 = a$ und $x_1 = b$ sowie den Stützwerten $y_0 = f(a)$ und $y_1 = f(b)$. Als Lösung ergibt sich die folgende Gerade:

$$P(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a). \quad (3.44)$$

Für den Fehler $|f(x) - P(x)|$ erhalten wir nach (3.42):

$$\begin{aligned} |f(x) - P(x)| &\leq M_2 \frac{(x - a)(b - x)}{2} \\ \text{mit } M_2 &= \max \left\{ |f''(\xi)| : \xi \in [a, b] \right\}. \end{aligned} \quad (3.45)$$

3.6 Numerische Integration (Trapezregel)

Bisher haben wir bestimmte Integrale $\int_a^b f(x) dx$ in der Regel dadurch berechnet, dass wir uns eine Stammfunktion $F(x)$ von $f(x)$ verschafft und $F(b) - F(a)$ bestimmt haben. Diese Methode lässt sich allerdings in vielen Fällen nicht in der gewohnten Art durchführen. Das liegt daran, dass man einerseits zwar von sämtlichen stetigen Funktionen $f(x)$ weiß, dass eine Stammfunktion $F(x)$ existiert und dass man diese Stammfunktion in der Form $\int_a^x f(x) dx$ darstellen kann¹²; andererseits kann man aber für viele stetige Funktionen $f(x)$ beweisen, dass sie „nicht elementar integrierbar“ sind. Das soll heißen: Es ist unmöglich, $F(x)$ aus den geläufigen sog. elementaren Funktionen (wie z.B. den rationalen Funktionen, e^x , $\sin x$, $\cos x$ sowie deren Umkehrfunktionen) durch endlich viele Additionen, Subtraktionen, Multiplikationen, Divisionen oder Hintereinanderausführungen zu gewinnen. (Man kann also $F(x)$ beispielsweise nicht auf die folgenden Arten darstellen: $e^{2x} \cdot \sin x + \sqrt{x}$ oder $\frac{x^3+1}{x-1} \sin x$ oder $\left(\ln x + \frac{1}{\sqrt{x^2+1}}\right)^2$.) Nicht elementar integrierbar sind schon so einfache Funktionen wie:

$$\frac{e^x}{x}, \quad \frac{\sin x}{x}, \quad \frac{1}{\ln x}, \quad e^{x^2}, \quad e^{-x^2} \quad \text{oder} \quad \frac{1}{\sqrt{1+x^4}}.$$

Bei komplizierteren Funktionen kann es geradezu als die Regel angesehen werden, dass sie nicht elementar integrierbar sind.

Zur Berechnung von bestimmten Integralen $\int_a^b f(x) dx$ nicht elementar integrierbarer Funktionen $f(x)$ kann man in vielen Fällen Potenzreihenentwicklungen von $F(x)$ heranziehen (siehe Abschnitt 4.2); man kann aber auch Methoden verwenden, die unmittelbar an die geometrische Deutung des Integrals als Flächeninhalt anknüpfen.

Zwei derartige Verfahren sind die **Trapezregel** und die **Simpsonregel**. Bei beiden Verfahren wird das Integral *näherungsweise* berechnet. Wir besprechen in dieser Vorlesung nur die Trapezregel; für die Simpsonregel verweisen wir auf die gängigen Lehrbücher der Analysis.

¹²Dies gilt aufgrund des Fundamentalsatzes der Differential- und Integralrechnung.

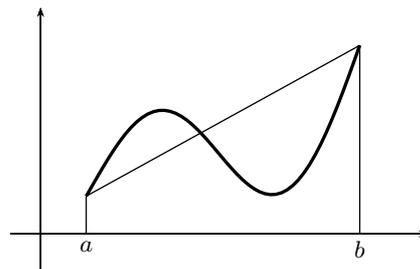
Grundidee der Trapezregel.

Man ersetzt $f(x)$ auf dem Intervall $[a, b]$ durch die Gerade $P(x)$, die durch die Punkte $(a, f(a))$ und $(b, f(b))$ geht und nimmt als Näherungswert für $\int_a^b f(x) dx$ das leicht zu berechnende Integral $\int_a^b P(x) dx$.

Es gilt:

$$\int_a^b P(x) dx = \frac{b-a}{2} (f(a) + f(b)). \quad (3.46)$$

Man kann (3.46) durch elementargeometrische Überlegungen erhalten oder auch dadurch, dass man für $P(x)$ die rechte Seite von (3.44) einsetzt und integriert. Sind $f(a)$ und $f(b)$ nicht negativ, so gibt (3.46) den Flächeninhalt des Trapez an, das durch die Punkte $(a, 0)$, $(a, f(a))$, $(b, 0)$ und $(b, f(b))$ gebildet wird.



Welchen Fehler macht man, wenn man $\int_a^b f(x) dx$ durch (3.46) ersetzt? Hierauf gibt der folgende Satz, der auf der Fehlerabschätzung (3.45) beruht, eine Antwort.

Satz 14.

Es sei f auf $[a, b]$ zweimal differenzierbar mit stetiger zweiter Ableitung f'' . Dann gilt:

$$\left| \int_a^b f(x) dx - \frac{b-a}{2} (f(a) + f(b)) \right| \leq M_2 \frac{(b-a)^3}{12} \quad (3.47)$$

$$\text{mit } M_2 = \max \{ |f''(\xi)| : \xi \in [a, b] \}.$$

Beweis. Es gilt:

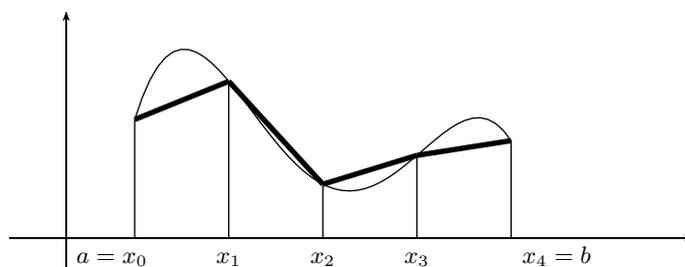
$$\begin{aligned} \left| \int_a^b f(x) dx - \frac{b-a}{2} (f(a) + f(b)) \right| &\stackrel{(3.46)}{=} \left| \int_a^b f(x) dx - \int_a^b P(x) dx \right| \\ &\stackrel{(3.10)}{=} \left| \int_a^b (f(x) - P(x)) dx \right| \\ &\stackrel{\text{Satz 5(b)}}{\leq} \int_a^b |f(x) - P(x)| dx \\ &\stackrel{(3.45)}{\leq} \frac{M_2}{2} \int_a^b (x-a)(x-b) dx \\ &= \frac{M_2}{2} \cdot \frac{(b-a)^3}{6} \quad (\text{nach eifrigem Ausrechnen}) \quad \square \end{aligned}$$

Verfeinerung der Grundidee.

Man zerlegt das Intervall $[a, b]$ äquidistant in n Teilintervalle:

$$a = x_0 < x_1 < \dots < x_n = b.$$

Auf jedem der Teilintervalle $[x_{i-1}, x_i]$ ersetzt man f durch die Gerade P_i , die durch die Punkte $(x_{i-1}, f(x_{i-1}))$ und $(x_i, f(x_i))$ geht.



Als Näherung für $\int_a^b f(x) dx$ nimmt man:

$$\begin{aligned} \sum_{i=1}^n \left(\int_{x_{i-1}}^{x_i} P_i(x) dx \right) &\stackrel{(3.46)}{=} \sum_{i=1}^n \left(\frac{x_i - x_{i-1}}{2} (f(x_{i-1}) + f(x_i)) \right) \\ &= \sum_{i=1}^n \left(\frac{b-a}{2n} (f(x_{i-1}) + f(x_i)) \right) \\ &= \frac{b-a}{2n} (f(x_0) + 2f(x_1) + \dots + 2f(x_{n-1}) + f(x_n)) \end{aligned} \quad (3.48)$$

Es ergibt sich:

Trapezregel

$$\int_a^b f(x) dx \approx \frac{b-a}{2n} (f(x_0) + 2f(x_1) + \dots + 2f(x_{n-1}) + f(x_n))$$

Hierbei bedeutet das Zeichen \approx , dass die rechts stehende Summe als Näherungswert für $\int_a^b f(x) dx$ anzusehen ist. (Hinweis: Die Terminologie ist nicht ganz einheitlich. Üblich ist beispielsweise auch, die Formel (3.46) als Trapezregel zu bezeichnen und die im Kasten aufgeführte Regel *zusammengesetzte* (oder *summierte*) *Trapezregel* zu nennen.

Satz 15 (Fehlerabschätzung bei der Trapezregel).

Es sei f auf $[a, b]$ zweimal differenzierbar mit stetiger zweiter Ableitung f'' . Dann gilt:

$$\left| \int_a^b f(x) dx - \frac{b-a}{2n} (f(x_0) + 2f(x_1) + \dots + 2f(x_{n-1}) + f(x_n)) \right| \leq M_2 \frac{(b-a)^3}{12n^2} \quad (3.49)$$

mit $M_2 = \max \{ |f''(\xi)| : \xi \in [a, b] \}$.

Beweis. Es gilt:

$$\begin{aligned}
 & \left| \int_a^b f(x) \, dx - \frac{b-a}{2n} (f(x_0) + 2f(x_1) + \dots + 2f(x_{n-1}) + f(x_n)) \right| \\
 \stackrel{(3.12), (3.48)}{=} & \left| \sum_{i=1}^n \left(\int_{x_{i-1}}^{x_i} f(x) \, dx \right) - \sum_{i=1}^n \left(\frac{x_i - x_{i-1}}{2} (f(x_{i-1}) + f(x_i)) \right) \right| \\
 = & \left| \sum_{i=1}^n \left(\int_{x_{i-1}}^{x_i} f(x) \, dx - \frac{x_i - x_{i-1}}{2} (f(x_{i-1}) + f(x_i)) \right) \right| \\
 \text{Dreiecks-} & \\
 \text{ungleichung} & \leq \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) \, dx - \frac{x_i - x_{i-1}}{2} (f(x_{i-1}) + f(x_i)) \right| \\
 \stackrel{(3.47)}{\leq} & \sum_{i=1}^n \left(M_2 \frac{(x_i - x_{i-1})^3}{12} \right) \\
 = & \sum_{i=1}^n \left(M_2 \frac{\left(\frac{b-a}{n}\right)^3}{12} \right) \\
 = & M_2 \frac{(b-a)^3}{12n^2} \quad \square
 \end{aligned}$$

Beispiel. Es sei $f(x) = \sin x$. Wir testen die summierte Trapezregel am folgenden Integral:

$$\int_0^1 \sin x \, dx.$$

$n = 1$

Dies ist der Fall der „einfachen“ Trapezregel $\frac{b-a}{2} (f(a) + f(b))$ (vgl. (3.46)). In unserem Beispiel erhält man als Näherungswert:

$$\int_0^1 \sin x \, dx \approx \frac{1}{2} (\sin(0) + \sin(1)) \approx 0.4207354924.$$

$n = 2$

Man erhält als Näherungswert:

$$\int_0^1 \sin x \, dx \approx \frac{1}{4} \left(\sin(0) + 2 \sin\left(\frac{1}{2}\right) + \sin(1) \right) \approx 0.4500805155.$$

$n = 3$

Man erhält als Näherungswert:

$$\int_0^1 \sin x \, dx \approx \frac{1}{6} \left(\sin(0) + 2 \sin\left(\frac{1}{3}\right) + 2 \sin\left(\frac{2}{3}\right) + \sin(1) \right) \approx 0.4554333307.$$

Wir vergleichen unseren Näherungswert mit dem exakten Wert:

$$\begin{aligned} \int_0^1 \sin x \, dx &= [-\cos x]_0^1 \\ &= -\cos(1) + \cos(0) \\ &= 1 - \cos(1) \\ &\approx 0.4596976941 \end{aligned}$$

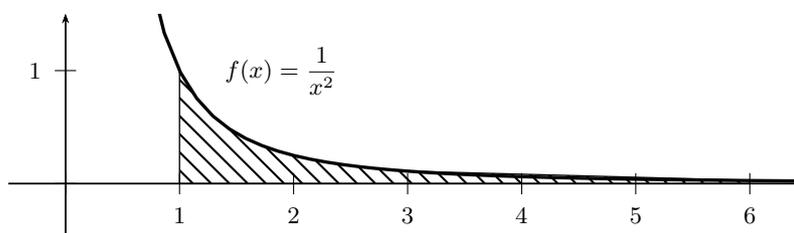
Für $n = 4, 5, \dots$ werden die Näherungswerte allmählich immer besser und konvergieren gegen den exakten Wert. Im Allgemeinen nähern sich die mit der summierten Trapezregel berechneten Werte allerdings *nur recht langsam* dem exakten Wert des Integrals. *Schneller geht es mit der Simpsonregel, deren Grundidee es ist, anstelle von Geraden (wie bei der Trapezregel) Parabelbögen zur Annäherung der Funktion zu verwenden.* Die Simpsonregel wird, wie bereits erwähnt, in diesem Skript jedoch nicht näher behandelt.

3.7 Uneigentliche Integrale

Wir haben bisher bei der Betrachtung von Integralen $\int_a^b f(x) \, dx$ stets vorausgesetzt, dass das Intervall $[a, b]$ endlich und f beschränkt ist. Solche Integrale nennt man *eigentliche Integrale*. Entsprechend sagt man auch, dass f in $[a, b]$ *eigentlich integrierbar* sei. Wir werden im Folgenden durch die Definition von sog. uneigentlichen Integralen den Integralbegriff auf unbeschränkte Funktionen und unendliche Intervalle ausdehnen.

Beispiele:

- Wir betrachten die Funktion $f(x) = \frac{1}{x^2}$ in $[1, \infty)$. Welchen Flächeninhalt hat die Fläche, die von der x -Achse, dem Graphen von $f(x)$ und der Geraden $x = 1$ eingeschlossen wird? Ist es überhaupt möglich, dieser Fläche einen endlichen Flächeninhalt in sinnvoller Weise zuzuordnen?



Offenbar müssen wir $\lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} \, dx$ berechnen. Es gilt:

$$\begin{aligned} \int_1^b \frac{1}{x^2} \, dx &= [-x^{-1}]_1^b = -\frac{1}{b} + 1 \rightarrow 1 \quad \text{für } b \rightarrow \infty \\ \implies \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} \, dx &= 1 \end{aligned}$$

Der gesuchte Flächeninhalt ist also 1. Anstelle von $\lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} \, dx$ schreibt man auch $\int_1^{\infty} \frac{1}{x^2} \, dx$ und nennt dies ein *uneigentliches Integral*.

2. Dieselbe Frage wie in Beispiel 1 für $f(x) = \frac{1}{x}$. Es gilt:

$$\int_1^b \frac{1}{x} dx = \left[\ln x \right]_1^b = \ln b - \ln 1 = \ln b$$

$$\implies \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x} dx = \lim_{b \rightarrow \infty} (\ln b) = \infty$$

Die entsprechende Fläche hat also keinen endlichen Flächeninhalt.

Diese Beispiele führen zur folgenden Definition.

Definition.

Für ein festes $a \in \mathbb{R}$ gelte: Die Funktion f sei (eigentlich) integrierbar in jedem Intervall $[a, b]$ mit $b > a$. Existiert der Grenzwert

$$\lim_{b \rightarrow \infty} \int_a^b f(x) dx,$$

so heißt f *uneigentlich integrierbar* in $[a, \infty)$ und man definiert:

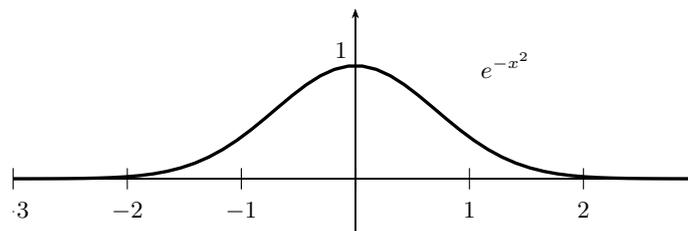
$$\int_a^\infty f(x) dx = \lim_{b \rightarrow \infty} \int_a^b f(x) dx.$$

Man nennt $\int_a^\infty f(x) dx$ ein *uneigentliches Integral*. Analog definiert man $\int_{-\infty}^b f(x) dx$. Falls f für ein $c \in \mathbb{R}$ sowohl in $(-\infty, c]$ als auch in $[c, \infty)$ uneigentlich integrierbar ist, so definiert man:

$$\int_{-\infty}^\infty f(x) dx = \int_{-\infty}^c f(x) dx + \int_c^\infty f(x) dx.$$

In der Wahrscheinlichkeitstheorie und der Statistik spielt das folgende uneigentliche Integral eine wichtige Rolle:

$$\int_{-\infty}^\infty e^{-x^2} dx.$$



Wir erwähnen (Beweis siehe W. Luh oder andere Lehrbücher der Analysis):

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Bisher haben wir uneigentliche Integrale betrachtet, bei denen über einem unendlichen Intervall (d.h. einem Intervall der Form $[a, \infty)$, $(-\infty, b]$ oder $(-\infty, \infty)$) integriert wurde. Darüber hinaus gibt es auch uneigentliche Integrale, bei denen das Integrationsintervall $[a, b]$ endlich ist, aber der Integrand f nicht beschränkt ist.

Definition.

Es seien a und b reelle Zahlen, für die $a < b$ gilt. Die Funktion $f : [a, b) \rightarrow \mathbb{R}$ sei auf dem halboffenen Intervall $[a, b)$ definiert; außerdem sei f eigentlich integrierbar in jedem Teilintervall $[a, \beta]$ mit $a < \beta < b$. Falls der Grenzwert

$$\lim_{\beta \rightarrow b} \int_a^{\beta} f(x) dx$$

existiert, so nennt man f *uneigentlich integrierbar* in $[a, b]$. Wir definieren

$$\int_a^b f(x) dx := \lim_{\beta \rightarrow b} \int_a^{\beta} f(x) dx$$

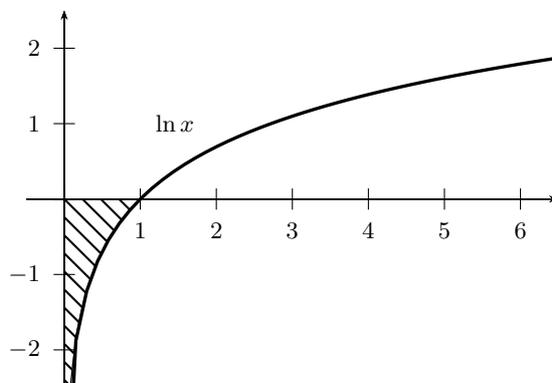
und bezeichnen $\int_a^b f(x) dx$ als *uneigentliches Integral*.

Analog definiert man: Ist $f : (a, b] \rightarrow \mathbb{R}$ eigentlich integrierbar in jedem Intervall $[\alpha, b]$ mit $a < \alpha < b$, so wird das uneigentliche Integral $\int_a^b f(x) dx$ durch

$$\int_a^b f(x) dx = \lim_{\alpha \rightarrow a} \int_{\alpha}^b f(x) dx$$

definiert, falls dieser Grenzwert existiert.

Beispiel. Wir betrachten $\int_0^1 \ln x dx$ und fragen uns, ob dieses uneigentliche Integral existiert (siehe Zeichnung).



Zu berechnen ist also $\lim_{\alpha \rightarrow 0} \int_{\alpha}^1 \ln x \, dx$. Eine Stammfunktion zu $\ln x$ ist $x \cdot \ln x - x$. Es folgt:

$$\begin{aligned} \int_0^1 \ln x \, dx &= \lim_{\alpha \rightarrow 0} \int_{\alpha}^1 \ln x \, dx \\ &= \lim_{\alpha \rightarrow 0} \left[x \ln x - x \right]_{\alpha}^1 \\ &= \lim_{\alpha \rightarrow 0} (1 \cdot \ln 1 - 1 - \alpha \ln \alpha + \alpha) \\ &= 0 - 1 - \lim_{\alpha \rightarrow 0} (\alpha \ln \alpha) + 0 \\ &\stackrel{(*)}{=} -1 \end{aligned}$$

An der Stelle (*) wurde benutzt, dass $\lim_{\alpha \rightarrow 0} \alpha \ln \alpha = 0$ (siehe Beispiel 1 in Abschnitt 2.7.2).

4 Reihen

4.1 Konvergenzkriterien für Reihen

Grundlegendes über Reihen haben wir bereits in Abschnitt 1.3 gesagt. Dort haben wir auch drei wichtige Beispiele für Reihen behandelt:

1) $\sum_{i=0}^{\infty} q^i$ (geometrische Reihe)

2) $\sum_{i=1}^{\infty} \frac{1}{i}$ (harmonische Reihe)

3) $\sum_{i=1}^{\infty} \frac{1}{i^2}$

Es seien $\sum_{i=0}^{\infty} a_i$ und $\sum_{i=0}^{\infty} b_i$ konvergente Reihen und $c \in \mathbb{R}$. Dann sind auch die Reihen

$$\sum_{i=0}^{\infty} (a_i + b_i) \quad \text{und} \quad \sum_{i=0}^{\infty} c \cdot a_i$$

konvergent und es gilt:

$$\sum_{i=0}^{\infty} (a_i + b_i) = \sum_{i=0}^{\infty} a_i + \sum_{i=0}^{\infty} b_i \quad (4.1)$$

$$\sum_{i=0}^{\infty} c \cdot a_i = c \sum_{i=0}^{\infty} a_i \quad (4.2)$$

Die Regeln (4.1) und (4.2) folgen unmittelbar aus den Rechenregeln (a) und (c) (Satz 8, Abschnitt 1.3) für Grenzwerte von Folgen.

Satz 1.

Konvergiert die Reihe $\sum_{i=0}^{\infty} a_i$, so gilt $\lim_{i \rightarrow \infty} a_i = 0$.

Beweis. $\sum_{i=0}^{\infty} a_i$ sei konvergent, d.h., für die Folge $s_n = \sum_{i=0}^n a_i$ der Partialsummen gilt $\lim_{n \rightarrow \infty} s_n = s$ für ein $s \in \mathbb{R}$. Für $n \geq 1$ gilt:

$$a_n = \sum_{i=0}^n a_i - \sum_{i=0}^{n-1} a_i = s_n - s_{n-1}.$$

Es folgt:

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (s_n - s_{n-1}) = \lim_{n \rightarrow \infty} s_n - \lim_{n \rightarrow \infty} s_{n-1} = s - s = 0. \quad \square$$

Nach Satz 1 ist es notwendig für die Konvergenz einer Reihe $\sum_{i=0}^{\infty} a_i$, dass die Folge ihrer Glieder gegen Null strebt. Umgekehrt ist diese Bedingung aber *keineswegs hinreichend* für die Konvergenz einer Reihe, wie

das Beispiel der harmonischen Reihe $\sum_{i=1}^{\infty} \frac{1}{i}$ zeigt: Diese Reihe ist nicht konvergent, obwohl ihre Glieder $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$ eine Nullfolge bilden.

Satz 1 liefert auch eine Begründung, weshalb die geometrische Reihe $\sum_{i=0}^{\infty} q^i$ nicht konvergiert, falls $|q| \geq 1$ gilt: In diesem Fall bilden die Glieder $a_i = q^i$ keine Nullfolge.

Wir haben bereits gesehen, dass die Reihe $\sum_{i=1}^{\infty} \frac{1}{i^2}$ konvergiert, während die Reihe $\sum_{i=1}^{\infty} \frac{1}{i}$ divergiert. Als Verallgemeinerung dieser Beobachtung erwähnen wir den folgenden Satz¹ (Beweis siehe Lehrbücher der Analysis).

Satz 2.

Die Reihe $\sum_{i=1}^{\infty} \frac{1}{i^\alpha}$ konvergiert für $\alpha > 1$ und divergiert für $\alpha \leq 1$.

Wir behandeln im Folgenden vier Kriterien für die Konvergenz von Reihen:

- das Majorantenkriterium
- das Quotientenkriterium
- das Wurzelkriterium
- das Leibnizsche Kriterium

4.1.1 Das Majorantenkriterium

Satz 3 (Majorantenkriterium).

Es sei $\sum_{i=0}^{\infty} b_i$ eine konvergente Reihe mit nichtnegativen Gliedern (d.h. $b_i \geq 0$ für $i = 0, 1, \dots$). Ist dann $\sum_{i=0}^{\infty} a_i$ eine Reihe, für die $|a_i| \leq b_i$ für alle $i = 0, 1, \dots$ gilt, so konvergiert auch die Reihe $\sum_{i=0}^{\infty} a_i$.

Bemerkungen.

(a) Unter den Voraussetzungen von Satz 3 nennt man die Reihe $\sum_{i=0}^{\infty} b_i$ eine *konvergente Majorante* von

$$\sum_{i=0}^{\infty} a_i.$$

(b) Das Majorantenkriterium wird manchmal auch *Vergleichskriterium* genannt.

(c) Ist in Satz 3 auch die Reihe $\sum_{i=0}^{\infty} a_i$ eine Reihe mit nichtnegativen Gliedern, so folgt Satz 3 aus dem

Satz über monotone, beschränkte Folgen (Satz 5, Abschnitt 1.1), nämlich so: Die Folge $s_n = \sum_{i=0}^n a_i$

der Partialsummen $\sum_{i=0}^n a_i$ ist wegen

$$\left| \sum_{i=0}^n a_i \right| \leq \sum_{i=0}^n |a_i| \leq \sum_{i=0}^n b_i \leq K := \sum_{i=0}^{\infty} b_i$$

beschränkt; sind die a_i nichtnegativ, so gilt $s_n \leq s_n + a_{n+1} = s_{n+1}$, d.h., die Folge der s_n ist auch monoton und somit konvergent.

¹Auch die in Satz 2 betrachteten Reihen $\sum_{i=1}^{\infty} \frac{1}{i^\alpha}$ werden oft als harmonische Reihen bezeichnet.

Beweis von Satz 3. Es sei $\sum_{i=0}^{\infty} a_i$ eine Reihe mit $|a_i| \leq b_i$ (für $i = 0, 1, \dots$). Wir bilden aus der Reihe $\sum_{i=0}^{\infty} a_i$ zwei neue Reihen

$$\sum_{i=0}^{\infty} a_i^* \quad \text{und} \quad \sum_{i=0}^{\infty} a_i^{**} \quad (4.3)$$

indem wir setzen:

$$a_i^* := \begin{cases} a_i, & \text{falls } a_i \geq 0 \\ 0, & \text{falls } a_i < 0 \end{cases} \quad a_i^{**} := \begin{cases} 0, & \text{falls } a_i \geq 0 \\ -a_i, & \text{falls } a_i < 0 \end{cases}$$

Diese neuen Reihen haben ausschließlich nichtnegative Glieder und es gilt (wegen $|a_i| \leq b_i$) $a_i^* \leq b_i$ und $a_i^{**} \leq b_i$. Nach der obigen Bemerkung (c) sind die Reihen (4.3) konvergent. Es gilt ferner $a_i = a_i^* - a_i^{**}$ (für $i = 0, 1, \dots$). Also (siehe (4.1) und (4.2)):

$$\sum_{i=0}^{\infty} a_i = \sum_{i=0}^{\infty} (a_i^* - a_i^{**}) = \sum_{i=0}^{\infty} a_i^* - \sum_{i=0}^{\infty} a_i^{**}.$$

D.h., $\sum_{i=0}^{\infty} a_i$ ist ebenfalls konvergent. \square

Oft vorkommender Spezialfall des Majorantenkriteriums.

Es sei $\sum_{i=0}^{\infty} a_i$ eine Reihe und für Konstanten $M > 0$ und q mit $0 < q < 1$ gelte:

$$|a_i| \leq Mq^i \quad (i = 0, 1, \dots). \quad (4.4)$$

Dann ist $\sum_{i=0}^{\infty} a_i$ konvergent, da $\sum_{i=0}^{\infty} Mq^i = M \sum_{i=0}^{\infty} q^i$ eine konvergente Majorante ist.

(Man beachte: $\sum_{i=0}^{\infty} q^i$ ist eine geometrische Reihe mit $0 < q < 1$.)

Definition.

Eine Reihe $\sum_{i=0}^{\infty} a_i$ heißt *absolut konvergent*, falls die Reihe $\sum_{i=0}^{\infty} |a_i|$ konvergiert.

Ist $\sum_{i=0}^{\infty} a_i$ absolut konvergent, so ist $\sum_{i=0}^{\infty} |a_i|$ eine konvergente Majorante von $\sum_{i=0}^{\infty} a_i$. Aus Satz 3 ergibt sich somit:

$$\text{Ist eine Reihe absolut konvergent, so ist sie auch konvergent.} \quad (4.5)$$

Die Umkehrung hiervon gilt nicht, wie wir an einem späteren Beispiel noch sehen werden.

4.1.2 Das Quotientenkriterium

Satz 4 (Quotientenkriterium).

Gegeben sei eine Reihe $\sum_{i=0}^{\infty} a_i$ mit $a_i \neq 0$ für alle bis auf höchstens endlich viele i .

- (a) Die Reihe konvergiert (und zwar sogar absolut), falls es ein $i_0 \geq 0$ und ein q mit $0 < q < 1$ gibt, so dass für alle $i \geq i_0$ gilt:

$$\left| \frac{a_{i+1}}{a_i} \right| \leq q < 1.$$

- (b) Die Reihe divergiert, falls es ein $i_0 \geq 0$ gibt, so dass für alle $i \geq i_0$ gilt:

$$\left| \frac{a_{i+1}}{a_i} \right| \geq 1.$$

Beweis.

- (a) Wir setzen $M = |a_{i_0}|$. Aus $\left| \frac{a_{i+1}}{a_i} \right| \leq q$ folgt $|a_{i+1}| \leq |a_i|q$ (für $i \geq i_0$). Also erhält man:

$$\begin{aligned} |a_{i_0+1}| &\leq |a_{i_0}|q = Mq \\ |a_{i_0+2}| &\leq |a_{i_0+1}|q \leq Mq^2 \\ |a_{i_0+3}| &\leq |a_{i_0+2}|q \leq Mq^3 \\ &\vdots \end{aligned}$$

Allgemein gilt:

$$|a_{i_0+j}| \leq Mq^j$$

Folglich ist die Reihe $\sum_{j=0}^{\infty} Mq^j$ eine konvergente Majorante der Reihe $\sum_{j=0}^{\infty} |a_{i_0+j}|$, d.h. die Reihe $\sum_{j=0}^{\infty} |a_{i_0+j}|$ ist konvergent. Da sich diese Reihe von der Reihe $\sum_{i=0}^{\infty} |a_i|$ nur um endlich viele Glieder unterscheidet², ist auch $\sum_{i=0}^{\infty} |a_i|$ konvergent.

- (b) Aus $\left| \frac{a_{i+1}}{a_i} \right| \geq 1$ folgt $|a_{i+1}| \geq |a_i|$ (für $i \geq i_0$). Also gilt:

$$0 < |a_{i_0}| \leq |a_{i_0+1}| \leq |a_{i_0+2}| \leq \dots$$

Die a_i bilden also keine Nullfolge. Damit kann die Reihe nicht konvergent sein. \square

Wir geben noch eine häufig benutzte Folgerung aus dem Quotientenkriterium an, die wir die *Limes-Version des Quotientenkriteriums* nennen:

Gegeben sei eine Reihe $\sum_{i=0}^{\infty} a_i$ mit $a_i \neq 0$ für alle bis auf höchstens endlich viele i . Dann gilt:

$$\text{Ist } \lim_{i \rightarrow \infty} \left| \frac{a_{i+1}}{a_i} \right| < 1, \text{ so konvergiert die Reihe absolut.}$$

$$\text{Ist } \lim_{i \rightarrow \infty} \left| \frac{a_{i+1}}{a_i} \right| > 1, \text{ so divergiert die Reihe.}$$

²Klarerweise gilt die folgende Feststellung, die wir oft benutzen werden ohne sie ausdrücklich zu erwähnen:

Ist $\sum_{i=0}^{\infty} a_i$ konvergent, so ist auch jede Reihe konvergent, die aus $\sum_{i=0}^{\infty} a_i$ durch Hinzufügen, Weglassen oder Abändern endlich vieler Glieder entsteht.

Der Deutlichkeit halber sei angemerkt, dass die zweite Bedingung den Fall einschließt, dass der genannte Limes gleich ∞ ist.

Aufgabe: Zeigen Sie, dass die Limes-Version aus Satz 4 folgt.

4.1.3 Das Wurzelkriterium

Satz 5 (Wurzelkriterium).

(a) Die Reihe $\sum_{i=0}^{\infty} a_i$ konvergiert (und zwar sogar absolut), falls es ein $i_0 \geq 0$ und ein q mit $0 < q < 1$ gibt, so dass für alle $i \geq i_0$ gilt:

$$\sqrt[i]{|a_i|} \leq q < 1.$$

(b) Die Reihe divergiert, falls es ein $i_0 \geq 0$ gibt, so dass für alle $i \geq i_0$ gilt:

$$\sqrt[i]{|a_i|} \geq 1.$$

Der Beweis von Satz 5 ist ähnlich dem Beweis von Satz 4. *Empfehlung:* Führen Sie diesen Beweis selbst aus. Orientieren Sie sich dabei am Beweis des Quotientenkriteriums.

Aus dem Wurzelkriterium lässt sich eine analoge Folgerung ziehen wie die aus dem Quotientenkriterium (*Limes-Version des Wurzelkriteriums*³). Für jede Reihe $\sum_{i=0}^{\infty} a_i$ gilt:

Ist $\lim_{i \rightarrow \infty} \sqrt[i]{|a_i|} < 1$, so konvergiert die Reihe absolut.

Ist $\lim_{i \rightarrow \infty} \sqrt[i]{|a_i|} > 1$, so divergiert die Reihe.

Die zweite Bedingung schließt wiederum den Fall ein, dass der genannte Limes gleich ∞ ist.

4.1.4 Das Leibnizsche Kriterium

Für Reihen der Form $\sum_{i=0}^{\infty} (-1)^i a_i$ mit $a_i \geq 0$ gibt es ein weiteres, oft verwendetes Konvergenzkriterium. Derartige Reihen, die abwechselnd positives und negatives Vorzeichen haben, nennt man *alternierende Reihen*.

Satz 6 (Leibnizsches Kriterium).

Eine Reihe $\sum_{i=0}^{\infty} (-1)^i a_i$ mit $a_i \geq 0$ ist konvergent, wenn (a_i) eine monotone Nullfolge ist (d.h., wenn $a_0 \geq a_1 \geq a_2 \geq \dots$ und $a_i \rightarrow 0$ gilt).

Beweis. Wir betrachten zunächst nur die Partialsummen mit einer geraden Gliederzahl, d.h.

$$\begin{aligned} s_1 &= a_0 - a_1 \\ s_3 &= (a_0 - a_1) + (a_2 - a_3) \\ &\vdots \\ s_{2n+1} &= (a_0 - a_1) + (a_2 - a_3) + (a_4 - a_5) + \dots + (a_{2n} - a_{2n+1}) \end{aligned}$$

Wegen $(a_i - a_{i+1}) \geq 0$ gilt: Die Folge s_1, s_3, s_5, \dots ist monoton wachsend. Sie ist auch beschränkt; um dies einzusehen, brauchen wir die Summe s_{2n+1} nur anders geklammert aufzuschreiben:

$$s_{2n+1} = a_0 - (a_1 - a_2) - (a_3 - a_4) - \dots - (a_{2n-1} - a_{2n}) - a_{2n+1} \leq a_0.$$

³Häufig wird in der Literatur statt „Limes-Version des Wurzelkriteriums“ auch einfach wieder „Wurzelkriterium“ gesagt; eine analoge Bemerkung gilt auch für das Quotientenkriterium.

Also ist die Folge s_1, s_3, s_5, \dots konvergent (nach Satz 5, Abschnitt 1.3). Es sei s der Grenzwert dieser Folge mit $s = \lim_{n \rightarrow \infty} s_{2n+1}$. Wir müssen jetzt nur noch zeigen, dass auch die Folge der Partialsummen s_0, s_2, s_4, \dots gegen denselben Grenzwert s strebt. Es gilt $s_{2n+2} = s_{2n+1} + a_{2n+2}$. Also:

$$\lim_{n \rightarrow \infty} s_{2n+2} = \lim_{n \rightarrow \infty} s_{2n+1} + \lim_{n \rightarrow \infty} a_{2n+2} = s + 0 = s. \quad \square$$

Wir wissen, dass jede absolut konvergente Reihe auch konvergent ist (vgl. (4.5)). Mithilfe des Leibniz-Kriteriums können wir nun leicht Beispiele für Reihen finden, die konvergent, aber nicht absolut konvergent sind: Die Reihe

$$\sum_{i=1}^{\infty} (-1)^{i+1} \cdot \frac{1}{i} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots \quad (4.6)$$

ist konvergent nach dem Leibniz-Kriterium; sie ist aber nicht absolut konvergent, da die Harmonische Reihe $\sum_{i=1}^{\infty} \frac{1}{i}$ divergent ist.

4.2 Potenzreihen

Unter einem *Polynom* versteht man bekanntlich einen Ausdruck der Form

$$\sum_{i=0}^n a_i x^i = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n,$$

wobei x eine reelle Variable ist; die reellen Zahlen a_i heißen *Koeffizienten* des Polynoms. Zu jedem Polynom gehört auf natürliche Art und Weise eine auf ganz \mathbb{R} definierte Funktion f , die jedem $x \in \mathbb{R}$ als Wert die reelle Zahl $f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$ zuordnet.

Unter einer Potenzreihe versteht man das „unendliche Analogon“ zu einem Polynom. Eine *Potenzreihe* ist ein Ausdruck der Form

$$\sum_{i=0}^{\infty} a_i x^i = a_0 + a_1 x + a_2 x^2 + \dots$$

Dabei ist x wie bei einem Polynom eine reelle Variable. Die Folge (a_i) von reellen Zahlen heißt die *Koeffizientenfolge* der Potenzreihe. Auch zu jeder Potenzreihe gehört eine Funktion $f(x) = \sum_{i=0}^{\infty} a_i x^i$, die allerdings nur für diejenigen $x \in \mathbb{R}$ definiert ist, für die die Reihe $\sum_{i=0}^{\infty} a_i x^i$ konvergiert.

Ein wichtiges Problem bei der Untersuchung von Potenzreihen ist die Bestimmung derjenigen $x \in \mathbb{R}$, für die die Reihe $\sum_{i=0}^{\infty} a_i x^i$ konvergiert, d.h. derjenigen x , für die die zugehörige Funktion f definiert ist. Zunächst betrachten wir einige **Beispiele**:

1. Wir betrachten die Potenzreihe $\sum_{i=0}^{\infty} x^i$, d.h., die Koeffizientenfolge dieser Potenzreihe ist konstant 1.

Es handelt sich hierbei natürlich um die geometrische Reihe. Diese divergiert bekanntlich für $|x| \geq 1$; sie konvergiert für $x \in (-1, 1)$ und es gilt:

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x} \quad (\text{für } x \in (-1, 1)).$$

Mit anderen Worten: Durch die Potenzreihe $\sum_{i=0}^{\infty} x^i$ wird die Funktion $f : (-1, 1) \rightarrow \mathbb{R}$ mit $f(x) = \frac{1}{1-x}$ dargestellt. (*Empfehlung*: Überlegen Sie sich zur Veranschaulichung, wie der Graph dieser Funktion aussieht und fertigen Sie eine Skizze an!)

2. Wir betrachten die Potenzreihe $\sum_{i=0}^{\infty} \frac{x^i}{i!}$. Die Koeffizientenfolge ist also in diesem Beispiel die Folge $\frac{1}{i!}$, d.h. $1, 1, \frac{1}{2}, \frac{1}{6}, \frac{1}{24}, \frac{1}{120}, \dots$. Wir bestimmen mit der Limes-Version des Quotientenkriteriums diejenigen $x \in \mathbb{R}$, für die diese Potenzreihe konvergiert. Es gilt für $x \neq 0$:

$$\lim_{i \rightarrow \infty} \left| \frac{x^{i+1}}{(i+1)!} \cdot \frac{i!}{x^i} \right| = |x| \cdot \lim_{i \rightarrow \infty} \left(\frac{1}{i+1} \right) = 0.$$

Das heißt (nach der Limes-Version des Quotientenkriteriums): Die Potenzreihe konvergiert für alle $x \neq 0$. Da jede Potenzreihe natürlich für $x = 0$ konvergiert, haben wir also gefunden: Die Potenzreihe

$$\sum_{i=0}^{\infty} \frac{x^i}{i!} \tag{4.7}$$

konvergiert für alle $x \in \mathbb{R}$. Mit anderen Worten: Durch die Potenzreihe (4.7) ist eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ dargestellt⁴. Im Gegensatz zum ersten Beispiel stellt die hier betrachtete Potenzreihe also eine Funktion dar, die auf der gesamten Menge \mathbb{R} definiert ist.

3. Wir betrachten die Potenzreihe $\sum_{i=0}^{\infty} i^i x^i$. Die Koeffizientenfolge ist also hier die Folge (i^i) (für $i = 0, 1, \dots$), d.h. $1, 1, 4, 27, 256, 5^5, 6^6, \dots$. Wir bestimmen mit dem Wurzelkriterium (genauer: mit der Limes-Version des Wurzelkriteriums) diejenigen $x \in \mathbb{R}$, für die diese Potenzreihe konvergiert. Es gilt für $x \neq 0$:

$$\lim_{i \rightarrow \infty} \sqrt[i]{|i^i x^i|} = |x| \cdot \lim_{i \rightarrow \infty} i = \infty.$$

Das heißt: Diese Potenzreihe ist nur für $x = 0$ konvergent; die zugehörige Funktion f ist also nur für $x = 0$ definiert (und deshalb uninteressant).

Satz 7.

Gegeben sei die Potenzreihe $\sum_{i=0}^{\infty} a_i x^i$. Konvergiert diese Potenzreihe für ein $x_0 \neq 0$ (d.h. $\sum_{i=0}^{\infty} a_i x_0^i$ konvergiert), so konvergiert diese Potenzreihe auch für jedes x mit $|x| < |x_0|$. Sie konvergiert für jedes x mit $|x| < |x_0|$ dann sogar absolut (d.h., $\sum_{i=0}^{\infty} |a_i x^i|$ konvergiert).

Beweis. Es sei $\sum_{i=0}^{\infty} a_i x_0^i$ konvergent ($x_0 \neq 0$). Dann ist $(a_i x_0^i)$ eine Nullfolge (Satz 1) und deshalb gibt es eine Konstante M mit $|a_i x_0^i| \leq M$ für alle $i \geq 0$. Für $|x| < |x_0|$ folgt hieraus:

$$|a_i x^i| = |a_i x_0^i| \cdot \left| \frac{x}{x_0} \right|^i \leq M \left| \frac{x}{x_0} \right|^i.$$

Wegen $|x| < |x_0|$ ist $\left| \frac{x}{x_0} \right| < 1$. Daher ist $\sum_{i=0}^{\infty} M \left| \frac{x}{x_0} \right|^i$ eine konvergente Majorante von $\sum_{i=0}^{\infty} |a_i x^i|$ (siehe Abschnitt 4.1.1, „Spezialfall des Majorantenkriteriums“). \square

Als Folgerung aus Satz 7 erhalten wir den folgenden Satz.

Satz 8.

Zu jeder Potenzreihe $\sum_{i=0}^{\infty} a_i x^i$ gibt es ein R mit $R \in \{r \in \mathbb{R} : r \geq 0\} \cup \{\infty\}$, so dass die Potenzreihe für alle x mit $|x| < R$ konvergiert (sogar absolut) und für alle x mit $|x| > R$ divergiert.

⁴Es handelt sich (wie wir später noch zeigen werden) um die Exponentialfunktion e^x .

Beweis. Wir definieren R durch

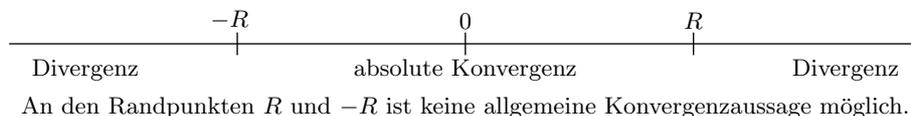
$$R = \sup \left\{ x \in \mathbb{R} : \sum_{i=0}^{\infty} a_i x^i \text{ konvergiert} \right\},$$

falls die Menge $\left\{ x \in \mathbb{R} : \sum_{i=0}^{\infty} a_i x^i \text{ konvergiert} \right\}$ beschränkt ist. Andernfalls definieren wir R durch:

$$R = \infty.$$

Die Behauptung von Satz 8 folgt dann unmittelbar aus Satz 7. \square

Man nennt R *Konvergenzradius* der Potenzreihe, das Intervall $(-R, R)$ heißt *Konvergenzintervall* der Potenzreihe. Nach Satz 8 konvergiert eine Potenzreihe also für alle x , die innerhalb des Konvergenzintervalls liegen und sie divergiert für alle x , deren Abstand vom Nullpunkt größer als R ist. Für $x = R$ bzw. $x = -R$ wird in Satz 8 keine Aussage gemacht. Es kann vorkommen, dass die Potenzreihe in beiden Randpunkten des Konvergenzintervalls $(-R, R)$ konvergiert; sie kann aber ebenso gut in beiden Randpunkten divergieren bzw. in einem Randpunkt konvergieren und im anderen divergieren.



Beispiele. Wir betrachten zunächst die obigen Beispiel 1, 2 und 3:

1. Die geometrische Reihe $\sum_{i=0}^{\infty} x^i$ hat den Konvergenzradius $R = 1$. An den Randpunkten $x = 1$ und $x = -1$ des Konvergenzintervalls liegt Divergenz vor.
2. Die Reihe $\sum_{i=0}^{\infty} \frac{x^i}{i!}$ hat den Konvergenzradius $R = \infty$.
3. Die Reihe $\sum_{i=0}^{\infty} i^i x^i$ hat den Konvergenzradius $R = 0$.
4. Die Reihe $\sum_{i=0}^{\infty} \frac{1}{i+1} x^i$ konvergiert für $x = -1$ und divergiert für $x = 1$. Also hat sie den Konvergenzradius $R = 1$. An den Randpunkten liegt einmal Konvergenz und einmal Divergenz vor.

In vielen Fällen lässt sich der Konvergenzradius einer Potenzreihe mithilfe der Limes-Version des Quotienten- bzw. Wurzelkriteriums berechnen. Beispiele hierfür stellen die obigen Beispiele 2 und 3 vor Satz 7 dar. Hier nun ein weiteres Beispiel zur Berechnung des Konvergenzradius R einer Potenzreihe.

Beispiel. Für $\sum_{i=0}^{\infty} i 5^i x^i$ soll der Konvergenzradius ermittelt werden.

Wir verwenden die *Limes-Version des Quotientenkriteriums* und betrachten dementsprechend (für $x \in \mathbb{R}$) den folgenden Grenzwert:

$$\lim_{i \rightarrow \infty} \left| \frac{(i+1)5^{i+1}x^{i+1}}{i5^i x^i} \right|.$$

Ist für ein $x \in \mathbb{R}$ dieser Grenzwert < 1 , so liegt Konvergenz vor; ist dieser Grenzwert > 1 , so liegt Divergenz vor. (Ist dieser Grenzwert $= 1$ oder existiert er gar nicht, so hat der Test kein Ergebnis geliefert.) Wir wählen ein beliebiges $x \in \mathbb{R}$; in der nachfolgenden Grenzwertbestimmung ist x also fest gewählt. Dann gilt:

$$\lim_{i \rightarrow \infty} \left| \frac{(i+1)5^{i+1}x^{i+1}}{i5^i x^i} \right| = \lim_{i \rightarrow \infty} \left| \frac{i+1}{i} \cdot 5x \right| = \lim_{i \rightarrow \infty} \left(\left| \frac{i+1}{i} \right| \cdot |5x| \right) \stackrel{(*)}{=} 5|x| \cdot \underbrace{\lim_{i \rightarrow \infty} \left| \frac{i+1}{i} \right|}_{=1} = 5|x|.$$

An der mit (\star) gekennzeichneten Stelle wurde benutzt, dass x fest gewählt wurde und der Term $|5x| = 5|x|$ somit konstant ist.

Des Weiteren gilt:

$$5|x| < 1 \iff |x| < \frac{1}{5}$$

$$5|x| > 1 \iff |x| > \frac{1}{5}$$

Folglich liegt Konvergenz vor, falls $|x| < \frac{1}{5}$ gilt; Divergenz liegt vor, falls $|x| > \frac{1}{5}$ gilt. Also gilt $R = \frac{1}{5}$.

Nach diesem Beispiel betrachten wir nun wieder den allgemeinen Fall einer Potenzreihe $\sum_{i=0}^{\infty} a_i x^i$ mit Konvergenzradius $R > 0$. Durch $\sum_{i=0}^{\infty} a_i x^i$ mit $R > 0$ wird (wie zuvor bereits ausgeführt) eine Funktion f mit Definitionsbereich $D(f) = (-R, R)$ definiert, nämlich durch die folgende Festsetzung:

$$f(x) = \sum_{i=0}^{\infty} a_i x^i.$$

Der folgende wichtige Satz, den wir ohne Beweis angeben, beschreibt, wie man eine solche Funktion f differenziert bzw. integriert. Etwas grob gesagt: Man differenziert bzw. integriert f auf dieselbe Art, wie man ein Polynom differenziert bzw. integriert, nämlich durch „gliedweise Anwendung“ der bekannten Regeln:

$$(a_i x^i)' = i a_i x^{i-1} \quad \text{und} \quad \int a_i x^i dx = \frac{1}{i+1} a_i x^{i+1}.$$

Genauer gilt der folgende Satz.

Satz 9.

Die Potenzreihe $f(x) = \sum_{i=0}^{\infty} a_i x^i$ habe den Konvergenzradius $R > 0$. Dann gilt:

(a) Die Funktion f ist in $(-R, R)$ differenzierbar. Ihre Ableitung f' erhält man durch:

$$f'(x) = \sum_{i=1}^{\infty} i a_i x^{i-1}. \quad (4.8)$$

(b) Eine Stammfunktion von f erhält man durch:

$$\int f(x) dx = \sum_{i=0}^{\infty} \frac{a_i}{i+1} x^{i+1}. \quad (4.9)$$

Ferner gilt: Die Potenzreihen (4.8) und (4.9) haben ebenfalls den Konvergenzradius R .

Für den Beweis von Satz 9 verweisen wir auf die Lehrbücher der Analysis.

Durch gliedweise Differentiation bzw. Integration kann man aus gegebenen Potenzreihendarstellungen von Funktionen Potenzreihendarstellungen anderer Funktionen finden.

Beispiel. Aus der geometrischen Reihe

$$\frac{1}{1-t} = \sum_{i=0}^{\infty} t^i \quad \text{für } t \in (-1, 1)$$

erhält man, wenn man $-x$ für t einsetzt, eine Darstellung der Funktion $\frac{1}{1+x}$ als Potenzreihe (für $x \in$

$(-1, 1)$):

$$\frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{i=0}^{\infty} (-x)^i = \sum_{i=0}^{\infty} (-1)^i x^i.$$

Also:

$$\frac{1}{1+x} = \sum_{i=0}^{\infty} (-1)^i x^i \quad \text{für } x \in (-1, 1). \quad (4.10)$$

Durch gliedweise Integration erhält man wegen $(\ln(1+x))' = \frac{1}{1+x}$:

$$\ln(1+x) = \sum_{i=0}^{\infty} \frac{(-1)^i}{i+1} x^{i+1} + c = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i} x^i + c = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} x^i + c.$$

mit einer Integrationskonstanten c . Setzt man hierin $x = 0$ ein, so ergibt sich $c = 0$. Als Ergebnis haben wir also die folgende Potenzreihendarstellung von $\ln(1+x)$ erhalten:

$$\ln(1+x) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} x^i = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \quad (\text{für } x \in (-1, 1)). \quad (4.11)$$

Wie wir wissen, konvergiert die Reihe $x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$ auch noch für $x = 1$ (nach dem Leibnizschen Kriterium). Man kann zeigen, dass auch für $x = 1$ die Formel (4.11) noch gültig ist. Es gilt also:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots \quad (4.11')$$

Für $x = -1$ divergiert die Reihe $x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$, da $-1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{4} - \dots$ bis auf das Vorzeichen die harmonische Reihe ist.

Ein weiteres **Beispiel**: In ähnlicher Weise, in der wir (4.11) gefunden haben, erhält man die Potenzreihendarstellung für $\arctan x$ (im Intervall $(-1, 1)$):

$$\frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = \sum_{i=0}^{\infty} (-x^2)^i = \sum_{i=0}^{\infty} (-1)^i x^{2i}.$$

Durch gliedweise Integration erhält man wegen $(\arctan x)' = \frac{1}{1+x^2}$:

$$\arctan x = \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} x^{2i+1} + c.$$

Durch Einsetzen von $x = 0$ folgt $c = 0$. (Man beachte, dass $\arctan 0 = 0$ gilt.) Also:

$$\arctan x = \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} x^{2i+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots \quad (\text{für } x \in (-1, 1)). \quad (4.12)$$

Nach dem Leibnizschen Kriterium konvergiert die Reihe $x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots$ auch für $x = 1$ und $x = -1$. Man kann zeigen, dass auch für $x = 1$ und $x = -1$ die Formel (4.12) gültig ist. Insbesondere folgt im Fall $x = 1$ (wegen $\arctan 1 = \frac{\pi}{4}$):

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \dots \quad (4.12')$$

Mithilfe der Formeln (4.11') und (4.12') ist es im Prinzip möglich, Näherungswerte für $\ln 2$ und π zu berechnen. (Nachteil: Die Konvergenz ist nur sehr langsam.)

Zusammenfassung der letzten beiden Beispiele.

1. Die Potenzreihe $\sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} x^i = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$ konvergiert für $-1 < x \leq 1$, ansonsten divergiert sie. Im Intervall $(-1, 1]$ stellt diese Reihe die Funktion $f(x) = \ln(1+x)$ dar⁵. Man nennt daher diese Reihe die *Logarithmusreihe*.
2. Die Potenzreihe $\sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} x^{2i+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots$ konvergiert für $-1 \leq x \leq 1$, ansonsten divergiert sie⁶. Im Intervall $[-1, 1]$ stellt diese Reihe die Funktion $f(x) = \arctan x$ dar. Man nennt daher diese Reihe die *Arcustangensreihe*.

Aus Satz 9 ergibt sich, dass die Funktion $f(x) = \sum_{i=0}^{\infty} a_i x^i$ in $(-R, R)$ beliebig oft differenziert werden kann, denn die Ableitung $f'(x) = \sum_{i=1}^{\infty} i a_i x^{i-1}$ ist wieder eine Potenzreihe mit Konvergenzradius R und kann deshalb wiederum gliedweise differenziert werden. Es ergibt sich:

$$\begin{aligned}
 f(x) &= \sum_{i=0}^{\infty} a_i x^i \\
 f'(x) &= \sum_{i=1}^{\infty} i \cdot a_i x^{i-1} \\
 f''(x) &= \sum_{i=2}^{\infty} i \cdot (i-1) \cdot a_i x^{i-2} \\
 f'''(x) &= \sum_{i=3}^{\infty} i \cdot (i-1) \cdot (i-2) \cdot a_i x^{i-3} \\
 &\vdots \\
 \text{allgemein: } f^{(k)}(x) &= \sum_{i=k}^{\infty} i \cdot (i-1) \cdot \dots \cdot (i-k+1) \cdot a_i x^{i-k}
 \end{aligned} \tag{4.13}$$

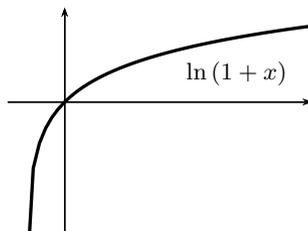
Setzt man in (4.13) speziell $x = 0$ ein, so erhält man:

$$f^{(k)}(0) = k \cdot (k-1) \cdot \dots \cdot 2 \cdot 1 \cdot a_k = k! \cdot a_k.$$

Für die Koeffizienten a_k (mit $k = 0, 1, 2, \dots$) einer Potenzreihe $\sum_{k=0}^{\infty} a_k x^k$ mit $R > 0$ haben wir somit die folgende Darstellung erhalten:

$$a_k = \frac{f^{(k)}(0)}{k!} \quad (\text{für } k = 0, 1, 2, \dots). \tag{4.14}$$

⁵Der Graph von $\ln(1+x)$ ist der um eine Einheit nach links verschobene Graph von $\ln x$:



⁶Für $|x| > 1$ bilden die Glieder dieser Reihe keine Nullfolge - daher liegt Divergenz vor.

Als Folgerung aus (4.14) erhalten wir einen wichtigen Satz, den sogenannten *Identitätssatz für Potenzreihen* (auch *Eindeutigkeitssatz für Potenzreihen* oder *Satz vom Koeffizientenvergleich* genannt).

Satz 10 (Identitätssatz für Potenzreihen).

Sind $\sum_{i=0}^{\infty} a_i x^i$ und $\sum_{i=0}^{\infty} b_i x^i$ zwei Potenzreihen, die beide in einem den Nullpunkt enthaltenden Intervall (a, b) konvergieren und dort dieselbe Funktion $f(x)$ darstellen, so sind die beiden Potenzreihendarstellungen identisch, d.h., für die Koeffizienten gilt $a_i = b_i$ (für $i = 0, 1, 2, \dots$).

Beweis. Es gelte auf dem Intervall (a, b) sowohl $f(x) = \sum_{i=0}^{\infty} a_i x^i$ als auch $f(x) = \sum_{i=0}^{\infty} b_i x^i$. Nach (4.14) folgt $a_i = \frac{f^{(i)}(0)}{i!} = b_i$ (für $i = 0, 1, 2, \dots$). \square

Mit anderen Worten besagt Satz 10 also: Es ist nicht möglich, eine Funktion $f : (a, b) \rightarrow \mathbb{R}$ durch zwei verschiedene Potenzreihen $\sum_{i=0}^{\infty} a_i x^i$ und $\sum_{i=0}^{\infty} b_i x^i$ darzustellen.

Wir haben hier nur Potenzreihen der folgenden Form betrachtet:

$$\sum_{i=0}^{\infty} a_i x^i. \tag{4.15}$$

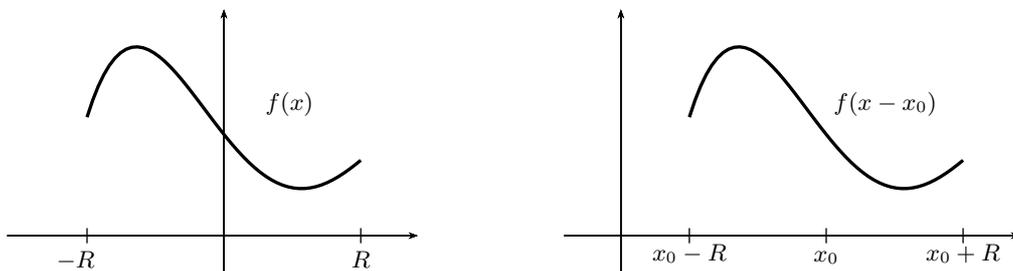
Darüber hinaus bezeichnet man auch Reihen der folgenden Form als Potenzreihen:

$$\sum_{i=0}^{\infty} a_i (x - x_0)^i. \tag{4.16}$$

(Erläuterung: In (4.16) ist x_0 fest und x variabel. (4.15) ist also der spezielle Fall, dass $x_0 = 0$ gilt.) Potenzreihen der Form (4.16) lassen sich durch die einfache Substitution $y = x - x_0$ in Potenzreihen der Form (4.15) überführen:

$$\sum_{i=0}^{\infty} a_i (x - x_0)^i = \sum_{i=0}^{\infty} a_i y^i.$$

Hieraus ergibt sich, dass die Graphen der durch (4.15) und (4.16) dargestellten Funktionen bis auf eine Parallelverschiebung gleich sind: Ist R der Konvergenzradius von (4.15) und ist f die durch (4.15) dargestellte Funktion, so konvergiert die Reihe (4.16) für alle x mit $|x - x_0| < R$; für alle x mit $|x - x_0| > R$ divergiert sie. Ferner ist $f(x - x_0)$ die durch (4.16) bestimmte Funktion⁷. Das bedeutet also: Die Graphen der durch (4.15) und (4.16) dargestellten Funktionen lassen sich durch eine Parallelverschiebung in Richtung der x -Achse ineinander überführen (siehe Zeichnung).



⁷Denn setzen wir wie oben $y = x - x_0$, so gilt:

$$\begin{aligned} |x - x_0| < R &\implies |y| < R \implies \sum_{i=0}^{\infty} a_i y^i = f(y) \implies \sum_{i=0}^{\infty} a_i (x - x_0)^i = f(x - x_0), \\ |x - x_0| > R &\implies |y| > R \implies \sum_{i=0}^{\infty} a_i y^i \text{ divergiert} \implies \sum_{i=0}^{\infty} a_i (x - x_0)^i \text{ divergiert.} \end{aligned}$$

Man nennt x_0 den *Mittelpunkt* (oder auch *Entwicklungspunkt*) der Potenzreihe $\sum_{i=0}^{\infty} a_i(x-x_0)^i$. Unsere obigen Überlegungen haben gezeigt, dass Potenzreihen mit beliebigem Mittelpunkt x_0 „nicht wesentlich verschieden“ von Potenzreihen mit Mittelpunkt $x_0 = 0$ sind. Daher konnten wir uns auf das Studium von Potenzreihen mit Mittelpunkt $x_0 = 0$ beschränken. Für Potenzreihen mit beliebigem Mittelpunkt x_0 gelten entsprechende Sätze wie für Potenzreihen mit Mittelpunkt $x_0 = 0$.

4.3 Der Satz von Taylor

Ist die Funktion f an der Stelle x_0 differenzierbar, so erhält man die Gleichung der Tangente $T(x)$ im Punkt $(x_0, f(x_0))$ durch (vgl. Abschnitt 2.1, Bemerkung zur geometrischen Interpretation der Ableitung):

$$T(x) = f(x_0) + f'(x_0)(x - x_0).$$

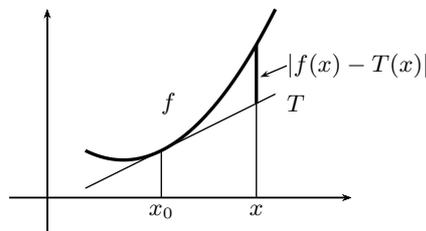
Unsere Anschauung legt nahe, dass $T(x)$ in der Nähe von x_0 eine gute Annäherung an $f(x)$ darstellt. Diese Aussage lässt sich folgendermaßen präzisieren: Wir fragen uns, welchen Fehler man begeht, wenn man $f(x)$ durch $T(x)$ ersetzt. Für den Fehler $|f(x) - T(x)|$ gilt:

$$|f(x) - T(x)| = |f(x) - f(x_0) - f'(x_0)(x - x_0)|.$$

Es folgt (für $x \neq x_0$):

$$\frac{|f(x) - T(x)|}{|x - x_0|} = \left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| \rightarrow 0 \quad \text{für } x \rightarrow x_0.$$

Das bedeutet, der Fehler $|f(x) - T(x)|$ *strebt schneller gegen Null als der Abstand* $|x - x_0|$. Diese Eigenschaft unterscheidet die Tangente $T(x)$ von jeder anderen Geraden durch den Punkt $(x_0, f(x_0))$. Dies ist die präzise Bedeutung der Feststellung, dass $T(x)$ in der Nähe von x_0 eine gute Annäherung an $f(x)$ darstellt.



Ferner gilt für die Funktion f und die Tangente T im Punkt $(x_0, f(x_0))$:

$$T(x_0) = f(x_0) \quad \text{und} \quad T'(x_0) = f'(x_0). \quad (4.17)$$

Mit anderen Worten: T und f stimmen an der Stelle x_0 in Funktionswert und erster Ableitung überein. Für viele Anwendungen reicht die Annäherung von f durch eine Gerade nicht aus. Um noch bessere Annäherungen von f in einer Umgebung von x_0 zu erhalten, kann man f beispielsweise durch ein Polynom höchstens zweiten Grades annähern:

$$T_2(x) = ax^2 + bx + c.$$

Unsere obigen Überlegungen legen uns nahe, dass $T_2(x)$ in der Umgebung von x_0 eine gute Annäherung an $f(x)$ darstellt, falls $T_2(x)$ so gewählt wird, dass (wie in (4.17)) $T_2(x_0) = f(x_0)$, $T_2'(x_0) = f'(x_0)$

und zusätzlich noch $T''(x_0) = f''(x_0)$ gilt. Allgemein: Man nähert f durch ein Polynom höchstens n -ten Grades $T_n(x)$ an, für das gilt:

$$T_n(x_0) = f(x_0), \quad T'_n(x_0) = f'(x_0), \quad \dots \quad T_n^{(n)}(x_0) = f^{(n)}(x_0). \quad (4.17')$$

Mit anderen Worten: (4.17') besagt, dass T_n und f an der Stelle x_0 in Funktionswert und den ersten n Ableitungen übereinstimmen ($n \in \mathbb{N} \cup \{0\}$). Dies setzt natürlich voraus, dass f n -mal differenzierbar ist. Es stellt sich nun die Frage, ob ein solches Polynom überhaupt immer existiert. Die Antwort lautet ja.

Das folgende Polynom erfüllt offenbar diese Bedingungen⁸:

$$\sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k. \quad (4.18)$$

Beispiel. Für $n = 3$ lautet dieses Polynom:

$$\sum_{k=0}^3 \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \frac{f'''(x_0)}{3!}(x - x_0)^3.$$

Man mache sich klar, dass der Funktionswert dieses Polynoms an der Stelle x_0 gerade $f(x_0)$ ist und dass die ersten drei Ableitungen dieses Polynoms an der Stelle x_0 die Werte $f'(x_0)$, $f''(x_0)$ und $f'''(x_0)$ besitzen!

Man nennt das Polynom (4.18) *n-tes Taylorpolynom von f an der Stelle x_0* . Wir bezeichnen es mit $T_n(x)$ oder (wenn wir die Abhängigkeit von der Stelle x_0 ausdrücken wollen) mit $T_n(x, x_0)$.

Um unsere Überlegungen möglichst übersichtlich zu halten, wollen wir ab jetzt $x_0 = 0$ voraussetzen. Dies bedeutet keine wesentliche Einschränkung, da man den allgemeinen Fall eines beliebigen x_0 ebenso wie den Fall $x_0 = 0$ behandeln kann. Das n -te Taylorpolynom von f an der Stelle $x_0 = 0$ lautet also:

$$T_n(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n. \quad (4.19)$$

Welchen Fehler begeht man, wenn man $f(x)$ durch $T_n(x)$ ersetzt? Wir setzen $f(x) = T_n(x) + R_n(x)$. Man nennt $R_n(x)$ *n-tes Restglied*. $|R_n(x)| = |f(x) - T_n(x)|$ ist der Fehler, den man begeht, wenn man $f(x)$ durch $T_n(x)$ ersetzt. Grundlage für die Abschätzung des Fehlers $|R_n(x)|$ ist der folgende *Satz von Taylor*.

Satz 11 (Satz von Taylor).

Die Funktion f sei im Intervall $[a, b]$ $(n + 1)$ -mal differenzierbar und die $(n + 1)$ -te Ableitung sei stetig auf $[a, b]$. Es gelte $0 \in (a, b)$. Für alle $x \in [a, b]$ gilt dann:

$$f(x) = T_n(x) + R_n(x), \quad (4.20)$$

wobei $T_n(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k$ das n -te Taylorpolynom (an der Stelle $x_0 = 0$) ist und das Restglied $R_n(x)$ die folgende Darstellung besitzt:

$$R_n(x) = \frac{1}{n!} \int_0^x (x - t)^n f^{(n+1)}(t) dt.$$

⁸Es lässt sich zeigen, dass dies das eindeutig bestimmte Polynom höchstens n -ten Grades ist, das die Bedingungen (4.17') erfüllt.

Beweis. (Induktion nach n)

(I) Für $n = 0$ gilt $T_n(x) = f(0)$ sowie

$$\frac{1}{n!} \int_0^x (x-t)^n f^{(n+1)}(t) dt = \int_0^x f'(t) dt = f(x) - f(0).$$

Die Behauptung (4.20) lautet für $n = 0$ demnach: $f(x) = f(0) + f(x) - f(0)$, was zweifellos richtig ist.

(II) Wir nehmen an, dass die Behauptung für ein festes $n \geq 0$ gelte, d.h.:

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + \frac{1}{n!} \int_0^x (x-t)^n f^{(n+1)}(t) dt. \quad (4.21)$$

Mittels partieller Integration folgt hieraus:

$$\begin{aligned} & \frac{1}{n!} \int_0^x (x-t)^n f^{(n+1)}(t) dt \\ &= \frac{1}{n!} \left[-\frac{(x-t)^{n+1}}{n+1} f^{(n+1)}(t) \right]_0^x + \frac{1}{n!} \frac{1}{n+1} \int_0^x (x-t)^{n+1} f^{(n+2)}(t) dt \\ &= \frac{f^{(n+1)}(0)}{(n+1)!} x^{n+1} + \frac{1}{(n+1)!} \int_0^x (x-t)^{n+1} f^{(n+2)}(t) dt \end{aligned}$$

Hieraus erhält man durch Einsetzen in (4.21), dass die Behauptung auch für $n+1$ gilt. \square

Es sei angemerkt, dass es auch andere Darstellungen des Restglieds gibt; die hier gewählte Darstellung

$$R_n(x) = \frac{1}{n!} \int_0^x (x-t)^n f^{(n+1)}(t) dt$$

nennt man die *Integralform* des Restglieds.

Abschätzung des Restglieds $R_n(x)$.

Da die Funktion $f^{(n+1)}$ in $[a, b]$ stetig ist, gibt es eine Konstante M_{n+1} , so dass $|f^{(n+1)}(t)| \leq M_{n+1}$ für alle $t \in [a, b]$ gilt⁹. Es sei $x \in [a, b]$. Wir betrachten zunächst den Fall, dass $x \geq 0$ gilt. Dann folgt:

$$\begin{aligned} |R_n(x)| &\leq \frac{1}{n!} \int_0^x (x-t)^n |f^{(n+1)}(t)| dt \\ &\leq \frac{M_{n+1}}{n!} \int_0^x (x-t)^n dt = \frac{M_{n+1}}{n!} \left[-\frac{(x-t)^{n+1}}{n+1} \right]_0^x = \frac{M_{n+1}}{(n+1)!} x^{n+1} \end{aligned}$$

(Die erste Ungleichung folgt nach Satz 5 (b) in Abschnitt 3.2, man beachte auch $x \geq t$; die zweite Ungleichung folgt nach Satz 5 (a) in Abschnitt 3.2.)

Gilt $x < 0$, so erhält man durch eine ähnliche Rechnung:

$$|R_n(x)| \leq \frac{M_{n+1}}{(n+1)!} (-x)^{n+1}.$$

⁹Ein bekannter Satz über stetige Funktionen besagt nämlich: Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist beschränkt. (Hierbei ist die Voraussetzung wichtig, dass das betrachtete Intervall $[a, b]$ abgeschlossen ist.)

Insgesamt haben wir also folgendes Ergebnis erhalten (für alle $x \in [a, b]$):

Abschätzung des Restglieds $R_n(x)$

$$\left| f(x) - T_n(x) \right| = \left| R_n(x) \right| \leq \frac{M_{n+1}}{(n+1)!} |x|^{n+1}$$

(4.22)

Insbesondere folgt aus (4.22) (für $x \rightarrow 0$):

$$0 \leq \frac{|f(x) - T_n(x)|}{|x|^n} \leq \frac{M_{n+1} \cdot |x|}{(n+1)!} \rightarrow 0.$$

Daraus folgt:

Für $x \rightarrow 0$ strebt der Fehler $|R_n(x)| = |f(x) - T_n(x)|$ sogar noch schneller gegen Null als $|x|^n$. (4.23)

Man kann zeigen, dass unter allen Polynomen $P_n(x)$ vom Grad höchstens n , für die $P_n(0) = f(0)$ gilt, das Taylorpolynom $T_n(x)$ das einzige ist, das die Eigenschaft (4.23) besitzt. In diesem Sinne ist das Taylorpolynom unter allen Polynomen $P_n(x)$ am besten zur näherungsweisen Berechnung von f in der Nähe von $x_0 = 0$ geeignet. (Entsprechendes gilt für jedes beliebige x_0 und das Taylorpolynom $T_n(x, x_0)$.) Man beachte, dass durch diese Feststellungen die Überlegungen vom Anfang dieses Abschnitts verallgemeinert werden.

Beispiele:

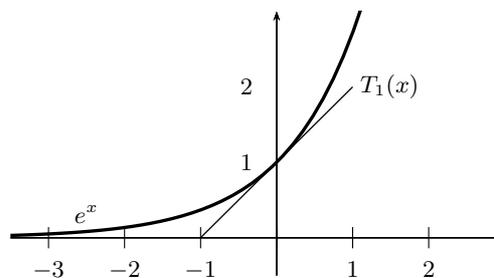
1. Es sei $f(x) = e^x$. Die Taylorpolynome von f an der Stelle $x_0 = 0$ lauten dann:

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= 1 + x \\ T_2(x) &= 1 + x + \frac{1}{2}x^2 \\ T_3(x) &= 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 \\ &\vdots \\ T_n(x) &= \sum_{k=0}^n \frac{x^k}{k!} \end{aligned}$$

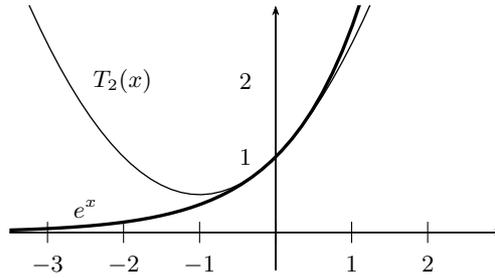
Nach (4.22) gilt für $x \in [-1, 1]$:

$$\left| R_n(x) \right| = \left| e^x - T_n(x) \right| \leq \frac{e}{(n+1)!} |x|^{n+1} \leq \frac{e}{(n+1)!} < \frac{3}{(n+1)!}.$$

Approximieren wir die Exponentialfunktion e^x im Intervall $[-1, 1]$ durch die Gerade $T_1(x) = 1 + x$ (siehe Zeichnung), so beträgt der Fehler $|e^x - T_1(x)|$ also höchstens $\frac{3}{2} = 1.5$.



Approximieren wir e^x in $[-1, 1]$ durch die quadratische Funktion (Parabel) $T_2(x) = 1 + x + \frac{1}{2}x^2$ (siehe Zeichnung), so beträgt der Fehler $|e^x - T_2(x)|$ höchstens $\frac{3}{6} = 0.5$.



Für die nächsten Taylorpolynome erhält man:

$$\begin{aligned} |e^x - T_3(x)| &\leq \frac{3}{24} = 0.125 \\ |e^x - T_4(x)| &\leq \frac{3}{120} = 0.025 \\ |e^x - T_5(x)| &\leq \frac{3}{720} \approx 0.0042 \\ |e^x - T_6(x)| &\leq \frac{3}{5040} \approx 0.0006 \end{aligned}$$

2. Es sei $f(x) = \sin x$.

Es gilt:

$$f'(x) = \cos x, \quad f''(x) = -\sin x, \quad f^{(3)}(x) = -\cos x, \quad f^{(4)}(x) = \sin x \quad \dots$$

Ferner: $\sin(0) = 0$ und $\cos(0) = 1$.

Hieraus ergibt sich:

$$\begin{aligned} T_0(x) &= 0 \\ T_1(x) &= T_2(x) = x \\ T_3(x) &= T_4(x) = x - \frac{x^3}{3!} \\ T_5(x) &= T_6(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} \\ T_7(x) &= T_8(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \\ &\vdots \\ T_{2n+1}(x) &= T_{2n+2}(x) = \sum_{k=0}^n \frac{(-1)^k}{(2k+1)!} x^{2k+1} \end{aligned}$$

Es gilt $|f^{(n)}(x)| \leq 1$, da sämtliche Ableitungen von $f(x) = \sin x$ entweder Sinus- oder Cosinusfunktionen sind. Also folgt nach (4.22) für alle $x \in \mathbb{R}$:

$$|\sin x - T_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

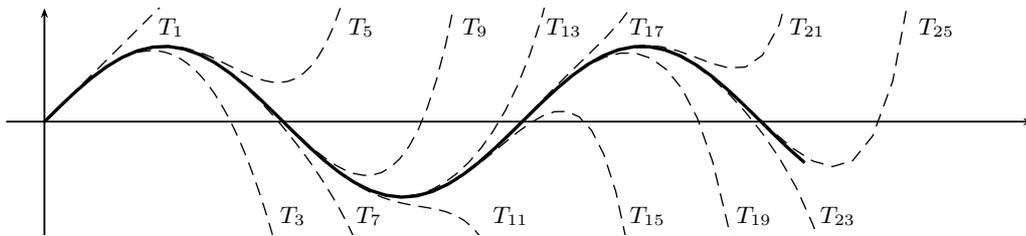
Wir können also abschätzen:

$$\begin{aligned} |\sin x - T_2(x)| &= |\sin x - x| \leq \frac{|x|^3}{3!} = \frac{|x|^3}{6} \\ |\sin x - T_4(x)| &= \left| \sin x - \left(x - \frac{x^3}{3!} \right) \right| \leq \frac{|x|^5}{5!} = \frac{|x|^5}{120} \end{aligned}$$

Hieraus ergeben sich beispielsweise die folgenden Fehlerabschätzungen:

$$\begin{aligned} |\sin x - x| &\leq \frac{1}{6} && \text{(für } |x| \leq 1) \\ |\sin x - x| &\leq \frac{1}{6.000} && \left(\text{für } |x| \leq \frac{1}{10} \right) \\ \left| \sin x - \left(x - \frac{x^3}{3!} \right) \right| &\leq \frac{1}{120} && \text{(für } |x| \leq 1) \\ \left| \sin x - \left(x - \frac{x^3}{3!} \right) \right| &\leq \frac{1}{12.000.000} && \left(\text{für } |x| \leq \frac{1}{10} \right) \end{aligned}$$

In der folgenden Zeichnung sind die Taylorpolynome T_1 bis T_{25} für $\sin x$ dargestellt:



Der Vollständigkeit halber formulieren wir noch den Satz von Taylor für ein beliebiges x_0 . (Der Beweis dieses Satzes geht analog zum Beweis von Satz 11.)

Satz 12 (Satz von Taylor in allgemeiner Form).

Die Funktion f sei im Intervall $[a, b]$ $(n + 1)$ -mal differenzierbar und die $(n + 1)$ -te Ableitung sei stetig auf $[a, b]$. Es gelte $x_0 \in (a, b)$. Dann folgt für alle $x \in [a, b]$:

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

Dabei ist $T_n(x) = \sum_{k=0}^n \left(\frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \right)$ das n -te Taylorpolynom (an der Stelle x_0) und für das Restglied gilt:

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

Die Formel für die Abschätzung des Restglieds lautet für ein beliebiges x_0 :

$$\left| f(x) - T_n(x) \right| = \left| R_n(x) \right| \leq \frac{M_{n+1}}{(n + 1)!} |x - x_0|^{n+1}. \quad (4.22')$$

M_{n+1} ist wie zuvor eine Konstante, für die $M_{n+1} \geq |f^{(n+1)}(t)|$ für alle $t \in [a, b]$ gilt.

4.4 Taylorreihen

Definition.

Die Funktion f sei im Intervall $[a, b]$ beliebig oft differenzierbar. Es gelte $0 \in (a, b)$. Dann heißt die Potenzreihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \quad (4.24)$$

die *Taylorreihe* von f . (Genauer müsste man sagen: Taylorreihe mit Entwicklungspunkt $x_0 = 0$. Statt „Taylorreihe“ sagt man auch „Taylor-Entwicklung“.)

Die Partialsummen der Taylorreihe sind also gerade die Taylorpolynome:

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

Da die Taylorreihe eine Potenzreihe ist, konvergiert sie in einem Intervall $(-R, R)$ für eine reelle Zahl $R \geq 0$ oder $R = \infty$ und divergiert für alle x mit $|x| > R$ (siehe Satz 8, Abschnitt 4.2). Konvergiert die Taylorreihe (4.24) für ein x , so braucht sie nicht notwendig gegen $f(x)$ zu konvergieren, d.h., die folgende Gleichheit muss nicht unbedingt gelten:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = f(x). \quad (4.25)$$

Man kann Beispiele konstruieren, in denen $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \neq f(x)$ gilt (siehe Bemerkung (1) weiter unten). Der nächste Satz gibt darüber Auskunft, wann (4.25) gilt.

Satz 13.

Die Funktion f sei in $[a, b]$ beliebig oft differenzierbar und es gelte $0 \in (a, b)$. Wir betrachten ein festes $x \in [a, b]$. Dann gilt

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

genau dann, wenn für das Restglied R_n aus dem Satz von Taylor gilt:

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

Beweis. Es gilt $R_n(x) = f(x) - T_n(x)$. Also ist $R_n(x) \rightarrow 0$ gleichbedeutend mit $T_n(x) \rightarrow f(x)$. \square

Um nachzuweisen, dass für eine Funktion f innerhalb eines bestimmten Intervalls die Gleichheit (4.25) gilt, muss man also zeigen, dass das Restglied $R_n(x)$ der Taylorschen Formel¹⁰ gegen Null strebt. Dies erfordert in manchen Fällen komplizierte Rechnungen, in anderen Fällen ist der Nachweis jedoch einfach.

Die Taylorreihen einiger Funktionen

1. $f(x) = e^x$:

Es gilt $f^{(k)}(0) = 1$. Also lautet die Taylor-Reihe

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$$

Wir zeigen, dass diese Reihe für jedes $x \in \mathbb{R}$ gegen $f(x) = e^x$ konvergiert (vgl. auch (4.7)).

¹⁰Unter der Taylorschen Formel versteht man die Identität (4.20) bzw. die entsprechende Formel aus Satz 12.

Es sei $x \in \mathbb{R}$ fest. Wir wählen das Intervall $[a, b]$ so, dass es x enthält und dass $0 \in (a, b)$ gilt. Es gilt für alle $n = 0, 1, \dots$ und für alle $t \in [a, b]$:

$$\left| f^{(n+1)}(t) \right| = e^t \leq e^b.$$

Nach (4.22) gilt (für alle $n = 0, 1, \dots$):

$$\left| R_n(x) \right| \leq \frac{e^b}{(n+1)!} \cdot |x|^{n+1}.$$

Also:

$$0 \leq \lim_{n \rightarrow \infty} \left| R_n(x) \right| \leq \lim_{n \rightarrow \infty} \left(\frac{e^b}{(n+1)!} |x|^{n+1} \right) = e^b \cdot \lim_{n \rightarrow \infty} \left(\frac{|x|^{n+1}}{(n+1)!} \right) = 0.$$

(Vgl. auch Beispiel (e), Abschnitt 1.3). Also gilt $\lim_{n \rightarrow \infty} \left| R_n(x) \right| = 0$ und somit nach Satz 13:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots \quad (\text{für alle } x \in \mathbb{R}). \quad (4.26)$$

Wir haben als Resultat die Potenzreihendarstellung von e^x gewonnen. Man nennt die Reihe (4.26) die *Exponentialreihe*. Wir erinnern daran, dass nach Satz 10 (Identitätssatz für Potenzreihen) diese Darstellung eindeutig ist¹¹.

2. $f(x) = \sin x$:

Es gilt $f^{(2k)}(0) = 0$ und $f^{(2k+1)}(0) = (-1)^k$ (mit $k = 0, 1, 2, \dots$). Also lautet die Taylor-Reihe:

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

Wir zeigen, dass diese Reihe für jedes $x \in \mathbb{R}$ gegen $f(x) = \sin x$ konvergiert: Es sei $x \in \mathbb{R}$ fest und das Intervall $[a, b]$ sei wie in 1. gewählt. Es gilt für alle $n = 0, 1, \dots$ und für alle $t \in [a, b]$:

$$\left| f^{(n+1)}(t) \right| \leq 1.$$

Nach (4.22) gilt also (für alle $n = 0, 1, \dots$):

$$\left| R_n(x) \right| \leq \frac{1}{(n+1)!} \cdot |x|^{n+1}.$$

Wie in 1. folgt $\lim_{n \rightarrow \infty} \left| R_n(x) \right| = 0$. Also gilt nach Satz 13:

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (\text{für alle } x \in \mathbb{R}). \quad (4.27)$$

Man nennt die Reihe (4.27) die *Sinusreihe*.

3. $f(x) = \cos x$:

Durch eine ähnliche Überlegung wie in 2. oder (etwas eleganter) durch gliedweise Differentiation von (4.27) erhält man:

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad (\text{für alle } x \in \mathbb{R}). \quad (4.28)$$

Man nennt die Reihe (4.28) die *Cosinusreihe*.

¹¹Dasselbe gilt natürlich auch für die im Folgenden behandelten Reihendarstellungen von $\sin x$, $\cos x$, etc.

4. $f(x) = \ln(1+x)$:

Diese Reihe wurde bereits mit der Methode der Reihenintegration gefunden (vgl. (4.11)):

$$\ln(1+x) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} x^i = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \quad (\text{für } x \in (-1, 1]). \quad (4.29)$$

Man nennt die Reihe (4.29) die *Logarithmusreihe* (vgl. Abschnitt 4.2).

5. $f(x) = \arctan x$:

Diese Reihe wurde bereits mit der Methode der Reihenintegration gefunden (vgl. (4.12)):

$$\arctan x = \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} x^{2i+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots \quad (\text{für } x \in [-1, 1]). \quad (4.30)$$

Man nennt die Reihe (4.30) die *Arcustangensreihe* (vgl. Abschnitt 4.2).

Potenzreihendarstellungen von weiteren Funktionen findet man in den Lehrbüchern der Analysis.

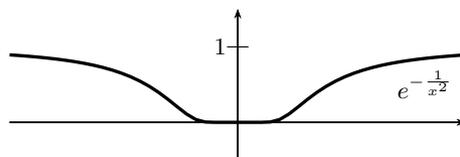
Zwei Bemerkungen zum Zusammenhang der Begriffe „Potenzreihe“ und „Taylorreihe“

- (1) Die Taylor-Reihe $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ einer im Nullpunkt beliebig oft differenzierbaren Funktion f ist eine Potenzreihe; als solche stellt die Taylorreihe auf einem Intervall $(-R, R)$ eine Funktion dar. In vielen Fällen ist diese Funktion gerade wieder die Ausgangsfunktion $f(x)$ (siehe etwa die obigen Beispiele $f(x) = e^x$, $f(x) = \sin x$ oder $f(x) = \cos x$). Es ist aber auch möglich, dass die Taylorreihe $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ auf ihrem Konvergenzintervall eine von $f(x)$ verschiedene Funktion darstellt; hierfür geben wir nun ein Beispiel.

Beispiel. Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ sei definiert durch

$$f(x) := \begin{cases} e^{-\frac{1}{x^2}}, & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases}$$

Der Graph dieser Funktion:



Für diese Funktion gilt¹² $f^{(k)}(0) = 0$ für alle $k = 0, 1, 2, \dots$. Für die Taylorreihe von f gilt also:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = 0 \quad (\text{für alle } x \in \mathbb{R}).$$

Dies bedeutet: Die Taylorreihe von f konvergiert auf ganz \mathbb{R} ; sie stellt dort aber keineswegs die Ausgangsfunktion $f(x)$ dar, sondern die konstante Funktion, die überall den Wert Null annimmt.

- (2) Es sei nun f eine Funktion, die in Form einer Potenzreihe vorliegt:

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \quad \text{für } x \in (-R, R) \text{ mit } R > 0.$$

¹²Auf den Beweis dieser Aussage, der nicht schwierig ist, aber etwas längere Rechnungen erfordert, wollen wir verzichten.

Wie sieht dann die Taylorreihe von f aus?

Antwort: Nach (4.14) gilt $a_k = \frac{f^{(k)}(0)}{k!}$ (für $k = 0, 1, 2, \dots$). Für die Taylorreihe $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ von f gilt also:

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

Das bedeutet, dass die Taylorreihe gleich der Potenzreihe ist, die f darstellt.

Beispiele hierzu: Wir haben die Potenzreihendarstellungen von $\ln(1+x)$ und $\arctan x$ für $x \in (-1, 1)$ durch gliedweise Integration aus der geometrischen Reihe gewonnen; nach obiger Bemerkung (2) können wir sicher sein, dass diese Potenzreihen gerade die Taylorreihen von $\ln(1+x)$ und $\arctan x$ sind.

Zur Übung: Man prüfe dies für $\ln(1+x)$ direkt nach! Anleitung: Man berechne für $f(x) = \ln(1+x)$ die Werte $f^{(k)}(0)$ (für $k = 0, 1, 2, \dots$) und stelle die Taylorreihe auf. Danach vergleiche man das Ergebnis mit (4.29).

5 Differentialrechnung für Funktionen mehrerer Variablen

5.1 Partielle Ableitungen

Zunächst betrachten wir nur Funktionen mit 2 Variablen. Für $D(f) \subseteq \mathbb{R}^2$ sei eine Funktion

$$\begin{aligned} f : D(f) &\rightarrow \mathbb{R} \\ (x, y) &\mapsto f(x, y) = z \end{aligned}$$

gegeben und $(x_0, y_0) \in D(f)$ sei ein fester Punkt.

Wir definieren eine Funktion φ durch

$$\varphi(x) = f(x, y_0).$$

φ ist eine Funktion der einen Variablen x , der Definitionsbereich von φ ist

$$D(\varphi) = \{x : (x, y_0) \in D(f)\}.$$

Wir können die Funktion φ auf Differenzierbarkeit an der Stelle x_0 untersuchen, d.h., wir fragen nach der Existenz des folgenden Grenzwerts:

$$\varphi'(x_0) = \frac{d\varphi}{dx}(x_0) = \lim_{h \rightarrow 0} \left(\frac{\varphi(x_0 + h) - \varphi(x_0)}{h} \right) = \lim_{h \rightarrow 0} \left(\frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \right).$$

Wenn der Grenzwert existiert, sagt man, die Funktion $f(x, y)$ ist an der Stelle (x_0, y_0) *partiell differenzierbar nach der Variablen x* .

Analog: Die Funktion $\psi(y) = f(x_0, y)$ mit $D(\psi) = \{y : (x_0, y) \in D(f)\}$ ist eine Funktion der einen Variablen y . Wenn der Grenzwert

$$\psi'(y_0) = \frac{d\psi}{dy}(y_0) = \lim_{h \rightarrow 0} \left(\frac{\psi(y_0 + h) - \psi(y_0)}{h} \right) = \lim_{h \rightarrow 0} \left(\frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} \right)$$

existiert, sagt man, die Funktion f ist an der Stelle (x_0, y_0) *partiell differenzierbar nach der Variablen y* . Wir fassen zusammen:

Definition.

Die Funktion f heißt an der Stelle $(x_0, y_0) \in D(f)$

(a) *partiell differenzierbar* nach der Variablen x , wenn der Grenzwert $\lim_{h \rightarrow 0} \left(\frac{f(x_0+h, y_0) - f(x_0, y_0)}{h} \right)$ existiert; man bezeichnet ihn mit

$$\frac{\partial f}{\partial x}(x_0, y_0) \quad \text{oder} \quad f_x(x_0, y_0).$$

(b) *partiell differenzierbar* nach der Variablen y , wenn der Grenzwert $\lim_{h \rightarrow 0} \left(\frac{f(x_0, y_0+h) - f(x_0, y_0)}{h} \right)$ existiert; man bezeichnet ihn mit

$$\frac{\partial f}{\partial y}(x_0, y_0) \quad \text{oder} \quad f_y(x_0, y_0).$$

Man nennt $\frac{\partial f}{\partial x}(x_0, y_0)$ die *partielle Ableitung von f nach x an der Stelle (x_0, y_0)*; analog heißt $\frac{\partial f}{\partial y}(x_0, y_0)$ die *partielle Ableitung von f nach y an der Stelle (x_0, y_0)*.

Beispiele:

1. Es sei $f(x, y) = x^3y - 2x^2y^4$ mit $D(f) = \mathbb{R}^2$.

An einer Stelle $(x, y) \in \mathbb{R}^2$ erhalten wir:

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= f_x(x, y) = 3x^2y - 4xy^4 \\ \frac{\partial f}{\partial y}(x, y) &= f_y(x, y) = x^3 - 8x^2y^3\end{aligned}$$

2. Es sei $f(x, y) = x \cdot e^{x+y^2}$ mit $D(f) = \mathbb{R}^2$.

An einer Stelle $(x, y) \in \mathbb{R}^2$ erhalten wir:

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= f_x(x, y) = e^{x+y^2} + x \cdot e^{x+y^2} = e^{x+y^2} \cdot (x+1) \\ \frac{\partial f}{\partial y}(x, y) &= f_y(x, y) = e^{x+y^2} \cdot 2xy\end{aligned}$$

3. Es sei $f(x, y) = \ln(xy)$ mit $D(f) = \{(x, y) : xy > 0\}$.

An einer Stelle $(x, y) \in D(f)$ erhalten wir:

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= f_x(x, y) = \frac{1}{xy} \cdot y = \frac{1}{x} \\ \frac{\partial f}{\partial y}(x, y) &= f_y(x, y) = \frac{1}{xy} \cdot x = \frac{1}{y}\end{aligned}$$

Die bisher betrachteten partiellen Ableitungen werden auch als *partielle Ableitungen 1. Ordnung* bezeichnet. Wenn man diese Funktionen erneut partiell differenziert, gewinnt man die *partiellen Ableitungen höherer Ordnung*.

Bezeichnungen für die partiellen Ableitungen 2. Ordnung:

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &:= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) \quad \text{oder} \quad f_{xx} := (f_x)_x \\ \frac{\partial^2 f}{\partial y \partial x} &:= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) \quad \text{oder} \quad f_{xy} := (f_x)_y \\ \frac{\partial^2 f}{\partial x \partial y} &:= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \quad \text{oder} \quad f_{yx} := (f_y)_x \\ \frac{\partial^2 f}{\partial y^2} &:= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) \quad \text{oder} \quad f_{yy} := (f_y)_y\end{aligned}$$

Analoge Bezeichnungen verwendet man für die partiellen Ableitungen k -ter Ordnung ($k \geq 3$).

Beispiel. Für $f(x, y) = x^3y + xe^y$ mit $D(f) = \mathbb{R}^2$ gilt:

$$\begin{aligned}f_x &= 3x^2y + e^y & f_y &= x^3 + xe^y & (\text{partielle Ableitungen 1. Ordnung}) \\ f_{xx} &= 6xy & f_{yx} &= 3x^2 + e^y & (\text{partielle Ableitungen 2. Ordnung}) \\ f_{xy} &= 3x^2 + e^y & f_{yy} &= xe^y \\ f_{xxx} &= 6y & f_{yxx} &= 6x & (\text{partielle Ableitungen 3. Ordnung}) \\ f_{xxy} &= 6x & f_{yyx} &= e^y \\ f_{xyx} &= 6x & f_{yyy} &= e^y \\ f_{xyy} &= e^y & f_{yyy} &= xe^y\end{aligned}$$

In diesem Beispiel kann die Reihenfolge, in der die partiellen Ableitungen vorgenommen werden, vertauscht werden; es gilt beispielsweise $f_{yx} = f_{xy}$. In vielen anderen Beispielen ist dies ebenso, aber immer gilt es nicht, wie folgendes Beispiel zeigt:

Beispiel. Für $f(x, y) = \begin{cases} xy \cdot \frac{x^2 - y^2}{x^2 + y^2} & , \text{ falls } (x, y) \neq (0, 0) \\ 0 & , \text{ falls } (x, y) = (0, 0) \end{cases}$ und $D(f) = \mathbb{R}^2$ gilt:

$$f_x(0, 0) = \lim_{h \rightarrow 0} \left(\frac{f(h, 0) - f(0, 0)}{h} \right) = 0$$

$$f_y(0, 0) = \lim_{h \rightarrow 0} \left(\frac{f(0, h) - f(0, 0)}{h} \right) = 0$$

Für $(x, y) \neq (0, 0)$ gilt:

$$f_x(x, y) = y \cdot \frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2}$$

$$f_y(x, y) = x \cdot \frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}$$

Hieraus erhält man:

$$f_{xy}(0, 0) = \lim_{h \rightarrow 0} \left(\frac{1}{h} \left(h \cdot \frac{-h^4}{(h^2)^2} \right) \right) = -1$$

$$f_{yx}(0, 0) = \lim_{h \rightarrow 0} \left(\frac{1}{h} \left(h \cdot \frac{h^4}{(h^2)^2} \right) \right) = +1$$

Definition.

(a) Es sei $(x_0, y_0) \in \mathbb{R}^2$ und $\delta \in \mathbb{R}$, $\delta > 0$. Dann nennen wir die Menge

$$U_\delta(x_0, y_0) := \left\{ (x, y) : |x - x_0| < \delta, |y - y_0| < \delta \right\}$$

eine (*offene*) δ -Umgebung von (x_0, y_0) .

(b) Es sei $(x_0, y_0) \in D \subseteq \mathbb{R}^2$. Man nennt (x_0, y_0) einen *inneren Punkt* von D , falls es ein $\delta > 0$ gibt, so dass $U_\delta(x_0, y_0) \subseteq D$ gilt.

Ohne Beweis geben wir den folgenden Satz an.

Satz 1 (Schwarz).

Für $D \subseteq \mathbb{R}^2$ sei $f : D \rightarrow \mathbb{R}$ eine Funktion und (x_0, y_0) ein innerer Punkt von D . Existieren für ein $k \geq 2$ auf einer δ -Umgebung $U_\delta(x_0, y_0)$ sämtliche partiellen Ableitungen der Ordnung $\leq k$ und sind diese auf $U_\delta(x_0, y_0)$ stetig, so ist auf $U_\delta(x_0, y_0)$ für die partiellen Ableitungen k -ter Ordnung die Reihenfolge der Differentiation vertauschbar.

Zum Begriff der Stetigkeit für Funktionen zweier Variablen: siehe Abschnitt 5.2.

Für Funktionen $f(x_1, x_2, \dots, x_n)$ mit n Variablen definiert man die partiellen Ableitungen analog zum Fall zweier Variablen.

Definition.

Für $D(f) \subseteq \mathbb{R}^n$ sei eine Funktion $f : D(f) \rightarrow \mathbb{R}$ gegeben und $(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}) \in D(f)$ sei ein fester Punkt. Die Funktion f heißt an der Stelle $(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$ *partiell nach der Variablen x_k differenzierbar*, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \left(\frac{f(x_1^{(0)}, \dots, x_{k-1}^{(0)}, x_k^{(0)} + h, x_{k+1}^{(0)}, \dots, x_n^{(0)}) - f(x_1^{(0)}, \dots, x_{k-1}^{(0)}, x_k^{(0)}, x_{k+1}^{(0)}, \dots, x_n^{(0)})}{h} \right)$$

existiert. Man bezeichnet ihn mit $\frac{\partial f}{\partial x_k}(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$ oder $f_{x_k}(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$.

Der Satz von Schwarz gilt entsprechend auch für Funktionen mit n Variablen (siehe Satz 1' am Ende von Abschnitt 5.2).

5.2 Extremstellen, Gradient und Hessesche Matrix

Im Folgenden sei mit $D = D(f)$ immer der Definitionsbereich der betrachteten Funktion f bezeichnet und es gelte stets $D \subseteq \mathbb{R}^n$.

Definition.

Eine Funktion $f : D \rightarrow \mathbb{R}$ hat an einem Punkt $(x_1^{(0)}, \dots, x_n^{(0)}) \in D$ ein *absolutes Maximum*, falls $f(x_1^{(0)}, \dots, x_n^{(0)}) \geq f(x_1, \dots, x_n)$ für alle $(x_1, \dots, x_n) \in D$ gilt. Analog definiert man den Begriff des *absoluten Minimums*. Statt absolutes Maximum (Minimum) sagt man auch *globales Maximum (Minimum)*.

Bevor wir mit der Definition der Begriffe „lokales Maximum“ und „strenges lokales Maximum“ fortfahren, wollen wir einige Begriffsbildungen einführen, die in der gesamten Analysis von grundlegender Bedeutung sind.

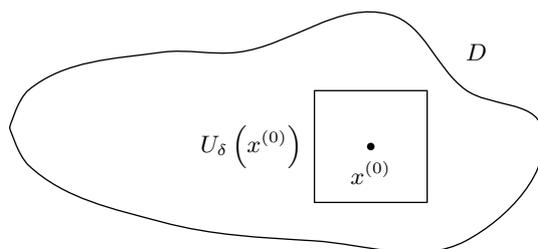
- (1) Es sei $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)}) \in \mathbb{R}^n$ und $\delta \in \mathbb{R}$, $\delta > 0$. Dann nennt man die Menge

$$U_\delta(x^{(0)}) := \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n : |x_i - x_i^{(0)}| < \delta \quad (i = 1, \dots, n) \right\}$$

eine (*offene*) δ -*Umgebung* von $x^{(0)}$.

- (2) Es sei $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)}) \in D \subseteq \mathbb{R}^n$. Man nennt $x^{(0)}$ einen *inneren Punkt* von D , falls es ein $\delta > 0$ gibt, für das $U_\delta(x^{(0)}) \subseteq D$ gilt.

- (3) Eine Menge $D \subseteq \mathbb{R}^n$ heißt *offen*, wenn es zu jedem $x^{(0)} \in D$ ein $\delta > 0$ gibt, so dass $U_\delta(x^{(0)}) \subseteq D$ gilt.



Man hätte ebenso gut sagen können: Eine Menge $D \subseteq \mathbb{R}^n$ heißt *offen*, wenn jedes $x^{(0)} \in D$ ein innerer Punkt von D ist.

- (4) Es sei $x^{(1)}, x^{(2)}, \dots$ eine unendliche Folge von Punkten des \mathbb{R}^n ; es gelte $x^{(i)} = (x_1^{(i)}, \dots, x_n^{(i)})$. Die Folge $x^{(1)}, x^{(2)}, \dots$ heißt *konvergent* gegen $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)}) \in \mathbb{R}^n$, falls für alle $j = 1, \dots, n$ gilt: $\lim_{i \rightarrow \infty} x_j^{(i)} = x_j^{(0)}$. Wir schreiben dann:

$$\lim_{i \rightarrow \infty} x^{(i)} = x^{(0)}$$

- (5) Es sei $f : D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^n$ und $x^{(0)} \in D$. f heißt *stetig an der Stelle $x^{(0)}$* , wenn für jede Folge $x^{(1)}, x^{(2)}, \dots$ von Punkten aus D gilt:

$$\lim_{i \rightarrow \infty} x^{(i)} = x^{(0)} \implies \lim_{i \rightarrow \infty} f(x^{(i)}) = f(x^{(0)}).$$

f heißt *stetig auf $X \subseteq D$* , falls f stetig an jeder Stelle $x^{(0)} \in X$ ist.

- (6) Es sei $f : D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^n$. f heißt *C^1 -Funktion auf X* (für $X \subseteq D$), falls auf X sämtliche partiellen Ableitungen erster Ordnung existieren und stetige Funktionen sind. Allgemeiner spricht man von einer *C^k -Funktion auf X* (für $k \geq 1$), wenn Entsprechendes für alle partiellen Ableitungen der Ordnung $\leq k$ gilt.

Definition.

Die Funktion f hat am Punkt $x^{(0)} \in D \subseteq \mathbb{R}^n$ ein *lokales Maximum*, wenn es ein $\delta > 0$ gibt, so dass $f(x^{(0)}) \geq f(x)$ für alle $x \in U_\delta(x^{(0)}) \cap D$ gilt.

Analog definiert man die Begriffe *lokales Minimum*, *strenges lokales Maximum* und *strenges lokales Minimum*.

Liegt ein lokales Maximum oder ein lokales Minimum vor, so spricht man von einem *lokalen Extremum*.

Satz 2.

Die Funktion $f : D \rightarrow \mathbb{R}$ sei an dem inneren Punkt $(x_1^{(0)}, \dots, x_n^{(0)}) \in D$ partiell differenzierbar nach jeder Variablen. Hat f bei $(x_1^{(0)}, \dots, x_n^{(0)})$ ein lokales Extremum, so gilt $f_{x_i}(x_1^{(0)}, \dots, x_n^{(0)}) = 0$ für alle $i = 1, \dots, n$.

Beweis. Da f an der Stelle $(x_1^{(0)}, \dots, x_n^{(0)})$ ein lokales Extremum hat, hat für $1 \leq i \leq n$ jede der Funktionen

$$\varphi_i(x_i) = f(x_1^{(0)}, \dots, x_{i-1}^{(0)}, x_i, x_{i+1}^{(0)}, \dots, x_n^{(0)})$$

ein lokales Extremum an der Stelle $x_i^{(0)}$. Es gilt also nach Satz 17 (Abschnitt 2.4): $\varphi_i'(x_i^{(0)}) = 0$ ($i = 1, \dots, n$). Wegen $\varphi_i'(x_i^{(0)}) = f_{x_i}(x_1^{(0)}, \dots, x_n^{(0)})$ folgt die Behauptung. \square

Ist die in Satz 2 angegebene notwendige Bedingung für das Vorliegen eines lokalen Extremums erfüllt, so nennt man $(x_1^{(0)}, \dots, x_n^{(0)})$ eine *stationäre* (oder *kritische*) *Stelle* von f . Mit anderen Worten: $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ ist eine stationäre (oder kritische) Stelle von f , wenn sämtliche partiellen Ableitungen erster Ordnung von f an der Stelle $x^{(0)}$ gleich Null sind.

Die in Satz 2 angegebene notwendige Bedingung für das Vorliegen eines lokalen Extremums ist nicht hinreichend, wie das folgende Beispiel zeigt.

Beispiel. Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ sei definiert durch $f(x, y) = xy$. Dann gilt $f_x(x, y) = y$ und $f_y(x, y) = x$. Also: $f_x(0, 0) = 0$ und $f_y(0, 0) = 0$. In $(0, 0)$ hat f aber kein lokales Extremum, da es in beliebiger Nähe von $(0, 0)$ sowohl positive als auch negative Funktionswerte gibt.

Eine stationäre Stelle, in der kein lokales Extremum vorliegt, nennt man einen *Sattelpunkt*. Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = xy$ hat also in $(0, 0)$ einen Sattelpunkt.

Der nächste Satz liefert ein hinreichendes Kriterium für das Vorliegen eines lokalen Extremums an einem inneren Punkt von $D \subseteq \mathbb{R}^2$.

Satz 3.

Für $D \subseteq \mathbb{R}^2$ sei eine Funktion $f : D \rightarrow \mathbb{R}$ gegeben und $(x_0, y_0) \in D$ sei ein innerer Punkt von D . In einer Umgebung $U_\delta(x_0, y_0)$ seien sämtliche partiellen Ableitungen der Ordnung ≤ 2 von f vorhanden und stetig (d.h., f sei eine C^2 -Funktion auf $U_\delta(x_0, y_0)$) und es gelte $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$. Wir setzen

$$\Delta(x_0, y_0) := f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - \left(f_{xy}(x_0, y_0)\right)^2.$$

Dann gilt:

(a) Ist $\Delta(x_0, y_0) > 0$, so hat f an (x_0, y_0) ein lokales Extremum; und zwar:

- ein strenges lokales Maximum, falls $f_{xx}(x_0, y_0) < 0$;
- ein strenges lokales Minimum, falls $f_{xx}(x_0, y_0) > 0$.

(b) Ist $\Delta(x_0, y_0) < 0$, so hat f an (x_0, y_0) kein lokales Extremum.

Bemerkung. Ist $\Delta(x_0, y_0) = 0$, so lässt sich mithilfe von Satz 3 nicht entscheiden, ob ein lokales Extremum vorliegt.

Verallgemeinerung von Satz 3 auf Funktionen $f : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$.

Wir benötigen einige Vorbereitungen; hierbei, ebenso wie in den folgenden Abschnitten, setzen wir Kenntnisse über Matrizen voraus (in dem Umfang, wie im ersten Semester behandelt). Darüber hinaus benötigen wir einige einfache Tatsachen aus der Linearen Algebra. Alle betrachteten Matrizen seien reell, d.h., die Einträge seien stets reelle Zahlen.

Es sei $A = (a_{ij})$ eine $n \times n$ -Matrix. Wir definieren eine Abbildung Q_A , die jedem n -Tupel $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ eine reelle Zahl $Q_A(x)$ zuordnet:

$$Q_A : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$x \mapsto Q_A(x) := \sum_{i=1}^n \sum_{j=1}^n (a_{ij} x_i x_j)$$

Man beachte, dass die auftretende Summe aus n^2 Summanden besteht. Man nennt die Abbildung Q_A die zu A gehörige *quadratische Form*. Zur Illustration betrachten wir ein Beispiel.

Beispiel. Es sei $n = 3$ und A sei die folgende 3×3 -Matrix:

$$A = \begin{pmatrix} 7 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & -1 & -2 \end{pmatrix}.$$

Ferner sei $x = (1, -2, 3)$. Man erhält

$$\begin{aligned} Q_A(x) &= 7 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot (-2) + 3 \cdot 1 \cdot 3 \\ &\quad + 4 \cdot (-2) \cdot 1 + 5 \cdot (-2) \cdot (-2) + 6 \cdot (-2) \cdot 3 \\ &\quad + 0 \cdot 3 \cdot 1 + (-1) \cdot 3 \cdot (-2) + (-2) \cdot 3 \cdot 3 \\ &= -30 \end{aligned}$$

Definition.

Eine symmetrische $n \times n$ - Matrix A heißt:

- *positiv definit*, falls $Q_A(x) > 0$ für alle $x \in \mathbb{R}^n$ mit $x \neq 0$ gilt.
- *negativ definit*, falls $Q_A(x) < 0$ für alle $x \in \mathbb{R}^n$ mit $x \neq 0$ gilt.
- *positiv semidefinit*, falls $Q_A(x) \geq 0$ für alle $x \in \mathbb{R}^n$ gilt.
- *negativ semidefinit*, falls $Q_A(x) \leq 0$ für alle $x \in \mathbb{R}^n$ gilt.
- *indefinit*, falls es $x, y \in \mathbb{R}^n$ gibt mit $Q_A(x) < 0 < Q_A(y)$.

Dabei sei Q_A die zu A gehörige quadratische Form.

Um festzustellen, ob positive oder negative Definitheit vorliegt, ist folgendes Kriterium nützlich, das wir ohne Beweis angeben.

Satz 4 (Definitheitskriterium).

$A = (a_{ij})$ sei eine symmetrische $n \times n$ - Matrix. Für $k = 1, \dots, n$ sei Δ_k die folgende Determinante, die man die *k-te Abschnittsdeterminante* von A nennt:

$$\Delta_k := \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}.$$

Dann gilt:

- 1) A ist positiv definit \Leftrightarrow alle Δ_k sind positiv
- 2) A ist negativ definit \Leftrightarrow die Vorzeichen der Δ_k alternieren, beginnend mit negativem Δ_1 , d.h., $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots$

Ferner gilt für 2×2 - Matrizen das folgende *Definitheits- und Indefinitheitskriterium*. Für symmetrische 2×2 - Matrizen $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ mit $\Delta := \det A = ac - b^2$ gilt:

- 1) A ist positiv definit $\Leftrightarrow a > 0$ und $\Delta > 0$
- 2) A ist negativ definit $\Leftrightarrow a < 0$ und $\Delta > 0$
- 3) A ist indefinit $\Leftrightarrow \Delta < 0$

Beispiele:

1. $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$

A ist positiv definit, da $\Delta_1 = 1 > 0, \Delta_2 = 2 > 0$ und $\Delta_3 = 3 > 0$

2. $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 2 \\ 0 & 2 & -3 \end{pmatrix}$

A ist negativ definit, da $\Delta_1 = -1 < 0, \Delta_2 = 2 > 0$ und $\Delta_3 = -2 < 0$

3. $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 4 & 3 \end{pmatrix}$

A ist indefinit, da es x, y gibt, so dass gilt: $Q_A(x) < 0 < Q_A(y)$. Beispielsweise gilt dies für $x = (1, -1, 1)$ und $y = (1, 1, 1)$.

Für $D \subseteq \mathbb{R}^n$ sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)}) \in D$.

Ist f in $x^{(0)}$ partiell nach allen Variablen differenzierbar, so nennt man den Vektor

$$\text{grad } f(x^{(0)}) := \left(\frac{\partial f}{\partial x_1}(x^{(0)}), \dots, \frac{\partial f}{\partial x_n}(x^{(0)}) \right)$$

den *Gradienten von f* an der Stelle $x^{(0)}$. Eine andere Schreibweise für $\text{grad } f(x^{(0)})$ ist $\nabla f(x^{(0)})$. Das Zeichen ∇ wird *Nabla-Operator* genannt. Es ist üblich, den Gradienten als Zeilenvektor anzusehen.

Die notwendige Bedingung aus Satz 2 können wir also auch so formulieren: Hat f an dem inneren Punkt $x^{(0)}$ von D ein lokales Extremum und existieren alle partiellen Ableitungen an $x^{(0)}$, so gilt $\text{grad } f(x^{(0)}) = 0$, d.h., der Gradient ist der Nullvektor.

Um die angekündigte hinreichende Bedingung zu formulieren, betrachten wir die partiellen Ableitungen zweiter Ordnung und bilden die sogenannte Hessesche Matrix. Wir setzen voraus, dass für f alle partiellen Ableitungen zweiter Ordnung an der Stelle $x^{(0)}$ existieren. Dann nennt man die folgende Matrix $H_f(x^{(0)})$ die *Hessesche Matrix von f* an der Stelle $x^{(0)}$:

$$H_f(x^{(0)}) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x^{(0)}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x^{(0)}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x^{(0)}) & \frac{\partial^2 f}{\partial x_2^2}(x^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(x^{(0)}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x^{(0)}) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x^{(0)}) \end{pmatrix}.$$

Beispiel. Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, die durch $f(x, y) = 2x^2 + \frac{1}{2}y^2 - xy + x + 2y + 3$ gegeben ist und berechnen zunächst die partiellen Ableitungen erster und zweiter Ordnung von f :

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 4x - y + 1 & \frac{\partial f}{\partial y}(x, y) &= y - x + 2 \\ \frac{\partial^2 f}{\partial x^2}(x, y) &= 4 & \frac{\partial^2 f}{\partial x \partial y}(x, y) &= -1 & \frac{\partial^2 f}{\partial y \partial x}(x, y) &= -1 & \frac{\partial^2 f}{\partial y^2}(x, y) &= 1. \end{aligned}$$

Hieraus erhält man den Gradienten und die Hessesche Matrix von f an der Stelle (x, y) :

$$\begin{aligned} \text{grad } f(x, y) &= (4x - y + 1, -x + y + 2), \\ H_f(x, y) &= \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}. \end{aligned}$$

Satz 1' (Schwarz, allgemeine Fassung).

Für $D \subseteq \mathbb{R}^n$ sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $x^{(0)}$ ein innerer Punkt von D . Existieren für ein $k \geq 2$ auf einer δ -Umgebung $U_\delta(x^{(0)})$ sämtliche partiellen Ableitungen der Ordnung $\leq k$ und sind diese auf $U_\delta(x^{(0)})$ stetig (d.h. f ist auf $U_\delta(x^{(0)})$ eine C^k -Funktion), so ist auf $U_\delta(x^{(0)})$ für die partiellen Ableitungen k -ter Ordnung die Reihenfolge der Differentiation vertauschbar. Insbesondere folgt, dass unter entsprechenden Voraussetzungen die Hessesche Matrix symmetrisch ist.

Satz 5 (Extremwertkriterium).

Für $D \subseteq \mathbb{R}^n$ sei eine Funktion $f : D \rightarrow \mathbb{R}$ gegeben und $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ sei ein innerer Punkt von D . Auf einer Umgebung $U_\delta(x^{(0)})$ von $x^{(0)}$ sei f eine C^2 -Funktion. Es gelte $\text{grad } f(x^{(0)}) = 0$. Dann gilt:

- (a) Ist $H_f(x^{(0)})$ positiv definit, so hat f an der Stelle $x^{(0)}$ ein strenges lokales Minimum.
- (b) Ist $H_f(x^{(0)})$ negativ definit, so hat f an der Stelle $x^{(0)}$ ein strenges lokales Maximum.
- (c) Ist $H_f(x^{(0)})$ indefinit, so hat f an der Stelle $x^{(0)}$ mit Sicherheit kein lokales Extremum.

Beweis. Siehe z.B.: Heuser, Analysis II.

Zur Erläuterung behandeln wir das folgende **Beispiel**. Die Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ sei gegeben durch:

$$f(x, y, z) = x^2 + y^2 + 2z^2 - 2xz - y.$$

Wir untersuchen f auf Extrema. Hierzu bilden wir zunächst den Gradienten. Es ergibt sich:

$$\text{grad } f(x, y, z) = (2x - 2z, 2y - 1, -2x + 4z).$$

Um die stationären Stellen von f zu finden (mögliche Extrema!), setzen wir

$$\text{grad } f(x, y, z) = (0, 0, 0).$$

Dies ist äquivalent zum folgenden Gleichungssystem:

$$\begin{aligned} 2x - 2z &= 0 \\ 2y - 1 &= 0 \\ -2x + 4z &= 0 \end{aligned}$$

Einzigste Lösung dieses Gleichungssystems (und damit einzige stationäre Stelle von f) ist:

$$(x, y, z) = \left(0, \frac{1}{2}, 0\right).$$

Wir testen nun mithilfe von Satz 5, ob ein lokales Extremum an dieser Stelle vorliegt und welcher Art es ggf. ist. Hierzu haben wir die Hessesche Matrix H_f von f an der Stelle $x^{(0)} = (0, \frac{1}{2}, 0)$ zu bilden. Es ergibt sich:

$$H_f = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 4 \end{pmatrix}.$$

Diese Matrix ist positiv definit, da die drei Abschnittdeterminanten positiv sind (Satz 4). Es folgt (nach Satz 5): f hat bei $(0, \frac{1}{2}, 0)$ ein strenges lokales Minimum. Weitere lokale Extrema gibt es nicht.

Aufgabe. Weiter oben (vor Satz 1') haben wir die Funktion $f(x, y) = 2x^2 + \frac{1}{2}y^2 - xy + x + 2y + 3$ betrachtet. Bestimmen Sie die kritischen Stellen dieser Funktion und entscheiden Sie für jede dieser Stellen, ob ein lokales Minimum oder Maximum vorliegt.

5.3 Die geometrische Bedeutung der partiellen Ableitungen und des Gradienten

Für $D \subseteq \mathbb{R}^n$ sei $f : D \rightarrow \mathbb{R}$ eine C^1 -Funktion und $x^{(0)}$ sei ein innerer Punkt von D . Wir wollen uns die anschauliche Bedeutung der partiellen Ableitungen vor Augen führen und uns mit der Frage nach der

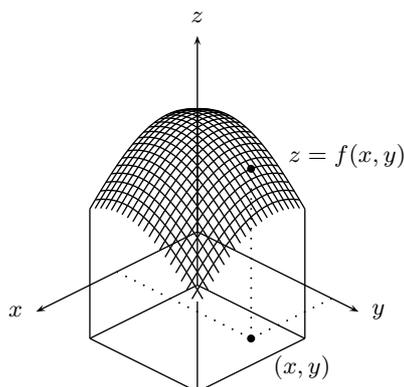
geometrischen Bedeutung des Gradienten

$$\text{grad } f(x^{(0)}) = \left(\frac{\partial f}{\partial x_1}(x^{(0)}), \dots, \frac{\partial f}{\partial x_n}(x^{(0)}) \right)$$

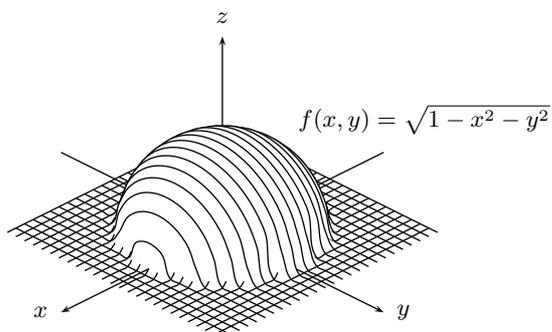
beschäftigen.

Wir betrachten hier nur den anschaulichen Fall $n = 2$, erwähnen aber, dass sich alle Überlegungen auf den allgemeinen Fall übertragen lassen.

Bis zum Ende des Abschnitts 5.3 sei also $n = 2$. Wir betrachten den *Graphen von f* , d.h. die Menge aller Punkte $(x, y, z) \in \mathbb{R}^3$ mit $(x, y) \in D$ und $z = f(x, y)$. Man kann sich den Graphen von f als ein Gebirge oder, wie man auch sagt, als eine *Fläche* vorstellen. Diese Vorstellung wird durch die folgenden Skizzen illustriert:

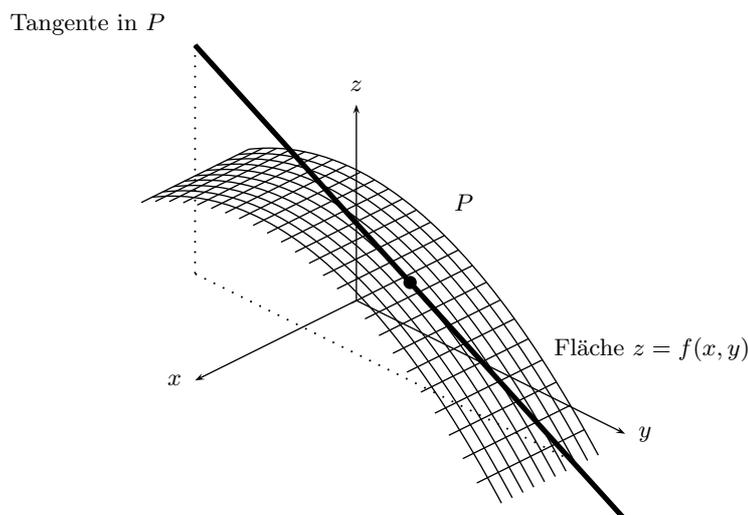


Der Graph einer Funktion von zwei Variablen.



Der Graph der Funktion $f(x, y) = \sqrt{1 - x^2 - y^2}$ (obere Hemisphäre der Einheitskugel).

Die folgende Abbildung veranschaulicht die geometrische Bedeutung der partiellen Ableitung $\frac{\partial f}{\partial y}(x, y)$:



Man erkennt, dass $\frac{\partial f}{\partial y}(x, y)$ die *Steigung des Graphen von f in Richtung der y -Achse* angibt (und zwar im Punkt (x, y, z) mit $z = f(x, y)$).

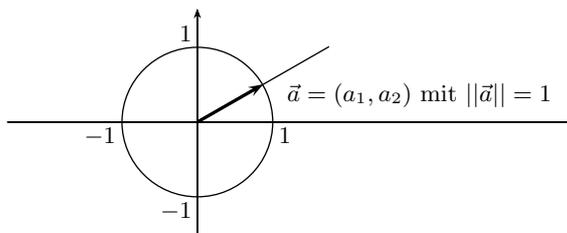
Entsprechend gibt $\frac{\partial f}{\partial x}(x, y)$ die *Steigung in Richtung der x -Achse* an (im Punkt (x, y, z) mit $z = f(x, y)$).

Damit haben wir die anschauliche Bedeutung der partiellen Ableitungen $\frac{\partial f}{\partial x}(x, y)$ und $\frac{\partial f}{\partial y}(x, y)$ beschrieben. Wir kommen nun zur Frage, welche geometrische Bedeutung dem Vektor

$$\text{grad } f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right)$$

zukommt.

Unter einer *Richtung* im \mathbb{R}^2 wollen wir einen Vektor $\vec{a} = (a_1, a_2)$ mit $\|\vec{a}\| = 1$ („Einheitsvektor“) verstehen¹.



Es sei $P = (x, y, z)$ mit $z = f(x, y)$ ein Punkt auf dem Graphen von f . Wir fragen nach dem *Anstieg des Graphen in diesem Punkt in der Richtung $\vec{a} = (a_1, a_2)$* . (Wir wollen hier nicht im Detail angeben, wie der Begriff „Anstieg des Graphen in Richtung \vec{a} “ mathematisch definiert wird, anschaulich sollte jedoch klar sein, was gemeint ist. Ähnlich, wie wir weiter oben in einer Zeichnung die Steigung in Richtung der y -Achse betrachtet haben, könnten wir auch die Steigung in einer beliebigen anderen Richtung \vec{a} betrachten.)

Für $u = (u_1, u_2)$ und $v = (v_1, v_2)$ sei mit $\langle u, v \rangle$ das gewöhnliche Skalarprodukt bezeichnet, d.h.

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2.$$

¹Mit $\|\vec{a}\|$ sei die Länge des Vektors $\vec{a} = (a_1, a_2)$ bezeichnet, also $\|\vec{a}\| = \sqrt{a_1^2 + a_2^2}$

Es gilt (vgl. Lehrbücher der Analysis, etwa E. Behrends, Analysis Band 2, Vieweg): *Der Anstieg des Graphen von f an der Stelle (x, y, z) (mit $z = f(x, y)$) in Richtung $\vec{a} = (a_1, a_2)$ ist gleich*

$$\langle \vec{a}, \text{grad } f(x, y) \rangle = a_1 \cdot \frac{\partial f}{\partial x}(x, y) + a_2 \cdot \frac{\partial f}{\partial y}(x, y). \quad (5.1)$$

Gilt $\text{grad } f(x, y) = (0, 0)$ an einer Stelle (x, y, z) , so ist der Anstieg in alle Richtungen gleich 0. Wir fragen für eine feste Stelle (x, y, z) auf dem Graphen von f , für die $\text{grad } f(x, y) \neq (0, 0)$ gilt: *Für welche Richtung \vec{a} ist der Anstieg des Graphen von f am größten?*

Antwort: $\varphi \in [0, \pi]$ sei die Größe des von \vec{a} und $\text{grad } f(x, y)$ gebildeten Winkels (im Bogenmaß). Dann gilt (man beachte $\|\vec{a}\| = 1$):

$$\begin{aligned} \langle \vec{a}, \text{grad } f(x, y) \rangle &= \|\vec{a}\| \cdot \|\text{grad } f(x, y)\| \cdot \cos \varphi \\ &= \|\text{grad } f(x, y)\| \cdot \cos \varphi \end{aligned} \quad (5.2)$$

Das Skalarprodukt $\langle \vec{a}, \text{grad } f(x, y) \rangle$ ist also am größten, wenn $\cos \varphi$ möglichst groß ist, d.h., wenn $\cos \varphi = 1$ gilt. Dies ist aber genau dann der Fall, wenn $\varphi = 0$ ist, d.h., wenn \vec{a} und $\text{grad } f(x, y)$ in dieselbe Richtung zeigen. Wir halten fest:

Der Gradient $\text{grad } f(x, y)$ zeigt in die Richtung des größten Anstiegs des Graphen von f im Punkt $(x, y, f(x, y))$.

Außerdem gilt (In (5.2) setze man $\varphi = 0$ ein!):

Die Länge des Gradienten $\text{grad } f(x, y)$ gibt die Größe des maximalen Anstiegs im Punkt $(x, y, f(x, y))$ an.

Entsprechend gilt: *Der negative Gradient $-\text{grad } f(x, y)$ zeigt in die Richtung des steilsten Abstiegs.* (Man beachte: Für $\varphi \in [0, \pi]$ wird $\cos \varphi$ am kleinsten, nämlich -1 , wenn $\varphi = \pi$ gilt.)

Beispiel (nach E. Behrends, Analysis Band 2). Eine rechteckige Platte sei aufgeheizt, die Temperatur bei (x, y) sei durch $T(x, y) = 50 - x^2y$ gegeben. Eine Maus befinde sich an der Stelle $(1, 2)$. Sie fühlt sich dort wegen $T(1, 2) = 48$ nicht sehr wohl. In welche Richtung empfehlen Sie der Maus wegzulaufen?



Lösung: T fällt am schnellsten in Richtung des negativen Gradienten:

$$\begin{aligned} -\text{grad } T(x, y) &= -(-2xy, -x^2) \\ &= (2xy, x^2) \end{aligned}$$

Für $(x, y) = (1, 2)$ gilt $-\text{grad } T(1, 2) = (4, 1)$.

Empfehlung: Die Maus sollte sich in Richtung des Vektors $(4, 1)$ entfernen.

5.4 Extrema mit Nebenbedingungen und Lagrangesche Multiplikatorenregel

In der Praxis stellt sich häufig die Aufgabe, lokale Extrema einer Funktion $f : D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^n$ unter Nebenbedingungen zu finden, die in der Form von m Gleichungen

$$g_1(x) = 0, \dots, g_m(x) = 0 \quad (*)$$

vorliegen, wobei $g_i : D \rightarrow \mathbb{R}$ ($i = 1, \dots, m$) gilt.

Definition.

Es seien f, g_1, \dots, g_m wie oben; wir setzen

$$M := \left\{ x \in D : g_1(x) = \dots = g_m(x) = 0 \right\}.$$

Die Funktion f hat an dem Punkt $x^{(0)}$ ein *lokales Maximum unter den Nebenbedingungen* (*), wenn $x^{(0)} \in M$ gilt und ein $\delta > 0$ existiert, so dass

$$f(x^{(0)}) \geq f(x) \text{ für alle } x \in U_\delta(x^{(0)}) \cap M.$$

Analog definiert man die Begriffe *lokales Minimum*, *strenges lokales Maximum*, *strenges lokales Minimum* sowie *Extremum unter der Nebenbedingung* (*).

Häufig löst man Extremwertaufgaben mit Nebenbedingungen mit der *Lagrangeschen Multiplikatorenregel*. (Man spricht in diesem Zusammenhang auch von *Lagrange-Optimierung* oder Lösung mithilfe des *Lagrange-Ansatzes*.) Im Folgenden stellen wir zunächst den zugrundeliegenden Satz vor und zeigen dann anhand eines Beispiels, wie der Satz zur Lösung konkreter Aufgaben eingesetzt werden kann.

Satz 6 (Lagrangesche Multiplikatorenregel).

Für $D \subseteq \mathbb{R}^n$ seien $f : D \rightarrow \mathbb{R}$ und $g_i : D \rightarrow \mathbb{R}$ ($i = 1, \dots, m$) Funktionen, wobei $m < n$ gelte. Es sei $x^{(0)}$ ein innerer Punkt von D und auf einer δ -Umgebung $U_\delta(x^{(0)})$ seien die Funktionen f und g_i ($i = 1, \dots, m$) C^1 -Funktionen. Der Rang der $m \times n$ -Matrix

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x^{(0)}) & \frac{\partial g_1}{\partial x_2}(x^{(0)}) & \dots & \frac{\partial g_1}{\partial x_n}(x^{(0)}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(x^{(0)}) & \frac{\partial g_m}{\partial x_2}(x^{(0)}) & \dots & \frac{\partial g_m}{\partial x_n}(x^{(0)}) \end{pmatrix}$$

sei m . Hat unter diesen Voraussetzungen die Funktion f an der Stelle $x^{(0)}$ ein lokales Extremum unter der Nebenbedingung (*), so gibt es Zahlen $\lambda_1, \dots, \lambda_m$ (genannt: *Lagrangesche Multiplikatoren*), für die gilt:

$$\frac{\partial f}{\partial x_j}(x^{(0)}) + \lambda_1 \frac{\partial g_1}{\partial x_j}(x^{(0)}) + \dots + \lambda_m \frac{\partial g_m}{\partial x_j}(x^{(0)}) = 0 \quad (j = 1, \dots, n). \quad (5.3)$$

Für den Beweis dieses Satzes sei auf die Lehrbücher der Analysis verwiesen, etwa:

- H. Heuser: Lehrbuch der Analysis, Teil 2
- K. Endl, W. Luh: Analysis II

Die in Satz 6 gemachte Voraussetzung, dass die dort angegebene Matrix den Rang m hat, nennt man auch *Regularitätsbedingung*. Die obige Funktion f wird häufig als *Zielfunktion* bezeichnet. In der Praxis ist es üblich, wie folgt vorzugehen: Zu den Unbestimmten x_1, \dots, x_n nimmt man weitere (zunächst) unbestimmte Zahlen $\lambda_1, \dots, \lambda_m$ hinzu und stellt die *Lagrange-Funktion* L auf. Dies ist die Funktion

$$L : \mathbb{R}^{n+m} \rightarrow \mathbb{R},$$

die gegeben ist durch:

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) + \lambda_1 g_1(x_1, \dots, x_n) + \dots + \lambda_m g_m(x_1, \dots, x_n). \quad (5.4)$$

Anschließend bildet man die $n + m$ partiellen Ableitungen von L und setzt diese gleich Null. Man erhält:

$$\begin{aligned} \frac{\partial L}{\partial x_1}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) &= 0 \\ &\vdots \\ \frac{\partial L}{\partial x_n}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) &= 0 \\ &\vdots \\ \frac{\partial L}{\partial \lambda_1}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) &= 0 \\ &\vdots \\ \frac{\partial L}{\partial \lambda_m}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) &= 0 \end{aligned} \quad (5.5)$$

Man beachte, dass Folgendes gilt (für $j = 1, \dots, n$ und $i = 1, \dots, m$):

$$\begin{aligned} \frac{\partial L}{\partial x_j}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) &= \frac{\partial f}{\partial x_j}(x_1, \dots, x_n) + \lambda_1 \frac{\partial g_1}{\partial x_j}(x_1, \dots, x_n) + \dots + \lambda_m \frac{\partial g_m}{\partial x_j}(x_1, \dots, x_n) \\ \frac{\partial L}{\partial \lambda_i}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) &= g_i(x_1, \dots, x_n) \end{aligned}$$

Die $n + m$ Gleichungen (5.5) sind also gerade die folgenden Gleichungen:

$$\begin{aligned} \frac{\partial f}{\partial x_j}(x_1, \dots, x_n) + \lambda_1 \frac{\partial g_1}{\partial x_j}(x_1, \dots, x_n) + \dots + \lambda_m \frac{\partial g_m}{\partial x_j}(x_1, \dots, x_n) &= 0 \quad (j = 1, \dots, n) \\ g_i(x_1, \dots, x_n) &= 0 \quad (i = 1, \dots, m) \end{aligned} \quad (5.6)$$

Aufgrund von Satz 6 gilt: Hat (unter den Voraussetzungen von Satz 6) die Funktion f an der Stelle $x^{(0)} = (x_1, \dots, x_n)$ ein lokales Extremum unter der Nebenbedingung (*), so erfüllt $x^{(0)} = (x_1, \dots, x_n)$ zusammen mit gewissen Zahlen $\lambda_1, \dots, \lambda_m$ das Gleichungssystem (5.6) bzw. (was dasselbe ist) das Gleichungssystem (5.5). Um herauszufinden, welche $x^{(0)} = (x_1, \dots, x_n)$ diese notwendige Bedingung für das Vorliegen eines lokalen Extremums erfüllen, hat man also das Gleichungssystem (5.5) zu lösen. Man beachte, dass es sich bei (5.5) um ein *Gleichungssystem mit $n + m$ Gleichungen für die $n + m$ Unbestimmten $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$* handelt.

Man nennt die Lösungen dieses Gleichungssystems *stationäre* (oder *kritische*) *Stellen* von f unter der Nebenbedingung (*). (Auch bei Extremwertaufgaben ohne Nebenbedingungen spricht man ja, wie bereits erwähnt, von stationären bzw. kritischen Stellen, nämlich genau dann, wenn die notwendige Bedingung aus Satz 2 erfüllt ist.)

Beispiel.

Es sei $K = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ („Kugeloberfläche“) und $E = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ („Ebene durch den Nullpunkt“). Gesucht ist derjenige Punkt $(x, y, z) \in K \cap E$, dessen Abstand zum Punkt $(1, 1, 0)$ minimal ist.

Anstelle des Abstands $\sqrt{(x-1)^2 + (y-1)^2 + z^2}$ können wir ebenso gut das Abstandsquadrat minimieren. Unsere Aufgabe lautet also: Finde das Minimum der Funktion

$$f(x, y, z) = (x-1)^2 + (y-1)^2 + z^2$$

unter den folgenden Nebenbedingungen:

$$\begin{aligned} g_1(x, y, z) &= x + y + z = 0 \\ g_2(x, y, z) &= x^2 + y^2 + z^2 - 1 = 0 \end{aligned} \quad (**)$$

Es gilt:

$$\begin{pmatrix} \frac{\partial g_1}{\partial x}(x, y, z) & \frac{\partial g_1}{\partial y}(x, y, z) & \frac{\partial g_1}{\partial z}(x, y, z) \\ \frac{\partial g_2}{\partial x}(x, y, z) & \frac{\partial g_2}{\partial y}(x, y, z) & \frac{\partial g_2}{\partial z}(x, y, z) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{pmatrix}.$$

Man erkennt, dass der Rang dieser Matrix gleich 2 ist für alle $(x, y, z) \in K \cap E$, d.h., die Regularitätsbedingung ist erfüllt. Lösung mithilfe der Lagrangeschen Multiplikatorenregel: Die Lagrange-Funktion lautet:

$$L(x, y, z, \lambda, \mu) = (x - 1)^2 + (y - 1)^2 + z^2 + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 1).$$

Partielles Ableiten und gleich Null Setzen ergibt folgendes Gleichungssystem:

$$\begin{aligned} 2(x - 1) + \lambda + 2\mu x &= 0 \\ 2(y - 1) + \lambda + 2\mu y &= 0 \\ 2z + \lambda + 2\mu z &= 0 \\ x + y + z &= 0 \\ x^2 + y^2 + z^2 - 1 &= 0 \end{aligned} \quad (5.7)$$

Lösung von (5.7): Addiert man die ersten drei Gleichungen und berücksichtigt die vierte, so erhält man $-4 + 3\lambda = 0$; also $\lambda = \frac{4}{3}$. Nach Einsetzen von λ erhält man aus den ersten drei Gleichungen $(1 + \mu)x = \frac{1}{3}$, $(1 + \mu)y = \frac{1}{3}$ und $(1 + \mu)z = -\frac{2}{3}$. Hieraus ergibt sich, dass für eine Lösung unseres Gleichungssystems nicht $1 + \mu = 0$ gelten kann. Es folgt

$$x = \frac{1}{3(1 + \mu)}, \quad y = \frac{1}{3(1 + \mu)} \quad \text{und} \quad z = \frac{-2}{3(1 + \mu)},$$

also $y = x$ und $z = -2x$. Wegen $x^2 + y^2 + z^2 - 1 = 0$ folgt $6x^2 - 1 = 0$. Also $x^2 = \frac{1}{6}$; es folgt $x = \frac{1}{\sqrt{6}}$ oder $x = -\frac{1}{\sqrt{6}}$ und man erhält, dass (x, y, z) gleich

$$\left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}} \right) \quad \text{oder} \quad \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right) \quad (5.8)$$

ist. Wir wissen somit aufgrund von Satz 6: Der gesuchte Punkt (x, y, z) muss einer der beiden Punkte (5.8) sein. Welcher es ist, erhält man durch Einsetzen in f . Es gilt:

$$\begin{aligned} f\left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right) &= 2\left(\frac{1}{\sqrt{6}} - 1\right)^2 + \frac{2}{3} \\ f\left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right) &= 2\left(\frac{1}{\sqrt{6}} + 1\right)^2 + \frac{2}{3} \end{aligned}$$

Da der erste Wert der kleinere ist, ist $\left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ der gesuchte Punkt auf $K \cap E$, dessen Abstand zum Punkt $(1, 1, 0)$ minimal ist. (Der Punkt $\left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)$ ist derjenige Punkt auf $K \cap E$, dessen Abstand von $(1, 1, 0)$ maximal ist.)

Hinweise.

- 1) Das Gleichungssystem (5.5) stellt nur eine notwendige Bedingung für das Vorhandensein eines Extremums dar. Das bedeutet: Nachdem man die stationären Stellen bestimmt hat, muss man sich noch überlegen, ob an diesen Stellen ein Minimum, ein Maximum oder überhaupt kein Extremum vorliegt. Bei vielen in der Praxis auftretenden Fragestellungen ist es möglich, dies (wie in obigem Beispiel) mit einfachen Mitteln zu entscheiden. Wer sich über hinreichende Bedingungen informieren möchte, findet eine detaillierte Beschreibung beispielsweise in
 - Alessandro Tomazic, *Wirtschaftsmathematik 2*, Folien zur gleichnamigen Vorlesung im Sommersemester 2009 an der Universität Wien.
- 2) Extremwertaufgaben mit Nebenbedingungen treten häufig in den Wirtschaftswissenschaften auf („Ertragsmaximierung unter gegebenen Produktionsbedingungen“).
- 3) Gelegentlich kann man Extremwertaufgaben *auf wesentlich einfachere Art* ohne Lagrangesche Multiplikatorenregel lösen, nämlich durch geeignetes Auflösen der Nebenbedingungen und anschließendes Einsetzen in die Zielfunktion. Dies läuft dann auf eine Extremwertaufgabe ohne Nebenbedingungen hinaus.

Beispiel. Gesucht sind die Extrema der Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) = x^2 + 3yz - 2xz$ unter den Nebenbedingungen $3x - z = 1$ und $y + 2z = 2$. Zur Einübung des Umgangs mit Satz 6 ist es sinnvoll, dies mithilfe des Lagrange-Ansatzes zu lösen. Dies läuft auf die Lösung eines linearen Gleichungssystems von 5 Gleichungen mit 5 Unbestimmten hinaus. Einfacher ist es jedoch, die Nebenbedingungen nach x bzw. y aufzulösen und das Ergebnis in die Zielfunktion $f(x, y, z)$ einzusetzen; man erhält $x = \frac{1}{3}(z + 1)$, $y = -2z + 2$ und durch Einsetzen:

$$f(x, y, z) = -\frac{59}{9}z^2 + \frac{50}{9}z + \frac{1}{9}.$$

Dies ist eine Funktion, die nur noch von der einen Unbestimmten z abhängt. Wir bezeichnen der Einfachheit halber die Funktion wieder mit f , schreiben also:

$$f(z) = -\frac{59}{9}z^2 + \frac{50}{9}z + \frac{1}{9}.$$

Man hat somit eine Extremwertaufgabe ohne Nebenbedingungen mit nur einer Unbestimmten erhalten.

Lösung: $f'(z) = -\frac{118}{9}z + \frac{50}{9} = 0 \Rightarrow z = \frac{25}{59} \Rightarrow x = \frac{28}{59}, y = \frac{68}{59}.$

Folglich ist $\frac{1}{59}(28, 68, 25)$ die einzige stationäre Stelle und wegen $f''\left(\frac{25}{59}\right) < 0$ handelt es sich um ein Maximum.

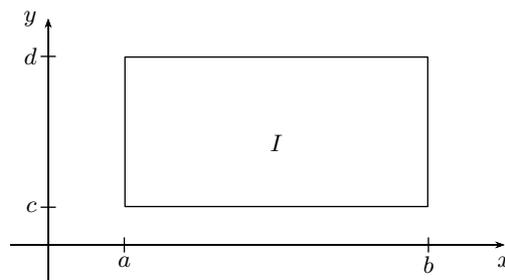
6 Integration von Funktionen mehrerer Variablen

Wir folgen in diesem Abschnitt der Darstellung in W. Luh: Mathematik für Naturwissenschaftler II; in diesem und anderen Lehrbüchern der Analysis findet man Ergänzungen (z.B. weitere Beispiele, Übungsaufgaben und Beweise). Wir beschränken uns hier auf zweidimensionale Integrale; drei- und mehrdimensionale Integrale werden ähnlich definiert und berechnet.

6.1 Definition zweidimensionaler Integrale

Unter einem *abgeschlossenen Intervall* im \mathbb{R}^2 verstehen wir hier immer eine Punktmenge:

$$I = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, c \leq y \leq d\} \quad \text{mit } a, b, c, d \in \mathbb{R}, a < b \text{ und } c < d.$$



Anstelle von abgeschlossenes Intervall im \mathbb{R}^2 sagen wir auch *achsenparalleles Rechteck* oder einfach nur *Rechteck* oder *Intervall*. Wir definieren zweidimensionale Integrale zunächst nur über Intervallen I . Es sei $f(x, y)$ eine im Intervall I beschränkte Funktion, d.h., es gelte $|f(x, y)| \leq K$ für eine Konstante $K \in \mathbb{R}$ und für alle $(x, y) \in I$. Um zur Definition des Integrals

$$\iint_I f(x, y) d(x, y)$$

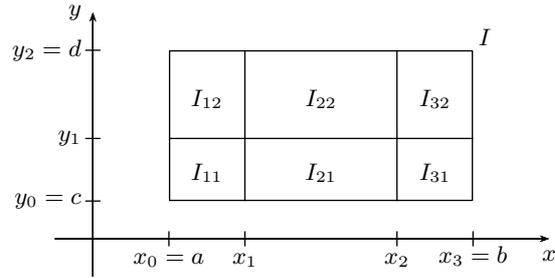
zu gelangen, betrachten wir Zerlegungen von I in achsenparallele Rechtecke und zugehörige Riemannsche Summen; wir gehen dabei ähnlich vor wie bei der Definition des Riemannsches Integrals im eindimensionalen Fall (siehe Abschnitt 3.2).

1. Z_1 sei eine Zerlegung von $[a, b]$ in n Teilintervalle und Z_2 sei eine Zerlegung von $[c, d]$ in m Teilintervalle:

$$\begin{aligned} Z_1 : a = x_0 < x_1 < x_2 < \dots < x_n = b \\ Z_2 : c = y_0 < y_1 < y_2 < \dots < y_m = d \end{aligned}$$

Hierdurch wird I in $n \cdot m$ Teilintervalle unterteilt (mit $i = 1, \dots, n$ und $j = 1, \dots, m$):

$$I_{ij} = \{(x, y) : x_{i-1} \leq x \leq x_i, y_{j-1} \leq y \leq y_j\}.$$



2. Zu jedem Teilintervall I_{ij} wählen wir einen Punkt $(\xi_{ij}, \eta_{ij}) \in I_{ij}$ und multiplizieren den Funktionswert $f(\xi_{ij}, \eta_{ij})$ mit dem Flächeninhalt $(x_i - x_{i-1})(y_j - y_{j-1})$ von I_{ij} :

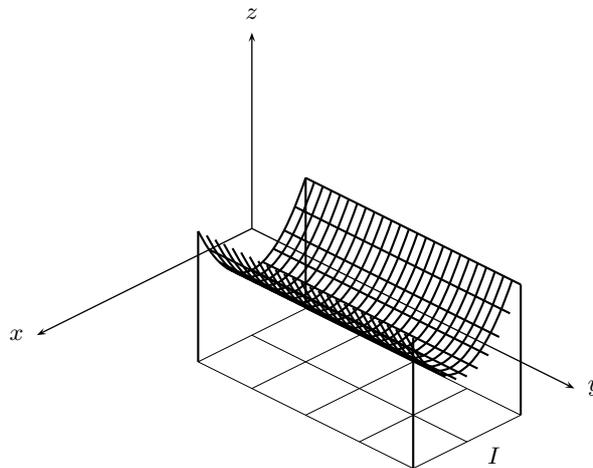
$$f(\xi_{ij}, \eta_{ij})(x_i - x_{i-1})(y_j - y_{j-1})$$

Unter der *Riemannschen Summe* von f bezüglich der Zerlegungen Z_1 und Z_2 sowie der Wahl der Zwischenpunkte (ξ_{ij}, η_{ij}) versteht man die Summe:

$$\sum_{i=1}^n \sum_{j=1}^m f(\xi_{ij}, \eta_{ij})(x_i - x_{i-1})(y_j - y_{j-1}). \quad (6.1)$$

Zur Abkürzung schreiben wir $S_Z(f)$ für die Summe (6.1), wobei $Z = \{I_{ij} : 1 \leq i \leq n, 1 \leq j \leq m\}$ die zugehörige Zerlegung von I sein soll.

Geometrische Bedeutung für den Fall, dass f nichtnegativ ist: Die Punktmenge $\{(x, y, z) : z = f(x, y), (x, y) \in I\}$ bildet ein „Flächenstück“ im Raum \mathbb{R}^3 , das über dem Rechteck I „aufgespannt“ ist. $S_Z(f)$ ist ein Näherungswert für das Volumen des Raumes zwischen diesem Flächenstück und dem Intervall I .



3. Unter dem *Feinheitsmaß* $|Z|$ von Z verstehen wir den größten Abstand, der unter den $n + m$ Abständen $|x_i - x_{i-1}|$ ($i = 1, \dots, n$) und $|y_j - y_{j-1}|$ ($j = 1, \dots, m$) vorkommt.

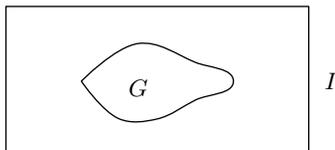
Wir gehen nun zu immer feineren Zerlegungen Z von I über, d.h., wir lassen $|Z| \rightarrow 0$ streben. Dies führt zur folgenden Definition (vgl. die entsprechende Definition in Abschnitt 3.2).

Definition.

Wir betrachten Folgen Z_1, Z_2, \dots von Zerlegungen von I in achsenparallele Rechtecke, für die $|Z_k| \rightarrow 0$ mit $k \rightarrow \infty$ gilt. Streben für alle derartigen Folgen (Z_k) und für jede mögliche Wahl der Zwischenpunkte (ξ_{ij}, η_{ij}) die Riemanschen Summen $S_{Z_k}(f)$ immer gegen denselben Grenzwert A , so heißt f *integrierbar* auf I und der Grenzwert A heißt *bestimmtes Integral* von f über I . Man bezeichnet A mit

$$\iint_I f(x, y) d(x, y).$$

Es sei nun $G \subseteq \mathbb{R}^2$ eine *beschränkte Menge*, d.h., es gibt ein abgeschlossenes Intervall $I \subseteq \mathbb{R}^2$, in dem G enthalten ist.

**Definition.**

Ist $f : G \rightarrow \mathbb{R}$ eine Funktion und gilt $G \subseteq I$ für ein Intervall $I \subseteq \mathbb{R}^2$, so bezeichnen wir mit $f_I : I \rightarrow \mathbb{R}$ diejenige Funktion, die wie folgt definiert wird:

$$f_I(x, y) = \begin{cases} f(x, y) & , \text{ falls } (x, y) \in G \\ 0 & , \text{ falls } (x, y) \notin G \end{cases}$$

Wir haben bisher zweidimensionale Integrale nur über sehr speziellen Mengen, nämlich über Intervallen I definiert. Mithilfe von $f_I(x, y)$ können wir nun Integrale über beliebigen beschränkten Mengen G definieren.

Definition.

Es sei $G \subseteq \mathbb{R}^2$ eine beschränkte Menge und I sei ein abgeschlossenes Intervall mit $G \subseteq I$. Die Funktion $f : G \rightarrow \mathbb{R}$ sei beschränkt. Existiert das Integral

$$\iint_I f_I(x, y) d(x, y),$$

so nennt man f *integrierbar* auf G und definiert:

$$\iint_G f(x, y) d(x, y) := \iint_I f_I(x, y) d(x, y).$$

Man kann zeigen, dass das so definierte Integral unabhängig von der Wahl des umgebenden Intervalls I ist. Ferner gilt:

$$\iint_G (f(x, y) + g(x, y)) d(x, y) = \iint_G f(x, y) d(x, y) + \iint_G g(x, y) d(x, y) \quad (6.2)$$

und (für $\alpha \in \mathbb{R}$):

$$\iint_G \alpha f(x, y) d(x, y) = \alpha \iint_G f(x, y) d(x, y). \quad (6.3)$$

Geometrische Veranschaulichung: Ist $f(x, y)$ auf G nichtnegativ, so gibt das Integral

$$\iint_G f(x, y) d(x, y)$$

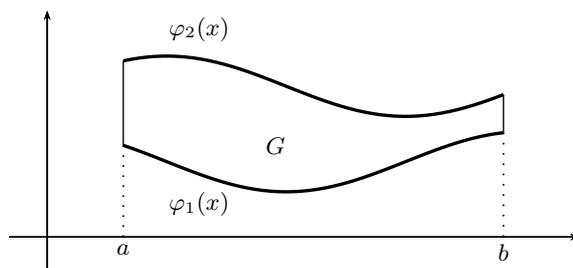
das Volumen des Raumstücks an, das zwischen G und dem durch f definierten „Flächenstück“ $F = \{(x, y, z) : (x, y) \in G, z = f(x, y)\}$ liegt.

6.2 Berechnung zweidimensionaler Integrale

Wir betrachten zunächst den Fall, dass der Integrationsbereich G durch die Graphen zweier Funktionen φ_1 und φ_2 berandet wird.

Es seien φ_1 und φ_2 stetige Funktionen $[a, b] \rightarrow \mathbb{R}$ und es gelte $\varphi_1(x) \leq \varphi_2(x)$ für alle $x \in [a, b]$. Es sei

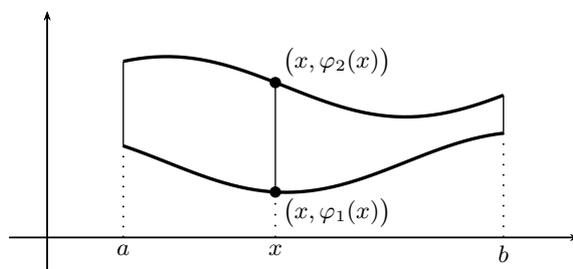
$$G = \{(x, y) : a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}. \quad (6.4)$$



$f : G \rightarrow \mathbb{R}$ sei eine stetige Funktion¹. Es sei nun $x \in [a, b]$ fest. Dann können wir für dieses feste x das Integral

$$\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy$$

bilden, wobei die Integrationsvariable y das Intervall zwischen $\varphi_1(x)$ und $\varphi_2(x)$ durchläuft (siehe Zeichnung).



Das Integral $\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy$ ist also ein gewöhnliches eindimensionales Integral; der Wert dieses Integrals ist davon abhängig, welches x wir gewählt haben. Auf diese Art haben wir also eine Funktion erhalten, die jedem $x \in [a, b]$ den Wert des Integrals $\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy$ zuordnet. Man kann zeigen, dass diese Funktion über $[a, b]$ integrierbar ist. Man erhält somit ein sogenanntes *iteriertes Integral*:

$$\int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right) dx.$$

¹Man definiert die Stetigkeit einer Funktion f mit zwei Variablen analog zum eindimensionalen Fall; siehe Abschnitt 5.2.

Durch den folgenden Satz wird die Berechnung eines zweidimensionalen Integrals auf die Berechnung eines iterierten Integrals (d.h. auf die Ausführung zweier eindimensionaler Integrationen) zurückgeführt:

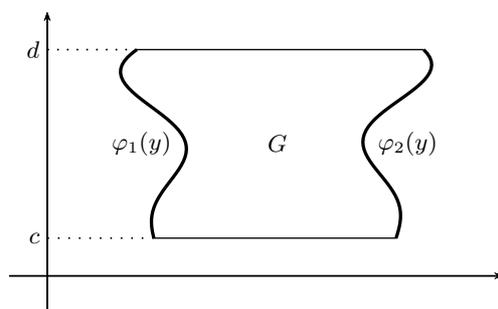
Satz 1.

$\varphi_1 : [a, b] \rightarrow \mathbb{R}$ und $\varphi_2 : [a, b] \rightarrow \mathbb{R}$ seien stetige Funktionen, für die $\varphi_1(x) \leq \varphi_2(x)$ für alle $x \in [a, b]$ gilt. Es sei $G = \{(x, y) : a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}$. Ist $f : G \rightarrow \mathbb{R}$ stetig, so gilt:

$$\iint_G f(x, y) d(x, y) = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right) dx. \quad (6.5)$$

Für den Beweis verweisen wir auf das Buch von W. Luh.

Man nennt eine Menge G , die wie oben durch zwei Funktionen $\varphi_1(x)$ und $\varphi_2(x)$ berandet ist, eine *y-projizierbare Menge*. Analog definiert man mittels zweier Funktionen $\varphi_1(y)$ und $\varphi_2(y)$, was eine *x-projizierbare Menge* $G \subseteq \mathbb{R}^2$ ist (siehe Zeichnung).



Für *x-projizierbare* Mengen gilt ein Satz 1 entsprechender Satz; die (6.5) entsprechende Formel lautet:

$$\iint_G f(x, y) d(x, y) = \int_c^d \left(\int_{\varphi_1(y)}^{\varphi_2(y)} f(x, y) dx \right) dy. \quad (6.6)$$

Eine Menge G , die sowohl *x-* als auch *y-projizierbar* ist, nennt man auch eine *Standardmenge*. Ein einfacher, aber wichtiger Spezialfall einer Standardmenge G ist ein Rechteck:

$$G = I = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}.$$

In diesem Fall folgt aus (6.5) und (6.6):

$$\iint_G f(x, y) d(x, y) = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy. \quad (6.7)$$

Beispiele:

1. Es sei $I = \{(x, y) : 0 \leq x \leq 3, 0 \leq y \leq 2\}$ mit $f(x, y) = xy + y^2$.

Dann gilt nach (6.7):

$$\iint_I (xy + y^2) d(x, y) = \int_0^3 \left(\int_0^2 (xy + y^2) dy \right) dx.$$

Für das innere Integral gilt:

$$\int_0^2 (xy + y^2) dy = \left[\frac{1}{2}xy^2 + \frac{1}{3}y^3 \right]_0^2 = 2x + \frac{8}{3}$$

Es folgt:

$$\begin{aligned} \iint_I (xy + y^2) d(x, y) &= \int_0^3 \left(2x + \frac{8}{3} \right) dx \\ &= \left[x^2 + \frac{8}{3}x \right]_0^3 \\ &= 9 + 8 \\ &= 17 \end{aligned}$$

Ebenso gut hätten wir (wegen (6.7)) erst nach x und dann nach y integrieren können:

$$\iint_I (xy + y^2) d(x, y) = \int_0^2 \left(\int_0^3 (xy + y^2) dx \right) dy.$$

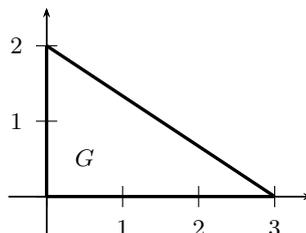
Für das innere Integral ergibt sich:

$$\begin{aligned} \int_0^3 (xy + y^2) dx &= \left[\frac{1}{2}x^2y + xy^2 \right]_0^3 \\ &= 3y^2 + \frac{9}{2}y \end{aligned}$$

Es folgt:

$$\begin{aligned} \iint_I (xy + y^2) d(x, y) &= \int_0^2 \left(3y^2 + \frac{9}{2}y \right) dy \\ &= \left[y^3 + \frac{9}{4}y^2 \right]_0^2 \\ &= 17 \end{aligned}$$

2. Es sei G das Dreieck, das durch die Punkte $(0, 0)$, $(3, 0)$ und $(0, 2)$ gebildet wird.



Gesucht ist $\iint_G (x+y) d(x,y)$. Die Gleichung der Geraden durch $(3,0)$ und $(0,2)$ lautet $y = -\frac{2}{3}x + 2$. Für $\varphi_1(x) = 0$ und $\varphi_2(x) = -\frac{2}{3}x + 2$ können wir (6.5) anwenden und erhalten:

$$\iint_G (x+y) d(x,y) = \int_0^3 \left(\int_0^{-\frac{2}{3}x+2} (x+y) dy \right) dx.$$

Für das innere Integral gilt:

$$\begin{aligned} \int_0^{-\frac{2}{3}x+2} (x+y) dy &= \left[xy + \frac{1}{2}y^2 \right]_0^{-\frac{2}{3}x+2} \\ &= -\frac{2}{3}x^2 + 2x + \frac{1}{2} \left(-\frac{2}{3}x + 2 \right)^2 \\ &= -\frac{4}{9}x^2 + \frac{2}{3}x + 2 \end{aligned}$$

Es folgt:

$$\begin{aligned} \iint_G (x+y) d(x,y) &= \int_0^3 \left(-\frac{4}{9}x^2 + \frac{2}{3}x + 2 \right) dx \\ &= \left[-\frac{4}{27}x^3 + \frac{1}{3}x^2 + 2x \right]_0^3 \\ &= -4 + 3 + 6 \\ &= 5 \end{aligned}$$

3. Es sei G das Dreieck $(0,0)$, $(a,0)$ und $(0,a)$. Gesucht ist $\iint_G e^{x+y} d(x,y)$. Die Gleichung der Geraden durch $(a,0)$ und $(0,a)$ lautet $y = -x + a$. Anwendung von (6.5) ergibt:

$$\begin{aligned} \iint_G e^{x+y} d(x,y) &= \iint_G e^x e^y d(x,y) \\ &= \int_0^a \left(\int_0^{a-x} e^x e^y dy \right) dx \\ &= \int_0^a \left(e^x \int_0^{a-x} e^y dy \right) dx \\ &= \int_0^a e^x (e^{a-x} - 1) dx \\ &= \int_0^a (e^a - e^x) dx \\ &= [e^a x - e^x]_0^a \\ &= e^a a - e^a + 1 \\ &= e^a(a-1) + 1 \end{aligned}$$

7 Komplexe Zahlen

Für viele Anwendungen (z.B. in der Elektrotechnik) reichen die reellen Zahlen nicht aus; man macht von einem größeren Zahlenbereich, der Menge \mathbb{C} der komplexen Zahlen, Gebrauch. Ein innermathematischer Grund für die Einführung komplexer Zahlen ist, dass schon so einfache Gleichungen wie $x^2 + 1 = 0$ in \mathbb{R} keine Lösung besitzen; um hier Abhilfe zu schaffen, erweitert man \mathbb{R} zum Körper \mathbb{C} der komplexen Zahlen. (Vergleichbare Erweiterungen von Zahlenbereichen hatten wir früher schon bei den Übergängen von \mathbb{N} zu \mathbb{Z} , von \mathbb{Z} zu \mathbb{Q} oder von \mathbb{Q} zu \mathbb{R} kennengelernt.)

7.1 Definition der komplexen Zahlen

Mit C wollen wir in diesem Abschnitt die Menge \mathbb{R}^2 (also die Menge aller geordneten Paare reeller Zahlen) bezeichnen, d.h.,

$$C = \{(a, b) : a, b \in \mathbb{R}\} = \mathbb{R}^2.$$

Wir definieren auf der Menge C eine Addition und eine Multiplikation.

Definition.

Für $(a, b), (c, d) \in C$ definieren wir:

$$(a, b) + (c, d) := (a + c, b + d), \quad (7.1)$$

$$(a, b) \cdot (c, d) := (ac - bd, ad + bc). \quad (7.2)$$

Satz 1.

Die Menge C bildet bezüglich der in (7.1) und (7.2) definierten Addition und Multiplikation einen Körper.

Beweis. Assoziativgesetze, Kommutativgesetze und Distributivgesetz lassen sich leicht nachrechnen; wir wollen hier darauf verzichten und empfehlen den Nachweis dieser Gesetze als Übungsaufgabe.

Existenz neutraler Elemente: Das neutrale Element bezüglich $+$ ist $(0, 0)$, das neutrale Element bezüglich \cdot ist $(1, 0)$, denn es gilt für alle $(a, b) \in C$:

$$(a, b) + (0, 0) \stackrel{(7.1)}{=} (a + 0, b + 0) = (a, b),$$

$$(a, b) \cdot (1, 0) \stackrel{(7.2)}{=} (a \cdot 1 - b \cdot 0, a \cdot 0 + b \cdot 1) = (a, b).$$

Existenz inverser Elemente: Es sei $(a, b) \in C$. Dann ist $(-a, -b)$ das zu (a, b) inverse Element der Addition, denn es gilt:

$$(a, b) + (-a, -b) \stackrel{(7.1)}{=} (a + (-a), b + (-b)) = (0, 0).$$

Ist $(a, b) \neq (0, 0)$, so ist $\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$ das zu (a, b) inverse Element der Multiplikation, denn es gilt:

$$(a, b) \cdot \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) \stackrel{(7.2)}{=} \left(a \cdot \frac{a}{a^2+b^2} - b \cdot \frac{-b}{a^2+b^2}, a \cdot \frac{-b}{a^2+b^2} + b \cdot \frac{a}{a^2+b^2}\right) = (1, 0). \quad \square$$

Wir können die Menge C zusammen mit den in (7.1) und (7.2) definierten Operationen $+$ und \cdot als eine vorläufige Definition der komplexen Zahlen ansehen. „Vorläufig“ deshalb, weil C noch einen Schönheitsfehler hat, der allerdings nur sehr klein ist und sich leicht beheben lässt: Die Menge \mathbb{R} ist offensichtlich nicht in C enthalten. Um hier Abhilfe zu schaffen, betrachten wir die Teilmenge $R = \{(a, 0) : a \in \mathbb{R}\}$ von C und beweisen den folgenden Hilfssatz.

Hilfssatz 1.

Für alle $(a_1, 0), (a_2, 0) \in R$ gilt:

$$(a_1, 0) + (a_2, 0) = (a_1 + a_2, 0) \quad \text{und} \\ (a_1, 0) \cdot (a_2, 0) = (a_1 \cdot a_2, 0).$$

Beweis.

$$(a_1, 0) + (a_2, 0) \stackrel{(7.1)}{=} (a_1 + a_2, 0 + 0) = (a_1 + a_2, 0) \quad \text{und} \\ (a_1, 0) \cdot (a_2, 0) \stackrel{(7.2)}{=} (a_1 \cdot a_2 - 0 \cdot 0, a_1 \cdot 0 + 0 \cdot a_2) = (a_1 \cdot a_2, 0). \quad \square$$

Durch die Zuordnung $a \mapsto (a, 0)$ wird \mathbb{R} bijektiv auf R abgebildet, wobei wegen Hilfssatz 1 gilt: Die Addition und Multiplikation in \mathbb{R} entspricht genau der Addition und Multiplikation in R . Man sagt: \mathbb{R} und R sind *isomorph* (strukturgleich). Wir kommen nun zur endgültigen Definition der Menge der komplexen Zahlen.

Definition.

Die Menge \mathbb{C} der komplexen Zahlen entsteht aus C dadurch, dass man R aus C entfernt und durch \mathbb{R} ersetzt. Dabei wird nur die Schreibweise geändert: Anstelle von $(a, 0)$ schreibt man a und lässt ansonsten alles beim Alten. Das Element $(0, 1)$ wird *imaginäre Einheit* genannt; als Abkürzung für $(0, 1)$ schreibt man i .

Mit diesen neuen Schreibweisen gilt Folgendes:

Hilfssatz 2.

Jedem Element $(a, b) \in C$ entspricht in \mathbb{C} das Element $a + ib$. Ferner gilt $i^2 = -1$.

Beweis. Für $(a, b) \in C$ gilt aufgrund von (7.1) und (7.2):

$$(a, b) \stackrel{(7.1)}{=} (a, 0) + (0, b) \stackrel{(7.2)}{=} (a, 0) + (0, 1)(b, 0).$$

Ersetzt man hierin (gemäß der Definition von \mathbb{C}) $(a, 0)$ durch a und $(b, 0)$ durch b und schreibt i statt $(0, 1)$, so erhält man die gewünschte Darstellung $a + ib$. Ferner gilt

$$i^2 = (0, 1)(0, 1) \stackrel{(7.2)}{=} (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0).$$

Ersetzt man hierin (gemäß der Definition von \mathbb{C}) $(-1, 0)$ durch -1 , so folgt $i^2 = -1$. \square

Damit haben wir die komplexen Zahlen in ihrer üblichen Darstellung $z = a + ib$ vorliegen. In dieser Schreibweise lassen sich Addition und Multiplikation zweier komplexer Zahlen $z_1 = a + ib$ und $z_2 = c + id$ so darstellen:

$$(a + ib) + (c + id) = (a + c) + i(b + d) \tag{7.3}$$

$$(a + ib) \cdot (c + id) = ac + iad + ibc + i^2bd = (ac - bd) + i(ad + bc). \tag{7.4}$$

(Die erste Gleichheit in (7.4) folgt durch „Ausmultiplizieren“ nach dem Distributivgesetz, das ja wegen Satz 1 gilt; die zweite Gleichheit gilt wegen $i^2 = -1$.)

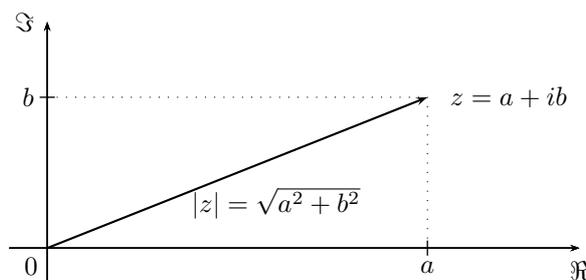
Definition.

Es sei $z = a + ib \in \mathbb{C}$. Dann heißt

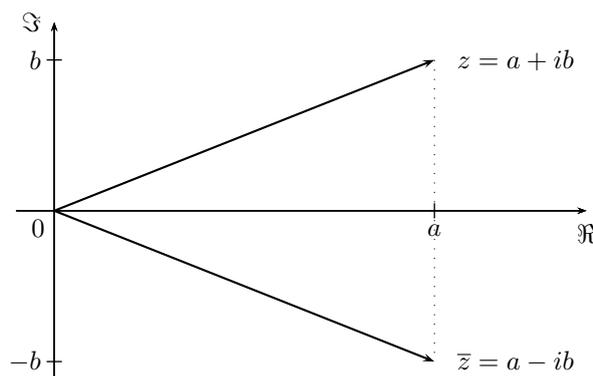
- (1) a *Realteil* von z (Bezeichnung: $a = \operatorname{Re} z$ oder $a = \Re z$);
- (2) b *Imaginärteil* von z (Bezeichnung: $b = \operatorname{Im} z$ oder $b = \Im z$);
- (3) $|z| = \sqrt{a^2 + b^2}$ *absoluter Betrag* von z ;
- (4) $\bar{z} = a - ib$ *konjugiert komplexe Zahl* zu z .

7.2 Veranschaulichung der komplexen Zahlen als Punkte der Gaußschen Zahlenebene

Die reellen Zahlen haben wir als Punkte einer Geraden veranschaulicht („Zahlengerade“). Entsprechend lassen sich die komplexen Zahlen als Punkte einer Ebene veranschaulichen. (Man spricht von der „Gaußschen Zahlenebene“.) Hierzu legt man auf die übliche Art ein rechtwinkliges kartesisches¹ Koordinatensystem in der Zeichenebene zugrunde. Komplexe Zahlen $z = a + ib$ und Punkte mit den Koordinaten (a, b) entsprechen sich dann umkehrbar eindeutig. Man stellt eine komplexe Zahl $a + ib$ oft auch als einen Ortsvektor dar, d.h., als einen Pfeil vom Punkt $(0, 0)$ zum Punkt (a, b) ; siehe Zeichnung. Man nennt die x -Achse *reelle Achse* und die y -Achse *imaginäre Achse*. Der Betrag $|z|$ einer komplexen Zahl lässt sich („Lehrsatz des Pythagoras“) als Abstand von $z = a + ib$ zum Punkt $(0, 0)$ veranschaulichen.

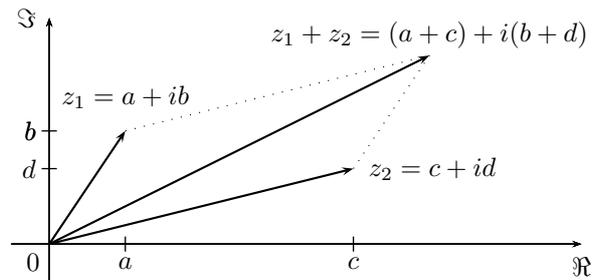


Die konjugiert komplexe Zahl \bar{z} zu z erhält man durch Spiegelung von z an der x -Achse:

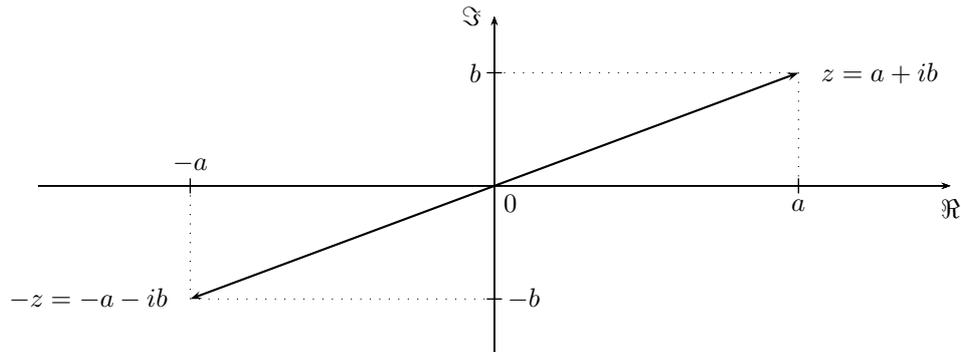


Die Addition komplexer Zahlen lässt sich als Vektoraddition („Parallelogramm der Kräfte“) veranschaulichen:

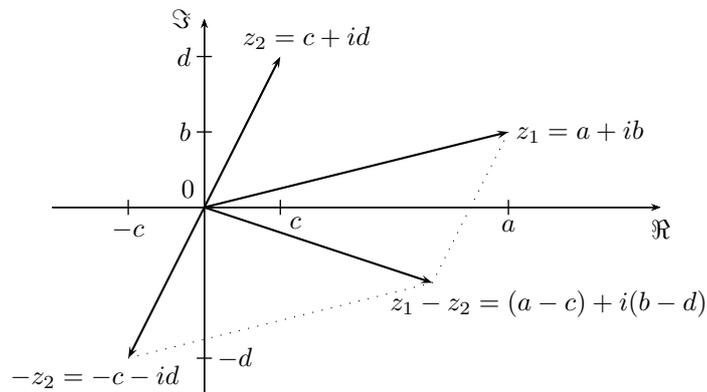
¹„Kartesisch“ bedeutet, dass auf beiden Achsen dieselbe Längeneinheit gewählt ist, d.h., der Punkt $(1, 0)$ hat denselben Abstand zum Ursprung wie der Punkt $(0, 1)$.



Veranschaulichung des Negativen $-z$ von z :



Zur Veranschaulichung der Differenz $z_1 - z_2$ zweier komplexer Zahlen betrachten wir zunächst die komplexe Zahl $-z_2$ und addieren dann die Zahlen z_1 und $-z_2$:



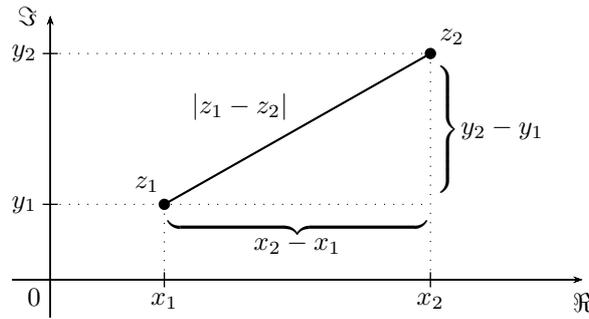
Definition.

Es seien $z_1, z_2 \in \mathbb{C}$. Unter dem *Abstand* von z_1 und z_2 versteht man die Zahl $|z_1 - z_2|$.

Diese Definition ist im Einklang mit der gewohnten Definition des Abstands zweier Punkte in der Ebene; es gilt nämlich für $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$:

$$|z_1 - z_2| = |x_1 - x_2 + i(y_1 - y_2)| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

Dies ist gerade die übliche Definition des Abstands zweier Punkte (x_1, y_1) und (x_2, y_2) in der Ebene („Lehrsatz des Pythagoras“).



Als Nächstes zeigen wir die folgende nützliche Identität, die einen Zusammenhang zwischen dem Betrag $|z|$ und dem Produkt $z \cdot \bar{z}$ herstellt:

$$z \cdot \bar{z} = a^2 + b^2 = |z|^2 \quad (\text{für } z = a + ib). \quad (7.5)$$

Beweis.

$$z \cdot \bar{z} = (a + ib)(a - ib) = a^2 - iab + iba + b^2 = a^2 + b^2 = |z|^2. \quad \square$$

Insbesondere gilt also: $z \cdot \bar{z}$ ist immer eine reelle Zahl.

Für den Quotienten zweier komplexer Zahlen $z_1 = a + ib$ und $z_2 = c + id$ (mit $c + id \neq 0$) gilt:

$$\frac{z_1}{z_2} = \frac{a + ib}{c + id} = \frac{(a + ib)(c - id)}{(c + id)(c - id)} = \frac{(ac + bd) + i(bc - ad)}{c^2 + d^2} = \frac{ac + bd}{c^2 + d^2} + i \frac{bc - ad}{c^2 + d^2}. \quad (7.6)$$

Hinweis. Anstatt die Formel $\frac{a+ib}{c+id} = \frac{ac+bd}{c^2+d^2} + i \frac{bc-ad}{c^2+d^2}$ auswendig zu lernen, ist es besser, sich den ersten Schritt ihrer Herleitung zu merken: *Man erweitert $\frac{a+ib}{c+id}$ mit der konjugiert komplexen Zahl des Nenners.* Wegen (7.5) steht dann eine reelle Zahl im Nenner – nämlich $c^2 + d^2$ – und man muss nur noch den Zähler ausmultiplizieren.

Beispiel. $z = \frac{3+4i}{2+3i}$ soll in die Form $z = a + ib$ gebracht werden. Wir erweitern mit $2 - 3i$:

$$z = \frac{(3 + 4i)(2 - 3i)}{(2 + 3i)(2 - 3i)} = \frac{6 - 9i + 8i + 12}{4 + 9} = \frac{18}{13} - \frac{1}{13}i.$$

Im nächsten Satz stellen wir einige Rechenregeln für das Rechnen mit konjugiert komplexen Zahlen zusammen.

Satz 2.

Es seien $z_1, z_2 \in \mathbb{C}$. Dann gilt:

- (a) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$
- (b) $\overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2$
- (c) $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$
- (d) $\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}$

Beweis. Als Übungsaufgabe empfohlen.

Im folgenden Satz werden die wichtigsten Eigenschaften des absoluten Betrags $|z|$ einer komplexen Zahl zusammengestellt.

Satz 3.Für alle $z_1, z_2 \in \mathbb{C}$ gilt:

- (a) $|z_1| \geq 0$; $|z_1| = 0$ gilt genau dann, wenn $z_1 = 0$
- (b) $|z_1 z_2| = |z_1| \cdot |z_2|$
- (c) $\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}$ (für $z_2 \neq 0$)
- (d) $|z_1 + z_2| \leq |z_1| + |z_2|$ (Dreiecksungleichung)
- (e) $|z_1 - z_2| \geq ||z_1| - |z_2||$

Beweis. Es seien $z_1 = a_1 + ib_1$ und $z_2 = a_2 + ib_2$.(a) Dies folgt unmittelbar aus der Definition $|z_1| = \sqrt{a_1^2 + b_1^2}$.

(b) Es gilt:

$$\begin{aligned}
 |z_1 z_2|^2 &= \left| (a_1 + ib_1)(a_2 + ib_2) \right|^2 \\
 &= \left| a_1 a_2 - b_1 b_2 + i(a_1 b_2 + a_2 b_1) \right|^2 \\
 &= (a_1 a_2 - b_1 b_2)^2 + (a_1 b_2 + a_2 b_1)^2 \\
 &= a_1^2 a_2^2 + b_1^2 b_2^2 + a_1^2 b_2^2 + a_2^2 b_1^2 \\
 &= (a_1^2 + b_1^2)(a_2^2 + b_2^2) \\
 &= |z_1|^2 \cdot |z_2|^2
 \end{aligned}$$

Zieht man die Quadratwurzel, so folgt die Behauptung.

(c) Wegen (b) gilt $|z_1| = \left| \frac{z_1}{z_2} \cdot z_2 \right| = \left| \frac{z_1}{z_2} \right| |z_2|$ woraus (c) folgt.

(d) Es gilt:

$$\begin{aligned}
 |z_1 + z_2|^2 &= (a_1 + a_2)^2 + (b_1 + b_2)^2 = |z_1|^2 + |z_2|^2 + 2(a_1 a_2 + b_1 b_2) \quad \text{sowie} \\
 (|z_1| + |z_2|)^2 &= |z_1|^2 + |z_2|^2 + 2|z_1||z_2| = |z_1|^2 + |z_2|^2 + 2\sqrt{a_1^2 + b_1^2}\sqrt{a_2^2 + b_2^2}.
 \end{aligned}$$

Ferner gilt (siehe unten):

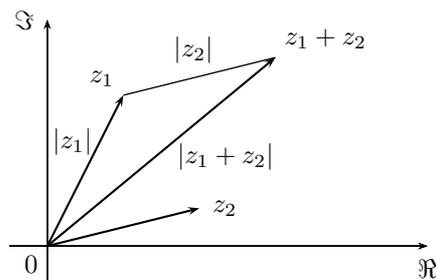
$$a_1 a_2 + b_1 b_2 \leq \sqrt{a_1^2 + b_1^2} \sqrt{a_2^2 + b_2^2}.$$

Insgesamt haben wir also $|z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$, woraus durch Ziehen der Quadratwurzel die Behauptung folgt.Es bleibt also $a_1 a_2 + b_1 b_2 \leq \sqrt{a_1^2 + b_1^2} \sqrt{a_2^2 + b_2^2}$ nachzuweisen. Falls $a_1 a_2 + b_1 b_2 < 0$ gilt, so ist dies gewiss richtig. Falls $a_1 a_2 + b_1 b_2 \geq 0$ gilt, folgt die Aussage durch äquivalente Umformungen:

$$\begin{aligned}
 a_1 a_2 + b_1 b_2 &\leq \sqrt{a_1^2 + b_1^2} \sqrt{a_2^2 + b_2^2} \\
 \Leftrightarrow (a_1 a_2 + b_1 b_2)^2 &\leq (a_1^2 + b_1^2)(a_2^2 + b_2^2) \\
 \Leftrightarrow a_1^2 a_2^2 + b_1^2 b_2^2 + 2a_1 a_2 b_1 b_2 &\leq a_1^2 a_2^2 + a_1^2 b_2^2 + a_2^2 b_1^2 + b_1^2 b_2^2 \\
 \Leftrightarrow 0 &\leq a_1^2 b_2^2 - 2a_1 a_2 b_1 b_2 + a_2^2 b_1^2 = (a_1 b_2 - a_2 b_1)^2.
 \end{aligned}$$

(e) Wegen (d) gilt $|z_1| = |z_1 - z_2 + z_2| \leq |z_1 - z_2| + |z_2|$, woraus (e) folgt. \square

Geometrische Interpretation von (d): In einem Dreieck ist die Summe der Längen zweier Seiten immer mindestens so groß wie die Länge der dritten Seite (siehe Zeichnung).



Es folgen noch einige **Beispiele** für das Rechnen mit komplexen Zahlen:

1. Aufgabe:

Man zerlege die folgenden komplexen Zahlen in Real- und Imaginärteil, d.h., man stelle die folgenden komplexen Zahlen in der Form $a + ib$ dar.

$$a) (3 + 7i)(2 - 2i) \quad b) \frac{3 + 7i}{2 - 2i} \quad c) \frac{1}{2 - i}$$

Lösung:

a) Man muss nur ausmultiplizieren und dabei $i^2 = -1$ berücksichtigen:

$$(3 + 7i)(2 - 2i) = 6 - 6i + 14i + 14 = 20 + 8i.$$

b) Man erweitert mit dem konjugiert komplexen Wert des Nenners:

$$\frac{3 + 7i}{2 - 2i} = \frac{(3 + 7i)(2 + 2i)}{(2 - 2i)(2 + 2i)} = \frac{6 + 6i + 14i - 14}{8} = -1 + \frac{5}{2}i.$$

c) Man erweitert mit dem konjugiert komplexen Wert des Nenners:

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

2. Aufgabe:

Man bestimme die Lösung z der folgenden Gleichung:

$$\left(\frac{3 + 3i}{2 + i} + \frac{8 - 3i}{2i} \right) \bar{z} = 2 + 6i.$$

Lösung:

Zunächst wird die Gleichung nach \bar{z} umgestellt:

$$\begin{aligned} & \frac{(3 + 3i)2i + (8 - 3i)(2 + i)}{(2 + i)2i} \bar{z} = 2 + 6i \\ \Leftrightarrow & \frac{6i - 6 + 16 + 8i - 6i + 3}{4i - 2} \bar{z} = 2 + 6i \\ \Leftrightarrow & \frac{13 + 8i}{4i - 2} \bar{z} = 2 + 6i \\ \Leftrightarrow & \bar{z} = \frac{(2 + 6i)(4i - 2)}{13 + 8i}. \end{aligned}$$

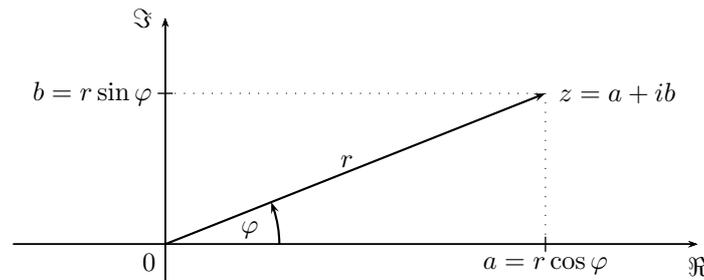
Anschließend wird \bar{z} in die Form $a + ib$ gebracht:

$$\begin{aligned} \bar{z} &= \frac{(2 + 6i)(4i - 2)}{13 + 8i} = \frac{8i - 4 - 24 - 12i}{13 + 8i} = \frac{-28 - 4i}{13 + 8i} \\ &= \frac{-(28 + 4i)(13 - 8i)}{(13 + 8i)(13 - 8i)} = \frac{-364 + 224i - 52i - 32}{13^2 + 8^2} = -\frac{396}{233} + \frac{172}{233}i. \end{aligned}$$

Hieraus ergibt sich sofort $z = -\frac{396}{233} - \frac{172}{233}i$.

7.3 Polarkoordinaten

Es sei $z = a + ib$ eine komplexe Zahl $\neq 0$; $r = |z|$ sei der Betrag von z (d.h. der Abstand von z zum Nullpunkt des Koordinatensystems); mit $\varphi \in [0, 2\pi)$ sei der Winkel (im Bogenmaß) bezeichnet, der von der positiven x -Achse und der Verbindungsstrecke zwischen dem Nullpunkt und z gebildet wird. Anders gesagt: φ ist der Winkel, um den man die positive x -Achse entgegen dem Uhrzeigersinn drehen muss, bis sie z enthält:



Dann gilt $\cos \varphi = \frac{a}{r}$ und $\sin \varphi = \frac{b}{r}$; es folgt:

$$a = r \cos \varphi \quad \text{und} \quad b = r \sin \varphi.$$

Hieraus erhält man die *Polarkoordinatendarstellung* von z :

Polarkoordinatendarstellung von z

$$z = r(\cos \varphi + i \sin \varphi)$$

Da \cos und \sin 2π -periodisch sind, gilt für alle $k \in \mathbb{Z}$:

$$z = r(\cos(\varphi + 2k\pi) + i \sin(\varphi + 2k\pi)). \tag{7.7}$$

Man nennt φ das *Argument* von z . Auch Darstellungen der Form (7.7) werden als Polarkoordinatendarstellungen von z bezeichnet; einen Winkel $\varphi + 2k\pi$ (der sich von φ ja nur um ein ganzzahliges Vielfaches des „vollen Winkels“ 2π unterscheidet), bezeichnet man ebenfalls als Argument von z . Als Abkürzung benutzt man $\arg z$.

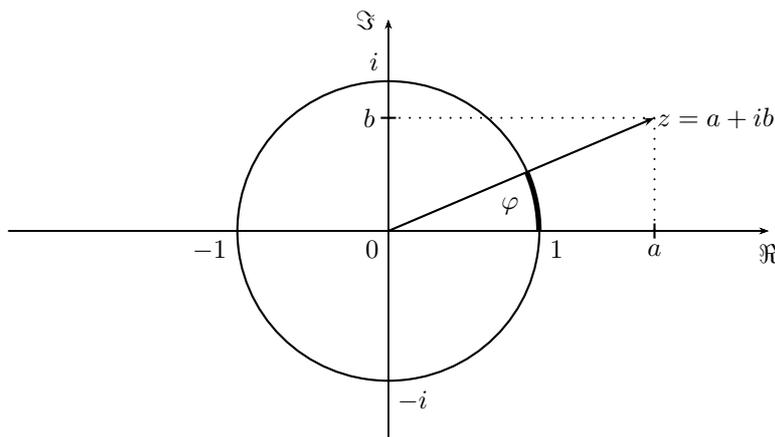
Berechnung von Betrag und Argument einer komplexen Zahl

Die komplexe Zahl $z \neq 0$ sei in der Form $z = a + ib$ gegeben. Um z in Polarkoordinatendarstellung anzugeben, sind der Betrag $r = |z|$ und das Argument φ von z zu berechnen. Wie man $r = |z|$ berechnet, ist klar:

$$r = |z| = \sqrt{a^2 + b^2}.$$

Die Berechnung des Arguments von z kann (wegen $\cos \varphi = \frac{a}{r}$) mithilfe der Arcuscosinus-Funktion erfolgen:

$$\varphi = \begin{cases} \arccos\left(\frac{a}{r}\right) & , \text{ falls } b \geq 0, \\ 2\pi - \arccos\left(\frac{a}{r}\right) & , \text{ falls } b < 0. \end{cases}$$



Geometrische Interpretation der Multiplikation komplexer Zahlen

Satz 4.

Es seien $z_1 = r_1(\cos \varphi_1 + i \sin \varphi_1)$ und $z_2 = r_2(\cos \varphi_2 + i \sin \varphi_2)$. Für das Produkt $z_1 z_2$ gilt dann:

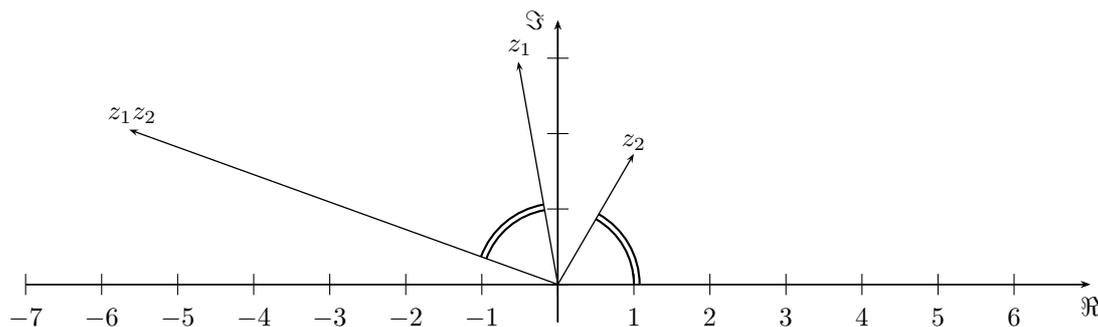
$$z_1 z_2 = r_1 r_2 \left(\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2) \right). \quad (7.8)$$

Beweis.

$$\begin{aligned} z_1 z_2 &= r_1 (\cos \varphi_1 + i \sin \varphi_1) \cdot r_2 (\cos \varphi_2 + i \sin \varphi_2) \\ &= r_1 r_2 (\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2 + i(\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2)) \\ &= r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)) \end{aligned}$$

Die letzte Gleichheit gilt aufgrund der Additionstheoreme für sin und cos. \square

Formuliert man das Ergebnis (7.8) in Worten, so erhält man die *geometrische Interpretation der Multiplikation komplexer Zahlen*: Bei der Multiplikation komplexer Zahlen werden die Beträge multipliziert und die Argumente addiert. (Siehe Zeichnung: Dort ist $|z_1| = 3$, $|z_2| = 2$, also $|z_1 z_2| = 6$. Sind die Argumente von z_1 und z_2 mit φ_1 und φ_2 bezeichnet, so ist $\varphi_1 + \varphi_2$ das Argument von $z_1 z_2$.)



Wie im Reellen werden in \mathbb{C} die Potenzen mit ganzzahligen Exponenten definiert. Es seien $n \in \mathbb{N}$ und $z \in \mathbb{C}$; man definiert:

$$\begin{aligned} z^n &:= \underbrace{z \cdot z \cdot \dots \cdot z}_{n \text{ Faktoren}} \\ z^{-n} &:= \frac{1}{z^n} \quad (\text{für } z \neq 0) \\ z^0 &:= 1 \end{aligned}$$

Satz 5.

Es seien $z \in \mathbb{C}$, $z \neq 0$, und $m \in \mathbb{Z}$. Es gelte $z = r(\cos \varphi + i \sin \varphi)$. Dann folgt:

$$z^m = r^m (\cos(m\varphi) + i \sin(m\varphi)). \quad (7.9)$$

Beweis. Für $m = 0$ und $m = 1$ gilt (7.9) klarerweise, für $m \geq 2$ folgt (7.9) durch Anwendung von Satz 4.

Es sei nun $m < 0$. Wir setzen $n = -m$. Dann gilt $n \in \mathbb{N}$ und es folgt (da wir (7.9) ja schon für natürliche Zahlen gezeigt haben) $z^n = r^n (\cos(n\varphi) + i \sin(n\varphi))$. Also:

$$\begin{aligned} z^m &= z^{-n} = \frac{1}{z^n} \\ &= \frac{1}{r^n (\cos(n\varphi) + i \sin(n\varphi))} \\ &= r^{-n} \frac{\cos(n\varphi) - i \sin(n\varphi)}{(\cos(n\varphi) + i \sin(n\varphi))(\cos(n\varphi) - i \sin(n\varphi))} \\ &= r^{-n} \frac{\cos(n\varphi) - i \sin(n\varphi)}{\cos^2(n\varphi) + \sin^2(n\varphi)} \\ &= r^{-n} (\cos(n\varphi) - i \sin(n\varphi)) \\ &= r^m (\cos(-m\varphi) - i \sin(-m\varphi)) \\ &= r^m (\cos(m\varphi) + i \sin(m\varphi)). \quad \square \end{aligned}$$

Ist in (7.9) speziell $r = |z| = 1$, so erhält man die *Moivresche Formel*:

$$(\cos \varphi + i \sin \varphi)^m = \cos(m\varphi) + i \sin(m\varphi) \quad (\text{für } m \in \mathbb{Z}). \quad (7.10)$$

7.4 Quadratische Gleichungen

In diesem Abschnitt geht es um die Lösung von quadratischen Gleichungen mit reellen Koeffizienten. Zunächst einige **Beispiele**:

1. $x^2 - 5 = 0$

Umstellen der Gleichung ergibt $x^2 = 5$, die beiden Lösungen lauten also $x_1 = \sqrt{5}$ und $x_2 = -\sqrt{5}$.

2. $x^2 + 5 = 0$

Umstellen der Gleichung ergibt $x^2 = -5$, gesucht ist also eine Zahl, deren Quadrat gleich -5 ist. In \mathbb{R} gibt es keine derartige Zahl x , in \mathbb{C} hingegen schon, nämlich $x_1 = i\sqrt{5}$ und $x_2 = -i\sqrt{5}$.

3. $x^2 + 3x + 1 = 0$

Wir verwenden die Methode der *quadratischen Ergänzung*:

$$\begin{aligned}x^2 + 3x + 1 &= 0 \\ \Leftrightarrow x^2 + 3x + \left(\frac{3}{2}\right)^2 + 1 - \left(\frac{3}{2}\right)^2 &= 0 \\ \Leftrightarrow \left(x + \frac{3}{2}\right)^2 - \frac{5}{4} &= 0 \\ \Leftrightarrow \left(x + \frac{3}{2}\right)^2 &= \frac{5}{4}\end{aligned}$$

Wurzelziehen ergibt $x + \frac{3}{2} = \sqrt{\frac{5}{4}}$ bzw. $x + \frac{3}{2} = -\sqrt{\frac{5}{4}}$; die beiden Lösungen lauten also $x_{1/2} = -\frac{3}{2} \pm \sqrt{\frac{5}{4}}$.

4. $x^2 + 4x + 9 = 0$

Quadratische Ergänzung ergibt:

$$\begin{aligned}x^2 + 4x + 9 &= 0 \\ \Leftrightarrow x^2 + 4x + \left(\frac{4}{2}\right)^2 + 9 - \left(\frac{4}{2}\right)^2 &= 0 \\ \Leftrightarrow (x + 2)^2 + 5 &= 0 \\ \Leftrightarrow (x + 2)^2 &= -5\end{aligned}$$

Wurzelziehen ergibt (Man vergleiche auch Beispiel 2!) $x + 2 = i\sqrt{5}$ bzw. $x + 2 = -i\sqrt{5}$; die beiden Lösungen lauten also $x_{1/2} = -2 \pm i\sqrt{5}$.

Nach diesen Beispielen betrachten wir den **allgemeinen Fall**: Zu lösen ist (für beliebige $p, q \in \mathbb{R}$) die quadratische Gleichung

$$x^2 + px + q = 0.$$

Quadratische Ergänzung ergibt:

$$\begin{aligned}x^2 + px + q &= 0 \\ \Leftrightarrow x^2 + px + \left(\frac{p}{2}\right)^2 + q - \left(\frac{p}{2}\right)^2 &= 0 \\ \Leftrightarrow \left(x + \frac{p}{2}\right)^2 + q - \frac{p^2}{4} &= 0 \\ \Leftrightarrow \left(x + \frac{p}{2}\right)^2 &= \frac{p^2}{4} - q\end{aligned}$$

1. Fall: $\frac{p^2}{4} - q \geq 0$:

Wie in Beispiel 3 erhält man durch Wurzelziehen $x + \frac{p}{2} = \sqrt{\frac{p^2}{4} - q}$ bzw. $x + \frac{p}{2} = -\sqrt{\frac{p^2}{4} - q}$. Die Lösungen lauten also:

$$x_{1/2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}.$$

2. Fall: $\frac{p^2}{4} - q < 0$:

Wie in Beispiel 4 erhält man durch Wurzelziehen die beiden Lösungen:

$$x_{1/2} = -\frac{p}{2} \pm i\sqrt{q - \frac{p^2}{4}}.$$

Bemerkungen:

- (a) Im ersten Fall ist der Sonderfall $\frac{p^2}{4} - q = 0$ enthalten; in diesem Fall gilt $x_1 = x_2$ („doppelte Nullstelle“).
- (b) Im 2. Fall liegt ein Paar von konjugiert komplexen Lösungen vor.
- (c) Häufig findet man die Lösungsformeln auch in der folgenden Form:

$$x_{1/2} = -\frac{p}{2} \pm \frac{\sqrt{p^2 - 4q}}{2} \quad (\text{im 1. Fall}) \quad \text{bzw.} \quad x_{1/2} = -\frac{p}{2} \pm i \frac{\sqrt{4q - p^2}}{2} \quad (\text{im 2. Fall}).$$

7.5 Der Fundamentalsatz der Algebra

Es sei $P(z) = a_0 + a_1z + a_2z^2 + \dots + a_nz^n$ ($a_n \neq 0$) ein Polynom vom Grad n , wobei $a_i \in \mathbb{C}$ ($i = 0, \dots, n$) gelte und z eine komplexe Unbestimmte sei. Man nennt $z_0 \in \mathbb{C}$ eine *Nullstelle* von P , falls $P(z_0) = 0$ gilt. Ist z_0 eine Nullstelle von P , so kann man den Linearfaktor $(z - z_0)$ „abdividieren“, d.h., man kann $P(z)$ zerlegen in $P(z) = P_1(z)(z - z_0)$, wobei $P_1(z)$ ein Polynom vom Grad $n - 1$ ist. Möglicherweise ist z_0 auch eine Nullstelle von $P_1(z)$; in diesem Falle kann man $(z - z_0)$ ein weiteres Mal „abdividieren“ und erhält so eine Darstellung $P(z) = P_2(z)(z - z_0)^2$, wobei $P_2(z)$ ein Polynom vom Grad $n - 2$ ist. Setzt man diese Überlegung fort, so wird man zu folgender Definition geführt.

Definition.

Die *Vielfachheit* einer Nullstelle z_0 von $P(z)$ ist die größte Zahl k , für die

$$P(z) = P_k(z)(z - z_0)^k$$

gilt, wobei $P_k(z)$ ein Polynom vom Grad $n - k$ ist.

Mit anderen Worten: z_0 hat die Vielfachheit k , wenn man $(z - z_0)$ k -mal „abdividieren“ kann, aber nicht $(k + 1)$ -mal.

Satz 6 (Nullstellensatz für Polynome).

Es sei $P(z) = a_0 + a_1z + a_2z^2 + \dots + a_nz^n$ ($a_n \neq 0$) ein Polynom vom Grad $n \geq 1$ mit $a_i \in \mathbb{C}$ ($i = 0, \dots, n$). Dann existiert ein $z_0 \in \mathbb{C}$, für das $P(z_0) = 0$ gilt. (Mit anderen Worten: P besitzt mindestens eine Nullstelle $z_0 \in \mathbb{C}$.)

Satz 6 wird auch **Fundamentalsatz der Algebra** genannt. Der Beweis dieses Satzes erfordert weitergehende Hilfsmittel – wir müssen ihn daher weglassen.

Als Folgerung aus Satz 6 erhält man²:

Satz 7 (Zerlegungssatz für Polynome).

Es sei $P(z) = a_0 + a_1z + a_2z^2 + \dots + a_nz^n$ ($a_n \neq 0$) ein Polynom vom Grad $n \geq 0$ mit $a_i \in \mathbb{C}$ ($i = 0, \dots, n$). Dann besitzt $P(z)$ die Darstellung

$$P(z) = a_n \cdot (z - z_0)^{k_0} \cdot (z - z_1)^{k_1} \cdot \dots \cdot (z - z_r)^{k_r}, \quad (7.11)$$

wobei z_0, \dots, z_r die verschiedenen Nullstellen von P sind und k_j ($j = 0, \dots, r$) die Vielfachheit der Nullstelle z_j ist.

Beweis. Hat P den Grad $n = 0$, so gilt $P(z) = a_0 \neq 0$; P hat also keine Nullstellen und die Darstellung (7.11) lautet einfach $P(z) = a_0$.

²Häufig wird auch diese Folgerung als „Fundamentalsatz der Algebra“ bezeichnet.

Hat P den Grad $n \geq 1$, so besitzt P nach Satz 6 eine Nullstelle z_0 . Es sei k_0 die Vielfachheit von z_0 . Dann gilt also $P(z) = Q(z) \cdot (z - z_0)^{k_0}$, wobei $Q(z)$ ein Polynom vom Grad $n - k_0$ ist. Ist $k_0 = n$, so gilt $Q(z) = a_n$ und wir sind fertig, da wir $P(z)$ wie in (7.11) dargestellt haben.

Ist $k_0 < n$, so ist $Q(z)$ ein Polynom vom Grad ≥ 1 . Wegen Satz 6 hat $Q(z)$ also eine Nullstelle z_1 , deren Vielfachheit mit k_1 bezeichnet sei. Hieraus ergibt sich $Q(z) = Q_1(z)(z - z_1)^{k_1}$, woraus man $P(z) = Q_1(z) \cdot (z - z_1)^{k_1} \cdot (z - z_0)^{k_0}$ erhält. Hierbei ist $Q_1(z)$ ein Polynom vom Grad $n - (k_1 + k_0)$. Falls $n = k_1 + k_0$ gilt, so folgt $Q_1(z) = a_n$ und wir sind fertig.

Andernfalls lässt sich von $Q_1(z)$ ein Faktor $(z - z_2)^{k_2}$ abspalten und man kann wie oben weiter verfahren. Nach endlich vielen Schritten erreicht man auf diese Art die Darstellung (7.11). \square

Da P den Grad n hat, gilt in (7.11) $k_0 + k_1 + \dots + k_r = n$.

Es sei noch auf Folgendes hingewiesen: Ist P ein Polynom mit *reellen Koeffizienten* (d.h., $a_i \in \mathbb{R}$, $i = 0, \dots, n$), so gilt: Ist z_0 eine Nullstelle von P , so ist immer auch \bar{z}_0 eine Nullstelle von P .

Der Grund hierfür ist einfach: Es gelte $P(z_0) = 0$, d.h., $a_0 + a_1 z_0 + a_2 z_0^2 + \dots + a_n z_0^n = 0$. Geht man in dieser Gleichung auf beiden Seiten zur konjugiert komplexen Zahl über, so ergibt sich (wegen Satz 2 und da $\bar{\bar{r}} = r$ für alle $r \in \mathbb{R}$ gilt):

$$\begin{aligned} 0 &= \bar{0} \\ &= \overline{a_0 + a_1 z_0 + a_2 z_0^2 + \dots + a_n z_0^n} \\ &= \bar{a}_0 + \bar{a}_1 \bar{z}_0 + \bar{a}_2 \bar{z}_0^2 + \dots + \bar{a}_n \bar{z}_0^n \\ &= a_0 + a_1 \bar{z}_0 + a_2 \bar{z}_0^2 + \dots + a_n \bar{z}_0^n \\ &= a_0 + a_1 \bar{z}_0 + a_2 (\bar{z}_0)^2 + \dots + a_n (\bar{z}_0)^n \\ &= P(\bar{z}_0). \end{aligned}$$

Wie behauptet folgt $P(\bar{z}_0) = 0$.

7.6 Die n-ten Wurzeln einer komplexen Zahl

Definition.

Es seien eine komplexe Zahl $a = |a|(\cos \varphi + i \sin \varphi)$ und $n \in \mathbb{N}$ gegeben. Jede Lösung $z \in \mathbb{C}$ der Gleichung $z^n = a$ heißt *n-te Wurzel* von a .

Satz 8.

Für $a = |a|(\cos \varphi + i \sin \varphi)$, $a \neq 0$, und $n \in \mathbb{N}$ hat die Gleichung $z^n = a$ genau n verschiedene Lösungen z_0, z_1, \dots, z_{n-1} ; diese sind gegeben durch:

$$z_k = \sqrt[n]{|a|} \left(\cos \left(\frac{\varphi}{n} + \frac{2\pi k}{n} \right) + i \sin \left(\frac{\varphi}{n} + \frac{2\pi k}{n} \right) \right) \quad (\text{für } k = 0, 1, \dots, n-1). \quad (7.12)$$

Beweis. Für $k \in \{0, \dots, n-1\}$ sei z_k wie in (7.12). Aus Satz 5 erhält man

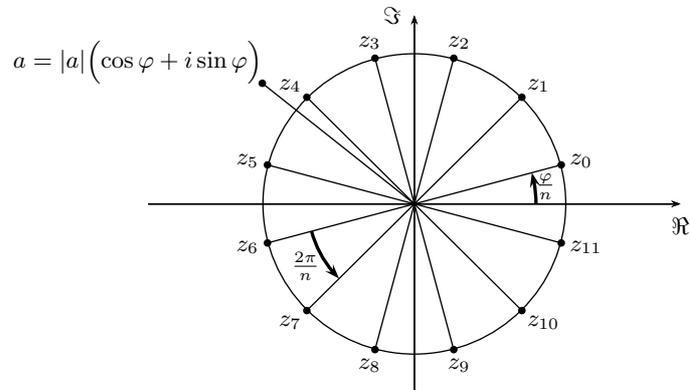
$$\begin{aligned} z_k^n &= |a| \left(\cos(\varphi + 2k\pi) + i \sin(\varphi + 2k\pi) \right) \\ &= |a| \left(\cos \varphi + i \sin \varphi \right) \\ &= a. \end{aligned}$$

Bei den in (7.12) angegebenen komplexen Zahlen z_0, \dots, z_{n-1} handelt es sich also um Lösungen der Gleichung $z^n = a$. Diese sind alle verschieden, da sie unterschiedliche Argumente besitzen. (Man beachte:

Bei den Zahlen $\frac{2k\pi}{n}$ ($k = 0, \dots, n-1$), die jeweils zum festen Wert $\frac{\varphi}{n}$ addiert werden, handelt es sich um n unterschiedliche Zahlen aus dem Intervall $[0, 2\pi)$. Da es in \mathbb{C} nicht mehr als n verschiedene Lösungen von $z^n = a$ geben kann (Wieso nämlich?), folgt die Behauptung. \square

Durch (7.12) haben wir die n -ten Wurzeln z_0, z_1, \dots, z_{n-1} von $a \in \mathbb{C}$ bestimmt. Insbesondere hat sich ergeben, dass z_0, z_1, \dots, z_{n-1} in gleichen Abständen auf einem Kreis mit Radius $\sqrt[n]{|a|}$ liegen.

Skizze für den Fall $n = 12$:



Spezialfall: $a = 1$

Man bezeichnet die Lösungen der Gleichung $z^n = 1$ als n -te Einheitswurzeln und bezeichnet sie mit $\omega_0, \omega_1, \dots, \omega_{n-1}$. Nach (7.12) gilt:

$$\omega_k = \cos\left(\frac{2\pi k}{n}\right) + i \sin\left(\frac{2\pi k}{n}\right) \quad (\text{für } k = 0, 1, \dots, n-1). \quad (7.13)$$

Die n -ten Einheitswurzeln bilden ein regelmäßiges n -Eck, dessen Eckpunkte auf dem Einheitskreis liegen.

7.7 Die komplexe Exponentialfunktion

Diese Funktion wird folgendermaßen definiert: Für $z = x + iy \in \mathbb{C}$ ($x, y \in \mathbb{R}$) definiert man

$$e^z = e^{x+iy} := e^x (\cos y + i \sin y). \quad (7.14)$$

Die komplexe Exponentialfunktion ordnet also jeder komplexen Zahl $x + iy$ diejenige komplexe Zahl zu, deren Betrag gleich e^x und deren Argument gleich y ist:

$$x + iy \mapsto e^{x+iy} = e^x (\cos y + i \sin y).$$

Es gilt:

- 1) Für reelle Zahlen stimmen die komplexe und die reelle Exponentialfunktion überein, denn für $y = 0$ gilt:

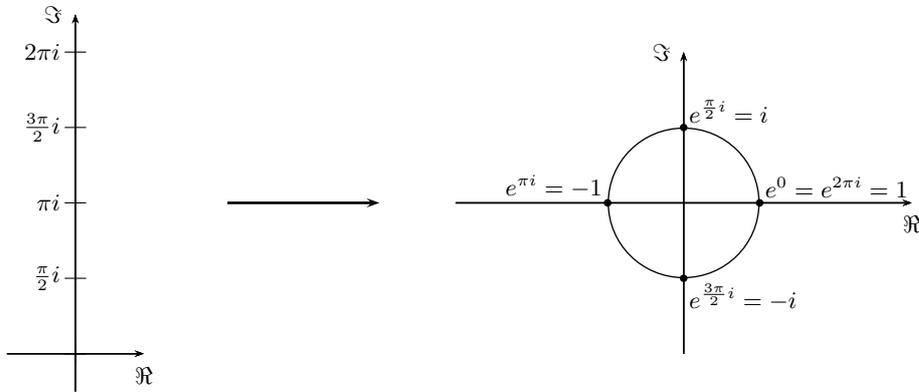
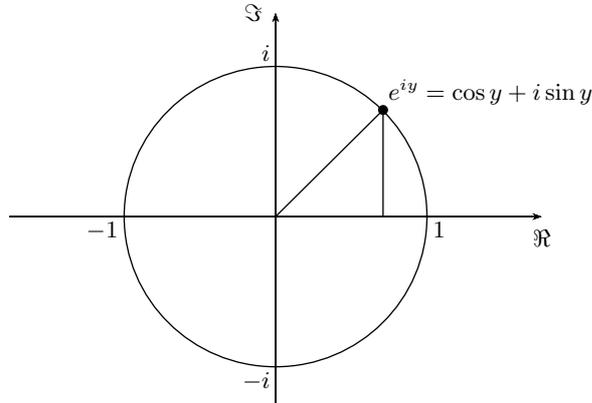
$$e^x (\cos y + i \sin y) = e^x (\cos 0 + i \sin 0) = e^x (1 + i \cdot 0) = e^x.$$

- 2) Es sei nun $x = 0$, d.h., $x + iy = iy$ liegt auf der imaginären Achse. (Man spricht in diesem Fall von einer *rein imaginären Zahl*.) Es gilt dann:

$$e^{iy} = \cos y + i \sin y. \quad (7.15)$$

Das heißt: Die imaginäre Achse wird auf den Einheitskreis abgebildet.

Insbesondere gilt: Der Abschnitt der Zahlen iy , $y \in [0, 2\pi)$ wird umkehrbar eindeutig auf den Einheitskreis abgebildet.



Insbesondere gilt also:

$$e^0 = e^{2\pi i} = 1, \quad e^{\frac{\pi}{2}i} = i, \quad e^{\pi i} = -1 \quad \text{und} \quad e^{\frac{3\pi}{2}i} = -i. \quad (7.16)$$

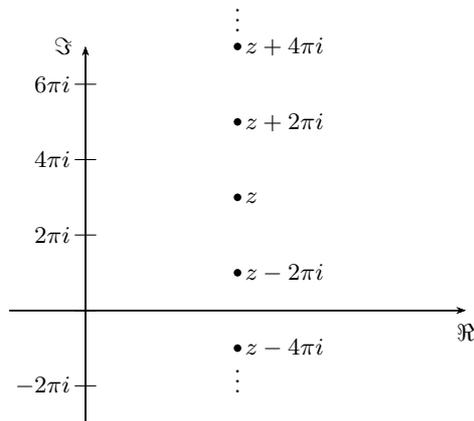
3) Es gilt für alle $z \in \mathbb{C}$:

$$e^{z+2\pi i} = e^z. \quad (7.17)$$

Beweis. Es sei $z = x + iy$. Dann ist $z + 2\pi i = x + i(y + 2\pi)$. Es folgt:

$$\begin{aligned} e^{z+2\pi i} &= e^{x+i(y+2\pi)} = e^x (\cos(y + 2\pi) + i \sin(y + 2\pi)) \\ &= e^x (\cos y + i \sin y) \\ &= e^{x+iy} \\ &= e^z \quad \square \end{aligned}$$

(7.17) besagt, dass die komplexe Exponentialfunktion *periodisch* ist: Komplexen Zahlen, deren Differenz $2\pi i$ beträgt, wird durch die Exponentialfunktion dasselbe Bild zugeordnet.



4) Für $z_1, z_2 \in \mathbb{C}$ gilt:

$$e^{z_1} \cdot e^{z_2} = e^{z_1+z_2}. \quad (7.18)$$

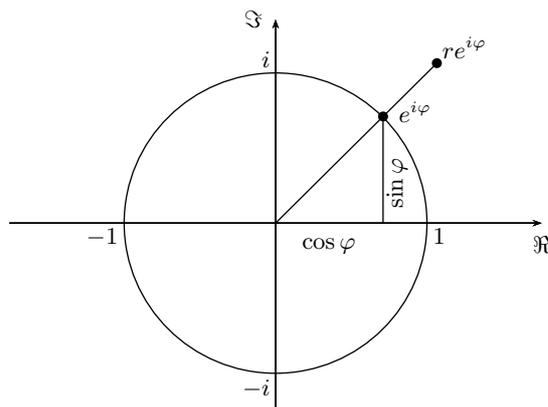
Beweis. Es sei $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$. Dann gilt:

$$\begin{aligned} e^{z_1} \cdot e^{z_2} &= e^{x_1} (\cos y_1 + i \sin y_1) \cdot e^{x_2} (\cos y_2 + i \sin y_2) \\ &= e^{x_1} \cdot e^{x_2} (\cos (y_1 + y_2) + i \sin (y_1 + y_2)) \\ &= e^{x_1+x_2} (\cos (y_1 + y_2) + i \sin (y_1 + y_2)) \\ &= e^{x_1+x_2+i(y_1+y_2)} \\ &= e^{z_1+z_2} \quad \square \end{aligned}$$

Aus (7.15) ergibt sich eine weitere oft benutzte Art, eine komplexe Zahl darzustellen. Ist r der Betrag und φ das Argument von z (mit $\varphi \in [0, 2\pi)$), so gilt $z = r(\cos \varphi + i \sin \varphi)$. Nach (7.15) können wir für $\cos \varphi + i \sin \varphi$ auch $e^{i\varphi}$ schreiben, so dass sich die folgende Darstellung von z ergibt:

$$z = r e^{i\varphi}. \quad (7.19)$$

In der folgenden Zeichnung ist (7.19) für den Fall $r = 1.5$ und $\varphi = \frac{\pi}{4}$ dargestellt.



8 Vektorräume über einem Körper K

8.1 Vorbemerkungen

In den Abschnitten 8 und 9 wird die **Lineare Algebra**, deren Behandlung unter Zugrundelegung des Lehrbuchs von Gramlich begonnen wurde, *weiter ausgebaut*. Dabei wird die Kenntnis der bislang behandelten Teile der Linearen Algebra vorausgesetzt - einiges wird aber auch wiederholt werden. Diese Wiederholung dient zum einen der Festigung der Kenntnisse, zum anderen aber auch der Vertiefung des Bisherigen, da wir einen *allgemeineren Standpunkt* einnehmen werden:

- Während im Gramlich ausschließlich die Räume \mathbb{R}^n (und deren Unterräume) behandelt wurden, werden wir nun *allgemeine Vektorräume* V studieren, d.h., wenn von einem Vektorraum V die Rede ist, so brauchen die Elemente von V nicht unbedingt n -Tupel zu sein, sondern können beispielsweise auch Matrizen, Polynome, unendliche Folgen oder Funktionen sein.
- Außerdem wird das meiste *über einem beliebigen Körper* K behandelt. Anders gesagt: Im Gramlich lag immer der Körper $K = \mathbb{R}$ zugrunde, jetzt kann K beispielsweise auch \mathbb{C} , \mathbb{Z}_2 oder \mathbb{Z}_3 sein.

Wir werden auch **neue Themen** behandeln, die im 1. Semester gar nicht zur Sprache kamen, insbesondere trifft dies zu auf die Abschnitte über:

- Lineare Abbildungen
- Eigenwerte¹

Wiederholungsfragen:

- a) Welche Beispiele für die algebraische Struktur eines Körpers kennen Sie?
- b) Richtig oder falsch: In einem Körper besitzt jedes Element ein Inverses der Multiplikation.
- c) Richtig oder falsch: Die Menge $M^*(n \times n, \mathbb{R})$ der invertierbaren $n \times n$ -Matrizen über \mathbb{R} bilden einen Körper bzgl. der üblichen Addition und Multiplikation von Matrizen.

8.2 Der Begriff des Vektorraums über einem Körper

Wir beginnen mit einigen **Beispielen**.

1. Für ein $n \in \mathbb{N}$ betrachten wir die Menge \mathbb{R}^n , d.h. die Menge aller n -Tupel reeller Zahlen. Auf der Menge \mathbb{R}^n können wir auf naheliegender Art eine Addition definieren:

Für zwei n -Tupel $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ definieren wir deren Summe durch

$$x + y = (x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n).$$

Die Summe zweier n -Tupel ist wieder ein n -Tupel; bei der Addition handelt es sich also um eine Abbildung vom Typ

$$+ : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Beispiel. $n = 4$, $x = (3, 4, 7, -1)$, $y = (4, 0, 2, 4)$. Dann gilt $x + y = (7, 4, 9, 3)$.

¹Den Abschnitten 8 und 9 liegt zum Teil das bekannte Lehrbuch von Jänich zugrunde (K. Jänich, Lineare Algebra, Springer).

Man sagt, die Addition zweier n -Tupel erfolgt *komponentenweise*, d.h., man addiert die erste Komponente von x zur ersten Komponente von y , die zweite Komponente von x zur zweiten Komponente von y usw.

Ähnlich (d.h. ebenfalls komponentenweise) definiert man, wie ein n -Tupel $x = (x_1, \dots, x_n)$ mit einer reellen Zahl λ multipliziert wird:

$$\lambda x = \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n).$$

D.h., jede Komponente wird mit λ multipliziert.

Das Ergebnis $(\lambda x_1, \dots, \lambda x_n)$ ist wieder ein n -Tupel, es liegt also eine Abbildung vom Typ

$$\cdot : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

vor.

Beispiel. $n = 4, x = (3, 4, 7, -1), \lambda = 3$. Dann gilt $\lambda x = (9, 12, 21, -3)$.

Für die soeben definierte Addition und Multiplikation gelten gewisse *Rechenregeln*, die alle direkt aus der Definition folgen (sowie aus der Tatsache, dass ähnliche Rechenregeln für \mathbb{R} gelten):

- (1) $(x + y) + z = x + (y + z)$ für alle $x, y, z \in \mathbb{R}^n$.
- (2) $x + y = y + x$ für alle $x, y \in \mathbb{R}^n$.
- (3) Schreiben wir (wie üblich) kurz 0 statt $(0, \dots, 0) \in \mathbb{R}^n$, so gilt $x + 0 = x$ für alle $x \in \mathbb{R}^n$.
- (4) Ist $x = (x_1, \dots, x_n)$ und schreiben wir $-x$ für $(-x_1, \dots, -x_n)$, so gilt $x + (-x) = 0$ für alle $x \in \mathbb{R}^n$.
- (5) $\lambda(\mu x) = (\lambda\mu)x$ für alle $\lambda, \mu \in \mathbb{R}, x \in \mathbb{R}^n$.
- (6) $1x = x$ für alle $x \in \mathbb{R}^n$.
- (7) $\lambda(x + y) = \lambda x + \lambda y$ für alle $\lambda \in \mathbb{R}, x, y \in \mathbb{R}^n$.
- (8) $(\lambda + \mu)x = \lambda x + \mu x$ für alle $\lambda, \mu \in \mathbb{R}, x \in \mathbb{R}^n$.

Die Regeln (1) - (4) beziehen sich ausschließlich auf die Addition; wir können diese ersten vier Regeln auch so zusammenfassen:

$$(\mathbb{R}^n, +) \text{ ist eine abelsche Gruppe}^2.$$

2. Mit $[-1, 1]$ bezeichnen wir das Intervall $I = \{x \in \mathbb{R} : -1 \leq x \leq 1\}$; wir betrachten die Menge M aller Funktionen $f : [-1, 1] \rightarrow \mathbb{R}$. Es sei also

$$M = \{f : f \text{ ist eine Funktion } [-1, 1] \rightarrow \mathbb{R}\}.$$

Beispiele:

- Die Funktion f , die durch $f(x) = x$ für alle $x \in [-1, 1]$ gegeben ist.
- Die Funktion g , die durch $g(x) = x^2$ für alle $x \in [-1, 1]$ gegeben ist.
- Die konstante Funktion h , für die $h(x) = 5$ für alle $x \in [-1, 1]$ gilt.

Man kann für 2 Funktionen $f, g \in M$ auf naheliegende Art eine Funktion $f + g$ definieren, die man die *Summe von f und g* nennt

$$(f + g)(x) := f(x) + g(x) \text{ für alle } x \in [-1, 1].$$

Die Funktion $f + g$ ordnet also jedem $x \in [-1, 1]$ die Summe von $f(x)$ und $g(x)$ zu. Die so definierte Funktion $f + g$ ist wieder in M ; bei der Addition handelt es sich also um eine Abbildung vom Typ

$$+ : M \times M \rightarrow M.$$

²Zur Erinnerung: Abelsche Gruppe bedeutet dasselbe wie kommutative Gruppe.

Ähnlich definiert man zu einer reellen Zahl λ und einer Funktion $f : [-1, 1] \rightarrow \mathbb{R}$ eine Funktion λf :

$$(\lambda f)(x) := \lambda f(x) \text{ für alle } x \in [-1, 1].$$

Die so definierte Funktion λf ist wieder in M ; die Abbildung, die jedem $\lambda \in \mathbb{R}$ und jedem $f \in M$ die neue Funktion λf zuordnet, ist also vom Typ

$$\cdot : \mathbb{R} \times M \rightarrow M.$$

Mit 0 wollen wir die konstante Funktion aus M bezeichnen, die jedes Element $x \in [-1, 1]$ auf die Zahl $0 \in \mathbb{R}$ abbildet³. Für jedes $f \in M$ sei mit $-f$ die Abbildung aus M bezeichnet, für die $(-f)(x) = -f(x)$ gilt. Mit diesen Bezeichnungen gelten genau die Rechenregeln, denen wir schon im ersten Beispiel begegnet sind: Für alle $f, g, h \in M, \lambda, \mu \in \mathbb{R}$ gilt:

$$(1) (f + g) + h = f + (g + h)$$

$$(2) f + g = g + f$$

$$(3) f + 0 = f$$

$$(4) f + (-f) = 0$$

$$(5) \lambda(\mu f) = (\lambda\mu)f$$

$$(6) 1f = f$$

$$(7) \lambda(g + f) = \lambda f + \lambda g$$

$$(8) (\lambda + \mu)f = \lambda f + \mu f$$

Die Regeln (1) - (4) lassen sich wieder zusammenfassen zu der Feststellung, dass $(M, +)$ eine abelsche Gruppe ist.

|| *Obwohl in den beiden obigen Beispielen die auftretenden Objekte recht unterschiedlich sind, wird doch mit diesen Objekten nach den gleichen formalen Regeln gerechnet.* ||

Wir wollen anstelle des ersten Beispiels nun ein etwas allgemeineres Beispiel betrachten. Es ist ohne Weiteres möglich, den Körper \mathbb{R} der reellen Zahlen durch einen beliebigen Körper K zu ersetzen.

3. Gegeben sei der Körper K . Für ein $n \in \mathbb{N}$ betrachten wir die Menge K^n aller n -Tupel $x = (x_1, \dots, x_n)$ mit Komponenten $x_i \in K$ ($i = 1, \dots, n$).

Für zwei n -Tupel $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ kann man dann wieder deren Summe definieren durch:

$$x + y = (x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n).$$

Beispiele zur soeben definierten Addition:

$$(i) K = \mathbb{Z}_2, x = (1, 0, 0, 1), y = (0, 1, 1, 1) \Rightarrow x + y = (1, 1, 1, 0).$$

$$(ii) K = \mathbb{C}, x = (1, 2 + i, 3, i), y = (0, 1 + 4i, 2i, 0) \Rightarrow x + y = (1, 3 + 5i, 3 + 2i, i).$$

Das Produkt eines Körperelements $\lambda \in K$ mit $x = (x_1, \dots, x_n)$ wird durch

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

definiert.

Beispiele hierzu:

$$(i) K = \mathbb{Z}_2, x = (1, 0, 0, 1) \Rightarrow 0x = (0, 0, 0, 0), 1x = (1, 0, 0, 1).$$

$$(ii) K = \mathbb{Z}_3, x = (2, 0, 1, 2), \lambda = 2 \Rightarrow \lambda x = 2x = (1, 0, 2, 1).$$

³Derartige Doppelbezeichnungen sind nicht unüblich in der Mathematik; wichtig ist, dass aus dem Zusammenhang immer eindeutig hervorgeht, was mit dem Zeichen 0 in einer bestimmten Situation gemeint ist.

(iii) $K = \mathbb{C}, x = (1, 2 + i, 3 - 4i, i), \lambda = i \Rightarrow \lambda x = ix = (i, -1 + 2i, 4 + 3i, -1)$.

Analog zum ersten Beispiel ist dann die Addition eine Abbildung vom Typ

$$+ : K^n \times K^n \rightarrow K^n$$

und die Multiplikation (die man übrigens *skalare Multiplikation* nennt) ist eine Abbildung vom Typ

$$\cdot : K \times K^n \rightarrow K^n.$$

Auch gelten dieselben Rechenregeln wie im ersten Beispiel: Man hat in den Regeln (1) bis (8) des Beispiels 1 nur das Symbol \mathbb{R} gegen K auszutauschen.

4. Ebenfalls kann man das zweite Beispiel verallgemeinern: Man kann ohne weiteres den Körper \mathbb{R} der reellen Zahlen durch einen beliebigen Körper K ersetzen und darüber hinaus an die Stelle der etwas speziellen Menge $I = [-1, 1]$ eine beliebige Menge A setzen („wobei noch nicht einmal $A \subseteq K$ gelten muss, sondern A völlig beliebig sein kann).

Man definiert dann ebenso wie zuvor für $M = \{f : f \text{ ist eine Funktion } A \rightarrow K\}$ eine Addition $+ : M \times M \rightarrow M$ und eine skalare Multiplikation $\cdot : K \times M \rightarrow M$, für die dann die Rechenregeln (1) - (8) gelten.

Diese und viele weitere Beispiele führen zum Begriff des Vektorraums über einem Körper K , der wie folgt definiert wird.

Definition.

Gegeben sei ein Körper K . Ein Tripel $(V, +, \cdot)$, bestehend aus einer Menge V , einer Abbildung (genannt *Addition*)

$$+ : V \times V \rightarrow V, \\ (x, y) \mapsto x + y$$

und einer Abbildung (genannt *skalare Multiplikation*)

$$\cdot : K \times V \rightarrow V, \\ (\lambda, x) \mapsto \lambda x$$

heißt ein *Vektorraum über K* , wenn für die Abbildungen $+$ und \cdot Folgendes gilt:

- (I) Bezüglich der Addition $+$ ist V eine abelsche Gruppe. Im Einzelnen bedeutet dies, dass die folgenden Axiome gelten:
- (1) $(x + y) + z = x + (y + z)$ für alle $x, y, z \in V$.
 - (2) $x + y = y + x$ für alle $x, y \in V$.
 - (3) Es gibt ein neutrales Element $0 \in V$ (genannt „Null“ oder „Nullvektor“) mit $x + 0 = x$ für alle $x \in V$.
 - (4) Zu jedem $x \in V$ gibt es ein Element $-x \in V$ mit $x + (-x) = 0$.
- (II) Darüber hinaus sollen vier weitere Axiome gelten:
- (5) $\lambda(\mu x) = (\lambda\mu)x$ für alle $\lambda, \mu \in K, x \in V$.
 - (6) $1x = x$ für alle $x \in V$.
 - (7) $\lambda(x + y) = \lambda x + \lambda y$ für alle $\lambda \in K, x, y \in V$.
 - (8) $(\lambda + \mu)x = \lambda x + \mu x$ für alle $\lambda, \mu \in K, x \in V$.

Die Elemente von V heißen *Vektoren* und die Elemente von K werden *Körper Elemente* oder *Skalare* genannt. Den Körper K nennt man auch den *zugrundeliegenden Körper*.

Ist der zugrundeliegende Körper \mathbb{R} oder \mathbb{C} , so spricht man von einem *reellen* bzw. *komplexen Vektorraum*.

Durch die obigen Beispiele ist bereits angedeutet worden, dass es Vektorräume ganz unterschiedlicher Art gibt. Hier einige weitere Beispiele, die zeigen, dass Vektorräume in den unterschiedlichsten Zusammenhängen auftreten.

5. Die Menge $K[x]$ der *Polynome* über einem Körper K bilden einen Vektorraum über K , wenn man die Addition zweier Polynome auf die übliche Art definiert und die skalare Multiplikation auf die naheliegende Art wie folgt erklärt: Für $a(x) \in K[x]$ mit $a(x) = a_0 + a_1x + \dots + a_nx^n$ und $\lambda \in K$ sei

$$\lambda a(x) := \lambda a_0 + \lambda a_1x + \dots + \lambda a_nx^n.$$

6. Die Menge $M(m \times n, K)$ der $m \times n$ -Matrizen über einem Körper K bilden einen Vektorraum über K , wenn Addition und skalare Multiplikation auf die übliche Art erklärt werden.
7. Es sei $U(K)$ die Menge der *unendlichen Folgen*

$$u_0, u_1, u_2, \dots,$$

wobei die Folgenglieder u_i aus K stammen ($i = 0, 1, \dots$). Besonders wichtig ist für uns dabei der Fall, dass $K = \mathbb{R}$ gilt, dass es sich also um *unendliche Folgen reeller Zahlen* handelt.

Als eine kompakte Schreibweise für Folgen u_0, u_1, u_2, \dots aus $U(K)$ wählen wir die Schreibweise $(u_i)_{i=0,1,\dots}$ oder (noch kürzer) (u_i) . Für $(u_i), (u'_i) \in U(K)$ und $\lambda \in K$ definiert man Addition und skalare Multiplikation *gliedweise* durch:

$$(u_i) + (u'_i) := (u_i + u'_i) \in U(K),$$

$$\lambda(u_i) := (\lambda u_i) \in U(K).$$

Dadurch wird $U(K)$ zu einem Vektorraum über K .

Wir haben bislang eine Reihe unterschiedlichster Vektorräume kennengelernt, u.a.

- K^n , den Vektorraum aller n -Tupel über K ,
- die Menge M aller *Funktionen* $f : [-1, 1] \rightarrow \mathbb{R}$,
- die Menge $K[x]$ aller *Polynome* über K ,
- die Menge $M(m \times n, K)$ aller $m \times n$ -*Matrizen* über K und
- die Menge $U(\mathbb{R})$ aller *reellen unendlichen Folgen*,

wobei die Addition und skalare Multiplikation jeweils wie zuvor beschrieben definiert sind. Oft interessiert man sich allerdings gar nicht so sehr für die Menge *aller* Vektoren eines Vektorraums V , sondern viel mehr für *interessante Teilmengen von V , die für sich genommen bereits einen Vektorraum bilden*; man spricht in diesem Zusammenhang von *Untervektorräumen*.

Beispiele für derartige Untervektorräume:

- die Menge aller *stetigen* Funktionen $f : [-1, 1] \rightarrow \mathbb{R}$
- die Menge aller *differenzierbaren* Funktionen $f : [-1, 1] \rightarrow \mathbb{R}$
- die Menge aller $(x_1, \dots, x_n) \in K^n$, die Lösung *eines gegebenen homogenen linearen Gleichungssystems* sind.
- die Menge der konvergenten Folgen aus $U(\mathbb{R})$.

8.3 Einige Folgerungen aus den Axiomen

In jedem Vektorraum V über einem Körper K gelten neben den Axiomen (1) - (8) noch weitere Rechenregeln, die nicht mit in die Liste der Axiome aufgenommen wurden, weil sie aus den Axiomen ableitbar sind. Beispielsweise finden sich nirgends die Regeln $0x = 0$ und $(-1)x = -x$. Diese Regeln sind dennoch gültig, da man sie (wie wir gleich besprechen werden) als Folgerungen aus den Axiomen erhalten kann.

Zunächst erinnern wir an einige einfache Tatsachen, die sich daraus ergeben, dass $(V, +)$ eine abelsche Gruppe ist: Wie in jeder Gruppe, ist in $(V, +)$ das neutrale Element 0 eindeutig bestimmt; ebenfalls eindeutig bestimmt ist für jedes $x \in V$ das Inverse bzgl. $+$: Es ist das Element $-x$, das *das Negative von x* genannt wird; aufgrund der Kommutativität von $+$ gilt nicht nur $x + 0 = x$ (wie in (3) gefordert), sondern auch $0 + x = x$; da in jeder Gruppe das Inverse des neutralen Elements immer das neutrale Element selber ist, gilt $-0 = 0$.

Abkürzung. Wie üblich schreiben wir $x - y$ anstelle von $x + (-y)$.

Nun zu einigen Folgerungen aus den Axiomen (1) - (8), die sich auch auf die skalare Multiplikation beziehen.

Feststellung 1.

In jedem Vektorraum V über einem Körper K gilt für alle $x \in V$ und alle $\lambda \in K$:

- (a) $\lambda 0 = 0$
- (b) $0x = 0$
- (c) $(-1)x = -x$
- (d) $\lambda x = 0 \Rightarrow \lambda = 0$ oder $x = 0$.

Beweis. Wir beweisen nur (a) und (d). Die Beweise von (b) und (c) gehen ähnlich; man findet sie im Übrigen auch im Buch von Jänich (Seite 31f).

(a) Zunächst halten wir fest, dass aufgrund von (3) und (7) gilt:

$$\lambda 0 = \lambda(0 + 0) = \lambda 0 + \lambda 0. \tag{*}$$

Addiert man zur linken und rechten Seite von (*) das Element $-(\lambda 0)$, so erhält man die Behauptung $0 = \lambda 0$. Unter genauer Angabe der benutzten Axiome sieht die Durchführung dieses Beweises so aus:

$$0 \stackrel{(4)}{=} \lambda 0 + \left(-(\lambda 0)\right) \stackrel{(*)}{=} (\lambda 0 + \lambda 0) + \left(-(\lambda 0)\right) \stackrel{(1)}{=} \lambda 0 + \left(\lambda 0 + \left(-(\lambda 0)\right)\right) \stackrel{(4)}{=} \lambda 0 + 0 \stackrel{(3)}{=} \lambda 0.$$

(b) Es gelte $\lambda x = 0$ und wir nehmen $\lambda \neq 0$ an. Zu zeigen ist dann $x = 0$.

Wegen $\lambda \neq 0$ existiert im Körper K das Inverse λ^{-1} bzgl. der Multiplikation in K . Es folgt

$$0 \stackrel{(a)}{=} \lambda^{-1} 0 \stackrel{\text{nach Voraussetzung}}{=} \lambda^{-1}(\lambda x) \stackrel{(5)}{=} (\lambda^{-1}\lambda)x \stackrel{\text{in } K \text{ gilt } \lambda^{-1}\lambda=1}{=} 1x \stackrel{(6)}{=} x. \quad \square$$

8.4 Untervektorräume

Definition.

Es sei V ein Vektorraum über K . Eine Teilmenge U von V nennt man einen *Untervektorraum* (oder *Unterraum*) von V , falls U in Bezug auf die Addition und skalare Multiplikation von V selber ein Vektorraum über K ist.

Feststellung 2.

Es sei V ein Vektorraum über K . Eine nichtleere Teilmenge U von V ist genau dann ein Untervektorraum von V , wenn für alle $x, y \in U$ und alle $\lambda \in K$ gilt:

$$x + y \in U \text{ und } \lambda x \in U.$$

Beweis.

- (I) Ist U in Bezug auf die Vektorraumoperationen⁴ von V selbst ein Vektorraum, so muss U insbesondere bezüglich $+$ und \cdot abgeschlossen sein, d.h., es muss $x+y \in U$ und $\lambda x \in U$ für alle $x, y \in U$ und alle $\lambda \in K$ gelten.
- (II) Es gelte nun umgekehrt für alle $x, y \in U$ und alle $\lambda \in K$: $x+y \in U$ und $\lambda x \in U$. Zu zeigen ist, dass U in Bezug auf $+$ und \cdot selbst ein Vektorraum über K ist, d.h., neben der Abgeschlossenheit bzgl. $+$ und \cdot (die wir ja voraussetzen) müssen noch sämtliche Axiome (1) - (8) gelten. Geht man die Axiome durch, so stellt man fest, dass sich die Axiome (1), (2), (5) - (8) selbstverständlich von V auf U übertragen. Es bleiben also noch (3) und (4) nachzuweisen. Dies geschieht durch Nachweis von (i) $0 \in U$ und (ii) $x \in U \Rightarrow -x \in U$.
- (i) Da $U \neq \emptyset$ gilt, können wir ein $x \in U$ wählen. Wegen $\lambda x \in U$ für alle $\lambda \in K$ folgt insbesondere $0x \in U$. Nach Feststellung 1(b) gilt $0x = 0$. Also folgt $0 \in U$.
- (ii) Es sei $x \in U$. Da dann $\lambda x \in U$ für alle $\lambda \in K$ gilt, folgt insbesondere $(-1)x \in U$. Nach Feststellung 1(c) gilt $(-1)x = -x$. Also folgt $-x \in U$. \square

Wir halten noch einmal ausdrücklich fest (siehe (i) im Beweis von Feststellung 2): Jeder Untervektorraum U muss den Nullvektor von V enthalten.

Besonders einfache Untervektorräume, die in jedem Vektorraum V enthalten sind, sind $\{0\}$ und V .

Ein **Beispiel**, das die Nützlichkeit von Feststellung 2 demonstriert:

Für $I \subseteq \mathbb{R}$ und einen inneren Punkt x_0 von I sei W die Menge der Funktionen

$$f : I \rightarrow \mathbb{R},$$

die an der Stelle x_0 differenzierbar sind. Wir möchten gerne nachweisen, dass W ein Vektorraum ist (bzgl. der naheliegenden Operationen). Schauen wir uns die Definition des Begriffs „Vektorraum“ in Abschnitt 8.2 an, so stellen wir fest, dass eine ganze Reihe von Eigenschaften nachzuweisen sind: Zunächst einmal ist die Abgeschlossenheit von W gegenüber $+$ nachzuweisen und außerdem die Gültigkeit der Axiome (1) - (4) zu prüfen; anschließend ist die gesamte Prozedur für die skalare Multiplikation in ähnlicher Weise zu wiederholen.

Einfacher geht es allerdings so: *Wir halten zunächst Ausschau nach einer geeigneten Obermenge V von W , von der wir bereits wissen, dass sie ein Vektorraum ist.* Ein solches V ist leicht gefunden: Wir wählen V einfach als die Menge *aller* Funktionen $f : I \rightarrow \mathbb{R}$ (vgl. Beispiel 4 in Abschnitt 8.2). Nun können wir Feststellung 2 anwenden: Es gilt $W \neq \emptyset$ und wir wissen aus der Analysis, dass die Summe zweier differenzierbarer Funktionen wieder differenzierbar ist und dass λf differenzierbar ist, falls f differenzierbar ist (alles an der Stelle x_0). *Es folgt, dass W ein Vektorraum ist*, da W nämlich (wie wir gezeigt haben) ein Untervektorraum von V ist.

Kurz zusammengefasst: Wenn von einer Menge W nachzuweisen ist, dass es sich um einen Vektorraum handelt, so geht das häufig am besten dadurch, dass man eine geeignete Obermenge von W wählt, von der man bereits weiß, dass sie ein Vektorraum ist und danach das *Unterraumkriterium* (Feststellung 2) anwendet.

8.5 Lineare Unabhängigkeit

Wenn wir im Folgenden Vektorräume V über einem Körper K studieren, so kann K , wenn nichts anderes gesagt ist, immer ein beliebiger Körper sein. Sie machen aber auch nichts verkehrt, wenn Sie sich vorstellen, dass K der Körper der reellen Zahlen ist, dass also $K = \mathbb{R}$ gilt; im Hinterkopf sollten Sie allerdings behalten, dass die Ergebnisse auch für jeden anderen Körper K gelten, beispielsweise für $K = \mathbb{C}$, aber auch (was für Informatiker besonders wichtig ist) für *endliche Körper* K , wie etwa $K = \mathbb{Z}_2$ oder $K = \mathbb{Z}_3$.

⁴Die Addition und skalare Multiplikation eines Vektorraums V über K werden auch die *Vektorraumoperationen* von V genannt.

Zur Erinnerung: Die Elemente eines Vektorraums werden *Vektoren*, die Elemente des zugrundeliegenden Körpers werden *Skalare* genannt.

Ist für einen Vektorraum V über K und ein $r \in \mathbb{N}$ ein r -Tupel (v_1, \dots, v_r) von Vektoren $v_i \in V$ gegeben und ist $(\lambda_1, \dots, \lambda_r)$ ein r -Tupel von Skalaren $\lambda_i \in K$, so kann man aus diesen Vektoren und Skalaren einen Vektor v wie folgt bilden:

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r = \sum_{i=1}^r \lambda_i v_i.$$

Einen auf diese Art gebildeten Vektor v nennt man eine *Linearkombination* der Vektoren v_1, \dots, v_r .

Beispiel. Es seien $K = \mathbb{Z}_7$ und $V = K^4$. Die Vektoren v_1, v_2 und v_3 seien gegeben durch $v_1 = (2, 1, 2, 3)$, $v_2 = (6, 0, 1, 0)$ und $v_3 = (5, 6, 1, 1)$. Eine Linearkombination der Vektoren ist dann beispielsweise:

$$\begin{aligned} v &= 3v_1 + 2v_2 + 6v_3 \\ &= (6, 3, 6, 2) + (5, 0, 2, 0) + (2, 1, 6, 6) \\ &= (6, 4, 0, 1). \end{aligned}$$

Eine andere Linearkombination dieser Vektoren ist $w = v_1 + v_2 + v_3 = (6, 0, 4, 4)$. Auch der Nullvektor ist eine Linearkombination dieser Vektoren: $0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 = (0, 0, 0, 0)$.

Gegeben seien ein Körper K , ein Vektorraum V über K und endlich viele Vektoren $v_1, \dots, v_r \in V$. Dann bezeichnen wir die Menge aller Linearkombinationen der Vektoren v_1, \dots, v_r mit

$$\text{Lin}(v_1, \dots, v_r).$$

Man nennt $\text{Lin}(v_1, \dots, v_r)$ die *lineare Hülle* des r -Tupels (v_1, \dots, v_r) von Vektoren. Es gilt also

$$\text{Lin}(v_1, \dots, v_r) = \left\{ v \in V : \text{Es gibt } \lambda_1, \dots, \lambda_r \in K \text{ mit } v = \lambda_1 v_1 + \dots + \lambda_r v_r \right\}.$$

Die vorausgegangenen Definitionen gelten alle für $r \geq 1$; wir wollen nun noch zusätzlich den Fall $r = 0$ einbeziehen. Wie Sie sich vielleicht erinnern (vgl. Abschnitt 2.6 des DM-Skripts), gibt es genau ein 0-Tupel, nämlich die leere Menge \emptyset . Für das 0-Tupel \emptyset definieren wir nun ebenfalls, was dessen lineare Hülle $\text{Lin}(\emptyset)$ ist, indem wir festsetzen:

$$\text{Lin}(\emptyset) = \{0\}.$$

Die lineare Hülle des 0-Tupels \emptyset soll also genau ein Element enthalten, nämlich den Nullvektor von V . Wenn wir im Folgenden von r -Tupeln von Vektoren v_1, \dots, v_r sprechen, so ist der Fall $r = 0$ mit einbezogen: Für $r = 0$ ist (v_1, \dots, v_r) dann nur eine andere Schreibweise für das 0-Tupel \emptyset und $\text{Lin}(v_1, \dots, v_r)$ ist für $r = 0$ eine andere Schreibweise für $\{0\}$.

Feststellung 3.

$\text{Lin}(v_1, \dots, v_r)$ ist ein Unterraum von V .

Beweis. Für $r = 0$ ist dies klar. (Wieso nämlich?) Sei nun $r \geq 1$ und $v, w \in \text{Lin}(v_1, \dots, v_r)$. Dann gilt $v = \lambda_1 v_1 + \dots + \lambda_r v_r$ und $w = \mu_1 v_1 + \dots + \mu_r v_r$ für gewisse $\lambda_i, \mu_i \in K$. Es folgt

$$\begin{aligned} v + w &= \lambda_1 v_1 + \dots + \lambda_r v_r + \mu_1 v_1 + \dots + \mu_r v_r \\ &= (\lambda_1 + \mu_1) v_1 + \dots + (\lambda_r + \mu_r) v_r, \end{aligned}$$

womit gezeigt ist, dass $v + w \in \text{Lin}(v_1, \dots, v_r)$ gilt.

Ähnlich: Für $v = \lambda_1 v_1 + \dots + \lambda_r v_r \in \text{Lin}(v_1, \dots, v_r)$ und $\lambda \in K$ gilt

$$\begin{aligned} \lambda v &= \lambda(\lambda_1 v_1 + \dots + \lambda_r v_r) \\ &= (\lambda \lambda_1) v_1 + \dots + (\lambda \lambda_r) v_r \end{aligned}$$

und deshalb auch $\lambda v \in \text{Lin}(v_1, \dots, v_r)$. \square

Beispiel. Es sei V der (reelle) Vektorraum der Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ (vgl. auch Abschnitt 8.2, Beispiel 4). Die Funktionen $\sin x$ und $\cos x$ sind Elemente von V ; Elemente von $U = \text{Lin}(\cos x, \sin x)$ sind dann beispielsweise die Funktionen $2 \cos x + \sin x$ und $\frac{2}{25} \cos x + \frac{11}{25} \sin x$, aber auch $\cos x$ und $\sin x$ selbst sowie alle Vielfachen $\lambda \cos x$ und $\mu \sin x$ der Funktionen $\cos x$ und $\sin x$. Es gilt

$$\text{Lin}(\cos x, \sin x) = \left\{ \lambda \cos x + \mu \sin x : \lambda, \mu \in \mathbb{R} \right\}.$$

Wir definieren nun, wann ein r -Tupel (v_1, \dots, v_r) von Vektoren linear unabhängig heißt. Für den Fall $r = 0$, wenn also das Nulltupel \emptyset vorliegt, definieren wir: *Das Nulltupel \emptyset ist linear unabhängig.* Für den Fall $r \geq 1$ gilt folgende Festlegung:

Definition.

Es sei $r \geq 1$ und V sei ein Vektorraum über K . Ein r -Tupel (v_1, \dots, v_r) von Vektoren in V heißt *linear unabhängig*, falls aus

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0$$

stets folgt

$$\lambda_1 = \dots = \lambda_r = 0.$$

Knapp zusammengefasst kann man den Kern dieser Definition auch so festhalten:

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0 \rightarrow \lambda_1 = \dots = \lambda_r = 0$$

Dass für $r \geq 1$ ein r -Tupel (v_1, \dots, v_r) linear unabhängig ist, bedeutet also: *Es gibt nur eine Möglichkeit, den Nullvektor 0 als Linearkombination von v_1, \dots, v_r darzustellen*, nämlich nur diejenige, bei der alle λ_i gleich Null sind:

$$0 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_r = 0.$$

Ist ein r -Tupel nicht linear unabhängig, so heißt es *linear abhängig*.

Dass ein r -Tupel (v_1, \dots, v_r) linear abhängig ist, bedeutet also, dass es außer der obigen *noch mindestens eine weitere Möglichkeit gibt, den Nullvektor 0 als Linearkombination von v_1, \dots, v_r darzustellen*, bei der dann mindestens ein λ_i von Null verschieden ist:

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0 \text{ mit } \lambda_i \neq 0 \text{ für mindestens ein } i \in \{1, \dots, r\}.$$

Sprechweise: Anstelle von „das r -Tupel (v_1, \dots, v_r) ist linear abhängig bzw. unabhängig“ sagen wir auch etwas vereinfacht: „Die Vektoren v_1, \dots, v_r sind linear abhängig bzw. unabhängig.“

Beispiele:

1. Es sei $V = K^3$ für $K = \mathbb{Z}_7$ und $v_1, v_2, v_3 \in V$ seien gegeben durch $v_1 = (1, 2, 3)$, $v_2 = (1, 1, 1)$ und $v_3 = (3, 4, 5)$. Die Vektoren v_1, v_2 und v_3 sind linear abhängig, da

$$\begin{aligned} v_1 + 2v_2 + 6v_3 &= (1, 2, 3) + 2(1, 1, 1) + 6(3, 4, 5) \\ &= (1, 2, 3) + (2, 2, 2) + (4, 3, 2) \\ &= (0, 0, 0). \end{aligned}$$

2. Im Gramlich werden in erster Linie die Vektorräume \mathbb{R}^n betrachtet. *Vieles von dem, was wir nach Gramlich (+ Ergänzungsskript) behandelt haben, überträgt sich jedoch unmittelbar auf die Vektorräume K^n für beliebige Körper K .* Ein Beispiel hierzu: Es sei $K = \mathbb{Z}_7$ und $V = K^3$. Es soll geklärt werden, ob die Vektoren $v_1 = (1, 2, 1)$, $v_2 = (6, 2, 3)$ und $v_3 = (4, 4, 0)$ linear abhängig oder unabhängig sind. Mit anderen Worten: Gibt es Skalare λ_1, λ_2 und λ_3 , die nicht alle gleich Null sind, so dass

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = (0, 0, 0) ? \tag{8.1}$$

Lösung: Zweckmäßigerweise schreiben wir die Vektoren v_1 , v_2 und v_3 in Spaltenform; (8.1) lautet dann

$$\lambda_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 6 \\ 2 \\ 3 \end{pmatrix} + \lambda_3 \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir haben also ein lineares Gleichungssystem mit Koeffizienten (und Unbestimmten) aus \mathbb{Z}_7 zu lösen:

$$\begin{aligned} \lambda_1 + 6\lambda_2 + 4\lambda_3 &= 0 \\ 2\lambda_1 + 2\lambda_2 + 4\lambda_3 &= 0 \\ \lambda_1 + 3\lambda_2 &= 0 \end{aligned}$$

Wir verwenden den *Gauß-Algorithmus*:

$$\begin{array}{ccc|c} 1 & 6 & 4 & 0 \\ 2 & 2 & 4 & 0 \\ 1 & 3 & 0 & 0 \\ \hline 1 & 6 & 4 & 0 \\ 0 & 4 & 3 & 0 \\ 0 & 4 & 3 & 0 \\ \hline 1 & 6 & 4 & 0 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Es ergibt sich als Lösung des Gleichungssystems ($t \in \mathbb{Z}_7$ ist ein frei wählbarer Parameter):

$$\begin{aligned} \lambda_3 &= t \\ \lambda_2 + 6\lambda_3 &= 0 \\ \rightarrow \lambda_2 &= -6\lambda_3 = -6t = t \\ \lambda_1 + 6\lambda_2 + 4\lambda_3 &= 0 \\ \rightarrow \lambda_1 &= -6\lambda_2 - 4\lambda_3 = \lambda_2 + 3\lambda_3 = 4t \end{aligned}$$

Die *allgemeine Lösung* des Gleichungssystems lautet also

$$(\lambda_1, \lambda_2, \lambda_3) = (4t, t, t) = t(4, 1, 1), \quad t \in \mathbb{Z}_7.$$

Insbesondere erhält man bei Wahl von $t = 1$ eine nichttriviale Darstellung des Nullvektors

$$4v_1 + v_2 + v_3 = (0, 0, 0).$$

Die Vektoren v_1 , v_2 und v_3 sind also linear abhängig.

Einige Bemerkungen zum obigen Beispiel:

Als Menge geschrieben lautet die allgemeine Lösung („Lösungsmenge“)

$$\mathcal{L} = \left\{ t(4, 1, 1) : t \in \mathbb{Z}_7 \right\},$$

wobei \mathcal{L} eine Teilmenge von K^3 für $K = \mathbb{Z}_7$ ist. In Analogie zum \mathbb{R}^3 nennt man \mathcal{L} eine *Ursprungsgerade* des K^3 . Man übernimmt also nicht nur die Methoden des Falls $K = \mathbb{R}$, sondern auch die *Sprechweisen*, wodurch (besonders wichtig!) anschauliche Interpretationen ins Spiel gebracht werden.

Fragen:

- (i) Wie viele Elemente enthält K^3 ? Wie viele dieser Elemente liegen auf der Geraden \mathcal{L} ? Tragen Sie die Antworten ein:

$$|K^3| =$$

$$|\mathcal{L}| =$$

Geben Sie sämtliche Elemente von \mathcal{L} („Punkte der Geraden“) an!

$$\mathcal{L} =$$

- (ii) Was versteht man wohl unter der Ursprungsebene des K^3 , die von den Vektoren $u = (2, 1, 0)$ und $v = (1, 0, 1)$ aufgespannt wird? Weshalb ist es angemessen, von einer „Ursprungsebene“ zu sprechen?

Weitere **Beispiele**:

3. $K = \mathbb{Z}_3$, $v_1 = (2, 2, 0)$, $v_2 = (1, 1, 1)$, $v_3 = (0, 1, 1)$. Sind diese Vektoren des K^3 linear abhängig? Wir verwenden wieder den Gauß-Algorithmus:

$$\begin{array}{ccc|c} 2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ \hline 1 & 2 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ \hline 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ \hline 1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{array}$$



Um eine führende 1 zu erhalten wird die erste Zeile mit dem Inversen von 2 multipliziert.

Es folgt, dass $\lambda_1 = \lambda_2 = \lambda_3 = 0$ die einzige Lösung ist, d.h., die Vektoren v_1 , v_2 und v_3 sind linear unabhängig.

4. Es sei V der (reelle) Vektorraum der Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$. Wir betrachten zwei Vektoren $v_1, v_2 \in V$, nämlich $v_1 = \cos x$ und $v_2 = \sin x$.

Behauptung: v_1 und v_2 sind linear unabhängig.

Bestätigung dieser Behauptung: Es gelte⁵

$$\lambda_1 v_1 + \lambda_2 v_2 = 0,$$

d.h., $\lambda_1 \cos x + \lambda_2 \sin x = 0$ für alle $x \in \mathbb{R}$.

Da dies für alle $x \in \mathbb{R}$ gilt, muss es insbesondere für $x = 0$ und $x = \frac{\pi}{2}$ gelten. Es folgt

$$\begin{aligned} \lambda_1 \cos 0 + \lambda_2 \sin 0 &= 0 \\ \lambda_1 \cos \frac{\pi}{2} + \lambda_2 \sin \frac{\pi}{2} &= 0, \end{aligned}$$

also $\lambda_1 = \lambda_2 = 0$ (Man beachte, dass $\cos 0 = 1$, $\sin 0 = 0$, $\cos \frac{\pi}{2} = 0$ und $\sin \frac{\pi}{2} = 1$ gilt!), d.h., die Vektoren v_1 und v_2 sind linear unabhängig.

⁵Beachten Sie, dass das Symbol „0“ im Folgenden mit wechselnder Bedeutung verwendet wird („Doppelbedeutung“).

Dass Vektoren linear abhängig sind, kann auch auf eine etwas andere Art formuliert werden. Dies ist der Inhalt der nachfolgenden Feststellung.

Feststellung 4.

Die beiden folgenden Aussagen sind äquivalent:

- (i) Die Vektoren v_1, \dots, v_r sind linear abhängig.
- (ii) Es gibt unter den Vektoren v_1, \dots, v_r mindestens einen, der Linearkombination der übrigen ist.

Beweis. Wir haben zwei Dinge zu zeigen, nämlich (i) \Rightarrow (ii) und (ii) \Rightarrow (i).

(i) \Rightarrow (ii): Wir setzen also voraus, dass die Vektoren v_1, \dots, v_r linear abhängig sind; demnach existieren $\lambda_1, \dots, \lambda_r \in K$, die nicht alle gleich 0 sind und für die gilt:

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0 \tag{8.2}$$

Es sei $\lambda_i \neq 0$. Wir können die Gleichung (8.2) also umstellen und erhalten

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \dots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} - \dots - \frac{\lambda_r}{\lambda_i} v_r,$$

was bedeutet, dass v_i eine Linearkombination der übrigen Vektoren ist.

(ii) \Rightarrow (i): Wir setzen jetzt umgekehrt voraus, dass es unter den Vektoren v_1, \dots, v_r einen gibt, etwa v_i , der Linearkombination der übrigen ist; es gelte also für gewisse $\mu_1, \dots, \mu_{i-1}, \mu_{i+1}, \dots, \mu_r \in K$:

$$v_i = \mu_1 v_1 + \dots + \mu_{i-1} v_{i-1} + \mu_{i+1} v_{i+1} + \dots + \mu_r v_r.$$

Wählen wir nun $\lambda_i = 1$ und $\lambda_j = -\mu_j$ für alle j mit $1 \leq j \leq r$ und $j \neq i$, so ergibt sich durch Umstellen der obigen Gleichung

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0,$$

wobei $\lambda_i = 1 \neq 0$ gilt. Also sind die Vektoren v_1, \dots, v_r linear abhängig. \square

In Hinblick auf die lineare Unabhängigkeit können wir Feststellung 4 auch so formulieren:

(v_1, \dots, v_r) ist genau dann linear unabhängig, wenn keiner dieser Vektoren eine Linearkombination der übrigen Vektoren ist.

Definition.

Es sei V ein Vektorraum über K .

(a) Ein r -Tupel (v_1, \dots, v_r) von Vektoren in V heißt *Erzeugendensystem* von V , falls $\text{Lin}(v_1, \dots, v_r) = V$ gilt. (Ist (v_1, \dots, v_r) ein Erzeugendensystem von V , so lässt sich also jeder Vektor $v \in V$ als Linearkombination der Vektoren v_1, \dots, v_r schreiben, man sagt hierzu auch: *Die Vektoren v_1, \dots, v_r erzeugen V oder spannen V auf.*)

(b) Ein Erzeugendensystem (v_1, \dots, v_r) von V heißt eine *Basis* von V , wenn es linear unabhängig ist.

Damit (v_1, \dots, v_r) eine Basis von V ist, muss also zweierlei erfüllt sein:

- (i) $\text{Lin}(v_1, \dots, v_r) = V$.
- (ii) Die Vektoren v_1, \dots, v_r sind linear unabhängig.

Liegt ein Erzeugendensystem (v_1, \dots, v_r) von V vor, so kann man jedes $v \in V$ als Linearkombination

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r$$

schreiben. Ist (v_1, \dots, v_r) sogar eine Basis von V (es gilt also zusätzlich (ii)), dann sind in der Darstellung $v = \lambda_1 v_1 + \dots + \lambda_r v_r$ die Koeffizienten $\lambda_1, \dots, \lambda_r$ *eindeutig bestimmt*. Dies ist der Inhalt der folgenden Feststellung.

Feststellung 5.

Ist (v_1, \dots, v_n) eine Basis von V , so gibt es zu jedem $v \in V$ genau ein n -Tupel $(\lambda_1, \dots, \lambda_n) \in K^n$, für welches $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ gilt.

Beweis. Dass es ein solches n -Tupel $(\lambda_1, \dots, \lambda_n)$ gibt, folgt aus $\text{Lin}(v_1, \dots, v_n) = V$. Wir haben zu zeigen, dass es kein weiteres derartiges n -Tupel gibt. Wir erschließen dies indirekt (Beweis durch Widerspruch): Angenommen, es gäbe noch ein anderes derartiges n -Tupel $(\mu_1, \dots, \mu_n) \in K^n$. Dann würde also

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n$$

gelten und es wäre $\lambda_i \neq \mu_i$ für mindestens ein $i \in \{1, \dots, n\}$.

Es folgt

$$(\lambda_1 - \mu_1)v_1 + \dots + \underbrace{(\lambda_i - \mu_i)}_{\neq 0} v_i + \dots + (\lambda_n - \mu_n)v_n = 0.$$

Damit hätten wir aber eine Darstellung des Nullvektors als Linearkombination von (v_1, \dots, v_n) , bei der wegen $\lambda_i \neq \mu_i$ mindestens ein Koeffizient verschieden von 0 wäre. Dies ist ein Widerspruch zur Annahme, dass (v_1, \dots, v_n) linear unabhängig ist. \square

Sprechweise: Anstelle von „Basis“ findet man gelegentlich auch die Bezeichnung *Fundamentalsystem*, insbesondere, wenn Funktionenräume betrachtet werden.

Beispiel. Es sei $K = \mathbb{Z}_7$ und $V = K^3$. Weiter oben hatten wir die Ursprungsebene betrachtet, die von $u = (2, 1, 0)$ und $v = (1, 0, 1)$ aufgespannt wird, also die Menge

$$\text{Lin}(u, v) = \left\{ s(2, 1, 0) + t(1, 0, 1) : s, t \in \mathbb{Z}_7 \right\}.$$

Wie viele Elemente enthält $\text{Lin}(u, v)$? Begründen Sie Ihre Meinung!

Es sei nun $V = K^n$. Dann bilden die Vektoren

$$\begin{aligned} e_1 &:= (1, 0, 0, \dots, 0) \\ e_2 &:= (0, 1, 0, \dots, 0) \\ &\vdots \\ e_n &:= (0, 0, \dots, 0, 1) \end{aligned}$$

eine Basis von V . Man nennt (e_1, \dots, e_n) die *kanonische Basis* von K^n .

Der Nachweis, dass es sich bei (e_1, \dots, e_n) tatsächlich um eine Basis von K^n handelt, ist einfach:

- (i) Es gilt $\text{Lin}(e_1, \dots, e_n) = K^n$, da man jedes $a = (a_1, \dots, a_n) \in K^n$ als Linearkombination der Vektoren e_1, \dots, e_n darstellen kann, nämlich so:

$$a = (a_1, \dots, a_n) = a_1 e_1 + \dots + a_n e_n.$$

- (ii) Außerdem sind die Vektoren e_1, \dots, e_n linear unabhängig: Aus $\lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_n e_n = (0, 0, \dots, 0)$ folgt nämlich $(\lambda_1, \lambda_2, \dots, \lambda_n) = (0, 0, \dots, 0)$ und somit $\lambda_i = 0$ für alle $i = 1, \dots, n$.

Statt kanonische Basis von K^n sagt man auch *Standardbasis* von K^n und die Vektoren e_1, \dots, e_n werden auch *Einheitsvektoren* genannt.

Zur Übung beantworten Sie bitte folgende **Fragen**: V sei ein Vektorraum über K und es gelte $v \in V$, $v \neq 0$. Ist dann das *1-Tupel*, das nur aus v besteht, linear abhängig oder linear unabhängig? Und wie lautet die Antwort für den Fall $v = 0$?

Noch eine *nützliche Sprechweise*: Gegeben sei eine Darstellung des Nullvektors als Linearkombination der Vektoren v_1, \dots, v_r :

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0.$$

Sind in dieser Darstellung alle λ_i gleich 0, so spricht man von einer *trivialen Linearkombination* oder von einer *trivialen Darstellung des Nullvektors* als Linearkombination der Vektoren v_1, \dots, v_r .

Ist hingegen mindestens eines der λ_i von Null verschieden, so spricht man entsprechend von einer *nicht-trivialen Linearkombination* oder einer *nichttrivialen Darstellung des Nullvektors*.

Unter Benutzung dieser Sprechweise können wir die Definition der linearen Unabhängigkeit für $r \geq 1$ auch so fassen: Für $r \geq 1$ ist ein r -Tupel (v_1, \dots, v_r) genau dann linear unabhängig, wenn es *nur die triviale Darstellung des Nullvektors* als Linearkombination der Vektoren v_1, \dots, v_r gibt.

Zur Übung: Es sei (v_1, \dots, v_r) ein linear unabhängiges r -Tupel von Vektoren aus V . Man zeige, dass dann niemals zwei gleiche Vektoren in (v_1, \dots, v_r) vorkommen. (Hinweis: Man führe einen Widerspruchsbeweis durch, der so beginnt: Angenommen, es gibt $v_i = v_j$ mit $i \neq j \dots$)

8.6 Die Dimension eines Vektorraums

In diesem und in den folgenden Abschnitten sei mit V immer ein Vektorraum über einem Körper K bezeichnet.

Ein grundlegendes Hilfsmittel beim Beweis von Sätzen über Vektorräume ist der folgende Satz.

Satz 1 (Basisergänzungssatz).

Es seien $v_1, \dots, v_r, w_1, \dots, w_s$ Vektoren in V , für die gilt:

- (i) (v_1, \dots, v_r) ist linear unabhängig.
- (ii) $\text{Lin}(v_1, \dots, v_r, w_1, \dots, w_s) = V$.

Dann gilt: Falls (v_1, \dots, v_r) nicht bereits eine Basis von V ist, so kann man (v_1, \dots, v_r) durch Hinzunahme geeigneter Vektoren aus (w_1, \dots, w_s) zu einer Basis von V ergänzen.

Ein Beweis dieses Satzes soll hier nicht gegeben werden, da es für unsere Zwecke ausreicht, die Aussage des Basisergänzungssatzes zu kennen. (Beweise findet man in Lehrbüchern der Linearen Algebra, z.B. im Jänich.)

Es sei darauf hingewiesen, dass im Basisergänzungssatz der Fall $r = 0$ keineswegs ausgeschlossen ist: In diesem Fall ist (v_1, \dots, v_r) das „Nulltupel“ \emptyset , das ja definitionsgemäß linear unabhängig ist.

Im wichtigen *Spezialfall* $r = 0$ besagt der Basisergänzungssatz Folgendes:

Es seien w_1, \dots, w_s Vektoren in V , für die $\text{Lin}(w_1, \dots, w_s) = V$ gilt, die also ein Erzeugendensystem von V bilden. Dann kann man durch Auswahl geeigneter Vektoren aus w_1, \dots, w_s eine Basis von V erhalten.

Etwas anders formuliert lässt sich dies auch so ausdrücken:

|| Liegt ein Erzeugendensystem (w_1, \dots, w_s) von V vor, dann gibt es eine Basis von V , deren ||
|| Vektoren alle aus (w_1, \dots, w_s) stammen. ||

Der folgende grundlegende Satz besagt, dass zwei Basen eines Vektorraums V immer gleich viele Elemente besitzen (oder, wie man auch sagt, gleiche Länge haben).

Satz 2.

Sind (v_1, \dots, v_n) und (w_1, \dots, w_m) Basen von V , so ist $n = m$.

Auf den Beweis dieses Satzes wollen wir verzichten; wir erwähnen nur, dass im Beweis in entscheidender Weise vom Basisergänzungssatz Gebrauch gemacht wird. (Falls Sie an Details interessiert sind: Im Jänich und auch in vielen anderen Lehrbüchern der Linearen Algebra können Sie alles finden.)

Satz 2 ermöglicht die folgende Definition. (Beachten Sie: *Ohne* Satz 2 könnte es sein, dass es in V Basen unterschiedlicher Länge gibt, wodurch die nachfolgende Definition unsinnig werden würde.)

Definition.

Es sei V ein Vektorraum über K und (v_1, \dots, v_n) sei eine Basis von V . Dann nennt man die Zahl n die *Dimension* von V , abgekürzt $\dim V$. Man schreibt in diesem Fall $\dim V = n$ und nennt V *endlichdimensional*.

Will man zum Ausdruck bringen, dass V endlichdimensional ist, so schreibt man auch

$$\dim V < \infty \text{ oder } \dim V = n < \infty.$$

Für $V = K^n$ gilt $\dim V = n$, da es in K^n beispielsweise die kanonische Basis gibt. Zur Erinnerung: Die kanonische Basis von K^n besteht aus den n Einheitsvektoren e_i ($i = 1, \dots, n$); es gilt:

$$e_i = (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \dots, 0)$$

Beispiel. Für $V = K^3$ ist auch $\tilde{e}_1 = (0, 0, 1)$, $\tilde{e}_2 = (0, 1, 1)$, $\tilde{e}_3 = (1, 1, 1)$ eine Basis. (Weisen Sie dies zur Übung nach und geben Sie weitere Beispiele für Basen von $V = K^3$ an!)

Stellen wir uns vor, dass von irgendwelchen Vektoren $v_1, \dots, v_r \in V = K^n$ nachzuweisen ist, dass sie linear abhängig sind; dann sind also Zahlen $\lambda_1, \dots, \lambda_r$ zu finden, die nicht alle gleich Null sind und für die

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_r v_r = 0$$

gilt. In vielen Fällen wird es sich nicht vermeiden lassen, eine Reihe von Rechnungen zur Bestimmung solcher λ_i anzustellen; in bestimmten Fällen ist Rechnen allerdings nicht nötig. Dies ergibt sich aus dem nächsten Satz.

Satz 3.

Für V gelte $\dim V = n$ und (v_1, \dots, v_r) sei ein r -Tupel von Vektoren in V mit $r > n$. Dann sind die Vektoren v_1, \dots, v_r linear abhängig.

Beweis. Wegen $\dim V = n$ gibt es eine Basis (w_1, \dots, w_n) von V . Es gilt dann $\text{Lin}(w_1, \dots, w_n) = V$ und somit auch $\text{Lin}(v_1, \dots, v_r, w_1, \dots, w_n) = V$. Angenommen, (v_1, \dots, v_r) wäre linear unabhängig. Dann könnten wir auf die Vektoren $v_1, \dots, v_r, w_1, \dots, w_n$ den Basisergänzungssatz anwenden und erhielten: Entweder ist (v_1, \dots, v_r) bereits eine Basis von V oder aber (v_1, \dots, v_r) lässt sich zu einer Basis ergänzen. In jedem Fall würde es eine Basis von V der Länge $\geq r$ geben, im Widerspruch zu $r > n = \dim V$. \square

Satz 4.

Für V gelte $\dim V = n$ und (v_1, \dots, v_r) sei ein r -Tupel von Vektoren in V mit $r < n$. Dann bilden die Vektoren v_1, \dots, v_r kein Erzeugendensystem von V .

Beweis. Wäre (v_1, \dots, v_r) ein Erzeugendensystem von V , würde also $\text{Lin}(v_1, \dots, v_r) = V$ gelten, so würde es (vgl. Bemerkungen im Anschluss an den Basisergänzungssatz) eine Basis von V der Länge $\leq r$ geben, im Widerspruch zu $r < n = \dim V$. \square

In Satz 3 galt $r > n$ und in Satz 4 wurde $r < n$ vorausgesetzt. Bleibt also noch der Fall $r = n$ zu betrachten. Gilt $\dim V = n$, so weiß man, dass es eine Basis (w_1, \dots, w_n) von V der Länge n gibt. Nun sei (v_1, \dots, v_n) irgendein anderes n -Tupel von linear unabhängigen Vektoren von V . Es stellt sich die folgende Frage: Kann man sicher sein, dass die Vektoren v_1, \dots, v_n ebenfalls eine Basis von V bilden? Die Antwort ist im folgenden Satz enthalten:

Satz 5.

Für V gelte $\dim V = n$ und (v_1, \dots, v_n) sei ein beliebiges n -Tupel von Vektoren aus V . Dann sind die folgenden beiden Aussagen äquivalent:

- (i) v_1, \dots, v_n sind linear unabhängig;
- (ii) v_1, \dots, v_n bilden ein Erzeugendensystem von V .

Beweis. Vorausgesetzt wird $\dim V = n$; es existiert also eine Basis (w_1, \dots, w_n) von V . Zu zeigen ist (i) \Leftrightarrow (ii). Wir erledigen dies wie üblich, indem wir die beiden Richtungen (i) \Rightarrow (ii) und (ii) \Rightarrow (i) separat behandeln.

- (I) (i) \Rightarrow (ii): Die Vektoren v_1, \dots, v_n seien linear unabhängig. Da (w_1, \dots, w_n) ein Erzeugendensystem ist, gilt $\text{Lin}(v_1, \dots, v_n, w_1, \dots, w_n) = V$. Anwendung des Basisergänzungssatzes ergibt, dass entweder (v_1, \dots, v_n) bereits eine Basis ist (In diesem Fall wären wir fertig!) oder dass man (v_1, \dots, v_n) zu einer Basis ergänzen kann. Letzteres ist jedoch unmöglich, da man eine Basis der Länge $\geq n + 1$ erhielte. Also ist (v_1, \dots, v_n) bereits eine Basis, womit (ii) gezeigt ist.
- (II) (ii) \Rightarrow (i): Es sei (v_1, \dots, v_n) ein Erzeugendensystem von V . Da (v_1, \dots, v_n) eine Basis von V enthält (vgl. Bemerkungen im Anschluss an den Basisergänzungssatz) und da (wegen $\dim V = n$) jede Basis von V die Länge n hat, muss (v_1, \dots, v_n) selbst eine Basis von V sein. Es folgt, dass die Vektoren v_1, \dots, v_n linear unabhängig sind, was zu zeigen war. \square

Zusammenfassung der Sätze 3-5.

Voraussetzung: $\dim V = n$ und (v_1, \dots, v_r) sei ein r -Tupel von Vektoren von V .

1. Falls $r > n$ gilt, so sind die Vektoren v_1, \dots, v_r linear abhängig.
2. Falls $r < n$ gilt, so ist (v_1, \dots, v_r) kein Erzeugendensystem von V .

In den beiden Fällen 1. und 2. ist (v_1, \dots, v_r) demnach keine Basis von V .

3. Falls $r = n$ gilt, so ist (v_1, \dots, v_r) möglicherweise eine Basis von V . Um nachzuweisen, dass tatsächlich eine Basis vorliegt, genügt es, Eigenschaft (i) oder Eigenschaft (ii) aus Satz 5 nachzuweisen.

Einen Fall haben wir bis jetzt noch nicht besprochen: In manchen Vektorräumen V kann es *beliebig lange* linear unabhängige r -Tupel (v_1, \dots, v_r) geben; anders gesagt: Es gibt in V für jedes $r \in \mathbb{N}$ ein linear unabhängiges r -Tupel (v_1, \dots, v_r) . Dann kann nach Satz 3 nicht $\dim V = n < \infty$ gelten. Derartige Vektorräume V wollen wir *unendlichdimensional* nennen und wir wollen in einem solchen Fall

$$\dim V = \infty$$

schreiben.

Beispiel. Es sei V die Menge aller Funktionen $f : [-1, 1] \rightarrow \mathbb{R}$. Dann ist (wie wir in Abschnitt 8.2 besprochen haben) V ein Vektorraum über \mathbb{R} . Für jedes $i \in \mathbb{N}$ sei $f_i : (\frac{1}{i}) = 1$ und $f_i(x) = 0$ für alle $x \in [-1, 1]$ mit $x \neq \frac{1}{i}$. Dann gilt für jedes $r \in \mathbb{N}$: Die Funktionen f_1, \dots, f_r sind linear unabhängig. (Nachweis zur Übung empfohlen!) Es folgt $\dim V = \infty$.

In der Mathematik hilft man sich im Fall unendlichdimensionaler Vektorräume damit, dass man den Basisbegriff erweitert und auch definiert, was eine unendliche Basis sein soll. Das führt dazu, dass auch jeder unendlichdimensionale Vektorraum eine Basis besitzt. Darauf wollen wir aber nicht weiter eingehen; wir werden im Folgenden in erster Linie endlichdimensionale Vektorräume betrachten.

Die folgende Feststellung enthält Aussagen über die Dimension von Unterräumen eines endlichdimensionalen Vektorraums.

Feststellung 6.

V sei endlichdimensional und U sei ein Untervektorraum von V . Dann ist U ebenfalls endlichdimensional und es gilt:

$$\dim U \leq \dim V.$$

Zusatz: $\dim U = \dim V$ gilt nur im Fall $U = V$.

Beweis. Es sei $\dim V = n$. Dann kann U nicht mehr als n linear unabhängige Vektoren enthalten. Wir wählen ein r -Tupel (u_1, \dots, u_r) von linear unabhängigen Vektoren aus U , für das r maximal ist. Dann gilt $r \leq n$.

Behauptung: (u_1, \dots, u_r) ist eine Basis von U .

Nachweis dieser Behauptung: Da (u_1, \dots, u_r) linear unabhängig ist, bleibt zu zeigen, dass (u_1, \dots, u_r) ein Erzeugendensystem von U ist. Zu diesem Zweck sei $u \in U$. Wegen der Maximalwahl von r ist

(u_1, \dots, u_r, u) linear abhängig, d.h., es gibt Skalare $\lambda_1, \dots, \lambda_r$ und λ , die nicht alle gleich Null sind und für die $\lambda_1 u_1 + \dots + \lambda_r u_r + \lambda u = 0$ gilt. Da (u_1, \dots, u_r) linear unabhängig ist, folgt $\lambda \neq 0$ und somit $u = (-\frac{\lambda_1}{\lambda}) u_1 + \dots + (-\frac{\lambda_r}{\lambda}) u_r$. Damit ist gezeigt, dass (u_1, \dots, u_r) eine Basis von U ist. Es folgt $\dim U = r \leq n = \dim V$.

Es bleibt der Zusatz zu beweisen. Hierzu setzen wir $\dim U = \dim V$ voraus und zeigen, dass daraus $U = V$ folgt. Es gelte also $\dim U = \dim V$. Aus $\dim V = n$ folgt dann $\dim U = n$, d.h., es gibt eine Basis (u_1, \dots, u_n) von U der Länge n . Wegen $U \subseteq V$ sind die Vektoren u_1, \dots, u_n also linear unabhängige Vektoren von V . Wegen $\dim V = n$ folgt nach Satz 5, dass die Vektoren u_1, \dots, u_n ein Erzeugendensystem von V bilden. Es gilt also $\text{Lin}(u_1, \dots, u_n) = V$. Da (u_1, \dots, u_n) eine Basis von U ist, gilt andererseits auch $\text{Lin}(u_1, \dots, u_n) = U$, also $U = \text{Lin}(u_1, \dots, u_n) = V$. \square

Für einen Körper K sei $V = K^n$. Es seien s Vektoren v_1, \dots, v_s gegeben. Wir betrachten

$$U = \text{Lin}(v_1, \dots, v_s)$$

und möchten gerne die Dimension und eine Basis von U bestimmen. Wie kann man das machen?

Antwort: Man schreibt die Vektoren v_1, \dots, v_s zeilenweise auf und wendet das Gauß-Verfahren an: Ist Z die entstandene Zeilenstufenmatrix, so kann man an Z eine Basis von U (und damit auch $\dim U$) ablesen. Wie nämlich? Lesen Sie es im Gramlich in Abschnitt 4.2 nach! Dort wird das Verfahren zwar nur für den Spezialfall $K = \mathbb{R}$ beschrieben, es überträgt sich aber alles auf den Fall eines beliebigen Körpers K .

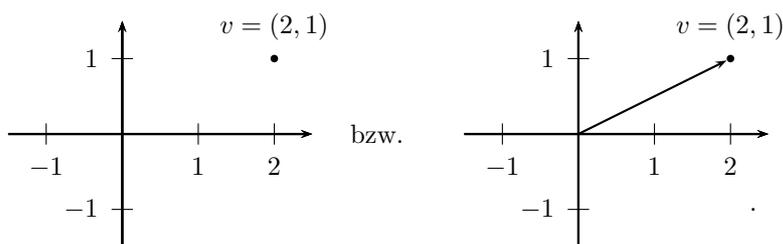
Der folgende Abschnitt 8.7 dient ausschließlich der Wiederholung und kann übersprungen werden.

8.7 Vektoren im \mathbb{R}^2 und \mathbb{R}^3

Unter den Beispielen von Vektorräumen, die wir bisher betrachtet haben (vgl. Abschnitt 8.2), sind die Vektorräume \mathbb{R}^n für $n = 1, 2, 3$ besonders anschaulich: Dem Fall $n = 1$ entspricht anschaulich die *Zahlengerade*, \mathbb{R}^2 stellt ein mathematisches Modell für eine Ebene (die *Zeichenebene*) dar, während \mathbb{R}^3 als Modell für den uns umgebenden Raum (dem *Anschauungsraum*) dient.

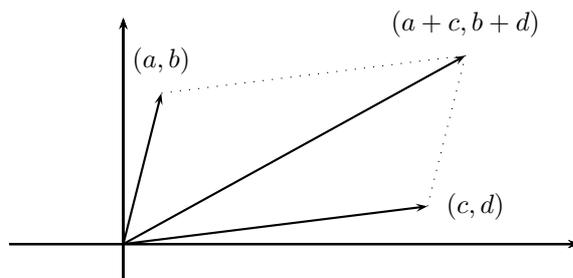
Im Folgenden wird auf die anschauliche Bedeutung eingegangen, die den bisher behandelten Grundbegriffen der Linearen Algebra in den Fällen $V = \mathbb{R}^2$ bzw. $V = \mathbb{R}^3$ zukommt.

In der Zeichenebene sei wie üblich ein rechtwinkliges Koordinatensystem zugrunde gelegt. Die Elemente von \mathbb{R}^2 entsprechen dann den Punkten in der Zeichenebene; dabei kann man sich, je nach Bedarf, ein Element $v = (a, b) \in \mathbb{R}^2$ als einen Punkt mit den Koordinaten (a, b) oder als einen Pfeil („Vektor“) vom Ursprung zu diesem Punkt vorstellen. Für $v = (2, 1)$ erhält man beispielsweise:



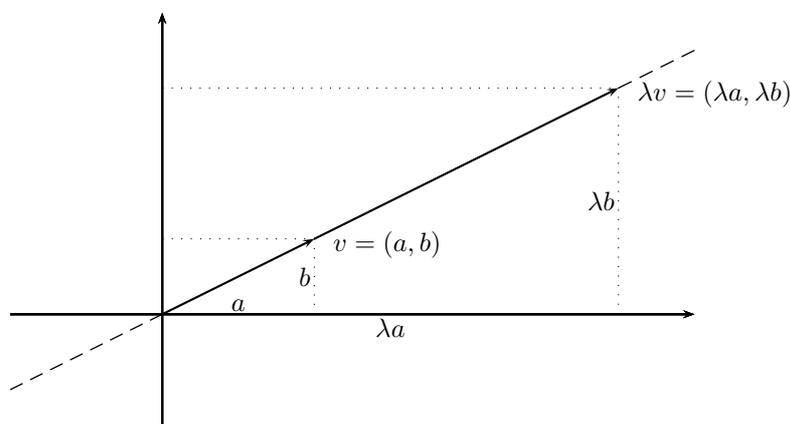
8.7.1 Addition für $V = \mathbb{R}^2$

Der komponentenweisen Addition $(a, b) + (c, d) = (a + c, b + d)$ entspricht anschaulich die aus der Schule bekannte Addition von Vektoren („Parallelogramm der Kräfte“, „man hängt die Pfeile aneinander“):



8.7.2 Skalare Multiplikation für $V = \mathbb{R}^2$

Es sei $v = (a, b) \neq (0, 0) \in \mathbb{R}^2$ und $\lambda \in \mathbb{R}$. Skalare Multiplikation von v mit λ ergibt $\lambda v = (\lambda a, \lambda b)$. Der Punkt $(\lambda a, \lambda b)$ liegt dann auf der Geraden, die durch $(0, 0)$ und (a, b) geht; den Übergang von v zu λv kann man sich als eine Streckung bzw. Stauchung um den Faktor λ vorstellen, wobei im Fall $\lambda < 0$ zusätzlich die Orientierung („Pfeilrichtung“) geändert wird.

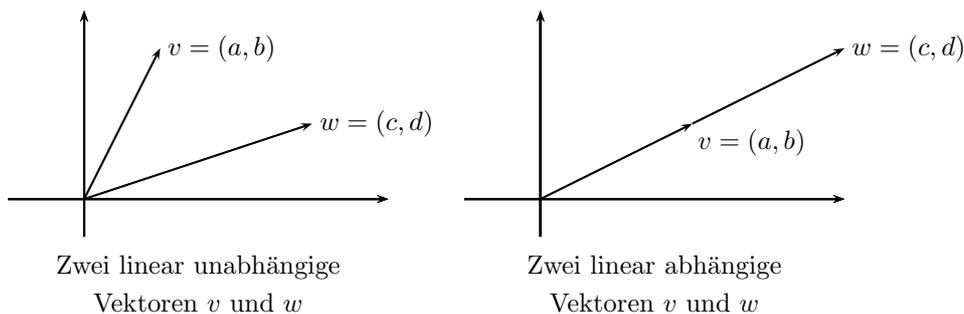


In der obigen Skizze wurde λv für $\lambda = 3$ angegeben. Wo liegen die Punkte λv für $\lambda = 2$, $\lambda = \frac{1}{2}$, $\lambda = 0$, $\lambda = -1$ und $\lambda = -2$?

Durch $\{\lambda(a, b) : \lambda \in \mathbb{R}\}$ wird die Menge aller Punkte beschrieben, die auf der Geraden liegen, die durch (a, b) und den Ursprung geht (für $(a, b) \neq (0, 0)$).

8.7.3 Linear abhängige und linear unabhängige Vektoren für $V = \mathbb{R}^2$

Es seien $v, w \in \mathbb{R}^2$, etwa $v = (a, b)$, $w = (c, d)$. Wir wissen (vgl. die Bemerkung nach dem Beweis von Feststellung 4 in Abschnitt 8.5): v und w sind genau dann linear unabhängig, wenn keiner dieser Vektoren ein skalares Vielfaches des anderen ist. Geometrisch bedeutet das: v und w sind genau dann linear unabhängig, wenn v und w nicht auf derselben Geraden durch den Ursprung liegen.

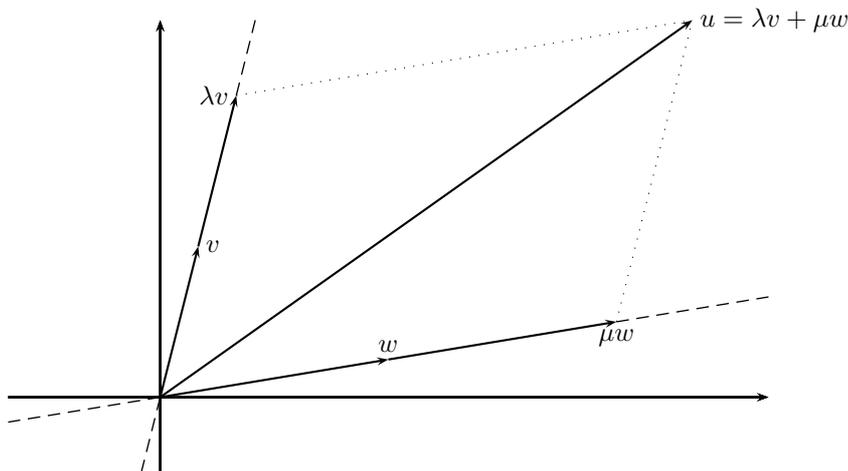


Aufgrund von Satz 5 gilt: Sind $v, w \in \mathbb{R}^2$ linear unabhängig, so bilden v und w eine Basis des \mathbb{R}^2 . Man kann dann also jeden Vektor $u \in \mathbb{R}^2$ als Linearkombination

$$u = \lambda v + \mu w$$

darstellen, wobei λ und μ (nach Feststellung 5) eindeutig bestimmt sind.

Durch die folgende Skizze werden die soeben gemachten Feststellungen illustriert:



8.7.4 Bestimmung der Unterräume des \mathbb{R}^2

Aufgrund von Feststellung 6 muss für jeden Unterraum U von \mathbb{R}^2 gelten: $\dim U \leq 2$. Es kann also nur 3 Typen von Unterräumen des \mathbb{R}^2 geben, nämlich diejenigen der Dimensionen 0, 1 und 2.

1) $\dim U = 2$.

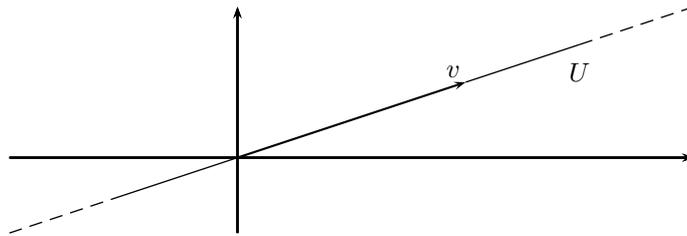
Aufgrund von Feststellung 6 gibt es nur einen Unterraum U von \mathbb{R}^2 , für den $\dim U = 2$ gilt, nämlich \mathbb{R}^2 selber.

2) $\dim U = 0$.

Der Unterraum $U = \{0\}$ ist der einzige, für den $\dim U = 0$ gilt. (Begründung: $\dim U = 0$ bedeutet, dass das Nulltupel \emptyset eine Basis von U ist. Da $\text{Lin}(\emptyset) = \{0\}$ gilt und da das Nulltupel \emptyset linear unabhängig ist, folgt die Behauptung.)

3) $\dim U = 1$.

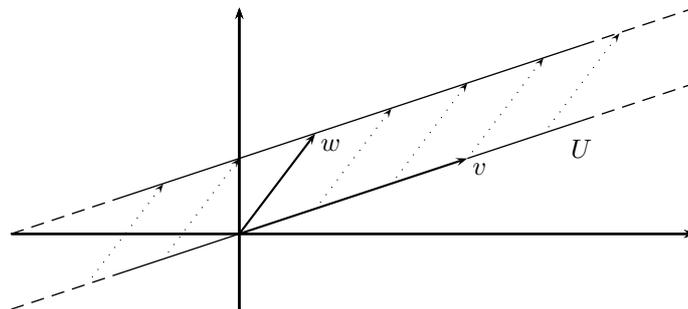
In diesem Fall besitzt U eine Basis, die nur aus einem Vektor $v \neq 0$ besteht; es folgt $U = \text{Lin}(v) = \{\lambda v : \lambda \in \mathbb{R}\}$, d.h., anschaulich handelt es sich bei U um eine Gerade durch den Ursprung.



Eine Bemerkung zu Geraden, die nicht durch den Ursprung gehen: Es liege wie oben eine Gerade $\text{Lin}(v) = \{\lambda v : \lambda \in \mathbb{R}\}$ durch den Ursprung vor und $w \in \mathbb{R}^2$ sei ein Vektor, für den $w \notin \text{Lin}(v)$ gilt. Dann wird durch

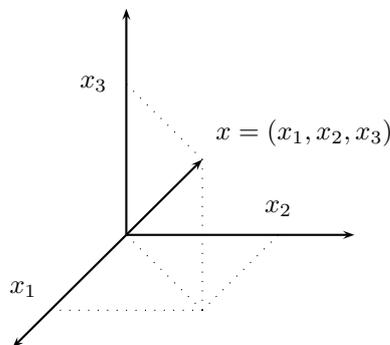
$$\{\lambda v + w : \lambda \in \mathbb{R}\}$$

eine Gerade beschrieben, die nicht durch den Ursprung geht (und dementsprechend kein Unterraum ist).



8.7.5 Vektoren im \mathbb{R}^3

In ähnlicher Weise wie \mathbb{R}^2 kann man sich den Vektorraum \mathbb{R}^3 veranschaulichen: An die Stelle eines 2-dimensionalen tritt nun ein 3-dimensionales rechtwinkliges Koordinatensystem.



Die Addition und die skalare Multiplikation veranschaulicht man sich dann analog zu \mathbb{R}^2 ; durch

$$\{\lambda(x_1, x_2, x_3) : \lambda \in \mathbb{R}\} \quad (8.3)$$

wird die Menge aller Punkte beschrieben, die auf der Geraden liegen, die durch (x_1, x_2, x_3) und den Ursprung geht (für $(x_1, x_2, x_3) \neq (0, 0, 0)$).

Es seien nun $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3) \in \mathbb{R}^3$ zwei linear unabhängige Vektoren. Wir betrachten die lineare Hülle von x und y (oder, wie man auch sagt, den *von x und y aufgespannten Untervektorraum*):

$$\text{Lin}(x, y) = \{\lambda x + \mu y : \lambda, \mu \in \mathbb{R}\} = \{\lambda(x_1, x_2, x_3) + \mu(y_1, y_2, y_3) : \lambda, \mu \in \mathbb{R}\}. \quad (8.4)$$

Anschaulich ist dies die Ebene, die durch den Ursprung geht und die beiden Geraden $\{\lambda(x_1, x_2, x_3) : \lambda \in \mathbb{R}\}$ und $\{\mu(y_1, y_2, y_3) : \mu \in \mathbb{R}\}$ enthält.

Ähnlich wie für \mathbb{R}^2 findet man, dass es folgende Typen von Unterräumen des \mathbb{R}^3 gibt:

- 1) $\{0\}$ ist der einzige Unterraum der Dimension 0 und \mathbb{R}^3 ist der einzige Unterraum der Dimension 3.
- 2) *Unterräume der Dimension 1*: Dies sind die Unterräume vom Typ (8.3) („Geraden durch den Ursprung“).
- 3) *Unterräume der Dimension 2*: Dies sind die Unterräume vom Typ (8.4) („Ebenen durch den Ursprung“).

Dass drei Vektoren im \mathbb{R}^3 linear unabhängig sind, bedeutet anschaulich, dass sie nicht in einer Ebene liegen, die durch den Ursprung geht.

Zur Übung: Es seien v_0, v_1 und v_2 linear unabhängige Vektoren des \mathbb{R}^3 . Man gebe eine geometrische Beschreibung der folgenden Teilmenge des \mathbb{R}^3 :

$$\{\lambda_1 v_1 + \lambda_2 v_2 + v_0 : \lambda_1, \lambda_2 \in \mathbb{R}\}.$$

9 Lineare Abbildungen und Matrizen

9.1 Lineare Abbildungen

Es seien V und W Vektorräume über K und

$$f : V \longrightarrow W$$

sei eine Abbildung. Wir nennen f *verträglich mit der Addition* von V und W , falls für alle $x, y \in V$ Folgendes gilt: Wenn f die Elemente x und y auf x' bzw. y' abbildet, so soll $x + y$ nicht irgendwohin abgebildet werden, sondern gerade auf $x' + y'$. Dasselbe, nur etwas kompakter formuliert: f heißt *verträglich mit der Addition* von V und W , falls für alle $x, y \in V$ gilt:

$$(i) \quad f(x + y) = f(x) + f(y)$$

Analog definiert man: f ist *verträglich mit der skalaren Multiplikation* von V und W , falls für alle $x \in V$ und $\lambda \in K$ gilt:

$$(ii) \quad f(\lambda x) = \lambda f(x)$$

Definition.

Es seien V und W Vektorräume über K . Eine Abbildung $f : V \rightarrow W$ heißt eine *lineare Abbildung* (oder ein *Homomorphismus*), falls f sowohl mit der Addition als auch mit der skalaren Multiplikation von V und W verträglich ist. Mit anderen Worten: Für alle $x, y \in V$ und $\lambda \in K$ muss Folgendes gelten:

$$(i) \quad f(x + y) = f(x) + f(y)$$

$$(ii) \quad f(\lambda x) = \lambda f(x)$$

Man kann (i) auch so aussprechen: Es ist gleichgültig, ob man zwei Elemente aus V zuerst in V addiert und dann deren Summe abbildet oder ob man erst die beiden Elemente abbildet und dann deren Bilder in W addiert. Kurz gesagt: *Das Bild einer Summe ist stets die Summe der Bilder.*

Entsprechendes gilt für (ii): *Das Bild des λ -fachen eines Vektors ist stets das λ -fache seines Bildes.*

Beispiele:

1. Die Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ sei gegeben durch $f(x_1, x_2, x_3) = (2x_1 + x_2, x_3)$. Dies ist eine lineare Abbildung, da für alle $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3) \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$ gilt:

$$\begin{aligned} (i) \quad f(x + y) &= f(x_1 + y_1, x_2 + y_2, x_3 + y_3) \\ &= (2(x_1 + y_1) + x_2 + y_2, x_3 + y_3) \\ &= (2x_1 + x_2, x_3) + (2y_1 + y_2, y_3) \\ &= f(x) + f(y), \end{aligned}$$

$$\begin{aligned} (ii) \quad f(\lambda x) &= f(\lambda x_1, \lambda x_2, \lambda x_3) \\ &= (2\lambda x_1 + \lambda x_2, \lambda x_3) \\ &= (\lambda(2x_1 + x_2), \lambda x_3) \\ &= \lambda(2x_1 + x_2, x_3) \\ &= \lambda f(x). \end{aligned}$$

2. $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ mit $f(x_1, x_2, x_3) = (x_1 \cdot x_2, x_3)$ ist *keine* lineare Abbildung, da es $x, y \in \mathbb{R}^3$ mit $f(x + y) \neq f(x) + f(y)$ gibt: Wählt man beispielsweise $x = (1, 0, 0)$ und $y = (0, 1, 0)$, so gilt $x + y = (1, 1, 0)$ und man erhält $f(x + y) = f(1, 1, 0) = (1 \cdot 1, 0) = (1, 0)$, aber $f(x) + f(y) = (1 \cdot 0, 0) + (0 \cdot 1, 0) = (0, 0) + (0, 0) = (0, 0)$.

3. **Zur Übung:** Die Abbildungen $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ und $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ seien gegeben durch $g(x_1, x_2) = (2x_1, x_2 + 1)$ und $h(x_1, x_2) = x_1 - x_2$. Prüfen Sie sowohl für g als auch für h , ob eine lineare Abbildung vorliegt.
4. Es sei V der (reelle) Vektorraum der unendlich oft differenzierbaren Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$. Wir betrachten die Abbildung

$$D : V \longrightarrow V$$

$$f \longmapsto D(f) := f',$$

die jeder Funktion $f \in V$ ihre Ableitung zuordnet. Aufgrund bekannter Ableitungsregeln gilt

$$(i) \quad D(f + g) = D(f) + D(g)$$

sowie

$$(ii) \quad D(\lambda f) = \lambda D(f).$$

D.h., die Abbildung D ist eine lineare Abbildung. (Da die Elemente von V , also die Vektoren, in diesem Beispiel selbst Abbildungen sind, wird D (der besseren Unterscheidung wegen) auch als *Operator* bezeichnet. Die Feststellungen (i) und (ii) besagen dann, dass D ein *linearer Operator* ist.)

Feststellung 1.

V , W und Y seien Vektorräume über K ; $f : V \rightarrow W$, $g : W \rightarrow Y$ seien lineare Abbildungen. Dann ist auch die Nacheinanderausführung $g \circ f$ eine lineare Abbildung.

Beweis. Zu zeigen ist, dass für alle $x, y \in V$ und $\lambda \in K$ gilt:

$$(i) \quad (g \circ f)(x + y) = (g \circ f)(x) + (g \circ f)(y),$$

$$(ii) \quad (g \circ f)(\lambda x) = \lambda(g \circ f)(x).$$

Nachweis von (i):

$$\begin{aligned} (g \circ f)(x + y) &= g(f(x + y)) \\ &\stackrel{1}{=} g(f(x) + f(y)) \\ &\stackrel{2}{=} g(f(x)) + g(f(y)) \\ &= (g \circ f)(x) + (g \circ f)(y). \end{aligned}$$

(Schauen Sie sich diese Rechnung, die (i) nachweist, genau an und beantworten Sie die folgende Frage: Welche Voraussetzungen wurden an den mit „1“ und „2“ gekennzeichneten Stellen benutzt?)

Nachweis von (ii): Diesen Beweis führen Sie bitte selbst; er geht ähnlich wie der Nachweis von (i). \square

Ist V ein Vektorraum über K , so bezeichnen wir mit

$$\text{Id}_V$$

die *Identität* auf V , d.h., Id_V ist diejenige Abbildung $V \rightarrow V$, die jedes Element von V auf sich selbst abbildet. Es gilt also $\text{Id}_V(v) = v$ für alle $v \in V$.

Feststellung 2.

Die Identität Id_V ist eine lineare Abbildung.

Beweis. Es ist also für alle $x, y \in V$ und $\lambda \in K$ zu zeigen:

$$(i) \quad \text{Id}_V(x + y) = \text{Id}_V(x) + \text{Id}_V(y),$$

$$(ii) \quad \text{Id}_V(\lambda x) = \lambda \text{Id}_V(x).$$

Das ist in beiden Fällen einfach; hier sei es für (ii) vorgemacht, (i) erledigen Sie bitte selbst.

Nachweis von (ii): Da die Abbildung Id_V alle Elemente von V auf sich selbst abbildet, gilt $\text{Id}_V(\lambda x) = \lambda x$ und $\text{Id}_V(x) = x$; also $\text{Id}_V(\lambda x) = \lambda x = \lambda \text{Id}_V(x)$. \square

Sind f und g lineare Abbildungen $V \rightarrow W$, so wird die Abbildung $f + g$ als diejenige Abbildung $V \rightarrow W$ definiert, für die gilt:

$$(f + g)(x) = f(x) + g(x) \text{ für alle } x \in V.$$

Ganz entsprechend definiert man für jedes $\lambda \in K$ die Abbildung $\lambda f : V \rightarrow W$ durch

$$(\lambda f)(x) = \lambda f(x) \text{ für alle } x \in V.$$

Feststellung 3.

Für alle linearen Abbildungen $f, g : V \rightarrow W$ und alle $\lambda \in K$ gilt: $f + g$ und λf sind ebenfalls lineare Abbildungen.

Beweis. Nachzuweisen ist, dass für alle $x, y \in V$ und alle $\mu \in K$ gilt:

- (i) $(f + g)(x + y) = (f + g)(x) + (f + g)(y)$,
- (ii) $(f + g)(\mu x) = \mu(f + g)(x)$,
- (iii) $(\lambda f)(x + y) = (\lambda f)(x) + (\lambda f)(y)$,
- (iv) $(\lambda f)(\mu x) = \mu(\lambda f)(x)$.

Wir führen nur (i) vor:

$$\begin{aligned} (f + g)(x + y) &= f(x + y) + g(x + y) \\ &\stackrel{*}{=} f(x) + f(y) + g(x) + g(y) \\ &= f(x) + g(x) + f(y) + g(y) \\ &= (f + g)(x) + (f + g)(y) \end{aligned}$$

An der mit * markierten Stelle wurde die Linearität sowohl von f als auch von g benutzt. \square

Zusammenfassung der vorangegangenen Feststellungen.

Nacheinanderausführung, Summe und skalares Vielfaches von linearen Abbildungen sind wieder lineare Abbildungen; außerdem ist die Identität eine lineare Abbildung.

Die Menge aller linearen Abbildungen¹ von V nach W bezeichnet man auch mit

$$\text{Hom}(V, W).$$

Mit dieser Schreibweise kann man die obigen Feststellungen wie folgt formulieren:

- 1) $f \in \text{Hom}(V, W)$, $g \in \text{Hom}(W, Y) \implies g \circ f \in \text{Hom}(V, Y)$
- 2) $\text{Id}_V \in \text{Hom}(V, V)$
- 3) $f, g \in \text{Hom}(V, W)$, $\lambda \in K \implies f + g \in \text{Hom}(V, W)$ und $\lambda f \in \text{Hom}(V, W)$

Ist $f : V \rightarrow W$ eine lineare Abbildung, so definiert man das *Bild von f* bekanntlich als

$$f(V) = \left\{ w \in W : \text{Es gibt ein } v \in V \text{ mit } f(v) = w \right\}.$$

Außerdem definiert man den *Kern von f* als

$$f^{-1}(\{0\}) = \left\{ v \in V : f(v) = 0 \right\}.$$

¹Zur Erinnerung: Statt „lineare Abbildung“ sagt man auch „Homomorphismus“.

Das Bild von f wollen wir mit

Bild f

und den Kern von f mit

Kern f

bezeichnen.

Feststellung 4.

Für jede lineare Abbildung $f : V \rightarrow W$ ist Bild f ein Untervektorraum von W und Kern f ist ein Untervektorraum von V .

Beweis. Zur Erinnerung: Um zu zeigen, dass U ein Untervektorraum eines Vektorraums über K ist, hat man *drei Dinge* nachzuweisen:

- 1) $U \neq \emptyset$.
- 2) $x, y \in U \implies x + y \in U$.
- 3) $x \in U, \lambda \in K \implies \lambda x \in U$.

Der Beweis, dass Bild f ein Untervektorraum ist, sei dem Leser überlassen; wir zeigen hier, dass Kern f ein Untervektorraum von V ist. Dabei werden vor allem die Linearitätseigenschaften

- (i) $f(x + y) = f(x) + f(y)$ für alle $x, y \in V$
- (ii) $f(\lambda x) = \lambda f(x)$ für alle $x \in V$ und $\lambda \in K$

benutzt.

- 1) Wir wollen Kern $f \neq \emptyset$ nachweisen und zeigen zu diesem Zweck $0 \in \text{Kern } f$: Es gilt $f(0) = f(0 + 0) \stackrel{(i)}{=} f(0) + f(0)$; addiert man auf beiden Seiten $-f(0)$, so folgt $f(0) = 0$. Also: $0 \in \text{Kern } f$.
- 2) $x, y \in \text{Kern } f \implies f(x) = 0, f(y) = 0 \implies f(x + y) \stackrel{(i)}{=} f(x) + f(y) = 0 + 0 = 0 \implies x + y \in \text{Kern } f$.
- 3) $x \in \text{Kern } f, \lambda \in K \implies f(x) = 0 \implies f(\lambda x) \stackrel{(ii)}{=} \lambda f(x) = \lambda \cdot 0 = 0 \implies \lambda x \in \text{Kern } f$. \square

Feststellung 5.

Eine lineare Abbildung $f : V \rightarrow W$ ist genau dann *injektiv*, wenn Kern $f = \{0\}$.

Beweis. Es sind zwei Dinge zu zeigen:

- (a) f ist injektiv $\implies \text{Kern } f = \{0\}$:

f sei injektiv. Da Kern f ein Untervektorraum ist, gilt $0 \in \text{Kern } f$, d.h. $f(0) = 0$. Da f injektiv ist, kann es kein $v \neq 0$ geben mit $f(v) = 0$. Also gilt Kern $f = \{0\}$.

- (b) Kern $f = \{0\} \implies f$ ist injektiv:

Es gelte nun umgekehrt Kern $f = \{0\}$. Nachweis der Injektivität von f : Sei $f(v) = f(w) \implies f(v - w) \stackrel{(*)}{=} f(v) - f(w) = 0$. Also liegt $v - w$ in Kern f , woraus (wegen Kern $f = \{0\}$) $v - w = 0$ folgt. Also $v = w$. Dies zeigt, dass f injektiv ist. \square

Die mit (*) gekennzeichnete Stelle leuchtet möglicherweise nicht jedem unmittelbar ein; sie lässt sich wie folgt rechtfertigen:

$$f(v - w) = f(v + (-1)w) = f(v) + (-1)f(w) = f(v) - f(w).$$

↑
Hier wird die Linearität von f benutzt (sowohl Bedingung (i) als auch (ii)).

Definition.

Eine lineare Abbildung $f : V \rightarrow W$ heißt ein

- *Monomorphismus*, falls f injektiv ist.
- *Epimorphismus*, falls f surjektiv ist.
- *Isomorphismus*, falls f bijektiv ist.
- *Endomorphismus*, falls $V = W$ gilt.
- *Automorphismus*, falls f bijektiv ist und $V = W$ gilt.

Feststellung 6.

Ist $f : V \rightarrow W$ ein Isomorphismus, so ist auch die Umkehrabbildung $f^{-1} : W \rightarrow V$ ein Isomorphismus.

Beweis. Siehe Jänich Seite 82. \square

Wir kommen nun zum *ersten Satz* des Abschnitts 9; er bezieht sich auf lineare Abbildungen $f : V \rightarrow W$, für die vorausgesetzt wird, dass V endlichdimensional ist, d.h., V besitzt eine Basis (v_1, \dots, v_n) . Der Satz enthält *zwei Teilaussagen*:

- (a) **Eindeutigkeitsaussage.** Diese Teilaussage stellt Folgendes fest: Gibt man für jeden der n Basisvektoren v_i vor, welches Element von W das Bild von v_i sein soll ($i = 1, \dots, n$), so existiert *höchstens eine* lineare Abbildung $f : V \rightarrow W$, die diese Vorgaben erfüllt.
- (b) **Existenzaussage.** Diese Teilaussage stellt fest: Gibt man für jeden der n Basisvektoren v_i vor, welches Element von W das Bild von v_i sein soll ($i = 1, \dots, n$), so existiert tatsächlich immer eine lineare Abbildung $f : V \rightarrow W$, die diese Vorgaben erfüllt.

Satz 1.

Es seien V und W Vektorräume über K und (v_1, \dots, v_n) sei eine Basis von V . Dann gibt es *zu jedem* n -Tupel (w_1, \dots, w_n) von Vektoren in W *genau eine* lineare Abbildung $f : V \rightarrow W$ mit $f(v_i) = w_i$ ($i = 1, \dots, n$).

Beweis.

- (a) **Beweis der Eindeutigkeit.** Für die lineare Abbildung $f : V \rightarrow W$ gelte $f(v_i) = w_i$ ($i = 1, \dots, n$). Wir haben zu zeigen, dass f dadurch schon vollständig festgelegt ist, dass also für jedes $v \in V$ schon feststeht, was das Bild $f(v)$ ist. Das ist aber ganz einfach einzusehen: Da (v_1, \dots, v_n) eine Basis von V ist, kann man v mit eindeutig bestimmten Koeffizienten λ_i als Linearkombination der v_i schreiben:

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n.$$

Da f linear ist, folgt:

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n) \\ &= f(\lambda_1 v_1) + f(\lambda_2 v_2) + \dots + f(\lambda_n v_n) \\ &= \lambda_1 f(v_1) + \lambda_2 f(v_2) + \dots + \lambda_n f(v_n) \\ &= \lambda_1 w_1 + \lambda_2 w_2 + \dots + \lambda_n w_n \end{aligned}$$

D.h., das Bild $f(v)$ steht tatsächlich bereits fest.

- (b) **Beweis der Existenz.** Aufgrund von a) ist klar, wie man eine Abbildung $f : V \rightarrow W$ definieren muss, damit f linear ist mit $f(v_i) = w_i$ ($i = 1, \dots, n$); man muss für $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ festlegen:

$$f(v) := \lambda_1 w_1 + \dots + \lambda_n w_n. \quad (*)$$

Wir zeigen, dass durch diese Festlegung tatsächlich eine Abbildung $f : V \rightarrow W$ definiert wird, für die gilt:

1. f ist linear,
2. $f(v_i) = w_i$ ($i = 1, \dots, n$).

Zu 1: Es sei $x, y \in V$ mit $x = \lambda_1 v_1 + \dots + \lambda_n v_n$ und $y = \mu_1 v_1 + \dots + \mu_n v_n$. Es folgt:

$$\begin{aligned} f(x+y) &= f\left((\lambda_1 + \mu_1)v_1 + \dots + (\lambda_n + \mu_n)v_n\right) \\ &\stackrel{(*)}{=} (\lambda_1 + \mu_1)w_1 + \dots + (\lambda_n + \mu_n)w_n \\ &= \lambda_1 w_1 + \dots + \lambda_n w_n + \mu_1 w_1 + \dots + \mu_n w_n \\ &= f(x) + f(y). \end{aligned}$$

Ähnlich zeigt man $f(\lambda x) = \lambda f(x)$ für alle $x \in V$ und $\lambda \in K$. Also ist die durch (*) definierte Abbildung f linear.

Zu 2: Es sei v_i ein Element der Basis (v_1, \dots, v_n) (für $i \in \{1, \dots, n\}$). Dann gilt $v_i = \lambda_1 v_1 + \dots + \lambda_n v_n$ mit $\lambda_i = 1$ und $\lambda_j = 0$ für alle $j \neq i$. Also folgt aus (*):

$$f(v_i) = \lambda_1 w_1 + \dots + \lambda_n w_n = w_i. \quad \square$$

Beispiele:

1. **Frage:** Wie viele lineare Abbildungen $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ gibt es?

Antwort: *unendlich viele.*

Begründung: Es gibt unendlich viele Möglichkeiten, den Vektoren e_1, \dots, e_n der kanonischen Basis von \mathbb{R}^n irgendwelche Bilder w_1, \dots, w_n aus \mathbb{R}^m zuzuordnen. Zu jeder solchen Zuordnung gehört nach Satz 1 eine lineare Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^m$.

2. Es sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ eine lineare Abbildung mit $f(1, 0, 0) = (1, 2)$, $f(0, 1, 0) = (-1, 0)$ und $f(0, 0, 1) = (3, 1)$. Damit hat man den Vektoren $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ und $e_3 = (0, 0, 1)$ der kanonischen Basis also Bilder $w_1 = (1, 2)$, $w_2 = (-1, 0)$ und $w_3 = (3, 1)$ aus \mathbb{R}^2 zugeordnet. Oben wurde gesagt, dass dadurch die Bilder sämtlicher Vektoren $v \in \mathbb{R}^3$ bereits festgelegt sind. Probieren wir es aus: Was ist $f(v)$ für $v = (-3, 7, 5)$?

Antwort: Es gilt $v = (-3, 7, 5) = (-3)e_1 + 7e_2 + 5e_3$; also (aufgrund der Linearität von f):

$$\begin{aligned} f(v) &= (-3)f(e_1) + 7f(e_2) + 5f(e_3) \\ &= (-3)(1, 2) + 7(-1, 0) + 5(3, 1) \\ &= (-3, -6) + (-7, 0) + (15, 5) \\ &= (5, -1). \end{aligned}$$

3. Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ eine lineare Abbildung mit $f(1, 0) = (3, 4, 5)$ und $f(0, 1) = (4, 2, 4)$. Berechnen Sie $f(-2, 2)$.
4. Es sei $K = \mathbb{Z}_p$; der Körper K hat also genau p Elemente. *Beantworten Sie nacheinander folgende Fragen:*
 - a) Wie viele Elemente enthält eine beliebige Basis von K^n ?
 - b) Wie viele Elemente enthält K^m ?
 - c) Wie viele verschiedene lineare Abbildungen $f: K^n \rightarrow K^m$ gibt es?
5. Für $K = \mathbb{Z}_2$: Wie viele Abbildungen $f: K^3 \rightarrow K^2$ gibt es und wie viele davon sind linear?

Es sei $f: V \rightarrow W$ eine lineare Abbildung und (v_1, \dots, v_n) sei eine Basis von V . Wie wir wissen, ist f durch Angabe der Bilder der Basisvektoren v_1, \dots, v_n vollständig beschrieben, d.h., durch

$$f(v_1) = w_1, f(v_2) = w_2, \dots, f(v_n) = w_n$$

ist f eindeutig festgelegt.

Frage: Kann man an den Bildern $f(v_1) = w_1, \dots, f(v_n) = w_n$ irgendwie erkennen, ob f injektiv ist?

Antwort: Ja, das geht. Wie es geht, steht in der folgenden Feststellung.

Feststellung 7.

Es seien V und W Vektorräume über K , (v_1, \dots, v_n) sei eine Basis von V und $f : V \rightarrow W$ sei eine lineare Abbildung. Dann gilt: f ist genau dann injektiv, wenn die Vektoren $f(v_1), \dots, f(v_n)$ linear unabhängig in W sind.

Beweis. Wir haben zwei Dinge zu zeigen:

(a) f ist injektiv $\implies (f(v_1), \dots, f(v_n))$ ist linear unabhängig:

f sei also injektiv. Es gelte $\lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = 0$. Zu zeigen ist: $\lambda_1 = \dots = \lambda_n = 0$. Aufgrund der Linearität von f gilt: $f(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = 0$. Außerdem weiß man: $f(0) = 0$. Aufgrund der Injektivität von f muss also $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ gelten. Da (v_1, \dots, v_n) linear unabhängig ist, folgt $\lambda_1 = \dots = \lambda_n = 0$.

(b) $(f(v_1), \dots, f(v_n))$ ist linear unabhängig $\implies f$ ist injektiv:

Umgekehrt sei also $(f(v_1), \dots, f(v_n))$ linear unabhängig. Aufgrund von Feststellung 5 haben wir Kern $f = \{0\}$ zu zeigen. Wir zeigen dies indirekt und nehmen daher an, dass Kern $f \neq \{0\}$ gilt; dann gibt es ein $v \neq 0$ mit $f(v) = 0$. Da (v_1, \dots, v_n) eine Basis von V ist, gibt es $\lambda_1, \dots, \lambda_n \in K$ mit $\lambda_1 v_1 + \dots + \lambda_n v_n = v$. Da $v \neq 0$ gilt, sind nicht alle diese λ_i gleich 0. Es folgt:

$$0 = f(v) = f(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n).$$

Dies ist dann aber eine nichttriviale Darstellung des Nullvektors, im Widerspruch zur Annahme, dass $(f(v_1), \dots, f(v_n))$ linear unabhängig ist. \square

Ein ähnliches Ergebnis gilt auch für die Surjektivität von f .

Feststellung 8.

Es seien V und W Vektorräume über K , (v_1, \dots, v_n) sei eine Basis von V und $f : V \rightarrow W$ sei eine lineare Abbildung. Dann gilt: f ist genau dann surjektiv, wenn $(f(v_1), \dots, f(v_n))$ ein Erzeugendensystem von W ist.

Der Beweis von Feststellung 8 sei als Übungsaufgabe empfohlen. Aus den Feststellungen 7 und 8 erhält man unmittelbar die folgende Feststellung.

Feststellung 9.

Es seien V und W Vektorräume über K , (v_1, \dots, v_n) sei eine Basis von V und $f : V \rightarrow W$ sei eine lineare Abbildung. Dann gilt: f ist genau dann bijektiv (oder anders gesagt, ein Isomorphismus), wenn $(f(v_1), \dots, f(v_n))$ eine Basis von W ist.

Zusammenfassung der Feststellungen 7 - 9.

Unter den Voraussetzungen dieser Feststellungen gilt:

$$\begin{aligned} f \text{ ist injektiv} &\iff (f(v_1), \dots, f(v_n)) \text{ ist linear unabhängig} \\ f \text{ ist surjektiv} &\iff (f(v_1), \dots, f(v_n)) \text{ ist ein Erzeugendensystem von } W \\ f \text{ ist bijektiv} &\iff (f(v_1), \dots, f(v_n)) \text{ ist eine Basis von } W \end{aligned}$$

Man nennt Vektorräume V und W *isomorph* (Bezeichnung: $V \cong W$), wenn es einen Isomorphismus $f : V \rightarrow W$ gibt. Aus Satz 1 und Feststellung 9 ergibt sich der folgende Satz.

Satz 2.

Es sei V ein Vektorraum über K mit $\dim V = n < \infty$. Dann ist V isomorph zu K^n .

Beweis. Wegen $\dim V = n < \infty$ gibt es eine Basis (v_1, \dots, v_n) von V . Es sei (e_1, \dots, e_n) die kanonische Basis von K^n . Nach Satz 1 gibt es eine lineare Abbildung $f : V \rightarrow K^n$ mit $f(v_i) = e_i$ ($i = 1, \dots, n$) und nach Feststellung 9 ist f ein Isomorphismus. \square

Es sei $f : V \rightarrow W$ eine lineare Abbildung. Wie wir wissen (vgl. Feststellung 4), ist $\text{Bild } f$ ein Untervektorraum von W . Ist der Untervektorraum $\text{Bild } f$ endlichdimensional, so nennt man die Dimension von $\text{Bild } f$ den *Rang* von f ; man bezeichnet den Rang von f mit $\text{rg } f$. Es gilt also:

$$\text{rg } f = \dim \text{Bild } f.$$

Wir kommen zur *Dimensionsformel für lineare Abbildungen*:

Satz 3 (Dimensionsformel für lineare Abbildungen).

Es sei V ein Vektorraum über K mit $\dim V = n < \infty$ und $f : V \rightarrow W$ sei eine lineare Abbildung. Dann ist $\text{Bild } f$ endlichdimensional und es gilt:

$$\dim \text{Kern } f + \text{rg } f = n.$$

Beweis. Es sei (v_1, \dots, v_r) eine Basis von $\text{Kern } f$. Nach dem Basisergänzungssatz können wir (durch Hinzunahme von Vektoren aus einer beliebigen Basis von V) das r -Tupel (v_1, \dots, v_r) zu einer Basis $(v_1, \dots, v_r, v_{r+1}, \dots, v_n)$ von V ergänzen. Es seien die Vektoren

$$w_1, \dots, w_{n-r}$$

definiert durch $w_1 := f(v_{r+1})$, $w_2 := f(v_{r+2})$, \dots , $w_{n-r} := f(v_n)$. Wie auf Seite 87 im Jänich kann man zeigen, dass w_1, \dots, w_{n-r} eine Basis von $\text{Bild } f$ ist. Es gilt also $\dim \text{Kern } f = r$ und $\text{rg } f = \dim \text{Bild } f = n - r$, woraus die behauptete Dimensionsformel folgt. \square

Folgerung aus Satz 3. Es seien V und W Vektorräume mit $\dim V = \dim W = n < \infty$ und $f : V \rightarrow W$ sei eine lineare Abbildung. Dann gilt: f ist genau dann surjektiv, wenn f injektiv ist.

Beweis.

$$\begin{aligned} f \text{ ist surjektiv} & \iff \text{Bild } f = W \\ & \iff \dim \text{Bild } f = \dim W \\ & \iff \text{rg } f = n \\ & \stackrel{\text{Satz 3}}{\iff} \dim \text{Kern } f = 0 \\ & \iff \text{Kern } f = \{0\} \\ & \stackrel{\text{Feststellung 5}}{\iff} f \text{ ist injektiv. } \square \end{aligned}$$

Obige Folgerung wird meist auf folgende Art angewendet werden: Für Vektorräume V, W mit $\dim V = \dim W < \infty$ soll gezeigt werden, dass eine bestimmte lineare Abbildung $f : V \rightarrow W$ bijektiv ist. Dann braucht man nur eines von beidem zu zeigen: Entweder, dass f surjektiv ist oder dass f injektiv ist. Die jeweils andere Eigenschaft folgt dann „automatisch“ aufgrund der Folgerung aus Satz 3.

9.2 Der Zusammenhang zwischen linearen Abbildungen und Matrizen

Matrizen spielen in den unterschiedlichsten Zusammenhängen sowohl innerhalb als auch außerhalb der Linearen Algebra eine wichtige Rolle. Hier interessieren wir uns für Matrizen, weil es einen sehr engen Zusammenhang zwischen linearen Abbildungen und Matrizen gibt. Mit

$$M(m \times n, K)$$

sei die Menge aller $m \times n$ -Matrizen über einem Körper K bezeichnet.

Grundbegriffe über Matrizen haben wir bereits im 1. Semester kennengelernt, wo wir insbesondere auch die Matrizenmultiplikation besprochen haben, an die noch einmal mithilfe einiger **Beispiele** erinnert werden soll.

1. $A \in M(3 \times 4, \mathbb{R})$ und $B \in M(4 \times 2, \mathbb{R})$ seien gegeben durch

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 3 & 0 \\ 1 & 1 & -1 & -1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 3 \\ 0 & 2 \\ 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Dann lässt sich das Produkt AB bilden und es gilt $AB \in M(3 \times 2, \mathbb{R})$; es ergibt sich

$$\begin{aligned} AB &= \begin{pmatrix} 1 \cdot 1 + 0 \cdot 0 + 2 \cdot 0 + 1 \cdot 1 & 1 \cdot 3 + 0 \cdot 2 + 2 \cdot (-1) + 1 \cdot 0 \\ (-1) \cdot 1 + 2 \cdot 0 + 3 \cdot 0 + 0 \cdot 1 & (-1) \cdot 3 + 2 \cdot 2 + 3 \cdot (-1) + 0 \cdot 0 \\ 1 \cdot 1 + 1 \cdot 0 + (-1) \cdot 0 + (-1) \cdot 1 & 1 \cdot 3 + 1 \cdot 2 + (-1) \cdot (-1) + (-1) \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ -1 & -2 \\ 0 & 6 \end{pmatrix}. \end{aligned}$$

Das Produkt AB kann gebildet werden, da die Zeilenlänge von A (= Anzahl der Spalten von A) mit der Spaltenlänge von B (= Anzahl der Zeilen von B) übereinstimmt:

$$\begin{array}{ccc} A \in M(3 \times 4, \mathbb{R}), & B \in M(4 \times 2, \mathbb{R}) \\ & \swarrow \quad \searrow \\ & \text{„=“} \end{array}$$

2. Es sei $A \in M(3 \times 4, \mathbb{R})$ und $B \in M(4 \times 1, \mathbb{R})$ mit A wie im ersten Beispiel und B gegeben durch

$$B = \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

Dann ist AB eine 3×1 -Matrix, die man wie folgt erhält:

$$A \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 3 & 0 \\ 1 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \\ 0 \end{pmatrix}.$$

Dasselbe können wir natürlich auch mit jeder beliebigen Spalte mit Einträgen $x_i \in \mathbb{R}$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

machen:

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 3 & 0 \\ 1 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 + 0x_2 + 2x_3 + x_4 \\ -x_1 + 2x_2 + 3x_3 + 0x_4 \\ x_1 + x_2 - x_3 - x_4 \end{pmatrix}.$$

Auf diese Art entsteht also eine Abbildung

$$f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3,$$

die jedem als Spalte geschriebenen 4-Tupel

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

ein als Spalte geschriebenes 3-Tupel

$$Ax = A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 + 0x_2 + 2x_3 + x_4 \\ -x_1 + 2x_2 + 3x_3 + 0x_4 \\ x_1 + x_2 - x_3 - x_4 \end{pmatrix} \quad (*)$$

zuordnet.

Da die Abbildung f mithilfe der Matrix A gebildet wurde, wollen wir sie mit f_A bezeichnen.

Eine weitere Bemerkung zur Bezeichnung: Es wird in manchen Situationen nicht so sehr darauf ankommen, ob ein n -Tupel $(x_1, \dots, x_n) \in \mathbb{R}^n$ als Zeile oder als Spalte aufgeschrieben wird; schon aus Platzgründen werden wir dann die Zeilenschreibweise bevorzugen. *Es gibt aber andererseits auch Situationen, in denen man ein n -Tupel unbedingt als Spalte schreiben muss*: Beispielsweise **muss** in der obigen Gleichung (*) das 4-Tupel x als Spalte geschrieben werden. (Wieso nämlich?)

Mithilfe der 3×4 - Matrix A haben wir im obigen Beispiel eine zu A gehörige Abbildung

$$f_A : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$$

erhalten, nämlich die Abbildung, die durch

$$f_A(x_1, x_2, x_3, x_4) = (x_1 + 0x_2 + 2x_3 + x_4, -x_1 + 2x_2 + 3x_3 + 0x_4, x_1 + x_2 - x_3 - x_4)$$

gegeben ist. Verallgemeinert man dies, so gelangt man zur folgenden Definition.

Definition.

Zu jeder $m \times n$ - Matrix A über K (also $A \in M(m \times n, K)$) mit

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

ist die zugehörige Abbildung

$$f_A : K^n \longrightarrow K^m$$

gegeben durch (für $x = (x_1, \dots, x_n)$):

$$f_A(x) = f_A(x_1, \dots, x_n) = \left(\sum_{i=1}^n a_{1i}x_i, \sum_{i=1}^n a_{2i}x_i, \dots, \sum_{i=1}^n a_{mi}x_i \right).$$

Fasst man das n -Tupel $x = (x_1, \dots, x_n)$ als einen Spaltenvektor auf (für den wir der Einfachheit halber ebenfalls x schreiben wollen), so kann man f_A auch mithilfe des Matrizenprodukts durch die Gleichung

$$f_A(x) = Ax$$

beschreiben; das Ergebnis ist dabei dasselbe wie oben, mit dem einzigen Unterschied, dass man Ax als einen Spaltenvektor anzusehen hat.

Feststellung 10.

Die Abbildung $f_A : K^n \rightarrow K^m$ ist eine lineare Abbildung.

Beweis. Es sei $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ und $\lambda \in K$; dann gilt:

$$\begin{aligned} f_A(x+y) &= \left(\sum_{j=1}^n a_{1j}(x_j+y_j), \dots, \sum_{j=1}^n a_{mj}(x_j+y_j) \right) \\ &= \left(\sum_{j=1}^n a_{1j}x_j + \sum_{j=1}^n a_{1j}y_j, \dots, \sum_{j=1}^n a_{mj}x_j + \sum_{j=1}^n a_{mj}y_j \right) \\ &= \left(\sum_{j=1}^n a_{1j}x_j, \dots, \sum_{j=1}^n a_{mj}x_j \right) + \left(\sum_{j=1}^n a_{1j}y_j, \dots, \sum_{j=1}^n a_{mj}y_j \right) \\ &= f_A(x) + f_A(y); \end{aligned}$$

die Gleichung $f_A(\lambda x) = \lambda f_A(x)$ zeigt man ähnlich. \square

Bemerkung: Man hätte die Linearität von f_A auch unter Berufung auf die uns bereits bekannten Matrixgleichungen

$$\begin{aligned} A(B_1 + B_2) &= AB_1 + AB_2 \\ A(\lambda B) &= \lambda(AB) \end{aligned}$$

zeigen können; aus diesen Gleichungen folgt die Linearität von f_A nämlich unmittelbar, wenn man speziell

$$B_1 = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, B_2 = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ und } B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

wählt.

Wir fragen: Wohin bildet f_A die Einheitsvektoren

$$\begin{aligned} e_1 &= (1, 0, \dots, 0) \\ &\vdots \\ e_i &= (0, \dots, 0, 1, 0, \dots, 0) \\ &\vdots \quad \uparrow \\ &\quad i\text{-te Stelle} \\ e_n &= (0, \dots, 0, 1) \end{aligned}$$

ab?

Beispiel. $A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 3 & 0 \\ 1 & 1 & -1 & -1 \end{pmatrix}$, $f_A : \mathbb{R}^4 \rightarrow \mathbb{R}^3$, $e_1 = (1, 0, 0, 0)$, $e_2 = (0, 1, 0, 0)$, $e_3 = (0, 0, 1, 0)$, $e_4 = (0, 0, 0, 1)$.

Dann gilt $f_A(e_1) = (1 \cdot 1 + 0 \cdot 0 + 2 \cdot 0 + 1 \cdot 0, -1 \cdot 1 + 2 \cdot 0 + 3 \cdot 0 + 0 \cdot 0, 1 \cdot 1 + 1 \cdot 0 + (-1) \cdot 0 + (-1) \cdot 0) = (1, -1, 1)$; das Bild des ersten Einheitsvektors ist also gerade die erste Spalte von A .

Entsprechend findet man für $i = 2, 3, 4$: Das Bild des i -ten Einheitsvektors ist die i -te Spalte. (Prüfen Sie dies zumindest für $i = 2$ nach!)

Allgemein gilt:

Feststellung 11.

Die Abbildung $f_A : K^n \rightarrow K^m$ bildet den j -ten Einheitsvektor e_j auf die j -te Spalte von A ab ($j = 1, \dots, n$).

Beweis. Es sei $A = (a_{ij}) \in M(m \times n, K)$. Dann gilt

$$f_A(x) = f_A(x_1, \dots, x_n) = \left(\sum_{i=1}^n a_{1i}x_i, \dots, \sum_{i=1}^n a_{mi}x_i \right).$$

Ist nun speziell $x = e_j$, d.h., es gilt $x_j = 1$ und $x_i = 0$ für alle $i \neq j$, so folgt

$$\sum_{i=1}^n a_{1i}x_i = a_{1j}, \dots, \sum_{i=1}^n a_{mi}x_i = a_{mj}.$$

Also: $f_A(e_j) = (a_{1j}, \dots, a_{mj})$, d.h., das Bild von e_j ist die j -te Spalte von A . \square

Die Bilder $f_A(e_1), \dots, f_A(e_n)$ lassen sich also unmittelbar an der Matrix A ablesen, *es sind gerade die Spalten von A* . Wir merken uns (vgl. Jänich Seite 92):

Die Spalten sind die Bilder der Einheitsvektoren.

Dass zu jeder Matrix $A \in M(m \times n, K)$ eine lineare Abbildung $f_A : K^n \rightarrow K^m$ gehört, haben wir nun ausführlich besprochen.

Wir fragen jetzt umgekehrt: Wenn man *irgendeine* lineare Abbildung

$$f : K^n \rightarrow K^m$$

vorliegen hat, gibt es dann immer eine Matrix $A \in M(m \times n, K)$, so dass

$$f(x) = f_A(x) \text{ für alle } x \in K^n$$

gilt?

Die Antwort wird ja sein! *Wie müsste eine solche Matrix A aussehen?* Um dies zu beantworten, braucht man nur die Bilder $f(e_j)$ der Einheitsvektoren e_j zu betrachten: Soll $f(x) = f_A(x)$ für alle $x \in K^n$ gelten, so muss insbesondere $f(e_j) = f_A(e_j)$ für $j = 1, \dots, n$ gelten, d.h., die Bilder $f(e_j)$ müssen die Spalten von A sein.

Sei A also die Matrix, deren Spalten gerade die Vektoren $f(e_1), \dots, f(e_n)$ sind. Dann gilt also

$$f_A(e_j) = f(e_j) \quad (j = 1, \dots, n),$$

d.h., für die Basisvektoren e_1, \dots, e_n stimmen die beiden linearen Abbildungen f_A und f überein. Nach Satz 1 (Abschnitt 9.1) stimmen sie dann aber auch ganz und gar überein, d.h., es gilt $f_A(x) = f(x)$ für alle $x \in K^n$.

Wir halten unser letztes Ergebnis noch einmal ausdrücklich fest.

Feststellung 12.

Ist $f : K^n \rightarrow K^m$ eine lineare Abbildung, so gibt es genau eine Matrix $A \in M(m \times n, K)$, so dass

$$f(x) = f_A(x) \text{ für alle } x \in K^n.$$

Zwischenbemerkung. Matrizen spielen nicht nur eine Rolle, wenn (wie bisher in diesem Abschnitt) lineare Abbildungen

$$f : K^n \rightarrow K^m$$

betrachtet werden, sondern auch, wenn es um lineare Abbildungen

$$f : V \rightarrow W$$

zwischen *beliebigen endlichdimensionalen Vektorräumen* V und W geht (d.h., es muss sich nicht unbedingt um K^n und K^m handeln). Auch in diesem allgemeineren Fall kann man f mithilfe einer Matrix A beschreiben. Ein wesentlicher Unterschied zum Bisherigen: *Man muss zunächst eine Basis* (v_1, \dots, v_n) *von* V *und eine Basis* (w_1, \dots, w_m) *von* W *fest wählen*. Die zu f gehörige Matrix A ist dann von dieser Wahl der Basen abhängig. *Weitere Einzelheiten hierzu findet man im Jänich auf den Seiten 73f.*

Nach dieser Zwischenbemerkung zu linearen Abbildungen zwischen beliebigen endlichdimensionalen Vektorräumen V und W befassen wir uns wieder mit linearen Abbildungen vom Typ $K^n \rightarrow K^m$.

Zur Erinnerung: Mit $M(m \times n, K)$ hatten wir die Menge der $m \times n$ -Matrizen über K und mit $\text{Hom}(K^n, K^m)$ die Menge der linearen Abbildungen $K^n \rightarrow K^m$ bezeichnet. Wir betrachten nun die Abbildung

$$\begin{aligned} M(m \times n, K) &\mapsto \text{Hom}(K^n, K^m) \\ A &\mapsto f_A, \end{aligned} \tag{*}$$

die also jeder Matrix $A \in M(m \times n, K)$ die zugehörige lineare Abbildung $f_A \in \text{Hom}(K^n, K^m)$ zuordnet. *Diese Abbildung von* $M(m \times n, K)$ *nach* $\text{Hom}(K^n, K^m)$ *ist bijektiv, da Folgendes gilt:*

- (i) Zu zwei verschiedenen Matrizen $A, B \in M(m \times n, K)$ gehören immer zwei verschiedene lineare Abbildungen f_A und f_B . (Wieso eigentlich? Man denke an die Aussage „Die Spalten sind die Bilder der Einheitsvektoren“!)
- (ii) Es gibt zu jedem $f \in \text{Hom}(K^n, K^m)$ eine Matrix A mit $f(x) = f_A(x)$ für alle $x \in K^n$ (siehe Feststellung 12).

Aussage (i) bedeutet, dass die Abbildung (*) injektiv ist, (ii) bedeutet, dass sie surjektiv ist. *Also haben wir wie behauptet gefunden, dass die durch (*) gegebene Abbildung bijektiv ist.*

Wir haben es bei (*) aber nicht nur mit einer Bijektion zwischen $M(m \times n, K)$ und $\text{Hom}(K^n, K^m)$ zu tun, sondern **es gilt noch mehr**, wie wir im Folgenden ausführen.

- 1) Bekanntlich kann man Matrizen $A, B \in M(m \times n, K)$ addieren und erhält als Summe $A + B$ wieder eine Matrix aus $M(m \times n, K)$. Ebenso kann man die beiden Abbildungen f_A und f_B addieren und erhält als Summe $f_A + f_B$ wieder eine Abbildung aus $\text{Hom}(K^n, K^m)$ (vgl. Feststellung 3, Abschnitt 9.1). Es ist nun nicht schwer, sich zu überlegen, dass die durch (*) gegebene Bijektion *sogar verträglich* mit den Additionen in $M(m \times n, K)$ und $\text{Hom}(K^n, K^m)$ ist, dass also für alle $A, B \in M(m \times n, K)$ gilt

$$f_{A+B} = f_A + f_B. \tag{9.1}$$

Nachweis von (9.1): Es ist zu zeigen, dass für alle $x \in K^n$ gilt: $f_{A+B}(x) = f_A(x) + f_B(x)$. Schreiben wir x als Spaltenvektor, so gilt $f_{A+B}(x) = (A+B)x$, $f_A(x) = Ax$ und $f_B(x) = Bx$. Wir haben also

$$(A+B)(x) = Ax + Bx$$

zu zeigen; aufgrund der Rechenregeln für Matrizen (vgl. Gramlich oder Jänich) ist das jedoch richtig. \square

- 2) Auf ähnliche Weise zeigt man, dass

$$f_{\lambda A} = \lambda f_A \tag{9.2}$$

gilt.

- 3) Es sei nun $A \in M(m \times n, K)$ und $B \in M(n \times \ell, K)$; dann gilt $AB \in M(m \times \ell, K)$. Für die zugehörigen linearen Abbildungen gilt somit $f_A \in \text{Hom}(K^n, K^m)$, $f_B \in \text{Hom}(K^\ell, K^n)$ und $f_{AB} \in \text{Hom}(K^\ell, K^m)$. Führt man f_B und f_A nacheinander aus, so erhält man $f_A \circ f_B$ und nach Feststellung 1 (Abschnitt 9.1) gilt $f_A \circ f_B \in \text{Hom}(K^\ell, K^m)$. Wir haben also auf unterschiedliche Art zwei Abbildungen aus $\text{Hom}(K^\ell, K^m)$ erhalten: einerseits f_{AB} und andererseits $f_A \circ f_B$. Wir überlegen uns nun, dass es sich hierbei um dieselbe Abbildung handelt, dass also gilt:

$$f_{AB} = f_A \circ f_B. \tag{9.3}$$

Nachweis von (9.3): Es ist $f_{AB}(x) = f_A(f_B(x))$ für alle $x \in K^\ell$ zu zeigen. Schreiben wir x als Spaltenvektor, so gilt $f_{AB}(x) = (AB)x$ und $f_A(f_B(x)) = A(Bx)$. Aufgrund der Rechenregeln für Matrizen (vgl. Gramlich oder Jänich) gilt:

$$(AB)x = A(Bx),$$

womit (9.3) gezeigt ist. \square

Die Gleichung (9.3) beschreibt einen Zusammenhang, den man auf den ersten Blick möglicherweise nicht vermutet hätte: In Worten ausgedrückt besagt Gleichung (9.3), dass das Matrizenprodukt gerade der Nacheinanderausführung der zugehörigen linearen Abbildungen entspricht.

Bemerkung zur Bezeichnung von linearen Abbildungen im Jänich: Wie wir gesehen haben, besteht ein sehr enger Zusammenhang zwischen den Matrizen aus $M(m \times n, K)$ und den linearen Abbildungen aus $\text{Hom}(K^n, K^m)$. Dies hat Jänich in seinem Buch „Lineare Algebra“ dazu veranlasst, die Abbildung f_A genauso wie die zugehörige Matrix zu bezeichnen, nämlich ebenfalls mit A . Jänich schreibt also immer einfach A , wo wir f_A schreiben würden.

9.3 Der Rang einer Matrix

Es sei A eine $m \times n$ -Matrix über K , also $A \in M(m \times n, K)$. Es gibt verschiedene Möglichkeiten, auszudrücken, was der *Rang von A* (abgekürzt: $\text{rg } A$) ist. Diese Möglichkeiten sollen im Folgenden vorgestellt werden; auf Beweise, dass diese verschiedenen Möglichkeiten tatsächlich alle auf dasselbe hinauslaufen, soll dabei teilweise verzichtet werden.

- (I) Die vielleicht *naheliegendste Möglichkeit* ist die folgende: Wie wir wissen, gehört zu A eine lineare Abbildung

$$f_A : K^n \rightarrow K^m,$$

für die, wenn man $x \in K^n$ als Spaltenvektor schreibt, $f_A(x) = Ax$ gilt. Was der Rang der Abbildung f_A ist, haben wir bereits definiert: Es ist die Dimension des Bildes von f_A . In Zeichen:

$$\text{rg } f_A = \dim \text{Bild } f_A.$$

Es ist also naheliegend, den Rang von A als $\text{rg } f_A$ zu definieren; in Zeichen:

$$\text{rg } A := \text{rg } f_A (= \dim \text{Bild } f_A).$$

Nach dieser Definition ist also der Rang $\text{rg } A$ einer Matrix A die *Dimension des Bildes* der zugehörigen linearen Abbildung.

- (II) Wir definieren nun, was der *Spaltenrang* der Matrix A ist: Darunter versteht man die *Maximalzahl linear unabhängiger Spalten* von A .

An einem Beispiel erklärt: A habe n Spalten s_1, \dots, s_n ; wenn man unter diesen n Spalten drei finden kann (etwa s_2, s_3 und s_5), die linear unabhängig sind, es aber unmöglich ist, vier linear unabhängige Spalten zu finden, so ist die Maximalzahl linear unabhängiger Spalten von A , also der Spaltenrang, gleich 3.

Oder (ähnliches Beispiel): Unter den Spalten von A kann man sieben finden, die linear unabhängig sind, aber acht linear unabhängige Spalten von A gibt es nicht. Dann ist der Spaltenrang von A gleich 7.

Man kann den Spaltenrang von A auch noch unter einem etwas anderen Blickwinkel betrachten: Die Spalten s_1, \dots, s_n der $m \times n$ -Matrix A sind Vektoren aus K^m und als solche spannen sie in K^m den Untervektorraum $\text{Lin}(s_1, \dots, s_n)$ auf. Es gilt nun

$$\text{Spaltenrang von } A = \dim \text{Lin}(s_1, \dots, s_n).$$

Beweis. Da (s_1, \dots, s_n) ein Erzeugendensystem des Vektorraums $\text{Lin}(s_1, \dots, s_n)$ ist, lässt sich durch Auswahl geeigneter Vektoren aus (s_1, \dots, s_n) eine Basis von $\text{Lin}(s_1, \dots, s_n)$ gewinnen (vgl. Abschnitt 8.6). Eine solche Basis besteht aus $\dim \text{Lin}(s_1, \dots, s_n)$ linear unabhängigen Spalten von A und mehr als $\dim \text{Lin}(s_1, \dots, s_n)$ linear unabhängige Spalten von A kann es nicht geben (vgl. Satz 3 in Abschnitt 8.6). Es folgt die Behauptung. \square

(III) Analog zum Spaltenrang definiert man den *Zeilenrang* von A : Dies ist die *Maximalzahl linear unabhängiger Zeilen* von A .

Für den von den Zeilen z_1, \dots, z_m aufgespannten Untervektorraum $\text{Lin}(z_1, \dots, z_m)$ von K^n gilt (analog zu (II)):

$$\text{Zeilenrang von } A = \dim \text{Lin}(z_1, \dots, z_m).$$

Wir wollen uns nun zunächst den Zusammenhang zwischen (I) und (II) anschauen: Sowohl in (I) als auch in (II) hatten wir es mit einem Untervektorraum von K^m zu tun, nämlich einerseits mit $\text{Bild } f_A$ und andererseits mit dem von den Spalten von A aufgespannten Vektorraum $\text{Lin}(s_1, \dots, s_n)$. Wir zeigen nun, dass beide Untervektorräume gleich sind.

Feststellung 13.

$$\text{Lin}(s_1, \dots, s_n) = \text{Bild } f_A.$$

In Worten ausgedrückt gilt also: *Die Spalten von A spannen das Bild von f_A auf.*

Beweis. Die Spalten selbst sind in $\text{Bild } f_A$ enthalten, denn sie sind ja gerade die Bilder der Einheitsvektoren; also $s_j = f_A(e_j) \in \text{Bild } f_A$ ($j = 1, \dots, n$).

Es sei nun $y \in \text{Bild } f_A$. Dann existiert also ein $x \in K^n$ mit $f_A(x) = y$. Da die Einheitsvektoren e_1, \dots, e_n eine Basis von K^n bilden („die kanonische Basis“), lässt sich x als Linearkombination $x = \lambda_1 e_1 + \dots + \lambda_n e_n$ schreiben. Es folgt unter Benutzung der Tatsache, dass f_A linear ist:

$$\begin{aligned} y = f_A(x) &= f_A(\lambda_1 e_1 + \dots + \lambda_n e_n) \\ &= \lambda_1 f_A(e_1) + \dots + \lambda_n f_A(e_n) \\ &= \lambda_1 s_1 + \dots + \lambda_n s_n. \end{aligned}$$

Also: Jedes $y \in \text{Bild } f_A$ ist eine Linearkombination der Spalten s_1, \dots, s_n .

Dies zeigt: $\text{Bild } f_A \subseteq \text{Lin}(s_1, \dots, s_n)$.

Dass auch umgekehrt $\text{Lin}(s_1, \dots, s_n) \subseteq \text{Bild } f_A$ gilt, folgt direkt aus der am Anfang dieses Beweises gemachten Feststellung, dass $s_j \in \text{Bild } f_A$ gilt ($j = 1, \dots, n$). Dann ist nämlich auch jede Linearkombination von (s_1, \dots, s_n) in $\text{Bild } f_A$ enthalten, da $\text{Bild } f_A$ ein Untervektorraum von K^m ist. \square

Aus Feststellung 13 folgt, dass $\text{rg } A$ gleich dem Spaltenrang von A ist, denn es gilt ja

$$\begin{aligned} \text{rg } A &= \dim \text{Bild } f_A \\ &= \dim \text{Lin}(s_1, \dots, s_n) \\ &= \text{Spaltenrang von } A. \end{aligned}$$

Für einen Beweis des folgenden Satzes verweisen wir auf Jänich (oder andere Lehrbücher der Linearen Algebra).

Satz 4.

$$\text{Zeilenrang } A = \text{Spaltenrang } A.$$

Zusammenfassung: Es gilt also

$$\begin{aligned} \text{rg } A &= \text{Spaltenrang von } A \\ &= \text{Zeilenrang von } A \end{aligned}$$

und deshalb spricht man auch meistens nur von dem *Rang von A*. Was der Rang von A ist, lässt sich aufgrund der Ergebnisse dieses Abschnitts auf vielfältige Art ausdrücken:

- rg A = Dimension des Bildes von f_A
- = Maximalzahl linear unabhängiger Spalten von A
- = Maximalzahl linear unabhängiger Zeilen von A
- = Dimension des von den Spalten von A aufgespannten Unterraums von K^m
- = Dimension des von den Zeilen von A aufgespannten Unterraums von K^n

Die **Bestimmung des Rangs einer Matrix** ist einfach: Ist A eine $m \times n$ -Matrix über einem Körper K und sind $z_1, \dots, z_m \in K^n$ die Zeilen von A , so können wir mit dem Gauß-Verfahren eine *Basis des Zeilenraums* von A (so nennt man den von z_1, \dots, z_m aufgespannten Untervektorraum $\text{Lin}(z_1, \dots, z_m)$ von K^n) bestimmen (vgl. Bemerkungen am Ende von Abschnitt 8.6); die Länge der gefundenen Basis ist gleich $\dim \text{Lin}(z_1, \dots, z_m)$. Wegen

$$\text{rg } A = \dim \text{Lin}(z_1, \dots, z_m)$$

hat man damit $\text{rg } A$ bestimmt.

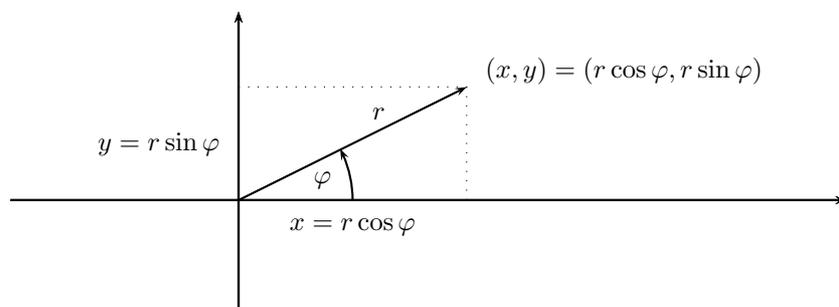
9.4 Einige Beispiele

Die folgenden Beispiele sollen den Begriff der linearen Abbildung etwas lebendiger machen. Wir betrachten zunächst nur lineare Abbildungen $f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die wir durch Angabe der zugehörigen Matrix A definieren.

1. $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Die zugehörige lineare Abbildung ist die *Identität* auf \mathbb{R}^2 : Jeder Vektor $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ wird auf sich selbst abgebildet.
2. $A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$. Es gilt $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}$. Jeder Vektor wird also auf sein λ -faches abgebildet. Ist $\lambda > 1$, so handelt es sich um eine *Streckung* mit dem Faktor λ . (Man gebe ähnliche Beschreibungen für die Fälle $0 < \lambda < 1$, $\lambda = 0$ und $\lambda < 0$.)
3. $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Es gilt $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$. Es handelt sich um die *Spiegelung* an der x -Achse.
4. $A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ für $0 \leq \alpha \leq 2\pi$. Es handelt sich um eine *Drehung* um den Winkel α (entgegen dem Uhrzeigersinn und mit dem Ursprung als Drehzentrum).

Um dies einzusehen, verwenden wir Polarkoordinaten: Es sei (siehe Skizze)

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}.$$



Es folgt

$$\begin{aligned} A \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \\ &= \begin{pmatrix} r(\cos \alpha \cdot \cos \varphi - \sin \alpha \cdot \sin \varphi) \\ r(\sin \alpha \cdot \cos \varphi + \cos \alpha \cdot \sin \varphi) \end{pmatrix} \\ &\stackrel{(*)}{=} \begin{pmatrix} r \cos(\varphi + \alpha) \\ r \sin(\varphi + \alpha) \end{pmatrix}. \end{aligned}$$

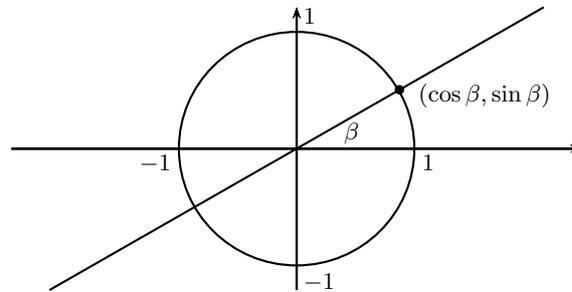
A bewirkt also eine Drehung der beschriebenen Art um den Winkel α .

Die Gleichheit (*) gilt aufgrund der bekannten *Additionstheoreme* für \sin und \cos , welche besagen (vgl. auch Abschnitt 8), dass für alle $x_1, x_2 \in \mathbb{R}$ gilt:

$$\sin(x_1 + x_2) = \sin x_1 \cos x_2 + \cos x_1 \sin x_2,$$

$$\cos(x_1 + x_2) = \cos x_1 \cos x_2 - \sin x_1 \sin x_2.$$

5. Mit $s_\beta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ wollen wir die Spiegelung an der Geraden durch die Punkte $(0,0)$ und $(\cos \beta, \sin \beta)$ bezeichnen (siehe Zeichnung). Bei s_β handelt es sich um eine lineare Abbildung².



Wir ändern nun die Matrix A aus Beispiel 4 etwas ab und betrachten stattdessen die Matrix

$$B = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}.$$

Während A eine Drehung um den Winkel α beschreibt, wird durch B eine Spiegelung beschrieben; genauer handelt es sich um die Spiegelung s_β mit $\beta = \frac{\alpha}{2}$.

Um einzusehen, dass dies tatsächlich so ist, genügt es nachzuweisen, dass die linearen Abbildungen $s_{\alpha/2}$ und f_B auf den Einheitsvektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

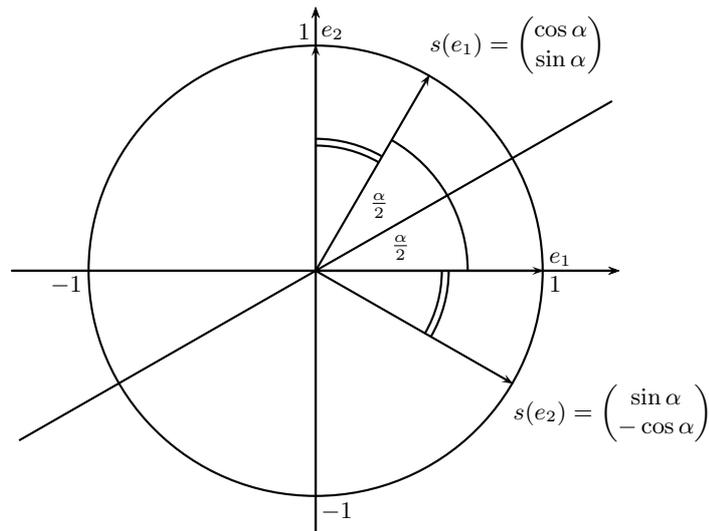
übereinstimmen. Es gilt:

$$f_B(e_1) = B e_1 = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \quad \text{und} \quad f_B(e_2) = B e_2 = \begin{pmatrix} \sin \alpha \\ -\cos \alpha \end{pmatrix}.$$

Entsprechendes gilt für $s = s_{\alpha/2}$ (siehe Zeichnung):

$$s(e_1) = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \quad \text{und} \quad s(e_2) = \begin{pmatrix} \sin \alpha \\ -\cos \alpha \end{pmatrix}.$$

²Man mache sich das anhand einer Zeichnung klar!



Die beiden linearen Abbildungen f_B und $s_{\alpha/2}$ stimmen also auf einer Basis des \mathbb{R}^2 überein. Nach Satz 1 stimmen f_B und $s_{\alpha/2}$ somit auf ganz \mathbb{R}^2 überein, d.h., bei f_B handelt es sich wie behauptet um die Spiegelung $s_{\alpha/2}$.

Aufgaben zu den Beispielen 4 und 5:

1. Berechnen Sie $\det(A)$ und $\det(B)$.
2. Berechnen Sie das Matrizenprodukt $B^2 = BB$. Entspricht das Ergebnis Ihren Erwartungen?

Mehr über Drehungen und Spiegelungen findet man beispielsweise in Jänich, Lineare Algebra.

Ein weiteres **Beispiel:**

6. $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Es gilt $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$. Es handelt sich um eine *Projektion*; genauer: Jeder Punkt des \mathbb{R}^2 wird in Richtung der y -Achse auf die x -Achse projiziert.

Abschließend sei noch eine lineare Abbildung $f_A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ betrachtet:

7. Es sei für $0 \leq \alpha \leq 2\pi$: $A = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Ähnlich wie in Beispiel 4 erkennt man, dass es sich bei der zugehörigen linearen Abbildung $f_A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ um eine Drehung handelt (α ist der Drehwinkel, Drehachse ist die z -Achse).

9.5 Eigenwerte und Eigenvektoren

Eine $n \times n$ -Matrix $A \in M(n \times n, K)$ kann man sich - wie wir wissen - auch als eine Abbildung vorstellen, die jedem Vektor $x \in K^n$ den Vektor Ax zuordnet, der ebenfalls im Raum K^n liegt³:

$$x \mapsto Ax.$$

Dabei ist zu beachten: Ax ist ein Matrizenprodukt („ $n \times n$ -Matrix mal $n \times 1$ -Matrix“) und deshalb ist x an dieser Stelle als ein *Spaltenvektor* aufzufassen, während man ansonsten aus Platzgründen meist die Zeilenschreibweise bevorzugt.

³Diese Abbildung hatten wir immer mit f_A bezeichnet. Wir wissen: f_A ist eine lineare Abbildung.

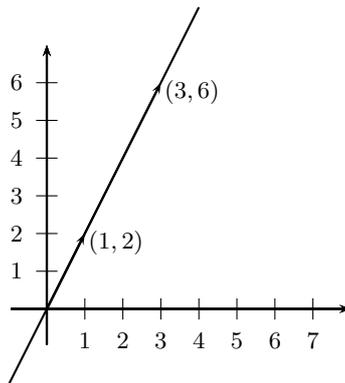
Beispiel. Es sei $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \in M(2 \times 2, \mathbb{R})$ und $x = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Dann ist

$$Ax = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

Für diese Matrix A berechnen wir außerdem noch die Bilder der Vektoren $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$:

$$A \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} -1 \\ -16 \end{pmatrix}, \quad A \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} \quad \text{und} \quad A \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}.$$

Für den Vektor $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ lässt sich *eine Besonderheit* beobachten: $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ wird auf ein skalares Vielfaches von sich selbst abgebildet.



Für die anderen betrachteten Vektoren $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ gilt Derartiges nicht.

Definition.

Es sei $A \in (n \times n, K)$. Unter einem *Eigenvektor* von A zum *Eigenwert* $\lambda \in K$ versteht man einen Vektor $x \neq 0$ auf K^n mit der folgenden Eigenschaft:

$$Ax = \lambda x$$

In obigem Beispiel ist also der Vektor $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ ein Eigenvektor von A zum Eigenwert $\lambda = 3$. Es stellt sich die Frage: *Besitzt A noch anderen Eigenwerte außer $\lambda = 3$?*

Allgemeiner fragen wir: *Wie berechnet man Eigenwerte und Eigenvektoren?*

Wir wollen uns zunächst um die *Berechnung der Eigenwerte* kümmern. Ist λ ein Eigenwert einer Matrix $A \in M(n \times n, K)$, so gibt es einen Vektor $x \neq 0$ aus K^n , für den

$$Ax = \lambda x$$

gilt. Um herauszufinden, wie man die Eigenwerte λ berechnen kann, formen wir diese Gleichung (etwas trickreich) um: Zunächst einmal beobachten wir, dass

$$(\lambda E_n)x = \lambda x$$

gilt. (E_n bezeichnet die n -reihige Einheitsmatrix.) Setzt man dies in die obige Gleichung ein, so erhält man

$$Ax = (\lambda E_n)x.$$

Äquivalente Umformung ergibt sodann

$$(A - \lambda E_n)x = 0.$$

Es folgt: λ ist genau dann ein Eigenwert von A , wenn das homogene lineare Gleichungssystem mit der Koeffizientenmatrix $A - \lambda E_n$ eine Lösung $x \neq 0$ („nicht triviale Lösung“) besitzt. Wir wissen (siehe Gramlich, Satz 5.9)⁴: Dies ist genau dann der Fall, wenn

$$\det(A - \lambda E_n) = 0$$

gilt. Man nennt diese Gleichung die *charakteristische Gleichung von A* . Der Ausdruck $\det(A - \lambda E_n)$ ist ein Polynom vom Grad n mit der Unbestimmten λ . Man nennt dieses Polynom das *charakteristische Polynom von A* . Nun wissen wir, was zu tun ist, um die Eigenwerte $\lambda \in K$ der Matrix A zu bestimmen:

Man muss die Nullstellen $\lambda \in K$ des charakteristischen Polynoms von A berechnen.

Wir greifen unser altes Beispiel wieder auf. Es sei $A \in M(2 \times 2, \mathbb{R})$ mit $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$. Die Matrix $A - \lambda E_2$ lautet dann:

$$A - \lambda E_2 = \begin{pmatrix} 1 - \lambda & 1 \\ -2 & 4 - \lambda \end{pmatrix}.$$

Das charakteristische Polynom lautet folglich:

$$\det(A - \lambda E_2) = (1 - \lambda)(4 - \lambda) - 1 \cdot (-2) = \lambda^2 - 5\lambda + 6.$$

Nullstellen des charakteristischen Polynoms:

$$\begin{aligned} \lambda_{1/2} &= \frac{5}{2} \pm \sqrt{\frac{25}{4} - 6} = \frac{5}{2} \pm \frac{1}{2} \\ \implies \lambda_1 &= 2, \quad \lambda_2 = 3 \end{aligned}$$

Die Matrix A besitzt also genau zwei Eigenwerte, nämlich 2 und 3.

Aufgabe: Berechnen Sie die Eigenwerte der Matrix $A = \begin{pmatrix} 1 & -4 \\ -2 & 3 \end{pmatrix} \in M(2 \times 2, \mathbb{R})$.

Ein Polynom vom Grad n kann niemals mehr als n verschiedene Nullstellen haben, deshalb gilt: *Eine $n \times n$ -Matrix A besitzt höchstens n verschiedene Eigenwerte.*

Hat man die Eigenwerte einer Matrix gefunden, so ist es nicht schwer, zu jedem Eigenwert λ die zugehörigen Eigenvektoren zu bestimmen: Man muss nur das homogene lineare Gleichungssystem

$$(A - \lambda E_n)x = 0$$

lösen. (Man beachte: $Ax = \lambda x$ ist äquivalent zu $(A - \lambda E_n)x = 0$.) *Die Menge aller nicht trivialen Lösungen dieses Gleichungssystems ist dann die Menge aller Eigenvektoren zum Eigenwert λ .*

Beispiel. Wie zuvor sei $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$. Wir wissen bereits: $\lambda_1 = 2$ und $\lambda_2 = 3$ sind die Eigenwerte von A . Es gilt

$$A - \lambda E_2 = \begin{pmatrix} 1 - \lambda & 1 \\ -2 & 4 - \lambda \end{pmatrix}.$$

⁴Satz 5.9 wird im Gramlich nur für $K = \mathbb{R}$ behandelt; die entsprechende Aussage gilt jedoch auch für einen beliebigen Körper K (siehe beispielsweise Jänich).

Um die *Eigenvektoren zum Eigenwert 2* zu bestimmen, setzen wir $\lambda = 2$ in die Matrix $A - \lambda E_2$ ein und lösen das zugehörige homogene lineare Gleichungssystem $(A - \lambda E_2)x = 0$; es lautet in expliziter Form:

$$\begin{aligned} -x_1 + x_2 &= 0 \\ -2x_1 + 2x_2 &= 0. \end{aligned}$$

Diese beiden Gleichungen sind äquivalent zu der einen Gleichung $x_1 = x_2$. Die Lösungsmenge des Gleichungssystem ist demnach eine Ursprungsgerade des \mathbb{R}^2 , nämlich

$$\left\{ \mu (1, 1) : \mu \in \mathbb{R} \right\}.$$

Man nennt diese Lösungsmenge den *Eigenraum zum Eigenwert $\lambda = 2$* .

Die Menge aller Eigenvektoren zum Eigenwert 2 ist demnach die Menge aller Vektoren $\mu(1, 1) \in \mathbb{R}^2$ mit $\mu \in \mathbb{R}$ und $\mu \neq 0$. (*Der Nullvektor muss ausgenommen werden, da er laut Definition kein Eigenvektor ist.*)

Eigenvektoren zum Eigenwert $\lambda = 3$: In diesem Fall lautet das Gleichungssystem

$$\begin{aligned} -2x_1 + x_2 &= 0 \\ -2x_1 + x_2 &= 0. \end{aligned}$$

Lösungsmenge („Eigenraum zum Eigenwert 3“):

$$\left\{ \mu \left(\frac{1}{2}, 1 \right) : \mu \in \mathbb{R} \right\}.$$

Die Menge aller Eigenvektoren zum Eigenwert 3 ist also die Menge aller Vektoren $\mu \left(\frac{1}{2}, 1 \right) \in \mathbb{R}^2$ mit $\mu \in \mathbb{R} \setminus \{0\}$.

Aufgabe: Bestimmen Sie zu jedem Eigenvektor der Matrix $A = \begin{pmatrix} 1 & -4 \\ -2 & 3 \end{pmatrix}$ den zugehörigen Eigenraum.

Beispiel. Es gelte $K = \mathbb{R}$ oder $K = \mathbb{C}$. Wir betrachten die Matrix $A \in M(2 \times 2, K)$ mit

$$A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

und fragen nach den Eigenwerten von A .

1. Fall: $K = \mathbb{R}$

Berechnung des charakteristischen Polynoms von A :

$$|A - \lambda E_2| = \begin{vmatrix} 1 - \lambda & -2 \\ 2 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 + 4 = \lambda^2 - 2\lambda + 5.$$

Nullstellen von $\lambda^2 - 2\lambda + 5$: $\lambda_1 = 1 + 2i$ und $\lambda_2 = 1 - 2i$.

Also: Für $K = \mathbb{R}$ besitzt A keine Eigenwerte.

2. Fall: $K = \mathbb{C}$

Wir sehen A jetzt also als eine Matrix über dem Körper \mathbb{C} an. Das charakteristische Polynom und seine Nullstellen ändern sich dadurch nicht, aber $\lambda_1 = 1 + 2i$ und $\lambda_2 = 1 - 2i$ sind nun Eigenwerte von A .

Bestimmung der Eigenvektoren zum Eigenwert $\lambda_1 = 1 + 2i$. Der zu λ_1 gehörige Eigenraum ist gleich der Lösungsmenge des folgenden linearen Gleichungssystems:

$$\begin{aligned} -2ix_1 - 2x_2 &= 0 \\ 2x_1 - 2ix_2 &= 0. \end{aligned}$$

Man beachte, dass die zweite Gleichung aus der ersten durch Multiplikation mit i hervorgeht. Die beiden Gleichungen sind somit äquivalent zu der einen Gleichung $x_1 - ix_2 = 0$. Als Lösungsmenge („Eigenraum zum Eigenwert $\lambda_1 = 1 + 2i$ “) erhält man

$$\{\mu(i, 1) : \mu \in \mathbb{C}\}.$$

Die Eigenvektoren zum Eigenwert λ_1 sind also genau die Vektoren $\mu(i, 1)$ mit $\mu \neq 0$.

Ähnlich erhält man, dass die Eigenvektoren zum Eigenwert $\lambda_2 = 1 - 2i$ genau die Vektoren $\mu(-i, 1)$ mit $\mu \neq 0$ sind.

10 Literatur

- E. Behrends:
Analysis, Band 1 und 2. Teubner. 2004.
- R. Courant:
Vorlesungen über Differential- und Integralrechnung. Springer. 1971.
- W. Dörfler, W. Peschek:
Einführung in die Mathematik für Informatiker. Hanser. 1988.
- H. D. Ebbinghaus et al.:
Zahlen. Springer. 1992.
- K. Endl, W. Luh:
Analysis I, II, III. Aula. 1989.
- G. Fischer:
Lineare Algebra. Vieweg-Verlag. 2014.
- H. Fischer, H. Kaul:
Mathematik für Physiker. Teubner. 2005.
- O. Forster:
Analysis 1 und 2. Vieweg. 2012.
- G. M. Gramlich:
Lineare Algebra. Hanser. 2011.
- K. Habetha:
Höhere Mathematik für Ingenieure und Physiker, Band I. Klett. 1982.
- D. Hachenberger:
Mathematik für Informatiker. Pearson Studium. 2008.
- H. Heuser:
Lehrbuch der Analysis, Teil 1 und 2. Teubner. 2009.
- K. Jänich:
Lineare Algebra. Springer-Verlag. 2013.
- K. Königsberger:
Analysis 1 und 2. Springer. 2004.
- B. Kreußler, G. Pfister:
Mathematik für Informatiker. Springer. 2009.
- W. Luh:
Mathematik für Naturwissenschaftler I und II. Aula. 1987.
- W. Schäfer, K. Georgi, G. Trippler:
Mathematik-Vorkurs. Teubner. 2006.

- G. Teschl, S. Teschl:
Mathematik für Informatiker, Band 1 und 2. Springer. 2008.
- B. L. van der Waerden:
Mathematik für Naturwissenschaftler. B.I.-Taschenbuch. 1975.

Index

- C^1 -Funktion, 142
- C^k -Funktion, 142
- δ -Umgebung, 141
 - offene, 141
- k -te Abschnittsdeterminante, 144
- n -te Einheitswurzel, 174
- n -te Wurzel, 173
- n -te Wurzel einer komplexen Zahl, 173
- n -tes Restglied, 129
- n -tes Taylorpolynom, 129
- q -te Wurzel, 37
- x -projizierbare Menge, 158
- y -projizierbare Menge, 158
- äquidistante Zerlegung, 69

- Abbildung
 - zugehörige, 207
- Abbildung, lineare, 198, 205
- abelsche Gruppe, 178
- abgeschlossenes Intervall, 21
- Ableitung
 - höhere, 32
 - partielle, 138, 139
- Ableitung einer Potenzreihe, 124
- Ableitung von f , 30
- Ableitungen der trigonometrischen Funktionen, 58
- Abschätzung des Restglieds, 130, 131
- Abschnittsdeterminante
 - k -te, 144
- absolut konvergent, 118
- absolute Konvergenz, 118
- absoluter Betrag, 7, 163
- absolutes Maximum, 48
- absolutes Minimum, 48, 141
- Abstand, 9, 164
- Achse
 - imaginäre, 163
 - reelle, 163
- achsenparalleles Rechteck, 154
- Addition, komponentenweise, 178
- Additionstheoreme, 57, 169, 214
- allgemeine Integrationsregeln, 90
- allgemeine Lösung eines Gleichungssystems, 186
- allgemeine Potenzfunktion, 43
- Allquantor, 11

- alternierende Reihe, 120
- angeordneter Körper, 5
- Anordnungsaxiome, 5
- Anpassung der Integrationsgrenzen, 96
- Anschauungsraum, 193
- Anstieg in einer Richtung, 148
- Archimedische Eigenschaft, 6
- Archimedisches Axiom, 6
- Arcuscosinus, 60
- Arcuscotangens, 61
- Arcusfunktion, 62
- Arcussinus, 60
- Arcustangens, 61
- Arcustangensreihe, 126, 136
- Asymptote von f für $x \rightarrow -\infty$, 50
- Asymptote von f für $x \rightarrow \infty$, 50
- Asymptote von f in x_0 , 50
- Automorphismus, 202
- Axiom
 - Archimedisches, 6
 - Vollständigkeits-, 5
- Axiome
 - Anordnungs-, 5
 - Körper-, 5
- Axiome der reellen Zahlen, 5

- Basis, 188
 - kanonische, 189
 - Standard-, 189
- Basisergänzungssatz, 190
- Berechnung des Rangs einer Matrix, 213
- Berechnung von Asymptoten, 51
- Berechnung von Eigenvektoren, 217
- Berechnung von Eigenwerten, 216
- Berechnung zweidimensionaler Integrale, 157
- Bernoullische Ungleichung, 14, 15, 18
- beschränkt, 13
 - nach oben, 13
 - nach unten, 13
- Bestimmen der Dimension, 193
- Bestimmen einer Basis, 193
- bestimmtes Integral, 71, 73
- Betrag, 7
 - absoluter, 7, 163
- Betragsfunktion, 31
- Bild von f , 200, 201

Bilder der Einheitsvektoren, 209
 Binomischer Lehrsatz, 12, 18

 Cauchyscher Konvergenztest, 18
 Cauchysches Konvergenzkriterium, 14
 charakteristische Gleichung, 217
 charakteristisches Polynom, 217
 Nullstellen, 217
 Cosinus, 56
 Cosinusreihe, 135
 Cotangens, 58

 de l'Hospital
 Regeln von, 64
 definit
 negativ, 144
 positiv, 144
 Definitheitskriterium, 144
 Differenzenquotient, 29
 differenzierbar, 30
 partiell, 138, 141
 Differenzierbarkeit, 30
 Dimension des Bildes, 211
 Dimension von V , 191
 Dimensionsformel für lineare Abbildungen, 205
 Dirichletsche Sprungfunktion, 74
 divergent, 10
 Division mit Rest, 98
 Division, Polynom-, 98
 Doppelbezeichnungen, 179
 Drehung, 213, 215
 Dreiecksungleichung, 7, 9, 166
 Durchschnittswert einer Funktion, 77

 Ebene, 197
 Ursprungs-, 187
 Eigenraum, 218
 eigentlich integrierbar, 112
 eigentliches Integral, 112
 Eigenvektor, 215
 Berechnung, 217
 Eigenwert, 215
 Berechnung, 216
 Eindeutigkeitsaussage, 202
 Eindeutigkeitsatz für Potenzreihen, 127
 Einheit
 imaginäre, 162
 Einheitskreis, 56
 Einheitsvektor, 189
 Einheitswurzel, n -te, 174
 Einschließungssatz, 17, 18
 endlichdimensionaler Vektorraum, 191
 endliches Intervall, 21
 Endomorphismus, 202
 Entwicklungspunkt, 128

 Epimorphismus, 202
 Erzeugendensystem, 188
 Eulersche Zahl, 16
 Existenzaussage, 202
 Existenzquantor, 11
 Exponentialfunktion zur Basis a , 39
 Exponentialfunktion, komplexe, 174
 Exponentialreihe, 135
 Extremstelle, 45, 141
 Extremstellen mit Nebenbedingungen, 150
 Extremum
 strenges lokales, 46
 Extremwertkriterium, 146

 Faktor, Linear-, 106
 Faktorzerlegung des Nenners, 99
 Fehlerabschätzung bei der Trapezregel, 110
 Feinheitsmaß, 71, 155
 Flächeninhaltsproblem, 69
 Flächenstück, 157
 Folge, 9
 konstante, 12
 Null-, 12
 Folgenglied, 9
 Folgerung aus dem Fundamentalsatz, 79
 Folgerung aus dem Wurzelkriterium, 120
 formales Rechnen, 87
 Formel
 Moivresche, 170
 Fundamentalsatz der Algebra, 172
 Fundamentalsatz der Differential- und Integral-
 rechnung, 78
 Fundamentalsystem, 189
 Funktion
 allgemeine Potenz-, 43
 Arcus-, 62
 Betrags-, 31
 Durchschnittswert einer, 77
 Exponentialfunktion zur Basis a , 39
 gerade, 57
 identische, 25
 komplexe Exponential-, 174
 Lagrange-, 150
 nichtintegrierbare, 74
 rationale, 26
 Stamm-, 69, 79, 81
 trigonometrische, 58
 Umkehr-, 36
 ungerade, 57
 Ziel-, 150
 zyklometrische, 62
 Funktionsgrenzwert, 22

 Gauß-Algorithmus, 186

geometrische Bedeutung der ersten Ableitung, 45
 geometrische Bedeutung der partiellen Ableitung, 146
 geometrische Bedeutung der zweiten Ableitung, 45
 geometrische Bedeutung des Gradienten, 146
 geometrische Reihe, 20, 117
 Gerade, 194, 196
 Ursprungs-, 186
 gerade Funktion, 57
 Gleichung der Tangente $T(x)$, 32
 Glied einer Folge, 9
 Glieder a_i einer Reihe, 19
 globales Maximum, 48, 141
 globales Minimum, 48, 141
 Größe des maximalen Anstiegs, 149
 Gradient, 141, 145
 Graph, 21
 graphische Darstellung, 21
 Grenze
 obere, 13
 untere, 13
 Grenzwert, 10
 Funktions-, 22
 linksseitiger, 27
 rechtsseitiger, 27
 uneigentlicher, 10
 Gruppe
 abelsche, 178
 kommutative, 178

 höhere Ableitung, 32
 Hülle, lineare, 184
 halboffenes Intervall, 21
 harmonische Reihe, 19, 117
 Hauptsatz der Differential- und Integralrechnung, 79
 Hessesche Matrix, 141, 145
 $\text{Hom}(V, W)$, 200
 Homomorphismus, 198

 identische Funktion, 25
 Identität, 25, 199, 213
 Identitätssatz für Potenzreihen, 127
 imaginäre Achse, 163
 imaginäre Einheit, 162
 Imaginärteil, 163
 indefinit, 144
 Indefinitheitskriterium, 144
 Infimum, 13
 injektiv, 201
 innerer Punkt, 141
 Integral
 bestimmtes, 71, 73
 eigentliches, 112
 iteriertes, 157
 Riemannsches, 76
 unbestimmtes, 81
 uneigentliches, 112, 113
 zweidimensionales, 154
 Integral von f als Funktion der oberen Grenze, 77
 Integralform des Restglieds, 130
 Integrand, 71, 73
 Integration
 numerische, 108
 partielle, 83
 Produkt-, 84
 Integration rationaler Funktionen, 96
 Integration von Funktionen mehrerer Variablen, 154
 Integrationsgrenzen, 73
 Anpassung der, 96
 Integrationsregeln, 90
 Integrationsvariable, 73
 integrierbar, 73
 eigentlich, 112
 nicht elementar, 93, 108
 uneigentlich, 113
 integrierbar auf G , 156
 Interpolation, 104
 Lagrange-, 104
 lineare, 107
 Newton-, 106
 Interpolationsfehler, 104
 Interpolationspolynom
 Lagrangesches, 105
 Newtonsches, 106
 Intervall, 21
 abgeschlossenes, 21
 endliches, 21
 halboffenes, 21
 Konvergenz-, 123
 offenes, 21
 unendliches, 21
 isomorph, 162, 204
 Isomorphismus, 202
 iteriertes Integral, 157

 Körper
 angeordneter, 5
 Körper, zugrundeliegender, 180
 Körperaxiome, 5
 kanonische Basis, 189
 Kern von f , 200
 Kettenregel, 34
 kleinste obere Schranke, 13
 kleinste untere Schranke, 13

Koeffizientenfolge der Potenzreihe, 121
 Koeffizientenvergleich, 100
 kommutative Gruppe, 178
 komplexe Exponentialfunktion, 174
 komplexe Zahlen, 161
 n -te Wurzel, 173
 Polarkoordinatendarstellung, 168
 komplexer Vektorraum, 180
 komponentenweise Addition, 178
 Komposition $g \circ f$, 34
 konjugiert komplexe Zahl, 163
 konkav, 49
 streng, 49
 konstante Folge, 12
 konvergent, 10, 142
 absolut, 118
 uneigentlich, 10
 konvergente Majorante, 117
 konvergente Reihe, 19
 Konvergenz, 9, 11
 absolute, 118
 uneigentliche, 10
 Konvergenzintervall, 123
 Randpunkte, 123
 Konvergenzkriterien für Reihen, 116
 Konvergenzkriterium
 Cauchysches, 14
 Konvergenzradius, 123
 Konvergenztest
 Cauchyscher, 18
 konvex, 49
 streng, 49
 Koordinaten
 Polar-, 168
 Kreiszahl π , 56
 Kriterium
 Definitheits-, 144
 Extremwert-, 146
 Indefinitheits-, 144
 Leibnizsches, 120
 Majoranten-, 117
 Quotienten-, 119
 Unterraum-, 183
 Vergleichs-, 117
 Wurzel-, 120
 kritische Stelle, 142
 kritische Stelle unter Nebenbedingungen, 151
 Kugeloberfläche, 151
 Kurvendiskussion, 45, 50

 Länge, 190
 Lösungsmenge, 186
 Lagrange-Ansatz, 150
 Lagrange-Funktion, 150
 Lagrange-Interpolation, 104
 Lagrange-Optimierung, 150
 Lagrangesche Multiplikatorenregel, 150
 Lagrangesches Interpolationspolynom, 105
 Lagrangesches Polynom, 104
 Lehrsatz
 Binomischer, 12, 18
 Leibniz, 29, 88
 Leibnizsche Schreibweise, 87
 Leibnizsches Kriterium, 120
 linear abhängig, 185
 linear unabhängig, 185
 lineare Abbildung, 198, 205
 Dimensionsformel, 205
 lineare Hülle, 184
 lineare Interpolation, 107
 lineare Substitution, 92
 linearer Operator, 199
 Linearfaktor, 106
 Linearkombination, 184
 nichttriviale, 190
 triviale, 190
 linksseitig stetig, 27
 linksseitige Stetigkeit, 27
 linksseitiger Grenzwert, 27
 logarithmisches Differenzieren, 44
 Logarithmus zur Basis a , 39
 Logarithmus, natürlicher, 40
 Logarithmusreihe, 126, 136
 lokales Extremum unter Nebenbedingungen, 150
 lokales Extremum, strenges, 46
 lokales Maximum, 46, 142
 strenges, 46, 142
 lokales Maximum unter Nebenbedingungen, 150
 lokales Minimum, 46, 142
 strenges, 46, 142
 lokales Minimum unter Nebenbedingungen, 150

 Majorante, konvergente, 117
 Majorantenkriterium, 117
 Spezialfall, 118
 Matrix, 143, 205
 Hessesche, 141, 145
 Rang einer, 211
 Matrizenmultiplikation, 206
 Matrizenprodukt, 211
 maximaler Anstieg
 Größe, 149
 Richtung, 149
 Maximum
 absolutes, 48
 globales, 48, 141
 lokales, 46, 142
 strenges lokales, 46, 142

Menge
 x -projizierbar, 158
 y -projizierbare, 158
 Lösungs-, 186
 Standard-, 158
 Minimum
 absolutes, 48, 141
 globales, 48, 141
 lokales, 46, 142
 strenges lokales, 46, 142
 Mittelpunkt, 128
 Mittelwertsatz der Differentialrechnung, 46
 Mittelwertsatz der Integralrechnung, 77
 Moivresche Formel, 170
 Monomorphismus, 202
 monoton, 14, 35
 stückweise, 73
 streng, 35
 monoton fallend, 14, 35
 streng, 35
 monoton steigend, 14
 monoton wachsend, 14, 35
 streng, 35
 Multiplikation, skalare, 180

 näherungsweise Berechnung von Nullstellen, 62
 Nabla-Operator, 145
 nach oben beschränkt, 13
 nach unten beschränkt, 13
 natürlicher Logarithmus, 40
 negativ definit, 144
 negativ semidefinit, 144
 Negatives, 182
 Newton, 29, 88
 Newton-Interpolation, 106
 Newtonsches Interpolationspolynom, 106
 Newtonsches Verfahren, 62
 nicht elementar integrierbar, 93, 108
 nichtintegrierbare Funktion, 74
 nichttriviale Darstellung des Nullvektors, 190
 nichttriviale Linearkombination, 190
 Nullfolge, 12
 Nullpolynom, 106
 Nullstelle, 50, 106, 172
 näherungsweise Berechnung einer, 62
 Nullstellen des charakteristischen Polynoms, 217
 Nullstellensatz für Polynome, 172
 Nullvektor, 180
 nichttriviale Darstellung, 190
 triviale Darstellung, 190
 numerische Integration, 108

 obere Grenze, 13
 obere Schranke, 13
 kleinste, 13

 Obersumme, 70
 offen, 141
 offene δ -Umgebung, 141
 offenes Intervall, 21
 Operator, 199
 linearer, 199
 Orientierung, 194

 Parallelogramm der Kräfte, 193
 Partialbruchzerlegung, 99
 Partialsumme, 18
 partiell differenzierbar, 138, 141
 partielle Ableitung, 138, 139
 partielle Ableitung 1. Ordnung, 139
 partielle Ableitung 2. Ordnung, 139
 partielle Ableitung höherer Ordnung, 139
 partielle Integration, 83
 partielle Integration für bestimmte Integrale, 93
 periodisch, 175
 Polarkoordinaten, 168, 213
 Polynom, 26
 charakteristisches, 217
 Lagrangesches, 104
 Lagrangesches Interpolations-, 105
 Null-, 106
 Polynomdivision, 98
 Polynome
 Nullstellensatz für Polynome, 172
 Zerlegungssatz für Polynome, 172
 positiv definit, 144
 positiv semidefinit, 144
 Potenz a^x , 37
 Potenzfunktion, allgemeine, 43
 Potenzreihe, 121, 136
 Ableitung einer, 124
 Koeffizientenfolge der, 121
 Stammfunktion einer, 124
 Potenzreihen
 Eindeutigkeitssatz für, 127
 Identitätssatz für, 127
 Produktintegration, 84
 Produktregel, 33
 Projektion, 215
 Punkt, innerer, 141

 quadratische Ergänzung, 97
 quadratische Form, 143
 Quantor
 All-, 11
 Existenz-, 11
 Quotientenkriterium, 119
 Quotientenregel, 33

 Randpunkte des Konvergenzintervalls, 123
 Rang, 205

- Bestimmung, 213
- Spalten-, 211
- Zeilen-, 212
- Rang einer Matrix, 211
- rationale Funktion, 26
 - Integration, 96
- Realteil, 163
- Rechenregeln, 37
- Rechenregeln für konvergente Folgen, 16
- Rechteck, achsenparalleles, 154
- rechtsseitig stetig, 27
- rechtsseitige Stetigkeit, 27
- rechtsseitiger Grenzwert, 27
- reelle Achse, 163
- reelle Zahlen, Axiome der, 5
- reeller Vektorraum, 180
- Regel
 - Ketten-, 34
 - Lagrangesche Multiplikatoren-, 150
 - Produkt-, 33
 - Quotienten-, 33
 - Simpson-, 112
 - Summen, 33
 - summierte Trapez-, 110
 - Trapez-, 108, 110
 - Umkehr-, 36
 - zusammengesetzte Trapez-, 110
- Regeln von de l'Hospital, 64
- Regularitätsbedingung, 150
- Reihe, 18, 116
 - alternierende, 120
 - Arcustangens-, 126, 136
 - Cosinus-, 135
 - Exponential-, 135
 - geometrische, 20, 117
 - Glieder a_i einer Reihe, 19
 - harmonische, 19, 117
 - konvergente, 19
 - Logarithmus-, 126, 136
 - Potenz-, 121, 136
 - Sinus-, 135
 - Taylor-, 134, 136
- Reihen, Konvergenzkriterien für, 116
- rein imaginäre Zahl, 175
- Restglied
 - n -tes, 129
 - Integralform, 130
 - Abschätzung, 130, 131
- Resubstitution, 86
- Richtung, 148
 - Anstieg in einer, 148
- Richtung des größten Anstiegs, 149
- Riemannsche Summe, 72, 154
- Riemannsches Integral, 76
- Sattelpunkt, 143
- Satz
 - Basisergänzungssatz, 190
 - Binomischer Lehr-, 12, 18
 - Eindeutigkeitssatz für Potenzreihen, 127
 - Einschließungs-, 17, 18
 - Fundamentalsatz der Algebra, 172
 - Fundamentalsatz der Differential- und Integralrechnung, 78
 - Hauptsatz der Differential- und Integralrechnung, 79
 - Identitätssatz für Potenzreihen, 127
 - Mittelwertsatz der Differentialrechnung, 46
 - Mittelwertsatz der Integralrechnung, 77
 - Nullstellensatz für Polynome, 172
 - Satz über monotone, beschränkte Folgen, 14, 17
 - Satz vom Koeffizientenvergleich, 127
 - Satz von Schwarz, 140, 145
 - Satz von Taylor, 129
 - Satz von Taylor in allgemeiner Form, 133
 - Zerlegungssatz für Polynome, 172
 - Zwischenwert-, 55
- Schaubild, 21
- Schranke
 - kleinste obere, 13
 - kleinste untere, 13
 - obere, 13
 - untere, 13
- Sekante, 29
- semidefinit
 - negativ, 144
 - positiv, 144
- Simpsonregel, 112
- Sinus, 56
- Sinusreihe, 135
- Skalar, 180
- skalare Multiplikation, 180
- Spaltenrang, 211
- Spezialfall des Majorantenkriteriums, 118
- Spiegelung, 213
- stückweise monoton, 73
- stückweise stetig, 73
- Stützstellen, 104
- Stützwerte, 104
- Stammfunktion, 69, 79, 81
- Stammfunktion einer Potenzreihe, 124
- Stammfunktionsproblem, 69
- Standardbasis, 189
- Standardmenge, 158
- stationäre Stelle, 142
- stationäre Stelle unter Nebenbedingungen, 151
- Steigung, 29
- Steigung des Graphen, 148

Stelle
 kritische, 142
 stationäre, 142
 stetig
 linksseitig, 27
 rechtsseitig, 27
 stückweise, 73
 stetig an der Stelle $x^{(0)}$, 142
 stetig an der Stelle x_0 , 24
 stetig auf X , 24, 142
 Stetigkeit
 linksseitige, 27
 rechtsseitige, 27
 Stetigkeit an der Stelle x_0 , 24
 Stetigkeit auf X , 24
 Streckung, 213
 streng konkav, 49
 streng konvex, 49
 streng monoton, 35
 streng monoton fallend, 35
 streng monoton wachsend, 35
 strenges lokales Extremum, 46
 strenges lokales Maximum, 46, 142
 strenges lokales Maximum unter Nebenbedingungen, 150
 strenges lokales Minimum, 46, 142
 strenges lokales Minimum unter Nebenbedingungen, 150
 Substitution, lineare, 92
 Substitutionsregel, 85, 86
 Substitutionsregel für bestimmte Integrale, 95
 Summe
 Ober-, 70
 Riemannsche, 72, 154
 Unter-, 70
 Zwischen-, 72
 Summe zweier Funktionen, 178
 Summenregel, 33
 summierte Trapezregel, 110
 Supremum, 13
 Supremumsprinzip, 13

 Tangens, 58
 Tangente, 30
 Gleichung der Tangente $T(x)$, 32
 Tangentengleichung, 62
 Taylor-Entwicklung, 134
 Taylorreihe, 134, 136
 Trapezregel, 108, 110
 Fehlerabschätzung, 110
 summierte, 110
 zusammengesetzte, 110
 trigonometrische Funktionen, 58
 Ableitungen, 58
 Umkehrfunktionen, 60
 triviale Darstellung des Nullvektors, 190
 triviale Linearkombination, 190

 Umkehrfunktion, 36
 Umkehrfunktionen der trigonometrischen Funktionen, 60
 Umkehrregel, 36
 unbestimmtes Integral, 81
 uneigentlich integrierbar, 113
 uneigentlich konvergent, 10
 uneigentliche Konvergenz, 10
 uneigentlicher Grenzwert, 10
 uneigentliches Integral, 112, 113
 unendlichdimensionaler Vektorraum, 192
 unendliches Intervall, 21
 ungerade Funktion, 57
 Ungleichung
 Bernoullische, 14, 15, 18
 Dreiecks-, 7, 9, 166
 Unstetigkeitsstelle, 24
 untere Grenze, 13
 untere Schranke, 13
 kleinste, 13
 Unterräume des \mathbb{R}^2 , 195, 197
 Unterraum, 182
 Unterraumkriterium, 183
 Untersumme, 70
 Untervektorraum, 181, 182
 Ursprungsebene, 187
 Ursprungsgerade, 186

 Vektor, 180
 Einheits-, 189
 Null-, 180
 Vektorraum
 endlichdimensionaler, 191
 komplexer, 180
 reeller, 180
 unendlichdimensionaler, 192
 Unter-, 181, 182
 Vektorraum über einem Körper, 177, 180
 Vektorraumoperationen, 183
 Vergleichskriterium, 117
 verträglich, 210
 Verträglichkeit mit der Addition, 198
 Verträglichkeit mit der skalaren Multiplikation, 198
 Vollständigkeitsaxiom, 5
 Vorzeichenwechsel, 48

 Wendepunkt, 49
 Wurzel
 n -te, 173
 n -te Einheits-, 174

- q-te, 37
- Wurzelkriterium, 120
 - Folgerung, 120
- Zahl
 - rein imaginäre, 175
 - konjugiert komplexe, 163
- Zahlen
 - komplexe, 161
- Zahlengerade, 193
- Zeichenebene, 193
- Zeilenrang, 212
- Zerlegung, äquidistante, 69
- Zerlegungssatz für Polynome, 172
- Zielfunktion, 150
- zugehörige Abbildung, 207
- zugrundeliegender Körper, 180
- zusammengesetzte Trapezregel, 110
- zweidimensionales Integral, 154
 - Berechnung, 157
- Zwischenpunkte, 156
- Zwischensumme, 72
- Zwischenwertsatz, 55
- zyklometrische Funktion, 62