

Kapitel 1

Grundlagen

1.1 ein wenig Logik

In der Mathematik und in den Naturwissenschaften argumentieren wir ganz ähnlich wie im Alltagsleben, nur halten wir uns eisern an die ausgewählten Argumentationsregeln. Besonders wichtig ist die Benutzung der Worte “nicht”, “und”, “oder”, “impliziert”, “dann und nur dann”, deren Gebrauch wir jetzt erläutern. Dazu benutzen wir als Beispiele die beiden Aussagen “Die Sonne scheint auf den Stein” (Aussage A) sowie “Der Stein wird warm” (Aussage B).

Nicht:

Die Verneinung “Die Sonne scheint nicht auf den Stein”, kurz: “nicht A”, in Zeichen: $\neg A$, ist wahr, wenn die Aussage “Die Sonne scheint auf den Stein” falsch ist und “nicht A” ist falsch, wenn A wahr ist.

Schreiben wir W für wahr und F für falsch, so ergibt sich folgende Tabelle:

A	$\neg A$
W	F
F	W

Sie besagt: Ist A wahr, so ist $\neg A$ falsch, Ist A falsch, so ist $\neg A$ wahr.

Und:

Die Aussage: “Die Sonne scheint auf den Stein **und** der Stein wird warm”, ist wahr, falls sowohl “Die Sonne scheint auf den Stein” als auch “Der Stein wird warm” beide wahr sind. In allen anderen Fällen (welche sind das?) ist die durch “und” verbundene komplexe Aussage falsch. Wir schreiben $A \wedge B$ für “A und B” und erhalten die folgende Tabelle:

A	B	$A \wedge B$
W	W	W
W	F	F
F	W	F
F	F	F

Oder:

Es hat sich für den Gebrauch des Wortes “oder” in der Mathematik und Informatik - abweichend vom

Umgangsdeutsch - die lateinische und angelsächsische Bedeutung durchgesetzt:

“Die Sonne scheint auf den Stein **oder** der Stein wird warm” ist falsch, wenn weder die Sonne auf den Stein scheint noch der Stein warm wird. Sie ist wahr, wenn entweder die Sonne auf den Stein scheint oder der Stein warm wird oder beides passiert.

Wir schreiben $A \vee B$ für “A oder B” und erhalten die Tabelle:

A	B	$A \vee B$
W	W	W
W	F	W
F	W	W
F	F	F

Versuchen wir nun einmal das *deutsche* “oder”, also das “entweder - oder” auszudrücken: es ist “Entweder A oder B” und ist nichts anderes als $(A \vee B) \wedge (\neg(A \wedge B))$.

Was bedeutet dabei “nichts anderes als”?

“Entweder A oder B” wird durch die folgende Tabelle charakterisiert

A	B	Entweder A oder B
W	W	F
W	F	W
F	W	W
F	F	F

Wenn Sie nun aufgrund der bisher aufgestellten Regeln alle Fälle von A und B bestimmen, bei denen $(A \vee B) \wedge (\neg(A \wedge B))$ wahr bzw. falsch wird, so erhalten Sie dieselbe Tabelle. Vollziehen Sie dies im einzelnen nach.

A	B	$A \vee B$	$A \wedge B$	$\neg(A \wedge B)$	$(A \vee B) \wedge (\neg(A \wedge B))$
W	W	W	W	F	F
W	F	W	F	W	W
F	W	W	F	W	W
F	F	F	F	W	F

Implikation:

Wir wollen den Gebrauch der Ausdrücke “Aus A folgt B”, bzw. “Wenn A, so auch B”, “A impliziert B” festlegen. Im Alltag ist der Satz “Wenn die Sonne auf den Stein scheint, wird der Stein warm” klar.

In der Mathematik und in den Naturwissenschaften sagt man, daß dieser Satz *falsch* ist, wenn die Sonne auf den Stein scheint, aber der Stein einfach kalt bleibt. In allen anderen Fällen sagt man, der Satz ist wahr. Das ist etwas gewöhnungsbedürftig. Wir schreiben $A \Rightarrow B$ für “Aus A folgt B” bzw. “Wenn A, so B” oder “A impliziert B” und fassen unsere Überlegung in der folgenden Tabelle zusammen:

A	B	$A \Rightarrow B$
W	W	W
W	F	F
F	W	W
F	F	W

Aufgabe: Zeigen Sie, daß " $A \Rightarrow B$ " und " $(\neg A) \vee B$ " dieselben Wahrheitswerte in Abhängigkeit von den Wahrheitswerten von A und B annehmen, d.h. daß Sie für beide Aussagen die gleiche Tabelle erhalten.

" $A \Rightarrow B$ " ist also nichts anderes als " $(\neg A) \vee B$ ".

Äquivalenz:

Es gilt zwar nach unserer Erfahrung der Satz "Wenn die Sonne auf den Stein scheint, wird der Stein warm", aber der umgekehrte Schluß "Wenn der Stein warm wird, scheint die Sonne auf ihn" gilt nicht. Zum Beispiel wird der Stein auch warm, wenn wir nachts ein Feuerchen auf ihm machen.

Sei nun A die Aussage: Die natürliche Zahl a ist durch 2 und 3 teilbar. B sei die Aussage: Die natürliche a ist durch 6 teilbar. Dann gilt offenbar $A \Rightarrow B$ und $B \Rightarrow A$. Wir nennen die Aussagen A und B **äquivalent**, wenn $A \Rightarrow B$ und $B \Rightarrow A$ gilt und schreiben dann $A \Leftrightarrow B$. Hierfür sagen wir "**A gilt genau dann, wenn B gilt**" oder umständlicher: "**A gilt dann und nur dann, wenn B gilt**".

Eine aus A und B durch mehrfache Anwendung von " \neg ", " \wedge ", " \vee ", " \Rightarrow ", " \Leftarrow " erhaltene Aussage $C = C(A, B)$ heißt **Tautologie**, wenn C immer wahr ist, egal ob A oder B wahr oder falsch sind. Ein berühmtes Beispiel ist: "Wenn der Hahn kräht auf dem Mist, ändert sich's Wetter oder bleibt, wie es ist".

Formalisiert handelt es sich um die Aussage $A \Rightarrow (B \vee (\neg B))$. Eine andere, seit Aristoteles gefeierte Tautologie ist $(A \wedge (\neg A)) \Rightarrow B$. Sie besagt: Aus falschen Voraussetzungen läßt sich alles folgern.

Nun können wir präzisieren, was es bedeutet, daß die Aussage $C(A, B)$ nichts anderes ist als die Aussage $D(A, B)$, wo beide aus $A, B, \neg, \vee, \wedge, \Rightarrow$ aufgebaut sind. " $C(A, B)$ ist nichts anderes als $D(A, B)$ " oder " $C(A, B)$ besagt dasselbe wie $D(A, B)$ ", wenn $C(A, B) \Leftrightarrow D(A, B)$ eine Tautologie ist. Wir schreiben hierfür $C(A, B) \equiv D(A, B)$.

In diesem Sinn besagt zum Beispiel $A \Rightarrow B$ dasselbe wie $\neg B \Rightarrow \neg A$, also $(A \Rightarrow B) \equiv (\neg B \Rightarrow \neg A)$. Das ist das aus unzähligen Kriminalromanen bekannte Prinzip des *indirekten Beweises*.

"Wenn der Stein nicht warm wird, scheint die Sonne nicht auf ihn."

Aufgaben: 1. Arbeiten Sie den Text durch anhand der beiden Aussagen A: Es regnet, B: Die Straße wird naß.

2. Wählen Sie jetzt: A: a ist durch 6 teilbar, B: a ist durch 3 teilbar. (a ist eine natürliche Zahl).

3. Zeigen Sie bitte $(\neg(\neg A)) \equiv A$

4. Zeigen Sie bitte die sog. de Morgan'schen Regeln: $\neg(A \wedge B) \equiv (\neg A) \vee (\neg B)$, $(\neg(A \vee B)) \equiv (\neg A) \wedge (\neg B)$.

Oft betrachten wir auch mehrere Aussagen $A_1, A_2, A_3, \dots, A_n$. Wir nennen sie äquivalent, wenn sie paarweise äquivalent sind, d.h. wenn $A_i \Rightarrow A_j$ für $i, j = 1 \dots n$ gilt. Um die Äquivalenz zu zeigen, genügt es, $A_1 \Rightarrow A_2, A_2 \Rightarrow A_3, \dots, A_{n-1} \Rightarrow A_n, A_n \Rightarrow A_1$ nacheinander zu zeigen.

Beispiel: Sei a eine natürliche Zahl. Die folgenden drei Aussagen sind äquivalent:

A_1 : Die Division von a durch 21 geht ohne Rest auf

A_2 : Es gibt eine natürliche Zahl b mit $a = 21 \cdot b$

A_3 : Sowohl die Division von a durch 3 als auch durch 7 geht ohne Rest auf.

1.2 Mengen

1.2.1 Einführung

Wir treiben hier keine Mengenlehre. Das ist ein anspruchsvoller Zweig der Grundlagen der Mathematik. Wir benutzen statt dessen Mengen und den Umgang mit ihnen zur Erleichterung der Formulierung mathematischer Sachverhalte, gewissermaßen also als *mathematische Kurzschrift*.

Eine **Menge** M ist für uns eine Gesamtheit von Objekten x mit einer bestimmten Eigenschaft E . Zum Beispiel bilden die deutschen Bürger eine Menge D . $x \in D$ besagt: x ist deutscher Bürger. Ist x ein Objekt

mit dieser Eigenschaft E , so sagen wir: x ist Element von M in Zeichen: $x \in M$. Hat x die Eigenschaft E nicht, so schreiben wir $x \notin M$.

In der Analysis häufig benutzte Mengen sind

- \mathbb{N} : Menge der natürlichen Zahlen $1, 2, 3, 4, \dots$
- \mathbb{N}_0 : Menge der natürlichen Zahlen zuzüglich 0
- \mathbb{Z} : Menge der ganzen Zahlen $0, \pm 1, \pm 2, \dots$
- \mathbb{Q} : Menge der rationalen Zahlen p/q mit $p \in \mathbb{Z}, q \in \mathbb{N}$
- \mathbb{R} : Menge der reellen Zahlen
- \mathbb{C} : Menge der komplexen Zahlen.

Aus Vereinfachungsgründen führt man auch die **leere Menge** \emptyset Sie enthält kein Element.

1.2.2 Mengen und Mengenrelationen

In Analogie zu unseren sprachlichen Ausdrücken “nicht”, “und”, “oder”, “impliziert”, “äquivalent”, mit denen wir Aussagen verknüpfen, führen wir nun Mengenoperationen und -relationen ein.

Der Implikation entspricht die Teilmengenrelation

Seien M, N Mengen. M ist **Teilmenge** von N , in Zeichen: $M \subset N$, wenn für alle x gilt $x \in M \Rightarrow x \in N$. In unserer Tabelle ist jede Menge Teilmenge der auf ihr folgenden, also $\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z}$, usw. Die leere Menge \emptyset ist Teilmenge jeder anderen Menge. Denn da $x \in \emptyset$ stets falsch ist, gilt stets $x \in \emptyset \Rightarrow x \in N$. (vergl. den Abschnitt über Tautologien)

Der Äquivalenz entspricht die Gleichheit von Mengen

$M = N$, wenn für alle x gilt: $x \in M \Leftrightarrow x \in N$.

Beispiel: Sei $M = \{x \in \mathbb{N} : x \text{ ist durch } 6 \text{ teilbar}\}$ und $N = \{x \in \mathbb{N} : x \text{ ist durch } 2 \text{ und } 3 \text{ teilbar}\}$ Dann ist $M = N$.

Dem “und” entspricht der Durchschnitt von Mengen

Der **Durchschnitt** $M \cap N$ von M und N ist die Menge aller x mit $(x \in M) \wedge (x \in N)$.

Beispiel: Wir schreiben $a \mid b$, falls die natürliche Zahl b durch die natürliche Zahl a teilbar ist.

Sei $M = \{x \in \mathbb{N} : 2 \mid x\}$, $N = \{x \in \mathbb{N} : 3 \mid x\}$. Dann ist $M \cap N = \{x \in \mathbb{N} : 6 \mid x\}$.

Dem “oder” entspricht die Vereinigung von Mengen

Die **Vereinigung** $M \cup N$ von M und N ist die Menge $M \cup N = \{x : (x \in M) \vee (x \in N)\}$.

Beispiel: $M = \{x \in \mathbb{N} : 2 \mid x\}$, $N = \{x \in \mathbb{N} : \neg(2 \mid x)\}$ $M \cup N = \mathbb{N}$.

Der Negation entspricht das Komplement

Sei $M \subset N$. Dann ist das **Komplement von** M in N die Menge $M^c = \{x \in N : x \notin M\}$. Man schreibt hierfür auch $N \setminus M$

Aufgaben: Sei $M \subset N$

1. Zeigen Sie bitte $(M^c)^c = M$

2. Seien $M_1, M_2 \subset N$. Zeigen Sie bitte die de Morganschen Regeln $(M_1 \cap M_2)^c = M_1^c \cup M_2^c$, $(M_1 \cup M_2)^c = M_1^c \cap M_2^c$.

Tip: Verwenden Sie die Aufgaben 3 und 4 3

Verallgemeinerungen

$$\begin{aligned}M_1 \cap M_2 \cap \cdots \cap M_n &= \{x : (x \in M_1) \wedge (x \in M_2) \wedge \cdots \wedge (x \in M_n)\} \\M_1 \cup M_2 \cup \cdots \cup M_n &= \{x : (x \in M_1) \vee \cdots \vee (x \in M_n)\}.\end{aligned}$$

1.2.3 Kartesisches Produkt und Abbildungen

Wir lesen von links nach rechts, also wissen wir, welches Zeichen vor welchem beim Lesen kommt. Wenn wir Personenlisten anfertigen, ist es wichtig, ob wir den Vornamen **vor** dem Nachnamen eintragen oder umgekehrt, weil sonst Irrtümer entstehen.

(Manfred, Wolff) ist in diesem Sinn ein *geordnetes Paar* von Daten. Ohne die zusätzliche Information, daß der Vorname **vor** dem Nachnamen kommt, könnten wir Datenpaare (Stephan, Paul) nicht eindeutig deuten.

Aus der Schule kennen Sie die Ebene. Einen Punkt in ihr findet man nach Festlegung eines Achsenkreuzes mit waagrechter und senkrechter Achse durch Angabe eines geordneten Zahlenpaares (x, y) , wo x sich auf die waagrechte, y auf die senkrechte Achse bezieht. Auch hier ist also die *Reihenfolge der Zahlen* wichtig. Damit ist uns klar, was ein geordnetes Paar ist. Ebenso ist uns klar, was ein *geordnetes Tripel* (x, y, z) bzw. ein *geordnetes 10-Tupel* (x_1, \dots, x_{10}) oder allgemeiner ein *geordnetes n -Tupel* (x_1, \dots, x_n) ist.

Definition 1.1 (Kartesisches Produkt)

Sei $n \in \mathbb{N}, n \geq 2$ und seien M_1, M_2, \dots, M_n nichtleere Mengen. Dann heißt die Menge der geordneten n -Tupel

$$M_1 \times M_2 \times \cdots \times M_n = \{(x_1, \dots, x_n) : x_1 \in M_1, x_2 \in M_2, \dots, x_n \in M_n\}$$

das **kartesische Produkt** der Mengen M_1, M_2, \dots, M_n .

Bemerkung 1.2 Statt $M_1 \times \cdots \times M_n$ schreibt man auch $\prod_{k=1}^n M_k$. Sind alle Mengen untereinander gleich, also $M_1 = \cdots = M_n = M$, so schreibt man M^n statt $\prod_{k=1}^n M_k$.

So ist die Ebene nach Einführung eines Koordinatensystems mit \mathbb{R}^2 identifizierbar.

In der Schule haben Sie mit Funktionen wie zum Beispiel $f(x) = x \sin(x)$ gearbeitet. Auf Reisen benutzen Sie Landkarten und machen Fotos. Und bei einer großen Feierlichkeit wird oft eine Sitzordnung mit Tischkarten festgelegt. Alle diese an sich völlig verschiedenen Dinge haben etwas gemeinsam, das wir im Begriff der **Abbildung** festlegen

Definition 1.3 (Abbildungen)

Seien M und N nichtleere Mengen.

- Eine Abbildung f der Menge M in die Menge N erhält man durch eine Zuordnung, die jedem Argument (Urbild) $x \in M$ eindeutig sein Bild $f(x)$ zuordnet. M heißt Definitionsbereich von f
- Die Menge $G_f = \{(x, f(x)) : x \in M\} \subset M \times N$ heißt Graph von f .

G_f ist das, was Sie (ausschnittsweise) in der Schule zeichneten. Beim obigen Beispiel $f(x) = x \sin x$ ist impliziert $M = \mathbb{R} = N$ angenommen. Bei der Landkarte ist M der Ausschnitt der Erde, den die Landkarte darstellt, N ist die Landkarte und f die spezielle Darstellungsweise (sog. Kartenprojektion). Bestimmen Sie selbst in den Beispielen "Foto" und "Tischordnung" die Definitionsbereiche M , die jeweilige Menge N und die Abbildung f . (Es gibt mehrere Möglichkeiten, weil die Umgangssprache nicht eindeutig ist).

Zwei Zuordnungsvorschriften können verschieden sein, aber zur gleichen Abbildung führen. Zum Beispiel liefern die Vorschriften $x \mapsto x^2 + x$ und $x \mapsto x(x + 1)$ dieselbe Abbildung (Funktion). Präziser erklären wir:

Definition 1.4 (Gleichheit von Abbildungen)

Zwei Abbildungen f von M_1 nach N_1 und g von M_2 nach N_2 heißen gleich, wenn $N_1 = N_2$ und $G_f = G_g$ gilt.

Im wesentlichen sind zwei Funktionen also gleich, wenn ihre Graphen gleich sind. Automatisch folgt daraus, daß ihre Definitionsbereiche gleich sind. Für eine Abbildung f von M nach N schreiben wir kurz $f : M \rightarrow N$. Ist $A \subset M$, so heißt $f(A) = \{f(x) : x \in A\}$ **Bild(menge) von A unter f**. Ist $B \subset N$, so heißt $f^{-1}(B) = \{x : f(x) \in B\}$ **Urbild von B unter f**. Trainieren Sie gleich den Umgang mit diesen Begriffen.

Aufgaben:

1. Sei $M = \mathbb{N}_0$, $N = \{0, 1, 2\}$ und $f(x) = x \bmod 3$, das ist der Rest nach Division durch 3. Was ist $f(\{3k + 1 : k \in \mathbb{N}\})$? Was ist $f^{-1}(\{0\})$?
2. Sei $M = N = \mathbb{R}$ und $f(x) = x^2$. Was ist $f(M)$? Was ist $f^{-1}(\{y : y \leq 0\})$. Was ist $f^{-1}(\{1\})$?
3. Sei $M = N = \mathbb{R}$ und $f(x) = \sin(x)$. Für zwei reelle Zahlen $s < t$ sei $[s, t]$ das Intervall aller Zahlen x mit $s \leq x \leq t$. Was ist $f([0, \pi])$? Was ist $f^{-1}([5, 6])$? Was ist $f^{-1}([0, 1])$?
4. Finden Sie alle Abbildungen der Menge $M = \{1, 2, 3\}$ in sich. (Es sind 24).
5. Ist $\{(x^3, x) : x \in \mathbb{R}\}$ Graph einer Abbildung von $\mathbb{R} \rightarrow \mathbb{R}$?
6. Ist $\{(x^2, x) : x \in \mathbb{R}\}$ Graph einer Abbildung von $\mathbb{R} \rightarrow \mathbb{R}$?
7. Sei $G = \{(x, y) : x \in M\} \subset M \times N$. Zeigen Sie bitte: G ist genau dann Graph einer Abbildung $f : M \rightarrow N$, wenn für alle $x \in M$ aus $(x, y) \in G$ und $(x, z) \in G$ stets $y = z$ folgt.

Definition 1.5 (Surjektivität, Injektivität, Bijektivität)

Sei $f : M \rightarrow N$ eine Abbildung.

- a) f heißt surjektiv, wenn $N = f(M)$.
- b) f heißt injektiv, wenn aus $x \neq y$ stets $f(x) \neq f(y)$ folgt.
- c) f heißt bijektiv, wenn f surjektiv und injektiv ist. Eine bijektive Abbildung heißt kurz Bijektion

Trainieren Sie gleich den Umgang mit diesen Begriffen:

Aufgaben:

1. Welche der folgenden Abbildungen ist surjektiv, welche injektiv, welche bijektiv? Hier ist $M = N = \mathbb{R}$. $f(x) = x$, $f(x) = x^2$, $f(x) = x^3$, $f(x) = \sin(x)$.
2. Bestimmen Sie alle Bijektionen der Menge $M = \{1, 2, 3\}$ auf sich.
3. Sei $f : M \rightarrow N$. Zeigen Sie bitte:

- a) f ist genau dann surjektiv, wenn die Gleichung $f(x) = y$ für jede rechte Seite $y \in N$ **mindestens** eine Lösung hat.
- b) f ist genau dann injektiv, wenn die Gleichung $f(x) = y$ für jede rechte Seite $y \in N$ **höchstens** eine Lösung hat.
- c) f ist genau dann bijektiv, wenn die Gleichung $f(x) = y$ für jede rechte Seite $y \in N$ **genau eine** Lösung hat. Dabei ist eine Lösung der Gleichung $f(x) = y$ bei vorgegebener rechter Seite $y \in N$ nichts anderes als ein Urbild von y unter f .

4. Zeigen Sie bitte: $f : M \rightarrow N$ ist genau dann injektiv, wenn $H := \{(f(x), x) : x \in M\}$ Graph einer Abbildung h von $f^{-1}(N)$ nach M ist. *Tip:* Verwenden Sie Aufgabe 7 6

In der Schule haben Sie bereits “zusammengesetzte” Funktionen behandelt, zum Beispiel $f(x) = \sin(x^2)$. Sie entsteht, indem man erst jedem x das x^2 zuordnet und *danach* den Sinus darauf anwendet, also zwei andere Funktionen hintereinander ausführt. Tatsächlich gilt der folgende allgemeine Satz.

Satz 1.6 (Hintereinanderausführung von Abbildungen)

- a) Seien $f : M \rightarrow N$ und $g : N \rightarrow P$ zwei Abbildungen. Dann erhält man durch die Zuordnung $x \mapsto e(x) = g(f(x))$ eindeutig eine Abbildung von M nach P , die Hintereinanderausführung $g \circ f$ der Abbildungen f und g .
- b) Ist zusätzlich $h : P \rightarrow R$ eine Abbildung, so gilt $(h \circ g) \circ f = h \circ (g \circ f)$. Man sagt, die Hintereinanderausführung ist **assoziativ**.

Beweis: a) ist klar.
 b) Es ist für $x \in M$

$$\begin{aligned} ((h \circ g) \circ f)(x) &= (h \circ g)(f(x)) = h(g(f(x))) \\ &= (h \circ (g \circ f))(x). \end{aligned}$$

Sei id_M die identische Abbildung von M auf sich, d.h. $id_M(x) = x$ für alle x . Bijektive Abbildungen lassen sich dadurch charakterisieren, daß sie eine Umkehrabbildung besitzen!

Satz 1.7 Sei $f : M \rightarrow N$ eine Abbildung. Die folgenden Aussagen sind äquivalent:

- (i) f ist bijektiv.
- (ii) Es gibt eine Abbildung $g : N \rightarrow M$ mit $g \circ f = id_M$ und $f \circ g = id_N$.

g ist eindeutig bestimmt und heißt Umkehrabbildung oder inverse Abbildung f^{-1} .

Den einfachen Beweis stellen wir als Aufgabe.

Aufgaben:

1. Beweisen Sie diesen Satz. *Tip:* Für (i) \Rightarrow (ii) betrachten Sie $H = \{(f(x), x) : x \in M\}$ und verwenden Sie Aufgabe 4 7.
2. Was ist die Umkehrabbildung zu $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^3$?
3. Zeigen Sie bitte: Die Hintereinanderausführung

$$\left\{ \begin{array}{l} \text{surjektiver} \\ \text{injektiver} \\ \text{bijektiver} \end{array} \right\} \text{ Abbildungen ist } \left\{ \begin{array}{l} \text{surjektiv} \\ \text{injektiv} \\ \text{bijektiv} \end{array} \right\}$$

1.2.4 Potenzmenge, beliebige Mengenoperationen

Sei M eine Menge. Unter der Potenzmenge $\mathcal{P}(M)$ verstehen wir die Menge aller Teilmengen von M , $\mathcal{P}(M) = \{A : A \subset M\}$. Anders ausgedrückt: $A \subset M \Leftrightarrow A \in \mathcal{P}(M)$.

Beispiele:

1. $M = \{1\}$, $\mathcal{P}(M) = \{\emptyset, \{1\}\}$.
2. $M = \{1, 2\}$, $\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.
3. $M = \{1, 2, 3\}$, $\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.
4. $M = \emptyset$, $\mathcal{P}(M) = \{\emptyset\}$.
5. $M = \mathbb{N}$. Hier kann man $\mathcal{P}(M)$ nicht mehr explizit hinschreiben (das ist klar), aber auch nicht durch Algorithmen beschreiben.

Für spätere Zwecke ist die folgende Verallgemeinerung von Durchschnitt und Vereinigung nützlich:

Definition 1.8 (Beliebige Durchschnitte und Vereinigungen)

Seien I und M Mengen, $I \neq \emptyset$ und sei $A : I \rightarrow \mathcal{P}(M)$, $i \mapsto A_i$ eine Abbildung von I in die Potenzmenge von M . Dann ist der Durchschnitt der A_i

$$\bigcap_{i \in I} A_i = \{x : \text{Für alle } i \in I \text{ ist } x \in A_i\}$$

und die Vereinigung der A_i

$$\bigcup_{i \in I} A_i = \{x : \text{Es gibt ein } i \in I \text{ mit } x \in A_i\}$$

Beispiele:

1. Aus der Schule wissen Sie: Ist $x \in \mathbb{R}$ und $0 \leq x < \frac{1}{n}$ für jede natürliche Zahl n , so ist $x = 0$. Sei $A_n = \{y : 0 \leq y < \frac{1}{n}\}$. Dann ist also $\bigcap_{n \in \mathbb{N}} A_n = \{0\}$.
2. Sei $A_n = \{x \in \mathbb{R} : x < \frac{1}{n}\}$ für $n \in \mathbb{N}$. Dann ist $\bigcap_{n \in \mathbb{N}} A_n = \{x \in \mathbb{R} : x \leq 0\}$.
3. Sei $I = \{1, 2\}$. Dann ist $\bigcap_{i \in I} A_i = A_1 \cap A_2$ (im früheren Sinn) und $\bigcup_{i \in I} A_i = A_1 \cup A_2$.
4. Die deMorganschen Regeln gelten in der folgenden Form. Seien I, M und A wie in der obigen Definition. Dann gilt $(\bigcap_{i \in I} A_i)^c = \bigcup_{i \in I} A_i^c$ und $(\bigcup_{i \in I} A_i)^c = \bigcap_{i \in I} A_i^c$. Beweisen Sie bitte selbst die Formeln. Sie benötigen dazu die aus dem Alltag entlehnten Aussageformeln:
-(für alle i $C(i)$) \equiv Es gibt i ($C(i)$) und
-(Es gibt i $C(i)$) \equiv Für alle i ($C(i)$),
wo $C(i)$ für jedes i eine Aussage ist.

1.3 Einiges über natürliche Zahlen

1.3.1 Die natürlichen Zahlen und das Induktionsprinzip

Sei A die Menge der geraden natürlichen Zahlen, $A = \{n : 2|n\}$. A hat ein kleinstes Element nämlich die 2. Ebenso hat die Menge B aller Primzahlen, die größer als 20 sind, ein kleinstes Element, nämlich 23.

Damit ist das folgende Axiom motiviert:

Induktionsaxiom

Jede nichtleere Teilmenge A der natürlichen Zahlen hat ein kleinstes Element.

Mit diesem Induktionsaxiom können wir nun sofort das Beweisprinzip der vollständigen Induktion und das Prinzip der rekursiven Definition von Funktionen einsehen. Dazu zunächst zwei Beispiele.

Beispiele:

- Wir wollen die Formel $1 + x + x^2 + \dots + x^n = (x^{n+1} - 1)/(x - 1)$ für alle $x \neq 1$ und $n \in \mathbb{N}$ beweisen. Wir wählen ein festes $x \in \mathbb{R}, x \neq 1$. Dann erhalten wir für jedes $n \in \mathbb{N}$ die Aussage $A(n) : 1 + x + \dots + x^n = (x^{n+1} - 1)/(x - 1)$. Für $n = 1$ lautet $A(1) : 1 + x = (x^2 - 1)/(x - 1)$. Sie ist wegen $x^2 - 1 = (x + 1)(x - 1)$ wahr. Wir nehmen nun an, $A(n)$ sei wahr und folgern daraus, daß $A(n + 1)$ wahr ist. Es ist

$$\begin{aligned}
 1 + x + \dots + x^n + x^{n+1} & \stackrel{\substack{= \\ A(n) \text{ ist wahr}}}{=} \frac{x^{n+1} - 1}{x - 1} + x^{n+1} \\
 & = \frac{x^{n+1} - 1 + x^{n+1}(x - 1)}{x - 1} \\
 & = \frac{x^{n+2} - 1}{x - 1}
 \end{aligned}$$

Also ist unter unserer Annahme der Wahrheit von $A(n)$ auch $A(n + 1)$ wahr. Können wir daraus folgern, daß $A(n)$ für jedes n wahr ist? Ja, mit Hilfe des *Induktionsaxioms*.

Sei $B = \{n \in \mathbb{N} : A(n) \text{ ist wahr}\}$. Angenommen $B \neq \mathbb{N}$. Dann ist $B^c =: C = \{n : A(n) \text{ ist falsch}\}$ nicht leer, hat also ein kleinstes Element n_C . Weil $A(1)$ wahr ist (s.o.), ist $n_C > 1$. Da n_C das kleinste Element von C ist, ist $A(n_C - 1)$ wahr. Aber dann ist nach dem 2. Argument auch $A(n_C)$ wahr im Widerspruch zu $n_C \in C$.

- Wir möchten nachweisen, daß durch $0! = 1$ und $n! = 1 \cdot 2 \cdot \dots \cdot n$ für $n \geq 1$ eindeutig eine Abbildung von \mathbb{N}_0 nach \mathbb{N} erklärt ist. Dies geschieht so: sei $F : \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{N}_0$ gegeben durch $F(n, \gamma) = (n+1) \cdot \gamma$. Wir erklären $g(0) = 1$ und $g(n+1) = F(n+1, g(n))$ für alle n in \mathbb{N}_0 . Damit ist $g : \mathbb{N}_0 \rightarrow \mathbb{N}$ eindeutig erklärt. Das Argument hierfür ist völlig analog zu dem im ersten Beispiel.

Wir formulieren beide Prinzipien etwas allgemeiner:

Beweisprinzip der vollständigen Induktion

Sei $n_0 \in \mathbb{N}_0$ fest. Für jedes $n \geq n_0$ sei $A(n)$ eine Aussage.

Es gelte

- a) $A(n_0)$ ist wahr (Induktionsanfang)
- b) Für jedes $n \geq n_0$ ist " $A(n) \Rightarrow A(n + 1)$ " wahr (Induktionsschritt)

Dann sind alle Aussagen $A(n)$ wahr.

Prinzip der eindeutigen Definition durch Rekursion

Sei $n_0 \in \mathbb{N}_0$ fest gewählt. Sei $A = \{n \in \mathbb{N}_0 : n \geq n_0\}$ und M eine beliebige, nicht leere Menge. Sei schließlich $F : A \times M \rightarrow M$ eine Abbildung und $s \in M$. Dann liefert die folgende Vorschrift eine eindeutige Funktion $g : A \rightarrow M$:

a) $g(n_0) = s$

Festlegung des Startgliedes

b) $g(n + 1) = F(n, g(n))$

Rekursionsschritt

Beispiele:

1. *Potenzen:* Sei $x \in \mathbb{R}$ fest $x^0 = 1$, $x^{n+1} = x \cdot x^n$. Hier ist $M = \mathbb{R}$, $A = \mathbb{N}_0$, $F(n, y) = yx$

2. *Summen:* Sei $a : \mathbb{N}_0 \rightarrow \mathbb{R}$, $n \mapsto a_n$ eine Abbildung.

Wir definieren $\sum_{k=0}^n a_k = a_0 + a_1 + \dots + a_n$ durch $\sum_{k=0}^0 a_k = a_0$, $\sum_{k=0}^{n+1} a_k = \sum_{k=0}^n a_k + a_{n+1}$

3. *Produkte:* $a_0 a_1 \dots a_n$ wird definiert durch $\prod_{k=0}^0 a_k = a_0$, $\prod_{k=0}^{n+1} a_k = \prod_{k=0}^n a_k \cdot a_{n+1}$

4. *Fakultät (s.O.):* $0! = 1$, $(n + 1)! = n! \cdot (n + 1)$

Aufgaben: Beweisen Sie mit vollständiger Induktion:

1. $\sum_{k=1}^n k = n(n + 1)/2$

2. $\sum_{k=1}^n k^2 = n(n + 1)(2n + 1)/6$

1.3.2 Etwas Kombinatorik

Als erstes Problem behandeln wir die Anzahl der Bijektionen von einer Menge M mit n Elementen auf eine Menge N (die dann auch n Elemente haben muß, wenn es mindestens eine Bijektion von M auf N gibt). Dazu bezeichnen wir mit $\#(A)$ die Anzahl der Elemente einer Menge A und mit $\mathcal{B}(M, N)$ die Menge aller Bijektionen von M auf N . Konkret erhalten wir zum Beispiel die Anzahl aller möglichen Tischordnungen für n Personen an einem Tisch mit n Stühlen.

Satz 1.9 Sei $n \in \mathbb{N}$ und M, N seien zwei Mengen mit n Elementen. Dann gibt es genau $n!$ Bijektionen von M auf N .

Beweis: (durch Induktion)

$n = 1$. $M = \{a\}$, $N = \{b\}$. Die einzige Bijektion ist $\sigma : a \mapsto b$.

Also $\#(\mathcal{B}(M, N)) = 1 = 1!$

Induktionsschritt: Annahme, der Satz ist richtig für Mengen mit n Elementen.

Sei $M = \{a_1, \dots, a_{n+1}\}$, $N = \{b_1, \dots, b_{n+1}\}$, $A = \{a_1, \dots, a_n\}$ und $B_k = N \setminus \{b_k\}$. Schließlich sei $\mathcal{B}_k = \{\sigma \in \mathcal{B}(M, N) : \sigma(a_{n+1}) = b_k\}$.

Ist σ aus \mathcal{B}_k , so ist die Einschränkung σ_A aus $\mathcal{B}(A, B_k)$.

Ist umgekehrt $\tau \in \mathcal{B}(A, B_k)$ und σ erklärt durch

$$\sigma(a_l) = \begin{cases} \tau(a_l) & : l \leq n \\ b_k & : l = n + 1, \end{cases} \quad \text{so ist } \sigma \in \mathcal{B}_k$$

Die Abbildung $\mathcal{B}_k \ni \sigma \mapsto \sigma|_A \in \mathcal{B}(A, B_k)$ ist also bijektiv. Damit ist nach Induktionsannahme

$n! = \#(\mathcal{B}(A, B_k)) = \#(\mathcal{B}_k)$. Wegen $\mathcal{B}_k \cap \mathcal{B}_l = \emptyset$ für $k \neq l$ (Warum?) und $\bigcup_{k=1}^{n+1} \mathcal{B}_k = \mathcal{B}(M, N)$ folgt

$$\#(\mathcal{B}(M, N)) = \sum_{k=1}^{n+1} \#(\mathcal{B}_k) = (n + 1) \cdot n! = (n + 1)!$$

Nach dem Prinzip der vollständigen Induktion ist der Satz bewiesen.

Ist $M = N$ so nennt man eine Bijektion auch **Permutation** der Ordnung n . Wir erhalten

Korollar 1.10 *Es gibt genau $n!$ Permutationen der Ordnung n .*

Als zweites Problem behandeln wir die Anzahl der k -elementigen Teilmengen einer n -elementigen Menge. Konkret taucht diese Frage in der Quantenmechanik auf: Fermi-Teilchen nehmen nicht gleichzeitig denselben Energiezustand an. Wenn wir $k(\leq n)$ Teilchen und n Energiezustände haben, wieviel verschiedene Konfigurationen gibt es dann? Nummerieren Sie die Energiezustände durch: $E = \{1, \dots, n\}$ und setzen Sie $E_k = \{\tau \in E: \text{Eines der } k \text{ Teilchen hat den Zustand } \tau\}$. Dann ist E_k eine k -elementige Teilmenge und jede Konfiguration wird durch eine solche Teilmenge dargestellt.

Sei $C(n, k)$ die Anzahl der k -elementigen Teilmengen einer n -elementigen Menge. Dann ist $C(n, k) = 0$ für $n < k$. Ferner ist $C(n, 0) = C(n, n) = 1$ und $C(n, k) = C(n, n - k)$.

Denn sei M eine Menge mit n Elementen und \mathcal{T}_k die Menge der k -elementigen Teilmengen ($0 \leq k \leq n$). Dann ist die Komplementabbildung

$\mathcal{T}_k \ni T \mapsto T^c = M \setminus T \in \mathcal{T}_{n-k}$ bijektiv (wegen $(T^c)^c = T$).

Satz 1.11 *Es ist $C(n, k) = \frac{n!}{k!(n-k)!}$ für $0 \leq k \leq n$ und $n \geq 1$.*

Beweis: (durch Induktion)

Die Aussage $A(n)$ ist: Für $0 \leq k \leq n$ ist $C(n, k) = \frac{n!}{k!(n-k)!}$. Sie ist für $n = 1$ nach dem letzten Absatz wahr.

Induktionsschritt: Sei $A(n)$ für ein $n \geq 1$ wahr. Sei $M = \{1, \dots, n + 1\}$ und $1 \leq k \leq n$ ($k = 0$ und $k = n + 1$ wurde oben behandelt). Ist T eine k -elementige Teilmenge, die $(n + 1)$ nicht enthält, so liegt T in $\{1, \dots, n\}$. Nach Induktionsannahme gibt es davon $\frac{n!}{k!(n-k)!}$. Ist T eine k -elementige Teilmenge, die $(n + 1)$ enthält, so ist $T \setminus \{n + 1\}$ eine $(k - 1)$ -elementige Teilmenge und davon gibt es nach Induktionsannahme $\frac{n!}{(k-1)!(n-(k-1))!}$. Die beiden Fälle liefern insgesamt alle k -elementigen Teilmengen. Wir erhalten

$$C(n + 1, k) = \frac{n!}{k!(n - k)!} + \frac{n!}{(k - 1)!(n - (k - 1))!} = \frac{(n + 1)!}{k!(n + 1 - k)!}$$

wobei die letzte Gleichung durch einfaches Ausrechnen folgt.

Die Zahlen $C(n, k)$ werden $\binom{n}{k}$ geschrieben (lies n **über** k). Sie heißen wegen des folgenden Satzes **Binominalkoeffizienten**.

Satz 1.12 *Für beliebige Zahlen $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt*

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

Für $n = 2$ kennen Sie diesen Satz aus der Schule.

Beweis: (durch Induktion)

Induktionsanfang: Der Satz ist richtig für $n = 1$.

Induktionsschritt: Angenommen der Satz ist richtig für ein $n \geq 1$.

Es ist

$$\begin{aligned}
 (a+b)^{n+1} &= (a+b)(a+b)^n \\
 &= (a+b) \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \\
 &= \sum_{k=0}^n \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^n \binom{n}{k} a^{n-k} b^{k+1} \\
 &= a^{n+1} + \sum_{k=1}^n \underbrace{\left(\binom{n}{k} + \binom{n}{k-1} \right)}_{= \binom{n+1}{k}} a^{n+1-k} b^k + b^{n+1} \\
 &= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k.
 \end{aligned}$$

Mit dem Prinzip der vollständigen Induktion folgt die Behauptung.

Aufgaben:

1. Wieviel verschiedene Spielausgänge gibt es beim Zahlenlotto 6 aus 49?
2. Zeigen Sie bitte durch Induktion $\sum_{k=1}^n k^3 = (\sum_{k=1}^n k)^2 = \frac{1}{4}n^2(n+1)^2$.
3. Zeigen Sie bitte durch günstige Wahl von a und b für $n \geq 1$ $\sum_{k=0}^n \binom{n}{k} = 2^n$. Wieviel Teilmengen hat eine Menge mit n Elementen insgesamt?
4. Für welche $n \geq 2$ gilt $2^n > n^2$. Geben Sie einen Beweis für Ihre Aussage

1.4 Der Körper der reellen Zahlen

1.4.1 Körperaxiome

Wir stellen zunächst die einfachsten Rechenregeln für reelle Zahlen zusammen:

Addition:

Die Summe $x + y$ zweier reeller Zahlen x, y ist eindeutig erklärt. Es gilt

- | | |
|---|---|
| 1. $(x + y) + z = x + (y + z)$ für alle x, y, z | <i>Assoziativgesetz</i> |
| 2. Es gibt ein Element 0 mit $x + 0 = 0 + x$ für alle x | <i>Existenz des neutralen Elements der Addition</i> |
| 3. Zu jedem x gibt es ein y mit $x + y = 0$ | <i>Existenz des additiven Inversen</i> |
| 4. $x + y = y + x$ für alle x, y | <i>Kommutativgesetz</i> |

Im folgenden geben wir einfache Sätze mit ihren Beweisen an. Dies dient zum Einüben im mathematischen Argumentieren. Informatiker sollten im Lauf ihres Studiums schauen, ob es Beweisprogramme gibt, die die Beweise liefern.

Satz 1.13 a) Allein aus Regel 2 folgt: Es gibt **genau ein** neutrales Element der Addition

b) Aus den Regeln 1 bis 4 folgt. Es gibt zu jedem x **genau ein** additives Inverses. Wir bezeichnen es mit $-x$ und schreiben $y - x = y + (-x)$.

c) Es ist $-(-x) = x$.

d) Es gilt stets $-(x + y) = -x - y$.

Beweis: a) Sei außer 0 noch z ein weiteres neutrales Element, d.h. es gilt $z + x = x$ für alle x . Dann ist speziell für $x = 0$ $0 = 0 + z = z$, letzteres weil 0 neutrales Element der Addition ist.

b) Sei x fest und seien u und y additive Inverse zu x , es gelte also $x + u = 0 = x + y$. Dann ist

$$y \underbrace{=}_{(2)} y + 0 = y + (x + u) \underbrace{=}_{(1)} (y + x) + u \underbrace{=}_{(4)} (x + y) + u = 0 + u \underbrace{=}_{(2)} u$$

c) Es ist $(-x) + (-(-x)) = 0$. Addition von x auf beiden Seiten liefert $\underbrace{x + (-x)}_{=0} + (-(-x)) = x$.

d) Analog

Multiplikation

Für je zwei reelle Zahlen x, y ist ihr Produkt eindeutig erklärt. Es gilt:

1. $(xy)z = x(yz)$ für alle x, y, z Assoziativgesetz
2. Es gibt ein Element 1 mit $1x = x \cdot 1$ für alle x Existenz des neutralen Elements der Multiplikation
3. Zu jedem $x \neq 0$ gibt es ein y mit $xy = 1$ Existenz des multiplikativen Inversen für $x \neq 0$
4. $xy = yx$ für alle x, y Kommutativgesetz

Den nächsten Satz können Sie genau so beweisen wie den obigen, Sie müssen nur $+$ durch \cdot ersetzen.

Satz 1.14 a) Allein aus Regel 2 der Multiplikation folgt: es gibt **genau ein** neutrales Element der Multiplikation

b) Aus den Regeln 1 bis 4 der Multiplikation folgt: es gibt zu jedem $x \neq 0$ **genau ein** multiplikatives Inverses, es wird mit x^{-1} oder $\frac{1}{x}$ bezeichnet. Wir schreiben $\frac{y}{x}$ für $y \cdot x^{-1}$.

Aufgabe: Beweisen Sie bitte diesen Satz

Verbindung von Addition und Multiplikation

Für alle x, y, z gilt

- $x(y + z) = xy + xz$ Distributivgesetz

Aus den Regeln für die Addition, die Multiplikation und aus dem Distributivgesetz lassen sich nun alle anderen Ihnen bekannten Formeln (Rechenregeln) ableiten. Hier sind einige Beispiele!

Satz 1.15 a) Es gilt $xy = 0$ genau dann, wenn mindestens einer der Faktoren x oder y gleich 0 ist

b) Es gilt stets $(-x)y = -xy$ und $(-x)(-y) = xy$

Beweis: a) (I) Für alle x ist $0x = 0$. Denn es ist $0x + x = (0 + 1)x = 1x = x$. Addition von $(-x)$ liefert $0x = 0x + x - x = x - x = 0$.

(II) Sei $xy = 0$ und $x \neq 0$. Multiplikation mit $1/x$ liefert nach (I) $y = 1/x \cdot (xy) = 1/x \cdot 0 = 0$.

b) Es ist $0 = 0y = (x + (-x))y = xy + (-x)y$ und $0 = xy + (-xy)$. Damit ist sowohl $(-x)y$ als auch $-xy$ additives Inverses zu xy , also $(-x)y = -xy$.

Schließlich ist nach dem eben Bewiesenen $(-x)(-y) = -x(-y) = -(-xy)$ das additive Inverse zu $-xy$. Aber dies ist nach Definition gerade xy .

Eine weitere Formel, die allein aus diesen Rechenregeln folgt, ist die Binominalformel $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$. Die bisher aufgelisteten Rechenregeln, einschließlich der abgeleiteten, gelten nicht nur für die reellen Zahlen, sondern auch für die rationalen Zahlen, die komplexen Zahlen, die wir gleich einführen werden, und die Booleschen Zahlen \mathbb{K}_2 die ebenfalls sofort erklärt werden. Wir fassen diesen allgemeinen Aspekt in einem Begriff zusammen

Definition 1.16 Sei \mathbb{K} eine beliebige Menge, $+$: $\mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$ und \cdot : $\mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$ seien zwei Abbildungen, so daß die Rechenregeln der Addition, der Multiplikation und das Distributivgesetz gelten. Dann nennt man $(\mathbb{K}, +, \cdot)$ oder kürzer \mathbb{K} einen Körper.

\mathbb{R} und \mathbb{Q} sind also Körper. Es folgen die beiden nächstwichtigen Körper.

Der Körper der komplexen Zahlen

Auf $\mathbb{C} = \mathbb{R}^2 = \{(a, b) : a, b \in \mathbb{R}\}$ erklären wir die Addition komponentenweise.

$$(a, b) + (c, d) := (a + c, b + d)$$

Die Multiplikation erklären wir durch

$$(a, b)(c, d) = (ac - bd, ad + bc)$$

. Die Bedeutung von \mathbb{C} liegt darin, daß dort die Gleichung $z^2 + 1 = 0$ eine Lösung hat und daß \mathbb{R} nach Identifizierung in \mathbb{C} liegt.

Theorem 1.17 a) $(\mathbb{C}, +, \cdot)$ ist ein Körper, der Körper der komplexen Zahlen. Sein Nullelement ist $(0, 0)$, sein Einselement $(1, 0)$.

b) Für $i = (0, 1)$ gilt $i^2 = (-1, 0) = -(1, 0)$. i heißt imaginäre Einheit.

c) Durch $\mathbb{R} \ni x \mapsto (x, 0)$ wird \mathbb{R} auf $\{(x, 0) : x \in \mathbb{R}\}$ abgebildet. Dabei gilt

$$(x + y) \mapsto (x, 0) + (y, 0) \quad \text{und} \quad xy \mapsto (x, 0)(y, 0).$$

Man kann also \mathbb{R} mit der Teilmenge $\{(x, 0) : x \in \mathbb{R}\} \subset \mathbb{C}$ identifizieren. In diesem Sinne ist \mathbb{R} ein Teilkörper von \mathbb{C} .

Aufgabe: Bitte führen Sie den Beweis des Theorems. Für a) genügt es, die Rechenregeln der Addition, der Multiplikation und das Distributivgesetz nachzuprüfen. Für $(x, y) \neq (0, 0)$ ist das Multiplikative Inverse $(\frac{x}{\sqrt{x^2+y^2}}, \frac{-y}{\sqrt{x^2+y^2}})$. Beweisen Sie dies.

Für $z = (x, y) = x + iy$ heißt x Realteil $\Re(z)$ und y Imaginarteil $\Im(z)$. Die Zahl $\bar{z} = x - iy$ heißt die zu z konjugiert komplexe Zahl. Sie können \mathbb{C} mit der Ebene identifizieren. Dazu wählen Sie $\{(x, 0) : x \in \mathbb{R}\}$ als waagrechte, $\{(0, y) : y \in \mathbb{R}\}$ als senkrechte Achse. Dann ist \bar{z} der Punkt, den Sie aus z durch Spiegelung an der waagrechten Achse erhalten. Es gelten die folgenden Formeln für die Konjugation $z \mapsto \bar{z}$:

Satz 1.18 a) $\bar{z}_1 + \bar{z}_2 = \overline{(z_1 + z_2)}$.

b) $\bar{z}_1 \cdot \bar{z}_2 = \overline{z_1 z_2}$.

c) $\bar{z} \cdot z = \Re(z)^2 + \Im(z)^2$.

Anleitung zum Applet "komplexe Zahlen" Schauen Sie sich die Addition und Multiplikation für verschiedene Paare von komplexen Zahlen an.

Sie werden feststellen, daß die Multiplikation einer Drehstreckung entspricht. Dieser Sachverhalt wird sich später sehr einfach ergeben, wenn Sie noch eine andere Darstellung komplexer Zahlen kennengelernt haben.

Der Körper der Booleschen Zahlen

Auf der Menge $\mathbb{K}_2 = \{0, 1\}$ erklären wir die Addition und die Multiplikation durch folgenden Tabellen:

+	0	1
0	0	1
1	1	0

•	0	1
0	0	0
1	0	1

Satz 1.19 $(\mathbb{K}_2, +, \cdot)$ ist ein Körper.

Wenn Sie statt 0 ein *F* (falsch), statt 1 ein *W* (wahr) schreiben, entspricht der Addition das "Entweder-Oder", der Multiplikation das "und". Denken Sie über diesen Zusammenhang in Ruhe nach. \mathbb{K}_2 ist nicht umsonst so wichtig in der Informatik.

1.4.2 Die Ordnung der reellen Zahlen

Zwei reelle Zahlen können wir der Größe nach vergleichen. $x < y$ bedeutet "*x ist kleiner als y*", $x > y$ bedeutet "*x ist größer als y*" und $x \leq y$ bedeutet $(x < y)$ oder $(x = y)$, also "*x kleiner gleich y*".

Hierfür gelten die folgenden **Grundregeln der Ordnung**

- | | |
|---|------------------------------|
| 1. Für alle x ist $x \leq x$ | <i>Reflexivität</i> |
| 2. Ist $x \leq y$ und $y \leq x$, so ist $x = y$ | <i>Antisymmetrie</i> |
| 3. Ist $x \leq y$ und $y \leq z$, so ist $x \leq z$ | <i>Transitivität</i> |
| 4. Je zwei reelle Zahlen sind vergleichbar,
d.h. es gilt entweder $x \leq y$ oder $x > y$ | <i>Totalität der Ordnung</i> |
| 5. $x \leq y$ impliziert stets $a + x \leq a + y$ | <i>Translationsinvarianz</i> |
| 6. $x \leq y$ und $0 \leq a$ impliziert stets $ax \leq ay$ | <i>Dehnungsinvarianz</i> |
| 7. Ist $0 < x < y$, so gibt es eine natürliche Zahl $n \in \mathbb{N}$ mit
$nx > y$ ($nx = \underbrace{x + x + \dots + x}_{n \text{ mal}}$) | <i>Archimedes' Axiom</i> |

Aus diesen Grundregeln lassen sich alle weiteren Ungleichungen der reellen Analysis ableiten. Wir geben hier die vorläufig Wichtigsten an:

Satz 1.20 a) $x \leq y$ gilt genau dann, wenn $0 \leq y - x$ gilt, und dies gilt genau dann, wenn $-y \leq -x$ gilt.

b) Ist $x \leq y$ und $a \leq 0$, so ist $ay \leq ax$.

c) Ist $0 < x \leq y$, so ist $0 < \frac{1}{y} \leq \frac{1}{x}$.

d) Zu jedem $x > 0$ gibt es ein $n \in \mathbb{N}$ mit $\frac{1}{n} < x$.

e) Für alle y ist $y^2 \geq 0$.

f) Seien $0 \leq x, y$. Für beliebiges $n \in \mathbb{N}$ gilt $x^n < y^n$ genau dann, wenn $x < y$.

Beweis: a) Sei $x \leq y$. Addition von $a = -x$ liefert nach 5) $0 \leq y - x$. Addition von $a = -y$ hierauf liefert $-y \leq -x$. Wie erhält man hieraus $x \leq y$?

b) $x \leq y$ und $a \leq 0$ impliziert nach a) $0 \leq -a$, also nach 6) $-ax \leq -ay$ und hiermit nach a) $ay \leq ax$.

c) (i) Es ist $0 \leq 1$. Denn aus $1 \leq 0$ würde nach b) für $a = 1$ folgen $0 \leq 1 \cdot 1 = 1$, also nach 2) $0 = 1$, ein Widerspruch.

(ii) $0 \leq x$ impliziert $0 < \frac{1}{x}$. Denn aus $\frac{1}{x} \leq 0$ folgt nach 5) für $a = x$ einfach $1 = \frac{1}{x} \cdot x \leq 0$, ein Widerspruch zu (i).

(iii) $0 < x \leq y$ impliziert nun nach 6) ($a = \frac{1}{y} > 0$) $0 < \frac{x}{y} \leq 1$ und nochmals mit 6) ($a = \frac{1}{x}$) $0 < \frac{1}{y} \leq \frac{1}{x}$.

d) Ist $x \geq 1$, so wähle $n = 2$ und verwende c).

Ist $0 < x < 1$, so gibt es nach 7) ein $n \in \mathbb{N}$ mit $nx > 1$ also nach c) $x > \frac{1}{n}$.

e) Ist $y \geq 0$, so folgt die Behauptung aus 6) (mit $a = y$).

Ist $y < 0$, so ist $(-y) > 0$ nach a), also $y^2 = (-y)^2 > 0$ nach dem eben Bewiesenen.

f)(I) Sei $0 < x < y$. Die Aussage $A(n) : 0 < x^n < y^n$ ist richtig für $n = 1$. Sie sei für n wahr. Dann ist nach 6) ($a = x$) $0 < x^{n+1} < y^n x$ und ebenso ($a = y^n$) $0 < y^n x < y^{n+1}$. Also folgt $A(n + 1)$. Das Prinzip der vollständigen Induktion liefert die Behauptung.

(II) Es sei $0 < x^n < y^n$, also $0 < y^n - x^n$.

Es ist $y^n - x^n = (y - x) \sum_{k=0}^{n-1} y^k x^{n-1-k}$. Nach Voraussetzung ist dieser Ausdruck > 0 , also ist weder x noch y gleich 0. Dann ist $0 < \sum_{k=0}^{n-1} y^k x^{n-1-k} =: z > 0$, also (nach 6) mit $a = \frac{1}{z}$ $0 < y - x$.

Zwei etwas anspruchsvollere Ungleichungen werden häufig gebraucht

Satz 1.21 a) Für $n \in \mathbb{N}$ und $x > -1$ ist $(1 + x)^n \geq 1 + nx$ (Bernoullische Ungleichung)

b) Sei $0 < q < 1$ und $x \in \mathbb{R}$. Für alle natürlichen Zahlen n gelte $x \leq q^n$. Dann ist $x \leq 0$.

Beweis: a) (Vollständige Induktion). $n = 1$. Die Formel ist richtig.

Induktionsschritt: Annahme die Formel gilt für n .

$$\begin{aligned} (1 + x)^{n+1} &= (1 + x)^n \cdot (1 + x) \geq (1 + nx)(1 + x) \\ &= 1 + nx + x + \underbrace{nx^2}_{\geq 0} \\ &\geq 1 + (n + 1)x \end{aligned}$$

Also gilt dann die Aussage für $n + 1$.

b) Es ist $1 = q + \tau$ mit $\tau = 1 - q > 0$. Sei $n \in \mathbb{N}$ beliebig.

Also ist $1 = (q + \tau)^{n+1} = q^{n+1} + (n + 1)\tau q^n + \dots + \tau^{n+1} > (n + 1)\tau q^n$. Also ist $x \leq q^n < \frac{1}{(n+1)\tau}$. Dies liefert $\tau x < \frac{1}{n+1}$ für alle n und damit nach Satz 1.20d) $\tau x \leq 0$. Wegen $\tau > 0$ folgt hieraus $x \leq 0$.

Das **Maximum** $\max(a, b)$ zweier reeller Zahlen ist die größere der beiden. Genauer:

$$\max(a, b) = \begin{cases} a & a > b \\ b & a \leq b \end{cases}.$$

Entsprechend erklären wir das **Minimum** $\min(a, b)$ durch:

$$\min(a, b) = \begin{cases} a & a < b \\ b & a \geq b \end{cases}.$$

Schließlich wird der **Absolutbetrag** $|a|$ durch $|a| = \max(a, -a)$ erklärt. Aus den Regeln (Axiomen) für die Ordnung ergeben sich die folgenden Beziehungen:

Satz 1.22 *Es gilt*

- a) $\min(a, b) = -\max(-a, -b)$
- b) $\max(a, b) = -\min(-a, -b)$
- c) $a^+ := \max(a, 0) = \frac{1}{2}(|a| + a)$, $a^- := -\min(a, 0) = \frac{1}{2}(|a| - a)$
- d) $-|a| \leq a \leq |a|$
- e) Für alle a ist $|a|^2 = a^2$

Aufgabe: Beweisen Sie bitte diesen Satz.

Der Absolutbetrag mißt gewissermaßen den Abstand der Zahl a vom Nullpunkt. Allgemeiner erhalten wir den Abstand $d(a, b)$ der beiden Zahlen voneinander als $d(a, b) = |a - b|$. Es gelten die folgenden zentralen Formeln:

Satz 1.23 (Eigenschaften des Abstandes)

- a) $|a| = 0 \Leftrightarrow a = 0$ Definitheit
- b) $|ab| = |a||b|$ absolute Homogenität
- c) $|a + b| \leq |a| + |b|$ Dreiecksungleichung
- d) $|a - b| \leq |a - c| + |c - b|$ allgemeine Dreiecksungleichung
- e) $||a| - |b|| \leq |a - b| \leq |a| + |b|$ verallgemeinerte Dreiecksungleichung

Beweis: Wir zeigen nur c), d) und e)

c)

$$\begin{aligned} |a + b|^2 & \stackrel{\text{nach Satz 1.20 e)}}{=} (a + b)^2 = a^2 + b^2 + 2ab \\ & \stackrel{\text{nach Satz 1.20d)}}{\leq} a^2 + b^2 + 2|a||b| \\ & \stackrel{\text{nach Satz 1.20e)}}{=} |a|^2 + |b|^2 + 2|a||b| = (|a| + |b|)^2 \end{aligned}$$

Die Behauptung folgt nun aus Satz 1.20 f).

d) $|a - b| = |(a - c) + (c - b)| \leq |a - c| + |c - b|$ nach c)

e) Nach c) ist $|a| = |(a - b) + b| \underset{\text{nach c)}}{\leq} |a - b| + |b|$

also $|a| - |b| \leq |a - b|$.

Analog erhält man $|b| - |a| \leq |b - a|$.

Nach b) ist $|b - a| = |(-1)(a - b)| = |a - b| \cdot |-1| = |a - b|$.

Also erhalten Sie

$\|, |a| - |b|, | = \max(|a| - |b|, |b| - |a|) \leq |a - b|$, d.h. die erste Ungleichung.

Es ist nach c) $|a - b| \leq |a| + |-b| = |a| + |b|$ wegen $|-b| = |(-1)b| = |-1||b| = |b|$.

Exkurs: Abstandsmessung in den komplexen Zahlen

Sei $z = x + iy$ mit $x = \Re(z)$, $y = \Im(z)$. Zeichnen wir z in die Ebene mit rechtwinkligen Koordinatenachsen ein, so ergibt sich nach dem Satz des Pythagoras anschaulich die Entfernung zum Nullpunkt als $d(0, z)^2 = x^2 + y^2 = \bar{z}z$. Aus der Schule wissen Sie, daß zu jeder reellen Zahl $a \geq 0$ genau eine reelle Zahl $b \geq 0$ mit $b^2 = a$ existiert. b heißt positive Wurzel $+\sqrt{a}$ aus a .

Wir definieren den Absolutbetrag $|z|$ der komplexen Zahl z also einfach als

$$|z| = +\sqrt{\bar{z}z} = +\sqrt{\Re(z)^2 + \Im(z)^2}.$$

Wir setzen $d(z_1, z_2) = |z_1 - z_2|$ und erhalten einen zum vorigen Satz völlig entsprechenden Satz:

Satz 1.24 Eigenschaften des Abstandes komplexer Zahlen

- a) $|z| = 0 \Leftrightarrow z = 0$. Definitheit
- b) $|z_1 z_2| = |z_1| |z_2|$. absolute Homogenität
- c) $|z_1 + z_2| \leq |z_1| + |z_2|$. Dreiecksungleichung
- d) $|z_1 - z_2| \leq |z_1 - u| + |u - z_2|$. allgemeine Dreiecksungleichung
- e) $||z_1| - |z_2|| \leq |z_1 - z_2| \leq |z_1| + |z_2|$ verallgemeinerte Dreiecksungleichung

Aufgabe: Beweisen Sie diesen Satz.

Tip: Gehen Sie analog zum Beweis des vorigen Satzes vor. Beim Beweis von c) benutzen Sie die Formel

$$|z_1 + z_2|^2 = (\overline{z_1 + z_2})(z_1 + z_2) = (\bar{z}_1 + \bar{z}_2)(z_1 + z_2)$$

sowie die Formel $\bar{u} + u = 2\Re(u) \leq 2|u|$, die Sie natürlich beweisen müssen.

1.4.3 beschränkte Mengen und das Vollständigkeitsaxiom

Wir kehren zurück zur Ordnung in \mathbb{R} . Für das Folgende stellen wir uns \mathbb{R} als senkrechte Gerade vor und zwar so, daß $a < b$ einfach "a liegt unterhalb von b" bedeutet. Dann wird die folgende Definition besonders anschaulich:

Definition 1.25 Sei M eine beliebige, nichtleere Teilmenge von \mathbb{R} .

- a) M heißt **nach oben beschränkt**, wenn es ein $b \in \mathbb{R}$ gibt mit $x \leq b$ für alle x aus M . Jedes solche b heißt dann obere Schranke von M .
- b) M heißt **nach unten beschränkt**, wenn es ein $a \in \mathbb{R}$ mit $a \leq x$ für alle x aus M gibt. Jedes solche a heißt untere Schranke von M .
- c) M heißt **beschränkt**, wenn M nach oben und unten beschränkt ist.

b ist also eine obere (bzw. untere) Schranke von M , wenn ganz M unterhalb (bzw. oberhalb) von b liegt. Jede endliche Menge $M = \{a_1, \dots, a_n\}$ ist beschränkt. Erklären wir induktiv $\max(a_1, \dots, a_{k+1}) = \max(\max(a_1, \dots, a_k), a_{k+1})$ und entsprechend $\min(a_1, \dots, a_{k+1}) = \min(\min(a_1, \dots, a_k), a_{k+1})$ so ist $\min(a_1, \dots, a_n) = a$ eine untere Schranke und $\max(a_1, \dots, a_n) = b$ eine obere Schranke von M . $b = \max\{a_1, \dots, a_n\}$ ist die *kleinste* obere Schranke von $M = \{a_1, \dots, a_n\}$. Das heißt präzise:

- i) b ist obere Schranke von M
- ii) Ist c obere Schranke von M , so ist $c \geq b$.

Dies verallgemeinern wir auf beliebige Mengen:

Definition 1.26 Sei $\emptyset \neq M \subset \mathbb{R}$.

- a) $b \in \mathbb{R}$ heißt **obere Grenze oder Supremum** von M , in Zeichen: $b = \sup(M)$, wenn b obere Schranke von M ist und für jede andere obere Schranke c von M stets $c \geq b$ gilt.
- b) $a \in \mathbb{R}$ heißt **untere Grenze oder Infimum** von M , in Zeichen: $a = \inf(M)$, wenn a untere Schranke von M ist und für jede andere untere Schranke d von M stets $d \leq a$ gilt.

Beispiele:

1. $M = \{a_1, \dots, a_n\}$. Dann ist $\max\{a_1, \dots, a_n\} = \sup(M)$ und $\min\{a_1, \dots, a_n\} = \inf(M)$.
2. Sei $M = [0, 1[= \{x \in \mathbb{R} : 0 \leq x < 1\}$.
Dann ist $0 = \inf(M) \in M$ und $1 = \sup(M) \notin M$.
Die obere (bzw. untere) Grenze einer Menge M muß also *nicht* in M liegen.
3. Sei $M = \{r \in \mathbb{Q} : r^2 \leq 2\}$. M ist nicht leer ($0 \in M$). Aber existiert $\sup(M)$? Wenn wir wissen, daß $\sqrt{2}$ eine reelle Zahl ist, so ist leicht $\sup(M) = \sqrt{2}$ zu zeigen. Es ist $\sqrt{2} \notin M$, weil $\sqrt{2}$ keine rationale Zahl ist.
4. Zur Bestimmung der Fläche des Kreises mit Radius r berechnet man den Flächeninhalt $F(E_n)$ des eingeschriebenen regelmäßigen n -Ecks. Er ergibt sich zu $F(E_n) = r^2 \frac{n}{2} \sin(2\pi/n)$ und setzt dann Flächeninhalt des Kreises ist $F(r) = \sup\{F(E_n) : n \in \mathbb{N}\}$.

Damit wir sicher wie in Beispiel 3 oder 4 vorgehen können, formulieren wir das folgende Axiom:

Vollständigkeitsaxiom:

In \mathbb{R} existiert für jede nach oben beschränkte Menge die obere Grenze.

Wir fassen die Eigenschaften der reellen Zahlen zusammen. \mathbb{R} ist ein Körper, der eine Ordnung mit

den Eigenschaften 1-7 15, hat und das Vollständigkeitsaxiom erfüllt.

Bevor wir Konsequenzen aus dem Vollständigkeitsaxiom ziehen, brauchen wir noch einige Rechenregeln für die obere und untere Grenze. Dazu sei $\emptyset \neq M$. Dann setzen wir für $z \in \mathbb{R}$

$$z + M = \{z + x : x \in M\}, \quad zM = \{zx : x \in M\}$$

Satz 1.27 (Rechenregeln für sup und inf)

- a) Sei b eine obere Schranke von M . b ist genau dann obere Grenze von M , wenn es zu jedem $n \in \mathbb{N}$ ein $x \in M$ mit $b - \frac{1}{n} < x \leq b$ gibt.
- b) Sei a untere Schranke von M . a ist genau dann untere Grenze von M , wenn es zu jedem $n \in \mathbb{N}$ ein $x \in M$ mit $a \leq x < a + \frac{1}{n}$ gibt.
- c) Es ist

$$\begin{aligned} z + \sup(M) &= \sup(z + M) \\ z + \inf(M) &= \inf(z + M) \\ z \inf(M) &= \begin{cases} \inf(zM) : z \geq 0 \\ \sup(zM) : z < 0 \end{cases}. \end{aligned}$$

Aufgabe: Beweisen Sie bitte den Satz (benutzen Sie alle Rechenregeln für Ungleichungen).

Damit erhalten wir die folgenden einfachen Konsequenzen aus dem Vollständigkeitsaxiom:

Satz 1.28 a) In \mathbb{R} existiert zu jeder nach unten beschränkten Menge M die untere Grenze.

b) Für alle $x \in \mathbb{R}$ gilt

$$\sup\{r \in \mathbb{Q} : r < x\} = x = \inf\{r \in \mathbb{Q} : r > x\}$$

c) Zu jedem $x \in \mathbb{R}$ und zu jedem $n \in \mathbb{N}$ gibt es eine rationale Zahl $r \in \mathbb{Q}$ mit $|x - r| < \frac{1}{n}$.

c) besagt anschaulich, daß \mathbb{Q} dicht liegt in \mathbb{R} .

Beweis: a) Nach Voraussetzung ist $(-1) \cdot M = -M$ nach oben beschränkt. Mit den Rechenregeln für die obere und untere Grenze erhält man $\inf(M) = -\sup(-M)$.

b) Sei $x > 0$. Dann ist $0 \in M : \{r \in \mathbb{Q} : r < x\}$, M ist also nicht leer und nach oben beschränkt durch x ; also existiert $a = \sup(M)$. Aus $a < x$ würde $0 < x - a$ folgen, also $a + \frac{1}{n} < x$.

Nach Satz 1.27 gibt es ein $r \in M$ mit $a - \frac{1}{2n} < r \leq a$. Also ist $a + \frac{1}{2n} < r + \frac{1}{n} \leq a + \frac{1}{n} < x$. Da $r \in M$, ist r und damit $r + \frac{1}{n}$ aus \mathbb{Q} . Wegen $r + \frac{1}{n} < x$ ist also $r + \frac{1}{n} \in M$, und damit $r + \frac{1}{n} \leq a$, ein Widerspruch zu $a + \frac{1}{2n} < r + \frac{1}{n}$. Um $x = \inf\{r \in \mathbb{Q} : r > x\}$ zu zeigen, müssen wir zuerst zeigen, daß $\{r \in \mathbb{Q} : r > x\}$ nicht leer ist. Aber das folgt aus Archimedes Axiom: es gibt ein n mit $n1 > x$. Dann geht man analog zum ersten Teil des Beweises vor. Ist schließlich $x < 0$, so gilt der Satz für $-x$. Aus den Rechenregeln für inf und sup folgt er für x .

c) Sei $x \in \mathbb{R}$ und $n \in \mathbb{N}$ beliebig. Nach a) gibt es ein $r \in \mathbb{Q}$ mit $x - \frac{1}{n} < r \leq x$, also $|x - r| = x - r < \frac{1}{n}$.

Kapitel 2

Folgen und Reihen

2.1 Einführung

Reelle Zahlenfolgen begegnen uns unter anderem in den folgenden Zusammenhängen.

Folgen von Meßwerten

Zur Bestimmung der Erdbeschleunigungskonstanten kann man eine Bleikugel fester Masse aus einer bestimmten Höhe h auf die Erde fallen lassen und die Fallzeit bestimmen. Man erhält die Meßdaten t_1, t_2, t_3, \dots , im Idealfall also eine unendliche Folge von Meßwerten.

Algorithmen zur Berechnung von Zahlen

Der bekannteste Algorithmus ist der zur Bestimmung der Quadratwurzel \sqrt{a} aus einer positiven reellen Zahl a . Es handelt sich um eine rekursiv definierte Abbildung:

$$a_0 = a, \quad a_{n+1} = \frac{1}{2} \left(a_n + \frac{a}{a_n} \right).$$

Man erhält also z.B. durch einen Rechner nacheinander die Zahlen $a_0, a_1, a_2, a_3, \dots$. Sie nähern sich sehr schnell \sqrt{a} . Zum Beispiel ist für $a = 4$ (das wählen wir, weil wir $\sqrt{4}$ exakt kennen) $a_7 = 2$ auf 30 Stellen genau.

Entwicklung diskreter Systeme

Sei S ein biologisches System, von dem wir im Takt einer Zeiteinheit eine charakteristische Größe beobachten. Ist S zum Beispiel eine Bakterienkultur einer bestimmten Art enthaltende Nährlösung, so können wir nach jeder Sekunde die Zahl z der Bakterien bestimmen. Ist T die Tragkapazität der Nährlösung, d.h. die maximal mögliche Bakterienzahl in ihr, so ist der Quotient $x = z/T$ der Bruchteil der Tragkapazität, der gerade vorhanden ist. Der Mathematiker P. Verhulst (1804 - 1849) hat für die zeitliche Entwicklung solch eines Systems das folgende Modell vorgeschlagen:

$x_{n+1} = qx_n(1 - x_n)$. Hier ist q eine die Güte der Nährlösung beschreibende Konstante.

Bei festem Anfangswert x_0 erhält man nacheinander die Zahlen x_1, x_2, x_3, \dots , die zur Voraussage des Systems benutzt werden können. D.h. stimmt das Modell mit der Wirklichkeit überein und kennt man q und x_0 , so läßt sich die Zahl $x_n = z_n/T$ nach n Zeiteinheiten vorausberechnen und damit lassen sich Voraussagen treffen.

Anleitung zum SeqPlot-Applet Wir können uns die zeitliche Entwicklung für verschiedene q anschauen. Damit wir besser vergleichen können, wählen wir immer denselben Anfangswert $x_0 = 0.3$. Schauen

Sie sich die ersten 100 Schritte der Entwicklung an für die Werte $q = 2.5$, $q = 3.2$, $q = 3.5$, $q = 3.6$, $q = 4.0$. Vergleichen Sie die Entwicklungen. Wodurch unterscheiden sie sich?

Mathematisch gesehen ist eine Folge offenbar nichts anderes als eine Abbildung von \mathbb{N}_0 in \mathbb{R} .

Definition 2.1 Sei $k \in \mathbb{Z}$ und $A_k = \{n \in \mathbb{Z} : n \geq k\}$.

Eine Abbildung $a : A_k \rightarrow \mathbb{R}$, $n \mapsto a_n$ heißt bei k beginnende Folge reeller Zahlen, in Zeichen: $(a_n)_{n \geq k}$.

Folgen erhalten wir z.B. durch eine Formel, d.i. eine Rechenvorschrift zur Berechnung von a_n oder durch Rekursion (oder natürlich durch eine beliebige Zuordnungsvorschrift)

Beispiele: In den Beispielen ist $A_k = \mathbb{N}$ oder $A_k = \mathbb{N}_0$,

1. $\mathbb{N} \ni n \mapsto \frac{1}{n}$ (andere Schreibweise: $a_n = \frac{1}{n}$)
2. $\mathbb{N} \ni n \mapsto 2^{-n}$ ($a_n = 2^{-n}$)
3. $\mathbb{N} \ni n \mapsto (n+1)^2 \cdot 2^{-n}$ ($a_n = (n+1)^2 2^{-n}$)
4. $a_0 = c$, $a_{n+1} = \frac{1}{2}(a_n + \frac{c}{a_n})$ für fest gewähltes $c > 0$
5. $a_n = (1 + (-1)^n)$
6. $a_n = \sum_1^n 1/k$
7. $a_n = \sum_0^n 2^{-k}$
8. $a_n = \sum_0^n (-2)^{-k}$
9. $a_n = \sum_1^n (-1)^k/k$
10. $a_n = \sum_{k=0}^n 2^k$
11. $a_0 = 0.3$, $a_{n+1} = qa_n(1 - a_n)$

Wählen Sie $q = 2.5, 3.2, 3.5, 3.6, 4.0$

Anleitung zum SeqPlot-Applet Schauen Sie sich nacheinander die ersten hundert Glieder einer jeden Folge an (bei 10. bitte nur die ersten 10 Glieder). Sie beobachten verschiedenes Verhalten. Hier eine Liste möglicher Verhaltensweisen (s.a. [Wolff], S. 43)

1. Die Folge häuft sich an genau einem Punkt.
2. Die Folge häuft sich an mehreren Punkten.
3. Ein Teil der Folge läuft nach ∞ , d.h. aus dem oberen Bildrand hinaus (das kann man am Computer nicht besser zeigen. Präzisieren Sie, was gemeint ist). Sie müssen evtl. dazu noch viel mehr Glieder betrachten, um das festzustellen. Beweisen können Sie so etwas nicht durch Bilder.
4. Ein Teil der Folge läuft nach $-\infty$ (Präzisieren Sie das!).
5. Die Folge liegt ganz zwischen zwei festen Zahlen $c < d$.
6. Die Folge ist periodisch, das heißt, es gibt ein $p \geq 1$ mit $a_{n+p} = a_n$ für alle n .

7. Die Folge kommt jedem Punkt eines Intervalls $[c, d] = \{x : c \leq x \leq d\}$ beliebig nahe. Sie füllt das Intervall scheinbar aus.
8. Die Folge steigt an: $a_n \leq a_{n+1}$ für alle n .
9. Die Folge fällt: $a_n \geq a_{n+1}$ für alle n .

Checken Sie bei jeder Folge, welche der genannten Eigenschaften sie hat. Wenn Sie durchweg annehmen, daß die Folgen die zeitliche Entwicklung von Systemen beschreiben, spüren Sie, wie wichtig diese Eigenschaften und deren Beweis ist.

Häufungspunkte einer Folge

Wir haben in den Bildern gesehen, daß sich manche Folgen bei bestimmten Zahlen häufen. Wir präzisieren dies nun:

Sei $c \in \mathbb{R}$ und $\varepsilon > 0$. Dann nennen wir das Intervall $U(c, \varepsilon) =]c - \varepsilon, c + \varepsilon[= \{x : c - \varepsilon < x < c + \varepsilon\}$ eine (offene) ε -Umgebung von c . ε kann in den Anwendungen beliebig groß sein. *Es ist aber gut, wenn Sie sich ε meist als klein vorstellen.*

Definition 2.2 (Häufungspunkt einer Folge)

$c \in \mathbb{R}$ heißt Häufungspunkt der Folge $(a_n)_{n \geq k}$ wenn in jeder ε -Umgebung $U(c, \varepsilon)$

unendlich viele Glieder der Folge liegen, d.h. wenn zu jedem $\varepsilon > 0$ die Menge $A(\varepsilon) = \{n : a_n \in U(c, \varepsilon)\}$ unendlich ist.

Aufgaben:

1. Zeigen Sie, daß 0 Häufungspunkt der Folgen $(1/n)_{n \in \mathbb{N}}$, $((-1)^n/n)_{n \in \mathbb{N}}$, $(q^n)_{n \in \mathbb{N}}$ ($|q| < 1$) ist.
2. Bestimmen Sie die Häufungspunkte der Folgen $(1 + (-1)^n)_{n \in \mathbb{N}}$, $((-1)^n \cdot (1 + \frac{1}{n}))_{n \in \mathbb{N}}$, $(n \bmod 4)_{n \in \mathbb{N}}$. Beweisen Sie Ihre Aussagen.
3. Bestimmen Sie die Häufungspunkte von $(a_n)_{n \geq 0}$ mit $a_n = \sum_{k=0}^n (-2)^{-k}$.

Beweisen Sie Ihre Aussagen.

Wir nennen eine Folge $a = (a_n)_{n \geq k}$ **nach oben beschränkt** wenn ihr Bildbereich $\{a_n : n \geq k\}$ nach oben beschränkt ist. Entsprechend heißt $a = (a_n)_{n \geq k}$ **nach unten beschränkt**, wenn ihr Bildbereich nach unten beschränkt ist. Die Folge heißt **beschränkt**, wenn sie nach oben und unten beschränkt ist.

Aufgabe: Bestimmen Sie welche der obigen Beispielfolgen beschränkt ist, welche nach oben bzw. nach unten unbeschränkt sind.

Wenn Sie noch einmal zurückblicken, werden Sie im Bild zu der Ansicht kommen, daß jede beschränkte Folge $(a_n)_{n \geq k}$ mindestens einen Häufungspunkt hat. Daß dies kein Computer-Artefakt ist, werden wir nun beweisen.

Theorem 2.3 (Satz von Bolzano-Weierstraß)

Sei $a = (a_n)_{n \geq k}$ eine beschränkte Folge. Dann besitzt a mindestens einen Häufungspunkt.

Beweis: Für jedes feste $l \geq k$ sei $A_l = \{a_n : n \geq l\}$. A_k selbst ist der Bildbereich der Folge, also beschränkt. Es gibt also Zahlen $c < d$ mit $A_k \subset [c, d] = \{x : c \leq x \leq d\}$. Wegen $A_l \subset A_k$ für $l \geq k$ ist A_l auch beschränkt, es existiert also $\bar{a}_l := \sup(A_l)$ und es gilt $c \leq \bar{a}_l \leq d$. Damit ist die so erhaltene Folge $(\bar{a}_l)_{l \geq k}$ ebenfalls beschränkt. Es existiert also $x = \inf\{\bar{a}_l : l \geq k\}$.

Behauptung: x ist ein Häufungspunkt von $(a_n)_{n \geq k}$.

Beweis: Sei $\varepsilon > 0$ beliebig. Nach Definition von x gibt es ein $l_0 \geq k$ mit $x \leq \bar{a}_{l_0} < x + \varepsilon$. Nun ist $A_l \subset A_{l_0}$ für alle $l \geq l_0$. Daraus folgt $x \leq \bar{a}_l \leq \bar{a}_{l_0} < x + \varepsilon$ für alle $l \geq l_0$. Nach Definition von \bar{a}_l als obere Grenze folgt nun wieder die Existenz eines $n(l) \geq l$ mit

$x - \varepsilon \leq \bar{a}_l - \varepsilon < a_{n(l)} \leq \bar{a}_l$. Damit ist insgesamt $x - \varepsilon < a_{n(l)} < x + \varepsilon$, also ist

$a_{n(l)} \in A(\varepsilon) = \{n : a_n \in U(x, \varepsilon)\}$. Wegen $n(l) \geq l$ ist $\{a_{n(l)} : l \geq k\}$ unendlich, also ist $A(\varepsilon)$ unendlich. Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.

Lemma 2.4 Ist $(a_n)_{n \geq k}$ eine reelle Zahlenfolge mit $c \leq a_n \leq d$ für alle n , so liegt auch jeder Häufungspunkt x der Folge $(a_n)_{n \geq k}$ zwischen c und d .

Beweis: Wir erinnern an $U(x, \varepsilon) = \{y : x - \varepsilon < y < x + \varepsilon\}$. Sei x ein Häufungspunkt und $\varepsilon > 0$ beliebig. Dann ist $A(\varepsilon, x) = \{n : a_n \in U(x, \varepsilon)\}$ unendlich. Also ist $U(x, \varepsilon) \cap [c, d] \neq \emptyset$. Wäre das Lemma falsch, so wäre entweder $x < c$ oder $x > d$. Im ersten Fall wäre $U(x, \varepsilon) \cap [c, d] = \emptyset$ für $\varepsilon = \frac{c-x}{2}$, in zweiten Fall für $\frac{x-d}{2}$. Beide Male erhielten wir also einen Widerspruch.

Satz 2.5 Sei $(a_n)_{n \geq k}$ eine beschränkte Zahlenfolge. Dann ist die Menge H aller Häufungspunkte nicht leer und beschränkt. Ihre obere Grenze $\sup(H)$ ist selbst ein Häufungspunkt, der größte Häufungspunkt $\limsup_{n \geq k} a_n$, gelesen **Limes Superior** von (a_n) .

Die untere Grenze $\inf(H)$ ist ebenfalls ein Häufungspunkt, der kleinste Häufungspunkt $\liminf_{n \geq k} a_n$, gelesen **Limes Inferior** von (a_n) .

Beweis: Daß H nicht leer ist, folgt aus dem Satz von Bolzano-Weierstraß. Daß H beschränkt ist, folgt aus dem vorausgegangenen Lemma. Damit existieren untere und obere Grenze von H und wir müssen nur noch zeigen, daß $\sup(H)$ und $\inf(H)$ selbst Häufungspunkte sind.

Behauptung: $\sup(H) =: y$ ist ein Häufungspunkt von $(a_n)_{n \geq k}$.

Beweis: Sei $\varepsilon > 0$ beliebig. Wir müssen zeigen, daß $A(\varepsilon, y) = \{n \geq k : y - \varepsilon < a_n < y + \varepsilon\}$ unendlich ist. Zu $\eta = \varepsilon/2$ gibt es nach Definition der oberen Grenze ein $z \in H$ mit $y - \eta < z \leq y$. Da z ein Häufungspunkt ist, ist $A(\eta, z) = \{n \geq k : z - \eta < a_n < z + \eta\}$ unendlich. Ist nun aber $n \in A(\eta, z)$, so ist

$$\begin{aligned} |y - a_n| &\leq |y - z| + |z - a_n| \\ &< |y - (y - \eta)| + \eta \\ &= 2\eta \\ &= \varepsilon, \end{aligned}$$

also $n \in A(\varepsilon, y)$. Damit ist $A(\eta, z) \subset A(\varepsilon, y)$, also ist auch diese Menge unendlich.

Behauptung: $\inf(H) =: x$ ist Häufungspunkt von $(a_n)_{n \geq k}$.

Beweis: Entweder ganz analog zum vorausgegangenen Beweisteil, oder aber Sie benutzen, daß $-H = \{-z : z \in H\}$ die Menge der Häufungspunkte von $(-a_n)_{n \geq k}$ ist. Dann ist $\sup(-H) = -\inf(H)$ Häufungspunkt von $(-a_n)_{n \geq k}$, also ist $\inf(H)$ Häufungspunkt von $(a_n)_{n \geq k}$.

Wir werden diesen Satz bei der Bestimmung des Konvergenzradius einer Potenzreihe benötigen, siehe Theorem 2.20

2.2 Konvergente Folgen

Von mathematischer sowie von systemtheoretischer Seite her sind diejenigen Folgen von höchstem Interesse, die beschränkt sind und genau einen Häufungspunkt haben. Ein Algorithmus zur Berechnung einer Zahl c soll ja Zahlen a_n liefern, die genau dieser Zahl und keiner anderen immer näher kommen. Ist $(a_n)_{n \geq 0}$ die Beschreibung der zeitlichen Entwicklung eines Systems, so kann man die ferne Zukunft des Systems dann besonders gut voraussagen, wenn man aufgrund der Information über $(a_n)_{n \geq 0}$ weiß, daß die Glieder a_n genau einer festen Zahl c beliebig nahe kommen. Wir wollen das im nächsten Abschnitt präzisieren.

2.2.1 Grundbegriffe und Grundtatsachen

Definition 2.6 Eine Folge $(a_n)_{n \geq k}$ heißt konvergent gegen c , wenn die Folge beschränkt ist und c als einziger Häufungspunkt hat. Man schreibt dann $c = \lim_{n \rightarrow \infty} a_n$ und nennt c Grenzwert oder Limes der Folge $(a_n)_{n \geq k}$.

Anleitung zum SeqPlot-Applet Checken Sie alle vorausgegangenen Beispielfolgen auf Konvergenz. Versuchen Sie zu beweisen (nicht zu lange), daß Ihr Check stimmt.

Wir erhalten ein erstes Konvergenzkriterium.

Satz 2.7 Eine Folge $(a_n)_{n \geq k}$ konvergiert genau dann gegen c , wenn es zu jedem $\varepsilon > 0$ ein $n(\varepsilon)$ gibt mit $|c - a_n| < \varepsilon$ für alle $n \geq n(\varepsilon)$.

Beweis: (I) Es gelte $c = \lim_{n \rightarrow \infty} a_n$. Sei $\varepsilon > 0$ beliebig und $B(\varepsilon) = \{n : |c - a_n| \geq \varepsilon\}$. Ist $B(\varepsilon)$ endlich, so erfüllt $n(\varepsilon) = \max(B(\varepsilon)) + 1$ die behauptete Bedingung $|c - a_n| < \varepsilon$ für alle $n \geq n(\varepsilon)$. Angenommen $B(\varepsilon)$ wäre unendlich. Dann ist keine der Mengen $B_n(\varepsilon) = \{r \in B(\varepsilon) : r \geq n\}$ leer, also hat jede dieser Mengen ein erstes Element $r(n)$. Die neue Folge $n \mapsto a_{r(n)} =: b_n$ ist beschränkt, da $(a_n)_{n \geq k}$ als konvergente Folge beschränkt ist. Sie besitzt nach dem Satz von Bolzano-Weierstraß also einen Häufungspunkt d mit $d \neq c$. Denn aus $d = c$ würde folgen, daß $\{n : b_n \in U(\varepsilon, c)\} = \{n : a_{r(n)} \in U(\varepsilon, c)\}$ unendlich ist, aber es gilt $a_{r(n)} \notin U(\varepsilon, c)$ für alle n , ein Widerspruch. Da d Häufungspunkt ist ist $\{n : a_{r(n)} \in U(\varepsilon, d)\}$ unendlich, d ist also Häufungspunkt von $(a_n)_{n \geq k}$, ein Widerspruch dazu, daß c der einzige Häufungspunkt ist.

(II) Es gelte: Für alle $\varepsilon > 0$ gibt es $n(\varepsilon)$, so daß für alle $n \geq n(\varepsilon)$ stets $|c - a_n| < \varepsilon$, also $a_n \in U(c, \varepsilon)$ ist. Dann ist $(a_n)_{n \geq k}$ zunächst beschränkt. Denn für $\varepsilon = 1$ und alle $n \geq n(1)$ ist $|a_n| \leq |a_n - c| + |c| \leq 1 + |c|$. Für $n < n(1)$ ist $|a_n| \leq \max(|a_1|, \dots, |a_{n(1)}|) = b$. Insgesamt ist also $|a_n| \leq 1 + |c| + b$ für alle n .

Daß c ein Häufungspunkt von $(a_n)_{n \geq k}$ ist, ist klar.

Weiterhin ist c einziger Häufungspunkt. Denn angenommen $d \neq c$ wäre ein weiterer Häufungspunkt.

Dann ist für $\varepsilon = \frac{|c-d|}{3}$ die Menge $A(\varepsilon) = \{n : a_n \in U(\varepsilon, d)\}$ unendlich.

Sei $n(\varepsilon)$ so bestimmt, daß $|c - a_n| < \varepsilon$ für alle $n \geq n(\varepsilon)$. Da $A(\varepsilon)$ unendlich ist, gibt es ein $n \geq n(\varepsilon)$ aus $A(\varepsilon)$. Dann gilt $|c - d| \leq |c - a_n| + |a_n - d| < \varepsilon + \varepsilon = \frac{2}{3}|c - d|$, ein Widerspruch.

Anleitung zum SeqPlot-Applet Konvergenztest am Bildschirm

Im Folgenden können Sie sich ganz leicht davon überzeugen, ob eine Folge konvergiert oder nicht. Bestimmen Sie gegebenenfalls den Grenzwert c , zeichnen Sie zu $\varepsilon > 0$ $U(c, \varepsilon)$ und lesen Sie am Bildschirm ein zugehöriges $n(\varepsilon)$ ab. Prüfen Sie es gegebenenfalls mit dem von Ihnen rechnerisch bestimmten.

1. $a_n = 1/n^2$

2. $a_n = 2^{-n}$
3. $a_0 = 4, \quad a_n = \frac{1}{2}(a_n + \frac{4}{a_n})$
4. $a_0 = 2, \quad a_n = \frac{1}{2}a_n + \frac{1}{a_n}$
5. $a_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \dots + (-1)^{n+1}/n$
6. $a_n = \sum_1^n \frac{1}{k}$
7. $a_n = \sum_0^n 2^{-k}$
8. mehrere eigene Folgen.

Das obige Konvergenzkriterium ist im allgemeinen unhandlich. Mit den folgenden Rechenregeln erhalten wir sofort die Konvergenz sehr komplizierter Folgen. Ein einfacher Trick ist die Benutzung von Nullfolgen. Eine gegen 0 konvergente Folge heißt **Nullfolge**. Die wesentliche Erleichterung zur Berechnung konvergenter Folgen ist die Beobachtung:

- Lemma 2.8** a) Seien $(a_n)_{n \geq k}$ und $(b_n)_{n \geq k}$ Nullfolgen und $c, d \geq 0$. Ist $(x_n)_{n \geq k}$ eine Folge mit $|x_n| \leq c|a_n| + d|b_n|$, für alle n , so ist $(x_n)_{n \geq k}$ eine Nullfolge.
- b) Eine Folge $(a_n)_{n \geq k}$ konvergiert genau dann gegen c , wenn die Folge $n \mapsto |c - a_n|$ eine Nullfolge ist.

Beweis: a) Sei $\varepsilon > 0$ beliebig und $\eta = \frac{\varepsilon}{c+d+1}$. Zu diesem η gibt es $n_1(\eta)$ mit $|a_n| = |0 - a_n| < \eta$ für alle $n \geq n_1(\eta)$ und ein $n_2(\eta)$ mit $|b_n| = |0 - b_n| < \eta$ für alle $n \geq n_2(\eta)$. Für $n \geq n(\varepsilon) = \max(n_1(\eta), n_2(\eta))$ gilt dann

$$|0 - x_n| = |x_n| \leq c|a_n| + d|b_n| \leq (c+d)\eta = \frac{(c+d)\varepsilon}{c+d+1} < \varepsilon,$$

also gilt $0 = \lim x_n$.

b) gilt wegen $|0 - |c - a_n|| = |c - a_n|$.

Anleitung zum SeqPlot-Applet Sie erhalten aus dem Lemma sofort, daß die folgenden Beispiele Nullfolgen darstellen. Schauen Sie sich an.

1. $a_n = \frac{1}{n}$. Zu $\varepsilon > 0$ gibt es nach Satz 1.20 ein $n(\varepsilon)$ mit $\frac{1}{n(\varepsilon)} < \varepsilon$. Dann ist für $n \geq n(\varepsilon)$ stets $0 < \frac{1}{n} \leq \frac{1}{n(\varepsilon)} < \varepsilon$.
2. $a_n = \frac{1}{n^3}$. Es ist $\frac{1}{n^3} < \frac{1}{n^2} < \frac{1}{n}$.
3. $a_n = q^n$ mit $0 \leq |q| < 1$, vergleiche Satz 1.21 b).

Satz 2.9 (Rechenregeln für konvergente Folgen)

- a) Ist $c = \lim a_n$, so ist $|c| = \lim |a_n|$
- b) Ist $c = \lim a_n$ und $d = \lim b_n$, so ist $c + d = \lim(a_n + b_n)$
- c) Ist $c = \lim a_n$ und $d = \lim b_n$, so ist $cd = \lim(a_n b_n)$

d) Ist $c = \lim a_n$, $d = \lim b_n \neq 0$ und sind alle $b_n \neq 0$, so ist $\frac{c}{d} = \lim \frac{a_n}{b_n}$

e) Ist $c = \lim a_n$, $d = \lim b_n$ und $a_n \leq b_n$ für alle n , so ist $c \leq d$.

Beweis: a) Es ist $|c| - |a_n| \leq |c - a_n| =: b_n$. (b_n) ist Nullfolge, also folgt der Satz.

b) Es ist

$$\begin{aligned} |(c+d) - (a_n + b_n)| &= |(c - a_n) + (d - b_n)| \\ &\leq |c - a_n| + |d - b_n|. \end{aligned}$$

Das Lemma liefert die Behauptung.

c) Es ist

$$\begin{aligned} |cd - a_n b_n| &= |(cd - a_n d) + (a_n d - a_n b_n)| \\ &\leq |c - a_n| \cdot |d| + |a_n| |d - b_n| \end{aligned}$$

Als konvergente Folge ist (a_n) beschränkt, es gibt also a mit $|a_n| \leq a$ für alle n . Dann ist

$$|cd - a_n b_n| \leq |d| |c - a_n| + a |d - b_n|$$

und der Satz folgt aus dem Lemma.

d) Es ist $|\frac{1}{d} - \frac{1}{b_n}| = \frac{1}{|d||b_n|} |b_n - d|$.

Zu $\varepsilon = \frac{|d|}{2} > 0$ gibt es $n(\varepsilon)$ mit $|d - b_n| < \varepsilon$ für alle $n \geq n(\varepsilon)$. Das bedeutet $-\frac{|d|}{2} < d - b_n < \frac{|d|}{2}$, also $|b_n| < \frac{|d|}{2}$ und damit $\frac{1}{|d||b_n|} \geq \frac{2}{|d|^2}$.

Das Lemma liefert $\lim \frac{1}{b_n} = \frac{1}{d}$. Der Rest folgt aus c).

e) Nach b) und c) ist $d - c = \lim(b_n - a_n)$.

Wegen $b_n - a_n \geq 0$ ist auch der Häufungspunkt $d - c \geq 0$ (s. Lemma 2.4).

Aufgaben: Bestimmen Sie durch geschicktes Zerlegen in einfachere Folgen die Grenzwerte der folgenden Beispiele:

1. $a_n = (4n^3 - (-1)^n n^2) / (5n + 2n^3)$

2. $a_n = \frac{(n^3 - 5n)^4 - n^{12}}{n^{11}}$

3. $a_n = \binom{n}{5} 2^{-n}$

2.2.2 Cauchys Konvergenzkriterium

Anleitung zum SeqPlot-Applet Schauen Sie sich die unten stehenden Folgen an und untersuchen Sie, welche wohl konvergieren, welche nicht. Was ist bei den konvergenten Folgen der Grenzwert?

Beispiele:

1. $a_n = \sum_{k=0}^n 2^{-k}$

2. $a_n = \sum_{k=1}^n 1/k^2$

3. $a_n = \sum_{k=1}^n (-1)^n / k$

4. $a_n = \sum_{k=1}^n 1/k!$

$$5. a_n = \sum_{k=2}^n \frac{1}{k(k-1)}$$

Obwohl wir bei den Folgen 2, 3, 4 den Verdacht haben, sie würden konvergieren, können wir es schon deshalb nicht beweisen, weil wir den Grenzwert nicht kennen. Wir können also keines der bisherigen Konvergenzkriterien (Häufungspunkt- Kriterium bzw. $\varepsilon - n(\varepsilon)$ -Kriterium) anwenden.

Wir sehen aber, daß die Folgenglieder mit wachsendem Index immer näher zusammenrücken. So ist das folgende Konvergenzkriterium, für das wir den Grenzwert *nicht* kennen müssen, fast offensichtlich:

Theorem 2.10 (Cauchys Konvergenzkriterium) Sei $(a_n)_{n \geq k}$ eine reelle Zahlenfolge. Die beiden folgenden Aussagen sind äquivalent.

a) $(a_n)_{n \geq k}$ ist konvergent.

b) Zu jedem $\varepsilon > 0$ gibt es ein $n(\varepsilon) \in \mathbb{N}$, so daß für alle $m, n \geq n(\varepsilon)$ stets $|a_m - a_n| < \varepsilon$ ist.

Beweis: a) \Rightarrow b): Sei $c = \lim_{n \rightarrow \infty} a_n$ und $\varepsilon > 0$ beliebig vorgegeben. Zu $\eta = \varepsilon/2$ gibt es ein $n(\eta)$ mit $|c - a_n| < \eta$ für alle $n \geq n(\eta)$. Setze $n(\varepsilon) = n(\eta)$. Seien $m, n \geq n(\varepsilon)$. Dann ist nach der Dreiecksungleichung

$$|a_m - a_n| \leq |a_m - c| + |c - a_n| < \eta + \eta = \varepsilon.$$

b) \Rightarrow a) Zunächst gibt es zu $\varepsilon = 1$ ein $n(1)$ mit $|a_m - a_n| < 1$ für alle $m, n \geq n(1)$, also für alle $n \geq n(1)$ $|a_n| \leq |a_n - a_{n(1)}| + |a_{n(1)}| < 1 + |a_{n(1)}|$. Also ist die ganze Folge beschränkt durch $1 + |a_{n(1)}| + |a_{n(1)-1}| + \dots + |a_k|$ (k : der erste Index der Folge). Damit hat die Folge nach dem Theorem 2.3 von Bolzano-Weierstraß einen Häufungspunkt c . Es ist der *einzig*e Häufungspunkt. Denn angenommen $d \neq c$ ist ein weiterer Häufungspunkt. Zu $\varepsilon = \frac{d-c}{4}$ gibt es ein $n(\varepsilon)$, mit $|a_m - a_n| < \varepsilon$ für alle $m, n \geq n(\varepsilon)$. Da d und c Häufungspunkte sind, sind die Mengen $A(\varepsilon, c) = \{n : |c - a_n| < \varepsilon\}$ und $A(\varepsilon, d) = \{n : |d - a_n| < \varepsilon\}$ unendlich. Es gibt also $r \geq n(\varepsilon)$ mit $|c - a_r| < \varepsilon$ und $s \geq n(\varepsilon)$ mit $|d - a_s| < \varepsilon$, d.h.

$$\begin{aligned} |d - c| &= |d - a_s + a_s - a_r + a_r - c| \\ &\leq |d - a_s| + |a_s - a_r| + |a_r - c| \\ &< \varepsilon + \varepsilon + \varepsilon = \frac{3}{4} \cdot |d - c| \end{aligned}$$

ein offensichtlicher Widerspruch.

Aufgabe: Untersuchen Sie mit diesem Kriterium, welche der vorangegangenen Beispielfolgen konvergiert, welche nicht.

2.2.3 Monotone Folgen

Eine reelle Zahlenfolge $(a_n)_{n \geq k}$ heißt **monoton wachsend** wenn für alle Indizes n stets $a_n \leq a_{n+1}$ ist, also $a_k \leq a_{k+1} \leq a_{k+2} \leq \dots$ gilt. Sie heißt **streng monoton wachsend**, wenn stets $a_n < a_{n+1}$ gilt. Sie heißt **monoton fallend**, wenn stets $a_n \geq a_{n+1}$ gilt und **streng monoton fallend**, wenn stets $a_n > a_{n+1}$ gilt. Eine Folge heißt **monoton** wenn sie entweder monoton wachsend oder monoton fallend ist.

Anleitung zum SeqPlot-Applet Schauen Sie sich die folgenden Beispiele an. Welche Folge ist monoton wachsend, welche monoton fallend, welche nichts von beiden? Beweisen Sie Ihre Vermutungen! Welche der monotonen Folgen konvergieren?

Beispiele:

1. $a_n = \frac{1}{n}$.
2. $a_n = (-1)^n/n$.
3. $a_n = (1 + \frac{1}{n})^n$.
4. $a_n = \frac{(1-q^{n+1})}{1-q}$ für $q = \frac{1}{2}$ bzw. für $q = (-\frac{1}{2})$.
5. $a_0 = 5, \quad a_{n+1} = \frac{a_n}{2} + \frac{2}{a_n}$.
6. $a_n = \sum_{k=0}^n 1/k!$
7. $a_n = n^2$.

Monotone Folgen sind besonders einfach auf Konvergenz hin zu untersuchen.

Satz 2.11 *Eine monotone Folge $(a_n)_{n \geq k}$ ist entweder unbeschränkt oder konvergent. In diesem Fall konvergiert sie gegen*

$$\left\{ \begin{array}{l} \sup\{a_n : n \geq k\} \\ \inf\{a_n : n \geq k\} \end{array} \right\} \text{ wenn sie } \left\{ \begin{array}{l} \text{monoton wachsend} \\ \text{monoton fallend} \end{array} \right\} \text{ ist.}$$

Beweis: Sei $(a_n)_{n \geq k}$ monoton wachsend und beschränkt. Dann existiert nach dem Vollständigkeitsaxiom 19 die obere Grenze $c = \sup\{a_n : n \geq k\}$. Weil c obere Grenze ist, gibt es zu jedem $r \in \mathbb{N}$ ein $n(r)$ mit $c - \frac{1}{r} < a_{n(r)} \leq c$. Aber da die Folge monoton wachsend ist, gilt für alle $n \geq n(r)$ stets $c - \frac{1}{r} < a_{n(r)} \leq a_n \leq c$, also $|c - a_n| = c - a_n < c - (c - \frac{1}{r}) = \frac{1}{r}$. Hieraus folgt $c = \lim a_n$.

Ist $(a_n)_{n \geq k}$ monoton fallend, so ist $(-a_n)_{n \geq k}$ monoton wachsend und der Satz folgt aus den Rechenregeln für sup und inf und für konvergente Folgen.

2.3 Teilfolgen

Anleitung zum SeqPlot-Applet Schauen Sie sich einmal die Folge $((-1)^n(1 + \frac{1}{n}))_{n \geq 1}$ an. Sie hat zwei Häufungspunkte. Sehen Sie sich im Bild die Glieder mit *geradem* Index an, also die Folge $(1 + \frac{1}{2n})_{n \geq 1}$ und dann die Glieder mit *ungeradem* Index, also die Folge $(-1 + \frac{1}{2n-1})_{n \geq 1}$. Eine nicht konvergente Folge kann also "Teilfolgen" haben, die konvergieren.

Wir präzisieren den Begriff der Teilfolge.

Definition 2.12 (Teilfolge)

Teilfolge] Sei $(a_n)_{n \geq k}$ eine Folge und $g : \{n : n \geq k\} \rightarrow \{n : n \geq k\}$ eine streng monoton wachsende Folge, d.h. es gelte $g(n) < g(n+1)$ für alle n . Dann heißt die Folge $(a_{g(n)})_{n \geq k}$ **Teilfolge** von $(a_n)_{n \geq k}$.

Es handelt sich genau genommen um die Hintereinanderausführung der beiden Abbildungen a und g . Die Teilfolge ist die Abbildung $a \circ g : \{n : n \geq k\} \rightarrow \mathbb{R}, \quad n \mapsto a_{g(n)}$.

Aufgabe: Bestimmen Sie in den folgenden Beispielen konvergente Teilfolgen:

1. $a_n = (1 + \frac{(-1)^n}{n})^n$
2. $a_n = n \bmod 3$

$$3. a_n = \left(1 + \frac{n \bmod 3}{n}\right)^n$$

Häufungspunkte einer Folge kann man durch Teilfolgen erreichen.

Satz 2.13 Sei c ein Häufungspunkt der Folge $(a_n)_{n \geq k}$. Dann gibt es eine Teilfolge $(a_{g(n)})_{n \geq k}$, die gegen c konvergiert.

Beweis: Wir setzen wieder $A_l(\varepsilon) = \{n \geq l : |c - a_n| < \varepsilon\}$. Da c ein Häufungspunkt der Folge $(a_n)_{n \geq k}$ ist, ist kein $A_l(\varepsilon)$ leer. Wir definieren nun induktiv

$g(k) = \min(A_k(2^{-k}))$, das ist das kleinste Element von $A_k(2^{-k})$

$g(n+1) = \min(A_{g(n)+1}(2^{-n}))$, also das kleinste Element von $A_{g(n)+1}(2^{-n})$. Wir erhalten $g(n+1) \geq g(n) + 1$, also ist g streng monoton wachsend. Ferner ist $|c - a_{g(n)}| < 2^{-n}$, also $((c - a_{g(n)}))_{n \geq k}$ eine Nullfolge. Das beweist $c = \lim_{n \rightarrow \infty} a_{g(n)}$.

Damit erhalten wir die folgende Variante des Satzes von Bolzano-Weierstraß (siehe Theorem 2.3):

Theorem 2.14 Jede beschränkte Folge besitzt eine konvergente Teilfolge.

Aufgabe: Finden Sie selbst den Beweis unter Benutzung des Satzes von Bolzano-Weierstraß Theorem 2.3

2.4 Unendliche Reihen

In der Schule haben Sie die geometrische Reihe kennengelernt (vielleicht im Zusammenhang mit dem Problem "Achilles und die Schildkröte"). Sie haben die Gleichung $\sum_{k=0}^{\infty} 1/2^k = 2$ oder sogar die Gleichung $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ für $|x| < 1$ kennengelernt. Wiederholen Sie, was damit gemeint war.

Definition 2.15 (Unendliche Reihen)

a) Sei $(a_n)_{n \geq k}$ eine reelle Zahlenfolge. Dann heißt die durch $s_n = \sum_{l=k}^n a_l$ gegebene Folge $(s_n)_{n \geq k}$ **unendliche Reihe**. Statt $(s_n)_{n \geq k}$ schreibt man $\sum_{l=k}^{\infty} a_l$. Die Folge $(a_n)_{n \geq k}$ heißt die Folge der Glieder der unendlichen Reihe.

b) Ist die Folge (s_n) konvergent mit $\lim_{n \rightarrow \infty} s_n = c$, so schreibt man $\sum_{l=k}^{\infty} a_l = c$.

Man verwendet also leider dasselbe Symbol $\sum_{n=k}^{\infty} a_n$ in zwei verschiedenen Bedeutungen, nämlich einmal als Symbol für die Folge der Teilsummen und einmal als Grenzwert dieser Folge.

Bei unendlichen Reihen eignet sich das Cauchy-Kriterium (siehe Theorem 2.10) besonders zum Nachweis der Konvergenz. Dazu berechnen wir die Differenz $|s_m - s_n|$, wobei wir ohne Beschränkung der Allgemeinheit $m > n$, also $m = n + p$ annehmen. Dann ist

$$|s_m - s_n| = \left| \sum_{l=l}^{n+p} a_l - \sum_{l=k}^n a_l \right| = \left| \sum_{l=n+1}^{n+p} a_l \right|.$$

Damit erhalten wir den Satz

Satz 2.16 a) Die unendliche Reihe $\sum_{l=k}^{\infty} a_l$ ist genau dann konvergent, wenn zu jedem $\varepsilon > 0$ ein $n(\varepsilon)$ existiert mit $|\sum_{l=n+1}^{n+p} a_l| < \varepsilon$ für alle $n \geq n(\varepsilon)$ und alle $p \geq 1$.

b) Ist die Reihe $\sum_{l=k}^{\infty} a_l$ konvergent, so ist die Folge $(a_n)_{n \geq k}$ eine Nullfolge.

Aufgabe: Führen Sie den Beweis aus. Beachten Sie, daß b) aus a) folgt, indem man $p = 1$ wählt.

Anleitung zum SeqPlot-Applet Schauen Sie sich die folgenden Reihen an. Welche konvergieren, welche nicht?

1. $\sum_{k=0}^{\infty} 1/2^k$
2. $\sum_{k=1}^{\infty} (-1)^k/k$
3. $\sum_{k=1}^{\infty} 1/k^2$
4. $\sum_{k=2}^{\infty} \frac{1}{k(k-1)}$
5. $\sum_{k=1}^{\infty} 1/k$
6. $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ für $x = 1, -1$ und 2 .

Beispiel: Geometrische Reihe

Die Reihe $\sum_{k=0}^{\infty} x^k$ konvergiert für $|x| < 1$ und divergiert für $|x| \geq 1$.

Denn es ist für $x \neq 1$ $s_n = \sum_{k=0}^n x^k = \frac{1-x^{n+1}}{1-x}$ und $(x^n)_{n \geq 0}$ ist genau dann eine Nullfolge, wenn $|x| < 1$ gilt.

Absolut konvergente Reihen

Die wichtigsten Reihen sind absolut konvergent. Genauer gilt:

Satz 2.17 Sei $(a_n)_{n \geq k}$ eine beliebige Folge reeller Zahlen, so daß die Reihe $\sum_{l=k}^{\infty} |a_l|$ mit den Absolutbeträgen $|a_l|$ als Glieder konvergiert. Dann konvergiert auch die Reihe $\sum_{l=k}^{\infty} a_k$ und es gilt $|\sum_{l=k}^{\infty} a_l| \leq \sum_{l=k}^{\infty} |a_k|$.

Ist $\sum_{l=k}^{\infty} |a_k|$ konvergent, so heißt die Reihe $\sum_{l=k}^{\infty} a_k$ **absolut konvergent**.

Beweis: Sei $\varepsilon > 0$. Nach Voraussetzung konvergiert $\sum_{l=k}^{\infty} |a_l|$, also gibt es nach dem Cauchy Kriterium Satz 2.16 ein $n(\varepsilon)$ mit $\sum_{l=n+1}^{n+p} |a_l| < \varepsilon$ für alle $n \geq n(\varepsilon)$ und $p \geq 1$. Aus der Dreiecksungleichung erhalten wir für dieselben n und p $|\sum_{l=n+1}^{n+p} a_l| \leq \sum_{l=n+1}^{n+p} |a_l| < \varepsilon$, also wiederum nach dem Cauchy Kriterium die Konvergenz der Reihe $\sum_{l=k}^{\infty} a_l$. Für alle n gilt $|\sum_{l=k}^n a_l| \leq \sum_{l=k}^n |a_l| \leq \sum_{l=k}^{\infty} |a_l|$, also erhält man nach Satz 2.9 a) $|\sum_{l=k}^{\infty} a_l| = \lim_{n \rightarrow \infty} |\sum_{l=k}^n a_l| = \lim_{n \rightarrow \infty} |\sum_{l=k}^n a_l| \leq \sum_{l=k}^{\infty} |a_l|$.

Aufgaben: Welche der folgenden Reihen sind absolut konvergent?

1. $\sum_{n=1}^{\infty} (-1)^n/n^2$
2. $\sum_{n=0}^{\infty} (-1)^n/n!$
3. $\sum_{n=1}^{\infty} (-1)^n/n$

$$4. \sum_{n=0}^{\infty} (-q)^n \qquad 0 < q < 1.$$

Anleitung zum SeqPlot-Applet Schauen Sie sich die Reihen $\sum_{n=k}^{\infty} a_n$ und $\sum_{n=k}^{\infty} |a_n|$ bei jeder Aufgabe am Bildschirm an.

Mit dem Begriff der absolut konvergenten Reihe erhalten wir sofort die folgenden Konvergenzkriterien für Reihen:

Satz 2.18 Sei $\sum_{k=0}^{\infty} c_k$ eine beliebige Reihe.

- (Wurzelkriterium)** Gilt für die Glieder von einem n_0 an stets $\sqrt[n]{|c_n|} \leq q < 1$ für ein festes q , so konvergiert die Reihe $\sum_0^{\infty} c_k$ absolut. Die Bedingung besagt $\limsup \sqrt[n]{|c_n|} < 1$.
- (Quotientenkriterium)** Gilt für die Glieder von einem n_0 an stets $|\frac{c_{n+1}}{c_n}| \leq q < 1$ für ein festes q , so konvergiert die Reihe $\sum_0^{\infty} c_k$ absolut.

Beweis: a) Es ist für $n \geq n_0$ stets $|c_n| \leq q^n$, also

$$\sum_{k=0}^n |c_k| \leq \sum_{k=0}^{n_0-1} |c_k| + \sum_{k=n_0}^n q^k \leq \sum_{k=0}^{n_0-1} |c_k| + \frac{1}{1-q},$$

woraus die Behauptung folgt

b) Für $n \geq n_0$ und $p \in \mathbb{N}$ ist

$$\left| \frac{c_{n+p}}{c_n} \right| = \left| \frac{c_{n+p}}{c_{n+p-1}} \cdot \frac{c_{n+p-1}}{c_{n+p-2}} \cdots \frac{c_{n+1}}{c_n} \right| \leq q^p$$

und damit folgt $|c_n| \leq |c_{n_0}| \cdot q^{-n_0} q^n$ für alle $n \geq n_0$.

Das ergibt für solche n

$$\begin{aligned} \sum_{k=0}^n |c_k| &\leq \sum_{k=0}^{n_0-1} |c_k| + |c_{n_0}| \cdot q^{-n_0} \sum_{k=n_0}^n q^k \\ &\leq \sum_{k=0}^{n_0-1} |c_k| + |c_{n_0}| q^{-n_0} \cdot \frac{1}{1-q}, \end{aligned}$$

woraus die Behauptung folgt.

Bitte lösen Sie unbedingt die folgenden Aufgaben, um die Konvergenzkriterien zu trainieren:

Aufgaben: Zeigen Sie, daß die folgenden Reihen absolut konvergieren.

- $\sum_{k=0}^{\infty} \frac{x^k}{k!}$.
- Sei $(b_n)_{n \geq 0}$ eine beschränkte Zahlenfolge und $|x| < 1$. Betrachten Sie die Reihe $\sum_{k=0}^{\infty} b_k x^k$.
- Es gelte $\lim \sqrt[n]{|a_n|} = 0$.
Dann konvergiert die Reihe $\sum_{k=0}^{\infty} a_k x^k$ für alle x absolut.
- Es sei die Folge $(\sqrt[n]{|a_n|})_{n \geq 0}$ beschränkt.
Dann konvergiert $\sum_{k=0}^{\infty} a_k x^k$ für alle x mit $|x| < \frac{1}{\limsup_n \sqrt[n]{|a_n|}}$ absolut ($\limsup \sqrt[n]{|a_n|} > 0$ wird vorausgesetzt).

Bevor wir die wichtige Klasse der Potenzreihen behandeln, wollen wir sehen, wie man Reihen miteinander multipliziert.

Theorem 2.19 (Cauchy-Produkt)

Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergente Reihen. Für jedes n sei $c_n = \sum_{k=0}^n a_k b_{n-k}$. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} c_n$ absolut und es gilt

$$\sum_{n=0}^{\infty} c_n = \sum_{k=0}^{\infty} a_k \cdot \sum_{l=0}^{\infty} b_l$$

Beweis: siehe [Wolff], S. 67/68.

Potenzreihen

Sei $(a_n)_{n \geq 0}$ eine beliebige reelle Zahlenfolge. Dann heißt die Reihe $\sum_{n=0}^{\infty} a_n x^n$ **Potenzreihe**. Es gilt der folgende zentrale Satz.

Theorem 2.20 Sei $(a_n)_{n \geq 0}$ eine beliebige reelle Zahlenfolge. Sei

$$R = \left\{ \begin{array}{ll} 0, & \text{falls } (\sqrt[n]{|a_n|})_{n \geq 0} \text{ unbeschränkt ist} \\ \infty, & \text{falls } \lim \sqrt[n]{|a_n|} = 0 \\ \frac{1}{\limsup \sqrt[n]{|a_n|}}, & \text{sonst} \end{array} \right\}$$

Dann gilt:

- a) Für alle $x \in \mathbb{R}$ mit $|x| < R$ konvergiert die Reihe $\sum_{n=0}^{\infty} a_n x^n$ absolut.
- b) Für kein $x \in \mathbb{R}$ mit $|x| > R$ konvergiert die Reihe. R heißt **Konvergenzradius** der Reihe.

Für den Fall $|x| = R$ lassen sich allgemein keine Aussagen machen.

Beweis: (I) Ist $\lim \sqrt[n]{|a_n|} = 0$, so konvergiert die Potenzreihe für alle $|x| < \infty$ absolut, (siehe Aufgabe 3 auf S. 32) für diesen Fall ist der Satz also bewiesen.

(II) Sei jetzt $\sqrt[n]{|a_n|}_{n \geq 0}$ beschränkt, aber keine Nullfolge. Dann ist $r := \limsup \sqrt[n]{|a_n|} > 0$. Sei $x \in \mathbb{R}$ mit $|x| < \frac{1}{r}$. Dann ist $\limsup \sqrt[n]{|a_n x^n|} = r|x| < 1$, also ist die Reihe $\sum_{n=0}^{\infty} a_n x^n$ nach dem Wurzelkriterium absolut konvergent.

(III) Sei jetzt $|x| > R$.

(i) Ist $R = 0$, so ist die Folge $(\sqrt[n]{|a_n|})_{n \geq 0}$ unbeschränkt. Wir zeigen, daß für $|x| > 0$ die Folge $(a_n x^n)_{n \geq 0}$ keine Nullfolge und daher $\sum_{n=0}^{\infty} a_n x^n$ nach dem Cauchy Kriterium Satz 2.16 nicht konvergent ist. Wäre $(a_n x^n)_{n \geq 0}$ eine Nullfolge, also $|a_n x^n| < 1/2$ für $n \geq n(1/2)$, so wäre für $n \geq n(1/2)$

$$\sqrt[n]{|a_n|} = \frac{1}{|x|} \cdot \sqrt[n]{|a_n x^n|} < \frac{1}{|x|} \sqrt[n]{1/2} \leq 1,$$

also $(\sqrt[n]{|a_n|})_{n \geq 0}$ beschränkt, ein Widerspruch.

(ii) Ist $R > 0$, so ist wegen $|x| > R$ $r = \limsup \sqrt[n]{|a_n|} > 0$ und aus $|x| > \frac{1}{r}$ folgt $\limsup \sqrt[n]{|a_n x^n|} = |x|r > 1$, damit ist $(a_n x^n)_{n \geq 0}$ keine Nullfolge, die Reihe $\sum_{n=0}^{\infty} a_n x^n$ also wieder nicht konvergent.

Korollar 2.21 Seien $\sum_0^\infty a_n x^n$ und $\sum_0^\infty b_n x^n$ Potenzreihen mit Konvergenzradien R bzw. S . Dann gilt für $|x| < \min(R, S)$ stets

$$\sum_0^\infty a_n x^n \cdot \sum_0^\infty b_n x^n = \sum_{n=0}^\infty \left(\sum_{k=0}^n a_k b_{n-k} \right) x^n$$

Der Konvergenzradius der rechts stehenden **Produktreihe** ist also größer oder gleich $\min(R, S)$.

Beweis: Für $|x| < \min(R, S)$ konvergieren beide Reihen absolut. Wir können also beide nach dem Cauchyprodukt miteinander multiplizieren. Rechts steht nichts anderes als das Cauchyprodukt siehe Theorem 2.19.

Die wichtigsten Potenzreihen

1. **Geometrische Reihe:** $\sum_{n=0}^\infty x^n$.

Der Konvergenzradius ist $R = 1$; für $|x| < 1$ kann man die Potenzreihe berechnen: $\sum_{n=0}^\infty x^n = \frac{1}{1-x}$

2. **Exponentialreihe:** $\sum_{n=0}^\infty x^n/n! =: \exp(x)$.

Nach dem Quotientenkriterium konvergiert diese Reihe für alle $x \in \mathbb{R}$, also ist der Konvergenzradius $R = \infty$. Aus der Cauchy-Produktformel Theorem 2.19 erhält man die fundamentale Formel

$$\exp(x + y) = \exp(x) \exp(y)$$

Also insbesondere

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

3. **Sinus-Reihe:** $\sin(x) := \sum_{k=0}^\infty (-1)^k \frac{x^{2k+1}}{(2k+1)!}$

Der Konvergenzradius ist $R = \infty$. Denn setzen Sie

$$c_k = \left\{ \begin{array}{ll} 0 & k \text{ gerade} \\ (-1)^{k-1/2} \cdot \frac{1}{k!} & k \text{ ungerade} \end{array} \right\}.$$

so ist $\sin(x) = \sum_{n=0}^\infty c_n x^n$. Weil die Exponentialreihe den Konvergenzradius $R = \infty$ hat, gilt $\lim \sqrt[n]{1/n!} = 0 = \lim \sqrt[n]{|c_n|}$, also ist auch der Konvergenzradius der Sinus-Reihe ∞ . Was diese Reihe mit dem Schulsinus zu tun hat, wird später behandelt.

4. **Cosinus-Reihe** $\cos(x) = \sum_{k=0}^\infty (-1)^k \frac{x^{2k}}{(2k)!}$.

Auch hier ist der Konvergenzradius $R = \infty$.

Weitere ebenso wichtige Potenzreihen lernen wir in den späteren Kapiteln kennen.

2.5 Komplexe Zahlenfolgen und Reihen

Die komplexen Zahlen haben sich nicht nur deshalb als nützlich erwiesen, weil man $x^2 + 1 = 0$ lösen kann, sondern sie offenbaren tiefe Zusammenhänge zwischen reellen Reihen und Funktionen, die sonst nicht verständlich werden.

Wir hatten in Kapitel 1 auf S. 18 den Abstand zweier komplexer Zahlen z_1 und z_2 erklärt als

$$|z_1 - z_2| = \sqrt{(\Re(z_1 - z_2))^2 + (\Im(z_1 - z_2))^2}.$$

Eine **komplexe Zahlenfolge** $(a_n)_{n \geq k}$ ist nun natürlich nichts anderes als eine Abbildung $a : \{n : n \geq k\} \rightarrow \mathbb{C}$, die jedem n das Element $a_n \in \mathbb{C}$ zuordnet. Eine **komplexe Reihe** $\sum_{i=k}^{\infty} a_k$ ist nichts anderes als die Folge $(s_n)_{n \geq k}$ mit $s_n = \sum_{i=k}^n a_i$.

Definition 2.22 a) Eine komplexe Zahlenfolge $(a_n)_{n \geq k}$ **konvergiert gegen die komplexe Zahl c** , wenn es zu jedem $\varepsilon > 0$ ein $n(\varepsilon)$ gibt mit $|a_n - c| < \varepsilon$ für alle $n \geq n(\varepsilon)$. Wir schreiben dann $c = \lim_{n \rightarrow \infty} a_n$.

b) Die unendliche Reihe $\sum_{i=k}^{\infty} a_k$ konvergiert gegen c , in Zeichen: $c = \sum_{i=k}^{\infty} a_k$, wenn die Folge $(s_n)_{n \geq k}$ ihrer Teilsummen gegen c konvergiert.

Anleitung zum Applet "komplexe Folgen" Wir zeigen in Bildern, wie die Konvergenz in \mathbb{C} aussieht. Da wir die waagrechte Achse für den Realteil, die senkrechte für den Imaginärteil brauchen, können wir den Index n nicht mehr abtragen. Benutzen Sie deshalb die Animation.

1. $a_n = (1 + 1/n) + i \cdot (1 - 1/n)$.
2. $a_n = (\cos(\pi/4) + i \cdot \sin(\pi/4))^n \cdot 2^{-n}$.
3. $a_n = \cos(1/n) + i \cdot \sin(\pi/2 - 1/n)$.
4. $s_n = \sum_{k=0}^n \left(\frac{\cos(\pi/4) + i \sin(\pi/4)}{2} \right)^k$.

Schauen Sie sich Realteil und Imaginärteil der Folgen an. Haben Sie auch den folgenden Satz entdeckt?

Satz 2.23 Sei $(a_n)_{n \geq k}$ eine komplexe Zahlenfolge. Die folgenden Aussagen sind äquivalent:

- a) $(a_n)_{n \geq k}$ konvergiert.
- b) Es gibt eine Zahl c , so daß die reelle Zahlenfolge $(|c - a_n|)_{n \geq k}$ eine Nullfolge ist.
- c) Die beiden reellen Zahlenfolgen $(\Re(a_n))_{n \geq k}$ und $(\Im(a_n))_{n \geq k}$ sind konvergent.
- d) Es gilt das Cauchy Kriterium: Zu jedem $\varepsilon > 0$ gibt es ein $n(\varepsilon)$ so daß für alle $m, n \geq n(\varepsilon)$ stets $|a_m - a_n| < \varepsilon$ gilt.

Ist eine (und damit alle) dieser äquivalenten Bedingungen erfüllt, so gilt

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \Re(a_n) + i \cdot \lim_{n \rightarrow \infty} \Im(a_n).$$

Beweis: a) \Rightarrow b) Nach Definition der Konvergenz gibt es ein $c \in \mathbb{C}$, so daß zu jedem $\varepsilon > 0$ ein $n(\varepsilon) \in \mathbb{N}$ existiert mit $b_n := |c - a_n| < \varepsilon$ für alle $n \geq n(\varepsilon)$. Das ist aber genau das Kriterium dafür, daß (b_n) eine Nullfolge in \mathbb{R} ist.

b) \Rightarrow c) Sei $c \in \mathbb{C}$ so gewählt, daß $(|c - a_n|)_{n \geq k}$ eine Nullfolge in \mathbb{R} ist. Für eine beliebige komplexe Zahl $z = \Re(z) + i \cdot \Im(z)$ gilt nun

$$|z| = \sqrt{(\Re(z))^2 + (\Im(z))^2} \geq \sqrt{(\Re(z))^2} = |\Re(z)|$$

und ebenso $|z| \geq |\Im(z)|$. Also ist $|\Re(c - a_n)| \leq |c - a_n|$ und $|\Im(c - a_n)| \leq |c - a_n|$. Damit sind $(|\Re(c - a_n)|)_{n \geq k}$ und $(|\Im(c - a_n)|)_{n \geq k}$ Nullfolgen.

Wegen $\Re(c - a_n) = \Re(c) - \Re(a_n)$ und $\Im(c - a_n) = \Im(c) - \Im(a_n)$ folgt nach Lemma 2.8 $\Re(c) = \lim \Re(a_n)$ und $\Im(c) = \lim \Im(a_n)$, also c .

$c) \Rightarrow d)$ Es ist zunächst für jede komplexe Zahl z stets $|z| \leq |\Re(z)| + |\Im(z)|$, wie man sofort durch Quadrieren der Ungleichung und Satz 1.20 erhält. Nach $c)$ sind die reellen Folgen $(\Re(a_n))_{n \geq k}$ und $(\Im(a_n))_{n \geq k}$ konvergent, erfüllen also nach Lemma 2.8 das Cauchy Kriterium.

Sei $\varepsilon > 0$ beliebig. Zu $\eta = \varepsilon/2$ gibt es also ein $n_1(\eta)$ mit $|\Re(a_m) - \Re(a_n)| < \eta$ für alle $m, n \geq n_1(\eta)$. Ferner gibt es ein $n_2(\eta)$ mit $|\Im(a_m) - \Im(a_n)| < \eta$ für alle $m, n \geq n_2(\eta)$. Sei $n(\varepsilon) = \max(n_1(\eta), n_2(\eta))$. Dann ist für alle $m, n \geq n(\varepsilon)$

$$\begin{aligned} |a_m - a_n| &\leq |\Re(a_m) - \Re(a_n)| + |\Im(a_m) - \Im(a_n)| \\ &< \eta + \eta = \varepsilon. \end{aligned}$$

$d) \Rightarrow a)$ Wegen $|\Re(z)| \leq |z|$ und $|\Im(z)| \leq |z|$ erfüllen $(\Re(a_n))_{n \geq k}$ und $(\Im(a_n))_{n \geq k}$ als reelle Folgen das Cauchy Kriterium, sind also nach Theorem 2.10 konvergent gegen u bzw. $v \in \mathbb{R}$.

Sei $c = u + iv$. Dann ist

$$|c - a_n| \leq |u - \Re(a_n)| + |v - \Im(a_n)|,$$

woraus leicht $c = \lim(a_n)$ folgt.

Der Rest ist klar.

Mit diesem Satz können wir sofort die Rechenregeln für die Grenzwerte konvergenter komplexer Folgen genau so beweisen, wie die reeller Zahlenfolgen. Es gilt also (siehe Satz 2.9):

$$\begin{aligned} |\lim_n a_n| &= \lim |a_n| \\ \lim_n a_n + \lim_n b_n &= \lim_n (a_n + b_n) \\ \lim_n a_n \cdot \lim_n b_n &= \lim_n (a_n \cdot b_n) \\ \frac{\lim_n a_n}{\lim_n b_n} &= \lim \frac{a_n}{b_n}, \text{ falls } b_n \neq 0 \forall n \text{ und } \lim b_n \neq 0. \end{aligned}$$

Zusätzlich gilt

$$\lim \overline{a_n} = \overline{\lim a_n},$$

wie man dem Satz sofort entnimmt. Vor allem gilt das Theorem 2.20 über die Konvergenz von Potenzreihen auch für komplexe Potenzreihen. So können wir die Exponentialreihe $\exp(z)$ auch für komplexe Zahlen z erklären und erhalten dieselbe Formel

$$\exp(z_1 + z_2) = \exp(z_1) \exp(z_2).$$

Aus ihr erhält man die besonders wichtige Formel:

$$\exp(z) = \exp(\Re(z)) \cdot \exp(i\Im(z))$$

Außerdem ergibt sich die Eulersche Formel

$$\begin{aligned} \exp(it) &= \cos(t) + i \sin(t) \quad \text{und damit} \\ |\exp(it)| &= 1 \quad \forall t \in \mathbb{R}, \text{ also} \\ |\exp(z)| &= \exp(\Re(z)). \end{aligned}$$

Beweis: Zum besseren Verständnis wiederholen Sie bitte die Reihen für Sinus und Kosinus siehe S. 34.
Es ist

$$\exp(it) = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} = \lim_{n \rightarrow \infty} \sum_{k=0}^{2n} \frac{(it)^k}{k!}.$$

Wegen $i^2 = -1, i^4 = 1$ ist

$$\sum_{k=0}^{2n} \frac{(it)^k}{k!} = \sum_{k=0}^n (-1)^k \frac{t^{2k}}{(2k)!} + i \sum_{k=1}^n (-1)^k \frac{t^{2k-1}}{(2k-1)!}.$$

Nach den Rechenregeln für Grenzwerte von Folgen ergibt sich die erste Behauptung. Die zweite folgt so:
Es ist

$$\exp(-it) = \sum_{k=0}^{\infty} (-i)^k \cdot \frac{t^k}{k!} = \sum_{k=0}^{\infty} \frac{\overline{(it)^k}}{k!} = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} = \overline{\exp(it)},$$

also

$$|\exp(it)|^2 = |\overline{\exp(it)} \cdot \exp(it)| = |\exp(-it) \exp(it)| = |1| = 1.$$

Der Rest folgt aus der Multiplikationsformel für die Exponentialreihe.

Kapitel 3

Einführung in die Theorie reeller Funktionen einer Veränderlichen

3.1 Grundbegriffe

3.1.1 Definitionen und Vereinbarungen

Reelle Funktionen einer Veränderlichen bringen die Abhängigkeit einer Größe von *einer einzigen* anderen zum Ausdruck. In der Realität hängt eine Größe (das ist im Moment etwas, das man experimentell messen kann) von *mehreren* anderen Größen ab. Aber in der Praxis nimmt man oft nur eine von ihnen als wesentlich an, die andere nimmt man als “konstant” an.

Beispiele bilden die folgenden Funktionen

- Blutdruck als Funktion der Tageszeit
- Zahl der Zellen als Funktion der Zeit (in Zellkulturen)
- Fallweg als Funktion der Zeit (beim Herunterfallen von Gegenständen)
- Luftdruck als Funktion der Höhe des Ortes über dem Meeresboden
- Auslenkung eines Pendels als Funktion der Zeit.

Finden Sie selbst unbedingt drei Beispiele für solche Abhängigkeiten.

Mathematisch sind reelle Funktionen einer Veränderlichen ein Spezialfall von Abbildungen (siehe 1.3)

Definition 3.1 *Eine reelle Funktion f einer Veränderlichen ist eine Abbildung einer Teilmenge D von \mathbb{R} nach \mathbb{R} .*

Hiernach sind auch reelle Folgen reelle Funktionen. Man nimmt $D = \mathbb{N}$.

Wir betrachten im folgenden im wesentlichen nur die folgenden 5 Typen des Definitionsbereichs D

1. $D = [a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ *abgeschlossenes Intervall*

2. $D =]a, b] = \{x \in \mathbb{R} : a < x \leq b\}$ *links offenes, rechts abgeschlossenes Intervall*
 3. $D = [a, b[= \{x \in \mathbb{R} : a \leq x < b\}$ *links abgeschlossenes, rechts offenes Intervall*
 4. $D =]a, b[= \{x \in \mathbb{R} : a < x < b\}$ *offenes Intervall*
 5. Sei J ein Intervall und $\{x_1, \dots, x_n\} \in J$. Dann ist $D = J \setminus \{x_1, \dots, x_n\}$

In 2. kann das linke Intervallende a auch $-\infty$ sein: $] - \infty, b] = \{x \in \mathbb{R} : x \leq b\}$, in 3. kann das rechte Intervallende b auch ∞ sein: $[a, \infty[= \{x \in \mathbb{R} : a \leq x\}$.

In 4. kann jedes der beiden Intervallenden unendlich sein, also

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

Legen wir uns nicht fest, ob die Intervallenden dazugehören sollen oder nicht, schreiben wir $\langle a, b \rangle$.

Anleitung zum FuncPlot-Applet Schauen Sie sich einige Funktionen an:

1. $f(x) = x^2$, $D = [-1, 1]$.
2. $f(x) = \sin(x)$, $D = [-\pi, \pi]$.
3. $f(x) = \exp(x)$, $D = [-2, 2]$.

3.1.2 Erzeugung von Funktionen

Wenn nichts anderes gesagt wird, nehmen wir $D = \mathbb{R}$ an. Die Sätze gelten jedoch für beliebiges D , es sei denn, es wird ausdrücklich etwas anderes gesagt.

Wir beginnen mit **konstanten Funktionen**:

Sei $c \in \mathbb{R}$. Dann ist $f(x) = c$ für alle $x \in D$ die Funktion, die konstant den Wert c annimmt.

Als zweites betrachten wir die **identische Abbildung** $f(x) = x$.

Ist $f : D \rightarrow \mathbb{R}$ eine Funktion und $c \in \mathbb{R}$ so ist das **skalare Vielfache** cf von f die Funktion

$$x \mapsto c \cdot f(x).$$

Zum Beispiel erhält man für $f(x) = x$: $2f : x \mapsto 2x$.

Sind $f, g : D \rightarrow \mathbb{R}$ Funktionen, so ist ihre **Summe** durch

$$f + g : x \mapsto f(x) + g(x)$$

gegeben. Ebenso ist ihr **Produkt** durch

$$fg : x \mapsto f(x)g(x)$$

erklärt.

Mit diesen Vereinbarungen und durch Induktion können wir nun allein aus den beiden Funktionen $1 : x \mapsto 1$ und $x \mapsto x$ alle **Polynome** konstruieren.

Ein Polynom ist also eine Abbildung der Form $P(x) = \sum_{k=0}^n a_k x^k$. Ist $a_n \neq 0$, so heißt n der **Grad** des Polynoms.

Summe und Produkt von Polynomen sind wieder Polynome

Quotient zweier Funktionen:

Seien $f, g : D \rightarrow \mathbb{R}$ zwei reelle Funktionen und $D_Q = \{x \in D : g(x) \neq 0\}$. Dann ist der Quotient $\frac{f}{g}$ auf D_Q erklärt durch $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$

Eine **rationale Funktion** ist in diesem Sinn der Quotient zweier Polynome. **Beispiele:**

1. $f(x) = x^2 + 2x - 1$.
2. $f(x) = x^3$.
3. $f(x) = 1 + x^2$.
4. $f(x) = \frac{1}{1+x^2}$
5. $f(x) = \frac{1}{1-x}$. Wo ist f **nicht** erklärt?
6. $f(x) = \frac{x^2-1}{x+1}$. Wo ist f **nicht** erklärt ?

Anleitung zum FuncPlot-Applet Schauen Sie sich die Funktionen an. Wie verhalten sie sich, wenn Sie den Ausschnitt auf der x -Achse vergrößern?

Hintereinanderausführung zweier Funktionen.

Seien $f : D_1 \rightarrow \mathbb{R}$ und $g : D_2 \rightarrow \mathbb{R}$ zwei reelle Funktionen. Das Bild $f(D_1)$ liege ganz in D_2 . Dann erhält man durch

$$g \circ f : D_1 \ni x \mapsto g(f(x))$$

wieder eine reelle Funktion.

Beispiele:

1. Seien f und g Polynome, also $D_1 = D_2 = \mathbb{R}$. Dann ist $g \circ f$ wieder ein Polynom. Prüfen Sie das nach!
2. $f(x) = x^2$, $g(x) = \sqrt{|x|}$, $g \circ f = |x|$.

Durch Potenzreihen erklärte Funktionen

Sei $(a_n)_{n \geq 0}$ eine reelle Zahlenfolge, so daß die Folge $(\sqrt[n]{|a_n|})_{n \geq 0}$ beschränkt ist. Dann ist der Konvergenzradius R (siehe Theorem 2.20) der Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ echt größer als 0. Auf $D =]-R, R[$ ist durch $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Funktion gegeben. Die Menge der so erklärten Funktionen (Polynome gehören dazu, warum?) ist extrem wichtig, wie wir anhand der folgenden Beispiele ahnen:

Beispiele:

1. $f(x) = \exp(x) = \sum_{k=0}^{\infty} x^k / k!$
2. $f(x) = \sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$
3. $f(x) = \cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$
4. $f(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k} = \frac{1}{1+x^2}$, $|x| < 1$

Die *Theorie der Taylorreihen* beschäftigt sich damit, alle Funktionen zu charakterisieren, die durch Potenzreihen gegeben sind (siehe der Abschnitt “Satz von Taylor”).

Weitere Methoden zur Konstruktion von Funktionen sind die *Umkehrung bijektiver Funktionen*, also zum Beispiel die Funktion $f(x) = \sqrt{x}$ auf $\mathbb{R}_+ = [0, \infty[$, die Konstruktion als *Lösung von Differentialgleichungen* und die Konstruktion als *Grenzwert einer Folge bekannter Funktionen*. Diese Methoden werden wir später kennen lernen.

Treppenfunktionen

Sei $A \subset \mathbb{R}$ eine Menge. Dann heißt die Funktion $1_A : x \mapsto \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$

Indikatorfunktion (“Anzeigefunktion”) der Menge A .

Man kann solche Funktionen auch “Einschalt-Funktionen” nennen. Versuchen Sie dies zu begründen.

$1_\emptyset = 0$, $1_{\mathbb{R}} = 1$ (konstante Funktionen).

Eine Treppenfunktion ist eine Funktion der Form $t(x) = \sum_{k=1}^n c_k \cdot 1_{A_k}$. Sie hat in der Regel Sprünge. Viele Funktionen kann man durch solche Funktionen annähern, ein wichtiger Aspekt in der Integrationstheorie und der graphischen Datenverarbeitung.

Anleitung zum Applet ”partiell definierte Funktionen” Treppenfunktionen kennen Sie als Histogramme. Hier ein paar Beispiele:

Beispiele: Wählen Sie den Ausschnitt der x -Achse so groß, daß Sie alle Sprünge sehen!

1. $f = 1_{[0,1]}$
2. $f = 2 \cdot 1_{[0,1]} + 3 \cdot 1_{[2,4]}$
3. $f = 2 \cdot 1_{[0,1]} - \frac{3}{2} \cdot 1_{[0.5,2]}$
4. $f = \sum_{k=1}^4 \left(\frac{k}{5}\right)^2 1_{[\frac{k}{5}, \frac{k+1}{5}[}$. Das ist eine erste Annäherung an $f(x) = x^2$.

Damit haben Sie die wichtigsten elementaren Regeln zur Bildung komplizierter Funktionen aus einfacheren kennengelernt. Ein **Parser** ist ein Programm, das bei einer gegebenen Funktion herauszufinden versucht, wie sie aus einfachen Funktionen und den angegebenen Regeln aufgebaut wurde, um sie berechenbar zu machen.

3.2 Grenzwert von Funktionswerten

Wenn die Funktion $f : \mathbb{R}_+ = [0, \infty[\rightarrow \mathbb{R}$ ein naturwissenschaftliches Phänomen beschreibt, möchte man wissen, wie verhält sich f für große Argumente, “für große Zeiten”. Strebt es da einer festen Zahl zu, die man dann als Prognose werten kann? Ganz ähnlich will man untersuchen was passiert, wenn man sich mit den Argumenten einem endlichen Wert nähert. Das kennen Sie aus der Schule: Sei $g :]a, b[\rightarrow \mathbb{R}$ und $a < c < b$. Um zu untersuchen, ob g eine “Tangente in c besitzt”, untersucht man die Funktion $S(x) = \frac{f(x)-f(c)}{x-c}$ auf $]a, b[\setminus\{c\}$. $S(x)$ ist die Steigung der Sekante an die Kurve $\{(x, g(x)) : x \in]a, b[\}$. Man will wissen, ob sich $S(x)$ einem festen Wert nähert, wenn x gegen c geht.

Wir präzisieren nun, was wir oben ausgeführt haben. Wir kennen schon den Begriff des Grenzwertes einer Folge. Also führen wir das Neue auf das Bekannte zurück. Zunächst brauchen wir den Begriff des Adhärenzpunktes einer Menge D .

Definition 3.2 Ein Punkt c (nicht notwendig in D) heißt **Adhärenzpunkt** der Menge $D \subset \mathbb{R}$, wenn es mindestens eine Folge $(a_n)_{n \geq 1}$ aus D gibt, die gegen c konvergiert.

c heißt **Häufungspunkt** der Menge D , wenn es mindestens eine Folge $(a_n)_{n \geq 1}$ aus D gibt, die gegen c konvergiert und außerdem noch $a_n \neq c$ für alle n erfüllt.

Jeder Häufungspunkt ist also Adhärenzpunkt, aber nicht umgekehrt.

Wir bezeichnen die Menge der Adhärenzpunkte einer Menge D mit \overline{D} und die Menge ihrer Häufungspunkte mit $\mathcal{H}(D)$

Beispiele:

1. Jeder Punkt des Intervalls $D =]a, b[$ ist Häufungspunkt von D , egal ob die Endpunkte dazugehören oder nicht.
2. Seien $a, b \in \mathbb{R}$, $a < b$. Dann sind beide Endpunkte a und b des Intervalls $D =]a, b[$ Häufungspunkte von D , egal ob die Endpunkte dazugehören oder nicht. Es gilt also $[a, b] \subseteq \mathcal{H}(]a, b[)$.
3. Sei $D \neq \emptyset$ beliebig. Dann ist jeder Punkt aus D ein Adhärenzpunkt von D . Denn sei $c \in D$. Dann konvergiert die konstante Folge (c, c, c, \dots) , also $(a_n)_{n \geq 1}$ mit $a_n = c$ für alle n gegen c und ist aus D . Als Formel: $D \subset \overline{D}$.
4. Sei $D = \mathbb{N}$. Dann ist jeder Punkt $n \in \mathbb{N}$ Adhärenzpunkt von \mathbb{N} , aber \mathbb{N} besitzt keinen einzigen Häufungspunkt.

Trainieren Sie die Begriffe in den folgenden Aufgaben. **Aufgaben:**

1. Sei $\emptyset \neq D \subset \mathbb{R}$. Zeigen Sie bitte: c ist genau dann Häufungspunkt von D , wenn c Adhärenzpunkt von $D_1 = D \setminus \{c\}$ ist.
2. Jede reelle Zahl x ist Häufungspunkt von \mathbb{Q} . *Tip:* Benutzen Sie Satz 1.28 c).
3. Zeigen Sie bitte: Sei $(a_n)_{n \geq 1}$ eine reelle Folge mit $a_m \neq a_n$ für $m \neq n$. c ist genau dann ein Häufungspunkt der Menge $\{a_n : n \in \mathbb{N}\} = D$, wenn c ein Häufungspunkt der Folge (im Sinne der Definition 2.2) ist.

Adhärenzpunkte und Häufungspunkte kann man auch so charakterisieren:

Satz 3.3 Sei $\emptyset \neq D \subset \mathbb{R}$.

- a) $c \in \mathbb{R}$ ist genau dann ein Adhärenzpunkt von D , wenn für jedes $\varepsilon > 0$ die Menge $]c - \varepsilon, c + \varepsilon[\cap D \neq \emptyset$.
- b) $c \in \mathbb{R}$ ist genau dann ein Häufungspunkt von D wenn für jedes $\varepsilon > 0$ die Menge $]c - \varepsilon, c + \varepsilon[\cap D$ unendlich viele Punkte enthält.

Beweis:

a) (I) Sei c ein Adhärenzpunkt von D . Dann gibt es eine Folge $(a_n)_{n \geq 1}$ aus D mit $c = \lim_n a_n$. Zu $\varepsilon > 0$ gibt es also ein $n(\varepsilon)$ mit $|c - a_n| < \varepsilon$ für alle $n \geq n(\varepsilon)$. Insbesondere ist $a_{n(\varepsilon)} \in]c - \varepsilon, c + \varepsilon[\cap D$, diese Menge also nicht leer.

(II) Für $\varepsilon = 1/n$ gibt es einen Punkt $a_n \in]c - 1/n, c + 1/n[\cap D$. Offensichtlich ist $c = \lim_n a_n$.

b) (I) Sei c ein Häufungspunkt von D . Dann gibt es eine Folge $(a_n)_{n \geq 1}$ aus D mit $c = \lim_n a_n$ und $a_n \neq c$

für alle n . Wir konstruieren induktiv:

Zu $\varepsilon = 1$ gibt es $n(1)$ mit $|c - a_{n(1)}| < 1$. Sei $a_{n(1)}, \dots, a_{n(k)}$ schon konstruiert mit $a_{n(j)} \neq a_{n(l)}$ für $j \neq l$ und $|c - a_{n(l)}| < 1/l$.

Sei $d = \min\{|c - a_{n(l)}| : 1 \leq l \leq k\}$. Es ist $d \neq 0$ nach Voraussetzung. Zu $\varepsilon = \min(\frac{1}{k+1}, \frac{d}{2})$ gibt es ein $n(\varepsilon)$ mit $|c - a_{n(\varepsilon)}| < \varepsilon$. Wir setzen $a_{n(k+1)} = a_{n(\varepsilon)}$. Es ist $a_{n(k+1)} \neq a_{n(l)}$ für $l \leq k$, weil sonst

$$\frac{d}{2} > |c - a_{n(k+1)}| = |c - a_{n(l)}| \geq d$$

für ein $l \leq k$ gelten würde. Außerdem gilt $|c - a_{n(k+1)}| < \frac{1}{k+1}$.

Sei $\varepsilon > 0$ beliebig. Dann gibt es ein k mit $1/k < \varepsilon$. Dann gilt aber

$a_{n(j)} \in]c - \varepsilon, c + \varepsilon[\cap D$ für alle $j \geq k$, also liegen in $]c - \varepsilon, c + \varepsilon[\cap D$ unendlich viele Elemente.

(II) Seien in jedem $]c - \varepsilon, c + \varepsilon[\cap D$ unendlich viele Elemente. Wähle $a_1 \in]c - 1, c + 1[\cap D$, $a_1 \neq c$ (das ist möglich, es gibt unendlich viele zur Auswahl).

Haben wir $a_1, \dots, a_n \in D$ schon konstruiert mit $c \neq a_k \in]c - 1/k, c + 1/k[\cap D$ für $1 \leq k \leq n$, so gibt es wegen der unendlich vielen Elemente in $]c - \frac{1}{k+1}, c + \frac{1}{k+1}[\cap D$ ein a_{n+1} in dieser Menge mit $a_{n+1} \neq c$. Dann gilt $|c - a_n| < 1/n$, also $\lim a_n = c$, c ist also Häufungspunkt von D .

Wir können nun den Grenzwert von Funktionswerten erklären:

Definition 3.4 Sei c ein Adhärenzpunkt der Menge $D \subset \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine reelle Funktion. Wir sagen, für x gegen c geht $f(x)$ gegen die Zahl $d \in \mathbb{R}$, oder $\lim_{x \rightarrow c} f(x) = d$ (lies Limes x gegen c von $f(x)$ gleich d), wenn für jeden gegen c konvergente Folge $(a_n)_{n \geq k}$ aus D stets die Bildfolge $(f(a_n))_{n \geq k}$ gegen d konvergiert. d heißt **Grenzwert von $f(x)$ für $x \rightarrow c$** .

Am Computer können Sie sich anschauen, was dieser Begriff bedeutet.

Anleitung zum SeqPlot-Applet Wählen Sie einen Adhärenzpunkt c und für diesen dann verschiedene gegen c konvergente Folgen $(a_n)_{n \geq 1}$ aus D (durch Formeln). Sie können dann am Bild sehen, ob die Folge $(f(a_n))_{n \geq 1}$ jeweils gegen ein und denselben Wert konvergiert.

Beispiele:

1. $f(x) = x^3$, $D = [0, 1]$, $c = 1/2$, zum Beispiel $a_n = 1/2 + (-1)^n/2^n$.
2. $f(x) = \frac{x^2-1}{x-1}$, $D = [0, 2] \setminus \{1\}$, $c = 1$, $a_n = (-1)^n/n + 1$.
3. $f(x) = \frac{x^2-1}{(x-1)^2}$, $D = [0, 2] \setminus \{1\}$, $c = 1$, $a_n = (-1)^n/n + 1$.
4. $f(x) = \sin(1/x)$, $D =]0, 1]$, $c = 0$, $a_n = 1/n$ oder $a_n = \frac{1}{n\pi}$ oder $a_n = \frac{1}{2n}$

Notieren Sie genau, wie die Bildfolge $(f(a_n))_{n \geq 1}$ sich jeweils verhält.

Gibt es einen Grenzwert d für $x \rightarrow 0$?

Aus den Rechenregeln für Grenzwerte konvergenter Folgen (siehe Satz 2.9 erhalten Sie nun sofort die Rechenregeln für Grenzwerte von Funktionswerten.

Satz 3.5 Sei c ein Adhärenzpunkt der Menge $D \subset \mathbb{R}$, alle auftretenden Funktionen seien auf D definiert und die Grenzwerte auf der linken Seite der Gleichungen mögen jeweils existieren. Dann existieren auch die Grenzwerte auf der rechten Seite der Gleichungen und es gilt jeweils Gleichheit:

$$a) \lim_{x \rightarrow c} f(x) \pm \lim_{x \rightarrow c} g(x) = \lim_{x \rightarrow c} (f(x) \pm g(x)).$$

b) $\lim_{x \rightarrow c} f(x) \cdot \lim_{x \rightarrow c} g(x) = \lim_{x \rightarrow c} (f(x)g(x)).$

c) Ist $g(x) \neq 0$ für alle x und $\lim_{x \rightarrow c} g(x) \neq 0$, so gilt

$$\frac{\lim_{x \rightarrow c} f(x)}{\lim_{x \rightarrow c} g(x)} = \lim_{x \rightarrow c} \frac{f(x)}{g(x)}.$$

d) $|\lim_{x \rightarrow c} f(x)| = \lim_{x \rightarrow c} |f(x)|.$

e) $\max(\lim_{x \rightarrow c} f(x), \lim_{x \rightarrow c} g(x)) = \lim_{x \rightarrow c} (\max(f(x), g(x))).$

f) $\min(\lim_{x \rightarrow c} f(x), \lim_{x \rightarrow c} g(x)) = \lim_{x \rightarrow c} (\min(f(x), g(x))).$

Beweis: Wir führen nur den Beweis für a), e) und f). Der Rest geht völlig analog.

a) Sei $(a_n)_{n \geq k}$ eine beliebige, gegen c konvergente Folge aus D . Nach Voraussetzung konvergieren dann die Bildfolgen $(f(a_n))_{n \geq k}$ gegen $\lim_{x \rightarrow c} f(x)$ und $(g(a_n))_{n \geq k}$ gegen $\lim_{x \rightarrow c} g(x)$. Nach Satz 2.9 a) konvergiert dann die Folge

$$((f + g)(a_n))_{n \geq k} = (f(a_n) + g(a_n))_{n \geq k}$$

gegen $d := \lim_{x \rightarrow c} f(x) + \lim_{x \rightarrow c} g(x)$. Da $(a_n)_{n \geq k}$ beliebig war, folgt die Behauptung. (Zum besseren Verständnis können Sie die Funktion $f + g$ einfach h nennen.)

e) Es ist $\max(u, v) = \frac{1}{2}(|u - v| + (u + v))$ (prüfen Sie das nach!). Damit folgt e) aus a), b) und d).

f) Es ist $\min(u, v) = -\max(-u, -v)$. Damit folgt f) aus b) und e).

Aufgaben: Berechnen Sie mit den Rechenregeln für Grenzwerte die folgenden Grenzwerte

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

2. $\lim_{x \rightarrow 0} \frac{x^3 + 3x - 1}{x^2 + 1}$

3. $\lim_{x \rightarrow 0} |\sin(x)|$. Dies ist etwas schwieriger. Es ist $\sin(x) = x - x^2 \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k-1}}{(2k+1)!}$. Zeigen Sie $|\sum_{k=1}^{\infty} (-1)^k \frac{x^{2k-1}}{(2k+1)!}| \leq \exp(x)$ und folgern Sie hieraus $\lim_{x \rightarrow 0} \sin(x) = 0$.

4. $\lim_{x \rightarrow 0} \frac{\sin(x)}{x}$. Vorsicht! Hier dürfen Sie die Rechenregeln nicht anwenden. Benutzen Sie die Formeln der vorigen Aufgabe.

Unsere Definition der Konvergenz von Funktionswerten benutzte diejenige für Konvergenz von Folgen. Um Konvergenz nachzuweisen, muß man theoretisch alle Folgen $(a_n)_{n \geq k}$ betrachten, die gegen die kritische Stelle c konvergieren und testen, ob die jeweilige Bildfolge $(f(a_n))_{n \geq k}$ gegen ein und denselben Grenzwert konvergiert. Das ist gerade bei Aufgabe 4 oben lästig, weil wir dort unsere Rechenregeln nicht zur Verfügung haben.

Anleitung zum FuncPlot-Applet Schauen wir uns $f(x) = \frac{\sin(x)}{x}$ für $x \neq 0$ an. f liegt offensichtlich dicht bei 1 für sehr kleine Werte von x . Das bedeutet genauer: legen wir waagrecht einen Streifen der Breite 2ε um die waagrechte Gerade $y = 1$, wo $\varepsilon > 0$ beliebig gewählt war, so finden wir ein kleines $\delta > 0$, so daß für $0 < |x| < \delta$ stets $f(x)$ im gewählten ε -Streifen liegt.

Lesen Sie aus der Zeichnung solch δ ab für $\varepsilon = 0.1$ und dann für $\varepsilon = 0.01$.

Behandeln Sie die anderen Beispiele aus der vorangegangenen Aufgabe entsprechend.

Es gilt tatsächlich das folgende $\varepsilon - \delta$ -Kriterium für Konvergenz.

Theorem 3.6 Sei $f : D \rightarrow \mathbb{R}$ eine reelle Funktion und c ein Adhärenzpunkt von D . Die beiden folgenden Aussagen sind äquivalent:

- a) $f(x)$ konvergiert gegen d für x gegen c , also $\lim_{x \rightarrow c} f(x) = d$.
- b) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle $x \in D$ mit $|x - c| < \delta$ stets $|f(x) - d| < \varepsilon$ gilt.
- c) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $f(]c - \delta, c + \delta[\cap D) \subset]d - \varepsilon, d + \varepsilon[$

Das Kriterium c) haben Sie sich oben gerade angeschaut.

Beweis: $c) \Leftrightarrow d)$ ist offensichtlich.

$a) \Rightarrow c)$ (indirekt): Gilt b) nicht, so gibt es ein $\varepsilon > 0$ und zu jedem δ , also insbesondere zu $\delta = 1/n$ ein Ausnahmeelement $a_n \in]c - 1/n, c + 1/n[\cap D$, das gerade *nicht* in $]d - \varepsilon, d + \varepsilon[$ abgebildet wird, das also $|f(a_n) - d| \geq \varepsilon$ erfüllt. Die Folge $(a_n)_{n \geq 1}$ konvergiert gegen c , die Folge $(f(a_n))_{n \geq 1}$ aber offensichtlich nicht gegen d , also gilt a) nicht.

$c) \Rightarrow a)$. Sei $(a_n)_{n \geq k}$ eine beliebige gegen c konvergente Folge aus D . Wir müssen zeigen, daß $(f(a_n))_{n \geq k}$ gegen d konvergiert. Sei dazu $\varepsilon > 0$ beliebig vorgegeben. Nach Voraussetzung existiert ein $\delta > 0$ mit $f(]c - \delta, c + \delta[\cap D) \subset]d - \varepsilon, d + \varepsilon[$. Zu diesem δ existiert ein $n(\delta)$ mit $|c - a_n| < \delta$ für alle $n \geq n(\delta)$, also $a_n \in]c - \delta, c + \delta[$ für all diese n . Dann gilt aber $|f(a_n) - d| < \varepsilon$ für all' diese $n \geq n(\delta)$. Wir können also $n(\delta) = n(\varepsilon)$ wählen. Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.

Neben der Konvergenz von $f(x)$ für $x \rightarrow c$ interessiert uns auch, ob $f(x)$ gegen ∞ oder gegen $-\infty$ geht für $x \mapsto c$. Zum Beispiel gilt für die Entwicklung der Weltbevölkerung $f(t) = \frac{1}{a-bt}$ für gewisse Konstanten $a, b > 0$ und $0 \leq t < \frac{a}{b}$ (der Nullpunkt $t = 0$ ist die Gegenwart). Was passiert für $t \rightarrow \frac{a}{b}$?

Definition 3.7 $f(x)$ **divergiert bestimmt gegen ∞** für $x \rightarrow c$, wenn zu jedem $L > 0$ ein $\delta > 0$ existiert, mit $f(x) > L$ für alle $x \in D$ mit $|x - c| < \delta$. Wir schreiben dann $\lim_{x \rightarrow c} f(x) = \infty$.

Entsprechend sagen wir $f(x)$ **divergiert bestimmt gegen $-\infty$** wenn zu jedem $L > 0$ ein $\delta > 0$ existiert mit $f(x) < -L$ für alle $x \in D$ mit $|x - c| < \delta$. Entsprechend schreiben wir dann $\lim_{x \rightarrow c} f(x) = -\infty$.

Aufgaben: Untersuchen Sie die folgenden Funktionen auf Konvergenz bzw. bestimmte Divergenz:

1. $f(t) = \frac{1}{a-bt}$, $a, b > 0$, $D = [0, \frac{a}{b}[$, $c = \frac{a}{b}$.
2. $f(x) = \frac{x^2-1}{x-1}$, $D = \mathbb{R} \setminus \{1\}$, $c = 1$.
3. $f(x) = \frac{x^2-1}{(x-1)^2}$, $D =]1, \infty[$, $c = 1$.
4. $f(x) = \frac{x^2-1}{(x-1)^2}$, $D =]-\infty, 1[$, $c = 1$.

Anleitung zum FuncPlot-Maplet mit Singularitäten Schauen Sie sich die Funktionen aus obiger Aufgabe an.

Genau so wichtig wie die Frage, was mit den Funktionswerten $f(x)$ passiert, wenn x sich einem Punkt c nähert, ist die Frage, wie verhält sich $f(x)$ für beliebig große Argumente x ? Wir präzisieren.

Definition 3.8 Sei $D =]b, \infty[$ ein rechts unbeschränktes Intervall und $f : D \rightarrow \mathbb{R}$ eine reelle Funktion. $f(x)$ konvergiert gegen d für x gegen ∞ , in Zeichen: $\lim_{x \rightarrow \infty} f(x) = d$, wenn zu jedem $\varepsilon > 0$ ein $L = L(\varepsilon) \in D$ existiert mit $|f(x) - d| < \varepsilon$ für alle $x \geq L$. Man sagt dann auch, der Grenzwert $\lim_{x \rightarrow \infty} f(x)$

existiert.

$f(x)$ divergiert bestimmt gegen ∞ (bzw. $-\infty$ für x gegen ∞ , wenn zu jedem $M > 0$ ein $L(M) \in D$ existiert mit $f(x) > M$ (bzw. $f(x) < -M$) für alle $x \geq L(M)$).

Wir schreiben dann $\lim_{x \rightarrow \infty} f(x) = \infty$ (bzw. $\lim_{x \rightarrow \infty} f(x) = -\infty$).

Anleitung zum FuncPlot-Applet Schauen Sie sich an, wie $\lim_{x \rightarrow \infty} f(x) = d$ aussieht, wenn $d \in \mathbb{R}$ ist. Der Intuition nach sollte f für große x fast konstant sein.

Beispiele:

1. $f(x) = \frac{1}{1+x}$, $D = [0, \infty[$.
2. $f(x) = \frac{\sin(x)}{1+x}$, $D = [0, \infty[$.
3. $f(x) = x \cdot \exp(-x)$, $D = \mathbb{R}$, $x \rightarrow \infty$.

Aufgaben:

1. Beweisen Sie die folgenden Formeln, vorausgesetzt, die Grenzwerte auf der linken Seite existieren:
 - a) $\lim_{x \rightarrow \infty} f(x) \pm \lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} (f(x) \pm g(x))$.
 - b) $\lim_{x \rightarrow \infty} f(x) \cdot \lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} f(x)g(x)$.
 - c) $|\lim_{x \rightarrow \infty} f(x)| = \lim_{x \rightarrow \infty} |f(x)|$.
 - d) Ist $f(x) \neq 0$ für alle $x \in D$ und $\lim_{x \rightarrow \infty} f(x) \neq 0$, so ist $\frac{1}{\lim_{x \rightarrow \infty} f(x)} = \lim_{x \rightarrow \infty} \frac{1}{f(x)}$.
2. $\lim_{x \rightarrow \infty} \frac{x^n}{\sum_{k=0}^{n+p} a_k x^k} = 0$ für $n \in \mathbb{N}$, $p \geq 1$, $a_k \geq 0$ und $a_{n+p} \neq 0$.
3. $\lim_{x \rightarrow \infty} x^n \cdot e^{-ax} = 0$ für $n \in \mathbb{N}$, $a > 0$.
Tip: $0 < e^{-ax} = \frac{1}{e^{ax}} \leq \frac{1}{\sum_{k=0}^{n+1} \frac{a^k x^k}{k!}}$ und Aufgabe
4. Formulieren Sie bitte selbst, was $\lim_{x \rightarrow -\infty} f(x) = d$ bzw. $\lim_{x \rightarrow -\infty} f(x) = \infty$ ($= -\infty$) heißen könnte und beweisen Sie die zu Aufgabe 1 analoge Aufgabe.

3.3 Stetigkeit

“natura non facit saltus” (die Natur macht keine Sprünge) ist ein altes Prinzip der Beschreibung von Naturgesetzen. Mit dem Hilfsmittel des Grenzwertes ist es einfach, dies Prinzip zu präzisieren.

Definition 3.9 (Stetigkeit)

Eine reelle Funktion $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ ist an der Stelle $x_0 \in D$ **stetig**, wenn $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ gilt. f ist (**überall**) **stetig**, wenn f in jedem Punkt $x_0 \in D$ stetig ist.

Aus den verschiedenen äquivalenten Bedingungen für die Existenz von Grenzwerten (Satz 3.6) erhalten wir die folgenden Bedingungen für die Stetigkeit. Der Beweis ist klar.

Theorem 3.10 Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $x_0 \in D$. Die folgenden Aussagen sind äquivalent.

- a) f ist in x_0 stetig.
- b) Für jede gegen x_0 konvergente Folge $(a_n)_{n \geq k}$ aus D konvergiert $(f(a_n))_{n \geq k}$ gegen $f(x_0)$.
- c) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ so daß $|f(x) - f(x_0)| < \varepsilon$ ist für alle $x \in D$ mit $|x - x_0| < \delta$.
- d) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ mit

$$f([x_0 - \delta, x_0 + \delta] \cap D) \subset]f(x_0) - \varepsilon, f(x_0) + \varepsilon[.$$

Anleitung zum FuncPlot-Applet Schauen Sie sich insbesondere das $\varepsilon - \delta$ -Kriterium (Bedingung c) bzw. d) im vorangegangenen Theorem) an den folgenden Beispielen an. Bestimmen Sie aus dem Bild ein geeignetes δ .

Beispiele:

1. $f(x) = x^2$, $D = [0, 2]$, $x_0 = 1$; $\varepsilon = 0.1$.
2. $f(x) = 2x^3 - x$, $D = [-1, 1]$, $x_0 = 0$; $\varepsilon = 0.1$.
3. $f(x) = \exp(x)$, $D = [-1, 1]$, $x_0 = 0$; $\varepsilon = 0.1$.
4. $f(x) = \sqrt{|x|}$, $D = [-1, 1]$, $x_0 = 0$; $\varepsilon = 0.1$.
5. $f(x) = 1_{[0,1]}(x) - 1_{]2,3]}(x)$, $x_0 = 1$; $\varepsilon = 0.1$.

Natürlich sind alle die genannten Stetigkeitskriterien unhandlich. Daher weisen wir die Stetigkeit schneller nach, wenn wir die Funktion daraufhin untersuchen, wie sie aus einfacheren aufgebaut ist, und dabei die folgenden Regeln benutzen, die sofort aus dem entsprechenden Satz 3.5 für Grenzwerte folgen.

Satz 3.11 (Rechenregeln für Stetigkeit)

Seien $f, g : D \rightarrow \mathbb{R}$ in $x_0 \in D$ stetig. Dann sind $f + g$, $f - g$, fg , $|f|$, $\max(f, g)$, $\min(f, g)$ in x_0 stetig. Ist $g(x) \neq 0$ für alle x , so ist auch $\frac{f}{g}$ in x_0 stetig.

Aufgaben:

1. Zeigen Sie: $f(x) = \frac{\exp(x)}{1+x^2}$ ist auf \mathbb{R} stetig.
Tip: Zeigen Sie zunächst, daß $\exp(x)$ stetig ist.
 Benutzen Sie dazu $|\exp(x) - \exp(x_0)| = \exp(x_0)|\exp(x - x_0) - 1|$ (siehe die Formeln für die Exponentialreihe 34) und zeigen Sie nun $|\exp(y) - 1| \leq |y| \exp(y) < |y| \cdot \exp(1)$ für $|y| < 1$ durch Heranziehen der Reihenformel für $\exp(y)$. (s. 34).
2. Zeigen Sie: $\sin(x)$ ist in 0 stetig.
Tip: benutzen Sie die Reihe für den Sinus 34.
3. Sei $f : D \rightarrow \mathbb{R}$ stetig und $f(x_0) \neq 0$. Zeigen Sie bitte: es gibt ein $\delta > 0$, so daß f auf $D \cap]x_0 - \delta, x_0 + \delta[$ keine Nullstelle hat.
Tip: Wählen Sie das Kriterium Theorem 3.10 c) für $\varepsilon = \frac{f(x_0)}{3}$ und unterscheiden Sie die zwei Fälle $f(x_0) > 0$ und $f(x_0) < 0$.

Anleitung zum FuncPlot-Applet Im folgenden sehen Sie Beispiele stetiger Funktionen $f :]a, b[\rightarrow \mathbb{R}$, die auf $]a, b[$ ihr Vorzeichen wechseln, also zum Beispiel $f(u) < 0$ und $f(v) > 0$ erfüllen ($u < v$). Können sie dazwischen die Null überspringen oder müssen sie irgendwo eine Nullstelle haben?

Beispiele:

1. $f(x) = x^2 - 1/2$, $D = [0, 1]$, $u = 0$, $v = 1$.
2. $f(x) = \sin(x) - \frac{1}{2}$, $D = [0, \pi/2]$, $u = 0$, $v = \pi/2$.
3. $f(x) = 1_{[0,1[} - 1_{[1,2]}$, $D = [0, 2]$, $u = 0$, $v = 2$. Gibt es hier eine Nullstelle? Was ist anders als bei den anderen Beispielen?

Tatsächlich gilt der folgende Nullstellensatz.

Theorem 3.12 (Nullstellensatz für stetige Funktionen)

Sei $f : D =]a, b[\rightarrow \mathbb{R}$ stetig, u und v seien aus D und $f(u)f(v) < 0$. Dann gibt es ein w echt zwischen u und v mit $f(w) = 0$.

Beweis: Der Beweis wird durch einen Algorithmus zur Nullstellenbestimmung geliefert, durch das **Bisektionsverfahren**. $f(x)f(y) < 0$ liegt genau dann vor, wenn $f(x)$ und $f(y)$ verschiedene Vorzeichen haben. Das nutzen wir aus.

Sei $u_0 = u$, $v_0 = v$; $J_0 = [u, v]$ liegt in D , hat die Länge $l_0 = (v_0 - u_0)/2^0$ und es gilt $f(u_0)f(v_0) < 0$. Seien u, u_1, \dots, u_n und v_1, \dots, v_n schon konstruiert mit den Eigenschaften:

$$\begin{aligned} u_0 &\leq u_1 \leq \dots \leq u_n < v_n \leq v_{n-1} \leq \dots \leq v_0 \\ l_n &= v_n - u_n = l_0/2^n \\ \text{und } f(u_n)f(v_n) &\leq 0. \end{aligned}$$

Ist $f(u_n)f(v_n) = 0$, so hat man bereits eine Nullstelle und das Verfahren bricht ab.

Andernfalls sei $c = (u_n + v_n)/2$. Dann gilt $u_n < c < v_n$. Ist $f(u_n)f(c) = 0$, so ist c eine Nullstelle, das Verfahren bricht ab. Ist $f(u_n)f(c) > 0$, so setze $u_{n+1} = c$, $v_{n+1} = v_n$, andernfalls setze $u_{n+1} = u_n$, $v_{n+1} = c$. Offensichtlich erfüllt u_{n+1}, v_{n+1} alle Formeln mit $(n + 1)$ statt n .

Angenommen das Verfahren bricht nicht ab.

Sei dann $u = \lim_{n \rightarrow \infty} u_n$. Da $(u_n)_{n \geq 0}$ monoton wachsend und durch v_0 beschränkt ist, existiert der Grenzwert. Wegen $0 < v_n - u_n < l_0/2^n$ ist $\lim_{n \rightarrow \infty} v_n = u$. Also ist (nach den Rechenregeln für konvergente Folgen) wegen der Stetigkeit von f : $f(u)^2 = \lim_{n \rightarrow \infty} f(u_n)f(v_n) \leq 0$, und damit $f(u)^2 = 0$.

Aufgabe: Schreiben Sie ein Programm für diesen Algorithmus und testen Sie es an den obigen Beispielen.

Dieser einfache Satz hat weitreichende Folgen.

Korollar 3.13 a) Sei $f : D =]a, b[\rightarrow \mathbb{R}$ stetig, $u, v \in D$ und $u < v$. Dann nimmt f auf $[u, v]$ jeden zwischen $f(u)$ und $f(v)$ liegenden Wert an.

b) Das Bild eines Intervalls unter einer stetigen Funktion ist ein Intervall.

Beweis: a) Sei $f(u) < c < f(v)$ (bzw. $f(u) > c > f(v)$). Wende dann den Nullstellensatz auf $g(x) = f(x) - c$ an.

b) Sei $D =]a, b[$ und $f : D \rightarrow \mathbb{R}$ stetig. Sei

$$c = \begin{cases} -\infty & \text{falls } f(D) \text{ nach unten unbeschränkt ist} \\ \inf f(D) & \text{sonst} \end{cases}$$

Sei analog

$$d = \begin{cases} \infty & \text{falls } f(D) \text{ nach oben unbeschränkt ist} \\ \sup f(D) & \text{sonst} \end{cases}$$

Dann ist $]c, d[\subset f(D) \subset [c, d]$.

Denn ist $c < t < d$, so gibt es $a < u < v < b$ mit $f(u) \leq t \leq f(v)$ oder $f(u) \geq t \geq f(v)$. Nach a) des Korollars gibt es ein x mit $u \leq x \leq v$ und $f(x) = t$ also ist $]c, d[\subset f(D)$. Der Rest ist klar.

Hieraus und aus dem Satz von Bolzano-Weierstraß für Folgen ergibt sich eine weitere Folgerung von besonderer Bedeutung.

Theorem 3.14 *Eine stetige Funktion f auf einem abgeschlossenen beschränkten Intervall $[a, b]$ nimmt dort ihre untere Grenze $\inf(f(D))$ und ihre obere Grenze $\sup(f(D))$ an. Das heißt es gibt ein $u \in [a, b]$ mit $f(u) = \inf(f(D))$ und $v \in [a, b]$ mit $f(v) = \sup(f(D))$. Es gibt also $f([a, b]) = [\inf(f([a, b])), \sup(f([a, b]))]$*

Beweis: (I) $f([a, b])$ ist beschränkt. Denn angenommen $f([a, b])$ wäre nach oben unbeschränkt. Dann gäbe es zu $n \in \mathbb{N}$ ein $u_n \in [a, b]$ mit $f(u_n) > n$. Nach dem Satz von Bolzano-Weierstraß für Folgen 2.14 hat die beschränkte Folge $(u_n)_{n \geq 1}$ eine in $[a, b]$ konvergente Teilfolge $(u_{g(n)})_{n \geq 1}$. Wegen der Stetigkeit konvergiert $(f(u_{g(n)}))_{n \geq 1}$ gegen $f(\lim_n u_{g(n)})$. Andererseits ist $f(u_{g(n)}) > g(n) \geq n$, ein Widerspruch. Analog zeigt man, daß $f([a, b])$ nach unten unbeschränkt ist. (II) Sei $c = \inf f([a, b])$. Zu $n \in \mathbb{N}$ gibt es dann nach Aufgabe 18, Blatt 5 ein $u_n \in [a, b]$ mit $c \leq f(u_n) < c + 1/n$. Wieder hat $(u_n)_{n \geq 1}$ eine gegen ein Element u konvergente Teilfolge $(u_{g(n)})_{n \geq 1}$. Damit ist $c = \lim f(u_n) = f(u)$, wegen der Stetigkeit von f . Analog beweist man die Aussage über die obere Grenze. Der Rest folgt aus dem vorangegangenen Korollar 3.13, Teil b).

3.3.1 Umkehrfunktionen

Wir hatten schon gezeigt, daß es zu jeder Zahl $x \geq 0$ genau eine Zahl $y \geq 0$ mit $y^2 = x$ gibt. Wir erhalten also eine neue Funktion $g :]0, \infty[\rightarrow]0, \infty[$, $x \mapsto \sqrt{x} = g(x)$. Es ist für $f(x) = x^2$: $g \circ f = f \circ g = id_{]0, \infty[}$, d.h. g ist die Umkehrfunktion zu f . Das allgemeine Problem ist: Wann existiert eine Umkehrfunktion zu einer stetigen Funktion und ist diese dann stetig?

Die erste Frage ist, wann sind stetige Funktionen injektiv. Dazu definieren wir:

Definition 3.15 *Sei $J \subset \mathbb{R}$ ein Intervall und $f : J \rightarrow \mathbb{R}$ eine Funktion.*

f heißt (streng) monoton wachsend, wenn für $x < y$ stets $f(x) \leq f(y)$ (bzw. $f(x) < f(y)$) gilt.

Sie heißt (streng) monoton fallend, wenn für $x < y$ stets $f(x) \geq f(y)$ (bzw. $f(x) > f(y)$) gilt.

Eine (streng) monoton wachsende oder fallende Folge nennt man (streng) monoton.

Eine streng monotone Funktion ist injektiv.

Anleitung zum FuncPlot-Applet Schauen Sie sich die folgenden Beispiele an. Welche Funktion ist streng monoton wachsend, welche streng monoton fallend, welche nichts von beiden?

1. $D = \mathbb{R}$, $f(x) = x$

2. $D = \mathbb{R}_+ = [0, \infty[$, $f(x) = x^2$
3. $D = \mathbb{R}$, $f(x) = x^2$
4. $D = [-\pi/2, \pi/2]$, $f(x) = \sin[x]$
5. $D = [-\pi, \pi]$, $f(x) = \sin(x)$.

Wie Sie sehen, hängt die Injektivität nicht nur davon ab, *welche* Funktion man betrachtet, sondern auch, *auf welchem Intervall* man sie betrachtet.

Lemma 3.16 *Eine stetige reelle Funktion f auf einem Intervall J ist genau dann injektiv, wenn sie streng monoton ist.*

Beweis: (I) Sei f streng monoton. *Beh.:* f ist injektiv. Denn sei zunächst f streng monoton wachsend. Dann ist für $x \neq y$, also o.B.d.A. $x < y$ $f(x) < f(y)$ also $f(x) \neq f(y)$. f ist also injektiv. Ist f streng monoton fallend, so ist $-f$ streng monoton wachsend, also injektiv, und damit ist $f = -(-f)$ injektiv.

(II) Sei f injektiv und $x < y$. Wir zeigen zunächst, daß f "keinen Buckel" hat. Genauer: gäbe es ein u mit $x < u < y$ mit $f(u) > f(x), f(y)$, so wäre $f(u) > \max(f(x), f(y)) = \alpha$. Nach dem Zwischenwertsatz, angewendet auf das Intervall $[x, u]$ gibt es ein v mit $x \leq v < u$ und $f(v) = \alpha$. Wendet man den Zwischenwertsatz auf $[u, y]$ an, so gibt es ein w mit $u < w \leq y$ und $f(w) = \alpha = f(v)$; wegen $v \neq w$ ist dies ein Widerspruch zur Injektivität.

f hat aber auch "keine Senke". Genauer: gäbe es ein u mit $x < u < y$ und $f(u) < f(x), f(y)$, so kommt man mit $\beta = \min(f(x), f(y)) > f(u)$ analog zu oben zum Widerspruch (Sie können auch $-f$ betrachten). Also gilt: *Ist $f(x) < f(y)$, so ist f auf $[x, y]$ streng monoton wachsend. Ist $f(x) > f(y)$, so ist f auf $[x, y]$ streng monoton fallend.*

(III) *Behauptung:* Ist f injektiv, so ist f streng monoton.

Beweis: Wäre f auf J nicht streng monoton, so gäbe es ein Intervall $[u, v] \subset J$, auf dem f streng monoton wachsend ist und ein Intervall $[x, y]$, auf dem f streng monoton fallend ist. Nach (II) ist $[u, v] \cap [x, y] = \emptyset$. Auf dem Intervall $[\min(u, x), \max(v, y)]$ ist dann f nicht streng monoton im Widerspruch zu (II). Also ist f auf ganz J streng monoton.

Daraus folgt

Theorem 3.17 Stetigkeit der Umkehrfunktion

Sei $f: J \subset \mathbb{R} \rightarrow \mathbb{R}$ ein Intervall und $f: J \rightarrow \mathbb{R}$ stetig und injektiv. Dann ist die Umkehrabbildung f^{-1} vom Bildintervall $f(J)$ auf J ebenfalls stetig.

Beweis: (I) Wir nehmen zunächst $J = \mathbb{R}$ an und weisen nach, daß das $\varepsilon - \delta$ -Kriterium erfüllt ist:

Sei $x_0 \in J, y_0 = f(x_0)$ und $\varepsilon > 0$ beliebig vorgegeben. Nach dem vorangegangenen Lemma ist f streng monoton; wir nehmen streng monoton wachsend an (sonst betrachte $-f$). Dann ist nach dem Zwischenwertsatz Korollar 3.13 das Intervall $]f(x_0 - \varepsilon), f(x_0 + \varepsilon)[$ enthalten im Bildintervall $f(]x_0 - \varepsilon, x_0 + \varepsilon[)$ und es gilt $f(x_0 - \varepsilon) < y_0 = f(x_0) < f(x_0 + \varepsilon)$. Sei $\delta = \min(y_0 - f(x_0 - \varepsilon), f(x_0 + \varepsilon) - y_0)$. Dann ist

$$]y_0 - \delta, y_0 + \delta[\subset]f(x_0 - \varepsilon), f(x_0 + \varepsilon)[\subset f(]x_0 - \varepsilon, x_0 + \varepsilon[)$$

also $f^{-1}(]y_0 - \delta, y_0 + \delta[) \subset]x_0 - \varepsilon, x_0 + \varepsilon[$. Das $\varepsilon - \delta$ -Kriterium für Stetigkeit ist damit erfüllt.

(II) Ist $J \neq \mathbb{R}$, aber offen, etwa $J =]a, b[$ so muß man, falls ε so groß war, daß $]x_0 - \varepsilon, x_0 + \varepsilon[$ nicht ganz in J liegt, ein $\varepsilon' < \varepsilon$ wählen, so daß $]x_0 - \varepsilon', x_0 + \varepsilon'[\subset J$. Dann geht alles genauso.

(III) Ist $J = [a, b[$ und $x_0 \neq a$, so verfährt man wie unter (II). Ist $x_0 = a$, so muß man nur $[a, a + \varepsilon[$ betrachten; analog schließt man, wenn $J =]a, b]$ bzw. $J = [a, b]$ ist.

Beispiele:

1. Sei n eine gerade und natürliche Zahl, $n = 2m$. Dann ist $f(x) = x^{2m}$ auf $\mathbb{R}_+ =]0, \infty[$ stetig und streng monoton wachsend nach Satz 1.20 f. Die Umkehrfunktion ist die n . Wurzel. Sie ist stetig.
2. Sei n eine ungerade Zahl, $n = 2m + 1$. $x \rightarrow x^n$ ist dann stetig und auf ganz \mathbb{R} streng monoton. Denn für $x < y$. können wir 3 Fälle unterscheiden: (i): Ist $0 \leq x$, so ist $0 \leq x^n < y^n$ nach Satz 1.20 f. (ii): ist $x < 0 < y$ so ist $x^n = (-|x|)^n = (-1)^n |x|^n = -|x|^n < 0 < y^n$, wobei das 3. Gleichheitszeichen gilt, weil n ungerade ist. (iii): ist $x < y \leq 0$, so ist $0 \leq -y < -x$, also $0 \leq -y^n = (-y)^n < (-x)^n = -x^n$. Multiplikation mit -1 liefert die Behauptung. Die Umkehrfunktion ist auf \mathbb{R} erklärt und ist die n .Wurzel $\sqrt[n]{y}$. Sie ist stetig.
3. Es ist $x \rightarrow \exp(x)$ stetig (siehe Satz 3.23) und streng monoton von \mathbb{R} auf $\mathbb{R}_+ \setminus \{0\} =]0, \infty[$. Die Umkehrfunktion ist also definiert auf $]0, \infty[$ und hat als Bild ganz \mathbb{R} . Sie ist stetig und heißt Logarithmus naturalis $\ln(x)$. Es gilt

$$\ln(xy) = \ln(x) + \ln(y).$$

4. Sei $a > 0$ beliebig. Man setzt $a^x = \exp(x \ln(a))$. Damit ist für $a \neq 1$ die Funktion $x \in \mathbb{R} \mapsto a^x \in]0, \infty[$ streng monoton und stetig (als Hintereinanderausführung stetige Funktionen, $x \rightarrow x \ln(a)$ und $y \rightarrow \exp(y)$). Die Umkehrfunktion heißt $\log_a(x)$, gelesen: Logarithmus x zur Basis a . Auch hier gilt

$$\log_a(xy) = \log_a(x) + \log_a(y).$$

Besonders wichtig für die Informatik ist $a = 2$.

Wiederholung:

- Folgenkriterium und $\varepsilon - \delta$ -Kriterium für Stetigkeit;
- Rechenregeln für stetige Funktionen (einschließlich Hintereinanderausführung);
- Nullstellensatz, Zwischenwertsatz, Minimum-Maximum-Satz;
- Umkehrfunktion einer stetigen Funktion.

3.4 Gleichmäßige Stetigkeit

Überblick: Sei $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ stetig. Dann gibt es zu jedem $x_0 \in D$ und zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß für alle $y \in D$ mit $|y - x_0| < \delta$ stets $|f(y) - f(x_0)| < \varepsilon$ ist. Im allgemeinen hängt das δ nicht nur von ε sondern auch noch von der Stelle x_0 ab, wie wir anhand von Beispielen zeigen. Gleichmäßig stetig heißt eine Funktion, wenn man zu jedem $\varepsilon > 0$ ein $\delta > 0$ unabhängig von der speziellen Stelle x_0 finden kann. Stetige Funktionen auf abgeschlossenen beschränkten Intervallen sind gleichmäßig stetig. Um das Folgende besser zu verstehen, wiederholen wir noch einmal das $\varepsilon - \delta$ -Kriterium für Stetigkeit (Theorem 3.10):

Sei $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ eine beliebige Funktion. f ist im Punkt $x \in D$ genau dann stetig, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß für alle $y \in D$ mit $|y - x| < \delta$ stets $|f(y) - f(x)| < \varepsilon$ gilt. Im allgemeinen hängt das δ nicht nur von ε ab, sondern auch vom betrachteten Punkt x , wie die folgenden Beispiele illustrieren:

Beispiele:

1. Sei $f(x) = x^2$ und $D = [1, \infty[$. Dann ist f nach den Rechenregeln für stetige Funktionen als Quadrat von $g(x) = x$ stetig. Sei $x \in \mathbb{R}$ beliebig. Zu $\varepsilon > 0$ gibt es also ein $\delta > 0$ mit $|y^2 - x^2| < \varepsilon$ für $|y - x| < \delta$. *Behauptung:*

$$\delta < \frac{\varepsilon}{\sqrt{x}}.$$

Beweis: Angenommen es ist $\delta \geq \frac{\varepsilon}{\sqrt{x}}$. Dann gilt für $y = x + \delta/2$ natürlich $|y - x| < \delta$ aber

$$|y^2 - x^2| = (y + x)(y - x) = (2x + \delta/2)\delta/2 > x\delta \geq \sqrt{x}\varepsilon \geq \varepsilon,$$

ein Widerspruch. Damit läßt sich kein von x unabhängiges $\delta > 0$ finden, so daß $|y^2 - x^2| < \varepsilon$ für $|y - x| < \delta$ gilt. Zum Beispiel muß für $x = 10^6$ und $\varepsilon = 1$ das δ kleiner als 10^{-3} , für $x = 10^{12}$ kleiner als 10^{-6} sein.

2. Sei $f(x) = \frac{1}{x}$ auf $[1, \infty[= D$. Dann genügt es wegen $x, y \geq 1$, für das vorgegebene $\varepsilon > 0$ $\delta = \varepsilon$ *unabhängig von der Stelle x* zu wählen: Ist $x, y \in D$ mit $|x - y| < \varepsilon$, so ist $|\frac{1}{x} - \frac{1}{y}| = \frac{1}{xy} |x - y| \leq |x - y| < \varepsilon$.
3. Genau so ist es, wenn wir uns in Beispiel 1 auf das Intervall $[-b, b]$ mit $b > 0$ fest gewählt beschränken. Ist dann $\varepsilon > 0$ und $\delta = \frac{\varepsilon}{2b}$, so gilt für alle $x, y \in [-b, b]$ stets: Ist $|x - y| < \delta$, so ist $|x^2 - y^2| < \varepsilon$. Denn

$$|x^2 - y^2| = |x - y| |x + y| \leq |x - y| (|x| + |y|) \leq 2b |x - y| < \varepsilon.$$

4. Im folgenden Beispiel kann man wie in Beispiel 1 kein von x unabhängiges δ finden: Sei $f(x) = 1/x$ auf dem Intervall $]0, 1] = D$. Sei $x \in D$ und $\varepsilon > 0$ vorgegeben. Ein zu ε passendes δ (das es ja gibt wegen der Stetigkeit von $1/x$) muß $\delta < 2x\varepsilon$ erfüllen. Denn angenommen, es ist $\delta \geq 2x\varepsilon$. Dann ist für $y = x - \delta/2$ einerseits $|y - x| < \delta$, andererseits

$$\begin{aligned} \left| \frac{1}{y} - \frac{1}{x} \right| &= \frac{1}{(x - \delta/2)x} (x - y) \\ &= \frac{1}{2(x - \delta/2)x} \\ &> 2x \frac{\varepsilon}{2x^2} \\ &= \varepsilon/x \\ &\geq \varepsilon, \end{aligned}$$

ein Widerspruch. Je näher x (bei festem ε) bei 0 liegt, desto kleiner fällt das passende $\delta < 2x\varepsilon$ aus, und da x beliebig nahe bei 0 liegen kann, gibt es kein von x unabhängiges δ .

Wir definieren

Definition 3.18 Sei $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion. f heißt *gleichmäßig stetig*, wenn es zu jedem $\varepsilon > 0$ ein $\delta(\varepsilon) > 0$ gibt, so daß für alle $x, y \in D$ aus $|x - y| < \delta(\varepsilon)$ stets $|f(x) - f(y)| < \varepsilon$ folgt.

Gleichmäßig stetige Funktionen sind natürlich stetig. Die Umkehrung gilt nicht (s. Beispiele 1 und 2 oben), aber sie gilt unter einer Einschränkung:

Theorem 3.19 (Stetigkeit auf abgeschlossenen beschränkten Intervallen)

Sei $D = [a, b]$ ein abgeschlossenes beschränktes Intervall. Dann ist jede stetige Funktion $f : D \rightarrow \mathbb{R}$ gleichmäßig stetig.

Beweis: (indirekt) Angenommen, f ist nicht gleichmäßig stetig. Dann gibt es $\varepsilon_0 > 0$ so daß für jedes $\delta > 0$ ein Paar (x, y) existiert mit $|x - y| < \delta$ aber $|f(x) - f(y)| \geq \varepsilon_0$. Zu $\delta_n = \frac{1}{n}$ sei (x_n, y_n) solch ein Paar. Nach dem Satz von Bolzano-Weierstraß (Theorem 2.14) gibt es eine konvergente Teilfolge $(x_{g(n)})_n$; deren Grenzwert $\lim x_{g(n)} = c$ wegen $a \leq x_{g(n)} \leq b$ für alle n selbst in $[a, b]$ liegt. Dann gilt aber wegen $|y_{g(n)} - x_{g(n)}| < \frac{1}{g(n)} \leq \frac{1}{n}$ auch $\lim y_{g(n)} = c$, also

$$0 = |f(c) - f(c)| = \lim_{n \rightarrow \infty} \underbrace{|f(y_{g(n)}) - f(x_{g(n)})|}_{\geq \varepsilon_0} \geq \varepsilon_0 > 0,$$

ein Widerspruch.

Aufgaben:

1. Zeigen Sie bitte: Ist f gleichmäßig stetig auf D und ist $(x_n)_n \subset D$ eine konvergente Folge, deren Grenzwert aber nicht in D liegen muß, so ist $(f(x_n))$ (dennoch) konvergent, d.h. $(f(x_n))_n$ erfüllt das Cauchy Kriterium.
2. Zeigen Sie bitte: Sei $D =]0, 1]$ und $f(x) = \frac{1}{\sqrt{x}}$. f ist nicht gleichmäßig stetig.
3. Zeigen Sie bitte:
 - a) $\cos(2t) = \cos(t)^2 - \sin(t)^2$
Tip zu a: $e^{2it} = (e^{it})^2$.
 - b) $|e^{is} - e^{it}| \leq 2|\sin(\frac{t-s}{2})|$.
Tip zu b: Beweisen Sie $|e^{is} - e^{it}|^2 = 2(1 - \cos(t-s))$. Setzen Sie nun $(t-s) = 2v$ und verwenden Sie a).
 - c) Benutzen Sie die Beziehung aus der Schule: $|\sin t| \leq |t|$ um zu zeigen: $|\sin(s) - \sin(t)| \leq |s - t|$.
 Leiten Sie hieraus ab: $\sin(t)$ ist auf \mathbb{R} gleichmäßig stetig.

Wiederholung:

- Definition und Theorem über gleichmäßige Stetigkeit;
- Beispiele für stetige, nicht gleichmäßig stetige Funktionen.

3.5 Gleichmäßige Konvergenz

Überblick: Wir haben bisher implizit schon Folgen von Funktionen betrachtet, als wir Potenzreihen betrachteten. $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ entspricht der Folge $(S_n(x))$ mit $S_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$. Wir wissen, daß die Folge $(S_n(x))_n$ für festes x konvergiert. Das *neue Problem* ist: ist die "Konvergenzgeschwindigkeit" unabhängig von x ? Ist das der Fall, so spricht man von gleichmäßiger Konvergenz der Folge $(S_n)_{n \in \mathbb{N}}$ von Funktionen. Konvergiert eine Folge stetiger Funktionen gleichmäßig gegen eine Funktion, so ist diese Grenzfunktion stetig. Damit zeigt man: durch Potenzreihen gegebene Funktionen sind stetig.

Definition 3.20 *Gibt es eine Zuordnungsvorschrift, die jeder natürlichen Zahl n eine Funktion $f_n : D \rightarrow \mathbb{R}$ zuordnet, so heißt $(f_n)_{n \in \mathbb{N}}$ eine Folge von Funktionen oder kurz: **Funktionsfolge**.*

Beispiele:

1. $D = \mathbb{R}, \quad f_n(x) = x^n$
2. $D = \mathbb{R}, \quad f_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$
3. $D =]-1, 1[, \quad f_n(x) = \frac{1-x^{n+1}}{1-x}$
4. $D = [-\frac{1}{2}, \frac{1}{2}], \quad f_n(x) = x^n$

Im 2. bzw. 4. Beispiel konvergiert die Folge $(f_n(x))$ für jedes feste x . Den Grenzwert bezeichnen wir mit $f(x)$. Es gibt also zu jedem $\varepsilon > 0$ ein $n(\varepsilon, x)$ mit $|f(x) - f_n(x)| < \varepsilon$ für alle $n \geq n(\varepsilon, x)$. Die Frage ist, ob es ein *von x unabhängiges* $n(\varepsilon)$ gibt.

Aufgaben: Testen Sie in den folgenden Beispielen, ob es zu jedem $\varepsilon > 0$ ein von $x \in D$ unabhängiges $n(\varepsilon)$ mit $|f(x) - f_n(x)| < \varepsilon$ für alle $n \geq n(\varepsilon)$ und für alle x ("gleichzeitig") gibt es :

1. $D = [-\frac{1}{2}, \frac{1}{2}], \quad f_n(x) = x^n.$
2. $D =]-1, 1[, \quad f_n(x) = x^n$
3. $D = [-\frac{1}{2}, \frac{1}{2}], \quad f_n(x) = \sum_{k=0}^n x^k.$
4. $D =]-1, 1[, \quad f_n(x) = \sum_{k=0}^n x^k.$
5. $D = [0, 1], \quad f_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$

Anleitung zum FuncPlot-Applet Schauen Sie sich ein paar der Funktionen der Aufgaben an, d.h. lassen Sie sich die Differenzen $f(x) - f_n(x)$ für $n = 1, 10, 20, 30$ zeichnen. Wie sehen die Differenzen aus, wenn das $n(\varepsilon)$ unabhängig von x existiert, wie, wenn dies nicht der Fall ist?

Um die Fragestellung zu präzisieren, definieren wir.

Definition 3.21 (Konvergenz punktweise, gleichmäßige Konvergenz)

Sei $(f_n)_n$ eine Folge reeller Funktionen auf der Menge $D \subset \mathbb{R}$.

- a) $(f_n)_n$ **konvergiert punktweise** gegen $f : D \rightarrow \mathbb{R}$, wenn für jedes $x \in D$ stets $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ gilt.
- b) $(f_n)_n$ **konvergiert gleichmäßig** gegen $f : D \rightarrow \mathbb{R}$, wenn es zu jedem $\varepsilon > 0$ ein (von x unabhängiges) $n(\varepsilon) \in \mathbb{N}$ gibt mit $|f(x) - f_n(x)| < \varepsilon$ für alle $n \geq n(\varepsilon)$ und alle x .

Anleitung zum FuncPlot-Applet Wählen Sie die Beispiele aus der oben stehenden Aufgabe. Zeichnen Sie für $\varepsilon = 0.1$ den " ε -Schlauch" um f , also die Funktionen $f_-(x) = f(x) - \varepsilon$, f und $f_+(x) = f(x) + \varepsilon$. Gleichmäßige Konvergenz bedeutet, daß ab $n(\varepsilon)$ alle f_n in diesem Streifen verlaufen. Testen Sie dies an Beispielen. Sie können stattdessen auch die konstante Funktion $g(\varepsilon) = \varepsilon$ und die Differenzen $|f(x) - f_n(x)|$ zeichnen lassen. Bei gleichmäßiger Konvergenz müssen die Differenzen ab $n(\varepsilon)$ ganz unterhalb von g verlaufen.

Der Begriff der gleichmäßigen Konvergenz ist wegen der folgenden beiden Sätze wichtig:

Theorem 3.22 (Gleichmäßige Konvergenz und Stetigkeit)

Sei $(f_n)_n$ eine Folge von Funktionen $f_n : D \rightarrow \mathbb{R}$, die gleichmäßig gegen die Funktion $f : D \rightarrow \mathbb{R}$ konvergiert. Dann ist f stetig.

Beweis: Wir benutzen das $\varepsilon - \delta$ -Kriterium. Sei $x_0 \in D$ und $\varepsilon > 0$ beliebig vorgegeben. Dann gibt es wegen der gleichmäßigen Konvergenz ein $n_0 = n(\varepsilon/3)$ mit $|f(x) - f_n(x)| < \varepsilon/3$ für alle $n \geq n_0$. Da f_{n_0} stetig ist, gibt es ein $\delta'(\varepsilon/3)$ mit $|f_{n_0}(x) - f_{n_0}(x_0)| < \varepsilon/3$ für alle $x \in D$ mit $|x - x_0| < \delta'(\varepsilon/3)$. Wir setzen $\delta(\varepsilon) = \delta'(\varepsilon/3)$ und erhalten für $|x - x_0| < \delta(\varepsilon)$

$$\begin{aligned} |f(x) - f(x_0)| &< |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(x_0)| + |f_{n_0}(x_0) - f(x_0)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\ &= \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.

Satz 3.23 (Stetigkeit von Potenzreihen)

Sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann ist f auf $] -R, R[$ stetig.

Beweis: Sei $|x_0| < R$ und $t \in \mathbb{R}$ mit $|x_0| < t < R$ fest gewählt. Dann ist für $S_n(x) = \sum_{k=0}^n a_k x^k$ mit $|x| \leq t$

$$\begin{aligned} |f(x) - S_n(x)| &= \left| \lim_{p \rightarrow \infty} \sum_{k=n+1}^{n+p} a_k x^k \right| \\ &\leq \lim_{p \rightarrow \infty} \sum_{k=n+1}^{n+p} |a_k| t^k \\ &= \sum_{k=n+1}^{\infty} |a_k| t^k; \end{aligned}$$

(beachten Sie, daß $t < R$, die Reihe $\sum_{n=0}^{\infty} a_n t^n$ also absolut konvergiert.) Zu $\varepsilon > 0$ gibt es ein $n(\varepsilon)$ mit $\sum_{k=n+1}^{n+p} |a_k| t^k < \varepsilon/2$ für alle $n \geq n(\varepsilon)$ und alle $p \geq 1$, also

$$\sum_{k=n+1}^{\infty} |a_k| t^k = \lim_{p \rightarrow \infty} \sum_{k=n+1}^{n+p} |a_k| t^k < \varepsilon.$$

Damit gilt für alle x mit $|x| \leq t$ und alle $n \geq n(\varepsilon)$ $|f(x) - S_n(x)| < \varepsilon$.

Das bedeutet: (S_n) konvergiert auf $[-t, t]$ gleichmäßig gegen f . Damit ist f nach Theorem 3.22 auf $[-t, t]$ stetig, insbesondere ist f in x_0 wegen $|x_0| < t$ stetig.

Als Anwendung des Theorems 3.22 über die gleichmäßige Konvergenz können wir statt Potenzreihen auch trigonometrische Reihen betrachten:

Aufgaben:

1. Sei $(a_n)_{n \geq 0}$ eine Folge, für die die Reihe $\sum_{n=0}^{\infty} |a_n|$ konvergiert. Zeigen Sie bitte:
Die Reihen $\sum_{n=1}^{\infty} a_n \sin(nt)$ und $\sum_{n=0}^{\infty} a_n \cos(nt)$ konvergieren gleichmäßig auf ganz \mathbb{R} . Die Grenzfunktionen f und g sind also stetig (und 2π -periodisch, d.h. $f(t + 2\pi) = f(t)$, ebenso g).
Tipp: $|\cos(t)| \leq 1$ und $|\sin(t)| \leq 1$. Wie immer: probieren Sie sich den Satz zunächst mal an einem Beispiel zu verdeutlichen, etwa $a_n = 1/2^n$.
2. Formulieren Sie den Stetigkeitsbegriff für komplexwertige Funktionen $f : D \subset \mathbb{R} \rightarrow \mathbb{C}$ und zeigen Sie: f ist in t_0 genau dann stetig, wenn $\Re f : t \rightarrow \Re f(t)$ und $\Im f : t \rightarrow \Im f(t)$ in t_0 stetig sind.
3. Zeigen Sie, daß $t \rightarrow \exp(it)$ stetig ist.

4. Sei $(a_n)_{n \in \mathbb{Z}}$ eine komplexwertige Folge mit Indizes aus \mathbb{Z} , also eine Abbildung von \mathbb{Z} in \mathbb{C} . Wir setzen jetzt $S_n(t) = \sum_{k=-n}^n a_k e^{-ikt}$ und schreiben

$$f(t) = \sum_{k=-\infty}^{\infty} a_k e^{-ikt} = \lim_{n \rightarrow \infty} S_n(t),$$

falls der Grenzwert existiert. Formulieren Sie, was es bedeuten soll, daß die Summe gleichmäßig auf \mathbb{R} gegen f konvergiert. Zeigen Sie dann, daß im Fall der gleichmäßigen Konvergenz die Funktion $f: \mathbb{R} \rightarrow \mathbb{C}$, $t \mapsto f(t)$ stetig ist.

Tipp: Es ist leichter, den Beweis von Theorem 3.22 zu imitieren, als sich auf den Realteil und Imaginärteil der Funktion zurückzuziehen. Aber auch dieser Weg ist gangbar. Solche Reihen heißen **(komplexe) Fourierreihen**. Sie spielen in der Physik und Informatik eine wichtige Rolle.

Wiederholung:

- Definition der gleichmäßigen Konvergenz;
- Beispiele für gleichmäßige und nicht gleichmäßige Konvergenz;
- Theorem über gleichmäßige Konvergenz und Stetigkeit;
- Stetigkeit von Potenzreihen.

Kapitel 4

Differentialrechnung

Überblick: Im ersten Abschnitt dieses Kapitels führen wir die Ableitung (Differentiation) ein und stellen die elementaren Regeln zur Berechnung der Ableitung zusammen. Schließlich führen wir höhere Ableitungen ein. Der zweite Abschnitt ist dem Kernstück der Differentialrechnung - dem *Mittelwertsatz* - gewidmet. Er ist grundlegend für alles folgende. Der dritte Abschnitt behandelt die "Entwicklung" einer Funktion um einen Punkt herum in ein Polynom mit Restglied, den *Satz von Taylor*. Im letzten Abschnitt werden wichtige Funktionsklassen untersucht.

4.1 Ableitung (Differentiation) einer Funktion

Überblick: Die Ableitung einer Funktion f in einem Punkt c ist der Grenzwert der Steigungen der Sekanten an den Graphen $G(f)$, die durch die Punkte $(c, f(c))$ und $(x, f(x))$, $x \neq c$ laufen. Äquivalente Charakterisierungen der Existenz der Ableitung erlauben die Begründung der elementaren Rechenregeln der Differentialrechnung. Wir schließen diesen Abschnitt mit dem Begriff der höheren Ableitung.

Die grundlegende Idee der Ableitung

Zunächst wiederholen wir aus der Schule die "Geradengleichung": eine Gerade durch den Punkt (c, d) der Ebene \mathbb{R}^2 mit Steigung m ist durch die Funktion $g(x) = d + m(x - c)$ gegeben, das heißt die Gerade ist der Graph $G(g) = \{(x, g(x)) : x \in \mathbb{R}\}$ der Funktion g . Eine Gerade, die durch die beiden Punkte (x_0, y_0) und (x_1, y_1) mit $x_1 \neq x_0$ geht, ist durch

$$g(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)$$

gegeben. Ihre Steigung ist $m = \frac{y_1 - y_0}{x_1 - x_0}$. Sei nun $J \subset \mathbb{R}$ ein Intervall und $f : J \rightarrow \mathbb{R}$ eine Funktion, ferner $c \in J$. Eine **Sekante** (Schneidende) der Kurve $G(f)$ durch die Punkte $(c, f(c))$ und $(x_1, f(x_1))$, $x_1 \neq c$ ist dann die Gerade durch die Punkte $(c, f(c))$ und $(x_1, f(x_1))$, gegeben durch die Funktion

$$s(x) = f(c) + \frac{f(x_1) - f(c)}{x_1 - c}(x - c).$$

Ihre Steigung ist also $m = \frac{f(x_1) - f(c)}{x_1 - c}$.

Als Steigung der **Tangente** an die Kurve $G(f)$ im Punkt $(c, f(c))$ wird man nun intuitiv die Steigung einer Sekante ansehen, deren zweiter Punkt $(x_1, f(x_1))$ "unendlich dicht" bei $(c, f(c))$ liegt. Um dies zu präzisieren, erklären wir:

Definition 4.1 Sei J ein Intervall und $f : J \rightarrow \mathbb{R}$ eine Funktion.

a) f heißt im Punkt $c \in J$ **differenzierbar** oder **ableitbar**, wenn der Grenzwert

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

existiert. Dieser Grenzwert heißt dann **Ableitung** oder **Differentialquotient** von f an der Stelle c und wird mit $f'(c)$ bezeichnet. Andere Bezeichnungen sind $\frac{df}{dx}(c)$ oder $\dot{f}(c)$.

b) f heißt **differenzierbar**, wenn f in jedem Punkt $c \in J$ differenzierbar ist.

Anleitung zum FuncPlot-Applet

a) Schauen Sie sich selbst den Übergang von Sekanten zur Tangente in den folgenden Beispielen an. Geben Sie dazu neben der Funktion einfach die Sekantengleichungen ein (zwei bis drei). Wählen Sie dann einen zweiten Punkt sehr dicht beim ersten Punkt. Sie "sehen" dann tatsächlich eine Tangente, sofern die Funktion differenzierbar ist.

1. $f(x) = 2x, c = 0, x_1 = 1, 1/2, 1/100$. Warum sieht man hier immer wieder dasselbe?
2. $f(x) = x^2, c = 0, x_1 = 1, -1/2, 1/10, -1/100$.
3. $f(x) = \sin(x), c, x_1$ wie unter 2.
4. $f(x) = \cos(x), c, x_1$ wie unter 2.

b) Zeichnen Sie nun einmal die oben genannten Funktionen in sehr kleiner Umgebung $[c - 1/100, c + 1/100]$ um c . Lassen Sie die Sekante mit $x_1 = c + 1/100$ zeichnen. Können Sie einen Unterschied zwischen der Funktion und der Sekante erkennen?

Die Zeichnungen aus b) legen nahe, daß in sehr kleinen Intervallen um c die Funktion durch ihre Tangente angenähert wird. Das ist tatsächlich im folgenden präzisen Sinn der Fall:

Satz 4.2 (Äquivalente Charakterisierung der Differenzierbarkeit)

Sei J ein Intervall, $f : J \rightarrow \mathbb{R}$ eine Funktion und $c \in J$ ein Punkt. Die folgenden Aussagen sind äquivalent:

- a) f ist in c differenzierbar.
- b) Es gibt eine in c stetige Funktion R auf J mit $R(c) = 0$, sowie eine Konstante d , so daß für alle $x \in J$

$$f(x) = \underbrace{f(c) + d(x - c)}_{\text{Gerade durch } (c, f(c)) \text{ mit Steigung } d} + R(x)(x - c)$$

gilt.

c) Es gibt ein in c stetige Funktion S auf J mit

$$f(x) = f(c) + S(x)(x - c) \text{ für alle } x \in J.$$

Gilt c) (und damit auch b) und a)), so ist $S(c) = f'(c)$. Die Funktion S ist durch f eindeutig bestimmt.

Gilt b) (und damit auch a) und c)), so ist $d = f'(c)$.

Erläuterung:

b) besagt gerade, wie groß der Fehler ist, wenn man statt der Funktion f die Tangente an $G(f)$ durch den Punkt $(c, f(c))$ betrachtet. Beachten Sie $R(c) = 0$. Das bedeutet, der Fehler $R(x)(x - c)$ ist eine "Größenordnung" kleiner als $x - c$.

c) klingt etwas theoretisch, ist aber die bequemste Form, um später die Rechenregeln zu begründen.

Beweis: a) \Rightarrow b): Wir setzen

$$R(x) = \begin{cases} 0 & x = c \\ \frac{f(x) - f(c)}{x - c} & x \neq c \end{cases}.$$

Es gilt dann nach Voraussetzung $\lim_{x \rightarrow c} R(x) = R(c) = 0$, das heißt R ist stetig in c . Weiter gilt offensichtlich die Formel b).

b) \Rightarrow c): Setze einfach $S(x) = R(x) + d$.

c) \Rightarrow a): Für $x \neq c$ ist

$$S(x) = \frac{f(x) - f(c)}{x - c}.$$

Da S in c stetig ist, muß auch der Grenzwert der rechten Seite existieren und gleich $S(c)$ sein. Der Rest ist klar.

Anleitung zum Applet "Animierte Differentiation" Schauen Sie sich das "Anschmiegen" der Sekante an eine differenzierbare Funktion in "animierter Form" an.

Korollar 4.3 ("differenzierbar impliziert stetig")

Sei $f : J \rightarrow \mathbb{R}$ in c differenzierbar. Dann ist f in c stetig.

Beweis: Nach c) ist $f(x) = f(c) + S(x)(x - c)$. Nach den Rechenregeln für stetige Funktionen folgt die Behauptung.

Beispiele:

1. $f(x) = 5$ für alle x . $f'(x) = 0$ für alle x . Allgemeiner: Ist f konstant, so ist $f' = 0$.
2. $f(x) = 2.7x + 10$. Dann ist $f'(x) = 2.7$ für alle x , das heißt, die Ableitung ist konstant. Allgemeiner: $f(x) = m(x - x_0) + d \Rightarrow f'(x) = m$ für alle x .
3. $f(x) = 3x^2 = 3c^2 + 3(x + c)(x - c)$, also $S(x) = 3(x + c)$ und damit $f'(c) = S(c) = 6c$. Allgemeiner: $f(x) = ax^2 \Rightarrow f'(c) = 2ac$ für alle $c \in \mathbb{R}$.
4. Noch allgemeiner: sei $f(x) = ax^n$ und $c \in \mathbb{R}$ beliebig. Dann ist

$$f(x) = ac^n + a(x^{n-1} + x^{n-2}c + \dots + c^{n-1})(x - c),$$

also $S(x) = a \sum_{k=0}^{n-1} x^{n-1-k} c^k$ und daher

$$f'(c) = S(c) = anc^{n-1}.$$

5. $f(x) = 1/x$ auf $\mathbb{R} \setminus \{0\}$. Sei $c \neq 0$ beliebig. Dann ist

$$1/x = 1/c - \frac{1}{xc}(x - c).$$

Also ist $S(x) = -1/xc$ und damit $f'(c) = S(c) = -1/c^2$.

Rechenregeln für die Differentiation

Die in den Beispielen erhaltenen Formeln für die Ableitung erhält man auch durch wiederholte Anwendung der folgenden Rechenregeln, die sich ihrerseits einfach aus den Rechenregeln für Grenzwerte ergeben.

Satz 4.4 Rechenregeln für die Differentiation

Sei J ein Intervall.

a) Sei $f : J \rightarrow \mathbb{R}$ konstant. Dann ist $f'(c) = 0$ für alle $c \in J$.

b) Seien $f, g : J \rightarrow \mathbb{R}$ in c differenzierbar und seien $\alpha, \beta \in \mathbb{R}$. Dann ist die Funktion $h = \alpha f + \beta g$ in c differenzierbar und es gilt die

$$\text{Linearität der Ableitung: } (\alpha f + \beta g)'(c) = \alpha f'(c) + \beta g'(c).$$

c) Seien wieder $f, g : J \rightarrow \mathbb{R}$ differenzierbar in c . Dann ist auch das Produkt $h = fg : x \mapsto f(x)g(x)$ in c differenzierbar und es gilt

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

d) Sei J_1 ein weiteres Intervall, $f : J \rightarrow J_1$ sei in $c \in J$ differenzierbar, $g : J_1 \rightarrow \mathbb{R}$ sei in $d = f(c)$ differenzierbar. Dann ist die Hintereinanderausführung $g \circ f : J \rightarrow \mathbb{R}, x \mapsto g(f(x))$ in c differenzierbar und es gilt die

$$\text{Kettenregel: } (g \circ f)'(c) = g'(f(c))f'(c).$$

Beweis: Wir beweisen nur die beiden interessanten Formeln c) und d). Das Prinzip ist nach dem Satz 4.2 c), eine in c stetige Funktion $S_h(x)$ zu finden mit $h(x) = h(c) + S_h(x)(x - c)$ und dann $S_h(c)$ auszurechnen.

c) Nach Voraussetzung gibt es stetige Funktionen S_f und S_g mit

$$\begin{aligned} f(x) &= f(c) + S_f(x)(x - c), \\ g(x) &= g(c) + S_g(x)(x - c). \end{aligned}$$

Daraus ergibt sich

$$h(x) = f(x)g(x) = f(c)g(c) + ((f(c)S_g(x) + g(c)S_f(x))(x - c).$$

Nach den Rechenregeln für stetige Funktionen ist

$$S_h(x) = f(c)S_g(x) + g(c)S_f(x)$$

stetig in c mit $S_h(c) = f(c)g'(c) + g(c)f'(c)$.

d) Nach Voraussetzung gibt es Funktionen $S_f : J \rightarrow \mathbb{R}$ und $S_g : J_1 \rightarrow \mathbb{R}$, die in c bzw. d stetig sind, so daß gilt:

$$\begin{aligned} f(x) &= f(c) + S_f(x)(x - c), \\ g(y) &= g(d) + S_g(y)(y - c) \quad (*) \end{aligned}$$

Dann ist $S_f(c) = f'(c)$ und $S_g(d) = g'(d)$. Setzt man in die Formel (*) für y speziell $f(x)$ ein, so erhält man wegen $d = f(c)$

$$\begin{aligned} g \circ f(x) &= g \circ f(c) + S_g(f(x))(f(x) - f(c)) \\ &= g \circ f(c) + S_g(f(x))S_f(x)(x - c). \end{aligned}$$

Da f in c differenzierbar ist, ist f nach Korollar 4.3 in c stetig. Damit ist auch die Hintereinanderausführung $S_g \circ f$ in c stetig. Nach den Rechenregeln für stetige Funktionen ist dann auch $x \mapsto S_g(f(x))S_f(x)$ in c stetig und es ist $S_g(f(c))S_f(c) = g'(f(c))f'(c)$. Mit Satz 4.2 c) folgt die Behauptung.

Korollar 4.5 (Quotientenregel)

Seien $f, g : J \rightarrow \mathbb{R}$ in c differenzierbar, ferner sei $g(x) \neq 0 \neq g'(c)$. Dann ist f/g in c differenzierbar und es gilt

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

Aufgaben:

1. Beweisen Sie dieses Korollar.

Tipp: Es ist $f/g = f \cdot 1/g$. Zeigen Sie also zuerst, daß $1/g$ in c differenzierbar ist (verwenden Sie dazu, daß $(1/x)' = -1/x^2$ (siehe die Beispiele oben) und die Kettenregel).

2. Sei $f(x) = \sum_{k=0}^n a_k x^k$. Zeigen Sie mit Begründung der einzelnen Schritte $f(x)' = \sum_{k=1}^n k a_k x^{k-1}$ (siehe auch die Beispiele oben).

Umkehrfunktion und Potenzreihen

Zur Vervollständigung des Formelarsenals für die Differentiation fehlt uns noch die Differenzierbarkeit der Umkehrfunktion und der Potenzreihen.

Satz 4.6 (Ableitung der Umkehrfunktion)

Sei f eine stetige Funktion, die das Intervall J bijektiv auf das Intervall J_1 abbildet. Sei f in c differenzierbar und $f'(c) \neq 0$. Dann ist die Umkehrfunktion $g = f^{-1} : J_1 \rightarrow J$ in $d = f(c)$ differenzierbar und es gilt

$$g'(d) = (f^{-1})'(d) = \frac{1}{f'(c)} = \frac{1}{f'(f^{-1}(d))}.$$

Beweis: g ist nach dem Satz über die Stetigkeit der Umkehrfunktion stetig. Nach Voraussetzung ist die Funktion

$$S_f(x) = \begin{cases} f'(c) & x = c \\ \frac{f(x) - f(c)}{x - c} & x \neq c \end{cases}$$

stetig auf J und stets ungleich 0, weil f injektiv und $f'(c) \neq 0$ ist. Es ist $f(x) = f(c) + S_f(x)(x - c)$, woraus

$$x - c = \frac{f(x) - f(c)}{S_f(x)}$$

folgt. Setzt man hier speziell $x = g(y)$, $c = g(d)$ ein, so erhält man

$$g(y) - g(d) = \frac{1}{S_f(g(y))}(y - d).$$

Die Funktion $S_g(y) := \frac{1}{S_f(g(y))}$ ist stetig in d und es ist

$$S_g(d) = \frac{1}{S_f(g(d))} = \frac{1}{S_f(c)} = \frac{1}{f'(c)} = \frac{1}{f'(g(d))}.$$

Die Behauptung folgt aus Satz 4.2 c).

Um den Satz über die Differenzierbarkeit von Potenzreihen beweisen zu können, wiederholen wir eine Folgerung aus der absoluten Konvergenz von Reihen.

Lemma 4.7 Sei $\sum_{n=0}^{\infty} b_n$ eine absolut konvergente Reihe. Dann konvergiert für jedes n auch die Reihe $\sum_{k=n}^{\infty} |b_k| =: R_n$ und $(R_n)_n$ ist eine Nullfolge.

Beweis: Die Folge $(\sum_{k=n}^{n+p} |b_k|)_p$ ist monoton wachsend und beschränkt durch $\sum_{k=0}^{\infty} |b_k|$, also konvergent nach Satz 2.11. Da $\sum_{k=0}^{\infty} |b_k|$ konvergiert, gibt es nach dem Cauchy Kriterium für Konvergenz zu beliebigem $\varepsilon > 0$ ein $n(\varepsilon/2)$ mit $0 \leq \sum_{k=n}^{n+p} |b_k| < \varepsilon/2$ für alle $n \geq n(\varepsilon/2)$ und alle p . Daraus folgt $0 \leq R_n \leq \varepsilon/2 < \varepsilon$ für all diese n .

Theorem 4.8 (Differentiation von Potenzreihen)

Sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann ist f auf ganz $] - R, R[$ differenzierbar und es ist

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1},$$

das heißt, man erhält die Ableitung durch "gliedweise" Differentiation.

Beweis:

(I) Sei $|c| < R$ beliebig. Wir setzen für $n \geq 1$

$$f_n(x) = \begin{cases} \frac{x^n - c^n}{x - c} & x \neq c \\ n c^{n-1} & x = c \end{cases}.$$

Dann ist $S_n(x) = \sum_{k=1}^n a_k f_k(x)$ stetig. Wir zeigen: $\lim_{n \rightarrow \infty} S_n(x) =: S(x)$ existiert, ist stetig und es gilt

$$f(x) = f(c) + S(x)(x - c).$$

Wegen $S(c) = \sum_{k=1}^{\infty} k a_k c^{k-1}$ folgt dann die Behauptung aus Satz 4.2 c).

(II) Sei $t > 0$ mit $|c| < t < R$ beliebig. Nach Aufgabe 26 haben die Reihen $\sum_{k=0}^{\infty} a_k x^k$ und $\sum_{k=1}^{\infty} k a_k x^k$ den gleichen Konvergenzradius. Also konvergiert die Reihe $\sum_{k=1}^{\infty} k a_k t^{k-1}$ absolut.

Es ist $f_n(x) = \sum_{l=0}^{n-1} x^{n-1-l} c^l$, also ist für $|x| \leq t$

$$|f_n(x)| \leq \sum_{l=0}^{n-1} |x|^{n-1-l} |c|^l \leq \sum_{l=0}^{n-1} t^{n-1-l} t^l = n t^{n-1}.$$

Damit erhält man für $|x| \leq t$

$$\begin{aligned}
|S_n(x)| &= \left| \sum_{k=1}^n a_k f_k(x) \right| \\
&\leq \sum_{k=1}^n |a_k| |f_k(x)| \\
&\leq \sum_{k=1}^n k |a_k| t^{k-1} \\
&\leq \sum_{k=1}^{\infty} k |a_k| t^{k-1}
\end{aligned}$$

Damit ist die Reihe $\sum_{k=1}^{\infty} a_k f_k(x)$ absolut konvergent. Ihr Grenzwert sei $S(x)$.

Behauptung: $(S_n(x))_n$ konvergiert gleichmäßig auf $[-t, t]$.

Beweis: Es ist für alle $x \in [-t, t]$ stets

$$\begin{aligned}
|S(x) - S_n(x)| &= \left| \lim_{p \rightarrow \infty} S_{n+p}(x) - S_n(x) \right| \\
&= \lim_{p \rightarrow \infty} \left| \sum_{k=n+1}^{n+p} a_k f_k(x) \right| \\
&\leq \lim_{p \rightarrow \infty} \sum_{k=n+1}^{n+p} k |a_k| t^{k-1} \\
&\leq \sum_{k=n}^{\infty} k |a_k| t^{k-1} \\
&=: R_n.
\end{aligned}$$

Die Folge (R_n) hängt nicht von x ab und ist nach dem vorangegangenen Lemma eine Nullfolge. Daraus folgt die Behauptung.

(III) Nach dem Theorem über die gleichmäßige Konvergenz stetiger Funktionen ist $S(x)$ stetig auf $[-t, t]$ und da $|c| < t < R$ beliebig war, ist $S(x)$ stetig auf ganz $] -R, R[$. Es ist

$$f(x) - f(c) = \lim_{n \rightarrow \infty} S_n(x)(x - c) = S(x)(x - c).$$

Nach Teil (I) folgt das Theorem.

Weitere wichtige Formeln

Wir stellen nun eine Reihe wichtiger Formeln zusammen, die Sie mit den angegebenen Regeln schnell ableiten können:

Aufgaben:

1. $\exp(x)' = \exp(x)$.
2. $\exp(ax)' = a \exp(ax)$.
3. Sei $a > 0$. Dann ist $(a^x)' = \ln(a)a^x$ (wegen $a^x = \exp(x \ln(a))$) nach Definition, siehe auch die untenstehenden Aufgaben über die Potenzen).
4. $\sin(x)' = \cos(x)$, $\cos(x)' = -\sin(x)$.
5. $(\ln(x))' = 1/x$ (s. Satz über die Umkehrfunktion und Aufg. 1).
6. $(\log_a(x))' = \frac{1}{x \ln(a)}$. (Beachte $\log_a(x) = \ln(x)/\ln(a)$.)
7. Man setzt $\tan(x) = \frac{\sin(x)}{\cos(x)}$ (lies: Tangens x) auf $] -\pi/2, \pi/2[$ (dort ist $\cos(x) > 0$, wie wir später sehen werden). Es ist $\tan(x)' = 1 + \tan(x)^2$.

Es folgen einige Aufgaben, die die Ableitungsroutine steigern.

Aufgaben:

1. Berechnen Sie bitte die Ableitungen der folgenden Funktionen:

(a) $f(x) = \sum_{k=0}^n a_k (x - c)^k$.

(b) $f(x) = \exp(\sin(x - c))$.

(c) $f(x) = \exp(\ln(x))$.

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

(e) $f(x) = \begin{cases} x^2 & x \geq 0 \\ -x^2 & x < 0 \end{cases}$

2. Beispiel einer überall differenzierbaren Funktion mit unstetiger Ableitung: Sei

$$h(x) = \begin{cases} \sqrt{|x|^3} \sin(1/|x|) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Zeigen Sie bitte: h ist überall differenzierbar, aber die Ableitung ist im Nullpunkt nicht stetig.

Tipp: Für $x > 0$ können Sie die Ableitung nach der Produkt- und Kettenregel bilden. Für $x < 0$ ist $h(x) = h(-x)$, also hilft wieder die Kettenregel. $h'(0) = 0$ muß man "zu Fuß" ausrechnen.

3. Die folgenden Aufgabenteile dienen zum Verständnis der Potenzfunktion und der Exponentialfunktion. Wir setzen die Formel $a^m a^n = a^{m+n}$ voraus.

(a) Seien $p, q \in \mathbb{N}$ und $a \in \mathbb{R}$. Zeigen Sie bitte: $(a^p)^q = (a^q)^p = a^{pq}$.

(b) Für alle weiteren Aufgaben benutzen Sie bitte die folgende Eigenschaft bijektiver Abbildungen (vergl. Abschnitt 1.1). Sei $f : M \mapsto N$ bijektiv. Für jedes $y \in N$ ist dann $f^{-1}(y)$ die *eindeutig bestimmte Lösung* der Gleichung $f(x) = y$.

Zeigen Sie bitte: Für $p, q \in \mathbb{N}$ und $x \geq 0$ gilt $\sqrt[p]{x^p} = (\sqrt[q]{x})^p$.

Tipp: Benutzen Sie die vorhergehende Aufgabe.

(c) Aufgrund von b) können wir für $p, q \in \mathbb{N}$ und $x > 0$ eindeutig erklären: $x^{p/q} = \sqrt[q]{x^p} = (\sqrt[q]{x})^p$. Ist $p \in \mathbb{Z}, p < 0$, so setzen wir $x^{p/q} = \frac{1}{x^{|p|/q}}$. Zeigen Sie bitte: Für $x > 0$ und $r, s \in \mathbb{Q}$ ist $x^r x^s = x^{r+s}$.

(d) Zeigen Sie bitte: sei $0 \neq p \in \mathbb{Z}, q \in \mathbb{N}$. Dann ist die Funktion $x \mapsto x^{p/q}$ stetig und bijektiv von $[0, \infty[$ auf sich. Die Umkehrfunktion ist $y \mapsto y^{q/p}$.

4. Für die folgenden Aufgaben wiederholen wir die Funktionalgleichung der Exponentialfunktion: $\exp(x + y) = \exp(x) \exp(y)$. Wir wollen den Zusammenhang zwischen der Exponentialfunktion und der vorigen Aufgabe herstellen. Sei $0 \neq p \in \mathbb{Z}$ und $q \in \mathbb{N}$.

(a) Sei $a = \exp(u)$ für ein festes $u \neq 0$. Zeigen Sie bitte:

$$a^p = \exp(pu) = \sum_{n=0}^{\infty} \frac{(pu)^n}{n!}.$$

(b) Zeigen Sie bitte $a^{1/q} = \exp(u/q)$.

(c) Zeigen Sie bitte $a^{p/q} = \exp(pu/q)$.

5. Zeigen Sie bitte: für $x > 0$ und $r \in \mathbb{Q}$ ist die Funktion $f(x) = x^r$ stetig differenzierbar und es gilt

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

Tipp: Am einfachsten ist es, die Formel $x^r = \exp(r \ln(x))$ zu benutzen.

Wegen der 4. Aufgabe kann man die Exponentialfunktion als *stetige Fortsetzung* der Funktion $r \in \mathbb{Q} \mapsto e^r$ ($e = \exp(1)$) von \mathbb{Q} auf ganz \mathbb{R} ansehen.

Höhere Ableitungen

Sei $f(x) = 3x^2$. Dann ist f nach den Rechenregeln in jedem Punkt differenzierbar. Damit erhält man eine neue Funktion f' , die jedem Punkt x die Steigung $f'(x)$ der Tangente im Punkt $(x, f(x))$ an die Krüve $G(F)$ zuordnet. Sie erhalten $f'(x) = 6x$. Sie sehen, daß Sie diese neue Funktion wieder differenzieren können; es ist $(f')'(x) = 6$ für alle x . (f') ist die "zweite Ableitung". Allgemein erklären wir:

Definition 4.9 Sei $J \subset \mathbb{R}$ ein Intervall und $f : J \rightarrow \mathbb{R}$ eine Funktion.

- Sei f in jedem Punkt $x \in J$ differenzierbar; dann heißt die neue Funktion $f' : J \rightarrow \mathbb{R}, x \mapsto f'(x)$ **erste Ableitung** von f .
- Sei für $n \geq 1$ die n . Ableitung $f^{(n)}$ als Funktion schon definiert. Ist die Funktion $f^{(n)}$ in jedem Punkt $x \in J$ differenzierbar, so heißt die Funktion

$$(f^{(n)})' : x \mapsto (f^{(n)})'(x)$$

$(n+1)$. **Ableitung** $f^{(n+1)}$ von f .

- Existiert die n . Ableitung für jedes n , so heißt f **unendlich oft** oder **beliebig oft differenzierbar**.

Statt $f^{(2)}$ schreibt man oft f'' oder auch \ddot{f} . Für $f^{(3)}$ ist auch f''' üblich. Statt $f^{(n)}$ schreibt man auch $\frac{d^n f}{dx^n}$.

Beispiele unendlich oft differenzierbarer Funktionen bilden die Potenzreihen:

Satz 4.10 Sei $f(x) = \sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann ist f unendlich oft differenzierbar und es gilt

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1) \cdots (k-n+1) a_k x^{k-n}.$$

Der Beweis dieses Satzes ergibt sich durch Induktion aus Theorem 4.8

Die dritte der folgenden Aufgaben spielt eine Rolle als Beispiel einer unendlich oft differenzierbaren Funktion, deren Taylorreihe zwar konvergiert, aber nicht gegen die Funktion (siehe den Abschnitt über die Taylorentwicklung). Sie ist auch Ausgangspunkt für die Konstruktion der "Diracschen Deltafunktion" (wichtig für Physiker).

Aufgaben:

- $\sin(x)'' = -\sin(x)$, $\cos(x)'' = -\cos(x)$.
- $\exp(-x^2/2)'' = (x^2 - 1) \exp(-x^2/2)$.
- Sei

$$h(x) = \begin{cases} \exp(-1/x^2) & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

Zeigen Sie bitte: h ist beliebig oft differenzierbar und es ist $h^{(n)}(0) = 0$ für alle n .

Tipp: Für $x \neq 0$ haben Sie dank der Rechenregeln (einschließlich der Kettenregel) keine Probleme.

Für $c = 0$ zeigen Sie zunächst durch Induktion $h^{(n)}(y) = (\sum_{k=0}^{n-1} a_k/y^k) \exp(-1/y^2)$ für $y \neq 0$, wobei Sie die a_k nicht explizit berechnen müssen. Damit ergibt sich für $x \neq 0$

$$\frac{h^{(n)}(x) - h^{(n)}(0)}{x - 0} = \left(\sum_{k=0}^{n-1} \frac{a_k}{x^{k+1}} \right) \exp(-1/x^2).$$

Nun zeigen Sie $\lim_{0 \neq x \rightarrow 0} \frac{\exp(-1/x^2)}{x^r} = 0$ für $r \in \mathbb{N}$, wobei Sie $\exp(-1/x^2) = \frac{1}{\exp(1/x^2)}$ benutzen und den Nenner in eine Reihe entwickeln, die Sie nach dem r . Glied abbrechen.

Wiederholung:

- Definition der Differenzierbarkeit und der Ableitung
- Äquivalente Formulierungen der Differenzierbarkeit
- Rechenregeln für das Differenzieren
- Ableitung der Umkehrfunktion
- Ableitung von Potenzreihen
- Höhere Ableitungen

4.2 Die Mittelwertsätze der Differentialrechnung

Überblick: Der Schlüsselsatz (Satz von Rolle) besagt, daß eine stetige Funktion $f : [a, b] \mapsto \mathbb{R}$ mit $f(a) = f(b) = 0$, die in $]a, b[$ differenzierbar ist, eine waagrechte Tangente besitzt. Aus diesem Satz leiten sich die beiden Mittelwertsätze durch eine einfache Rechnung ab. Der erste von ihnen besagt, daß es zu jeder Sekante an die Kurve $G(f)$ eine parallele Tangente gibt, falls f differenzierbar ist. Der zweite Mittelwertsatz erlaubt komplizierte Grenzwertberechnungen und ist der Schlüssel für den Taylor'schen Entwicklungssatz für genügend oft differenzierbare Funktionen.

Der Satz von Rolle

Anschaulich klar und aus der Schule bekannt ist der folgende Sachverhalt:

Lemma 4.11 Sei $J =]a, b[$ ein Intervall und $f : J \rightarrow \mathbb{R}$ sei eine differenzierbare Funktion. Hat f in c ein Maximum oder ein Minimum, so ist $f'(c) = 0$.

Die Voraussetzung bedeutet genauer: es gilt $f(c) \geq f(x)$ für alle x oder $f(c) \leq f(x)$ für alle x .

Beweis: Sei o. B. d. A. $f(c) \geq f(x)$ für alle x (sonst betrachte $g = -f$). Nach Voraussetzung existiert $f'(c) = \lim_{c \neq x \rightarrow c} \frac{f(x) - f(c)}{x - c}$. Für die Folge $(a_n)_n = (c + 1/n)_n$, also $a_n - c = 1/n$ gilt dann

$$f'(c) = \lim_{n \rightarrow \infty} n \underbrace{f(c + 1/n) - f(c)}_{\leq 0} \leq 0.$$

Für $(b_n)_n = (c - 1/n)_n$ ergibt sich

$$f'(c) = \lim_{n \rightarrow \infty} \underbrace{(-n)}_{<0} \underbrace{f(c - 1/n) - f(c)}_{\leq 0} \geq 0.$$

Damit folgt $f'(c) = 0$.

Geometrisch besagt der folgende Satz, daß eine differenzierbare Funktion, die in 0 startet und endet, irgendwo dazwischen eine waagrechte Tangente haben muß.

Satz 4.12 (Satz von Rolle)

Sei $J = [a, b]$ und $f : J \rightarrow \mathbb{R}$ sei stetig und im offenen Intervall $]a, b[$ differenzierbar. Sei ferner $f(a) = f(b) = 0$. Dann gibt es ein $c \in]a, b[$ (also $a < c < b$) mit $f'(c) = 0$.

Beweis: Ist f konstant gleich 0, so ist $f'(x) = 0$ für alle x , also kann man als c jedes $x \in]a, b[$ wählen. Sei f nicht konstant. Die Funktion $g = |f|$ ist stetig auf $[a, b]$ mit $g(a) = g(b) = 0$. Sie nimmt nach Theorem 3.14 ihr Maximum in einer Stelle $c \in [a, b,]$ an. Da sie nicht konstant ist, ist $g(c) > 0$, also $a < c < b$. Wir unterscheiden zwei Fälle:

(I) $f(c) = g(c)$, oder kurz: $f(c) > 0$. Dann ist für alle x stets

$$f(c) = |f(c)| = g(c) \geq |f(x)| \geq f(x),$$

also $f'(c) = 0$ nach dem vorangegangenen Lemma.

(II) $f(c) = -g(c)$, also $f(c) < 0$. Nach (I) hat $h = -f$ eine waagrechte Tangente, also auch f . Damit ist der Satz bewiesen.

Die Mittelwertsätze

Theorem 4.13 (Erster Mittelwertsatz)

Sei $F : [a, b] \rightarrow \mathbb{R}$ stetig und im offenen Intervall $]a, b[$ differenzierbar. Dann gibt es ein $c \in]a, b[$ mit

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

Meist wird dieser Satz in einer geometrisch noch anschaulicheren Version benutzt: Sei $f : J \rightarrow \mathbb{R}$ differenzierbar und $u, v \in J$ mit $u < v$ beliebig. Dann gibt es zur Sekante durch die Punkte $(u, f(u))$ und $(v, f(v))$ eine parallele Tangente an die Kurve $G(f)$ an einer zwischen u und v liegenden Stelle c . Noch anders ausgedrückt:

Korollar 4.14 Sei J ein beliebiges Intervall und $f : J \rightarrow \mathbb{R}$ sei differenzierbar auf J . Seien $u < v$ zwei beliebige Punkte aus J . Dann gibt es ein c mit $u < c < v$ und

$$\frac{f(v) - f(u)}{v - u} = f'(c).$$

Beweis: Beweis vom Theorem:

Wir ziehen die Sekante $s(x) = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$ von f ab und wenden den Satz von Rolle auf die Differenz $h(x) = f(x) - s(x)$ an. Es gibt danach ein $c \in]a, b[$ mit $0 = h'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}$. Das Korollar erhält man durch Anwendung des Theorems auf das Intervall $[u, v]$.

Anleitung zum FuncPlot-Applet Verfolgen Sie das geometrisch so anschauliche Vorgehen im Beweis anhand der beiden folgenden Beispiele:

1. $f(x) = x^2$, $a = -1$, $b = 0.5$. Zeichnen Sie f , die Sekante s und die Differenz h und lesen Sie den Punkt c aus der Zeichnung ab. Berechnen Sie c !
2. $f(x) = \sin(x)$, $a = -\pi/4$, $b = \pi/2$ ($\pi = 3.14159$). Es ist $f(a) = -1/\sqrt{2}$, $f(b) = 1$. Zeichnen Sie wieder f , s , und die Differenz h . Hier können Sie die Stelle c nur aus der Zeichnung ablesen.

Eine erste unmittelbare Anwendung des Korollars ist die folgende Charakterisierung der strengen Monotonie einer Funktion:

Satz 4.15 (Monotoniekriterium)

Sei $a < b$ und $g : [a, b] \rightarrow \mathbb{R}$ sei stetig und auf $]a, b[$ differenzierbar. Es habe g' keine Nullstelle. Dann ist g streng monoton auf $[a, b]$.

Beweis: Annahme der Satz ist falsch. Dann ist g nicht injektiv. Es gibt also $u < v$ in $[a, b]$ mit $g(u) = g(v)$. Für die Funktion $h(x) = g(x) - g(u)$ gibt es nach dem Satz 4.12 von Rolle eine Stelle c mit $0 = h'(c) = g'(c)$, ein Widerspruch.

Daraus erhalten wir den zweiten Mittelwertsatz:

Theorem 4.16 (zweiter Mittelwertsatz)

Seien $f, g : [a, b] \mapsto \mathbb{R}$ stetig und auf dem offenen Intervall $]a, b[$ differenzierbar. Es habe g' keine Nullstelle. Dann ist $g(a) \neq g(b)$ und es gibt eine Zwischenstelle c mit $a < c < b$ und

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

Beweis: Da g nach dem vorangegangenen Satz streng monoton ist, ist $g(a) \neq g(b)$. Wir wenden nun den Satz 4.12 von Rolle auf die Differenz

$$h(x) = f(x) - \left(f(a) + \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)) \right)$$

an. Wir erhalten eine Zwischenstelle c mit

$$0 = h'(c) = f'(c) - \left(\frac{f(b) - f(a)}{g(b) - g(a)} g'(c) \right).$$

Kurvendiskussion

Die ersten Anwendungen zur Kurvendiskussion können Sie selbst in den Aufgaben entwickeln.

Aufgaben:

1. Sei $f : [a, b] \mapsto \mathbb{R}$ stetig und in $]a, b[$ differenzierbar. Ist $f'(x) = 0$ für alle x , so ist f konstant.
Tipp: $\frac{f(x) - f(a)}{x - a} = ?$.
2. Sei f wie in Aufg. 1, nur sei jetzt $f'(x) > 0$ für alle x . Dann ist f streng monoton wachsend.
Tipp: Sei $u < v$. Was ist $\frac{f(v) - f(u)}{v - u}$? Löse nach $f(v)$ auf.
3. Sei f wie in 1. Zeigen Sie bitte: f ist genau dann monoton wachsend, wenn $f'(x) \geq 0$ ist für alle x .
4. Formulieren und beweisen Sie zu Aufg. 2 und 3 entsprechende Aussagen für "monoton fallend".

Grenzwertbestimmungen

Wir hatten schon früher gesehen, daß $\lim_{0 \neq x \rightarrow 0} \frac{\sin(x)}{x} = 1$ gilt (der Grenzwert ist die Ableitung von $\sin(x)$ an der Stelle 0). Der Grenzwert existiert also, obwohl Zähler und Nenner für sich betrachtet gegen 0 konvergieren. Mit Hilfe des zweiten Mittelwertsatzes können wir die allgemeine Situation behandeln.

Satz 4.17 (1. Regel von de L'Hopital: "0/0")

Sei J ein Intervall, $c \in J$ und $f, g : J \setminus \{c\} \rightarrow \mathbb{R}$ seien differenzierbar mit $g'(x) \neq 0$ für alle x . Es gelte $\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = 0$ und es existiere $\lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} =: L$. Dann existiert auch $\lim_{x \rightarrow c} \frac{f(x)}{g(x)}$ und ist gleich L . Kurz:

$$\lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow c} \frac{f(x)}{g(x)}.$$

Beweis:

- (I) Wir setzen f und g auf ganz J durch $f(c) = g(c) = 0$ fort. Die neuen, so fortgesetzten Funktionen sind auf ganz J nach Voraussetzung über die Existenz der Limites stetig und (nach wie vor) in $J \setminus \{c\}$ differenzierbar. Ist also $x \in J$ und $x < c$ bzw. $x > c$, so kann man den 2. Mittelwertsatz auf das Intervall $[x, c]$ bzw. $[c, x]$ anwenden.
- (II) Sei $\varepsilon > 0$ vorgegeben. Nach Voraussetzung gibt es ein $\delta > 0$ mit $|L - \frac{f'(y)}{g'(y)}| < \varepsilon$ für alle y mit $0 < |y - c| < \delta$. Sei $x \in J$ mit $0 < |x - c| < \delta$. Dann ist nach dem 2. Mittelwertsatz

$$\frac{f(x)}{g(x)} = \frac{f(x) - 0}{g(x) - 0} = \frac{f(x) - f(c)}{g(x) - g(c)} = \frac{f'(y)}{g'(y)}$$

für ein y echt zwischen x und c . Dies y erfüllt also insbesondere $0 < |y - c| < \delta$. Also ist

$$\left| L - \frac{f(x)}{g(x)} \right| = \left| L - \frac{f'(y)}{g'(y)} \right| < \varepsilon \text{ für } 0 < |x - c| < \delta.$$

Ganz ähnlich beweist man die folgende Regel:

Satz 4.18 (2. Regel von de L'Hopital: " ∞/∞ ")

Sei J ein Intervall, $c \in J$ und $f, g : J \setminus \{c\} \rightarrow \mathbb{R}$ seien differenzierbar mit $g'(x) \neq 0$ für alle x . Es möge $g(x)$ für $x \rightarrow c$ bestimmt gegen ∞ divergieren. Es existiere $\lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} =: L$. Dann existiert auch $\lim_{x \rightarrow c} \frac{f(x)}{g(x)}$ und ist gleich L . Kurz:

$$\lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow c} \frac{f(x)}{g(x)}.$$

Beispiele:

$$1. \lim_{x \rightarrow 0} \frac{\ln(1+ax)}{x} = \lim_{x \rightarrow 0} \frac{a}{(1+ax) \cdot 1} = a.$$

$$2. \lim_{x \rightarrow 2} \frac{\sqrt{x} - \sqrt{2}}{\sqrt{|x-2|}} = \pm \lim_{x \rightarrow 2} \frac{1 \cdot 2 \cdot \sqrt{|x-2|}}{2\sqrt{x}} = 0 \text{ ("+" für } x > 2, \text{ "-" für } x < 2).$$

$$3. \lim_{x \rightarrow 0} \frac{\sin(x)^2}{\cos(x) - 1} = \lim_{x \rightarrow 0} \frac{2 \sin x \cos x}{-\sin(x)} = -2.$$

$$4. \lim_{x \rightarrow 0} (x \ln(x)) = \lim_{x \rightarrow 0} \frac{\ln(x)}{\frac{1}{x}} = \lim_{x \rightarrow 0} \frac{-1 \cdot x^2}{x} = 0.$$

$$5. \lim_{x \rightarrow \infty} (1 + 1/x)^x = \lim_{y \rightarrow 0} (1 + y)^{1/y} = \lim_{y \rightarrow 0} \exp\left(\frac{\ln(1+y)}{y}\right) = \exp(1) = e \text{ nach Aufgabe 1 und wegen der Stetigkeit der Exponentialfunktion.}$$

Wiederholung:

- Satz von Rolle, 1. und 2. Mittelwertsatz.
- Monotonie und Ableitung.
- Grenzwertbestimmungen nach de L'Hopital.

4.3 Der Entwicklungssatz von Taylor

Überblick: Als Konsequenz des 2. Mittelwertsatzes ergibt sich: Eine $(n + 1)$ mal differenzierbare Funktion f kann um einen Punkt c durch das Polynom $P_n(x) = \sum_{k=0}^n \frac{f^{(k)}(c)}{k!} (x-c)^k$ sehr gut angenähert werden. Daraus leiten wir ab, wann man eine Funktion f durch eine Potenzreihe darstellen kann.

Der allgemeine Entwicklungssatz

Im ersten Abschnitt dieses Kapitels haben wir gesehen, daß und wie genau eine im Punkt c differenzierbare Funktion f durch die Tangente im Punkt $(c, f(c))$ an den Graphen $G(f)$ angenähert werden kann. Ist die Funktion nun n mal differenzierbar, so bietet sich die Annäherung durch ein "Schmiegpolyinom" n -ten Grades an. Die Tangente ist ja durch ein Polynom 1. Grades gegeben.

Anleitung zum FuncPlot-Applet Schauen Sie sich die Annäherung durch Polynome einmal an.

1. $f(x) = \sin(x)$ auf $[-2, 2]$. Der "Schmiegpunkt" ist $c = 0$. Wählen Sie $P_1(x) = x$, $P_2(x) = x - x^3/6$, $P_3(x) = x - x^3/6 + x^5/120 - x^7/5040$.
2. Dasselbe Beispiel, aber jetzt auf dem Intervall $[-6, 6]$.
3. $f(x) = \sqrt{x}$ auf $[0.5, 2]$. Der Schmiegpunkt ist $c = 1$. Wählen sie $P_1(x) = 1 + (x - 1)/2$, $P_2(x) = 1 + (x - 1)/2 - (x - 1)^2/8$, $P_3(x) = P_2(x) + \frac{3(x - 1)^3}{16} - \frac{5(x - 1)^4}{128} + \frac{7(x - 1)^5}{256}$.

Der Hauptsatz lautet:

Theorem 4.19 (Entwicklungssatz von Taylor)

Sei J ein Intervall und $f : J \rightarrow \mathbb{R}$ sei eine $(n + 1)$ mal differenzierbare Funktion auf J . Sei $c \in J$ beliebig. Dann gibt es zu jedem $x \in J$ eine Zwischenstelle y zwischen x und c mit

$$\begin{aligned} f(x) &= f(c) + \frac{f'(c)}{1!}(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \dots + \frac{f^{(n)}(c)}{n!}(x - c)^n + \frac{f^{(n+1)}(y)}{(n+1)!}(x - c)^{n+1} \\ &= \sum_{k=0}^n \frac{f^{(k)}(c)}{k!}(x - c)^k + \frac{f^{(n+1)}(y)}{(n+1)!}(x - c)^{n+1}. \end{aligned}$$

Das Polynom $\sum_{k=0}^n \frac{f^{(k)}(c)}{k!}(x - c)^k$ heißt **Taylorpolynom n . Grades um den Entwicklungspunkt c** , der Term $\frac{f^{(n+1)}(y)}{(n+1)!}(x - c)^{n+1}$ **Restglied**.

Beweis: Sei $T_n(x) = \sum_{k=0}^n \frac{f^{(k)}(c)}{k!}(x - c)^k$ das Taylorpolynom, sei $h(x) = f(x) - T_n(x)$ die Differenz zwischen f und dem Taylorpolynom und $g(x) = (x - c)^{n+1}$. Dann ist $h^{(k)}(c) = g^{(k)}(c) = 0$ für alle $k \leq n$. Wir wenden den 2. Mittelwertsatz nacheinander auf die Funktionen $f^{(k)}$ und $g^{(k)}$ an und erhalten

$$\begin{aligned} \frac{h(x)}{g(x)} &= \frac{h(x) - h(c)}{g(x) - g(c)} \\ &= \frac{h'(x_1)}{g'(x_1)} \\ &= \frac{h'(x_1) - h'(c)}{g'(x_1) - g'(c)} \\ &= \frac{h''(x_2)}{g''(x_2)} \\ &= \dots = \frac{h^{(n)}(x_n) - h^{(n)}(c)}{g^{(n)}(x_n) - g^{(n)}(c)} \\ &= \frac{h^{(n+1)}(x_{n+1})}{(n+1)!}, \end{aligned}$$

wobei x_1 zwischen x und c , x_2 zwischen x_1 und c , \dots , $x_{n+1} =: y$ zwischen x_n und c liegt. Multiplikation mit $g(x)$ liefert die Behauptung.

Korollar 4.20 Die Funktion $f : J \rightarrow \mathbb{R}$ ist genau dann ein Polynom n . Grades, wenn f $(n + 1)$ mal differenzierbar ist, $f^{(n)} \neq 0$ und $f^{(n+1)} = 0$ ist.

Lokale Extremwerte

Die Entwicklung einer Funktion bis zum 2. Glied, $f(x) = f(c) + f'(c)(x - c) + f''(y)/2 \cdot (x - c)^2$, spielt bei der Bestimmung lokaler Extremwerte eine Rolle.

Definition 4.21 Sei J ein Intervall und $f : J \rightarrow \mathbb{R}$ eine reelle Funktion. Sei $c \in J$, so daß noch ein ganzes Intervall $]c - \delta_0, c + \delta_0[$ in J enthalten ist ($\delta_0 > 0$). $f(c)$ heißt **lokales Maximum**, wenn es ein $\delta > 0$ gibt mit $f(c) \geq f(x)$ für alle $x \in J$ mit $|x - c| < \delta$. $f(c)$ heißt **lokales Minimum**, wenn es ein $\delta > 0$ gibt mit $f(c) \leq f(x)$ für alle $x \in J$ mit $|x - c| < \delta$. Liegt einer der beiden Fälle vor, so heißt $f(c)$ **lokaler Extremwert** und c eine **(lokale) Extremalstelle**.

Wie Sie sich aus der Schule vielleicht noch erinnern, gilt der folgende Satz:

Satz 4.22 (Lokale Extrema)

Sei J ein Intervall und $c \in J$, so daß noch ein ganzes Intervall $]c - \delta_0, c + \delta_0[$ in J enthalten ist. Sei $f : J \rightarrow \mathbb{R}$ differenzierbar.

- a) Ist c eine Extremalstelle, so ist $f'(c) = 0$.
- b) Sei f zweimal stetig differenzierbar. Ist $f'(c) = 0$ und $f''(c) < 0$, so ist $f(c)$ ein lokales Maximum. Ist $f'(c) = 0$ und $f''(c) > 0$, so ist $f(c)$ ein lokales Minimum.

Beweis:

- a) ist schon in Lemma 4.11 bewiesen worden.
- b) Sei $f''(c) > 0$. Zu $\varepsilon = f''(c)/2$ gibt es ein $\delta > 0$ mit

$$-f''(c)/2 < f''(x) - f''(c) < f''(c)/2,$$

also $0 < f''(c)/2 < f''(x)$ für alle x mit $|x - c| < \delta$. Für solche x gibt es ein y zwischen x und c (also insbesondere $|y - c| < \delta$) mit

$$f(x) = f(c) + f'(c)(x - c) + \underbrace{f''(y)(x - c)^2/2}_{>0} > f(c).$$

$f(c)$ ist also ein lokales Minimum. Ist $f''(c) < 0$ so betrachte $g = -f$.

Näherungsformeln

Aus dem Satz von Taylor erhält man die folgenden in der Praxis häufig verwendeten Formeln, indem man die Taylorreihe nach dem 1. oder 2. Glied abbricht (der Funktionswert an der Entwicklungsstelle c gilt als nulltes Glied). **Anleitung zum FuncPlot-Applet** Zeichnen Sie im Folgenden in der Nähe des Entwicklungspunktes $c = 0$ (also zum Beispiel auf dem Intervall $[-0.1, 0.1]$ oder sogar $[-0.5, 0.5]$) die gegebene Funktion und die Näherungsfunktion. Wählen Sie auf der y -Achse einen vernünftigen Maßstab, damit Sie die Abweichungen der Näherungsfunktion von der gegebenen überhaupt erkennen können!

1. $f(x) = \frac{1}{1+x} \approx 1 - x$; y -Intervall $[0.9, 1.1]$.
2. $f(x) = \frac{1}{1+x^2} \approx 1 - x^2$; y -Intervall $[0.9, 1.1]$.
3. $f(x) = \ln(1+x) \approx x - x^2/2$; y -Intervall $[-0.1, 0.1]$.
4. $f(x) = \sin(x) \approx x$; y -Intervall $[-0.3, 0.3]$.
5. $f(x) = \cos(x) \approx 1 - x^2/2$; y -Intervall $[0.8, 1.1]$.
6. $f(x) = \sqrt{1+x} \approx 1 + x/2$; y -Intervall $[0.8, 1.1]$.

Die Taylorreihe

Für die 2. Anwendung des Satzes von Taylor - die Entwickelbarkeit einer Funktion in eine Potenzreihe - benötigen wir die folgende Definition:

Definition 4.23 Sei $g : [a, b] \rightarrow \mathbb{R}$ eine beliebige beschränkte Funktion. Dann heißt

$$\|g\|_\infty := \sup\{|g(x)| : x \in [a, b]\}$$

die Supremumsnorm von g .

Bemerkung: Jede auf $[a, b]$ stetige Funktion g ist nach Theorem 3.14 beschränkt. Da auch $|g|$ stetig ist, gibt es sogar ein $x_{max} \in [a, b]$ mit $\|g\|_\infty = |g(x_{max})|$.

Satz 4.24 (Entwickelbarkeit in eine Taylorreihe)

Sei $f : [a, b] \mapsto \mathbb{R}$ beliebig oft differenzierbar. Es gelte $\lim_{n \rightarrow \infty} \frac{(b-a)^n}{n!} \|f^{(n)}\|_\infty = 0$. Sei $c \in [a, b]$ beliebig.

Dann konvergiert die Taylorreihe $\sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!} (x-c)^k$ gleichmäßig auf $[a, b]$ gegen f .

Beweis: Sei $T_n = \sum_{k=0}^n \frac{f^{(k)}(c)}{k!} (x-c)^k$ das n . Taylorpolynom. Dann ist für ein y zwischen x und c

$$\begin{aligned} |f(x) - T_n(x)| &= \left| \frac{f^{(n+1)}(y)(x-c)^{n+1}}{(n+1)!} \right| \\ &\leq \frac{|f^{(n+1)}(y)|}{(n+1)!} (b-a)^{n+1} \\ &\leq \frac{\|f^{(n+1)}\|_\infty}{(n+1)!} (b-a)^{n+1}. \end{aligned}$$

Da f beliebig oft differenzierbar, als $f^{(n+1)}$ differenzierbar ist, ist $f^{(n+1)}$ stetig, der rechts stehende Ausdruck also wohldefiniert. Er hängt nicht mehr von x ab und die Folge $(\frac{\|f^{(n+1)}\|_\infty}{(n+1)!} (b-a)^{n+1})_n$ konvergiert gegen 0. Daraus folgt, daß die Folge $(T_n)_n$ gleichmäßig gegen f konvergiert.

Korollar 4.25 Es gelte $\limsup_{n \rightarrow \infty} \sqrt[n]{\frac{\|f^{(n)}\|_\infty}{n!}} < \frac{1}{b-a}$. Dann konvergiert die Taylorreihe gleichmäßig auf $[a, b]$ gegen f .

Beweis: Die Voraussetzung impliziert

$$\limsup_{n \rightarrow \infty} (b-a) \sqrt[n]{\frac{\|f^{(n)}\|_\infty}{n!}} = q < 1,$$

also gibt es ein n_0 , so daß für alle $n > n_0$

$$(b-a) \sqrt[n]{\frac{\|f^{(n)}\|_\infty}{n!}} < (1+q)/2 < 1$$

ist. Damit ist aber für solche n

$$0 \leq \frac{(b-a)^n}{n!} \|f^{(n)}\|_\infty < ((1+q)/2)^n,$$

und die rechte Seite konvergiert für n gegen ∞ gegen 0.

Korollar 4.26 Sei $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius $R > 0$ und $f(x) = \sum_{k=0}^{\infty} a_k x^k$ auf $] -R, R[$. Dann ist die Potenzreihe die Taylorreihe von f , das heißt es gilt $f^{(n)}(0) = n! a_n$ für alle n .

Beweis: Rechnen Sie einfach $f^{(n)}(0)$ aus. Sie dürfen ja gliedweise differenzieren.

Beispiele:

1. Die Taylorreihen von $\exp(x)$, $\sin(x)$ und $\cos(x)$ sind also einfach die bekannten entsprechenden Potenzreihen, mit denen diese Funktionen definiert wurden.
2. $f(x) = \ln(1+x)$ auf $] -1, 1[$. Der Entwicklungspunkt sei 0. Es ergibt sich $f(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}$ und diese Taylorreihe konvergiert gegen f .
3. Entsprechend erhält man $\ln(1-x) = -\sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1}$. Subtrahiert man diese Reihe von der für $\ln(1+x)$, so erhält man

$$\ln\left(\frac{1+x}{1-x}\right) = \ln(1+x) - \ln(1-x) = 2 \sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1}.$$

4. Sei $\alpha \neq 0, -1$ und für $|x| < 1$ sei $f(x) = (1+x)^\alpha$. Wir setzen $\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-(k-1))}{k!}$ und erhalten für die Taylorreihe um $c=0$

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{k} x^k.$$

Auf $] -1, 1[$ konvergiert die Reihe gegen die Funktion.

5. Sei $h(x) = \begin{cases} \exp(-1/x^2) & x \neq 0 \\ 0 & x = 0 \end{cases}$. Diese Funktion ist beliebig oft differenzierbar, ihre Taylorreihe um den Entwicklungspunkt ist wegen $h^{(n)}(0) = 0$ gleich 0. Diese Reihe konvergiert also trivialerweise überall, aber nur im Nullpunkt gegen die Funktion.

Wiederholung:

- Taylorpolynom einer Funktion und Taylorsche Formel.
- Lokale Maxima und Minima einer Funktion.
- Näherungsformeln
- Konvergenz von Taylorreihen.
- Zusammenhang zwischen Potenzreihen und Taylorreihen.

4.4 Spezielle Funktionen

Überblick: Wir werden die Exponentialfunktion, die Winkelfunktionen und die Hyperbelfunktionen einschließlich ihrer Umkehrfunktionen ausführlich diskutieren. Dabei werden wir eine neue Methode zur Erzeugung von Funktionen kennenlernen: Lösungen von Differentialgleichungen.

Evolutionsgleichung und Exponentialfunktion

Die Gleichung

$$f'(t) = af(t) \quad \text{mit} \quad (0 \neq a)$$

beschreibt die sogenannte **lineare Evolution** von Systemen. Die Gleichung bedeutet: *die relative Änderung $\frac{f'(t)}{f(t)}$ pro Zeiteinheit und vorhandener Masse (Menge, Substanz, etc.) ist konstant.*

Gleichungen zwischen Funktionen, die für jedes Argument gelten sollen und neben der Funktion ihre Ableitungen bis zur Ordnung n enthalten, heißen **gewöhnliche Differentialgleichungen n-ter Ordnung**. Jede Funktion, die eine solche Gleichung erfüllt, heißt deren **Lösung**.

Satz 4.27 (Lineare Evolutionsgleichung)

Zu jedem Paar $(t_0, y_0) \in \mathbb{R}^2$ gibt es genau eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, die die Differentialgleichung

$$f'(t) = af(t)$$

löst und den Anfangswert

$$f(t_0) = y_0$$

hat, nämlich

$$f(t) = \exp(a(t - t_0))y_0.$$

Beweis: Die angegebene Funktion erfüllt beide Gleichungen, wie man durch Nachrechnen bestätigt. Sei g eine weitere Funktion, die beide Gleichungen erfüllt. Die Funktion $h(t) = \exp(-a(t - t_0))g(t)$ erfüllt dann $h'(t) = 0$ für alle t , also ist $h(t)$ konstant gleich y_0 . Damit ist $g(t) = \exp(a(t - t_0))h(t) = f(t)$.

Aufgabe: Man kann die Differentialgleichung auch mit dem Satz von Taylor lösen:

- Zeigen Sie bitte, daß jede Lösung der Differentialgleichung beliebig oft differenzierbar ist mit $f^{(n)}(t) = a^n f(t)$.
- Sei $L > 0$ beliebig. Zeigen Sie, daß $\limsup_{n \rightarrow \infty} \sqrt[n]{\frac{\|f^{(n)}\|_\infty}{n!}} = 0$ gilt. Damit konvergiert die Taylorreihe der Lösung f auf $[-L, L]$ gleichmäßig gegen f .
- Bestimmen Sie die Taylorreihe einer Lösung der beiden Gleichungen oben.

Die Differentialgleichung des mathematischen Pendels (Die Schwingungsgleichung)

Die zweite besonders wichtige Differentialgleichung ist die des mathematischen Pendels, in der die rücktreibende Kraft eines Pendels proportional zur Auslenkung ist.

$$f''(t) = -\omega^2 f(t), \quad \omega \neq 0.$$

Um die Eindeutigkeit der Lösung zu erhalten, müssen wir - wie es physikalisch sinnvoll ist - noch die Startauslenkung $f(t_0) = y_0$ und die Startgeschwindigkeit $f'(t_0) = z_0$ vorgeben.

Satz 4.28 (Schwingungsgleichung)

Zu jedem Tripel $(t_0, y_0, z_0) \in \mathbb{R}^3$ gibt es genau eine Lösung der Schwingungsgleichung

$$f''(t) = -\omega^2 f(t),$$

mit den Anfangswerten

$$\begin{aligned} f(t_0) &= y_0, \\ f'(t_0) &= z_0, \end{aligned}$$

nämlich

$$f(t) = y_0 \cos(\omega(t - t_0)) + z_0 \sin(\omega(t - t_0)).$$

Beweis: Die angegebene Funktion erfüllt die Differentialgleichung und die beiden “Anfangswerte”, wie Sie durch Nachrechnen bestätigen.

Eindeutigkeit: Sei g eine weitere Lösung mit den gleichen Anfangswerten und $h = f - g$. Dann ist $h'' + \omega^2 h = 0$. Multipliziert man diese Gleichung mit $2h'$ und benutzt die Kettenregel, so erhält man

$$((h')^2)' + \omega^2 (h^2)' = 0,$$

also ist

$$(h')^2(t) + \omega^2 h^2(t) = (h')^2(t_0) + \omega^2 h^2(t_0) = 0,$$

woraus $h^2 = 0$, also $f = g$ folgt.

Aufgabe: Zeigen Sie analog zur vorigen Aufgabe, daß jede Lösung der Differentialgleichung eine gegen sie konvergente Taylorreihe hat und bestimmen Sie diese Reihe aus den Anfangswerten und der Differentialgleichung. Sie erhalten die angegebene Lösung.

Die Winkelfunktionen

Wir haben Sinus und Cosinus durch Potenzreihen erklärt. In der Schule haben Sie Funktionen gleichen Namens geometrisch eingeführt. Wir machen absichtlich einen Querstrich über die “Schulfunktionen” und zeigen dann daß sie mit den durch die Potenzreihen gegebenen Funktionen übereinstimmen. Sinus bzw. Cosinus wurden als Gegenkathete $\overline{\sin}$ bzw. Ankathete $\overline{\cos}$ eines rechtwinkligen Dreiecks mit Winkel x im Bogenmaß und Hypotenuse 1 erklärt. (siehe Analysis Alive S. 134 ff). Dann zeigt sich, daß $\overline{\sin}(x)' = \overline{\cos}(x)$ und $\overline{\cos}(x)' = -\overline{\sin}(x)$ ist. Damit erhält man $\overline{\sin}(x)'' = -\overline{\sin}(x)$ und $\overline{\sin}(0) = 0$, $\overline{\sin}(0)' = 1$ ist. Der Satz 4.28 zeigt, daß $\overline{\sin} = \sin$ gilt. Ebenso erhält man, daß $\overline{\cos} = \cos$ ist. Damit können wir einige Eigenschaften der Winkelfunktionen direkt aus ihrer geometrischen Bedeutung ablesen. Zum Beispiel ist $\pi/4$ die erste Nullstelle > 0 des Cosinus und ferner gilt

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

Die Winkelfunktionen Cosinus und Sinus sind also **periodisch** mit kleinster Periode 2π .

Die **Tangensfunktion** $\tan(x) = \frac{\overline{\sin}(x)}{\overline{\cos}(x)}$ ist auf $]-\pi/2, \pi/2[$ erklärt und wegen $\tan(x)' = 1 + \tan(x)^2$ streng monoton wachsend mit Bild \mathbb{R} . Die Umkehrfunktion **Arcustangens** $\arctan(x)$ bildet \mathbb{R} auf $]-\pi/2, \pi/2[$ ab und erfüllt $\arctan x' = \frac{1}{1+x^2}$.

Es ist $\cos(x) > 0$ auf $] -\pi/2, \pi/2[$, also ist dort $\sin(x)$ streng monoton wachsend. Er bildet das abgeschlossene Intervall $[-\pi/2, \pi/2]$ also bijektiv auf $[-1, 1]$ ab. Die Umkehrfunktion **Arcus Sinus** $\arcsin(x)$ bildet daher $[-1, 1]$ stetig auf $[-\pi/2, \pi/2]$ ab und ist auf $] -1, 1[$ stetig differenzierbar mit

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

Ähnlich ist der Cosinus streng monoton fallend von $[0, \pi]$ auf $[-1, 1]$. Die Umkehrfunktion **Arcus Cosinus** $\arccos(x)$ bildet $[-1, 1]$ streng monoton fallend auf $[0, \pi]$ ab und ist auf dem offenen Intervall $] -1, 1[$ stetig differenzierbar mit

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

Die Hyperbelfunktionen

Die die Hyperbelfunktionen definierende Differentialgleichung lautet

$$f''(x) = f(x).$$

Um diese der Gleichung des mathematischen Pendels ähnliche Gleichung zu lösen, führen wir die folgenden Funktionen ein:

$$\begin{aligned} \text{Cosinus Hyperbolicus } \cosh(t) &= \frac{1}{2}(e^t + e^{-t}), \\ \text{Sinus Hyperbolicus } \sinh(t) &= \frac{1}{2}(e^t - e^{-t}). \end{aligned}$$

Diese Funktionen spielen für die Hyperbel $\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 = 1\}$ eine zu den Winkelfunktionen für den Einheitskreis $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ analoge Rolle.

Satz 4.29 (Differentialgleichung der Hyperbelfunktionen)

Zu jedem Tripel $(t_0, y_0, z_0) \in \mathbb{R}^3$ gibt es genau eine Lösung $f : \mathbb{R} \rightarrow \mathbb{R}$ der Differentialgleichung

$$f''(t) = f(t)$$

mit den Anfangswerten

$$\begin{aligned} f(t_0) &= y_0 \\ f'(t_0) &= z_0, \end{aligned}$$

nämlich

$$f(t) = y_0 \cosh(t - t_0) + z_0 \sinh(t - t_0).$$

Beweis: Daß die angegebene Funktion Lösung der Differentialgleichung mit den gegebenen Anfangswerten ist, erhält man durch einfaches Nachrechnen.

Eindeutigkeit: Wir können wieder zeigen, daß jede Lösung des Problems in eine Taylorreihe entwickelbar ist und diese gerade die angegebene Funktion darstellt. Ein anderer, einfacherer Weg ist der folgende: Sei g eine Lösung mit den vorgegebenen Anfangswerten. Dann erfüllt $h = f - g$ die Differentialgleichung

mit den Anfangswerten $(t_0, 0, 0)$. Sei $h_+ = h + h'$. Dann gilt $h'_+ = 2h_+$, $h_+(t_0) = 0$. Nach dem Eindeutigkeitssatz für lineare Evolutionsgleichungen (Satz 4.27) folgt $h_+ = 0$. Analog zeigt man $h_- = h - h' = 0$. Daraus folgt $2h = h_+ + h_- = 0$, also $f = g$.

Wir haben bei allen drei Differentialgleichungen einen dem jeweiligen Typ angepaßten Trick benutzt, um die Eindeutigkeit der Lösungen zu zeigen. *Es gibt einen ganz allgemeinen Satz über die Lösungen von Differentialgleichungen, aus dem die Eindeutigkeit sofort in jedem unserer Fälle folgen würde. Dies übersteigt jedoch den Stoff dieser Vorlesung.*

Wir diskutieren die Hyperbelfunktionen ausführlicher. Ihre Umkehrfunktionen spielen in der Integralrechnung eine wichtige Rolle.

Wegen $\cosh(x) > 0$ und $\sinh(x)' = \cosh(x)$ ist \sinh streng monoton wachsend und bildet \mathbb{R} auf sich ab. Die Umkehrfunktion **Area Sinus Hyperbolicus** $\operatorname{arsinh}(x)$ hat als Ableitung

$$\operatorname{arsinh}(x)' = \frac{1}{\sqrt{1+x^2}}.$$

Entsprechend erhält man, daß \cosh das Intervall $[0, \infty[$ streng monoton auf $[1, \infty[$ abbildet. Die Umkehrfunktion **Area Cosinus Hyperbolicus** $\operatorname{arcosh}(x)$ erfüllt

$$\operatorname{arcosh}(x)' = \frac{1}{\sqrt{x^2-1}} \text{ für } x > 1.$$

Schließlich führt man noch den **Tangens Hyperbolicus** $\tanh(x)$ durch

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)}$$

ein. Es ist

$$\tanh(x)' = 1 - \tanh(x)^2.$$

Wegen $|\sinh(x)| \leq \cosh(x)$ bildet \tanh ganz \mathbb{R} streng monoton auf $] -1, 1[$ ab. Die Umkehrfunktion **Area Tangens Hyperbolicus** $\operatorname{artanh}(x)$ erfüllt

$$\operatorname{artanh}(x)' = \frac{1}{1-x^2} \quad (|x| < 1).$$

Applet "Symbolische Integration und Differentiation" Ihnen steht ein Applet zur Verfügung mit dem Sie mittels eines Computer-Algebra-Programms die Ableitungen beliebiger (differenzierbarer) Funktionen bestimmen können.

Wiederholung:

- Exponentialfunktion und die lineare Evolutionsgleichung.
- Winkelfunktionen und die Schwingungsgleichung.
- Hyperbelfunktionen

Kapitel 5

Integralrechnung

Überblick: Im ersten Abschnitt führen wir das Integral von Treppenfunktionen und die Rechenregeln für das Integral ein. Im zweiten Abschnitt setzen wir das Integral mit dem Ausschöpfungsprinzip auf die "integrierbaren" Funktionen fort. Stetige Funktionen sind integrierbar. Der dritte Abschnitt stellt über das unbestimmte Integral den Zusammenhang zur Differentialrechnung her. Im vierten Abschnitt behandeln wir Fourierreihen und uneigentliche Integrale.

5.1 Das Integral von Treppenfunktionen

Überblick: Wir führen erst Treppenfunktionen zu einer festen Zerlegung eines Intervalls ein. Für sie ist das Integral der Flächeninhalt der zwischen der x -Achse und der Funktion liegenden Fläche. Wir erweitern das Integral auf Treppenfunktionen zu beliebigen Zerlegungen eines Intervalls und stellen die Rechenregeln für das Integral zusammen.

Sei $J = [0, 1]$. Wie groß ist der Inhalt der Fläche F die zwischen der x -Achse und der Kurve $\{(x, x^2) : 0 \leq x \leq 1\}$, also dem Graphen der Funktion $f(x) = x^2$, liegt? Wir können ihn ganz naiv erhalten, indem wir das Intervall sehr fein unterteilen durch die Einteilungspunkte $Z = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\}$, wo n sehr groß ist, und dann als Annäherung von unten die Summe der Flächeninhalte der Rechtecke mit den Grundlinien $[\frac{k}{n}, \frac{k+1}{n}]$ und der entsprechenden Höhe $\frac{k^2}{n^2}$, wählen.

Also kurz: Flächeninhalt:

$$F \approx \frac{1}{n} \cdot \left(\frac{1}{n^2} + \frac{4}{n^2} + \frac{[n-1]^2}{n^2} \right) = \frac{1}{n^3} \sum_{k=1}^{n-1} k^2 = \frac{1}{n^3} \cdot \frac{(n-1) \cdot n(2n-1)}{6} \approx \frac{1}{3}$$

Anleitung zum Applet "partiell definierte Funktionen" Zeichnen Sie die Funktion $f(x) = x^2$ auf dem Intervall $J = [0, 1]$.

Wählen sie nun $n = 4$ und geben die Funktion

$$f_4 = 1_{[1/4, 1/2]} \cdot 1/16 + 1_{[1/2, 3/4]} \cdot 1/4 + 1_{[3/4, 1]} \cdot \frac{9}{16} = \sum_{k=1}^3 \frac{k^2}{16} \cdot 1_{[k/4, (k+1)/4]}$$

ein.

Danach geben Sie bitte

$f_8 = \sum_{k=1}^7 \frac{k^2}{64} 1_{] \frac{k}{8}, \frac{k+1}{8}]}$ ein. Sie sehen wie sich die Flächen immer besser von unten an die Fläche f annähern.

Ebenso könnte man die Fläche von oben einschließen, indem man für die Rechtecke dieselbe Grundlinie, aber die Höhe $\frac{(k+1)^2}{n^2}$ wählt. Man erhält

$$F \approx \frac{1}{n} \left(\frac{1}{n^2} + \frac{4}{n^2} + \dots + \frac{n^2}{n^3} \right) = \frac{1}{n^3} \frac{n(n+1)(2n+1)}{6} \approx \frac{1}{3}.$$

Anleitung zum Applet "partiell definierte Funktionen" Zeichnen Sie wieder erst $f(x) = x^2$ auf $[0, 1]$.

Zeichnen Sie nun

$$F_4 = \frac{1}{16} 1_{[0, \frac{1}{4}]} + \sum_{k=1}^{3-1} \frac{(k+1)^2}{16} 1_{] \frac{k}{4}, \frac{k+1}{4}]} \quad \text{und} \quad F_8 = \frac{1}{64} 1_{[0, \frac{1}{8}]} + \sum_{k=1}^7 \frac{(k+1)^2}{64} 1_{] \frac{k}{8}, \frac{k+1}{8}]}$$

Sie erhalten Flächen, die sich von oben immer besser an F schmiegen.

Um dieses Vorgehen zu rechtfertigen, beginnen wir, den Inhalt von Flächen zu berechnen, die sich aus Rechtecken mit Grundlinien $[a_0, a_1], [a_1, a_2], \dots, [a_{n-1}, a_n]$ und Höhen h_0, h_1, \dots, h_{n-1} zusammensetzen. Eine solche Fläche hat den Inhalt $h_0 \cdot (a_1 - a_0) + h_1(a_2 - a_1) + \dots + h_{n-1}(a_n - a_{n-1})$. Um die Rechenregeln hier zu beherrschen, beginnen wir ganz einfach. Um die Übersicht zu behalten, wählen wir zunächst eine "unregelmäßige" Zerlegung $Z = \{0, \frac{1}{4}, \frac{3}{4}, 1\}$, also die Intervalle $[0, \frac{1}{4}],] \frac{1}{4}, \frac{3}{4}],] \frac{3}{4}, 1]$ und betrachten also Funktionen der Bauart

$$f = h_0 \cdot 1_{[0, \frac{1}{4}]} + h_1 1_{] \frac{1}{4}, \frac{3}{4}]} + h_2 1_{] \frac{3}{4}, 1]}.$$

Anleitung zum Applet "partiell definierte Funktionen" Zeichnen Sie mindestens 3 solche Funktionen

Sei $\mathcal{T}(Z)$ die Menge aller dieser Funktionen.

Aufgaben: 1. Zeigen Sie bitte:

- $f \in \mathcal{T}(Z) \Rightarrow |f| \in \mathcal{T}(Z)$.
- $f, g \in \mathcal{T}(Z)$ und $\alpha, \beta \in \mathbb{R} \Rightarrow \alpha f + \beta g \in \mathcal{T}(Z)$.
- $f, g \in \mathcal{T}(Z) \Rightarrow fg \in \mathcal{T}(Z)$.
- $f, g \in \mathcal{T}(Z) \Rightarrow \max(f, g)$ und $\min(f, g) \in \mathcal{T}(Z)$. Insbesondere $f^+ \in \mathcal{T}(Z)$, $f^- \in \mathcal{T}(Z)$.

2. Sei jetzt $Z = \{a_0, \dots, a_n\} \subset [0, 1]$ eine beliebige Zerlegung mit $0 = a_0 < a_1 < \dots < a_n = 1$ und $\mathcal{T}(Z) = \{f : [0, 1] \rightarrow \mathbb{R} : f = h_0 1_{[a_0, a_1]} + \sum_{k=1}^{n-1} h_k 1_{] a_k, a_{k+1}]}\}$. Zeigen Sie a) - d) aus Aufgabe 1.

Wir abstrahieren:

Definition 5.1 Eine endliche Teilmenge $Z = \{a_0, \dots, a_n\}$ des Intervalls $[a, b]$, die die Endpunkte a und b enthält, heißt **Zerlegung von J** .

Da Z endlich ist, können wir diese Menge (evtl. nach Umbenennung) immer so gegeben denken:

$$a = a_0 < a_1 < a_2 < \dots < a_n = b. \quad (5.1)$$

Satz 5.2 Sei $Z = \{a_0, \dots, a_n\}$ eine Zerlegung des Intervalls $J = [a, b]$ wie in Formel (5.1).

Sei $\mathcal{T}(Z) = \{f : J \rightarrow \mathbb{R} : f = h_0 1_{[a_0, a_1]} + \sum_{k=1}^{n-1} h_k 1_{[a_k, a_{k+1}]}\}$.

Dann gilt:

- a) $f \in \mathcal{T}(Z) \Rightarrow |f| \in \mathcal{T}(Z)$
- b) $f, g \in \mathcal{T}(Z)$ und $\alpha, \beta \in \mathbb{R} \Rightarrow \alpha f + \beta g \in \mathcal{T}(Z)$
- c) $f, g \in \mathcal{T}(Z) \Rightarrow fg \in \mathcal{T}(Z)$
- d) $f, g \in \mathcal{T}(Z) \Rightarrow \max(f, g)$ und $\min(f, g) \in \mathcal{T}(Z)$, insbesondere $f^+ \in \mathcal{T}(Z)$, $f^- \in \mathcal{T}(Z)$.

Der Beweis ist analog zu Aufgabe 2 oben. Wir können nun sofort die Fläche berechnen, die zwischen $f \in \mathcal{T}(Z)$ und der x -Achse liegt. Dabei rechnen wir Flächeninhalte unter der x -Achse negativ.

Satz 5.3 Für $f \in \mathcal{T}(Z)$, $f = h_0 1_{[a_0, a_1]} + \sum_{k=1}^{n-1} h_k 1_{[a_k, a_{k+1}]}$ sei $S_Z(f) = \sum_{k=0}^{n-1} h_k (a_{k+1} - a_k)$. Dann gilt:

- a) $S_Z(\alpha f + \beta g) = \alpha S_Z(f) + \beta S_Z(g)$
- b) $0 \leq f \Leftrightarrow 0 \leq S_Z(f)$; $f \leq g \Rightarrow S_Z(f) \leq S_Z(g)$
- c) $|S_Z(f)| \leq S_Z(|f|)$

Den Beweis stellen wir als Aufgabe.

Aufgabe: Beweisen Sie bitte den oben stehenden Satz.

Definition 5.4 Eine Zerlegung Z' des Intervalls $J = [a, b]$ heißt **feiner** als Z , wenn $Z \subset Z'$ gilt.

Zu je zwei Zerlegungen Z, Z' gibt es also eine "gemeinsame" Verfeinerung $Z'' = Z \cup Z'$.

Um den nächsten Schritt verständlich zu machen, wählen wir $J = [0, 1]$, $Z = \{0, \frac{1}{4}, \frac{3}{4}, 1\}$ und $Z' = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$.

Behauptung: $\mathcal{T}(Z) \subset \mathcal{T}(Z')$ (aber $\mathcal{T}(Z') \not\subset \mathcal{T}(Z)$) und für $f \in \mathcal{T}(Z)$ ist $S_Z(f) = S_{Z'}(f)$.

Beweis: Sei $f \in \mathcal{T}(Z)$, also $f = h_0 1_{[0, \frac{1}{4}]} + h_1 1_{[\frac{1}{4}, \frac{3}{4}]} + h_2 1_{[\frac{3}{4}, 1]}$. Dann ist

$$f = h_0 1_{[0, \frac{1}{4}]} + h_1 1_{[\frac{1}{4}, \frac{1}{2}]} + h_1 1_{[\frac{1}{2}, \frac{3}{4}]} + h_2 1_{[\frac{3}{4}, 1]}$$

weil

$$1_{[\frac{1}{4}, \frac{3}{4}]} = 1_{[\frac{1}{4}, \frac{1}{2}]} + 1_{[\frac{1}{2}, \frac{3}{4}]}.$$

Weiter ist wegen $S_Z(1_{[\frac{1}{4}, \frac{3}{4}]}) = \frac{3}{4} - \frac{1}{4} = \frac{3}{4} - \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = S_{Z'}(1_{[\frac{1}{4}, \frac{3}{4}]})$ stets $S_Z(f) = S_{Z'}(f)$. Das gilt natürlich ganz allgemein.

Satz 5.5 Sei $J = [a, b]$ ein Intervall und Z, Z' zwei Zerlegungen mit $Z \subset Z'$, Z' also feiner als Z .

- a) $\mathcal{T}(Z) \subset \mathcal{T}(Z')$.

b) Sei $f \in \mathcal{T}(Z)$. Dann ist $S_{Z'}(f) = S_Z(f)$.

Beweis: Den Beweis führt man durch Induktion nach der Anzahl der zu Z hinzukommenden Punkte aus Z' . Hat Z' nur einen Punkt mehr, so argumentiert man wie oben im Beispiel. Der Rest ist klar (s. Analysis Alive, S. 146-148)

Definition 5.6 Sei $J = [a, b]$ und \mathcal{Z} die Menge aller Zerlegungen von J .

- a) $f : J \mapsto \mathbb{R}$ heißt **Treppenfunktion**, wenn es ein $Z \in \mathcal{Z}$ mit $f \in \mathcal{T}(Z)$ gibt.
- b) Die Menge $\cup_{Z \in \mathcal{Z}} \mathcal{T}(Z)$ aller Treppenfunktionen wird mit $\mathcal{T}([a, b])$ bzw. $\mathcal{T}(J)$ bezeichnet.

Satz 5.7 Sei $J = [a, b]$. Dann gilt

- a) $f \in \mathcal{T}(J) \Rightarrow |f| \in \mathcal{T}(J)$
- b) $f, g \in \mathcal{T}(J)$ und $\alpha, \beta \in \mathbb{R} \Rightarrow \alpha f + \beta g \in \mathcal{T}(J)$
- c) $f, g \in \mathcal{T}(J) \Rightarrow \max(f, g)$ und $\min(f, g) \in \mathcal{T}(J)$.

Insbesondere ist f^+ und $f^- \in \mathcal{T}(J)$.

Beweis: Sind $f, g \in \mathcal{T}(J)$, so gibt es Zerlegungen Z_f, Z_g mit $f \in \mathcal{T}(Z_f), g \in \mathcal{T}(Z_g)$, also nach Satz 5.5 $f, g \in \mathcal{T}(Z_f \cup Z_g)$. Damit folgt alles aus Satz 5.3.

Satz 5.8 Für $f \in \mathcal{T}(J)$, also $f \in \mathcal{T}(Z_f)$ sei $\int_a^b f(x) dx := S_{Z_f}(f)$ das Integral der Treppenfunktion f über $J = [a, b]$.

- a) Das Integral hängt nicht von der speziellen Wahl $Z_f \in \mathcal{Z}$ ab, ist also eindeutig definiert.
- b) Es gilt:

(i) Seien $\alpha, \beta \in \mathbb{R}, f, g \in \mathcal{T}(J)$. Dann ist

$$\int_a^b (\alpha f + \beta g)(x) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x)$$

(Linearität des Integrals)

(ii) $0 \leq f \Rightarrow 0 \leq \int_a^b f(x) dx; f \leq g \Rightarrow \int_a^b f(x) dx \leq \int_a^b g(x) dx$

(Positivität des Integrals)

(iii) $|\int_a^b f(x) dx| \leq \int_a^b |f(x)| dx$

Beweis:

- a) Sind Z, Z' zwei Zerlegungen mit $f \in \mathcal{T}(Z), f \in \mathcal{T}(Z')$, so folgt $f \in \mathcal{T}(Z \cup Z')$ also nach Satz 5.5 wegen $Z \subset Z \cup Z'$ und $Z' \subset Z \cup Z'$

$$S_Z(f) = S_{Z \cup Z'}(f) = S_{Z'}(f).$$

Der Rest folgt nun leicht so:

- b) (i) $f \in \mathcal{T}(Z_f), g \in \mathcal{T}(Z_g) \Rightarrow f, g \in \mathcal{T}(Z_f \cup Z_g)$ also folgt (i) aus Satz 5.3.
 (i),(ii) analog.

Man kann das Integral einer Treppenfunktion einfach abschätzen:

Satz 5.9 Sei $f \in \mathcal{T}([a, b])$ und

$$m = \min\{f(x) : x \in [a, b]\}, \quad M = \max\{f(x) : x \in [a, b]\}.$$

a) Es gilt $m(b-a) \leq \int_a^b f(x) dx \leq M(b-a)$.

b) Sei $\|f\|_\infty = \sup\{|f(x)| : x \in [a, b]\}$.
 Dann ist $|\int_a^b f(x) dx| \leq \|f\|_\infty (b-a)$.

Beweis: Es ist $m1_{[a,b]} \leq f \leq M1_{[a,b]}$ und $1_{[a,b]} \in \mathcal{T}([a, b])$ ($Z = \{a, b\}$). Also folgt die Behauptung a) aus Satz 5.8.

b) Es ist $|f| \leq \|f\|_\infty 1_{[a,b]}$. Verwende obiges Argument.

Schließlich benötigen wir noch die **Intervalladditivität**.

Satz 5.10 Sei $a < c < b$ und $f \in \mathcal{T}([a, b])$. Dann ist $f|_{[a,c]} \in \mathcal{T}([a, c]), f|_{[c,b]} \in \mathcal{T}([c, b])$ und es gilt $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$.

Beweis: Sei $f \in \mathcal{T}(Z)$, wo Z eine Zerlegung von $[a, b]$ ist. Sei $Z' = Z \cup \{c\}$. Jetzt folgt die Behauptung sofort, wie Sie sich an einem einfachen Beispiel klar machen können.

Wiederholung:

- Zerlegung eines Intervalls
- Treppenfunktion
- Integral einer Treppenfunktion
- Rechenregeln für das Integral (einschließlich Abschätzungen und Intervalladditivität)

5.2 Das allgemeine Integral

Überblick: Das allgemeine Integral wird anschaulich durch Archimedes' Ausschöpfungsmethode definiert. Stetige Funktionen sind integrierbar. Wir stellen die Rechenregeln zusammen und zeigen, wann man Integral und Konvergenz vertauschen kann.

Wir nennen eine endliche Vereinigung $E = \cup_{k=1}^n R_k$ von Rechtecken $R_k =]a_k, b_k[\times [0, h_k]$, wo $]a_k, b_k[\subset [a, b]$ liegt, eine Elementarfläche.

Sie liegt zwischen der x -Achse und dem Graphen der Treppenfunktion $g = \sum_{k=1}^n h_k 1_{]a_k, b_k]}$. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beliebige Funktion mit $f(x) \geq 0$ für alle x . Dann kann man die Fläche $F = \{(x, y) : 0 \leq y \leq f(x)\}$ zwischen dem Graphen von f und der x -Achse von unten durch Elementarflächen und von oben durch Elementarflächen ausschöpfen, indem man $[a, b]$ immer feiner zerlegt und die Rechteckshöhen möglichst dicht unterhalb oder oberhalb von $f(x)$ wählt. Das legt die folgende, auf Archimedes zurückgehende Definition des Flächeninhalts durch Ausschöpfung von außen und innen nahe:

Definition 5.11 Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Dann heißt $\sup\{\int_a^b g(x)dx : g \in \mathcal{T}([a, b]), g \leq f\} =: \int_a^b f(x)dx$ **unteres Integral** von f über $[a, b]$ und $\inf\{\int_a^b h(x)dx : h \in \mathcal{T}([a, b]), h \geq f\} =: \overline{\int}_a^b f(x)dx$ **oberes Integral** von f über $[a, b]$.

Ist $\int_a^b f(x)dx = \overline{\int}_a^b f(x)dx$, so heißt f **integrierbar** über $[a, b]$ und der gemeinsame Wert heißt **Integral** $\int_a^b f(x)dx$.

Bemerkung: Da f als beschränkt vorausgesetzt wird, gibt es ein $M \geq 0$ mit $|f(x)| \leq M$ für alle $x \in [a, b]$, oder $-M1_{[a,b]} \leq f \leq M1_{[a,b]}$. Damit ist

$$-M(b-a) \leq \int_a^b f(x)dx \leq \overline{\int}_a^b f(x)dx \leq M(b-a).$$

Wir schauen uns das einführende motivierende Beispiel aus dem ersten Abschnitt dieses Kapitels an: $[a, b] = [0, 1]$, $f(x) = x^2$.

Behauptung: f ist integrierbar und $\int_0^1 f(x)dx = \frac{1}{3}$.

Beweis: Wir benutzen die Formel $\sum_{k=1}^r k^2 = \frac{r(r+1)(2r+1)}{6}$, die Sie früher einmal durch Induktion bewiesen haben. Sei $n \in \mathbb{N}$ beliebig und $g_n = \sum_{k=1}^{n-1} \frac{k^2}{n^2} 1_{\left[\frac{k}{n}, \frac{k+1}{n}\right]}$. Dann ist $g_n \in \mathcal{T}([0, 1])$, $g_n(x) \leq x^2$ und daher

$$\int_0^1 x^2 dx \geq \int_0^1 g_n(x)dx = \frac{1}{n^3} \sum_{k=1}^{n-1} k^2 = \frac{1}{6n^3} (n-1) \cdot n(2n-1) = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}.$$

Sei $h_n = 1_{\left[0, \frac{1}{n}\right]} \cdot \frac{1}{n^2} + \sum_{k=1}^{n-1} \frac{(k+1)^2}{n^2} 1_{\left[\frac{k}{n}, \frac{k+1}{n}\right]}$. Dann ist $h_n \in \mathcal{T}([0, 1])$, $h_n(x) \geq x^2$ und daher

$$\begin{aligned} \overline{\int}_0^1 x^2 dx &\leq \int_0^1 h_n dx = \frac{1}{n^3} + \frac{1}{n} \sum_{k=2}^n \frac{k^2}{n^2} \\ &= \frac{1}{6n^3} \cdot n(n+1)(2n+1) \\ &= \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2} \end{aligned}$$

also

$$\frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2} \leq \int_0^1 x^2 dx \leq \overline{\int}_0^1 x^2 dx \leq \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}.$$

Da n beliebig war, ist $\int_0^1 x^2 dx = \frac{1}{3}$. Wir verallgemeinern dieses Beispiel auf beliebige stetige Funktionen.

Theorem 5.12 Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist integrierbar.

Beweis:

- (I) Nach Theorem 3.14 ist f beschränkt und nach Theorem 3.19 ist f gleichmäßig stetig. Sei $\varepsilon > 0$ beliebig. Dann gibt es also ein $\delta > 0$ mit $|f(x) - f(y)| < \varepsilon$ für alle x, y mit $|x - y| < \delta$. Wir wählen ein festes n mit $\frac{1}{n} < \delta$ und betrachten die Zerlegung $Z = \{a_k : k = 0, \dots, n\}$ mit $a_k = a + \frac{k(b-a)}{n}$. Nach Theorem 3.19 nimmt f auf dem Intervall $[a_k, a_{k+1}]$ sein Minimum \underline{f}_k an einer Stelle x_k und sein Maximum \overline{f}_k an einer Stelle y_k an. Wegen $|x_k - y_k| \leq (a_{k+1} - a_k) = \frac{1}{n} < \delta$ ist $|\overline{f}_k - \underline{f}_k| < \varepsilon$.

(II) Sei $g_n = 1_{[a_0, a_1]} \cdot f_{-0} + \sum_{k=1}^{n-1} \underline{f}_k 1_{]a_k, a_{k+1}]}$.
 Dann ist $g_n \leq f$. Sei $h_n = 1_{[a_0, a_1]} \cdot \overline{f}_0 + \sum_{k=1}^{n-1} \overline{f}_k 1_{]a_k, a_{k+1}]}$. Dann ist $h_n \geq f$.
 Dann ist $\int_a^b g_n(x) dx \leq \int_a^b f(x) dx \leq \int_a^b h_n(x) dx$ und

$$\int_a^b (h_n(x) - g_n(x)) dx = \frac{(b-a)}{n} \sum_{k=0}^{n-1} \underbrace{(\overline{f}_k - \underline{f}_k)}_{\leq \varepsilon} \leq \frac{(b-a)}{n} \cdot n \cdot \varepsilon = (b-a)\varepsilon.$$
 Also ist $0 \leq \int_a^b f(x) dx - \int_a^b g_n(x) dx \leq (b-a)\varepsilon$ und da $\varepsilon > 0$ beliebig war, folgt die Behauptung.

Wir stellen nun die Rechenregeln für das Integral zusammen:

Satz 5.13 (Rechenregeln für das Integral)

a) Seien $f, g : [a, b]$ integrierbar und $\alpha, \beta \in \mathbb{R}$. Dann ist $\alpha f + \beta g$ integrierbar und es gilt

$$\int_a^b (\alpha f + \beta g)(x) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx \quad (\text{Linearität des Integrals}).$$

b) Sei $f : [a, b] \rightarrow \mathbb{R}$ integrierbar und $f \geq 0$. Dann ist $\int_a^b f(x) dx \geq 0$. (Positivität des Integrals)

Beweis:

a) (I) Sei $\alpha > 0$. Da f integrierbar ist, gilt $\int_a^b f(x) dx = \overline{\int}_a^b f(x) dx$. Zu $\varepsilon > 0$ gibt es daher $g, h \in \mathcal{T}([a, b])$ mit $g \leq f \leq h$ und

$$\int_a^b f(x) dx - \frac{\varepsilon}{\alpha} < \int_a^b g(x) dx \leq \int_a^b h(x) dx < \int_a^b f(x) dx + \frac{\varepsilon}{\alpha}.$$

Multiplikation mit α liefert unter Ausnutzung der Rechenregeln für das Integral von Treppenfunktionen wegen $\alpha g \leq \alpha f \leq \alpha h$

$$\begin{aligned} \alpha \int_a^b f(x) dx - \varepsilon &< \int_a^b \alpha g(x) dx \\ &\leq \int_a^b \alpha f(x) dx \\ &\leq \int_a^b \alpha f(x) dx \\ &\leq \int_a^b \alpha h(x) dx \\ &< \alpha \int_a^b f(x) dx + \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt

$$\alpha \int_a^b f(x) dx \leq \int_a^b \alpha f(x) dx \leq \overline{\int}_a^b \alpha f(x) dx \leq \alpha \int_a^b f(x) dx,$$

also $\alpha \int_a^b f(x) dx = \int_a^b \alpha f(x) dx$.

Für $\alpha = 0$ ist diese Formel klar, für $\alpha < 0$ beweist man sie analog (die Vorzeichen drehen sich bei Multiplikationen mit α einfach um).

(II) Wir zeigen $\int_a^b f_1(x) dx + \int_a^b f_2(x) dx = \int_a^b (f_1 + f_2)(x) dx$.

Zu $\varepsilon > 0$ gibt es $g_1, g_2, h_1, h_2 \in \mathcal{T}([a, b])$ mit $g_j \leq f_j \leq h_j$ und $\int_a^b f_j(x) dx - \frac{\varepsilon}{2} < \int_a^b g_j(x) dx \leq \int_a^b h_j(x) dx < \int_a^b f_j(x) dx + \frac{\varepsilon}{2}$ für $j = 1, 2$.

Addition beider Ungleichungen liefert unter Berücksichtigung der Rechenregeln für das Integral von Treppenfunktionen

$$\begin{aligned} \int_a^b f_1(x)dx + \int_a^b f_2(x)dx - \varepsilon &< \int_a^b (g_1 + g_2)(x)dx \\ &\leq \int_a^b (h_1 + h_2)(x)dx \\ &< \int_a^b f_1(x)dx + \int_a^b f_2(x)dx + \varepsilon. \end{aligned}$$

Wegen $g_1 + g_2 \leq f_1 + f_2 \leq h_1 + h_2$ folgt

$$\begin{aligned} \int_a^b f_1(x)dx + \int_a^b f_2(x)dx - \varepsilon &\leq \int_a^b (f_1 + f_2)(x)dx \\ &\leq \int_a^b (f_1 + f_2)(x)dx \\ &\leq \int_a^b f_1(x)dx + \int_a^b f_2(x)dx + \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt

$$\begin{aligned} \int_a^b f_1(x)dx + \int_a^b f_2(x)dx &\leq \int_a^b (f_1 + f_2)(x)dx \\ &\leq \int_a^b (f_1 + f_2)(x)dx \\ &\leq \int_a^b f_1(x)dx + \int_a^b f_2(x)dx, \end{aligned}$$

also die Behauptung.

(III) Die Formel a) aus dem Satz folgt nun für $f_1 = \alpha f$, $f_2 = \beta g$ und Teil (I) und (II).

b) $g = 0$ ist eine Treppenfunktion auf $[a, b]$ mit $g \leq f$ und $0 = \int_a^b g(x)dx \leq \int_a^b f(x)dx$.

Satz 5.14 Sei $f : [a, b] \rightarrow \mathbb{R}$ integrierbar. Dann ist auch $|f|$ integrierbar und es gilt

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)| dx \leq (b-a) \sup\{|f(x)| : a \leq x \leq b\}$$

Beweis: Wir benutzen die Formeln $f^+(x) = \max(f(x), 0)$ und $f^-(x) = \max(-f(x), 0)$.

Sei f integrierbar. Wir zeigen, daß f^+ integrierbar ist. Zu $\varepsilon > 0$ gibt es $g, h \in \mathcal{T}([a, b])$ mit $g \leq f \leq h$ und

$$\int_a^b f(x)dx - \varepsilon/2 < \int_a^b g(x)dx \leq \int_a^b h(x)dx < \int_a^b f(x)dx + \varepsilon/2.$$

Dann ist $g^+ \leq f^+ \leq h^+$ und wegen $0 \leq h^+ - g^+ \leq h - g$ (einfaches Nachrechnen punktweise) ist nach den Rechenregeln für das Integral von Treppenfunktionen

$$0 \leq \int_a^b h^+(x)dx - \int_a^b g^+(x)dx \leq \int_a^b h(x)dx - \int_a^b g(x)dx < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Also ist

$$0 \leq \int_a^b f^+(x)dx - \int_a^b f^+(x)dx \leq \int_a^b h^+(x)dx - \int_a^b g^+(x)dx < \varepsilon$$

und die Behauptung folgt.

Damit ist nach Satz 5.13 $f^- = f - f^+$, also auch $|f| = f^+ + f^-$ integrierbar. Wegen $|f| \pm f \geq 0$ folgt $\int_a^b |f|(x)dx \pm \int_a^b f(x)dx \geq 0$, also $\pm \int_a^b f(x)dx \leq \int_a^b |f(x)| dx$ und hieraus $|\int_a^b f(x)dx| = \max(\pm \int_a^b f(x)dx) \leq \int_a^b |f(x)| dx$. Der Rest folgt wegen

$$|f| \leq 1_{[a,b]} \cdot \sup\{|f(x)| : x \in [a, b]\}.$$

Satz 5.15 (Integrierbarkeit des Produkts)

Seien $f, g : [a, b] \rightarrow \mathbb{R}$ integrierbar. Dann ist auch $fg : x \mapsto f(x)g(x)$ integrierbar.

Beweis:

(I) Sei $0 \leq f$. *Behauptung:* f^2 ist integrierbar.

Beweis: Sei $M = \sup\{f(x) : x \in [a, b]\}$. Da f als integrierbare Funktion beschränkt ist, ist M wohldefiniert mit $f \leq M1_{[a, b]}$.

Zu $\varepsilon > 0$ gibt es $g, h \in \mathcal{T}([a, b])$ mit $g \leq f \leq h$ und

$$\begin{aligned} cl \quad & \int_a^b f(x)dx - \frac{\varepsilon}{2M} \\ & < \int_a^b g(x)dx \\ & \leq \int_a^b f(x)dx \\ & \leq \int_a^b h(x)dx \\ & < \int_a^b f(x)dx + \frac{\varepsilon}{2M}. \end{aligned}$$

Wir können o.B.d.A $0 \leq g$ und $h \leq M1_{[a, b]}$ annehmen (sonst ersetzen Sie g durch $g^+ \geq g, 0$ und h durch $\inf(h, M1_{[a, b]}) \leq h$. Wegen der Monotonie des Integrals bleiben die Ungleichungen erhalten).
Wegen

$$0 \leq \int_a^b (h^2 - g^2)dx = \int_a^b \underbrace{(h+g)}_{\leq 2M1_{[a, b]}} \underbrace{(h-g)}_{\geq 0} dx \leq 2M \int_a^b (h-g)(x)dx < \varepsilon,$$

also ist wegen $\int_a^b g(x)^2 dx \leq \int_a^b f(x)^2 dx \leq \int_a^b f(x)dx \leq \int_a^b h^2(x)dx$

dann $0 \leq \int_a^b f(x)^2 dx - \int_a^b g(x)^2 dx < \varepsilon$.

Da $\varepsilon > 0$ beliebig war, ist f^2 integrierbar.

(II) Es ist

$$fg = \frac{1}{2}((f+g)^2 - (f-g)^2) = \frac{1}{2}(|f+g|^2 - |f-g|^2).$$

Nach Satz 5.14 und Satz 5.13 sind $|f \pm g|$ integrierbar, also nach (I) auch $|f \pm g|^2$ und damit nach Satz 5.13 auch fg .

Beispiel: Sei $f : [-\pi, \pi] \rightarrow \mathbb{R}$ integrierbar. Nach Theorem 5.12 sind $\sin(kx)$ und $\cos(kx)$ für alle $k \in \mathbb{N}_0$ integrierbar, also auch $x \mapsto f(x)\sin(kx)$ und $x \mapsto f(x)\cos(kx)$. Dieses Beispiel ist fundamental für die Theorie der Fourierreihen.

Das Integral ist Intervall-additiv.

Satz 5.16 Sei $f : [a, b] \rightarrow \mathbb{R}$ und $a < c < b$. f ist genau dann über $[a, b]$ integrierbar, wenn die Einschränkungen $f|_{[a, c]}$ und $f|_{[c, b]}$ über $[a, c]$ bzw. $[c, b]$ integrierbar sind.

Dann gilt $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$.

Beweis:

(I) Sei f über $[a, b]$ integrierbar. Dann ist $f1_{[a,c]}$ und $f1_{]c,b]}$ nach Satz 5.15 integrierbar und $\int_a^c f(x)dx = \int_a^b f(x)1_{[a,c]}(x)dx$, $\int_c^b f(x)dx = \int_a^b f(x)1_{]c,b]}(x)dx$, also gilt

$$\int_a^b f(x)dx = \int_a^b (f(x)1_{[a,c]}(x) + f(x)1_{]c,b]}(x))dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

(II) Sei $f|_{[a,c]}$ und $f|_{]c,b]}$ integrierbar. Dann ist nicht schwer zu sehen, daß $f1_{[a,c]}$ und $f1_{]c,b]}$ über $[a, b]$ integrierbar sind und der Rest folgt wie unter (I).

Als letztes dieser allgemeinen Ergebnisse zeigen wir, wie sich Integral und Konvergenz von Funktionenfolgen verhalten:

Theorem 5.17 Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge integrierbarer Funktionen auf dem Intervall $[a, b]$ und $(f_n)_n$ konvergiere gleichmäßig gegen die Funktion f . Dann ist f integrierbar und es gilt

$$\int_a^b f(x)dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x)dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x)dx.$$

Beweis:

(I) Wir zeigen, daß f integrierbar ist.

Sei $\varepsilon > 0$ beliebig. Zu $\eta = \frac{\varepsilon}{4(b-a)}$ gibt es ein $n(\eta)$ mit $|f(x) - f_n(x)| < \eta$ für alle $x \in [a, b]$ und $n \geq n(\eta)$. Da $\tilde{f} = f_{n(\eta)}$ integrierbar ist, gibt es $g, h \in \mathcal{T}([a, b])$ mit $g \leq \tilde{f} \leq h$ und $\int_a^b (h(x) - g(x))dx < \eta$. Sei $g_1 = g - \eta 1_{[a,b]}$ und $h_1 = h + \eta 1_{[a,b]}$. Dann ist wegen $\tilde{f}(x) = f_{n(\eta)}(x)$, also $\tilde{f}(x) - \eta < f(x) < \tilde{f}(x) + \eta$ für alle x auch

$$g_1(x) = g(x) - \eta \leq \tilde{f}(x) - \eta < f(x) < \tilde{f}(x) + \eta \leq h(x) + \eta = h_1(x),$$

also ist wegen $g_1, h_1 \in \mathcal{T}([a, b])$

$$0 \leq \overline{\int_a^b f(x)dx} - \underline{\int_a^b f(x)dx} \leq \int_a^b h_1(x)dx - \int_a^b g_1(x)dx < \eta + 2\eta(b-a) < \varepsilon.$$

Damit ist f integrierbar, da $\varepsilon > 0$ beliebig war.

(II) Zu $\varepsilon > 0$ gibt es $n(\varepsilon)$ mit $|f(x) - f_n(x)| < \frac{\varepsilon}{b-a}$ für alle $x \in [a, b]$ und $n \geq n(\varepsilon)$. Das heißt $|f - f_n| < \frac{\varepsilon}{(b-a)} 1_{[a,b]}$.

Aus Satz 5.13 und Satz 5.14 folgt

$$\left| \int_a^b f(x)dx - \int_a^b f_n(x)dx \right| \leq \int_a^b |f(x) - f_n(x)| dx < \frac{\varepsilon}{b-a} \int_a^b 1_{[a,b]}(x)dx = \varepsilon$$

für alle $n \geq n(\varepsilon)$. Die Behauptung folgt.

Beispiele:

1. Sei $f(x) = \sum_{n=1}^{\infty} \frac{\sin(nx)}{2^n}$. Dann ist $\int_{-\pi}^{\pi} f(x)dx = \sum_{n=1}^{\infty} \frac{1}{2^n} \int_{-\pi}^{\pi} \sin(nx)dx$. Daß hierbei 0 herauskommt, wird später aus $\int_{-\pi}^{\pi} \sin(nx)dx = 0$ folgen.

2. Sei $f(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$. Dann gilt für $0 < b < 1$

$$\int_0^b f(x) dx = \sum_{n=0}^{\infty} \int_0^b x^n dx = \sum_{n=0}^{\infty} \frac{b^{n+1}}{n+1} \quad (\text{Schulwissen}) = \ln(1-x),$$

wie wir von Beispiel 3 im Anschluß an Korollar 4.26 wissen.

Weitere vernünftige Beispiele können wir erst nach dem nächsten Abschnitt bringen.

Wiederholung:

- Oberes Integral, unteres Integral, Integrierbarkeit und Integral
- Rechenregeln für Integrale
- Ungleichungen für Integrale
- Integral und gleichmäßige Konvergenz.

5.3 Der Hauptsatz der Differential- und Integralrechnung

Überblick: Der wichtigste Satz der elementaren Integralrechnung ist der Hauptsatz. Er besagt: zu jeder stetigen Funktion $f : [a, b] \rightarrow \mathbb{R}$ gibt es genau eine differenzierbare Funktion $F : [a, b]$ mit $F(a) = 0$ und $F'(x) = f(x)$ für alle x . Aus ihm ergeben sich Integrationstechniken, die es erst erlauben, Integrale zu berechnen. Als Anwendungen berechnen wir Volumina von Rotationskörpern, und zeigen die Vertauschbarkeit von Konvergenz und Differentiation unter bestimmten Voraussetzungen.

Wir hatten schon $\int_0^1 x^2 dx = \frac{1}{3}$ im vorigen Abschnitt als Beispiel berechnet. Sei jetzt $t > 0$ beliebig. Wir teilen das Intervall $[0, t]$ in n gleiche Teile $a_k = \frac{tk}{n}$ ($k = 0, \dots, n$). Dann ist $g_n = \sum_{k=1}^{n-1} \frac{k^2 t^2}{n^2} 1_{[a_k, a_{k+1}]}$ eine Treppenfunktion $\leq x^2$ auf $[0, t]$ und

$h_n = \frac{t^2}{n^2} 1_{[0, \frac{t}{n}]} + \sum_{k=1}^{n-1} \frac{(k+1)^2 t^2}{n^2} 1_{[a_k, a_{k+1}]}$ eine Treppenfunktion $\geq x^2$ auf $[0, t]$. Wie früher ergibt sich

$$\int_0^t g_n(x) dx = \frac{t \cdot t^2}{n^3} \sum_{k=1}^{n-1} k^2 = \frac{t^3}{n^3} (n-1)n \cdot (2n-1)$$

und $\int_0^t h_n(x) dx = \frac{t^3}{n^3} n(n+1)(2n+1)$. Es ist

$$\lim_{n \rightarrow \infty} \int_0^t h_n(x) dx = \frac{t^3}{3} = \lim_{n \rightarrow \infty} \int_0^t g_n(x) dx,$$

also ist $\int_0^t x^2 dx = \frac{t^3}{3}$.

Damit ergibt sich, daß die Funktion $t \mapsto \int_0^t x^2 dx = F(t) = \frac{t^3}{3}$, also der Flächeninhalt zwischen der Kurve (x, x^2) und der x -Achse in Abhängigkeit von der Länge t des Intervalls $[0, t]$ nach t differenzierbar ist und die Ableitung gerade den Integranden $t \mapsto t^2$ ergibt. **Aufgabe:** Zeigen Sie bitte auf die gleiche Weise

$$\int_0^t x dx = \frac{1}{2} t^2.$$

Dies ist kein Zufall. Vielmehr gilt der folgende Hauptsatz der Differential- und Integralrechnung: Um ihn in voller Allgemeinheit formulieren zu können, definieren wir:

Ist $b < a$, so sei $\int_a^b f(x)dx = -\int_b^a f(x)dx$.

Definition 5.18 Sei $J = \langle a, b \rangle$ ($a < b$) ein beliebiges Intervall, dessen Endpunkte dazu gehören können, aber nicht müssen und $f : J \mapsto \mathbb{R}$ eine Funktion, die über jedes beschränkte Teilintervall $[u, v] \subset J$ ($u < v$) integrierbar ist. Eine Funktion $F : J \rightarrow \mathbb{R}$ heißt **Stammfunktion** von f , wenn für alle $u, v \in J$ stets $\int_u^v f(x)dx = F(v) - F(u)$ gilt.

Beispiele:

1. Sei $J = \mathbb{R}$, $f = 1_{[0, \infty]}$, also $f(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$
Dann ist $F(x) = xf(x)$ eine Stammfunktion von f . f heißt Heavisidefunktion und repräsentiert Einschaltvorgänge.
2. Sei $J = [0, \infty[$ und $f(x) = x$. Dann ist $f(x) = \frac{x^2}{2}$ eine Stammfunktion von f .
3. Sei $f(x) = 1_{[0, \infty[} - 1_{]-\infty, 0]}$. Dann ist $F(x) = |x|$ eine Stammfunktion von f .

Bevor wir den Hauptsatz beweisen können, benötigen wir das folgende Lemma

Lemma 5.19 (Mittelwertsatz der Integralrechnung)

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann gibt es ein u mit $\int_a^b f(x)dx = f(u)(b - a)$.

Beweis: Sei $m = \min\{f(x) : a \leq x \leq b\}$ und $M = \max\{f(x) : a \leq x \leq b\}$ (Minimum und Maximum existieren nach Theorem 3.14).

Dann ist $m1_{[a, b]} \leq f \leq M1_{[a, b]}$, also erhält man durch Integration $m(b - a) \leq \int_a^b f(x)dx \leq M(b - a)$, insbesondere liegt der Wert $\frac{1}{b - a} \int_a^b f(x)dx$ zwischen m und M . Nach dem Zwischenwertsatz Korollar 3.13 gibt es ein u mit $f(u) = \frac{1}{b - a} \int_a^b f(x)dx$.

Theorem 5.20 (Hauptsatz der Differential- und Integralrechnung)

Sei $J = \langle a, b \rangle$ ein Intervall, dessen Endpunkte $a < b$ dazu gehören können, aber nicht müssen. Sei $f : J \mapsto \mathbb{R}$ eine stetige Funktion.

- a) Sei $x_0 \in J$ beliebig. Dann ist $F_0 : x \rightarrow \int_{x_0}^x f(s)ds$ eine Stammfunktion zu f .
- b) Eine Funktion G ist genau dann eine Stammfunktion von f , wenn G auf J differenzierbar ist und $G' = f$ gilt.

Korollar 5.21 Die Differenz zweier Stammfunktionen der stetigen Funktion f ist konstant.

Beweis: Beweis des Theorems

- a) Die Behauptung folgt aus der Intervalladditivität des Integrals (s. Satz 5.16)

b) Sei G eine Stammfunktion von F . Nach dem Mittelwertsatz der Integralrechnung Lemma 5.19 gilt für ein $c \in J$ und $x \neq c$

$\frac{G(x) - G(c)}{x - c} = \frac{1}{x - c} \int_c^x f(s) ds = f(\theta(x))$ mit $\theta(x)$ zwischen x und c , also $|\theta(x) - c| \leq |x - c|$. Zu $\varepsilon > 0$ existiert ein $\delta > 0$ mit $|f(y) - f(c)| < \varepsilon$ für alle y mit $|y - c| < \delta$, also ist für $|x - c| < \delta$

$$\left| \frac{G(x) - G(c)}{x - c} - f(c) \right| = |f(\theta(x)) - f(c)| < \varepsilon.$$

Damit folgt $G'(c) = f(c)$.

Sei umgekehrt G auf J differenzierbar mit $G' = f$. Dann ist $(G - F_0)' = G' - F_0' = 0$, da $f = F_0'$ nach a) und dem bisher Bewiesenen. Damit ist $G - F_0 = d$ (Konstante), also $G(v) - G(u) = F_0(v) - F_0(u) = \int_u^v f(x) dx$.

Integrationstechniken

Nach dem Hauptsatz muß man nur eine passende Stammfunktion kennen, um Flächeninhalte von Flächen zwischen der x -Achse und dem Graphen einer stetigen Funktion berechnen zu können. Um zu zeigen, daß F eine Stammfunktion zu f ist, muß man nur $F' = f$ beweisen. Wir schreiben statt F auch $\int f(x) dx$ oder $\int f dx$. *Integrale ohne Integrationsgrenzen bezeichnen also unbestimmte Integrale oder Stammfunktionen.* Die Bestimmung von Stammfunktionen läuft also auf die "Umkehrung" der Formeln für die Differentiation hinaus. Bitte beachten Sie, daß Stammfunktionen nur bis auf eine Konstante *const* eindeutig bestimmt sind! Denken Sie immer auch an die Konstante!

Grundformeln

Beweisen Sie die Grundformeln einfach durch Differentiation der rechten Seite.

1. $\int x^\alpha dx = \frac{1}{\alpha + 1} x^{\alpha+1} + \text{const}$ für $\alpha \neq -1$.
2. $\int \frac{1}{x} dx = \ln|x| + \text{const}$; unterscheiden Sie stets $x < 0$ und $x > 0$!
3. $\int \exp(x) dx = \exp(x) + \text{const}$.
4. $\int \sin(x) dx = -\cos(x) + \text{const}$ und $\int \cos(x) dx = \sin(x) + \text{const}$.
5. $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin(x) + \text{const}$ auf dem Intervall $] -1, 1[$.
6. $\int \frac{dx}{-\sqrt{1-x^2}} = \arccos(x) + \text{const}$ auf dem Intervall $] -1, 1[$.
7. $\int \frac{dx}{1+x^2} = \arctan(x) + \text{const}$.
8. $\int \frac{dx}{\sqrt{x^2+1}} = \text{Arsinh}(x) + \text{const}$.
9. $\int \frac{dx}{\sqrt{x^2-1}} = \text{Arcosh}(x) + \text{const}$ auf dem Intervall $]1, \infty[$.
10. $\int \frac{dx}{1-x^2} = \text{Artanh}(x) + \text{const}$ auf dem Intervall $] -1, 1[$.

11. $\int (\alpha f(x) + \beta g(x)) dx = \alpha \int f(x) dx + \beta \int g(x) dx$ für beliebiges $\alpha, \beta \in \mathbb{R}$.

Mit diesen Grundformeln können Sie bereits die folgenden Integrale berechnen. *Bitte üben Sie die Integration!* Sie brauchen die Integration später ähnlich als Basis für die Analysis in Physik und Bioinformatik.
Aufgaben: Berechnen Sie die folgenden Integrale:

1. $\int (x^2 - 3x) dx$
2. $\int (1/x - 1/x^2) dx$
3. $\int (49x^{48} - 7 \exp(x)) dx$
4. $\int (2x + \frac{25}{(1+x^2)}) dx$

Partielle Integration

Die partielle Integration ist die “Umkehrung” der Produktregel.
Für stetig differenzierbare Funktionen f und g gilt:

$$\int f(x)g'(x) dx = fg - \int f'(x)g(x) dx$$

Für das bestimmte Integral $\int_a^b f(x)g'(x) dx$ bedeutet dies:

$$\begin{aligned} \int_a^b f(x)g'(x) dx &= f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x) dx \\ &= [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx \end{aligned}$$

Somit bekommen wir für $f(x) = x$ und $g(x) = \sin(x)$:

$$\int x \cos(x) dx = x \sin(x) - \int \sin(x) dx = x \sin(x) + \cos(x) + \text{const.}$$

Das Problem besteht im wesentlichen darin, einen Integranden geschickt in zwei Faktoren zu zerlegen.

Aufgaben:

1. $\int \sqrt{1-x^2} dx$ im Intervall $[-1, 1]$. Tip: $g' = 1$.
2. $\int \sqrt{1+x^2} dx$ in \mathbb{R} . Tip: $g' = 1$.
3. $\int \sqrt{x^2-1} dx$ im Intervall $[1, \infty[$. Tip: $g' = 1$.
4. $\int x^2 \sin(x) dx$. Tip: Mehrfache Anwendung der partiellen Integration.
5. $\int x \ln(x) dx$
6. $\int x^2 \exp(x) dx$.
7. $\int \cos(x)^2 dx$. Tip: $g'(x) = \cos(x)$.
8. $\int \sin(x)^3 dx$.

Integration durch Substitution

Die Umkehrung der Kettenregel ergibt die folgende Integrationsformel, die Sie einfach durch Ableiten der rechten Seite nachprüfen:

Sei $f : [a, b] \rightarrow [c, d]$ stetig differenzierbar, und $g : [c, d] \rightarrow \mathbb{R}$ stetig mit Stammfunktion G . Dann ist

$$\int g(f(x))f'(x)dx = G(f(x)) + \text{const.}$$

Beispiel: In $\int \frac{f'(x)}{f(x)} dx = \ln(|f(x)|) + \text{const}$ ist $g(x) = 1/x$. Für das bestimmte Integral bedeutet dies:

$$\int_a^b g(f(x))f'(x)dx = G(f(b)) - G(f(a)) = \int_{f(a)}^{f(b)} g(u)du$$

Die Integrationsgrenzen müssen also mittransformiert werden. **Aufgaben:**

1. Zeigen Sie bitte: Das Integral ist *translationsinvariant*, das heißt $\int_a^b f(x+c)dx = \int_{a+c}^{b+c} f(x)dx$.
2. Zeigen Sie bitte: Das Integral berücksichtigt Umskalierungen: $\int_a^b f(vx)dx = 1/v \int_{av}^{bv} f(x)dx$.
3. Berechnen Sie bitte $\int \frac{dx}{x^2+a^2}$. Tip: $g(u) = 1/(u^2+1)$.
4. Berechnen Sie bitte $\int \ln(x)/x dx$ für $x > 0$. Tip: Substitution $t = \ln(x)$.
5. Berechnen Sie bitte $\int \cos(\ln(x))dx$ für $x > 0$. Tip: Substitution $t = \ln(x)$ mit anschließender zweimaliger partieller Integration.
6. Berechnen Sie bitte noch einmal das Integral $\int \sqrt{1-x^2}dx$, jetzt mit der Substitution $x = \sin(t)$.

Integration rationaler Funktionen

Im folgenden bringen wir die wichtigsten Beispiele. Weitere Formeln finden Sie in jedem Nachschalgerwerk oder mit Computeralgebrasystemen. Seien Polynome $P(x) = \sum_{k=0}^m a_k x^k$ und $Q(x) = \sum_{k=0}^n b_k x^k$ gegeben.

wir wollen $\int \frac{P(x)}{Q(x)} dx$ berechnen. **Beispiele:**

1. $\int \frac{dx}{x^2-1}$; es ist $\frac{1}{x^2-1} = \frac{1}{(x+1)(x-1)} = (\frac{1}{x-1} - \frac{1}{x+1})/2$. Mit Integration durch Substitution ergibt sich aus der Grundformel $\int \frac{1}{u} du = \ln(|u|) + \text{const}$:

$$\int \frac{dx}{x^2-1} = \frac{1}{2}(\ln|x-1| - \ln|x+1|) + \text{const} = \ln \sqrt{\frac{|x-1|}{|x+1|}} + \text{const.}$$

2. Allgemeiner: Ist $Q(x) = (x-x_1) \dots (x-x_n)$ mit $x_i \neq x_j$ für $i \neq j$, so ist $\frac{1}{Q(x)} = \sum_{k=1}^n \frac{A_k}{x-x_k}$ mit $A_k = \frac{1}{\prod_{i \neq k} (x_k - x_i)}$. Das beweisen Sie durch Nachrechnen. Dann ist $\int \frac{dx}{Q(x)} = \sum_{k=1}^n A_k \ln|x-x_k| + \text{const}$.

3. $\int \frac{x dx}{x^2 + 1} = \frac{1}{2} \int \frac{2x dx}{x^2 + 1} = \ln \sqrt{x^2 + 1} + \text{const}$ (nach Integration durch Substitution).

4. $\int \frac{dx}{x^2 + bx + c} = ?$. Es ist $x^2 + bx + c = (x + \frac{b}{2})^2 + (c - \frac{b^2}{4})$. Ist $c - \frac{b^2}{4} < 0$, so liegt Beispiel 2 vor,

genauer: $x^2 + bx + c = (u + d)(u - d)$ mit $u = (x + \frac{b}{2})$ und $d = \sqrt{\frac{b^2}{4} - c}$.

Ist $c - \frac{b^2}{4} > 0$, so ist $x^2 + bx + c = (x + \frac{b}{2})^2 + d^2$ mit $d = \sqrt{c - \frac{b^2}{4}}$. Man erhält aus der Grundformel 7 mit Integration durch Substitution

$$\int \frac{dx}{x^2 + bx + c} = \frac{1}{d} \arctan\left(\frac{x + \frac{b}{2}}{d}\right) + \text{const.}$$

Integration von Potenzreihen

Sei $f(x) = \sum_{k=0}^{\infty} a_k x^k$ mit Konvergenzradius $R > 0$. Dann ist

$$\int \sum_{k=0}^{\infty} a_k x^k dx = \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1} + \text{const.}$$

Unmöglichkeit der formelmäßigen Integration aller stetigen Funktionen

Es gibt eine Reihe wichtiger Funktionen, deren Stammfunktion sich nicht durch Formeln mit den bisher bekannten Funktionen ausdrücken lassen. Ein solches Beispiel ist die Stammfunktion zur "Dichte" $\frac{1}{\sqrt{2\pi}} \int \exp(-x^2) dx$, die sog. *Normalverteilung* in der Wahrscheinlichkeitstheorie

Applet "Symbolische Integration und Differentiation" Mit dem selben Applet können sie analog zu Ableitungen differenzierbarer Funktionen Stammfunktionen integrierbarer Funktionen bestimmen

Kapitel 6

Anwendungen

Überblick: In diesem Kapitel wenden wir das bisher Behandelte auf elementargeometrische Probleme, auf komplexe Differentiation und Integration nach einem reellen Parameter und auf periodische Funktionen an und behandeln schließlich Integrale über offenen Mengen und Integrale unbeschränkter Funktionen.

6.1 Ergänzungen

Überblick: In diesem Abschnitt behandeln wir die Vertauschbarkeit von Differentiation und Konvergenz, die Bogenlänge von Graphen und die Oberfläche und den Rauminhalt von Drehkörpern.

Vertauschbarkeit von Differentiation und Konvergenz

Wir wissen schon, daß man Potenzreihen gliedweise differenzieren kann. In der gesamten Physik, vor allem in der Nachrichtentechnik und Signalverarbeitung, spielen Reihen der Form $\sum_0^\infty a_k \cos(kt)$ und $\sum_1^\infty b_k \sin(kt)$ eine herausragende Rolle (s. den Abschnitt über periodische Funktionen). Für all diese Anwendungen ist der folgende Satz von entscheidender Bedeutung, den wir ganz einfach aus dem Hauptsatz der Differential- und Integralrechnung gewinnen.

Theorem 6.1 *Sei $(g_n)_{n \geq 0}$ eine Folge stetig differenzierbarer Funktionen auf dem abgeschlossenen Intervall $[a, b]$. Es konvergiere die Reihe $\sum_{n=0}^\infty g'_n(x)$ der Ableitungen gleichmäßig auf $[a, b]$ gegen die Funktion h . Es konvergiere außerdem die Reihe $\sum_{n=0}^\infty g_n(c)$, wo $a \leq c \leq b$ ein fester Punkt ist. Dann konvergiert die Reihe $\sum_{n=0}^\infty g_n(x)$ gleichmäßig auf $[a, b]$. Die Grenzfunktion f ist differenzierbar und es ist $f' = h = \sum_{n=0}^\infty g'_n$.*

Kurz zusammengefaßt: Eine solche Reihe $\sum_{n=0}^\infty g_n$ konvergiert gleichmäßig und man darf sie gliedweise differenzieren, falls die Reihe in einem Punkt konvergiert und die **abgeleitete Reihe** $\sum_{n=0}^\infty g'_n$ **gleichmäßig konvergiert**.

Beweis: Nach dem Hauptsatz der Differential- und Integralrechnung ist $g_n(x) = g_n(c) + \int_c^x g'_n(s) ds$. Nach dem Satz über die gleichmäßige Konvergenz einer Folge stetiger Funktionen (Theorem 3.22) ist h

stetig, also nach dem Hauptsatz integrierbar und nach dem Theorem 5.17 über die Vertauschbarkeit von Konvergenz und Integration erhält man

$$H(x) := \int_c^x h(s) ds = \sum_{n=0}^{\infty} \int_c^x g'_n(s) ds = \sum_{n=0}^{\infty} (g_n(x) - g_n(c)).$$

Nach Voraussetzung konvergiert $\sum_{n=0}^{\infty} g_n(c)$, nach den Rechenregeln für Konvergenz demnach auch $\sum_{n=0}^{\infty} g_n(x) = \sum_{n=0}^{\infty} (g_n(x) - g_n(c)) + \sum_{n=0}^{\infty} g_n(c)$. Damit ist $f(x) = H(x) + \sum_{n=0}^{\infty} g_n(c) = \sum_{n=0}^{\infty} g_n(x)$ und $f'(x) = H'(x) = h(x) = \sum_{n=0}^{\infty} g'_n(x)$.

Wir zeigen nun noch, daß die Reihe $\sum_{n=0}^{\infty} g_n(x)$ *gleichmäßig* gegen $f(x)$ konvergiert. Sei $S_n(x) = \sum_{k=0}^n g_k(x)$. Dann ist

$$\begin{aligned} |f(x) - S_n(x)| &= |f(x) - f(c) + f(c) - S_n(c) - (S_n(x) - S_n(c))| \\ &\leq |(f(x) - f(c)) - (S_n(x) - S_n(c))| + |f(c) - S_n(c)| \\ &= \left| \int_c^x h(x) - \sum_{k=0}^n \int_c^x g'_k(x) dx \right| + |f(c) - S_n(c)| \\ &\leq \int_a^b |h(x) - \sum_{k=0}^n g'_k(x)| dx + |f(c) - S_n(c)|. \end{aligned}$$

Nach Voraussetzung gibt es zu $\varepsilon > 0$ ein $n_1(\varepsilon)$ mit $|f(c) - S_n(c)| < \frac{\varepsilon}{2}$ für alle $n \geq n_1(\varepsilon)$ und ein $n_2(\varepsilon)$ mit $|h(x) - \sum_{k=0}^n g'_k(x)| < \frac{\varepsilon}{2(b-a)}$ für alle $n \geq n_2(\varepsilon)$ und alle x . Für $n \geq n(\varepsilon) := n_1(\varepsilon) + n_2(\varepsilon)$ ist dann für alle x : $|f(x) - S_n(x)| \leq \int_a^b \frac{\varepsilon}{2(b-a)} dx + \frac{\varepsilon}{2} = \varepsilon$.

Beispiele:

1. Aus diesem Satz folgt noch einmal, daß man Potenzreihen gliedweise differenzieren kann: ist $f(x) = \sum_{n=0}^{\infty} a_n x^n$, so setze $g(x) = a_n x^n$. $g'_n(x) = n a_n x^{n-1}$. Die Potenzreihe $\sum_{n=1}^{\infty} n a_n x^{n-1}$ hat nach Aufgabe 29 den gleichen Konvergenzradius R wie die gegebene. Auf jedem Intervall $[-b, b]$ mit $b < R$ konvergiert sie also gleichmäßig. Die Behauptung folgt.

2. $\sum_{n=0}^{\infty} \frac{n \cos(nx)}{2^n}$ ist gleichmäßig konvergent wegen $|\cos(nx)| \leq 1$. Der Grenzwert ist einfach zu berechnen: $g_n(x) = \frac{\sin(nx)}{2^n}$ erfüllt $g'_n(x) = \frac{n \cos(nx)}{2^n}$.

$$\text{Es ist } \sum_{n=0}^{\infty} g_n(x) = \frac{1}{i} \Im \sum_{n=0}^{\infty} \frac{e^{inx}}{2^n} = \frac{1}{i} \Im \left(\frac{1}{1 - \frac{e^{ix}}{2}} \right) = \frac{4 + 2 \sin(x)}{5}.$$

$$\text{Also ist } \sum_{n=0}^{\infty} \frac{n \cos(nx)}{2^n} = \left(\frac{4 + 2 \sin(x)}{5} \right)' = \frac{-2 \cos(x)}{5}.$$

3. Hier ist eine Reihe, für die der Satz nicht gilt:

$$g_n(x) = \frac{1}{(n+1)n} ((n+1) \cos(n^3 x) - n \cos((n+1)^3 x)), \quad (n \geq 1) \text{ d.h.}$$

$\sum_{k=1}^n g_k(x) = -\frac{\cos((n+1)^3 x)}{(n+1)} + \cos(x)$. Die Reihe konvergiert gleichmäßig gegen $\cos(x)$ die Grenzfunktion ist also differenzierbar, aber die gliedweise abgeleitete Reihe konvergiert nur in den Punkten $\frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$.

4. Hier ist ein Beispiel einer gleichmäßig konvergenten Reihe stetig differenzierbarer Funktionen, deren

Grenzfunktion in 0 nicht differenzierbar ist:

$$g_n(x) = \sqrt{\frac{1}{n+1} + x^2} - \sqrt{\frac{1}{n} + x^2} \quad \text{auf } [-1, 1]$$

$$\sum_{n=0}^{\infty} g_n(x) = |x| - \sqrt{1+x^2}$$

Geometrische Probleme

Flächeninhalt einer Kreisscheibe mit Radius r

Der Rand eines Kreises mit Radius r ist gegeben durch $x = r \cos \rho$, $y = r \sin \rho$, das heißt durch die Menge $R = \left\{ \begin{pmatrix} r \cos \rho \\ r \sin \rho \end{pmatrix} : 0 \leq \rho \leq 2\pi \right\}$.

Wir teilen ihn in vier gleiche Teile $0 \leq \rho \leq \pi/2$, $\pi/2 \leq \rho \leq \pi$, usw. Der Rand des ersten Teiles ist der Graph von $f(u) = \sqrt{r^2 - u^2}$, ($0 \leq u \leq r$). Also ist $4 \int_0^r \sqrt{r^2 - u^2} du$ der gesuchte Flächeninhalt.

$u = r \cos(\rho)$, ($0 \leq \rho \leq \pi/2$), also $\frac{du}{d\rho} = -r \sin \rho$ ergibt

$$\begin{aligned} \int_0^r \sqrt{r^2 - u^2} du &= -r^2 \int_{\pi/2}^0 \sin(\rho)^2 d\rho \\ &= -\frac{r^2}{2} \int_{\pi/2}^0 (1 - \cos(2\rho)) d\rho \\ &= -\frac{r^2}{2} \left[\rho - \frac{1}{2} \sin(2\rho) \right]_{\pi/2}^0 \\ &= \frac{\pi r^2}{4}. \end{aligned}$$

Also ist der Flächeninhalt des Kreises mit Radius r gleich πr^2 .

Bogenlänge des Graphen einer stetig differenzierbaren Funktion

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar. Ihr Graph ist dann $\{(x, f(x)) : a \leq x \leq b\}$. Wir teilen das Intervall $[a, b]$ in n gleiche Teile ein und ersetzen die Kurven durch den Sekantenzug durch die Punkte $(a_k, f(a_k))$ mit $a_k = a + \frac{k(b-a)}{n}$. Das k -te Sekantenstück wird durch die Gleichung $s_k(t) = f(a_k) + \frac{f(a_{k+1}) - f(a_k)}{a_{k+1} - a_k} \cdot (t - a_k)$ mit $a_k \leq t \leq a_{k+1}$ gegeben, der Graph G_k ist also $\{(t - a_k, s_k(t)) : a_k \leq t \leq a_{k+1}\}$. Nach dem Satz des Pythagoras ist die Länge ℓ_k dieses Stückchens $\sqrt{(a_{k+1} - a_k)^2 + (s(a_{k+1}) - s(a_k))^2}$. Einsetzen liefert

$$\begin{aligned} \ell_k^2 &= \frac{(b-a)^2}{n^2} + (f(a_{k+1}) - f(a_k))^2 \\ &= \frac{(b-a)^2}{n^2} \left(1 + \frac{(f(a_{k+1}) - f(a_k))^2}{(a_{k+1} - a_k)^2} \right) \end{aligned}$$

also nach dem 1. Mittelwertsatz

$$\begin{aligned} \ell_k &= \frac{(b-a)}{n} \sqrt{1 + \left(\frac{f(a_{k+1}) - f(a_k)}{a_{k+1} - a_k} \right)^2} \\ &= \frac{(b-a)}{n} \sqrt{1 + f'(t_k)^2}, \end{aligned}$$

mit $a_k < t_k < a_{k+1}$. Summiert man die Teillängen ℓ_k auf, so erhält man für die Gesamtlänge

$$\ell = \frac{(b-a)}{n} \sum_{k=0}^{n-1} \sqrt{1+f'(t_k)^2} =: \ell(n).$$

Da die Funktion $g(t) = \sqrt{1+f'(t)^2}$ stetig, also integrierbar ist, folgt

$$\lim_{n \rightarrow \infty} \ell(n) = \int_a^b \sqrt{1+f'(t)^2} dt$$

und das setzen wir dann per Definition als die Länge der Kurve $\{(x, f(x)) : a \leq x \leq b\}$. Für den Viertelkreis mit Radius r im positiven Quadranten erhalten wir mit $f(t) = \sqrt{r^2 - t^2}$ für $0 \leq t < r$: $f'(t) = -\frac{t}{\sqrt{r^2 - t^2}}$, also

$$1 + f'(t)^2 = 1 + \frac{t^2}{r^2 - t^2} = \frac{r^2}{r^2 - t^2}, \text{ also}$$

$\sqrt{1+f'(t)^2} = \frac{r}{\sqrt{r^2 - t^2}}$. Dies ist für $0 \leq t < r$ die Ableitung der Funktion $r \arcsin(\frac{t}{r})$. Also ist für

$0 < \varepsilon < r$: $\int_0^{r-\varepsilon} \sqrt{1+f'(t)^2} dt = r \arcsin(\frac{r-\varepsilon}{r})$. Für $\varepsilon \rightarrow 0$ erhält man $\int_0^r \sqrt{1+f'(t)^2} dt = r\pi/2$. Multipliziert man dies mit 4, so folgt:

$$\text{der Kreisumfang ist } 2\pi r$$

.

6.1.1 Oberfläche und Volumen von Rotationskörpern

Drehen wir den Graphen $\{(x, f(x)) : a \leq x \leq b\}$ einer Funktion $f : [a, b] \rightarrow \mathbb{R}_+ = [0, \infty[$ um die x -Achse so erhalten wir einen **Rotationskörper** K_f . Zu einer Einteilung $Z_n = \{a + \frac{k(b-a)}{n} : 0 \leq k \leq n\}$ erhalten wir als Annäherung an K_f einen aus lauter Zylindern $Z_n(k)$ mit Radius $f(\frac{k(b-a)}{n})$ und Höhe $\frac{(b-a)}{n}$ zusammengesetzten Körper. Jeder einzelne Zylinder $Z_n(k)$ hat das Volumen $\pi(f(\frac{k(b-a)}{n}))^2 \cdot \frac{(b-a)}{n}$ und den Mantel $2\pi f(\frac{k(b-a)}{n}) \frac{(b-a)}{n}$. Damit ist das Volumen des aus den Zylindern zusammengesetzten Körpers $V(\cup_{k=0}^{n-1} Z_n(k))$ gegeben durch $\sum_{k=0}^{n-1} \pi(f(\frac{k(b-a)}{n}))^2 \frac{(b-a)}{n}$ und die Mantelfläche $O(\cup_{k=0}^{n-1} Z_n(k))$ durch $\sum_{k=0}^{n-1} 2\pi f(\frac{k(b-a)}{n}) \frac{(b-a)}{n}$.

Ist f stetig so konvergiert $V(\cup_{k=0}^{n-1} Z_n(k))$ für n gegen unendlich gegen das Integral $\pi \int_a^b f(x)^2 dx$, das dann vernünftigerweise als **Volumen** von K_f bezeichnet wird. Die Mantelflächen konvergieren gegen $2\pi \int_a^b f(x) dx$. Das Integral ist also die **Oberfläche** von K_f . Wir fassen zusammen.

Satz 6.2 Sei $f : [a, b] \rightarrow \mathbb{R}_+$ stetig und K_f der durch Drehung der Kurve um die x -Achse erhaltene (geometrische) Rotationskörper K_f . Dann ist sein Volumen $V(K_f) = \pi \int_a^b f(x)^2 dx$, seine Mantelfläche $O(K_f) = 2\pi \int_a^b f(x) dx$.

Beispiele:

1. $f(x) = \sqrt{r^2 - x^2}$ für $0 \leq x \leq r$. Der Rotationskörper ist die Halbkugel mit Radius r . Ihr Volumen ist $\pi \int_0^r (r^2 - x^2) dx = \frac{2}{3}\pi r^2$. Ihre Mantelfläche ist $2\pi \int_0^r \sqrt{r^2 - x^2} dx$. Durch die Substitution $\sqrt{r^2 - x^2} = r \sin t$ erhält man $\int_0^r \sqrt{r^2 - x^2} dx = \frac{\pi}{2} r$, also $O(K_f) = \pi^2 r$.

2. $f(x) = \sqrt{x}$ auf $[1, 2]$. Man erhält $V(K_f) = \pi \int_1^2 x dx = 3\pi/2$ und

$$\begin{aligned} O(K_f) &= 2\pi \int_1^2 \sqrt{x} dx = 2\pi \left[\frac{2}{3} x^{3/2} \right]_1^2 \\ &= \frac{4\pi}{3} (2\sqrt{2} - 1). \end{aligned}$$

6.2 Periodische Funktionen und Fourier-Reihen

Differentiation und Integration komplexwertiger Funktionen auf Intervallen

Sei $f : [a, b] \rightarrow \mathbb{C}$ eine komplexwertige Funktion. Wir definieren, daß f stetig ist in Einklang mit 2.5, und Kapitel 3 bis 5:

f ist **stetig**, wenn die Funktionen $t \rightarrow \Re(f(t))$ und $t \rightarrow \Im(f(t))$ stetig sind. f ist **differenzierbar**, wenn die Funktionen $t \mapsto \Re(f(t))$ und $t \mapsto \Im(f(t))$ differenzierbar sind. Dann setzen wir $f'(t) = (\Re f)'(t) + i(\Im f)'(t)$. f ist **integrierbar** über $[a, b]$, wenn $t \mapsto \Re f(t)$ und $t \mapsto \Im f(t)$ integrierbar sind. Dann setzen wir $\int_a^b f(t) dt = \int_a^b \Re f(t) dt + i \int_a^b \Im f(t) dt$. **Beispiele:**

1. $f(t) = \exp(it) = \cos(t) + i \sin(t)$. f ist stetig differenzierbar und es ist $f'(t) = -\sin(t) + i \cos(t) = i \exp(it)$.
2. Genauso zeigt man $f'(\exp(int)) = in \exp(int)$.
3. $\int_0^{2\pi} \exp(int) dt = \int_0^{2\pi} \cos(nt) dt + i \int_0^{2\pi} \sin(nt) dt = 0, n \neq 0$
4. $\frac{1}{2\pi} \int_0^{2\pi} \exp(-int) \exp(imt) dt = \begin{cases} 1 & n = m \\ 0 & \text{sonst} \end{cases}$

Definition 6.3 Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt **periodisch** mit Periode $\omega \neq 0$, wenn stets $f(t + \omega) = f(t)$ gilt.

Bemerkungen:

1. Die Funktionen $\exp(int)$ sind $\frac{2\pi}{n}$ -periodisch.
2. Ist f periodisch mit Periode ω , so ist die Funktion $\tilde{f}(t) = f\left(\frac{\omega t}{2\pi}\right)$ 2π -periodisch. Es genügt also 2π -periodische Funktionen zu betrachten.

Definition 6.4 a) Eine Funktion f der Form $f(t) = \sum_{k=-m}^n a_k \exp(ikt)$ heißt **trigonometrisches Polynom** ($m, n \in \mathbb{Z}, a_k \in \mathbb{C}$).

b) Eine Reihe der Form $\sum_{k=-\infty}^{\infty} a_k \exp(ikt)$ heißt **Fourier-Reihe** oder auch **trigonometrische Reihe**. Sie heißt **konvergent**, wenn die beiden Teilreihen $\sum_{k=0}^{\infty} a_k \exp(ikt)$ und $\sum_{\ell=1}^{\infty} a_{-\ell} \exp(-i\ell t)$ konvergieren. Der Grenzwert ist dann

$$\sum_{k=-\infty}^{\infty} a_k \exp(ikt) := \sum_{k=0}^{\infty} a_k \exp(ikt) + \sum_{\ell=1}^{\infty} a_{-\ell} \exp(-i\ell t).$$

c) Die Reihe konvergiert gleichmäßig, wenn die beiden Teilreihen gleichmäßig konvergieren.

Bemerkung: Offensichtlich konvergiert die Fourierreihe genau dann gleichmäßig, wenn Real- und Imaginärteil gleichmäßig konvergieren. Dann darf man also gliedweise differenzieren und integrieren (nach den Sätzen des ersten Abschnitts dieses Kapitels).

Die Koeffizienten a_k lassen sich leicht berechnen.

Satz 6.5 Sei $f(t) = \sum_{-\infty}^{\infty} a_k \exp(ikt)$, und die Reihe konvergiere gleichmäßig. Dann ist

$$\frac{1}{2\pi} \int_0^{2\pi} f(t) \exp(-int) dt = a_n \quad \forall n.$$

a_n heißt der n -te **Fourierkoeffizient**.

Beweis: Da die Reihe gleichmäßig konvergiert, ist f stetig nach Theorem 3.22. Damit ist sie über $[0, 2\pi]$ integrierbar, also ist auch $f(t) \exp(-int)$ integrierbar (Satz 5.15). Wir multiplizieren die Reihe aus und integrieren gliedweise:

$$\frac{1}{2\pi} \int_0^{2\pi} f(t) \exp(-int) dt = \sum_{k=-\infty}^{\infty} a_k \frac{1}{2\pi} \int_0^{2\pi} \exp(ikt) \exp(-int) dt = a_n$$

nach Beispiel 4 oben.

Für unser Ziel, stetige 2π -periodische Funktionen in Fourier-Reihen zu entwickeln, benötigen wir den folgenden Satz, den wir ohne Beweis angeben.

Theorem 6.6 (Satz von Weierstraß)

Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig und 2π -periodisch. Dann gibt es eine Folge $(T_n)_n$ trigonometrischer Polynome T_n , die gleichmäßig gegen f konvergiert.

Wir ziehen sofort die Folgerung:

Satz 6.7 Seien f, g stetig und 2π -periodisch.

Gilt für die Fourierkoeffizienten

$$a_k(f) := \frac{1}{2\pi} \int_0^{2\pi} f(t) \exp(-ikt) dt = a_k(g) := \frac{1}{2\pi} \int_0^{2\pi} g(t) \exp(-ikt) dt$$

für alle $k \in \mathbb{Z}$, so ist $f = g$.

Der Satz besagt also, daß eine Messung der Fourierkoeffizienten die Funktion schon **eindeutig** bestimmt.

Beweis: Sei $h = f - g$. Wir zeigen $h(t) = 0$ für alle t . Nach Voraussetzung ist $a_k(h) = 0$ für alle k . Damit ist aber $\int_0^{2\pi} h(t) \sum_{k=m}^n b_k \exp(ikt) dt = 0$. Mit h ist auch die konjugiert komplexe Funktion \bar{h} stetig und 2π -periodisch. Also gibt es nach dem Theorem von Weierstraß eine Folge (T_n) trigonometrischer Funktionen, die gleichmäßig gegen \bar{h} konvergiert. Also konvergiert hT_n gleichmäßig gegen $h\bar{h} = |h|^2$. Mit Theorem 5.17 folgt

$$0 = \lim \int_0^{2\pi} h(t) T_n(t) dt = \int_0^{2\pi} |h(t)|^2 dt.$$

Wäre $|h(t_0)|^2 > 0$, so gäbe es $\delta > 0$ mit $|h(t)|^2 \geq |h(t_0)|^2/2$ für $|t - t_0| < \delta$, also wäre $\int_0^{2\pi} |h(t)|^2 dt \geq \int_{t_0-\delta}^{t_0+\delta} \frac{|h(t_0)|^2}{2} dt = \delta|h(t_0)|^2 \neq 0$. Also ist $|h(t)|^2 = 0$ für alle t und damit $h(t) = 0$, also $f(t) = g(t)$.

Wir ordnen jeder stetigen 2π -periodischen Funktion f ihre Fourierreihe $\sum_{k=-\infty}^{\infty} a_k(f) \exp(ikt)$ zu, wo $a_k(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \exp(-ikt) dt$ ist.

Korollar 6.8 Sei f stetig und 2π -periodisch. Konvergiert die zugeordnete Fourierreihe gleichmäßig gegen eine Funktion g , so ist $f = g$.

Beweis: Nach Satz 6.5 haben f und g die gleichen Fourierkoeffizienten. Also ist $f = g$ nach dem gerade bewiesenen Satz.

Es gibt eine Abschätzung für die Fourierkoeffizienten, die vom Astronom F.W. Bessel stammt, und die wir für das nächste Ergebnis brauchen:

Satz 6.9 (Besselsche Ungleichung) Sei $f : \mathbb{R} \mapsto \mathbb{R}$ $2 - \pi$ -periodisch und stetig. Dann gilt für die Fourier-Koeffizienten $a_k(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \exp(-ikt) dt$

$$\sum_{k=-\infty}^{\infty} |a_k|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt.$$

Beweis: Sei $S_n(t) = \sum_{k=-n}^n a_k(f) \exp(-ikt)$.

Dann ist

$$\overline{(f(t) - S_n(t))} (f(t) - S_n(t)) = |f(t)|^2 + \overline{S_n(t)} S_n(t) - \overline{f(t)} S_n(t) - \overline{S_n(t)} f(t).$$

Nun ist $\overline{S_n(t)} S_n(t) = \sum_{\ell, k=-n}^n \overline{a_k(f)} a_\ell(f) \exp(-ikt) \exp(i\ell t)$. Also ergibt sich aus dem 4. Beispiel in (6.2) $\frac{1}{2\pi} \int_0^{2\pi} \overline{S_n(t)} S_n(t) dt = \sum_{k=-n}^n |a_k(f)|^2$, weil das Integral über gemischte Terme 0 ist. Es ist $\frac{1}{2\pi} \int_0^{2\pi} \overline{S_n(t)} f(t) dt = \sum_{k=-n}^n |a_k(f)|^2 = \frac{1}{2\pi} \int_0^{2\pi} S_n(t) \overline{f(t)} dt$.

Insgesamt erhält man

$$\begin{aligned} 0 &\leq \frac{1}{2\pi} \int_0^{2\pi} |f(t) - S_n(t)|^2 dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} (\overline{f(t)} - \overline{S_n(t)}) (f(t) - S_n(t)) dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt - 2 \sum_{k=-n}^n |a_k(f)|^2 + \sum_{k=-n}^n |a_k(f)|^2 \end{aligned}$$

und hieraus folgt sofort die Behauptung. Hieraus folgt:

Satz 6.10 Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig differenzierbar und 2π -periodisch. Dann konvergiert die Fourier-Reihe $\sum_{k=-\infty}^{\infty} a_k(f) e^{ikt}$ gleichmäßig und absolut gegen $f(t)$.

Beweis: Natürlich ist auch f' 2π -periodisch und außerdem stetig nach Voraussetzung.

Wir erhalten durch partielle Integration

$$a_k(f') = \frac{1}{2\pi} \int_0^{2\pi} f'(t) e^{-ikt} dt = \frac{1}{2\pi} [f(t) e^{-ikt}]_0^{2\pi} + \frac{ik}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt = ika_k(f).$$

Daraus folgt für $k \neq 0$

$$|a_k(f)| = \frac{1}{k} \cdot |a_k(f')| \leq \frac{1}{k^2} + |a_k(f')|^2$$

(es ist immer $|uv| \leq |u|^2 + |v|^2$ wegen $(|u|^2 + |v|^2 - 2|uv|) = (|u| - |v|)^2 \geq 0$).

Hiermit ergibt sich nach der Besselschen Ungleichung

$$\begin{aligned}
\left| \sum_{k=-n}^n a_k(f) e^{ikt} \right| &\leq \sum_{k=-n}^n |a_k(f)| \cdot 1 \\
&\leq \sum_{k=-n}^n |a_k(f)|^2 + \sum_{k \neq 0} \frac{1}{k^2} \\
&\leq \int_0^{2\pi} |f'(t)|^2 dt + \sum_{k \neq 0} \frac{1}{k^2} \\
&\leq M < \infty,
\end{aligned}$$

weil die Reihe $\sum_{k \neq 0} \frac{1}{k^2}$ konvergiert. Da n ganz links beliebig ist, folgt die absolute Konvergenz der Reihe und damit die Behauptung nach Korollar 6.8.

Beispiele:

$$1. f(t) = 1 \text{ für alle } t \Rightarrow a_k(f) = \begin{cases} 1 & k = 0 \\ 0 & \text{sonst} \end{cases}$$

$$2. f(t) = \begin{cases} t & 0 \leq t \leq \pi \\ 2\pi - t & \pi \leq t \leq 2\pi \end{cases}$$

und periodisch fortgesetzt durch $f(t + 2k\pi) = f(t)$. Es entsteht eine Sägezahnkurve. Eine Stammfunktion zu te^{-ikt} erhält man durch partielle Integration

$$\int te^{-ikt} = \left[-\frac{te^{-ikt}}{ik} \right] + \frac{1}{ik} \int e^{-ikt} dt.$$

Einsetzen liefert dann die Koeffizienten.

Aufgabe: Berechnen Sie bitte die Fourierkoeffizienten für das 2. Beispiel.

Es gibt stetige Funktionen, deren Fourierreihe nicht überall konvergiert, erst recht nicht gegen die jeweilige Funktion. Man kann jedoch den folgenden Abstand zwischen 2π -periodischen Funktionen einführen:

$$\text{Abstand}_2(f, g) = \left(\frac{1}{2\pi} \int_0^{2\pi} |f(t) - g(t)|^2 dt \right)^{1/2}.$$

Das bedeutet die "mittlere quadratische Abweichung der Funktionswerte". Dafür gilt

Theorem 6.11 Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig und 2π -periodisch und $f_n(t) = \sum_{-n}^n a_k(f) e^{ikt}$. Dann gilt:

$$\lim_{n \rightarrow \infty} \text{Abstand}_2(f - f_n) = 0$$

Dieser tiefer liegende Satz kann mit den uns zur Verfügung stehenden Mitteln nicht bewiesen werden

Kapitel 7

Geometrie und lineare Algebra

Überblick: Wir behandeln einfache geometrische Sachverhalte in der Ebene und im Raum. Grundlegend dabei ist der Begriff des Vektors, das ist ein Objekt, das durch Länge und Richtung charakterisiert ist. Wir verallgemeinern die gewonnenen Erkenntnisse auf Räume höherer Dimension und behandeln lineare Gleichungssysteme und Eigenwertprobleme.

7.1 Geometrie in Vektorräumen

Überblick: Wir führen Vektoren in der Ebene und im Raum ein. Wir erklären Geraden in der Ebene, und Geraden und Ebenen im Raum. Wir verallgemeinern unsere Erkenntnisse auf abstrakte Vektorräume und führen die fundamentalen Begriffe der linearen Unabhängigkeit und der Basis ein.

Wir definieren das Skalarprodukt, mit dem man Winkelmessung, Rechtwinkligkeit (Orthogonalität) und Länge eines Vektors (in Verallgemeinerung des Satzes von Pythagoras) erklären kann. Im \mathbb{R}^3 gibt es noch das Vektorprodukt und das Spatprodukt, die hilfreich bei geometrischen Überlegungen im Raum sind.

7.1.1 Vektorräume

Überblick: Wir beginnen mit dem physikalisch anschaulichen fundamentalen Begriff des Vektors in der Ebene und im Raum. In Verallgemeinerung von Ebene und Raum führen wir den Begriff eines Vektorraumes ein. Wir behandeln insbesondere Geraden und Ebenen als erste Beispiele.

Ein **Vektor** ist anschaulich gesprochen ein Objekt, das durch *Richtung und Länge* beschrieben wird. Beispiele sind Kraft, elektrische Feldstärke, Geschwindigkeit, Beschleunigung, usw.

Legen wir zwei Vektoren *verschiedener* Richtungen, die nicht entgegengesetzt sind, fest, so bestimmen diese ein Koordinatensystem der von ihnen aufgespannten **Ebene**. Wir können jeden Punkt der Ebene nach Festlegung eines Ursprungspunktes durch zwei Koordinaten x, y beschreiben, die wir zu einer Spalte $\begin{pmatrix} x \\ y \end{pmatrix}$ zusammenfassen. Den Punkt findet man, indem man vom Ursprung aus x Einheiten des ersten Vektors in seine Richtung und y Einheiten des zweiten Vektors in dessen Richtung geht.

Wir können also eine Ebene durch $\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\}$ beschreiben.

Anleitung zum Applet "Koordinatentransformation" Schauen Sie sich dies am Computer an. Die frei wählbaren Vektoren a und b definieren zusammen mit dem (frei wählbaren) Ursprung u das

Koordinatensystem. Durch den schwarzen Punkt werden die Streckfaktoren x, y festgelegt. Der rote Punkt ist dann der entsprechende Punkt im gewählten Koordinatensystem.

Zwei beliebige **Vektoren** der Ebene werden **addiert**, indem man den zweiten an die Spitze des ersten legt. In Koordinaten sieht das so aus:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}.$$

Ein Vektor wird mit dem Faktor λ gestreckt (gestaucht, das ist Geschmackssache), wenn man die Richtung beibehält, aber die Länge mit λ multipliziert. In Koordinaten sieht das so aus:

$$\lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}.$$

Anleitung zum Applet "Vektoren in der Ebene" Schauen Sie sich die Vektoraddition und Multiplikation mit Skalaren an. Wählen Sie zum Beispiel $x^\downarrow = (1, 2)^t$, $y^\downarrow = (14, 3)^t$ (t bedeutet: man meint den Spaltenvektor statt der Zeile). Wählen Sie x^\downarrow und den Skalar $\lambda = 2.5$.

Damit haben wir in \mathbb{R}^2 eine Addition und Multiplikation mit Skalaren (das sind reelle Zahlen) erklärt, die die folgenden Eigenschaften haben. Wir schreiben dabei x^\downarrow für $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Satz 7.1 Sei $\mathbb{R}^2 = V$ gesetzt. Dann gilt:

- a) Die Addition ist assoziativ, das heißt es gilt $(x^\downarrow + y^\downarrow) + z^\downarrow = x^\downarrow + (y^\downarrow + z^\downarrow)$
- b) Es gibt einen Nullvektor $0^\downarrow (= \begin{pmatrix} 0 \\ 0 \end{pmatrix})$ mit $x^\downarrow + 0^\downarrow = 0^\downarrow + x^\downarrow = x^\downarrow$ für alle x^\downarrow
- c) Zu jedem x^\downarrow gibt es ein y^\downarrow (nämlich $\begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$) mit $x^\downarrow + y^\downarrow = 0$. Wir setzen $y^\downarrow = -x^\downarrow$.
- d) Die Addition ist kommutativ, das heißt es gilt $x^\downarrow + y^\downarrow = y^\downarrow + x^\downarrow$
- e) Für jedes $\lambda \in \mathbb{R}$ und $x^\downarrow \in V$ ist λx^\downarrow wieder aus V und es gilt

$$\begin{aligned} (\lambda + \mu)x^\downarrow &= \lambda x^\downarrow + \mu x^\downarrow \\ \lambda(x^\downarrow + y^\downarrow) &= \lambda x^\downarrow + \lambda y^\downarrow \\ (\lambda\mu)x^\downarrow &= \lambda(\mu x^\downarrow) \\ 1 \cdot x^\downarrow &= x^\downarrow \end{aligned}$$

Wir sparen uns den Beweis, den man durch einfaches Nachrechnen erhält.

Sehen wir von der speziellen Wahl $V = \mathbb{R}^2$ ab und betrachten nun abstrakt eine Menge V , auf der eine Addition $(x, y) \mapsto x + y \in V$ und eine Multiplikation mit Skalaren $(\lambda, x) \mapsto \lambda x \in V$ gegeben ist.

Definition 7.2 (Vektorraum über \mathbb{R}) Sei V eine Menge, auf der eine Addition

$$V \times V \rightarrow V, (x, y) \mapsto x + y$$

und eine Multiplikation mit Skalaren

$$\mathbb{R} \times V \rightarrow V, (\lambda, x) \mapsto \lambda x$$

gegeben sind, die die Eigenschaften a) bis d) von Satz 6.1 haben. Dann heißt V (abstrakter) **Vektorraum über \mathbb{R}** .

Beispiele:

1. Eine **Gerade** G in der Ebene durch den Punkt P in Richtung $x^\downarrow \neq 0$ erhält man (nach Einführung von Koordinaten) durch $G = \{p^\downarrow + \lambda x^\downarrow : \lambda \in \mathbb{R}\}$. Hierbei ist p^\downarrow der Koordinatenvektor von P . Seien p^\downarrow und q^\downarrow zwei feste Punkte. Man erhält die Gerade durch p^\downarrow und q^\downarrow als $G = \{p^\downarrow + t(q^\downarrow - p^\downarrow) : t \in \mathbb{R}\}$. Ist $p^\downarrow = 0^\downarrow$, so ist G selbst ein Vektorraum. (für $p \neq 0$ nicht, warum?) Prüfen Sie das bitte nach!

2. $V = \mathbb{R}^p = \{x : x = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, x_j \in \mathbb{R}\}$. Die Addition ist koordinatenweise erklärt:

$$x + y = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_p + y_p \end{pmatrix}. \text{ Ebenso ist } \lambda x \text{ durch } \lambda x := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_p \end{pmatrix} \text{ erklärt.}$$

Offensichtlich ist V damit ein Vektorraum, der **arithmetische Vektorraum der Dimension p über \mathbb{R}** .

3. $V = C([a, b]) = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ ist stetig}\}$. Die Addition ist wie üblich punktweise erklärt: $(f + g)(t) = f(t) + g(t)$. Ebenso setzt man $(\lambda f)(t) = \lambda(f(t))$. Damit ist V ein Vektorraum über \mathbb{R} .
4. Sei $V = \mathbb{R}^p$, ferner $0 \neq x \in V$ und $G = \{tx : t \in \mathbb{R}\}$. Dann ist G bezüglich der aus V geerbten Addition und Multiplikation mit Skalaren selbst ein Vektorraum; denn es gilt für $y, z \in G$, $y = sx$, $z = tx$ eben $y + z = sx + tx = (s + t)x \in G$ und $\lambda y = \lambda(sx) = (\lambda s)x \in G$. Die übrigen der Eigenschaften a) bis d) erbt G von V . G ist eine **Gerade durch den Nullpunkt** oder ein **eindimensionaler Untervektorraum**. Zum Beispiel ist $D = \{t \begin{pmatrix} 1 \\ 1 \end{pmatrix} : t \in \mathbb{R}\}$ die Hauptdiagonale in der Ebene \mathbb{R}^2 , also ein eindimensionaler Untervektorraum der Ebene.
5. Im Raum gibt es drei verschiedene unabhängige Richtungen. Also können wir ihn durch \mathbb{R}^3 ähnlich wie die Ebene durch \mathbb{R}^2 koordinatisieren. Im Raum gibt es viele Ebenen, allein durch den Nullpunkt bereits. Seien nämlich e^\downarrow und f^\downarrow zwei Vektoren in \mathbb{R}^3 , die auf *verschiedenen* Geraden durch den Nullpunkt liegen. Dann ist $E = \{se^\downarrow + tf^\downarrow : s, t \in \mathbb{R}\}$ eine **Ebene durch den Nullpunkt**, aufgespannt von e^\downarrow und f^\downarrow . E ist mit der von \mathbb{R}^3 ererbten Addition und Multiplikation ein Vektorraum der Dimension 2, ein **zweidimensionaler Untervektorraum**. Eine Ebene durch den Punkt P mit Koordinatenvektor p^\downarrow erhalten wir hieraus durch Parallelverschiebung:

$$E_p = \{p^\downarrow + se^\downarrow + tf^\downarrow : s, t \in \mathbb{R}\}.$$

Anleitung zum Applet "Geraden und Ebenen im Raum" Schauen Sie sich Geraden und Ebenen im Raum an. Hier sind Beispiele:

1. $G = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} : t, s \in \mathbb{R} \right\}$.
2. $E = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} : t, s \in \mathbb{R} \right\}$.

3. Seien $p^\downarrow = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $q^\downarrow = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $r^\downarrow = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

Lassen Sie erst diese drei Vektoren zeichnen (ausgehend vom Nullpunkt).
Die Ebene durch diese drei Punkte ist

$$E = \{p^\downarrow + s(q^\downarrow - p^\downarrow) + t(r^\downarrow - p^\downarrow) : s, t \in \mathbb{R}^2\}.$$

Berechnen Sie die Vektoren. Für welches Paar (s, t) erhält man p^\downarrow , für welches q^\downarrow und für welches r^\downarrow ? Lassen Sie sich die Ebene zeichnen.

Ein Vorteil der Vektorrechnung zeigt sich im folgenden Beispiel:

Beispiel: *Wir beweisen kurz, daß sich die Diagonalen eines Parallelogramms stets halbieren.*

Seien e^\downarrow und f^\downarrow die Grundkanten eines Parallelogramms.

Dann ist die Hauptdiagonale $h = \{t(e^\downarrow + f^\downarrow) : t \in \mathbb{R}\}$,

die Nebendiagonale $N = \{e^\downarrow + s(f^\downarrow - e^\downarrow) : s \in \mathbb{R}\}$.

Für einen Schnittpunkt s^\downarrow muß gelten

$$s^\downarrow = t(e^\downarrow + f^\downarrow) = e^\downarrow + s(f^\downarrow - e^\downarrow)$$

oder

$$(1 - t - s)e^\downarrow = (t - s)f^\downarrow.$$

Die Geraden durch e^\downarrow und f^\downarrow haben als einzigen Punkt den Nullpunkt gemeinsam. Also ist $1 - t - s = 0 = t - s$ und damit $t = s = \frac{1}{2}$.

Also gibt es genau einen Schnittpunkt

$$s^\downarrow = \frac{1}{2}(e^\downarrow + f^\downarrow) = e^\downarrow + \frac{1}{2}(f^\downarrow - e^\downarrow).$$

Wiederholung:

- Ebene und Raum und ihre Koordinatisierung.
- Vektorräume, Axiomensystem.
- \mathbb{R}^p , Untervektorräume, Geraden, Ebenen.

7.1.2 Lineare Abhängigkeit

Überblick: In Verallgemeinerung der Parallelität zweier Vektoren führen wir den Begriff der linearen Abhängigkeit ein. Lineare Unabhängigkeit ist die Verallgemeinerung der grundlegenden Auszeichnung von endlich vielen Vektoren (besser: Richtungen), nach denen man alle Vektoren eindeutig zerlegen kann. Damit können wir in Verallgemeinerung von Ebene und Raum Basen und die Dimension eines Vektorraumes einführen.

Wir sagen, daß 2 Vektoren a^\downarrow und b^\downarrow im Vektorraum **parallel** sind, wenn sie in die gleiche oder entgegengesetzte Richtung zeigen, wenn also $a^\downarrow = tb^\downarrow$ für ein $t \in \mathbb{R}$ ist, oder kürzer, wenn a^\downarrow auf der durch den Nullpunkt in Richtung b^\downarrow laufenden Geraden liegt. Dann liegt auch b^\downarrow auf der durch den Nullpunkt in Richtung a^\downarrow laufenden Geraden, denn es ist ja (vorausgesetzt, daß $a^\downarrow \neq 0^\downarrow \neq b^\downarrow$ gilt) $b^\downarrow = \frac{1}{t}a^\downarrow$. Damit

wir auch den Fall $a^\perp = 0^\perp$ bzw. $b^\perp = 0^\perp$ erfassen - der Nullvektor ist sozusagen parallel zu jedem anderen Vektor - können wir sagen:

a^\perp und b^\perp sind parallel, wenn es zwei Zahlen $s, t \in \mathbb{R}$ gibt, die nicht beide gleich 0 sind mit $sa^\perp + tb^\perp = 0$. Man nennt die beiden Vektoren dann auch **linear abhängig**.

Wir verallgemeinern:

Definition 7.3 Sei V ein Vektorraum über \mathbb{R} . Die Vektoren $a_1, \dots, a_n \in V$ heißen **linear abhängig**, wenn es Zahlen t_1, \dots, t_n gibt, die **nicht alle gleich 0** sind, so daß $t_1 a_1 + \dots + t_n a_n = 0$ gilt. Gibt es solche Zahlen nicht, so heißen die Vektoren **linear unabhängig**.

a_1, \dots, a_n sind also linear unabhängig, wenn die Gleichung $t_1 a_1 + \dots + t_n a_n = 0$ nur für $t_1 = t_2 = \dots = t_n = 0$ gilt. a_1, \dots, a_n sind dagegen linear abhängig, wenn $t_1 a_1 + \dots + t_n a_n = 0$ für Zahlen t_1, \dots, t_n gilt, von denen mindestens eine ungleich Null ist. Ist etwa $t_n \neq 0$, so ist $a_n = -\frac{t_1}{t_n} a_1 - \frac{t_2}{t_n} a_2 - \dots - \frac{t_{n-1}}{t_n} a_{n-1}$, a_n liegt also in dem Untervektorraum

$$W = \{s_1 a_1 + \dots + s_{n-1} a_{n-1} : s_j \in \mathbb{R} \text{ für } j = 1, \dots, n-1\}.$$

(Siehe hierzu auch Definition 7.4)

Beispiele:

1. In \mathbb{R}^3 sind $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ linear abhängig.

Hingegen sind $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ linear unabhängig.

2. In \mathbb{R}^2 sind $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ linear unabhängig, aber je drei Vektoren sind linear abhängig.

3. In \mathbb{R}^p gibt es p linear unabhängige Vektoren. Je $(p+1)$ Vektoren sind linear abhängig.

4. In $C([0, 1])$ sind die Funktionen $1, x, x^2, \dots, x^{20}$ linear unabhängig. Gilt das auch für beliebiges n statt 20?

5. Seien $L = \{a_1, \dots, a_n\} \subset V$ linear unabhängig und $\emptyset \neq M \subset \{a_1, \dots, a_n\}$ eine Teilmenge, also $M = \{a_{i_1}, \dots, a_{i_r}\}$ mit $r \leq n$ und $a_{i_j} \neq a_{i_k}$ für $j \neq k$. Dann ist auch M linear unabhängig. Denn wäre $0 = t_{i_1} a_{i_1} + \dots + t_{i_r} a_{i_r}$ mit einem $t_{i_j} \neq 0$, so wäre

$$0 = t_{i_1} a_{i_1} + \dots + t_{i_r} a_{i_r} + 0 \cdot a_{\ell_1} + \dots + 0 \cdot a_{\ell_{n-r}}$$

(wo $\{a_{\ell_1}, \dots, a_{\ell_{n-r}}\} = L \setminus M$) mit mindestens einem $t_{i_j} \neq 0$, also wäre L linear abhängig.

6. Sei $L = \{a_1, \dots, a_n\} \subset V$ linear abhängig und N irgend eine endliche Menge aus V , die L enthält. Dann ist N linear abhängig.

Wir haben bereits gesehen, daß es in \mathbb{R}^3 , also auch im geometrischen bzw. euklidischen Raum drei linear unabhängige Vektoren gibt (z.B. $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$) und sich jeder Vektor aus

diesen
 "linear kombinieren" läßt.

$$x^\downarrow = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 e_1^\downarrow + x_2 e_2^\downarrow + x_3 e_3^\downarrow.$$

Definition 7.4 Seien a_1, \dots, a_p Vektoren in einem beliebigen Vektorraum und $\lambda_1, \dots, \lambda_p \in \mathbb{R}$. Dann nennt man die Summe $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_p a_p = \sum_{j=1}^p \lambda_j a_j$ der skalaren Vielfachen $\lambda_j a_j$ der Vektoren a_j eine **Linearkombination** von a_1, \dots, a_p . Eine Teilmenge W des Vektorraumes V heißt **Untervektorraum** oder **linearer Teilraum** von V , wenn mit $a, b \in W$ auch jede Linearkombination $\lambda a + \mu b$ in W liegt. W ist dann tatsächlich ein Vektorraum. Denn daß die Addition assoziativ ist usw. erbt W ja von V .

Sei $a_1, \dots, a_p \in V$. Dann ist die Menge $\text{span}(a_1, \dots, a_p)$ aller Linearkombinationen von a_1, \dots, a_p ein Untervektorraum, der von a_1, \dots, a_p **erzeugte Untervektorraum**. Beispiele sind Geraden durch den Nullpunkt ($p = 1, 0 \neq a_1$) und Ebenen durch den Nullpunkt ($p = 2, a_1, a_2$ linear unabhängig).

Definition 7.5 Eine Menge $B = \{b_1, \dots, b_p\}$ heißt **Basis** des Vektorraumes V , wenn $V = \text{span}(B)$ und B linear unabhängig ist.

Wir erhalten

Satz 7.6 a) $B = \{b_1, \dots, b_p\} \subset V$ ist genau dann Basis von V , wenn sich jedes x auf genau eine Weise als Linearkombination $x = \sum \xi_j b_j$ der b_1, \dots, b_p schreiben läßt.

b) Hat V eine Basis B aus p Elementen, so hat jede Basis von V p Elemente.

Beweis:

a) Sei B Basis und $x \in V$ beliebig. Wegen $V = \text{span}(B)$ ist $x = \sum \xi_j b_j$ für gewisse $\xi_j \in \mathbb{R}$. Gilt außerdem $x = \sum \mu_j b_j$, so ist $0 = \sum (\xi_j - \mu_j) b_j$. Da B linear unabhängig ist, ist $0 = \xi_j - \mu_j$, also $\xi_j = \mu_j$ für $j = 1, \dots, p$. Die x darstellende Linearkombination ist eindeutig. Sei umgekehrt jedes $x \in V$ eindeutige Linearkombination von B . Dann ist $V = \text{span}(B)$. Außerdem ist $x = 0$ *eindeutige* Linearkombination von B . Ist also $0 = \sum \xi_j b_j$, so muß $\xi_1 = \dots = \xi_p = 0$ gelten. Also ist B linear unabhängig

b) ist sehr viel aufwendiger zu beweisen.

Definition 7.7 Ein Vektorraum V über \mathbb{R} heißt **unendlichdimensional**, in Zeichen $\dim(V) = \infty$, wenn er keine endliche Basis hat. Im anderen Fall haben alle möglichen Basen die gleiche Mächtigkeit. Diese Zahl heißt die **Dimension** $\dim(V)$ von V .

Beispiele:

- $\dim(\mathbb{R}^p) = p$.
- Sei $G = \{tb^\downarrow : t \in \mathbb{R}\}$, $b^\downarrow \neq 0^\downarrow$ eine Gerade durch den Nullpunkt in \mathbb{R}^p . Dann ist G ja ein Untervektorraum. Es gilt $\dim(G) = 1$. Alle Basen sind eindeutig und haben die Form $B = \{t_\circ b^\downarrow\}$, wo t_\circ eine feste Zahl ist.

3. Sei $E = \{sa^\perp + tb^\perp : s, t \in \mathbb{R}\}$ (mit a^\perp, b^\perp linear unabhängig) eine Ebene durch den Nullpunkt in \mathbb{R}^p . Dann ist $\{a^\perp, b^\perp\}$ eine Basis des Untervektorraumes E . Es gilt also $\dim E = 2$. Sind $\begin{pmatrix} u \\ v \end{pmatrix}$ und $\begin{pmatrix} s \\ t \end{pmatrix}$ linear unabhängig in \mathbb{R}^2 , so ist $B = \{ua^\perp + vb^\perp, sa^\perp + tb^\perp\}$ eine weitere Basis in E .

Wiederholung:

- Lineare Abhängigkeit, lineare Unabhängigkeit
- Untervektorraum
- Basis, Dimension eines Vektorraumes

7.1.3 Länge oder Norm eines Vektors, Skalarprodukt auf \mathbb{R}^p

Überblick: Ausgehend von der Anschauung in der Ebene definieren wir die Länge eines Vektors (seine Norm) und den Abstand zweier Vektoren. Dies führt zur Einführung des Skalarprodukts, mit dessen Verallgemeinerung wir nun auch eine Norm in \mathbb{R}^p einführen können.

Betrachten wir eine Ebene, die physikalisch gesprochen von zwei Vektoren der Länge 1 aufgespannt werden, die senkrecht aufeinander stehen, und koordinatisieren wir die Ebene durch \mathbb{R}^2 wie oben beschrieben, so hat ein Vektor $x^\perp = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ die Länge $\|x^\perp\| = \sqrt{x_1^2 + x_2^2}$. Zwei Punkte P und Q mit den Koordinaten p^\perp und q^\perp haben dann den Abstand $A(P, Q) = \|p^\perp - q^\perp\|$.

Die Länge, auch **Norm** genannt, hat die folgenden Eigenschaften:

Satz 7.8 Es gilt

- | | |
|---|-----------------------------|
| (i) $\ x^\perp\ = 0$ genau dann wenn $x^\perp = 0$ ist. | <i>Definitheit</i> |
| (ii) $\ tx^\perp\ = t \ x^\perp\ $ | <i>Absolute Homogenität</i> |
| (iii) $\ x^\perp + y^\perp\ \leq \ x^\perp\ + \ y^\perp\ $ | <i>Dreiecksungleichung</i> |
| (iv) $ \ x^\perp\ - \ y^\perp\ \leq \ x^\perp - y^\perp\ $ | |

Während man die ersten beiden Gleichungen sowohl aus der Anschauung als auch durch einfache Rechnung erhält, ergibt sich (iii) zunächst nur aus der Anschauung. Ihr rechnerischer Beweis ist dagegen komplizierter, man zeigt

$$\|x^\perp + y^\perp\|^2 \leq \|x^\perp\|^2 + \|y^\perp\|^2 + 2\|x^\perp\| \|y^\perp\|.$$

(siehe Theorem 7.12)

Das Skalarprodukt

Zeichnet man sich ein Dreieck mit den Seiten x^\perp, y^\perp und $x^\perp + y^\perp$ auf und projiziert $x^\perp + y^\perp$ senkrecht auf die durch x^\perp laufende Gerade, so erhält man nach dem Satz des Pythagoras

$$\|x^\perp + y^\perp\|^2 = (\|x^\perp\| + \|y^\perp\| \cos \varphi)^2 + (\|y^\perp\| \sin \varphi)^2 = \|x^\perp\|^2 + \|y^\perp\|^2 + 2\|x^\perp\| \|y^\perp\| \cos \varphi.$$

(φ ist hier der Winkel zwischen x^\perp und y^\perp)

Rechnet man andererseits $\|x^\perp + y^\perp\|^2$ einfach aus, so ergibt sich

$$\|x^\perp + y^\perp\|^2 = (x_1 + y_1)^2 + (x_2 + y_2)^2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 + 2x_1y_1 + 2x_2y_2.$$

Hieraus erhält man durch Vergleich mit der vorigen Gleichung

$$x_1y_1 + x_2y_2 = \|x\| \|y\| \cos \varphi.$$

Wir nennen diesen Ausdruck das **Skalarprodukt** von x^\perp mit y^\perp .
Allgemeiner erklären wir auch das Skalarprodukt auf \mathbb{R}^p :

Definition 7.9 Seien $x^\perp = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ und $y^\perp = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$ in \mathbb{R}^p beliebig. Wir definieren das **Skalarprodukt** durch

$$(x^\perp | y^\perp) = x_1y_1 + x_2y_2 + \dots + x_py_p = \sum_{j=1}^p x_jy_j \dots \in \mathbb{R}.$$

Es hat folgende Eigenschaften:

Satz 7.10 Es gelten die Formeln:

- | | |
|---|--|
| (i) $(x^\perp + y^\perp z^\perp) = (x^\perp z^\perp) + (y^\perp z^\perp)$ | <i>Linearität im
ersten Argument</i> |
| $(\lambda x^\perp z^\perp) = \lambda(x^\perp z^\perp)$ | |
| (ii) $(x^\perp y^\perp) = (y^\perp x^\perp)$ | <i>Symmetrie</i> |
| (iii) $(x^\perp x^\perp) > 0$ für alle $x^\perp \neq 0$ | <i>Definitheit</i> |

Sie erhalten all diese Eigenschaften durch Ausrechnen. Wegen (ii) ist das Skalarprodukt auch linear im zweiten Argument, es gilt also $(x^\perp | y^\perp + z^\perp) = (x^\perp | y^\perp) + (x^\perp | z^\perp)$ und $(x^\perp | \lambda y^\perp) = \lambda(x^\perp | y^\perp)$.

Aber *beachten Sie!* Das Skalarprodukt ist *kein Produkt im üblichen Sinn*, weil es seinen zwei Argumenten, die Vektoren sind, eine *Zahl* und keinen Vektor zuordnet.

Das Skalarprodukt erfüllt eine wichtige Ungleichung, die im Fall der Ebene wegen $|\cos \varphi| \leq 1$ klar ist (s.o.)

Satz 7.11 (Cauchy-Schwarzsche Ungleichung)

Es gilt

$$|(x^\perp | y^\perp)|^2 \leq (x^\perp | x^\perp)(y^\perp | y^\perp).$$

Das Gleichheitszeichen gilt genau dann, wenn x^\perp und y^\perp linear abhängig sind.

Beweis: Ist $y^\perp = 0$, so steht links und rechts 0. (Dann sind x^\perp und y^\perp auch linear abhängig).

Sei $y^\perp \neq 0$ und $t = \frac{(x^\perp | y^\perp)}{(y^\perp | y^\perp)}$.

Nach 7.10 gilt

$$\begin{aligned}
0 \leq (x^\perp - ty^\perp | x^\perp - ty^\perp) &= (x^\perp | x^\perp) - 2t(x^\perp | y^\perp) + t^2(y^\perp | y^\perp) \\
&= (x^\perp | x^\perp) - \frac{2(x^\perp | y^\perp)^2}{(y^\perp | y^\perp)} + \frac{(x^\perp | y^\perp)^2}{(y^\perp | y^\perp)} \\
&= (x^\perp | x^\perp) - \frac{(x^\perp | y^\perp)^2}{(y^\perp | y^\perp)}
\end{aligned}$$

Multipliziert man diese Ungleichung mit $(y^\perp | y^\perp)$ durch, so folgt die erste Aussage. Nach 7.10 gilt $0 = (x^\perp - ty^\perp | x^\perp - ty^\perp)$ genau dann, wenn $x^\perp - ty^\perp = 0$, x^\perp und y^\perp also linear abhängig sind.

Motiviert durch den Fall der Ebene sagen wir, **zwei Vektoren x^\perp und y^\perp stehen senkrecht aufeinander**, oder sind **orthogonal**, wenn $(x^\perp | y^\perp) = 0$ gilt. Genau wie in der Ebene erhalten wir auch im \mathbb{R}^p eine Abstandsmessung aus dem Skalarprodukt:

Theorem 7.12 Für $x^\perp \in \mathbb{R}^p$ sei die Norm erklärt durch

$$\|x^\perp\| = \sqrt{(x^\perp | x^\perp)}.$$

Dann gilt

- | | |
|---|----------------------|
| (i) $\ x^\perp\ = 0$ genau dann wenn $x^\perp = 0$ ist. | Definitheit |
| (ii) $\ tx^\perp\ = t \ x^\perp\ $ | Absolute Homogenität |
| (iii) $\ x^\perp + y^\perp\ \leq \ x^\perp\ + \ y^\perp\ $ | Dreiecksungleichung |
| (iv) $ \ x^\perp\ - \ y^\perp\ \leq \ x^\perp - y^\perp\ $ | |

Beweis: (i) und (ii) sind unmittelbar klar.

(iii) Es ist

$$\begin{aligned}
\|x^\perp + y^\perp\|^2 &= (x^\perp + y^\perp | x^\perp + y^\perp) \\
&= (x^\perp | x^\perp) + (y^\perp | y^\perp) + 2(x^\perp | y^\perp) \\
&\leq \underbrace{(x^\perp | x^\perp) + (y^\perp | y^\perp)}_{\text{Satz 7.11}} + 2\sqrt{(x^\perp | x^\perp)(y^\perp | y^\perp)} \\
&= (\|x^\perp\| + \|y^\perp\|)^2,
\end{aligned}$$

woraus (iii) durch Wurzelziehen folgt.

(iv) Es ist nach (iii)

$$\|x^\perp\| = \|(x^\perp - y^\perp) + y^\perp\| \leq \|x^\perp - y^\perp\| + \|y^\perp\|,$$

also

$$\|x^\perp\| - \|y^\perp\| \leq \|x^\perp - y^\perp\|.$$

Vertauscht man x^\perp und y^\perp so erhält man

$$\|y^\perp\| - \|x^\perp\| \leq \|y^\perp - x^\perp\|.$$

Wegen $\|y^\perp - x^\perp\| = \|x^\perp - y^\perp\|$ (nach (ii)) folgt die Behauptung.

Beispiele:

1. In \mathbb{R}^2 sind $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ senkrecht aufeinander.
2. Sei $x^\perp = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. Dann ist die Gerade senkrecht zu x^\perp durch den Nullpunkt gerade $G = \{t \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} : t \in \mathbb{R}\}$. Die Gerade G läßt sich also auch durch die Gleichung $(y^\perp | x^\perp) = y_1 x_1 + y_2 x_2 = 0$ beschreiben.

Wiederholung:

- Skalarprodukt in \mathbb{R}^p
- Cauchy-Schwarzschen Ungleichung
- Norm eines Vektors, Abstand zweier Vektoren
- Orthogonalität von Vektoren

7.1.4 Vektorprodukt und Spatprodukt im \mathbb{R}^3

Überblick: Im Raum (also im \mathbb{R}^3) stehen uns zwei weitere "Produkte" von geometrischer Bedeutung zur Verfügung, die wir erfolgreich bei geometrischen Untersuchungen einsetzen können: das Vektorprodukt und das Spatprodukt. Wir liefern geometrischen Anwendungen.

Als erste Aufgabe suchen wir im Raum zu zwei Vektoren x^\perp und y^\perp einen Vektor, der senkrecht auf der durch x^\perp und y^\perp aufgespannten Ebene $E = \{s x^\perp + t y^\perp : s, t \in \mathbb{R}\}$ steht und außerdem auch den Flächeninhalt des von x^\perp und y^\perp aufgespannten Parallelogramms $P(x^\perp, y^\perp) = \{s x^\perp + t y^\perp : 0 \leq s, t \leq 1\}$ charakterisiert.

Definition 7.13 Das **Vektorprodukt** $x^\perp \times y^\perp$ zweier Vektoren x^\perp, y^\perp ist ein Vektor mit den *Eigenschaften*

1. $x^\perp \times y^\perp$ steht senkrecht auf x^\perp und y^\perp
2. $\|x^\perp \times y^\perp\| = \|x^\perp\| \|y^\perp\| \sin(\angle(x^\perp, y^\perp))$, das ist der Flächeninhalt von $P(x^\perp, y^\perp)$
3. $x^\perp, y^\perp, x^\perp \times y^\perp$ soll in dieser Reihenfolge ein Rechtssystem bilden, das heißt: steht ein Mensch mit dem rechten Fuß auf x^\perp und mit dem linken Fuß auf y^\perp , so zeigt der Kopf in Richtung $x^\perp \times y^\perp$. Oder "physikalischer": halte die rechte Faust mit der durch den kleinen Finger gebildeten Kante auf die Ebene E , so daß die Fingerspitzen der Faust von x^\perp nach y^\perp zeigen. Dann zeigt der Daumen in Richtung $x^\perp \times y^\perp$.

Anleitung zum Applet "Der Matrizenrechner" Wählen Sie Paare von Vektoren und lassen Sie sich ihr Vektorprodukt berechnen.

Es gilt z.B.:

Beispiele:

1. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$$2. \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

$$3. \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

Das Vektorprodukt hat die folgenden Eigenschaften

Satz 7.14 1. $y^\perp \times x^\perp = -x^\perp \times y^\perp$

2. Genau dann sind x^\perp und y^\perp linear abhängig (parallel), wenn $x^\perp \times y^\perp = 0$

3. $(x^\perp + y^\perp) \times z^\perp = x^\perp \times z^\perp + y^\perp \times z^\perp$ *Linearität im 1. Argument*
 $(\lambda x^\perp) \times y^\perp$

Wegen 1) ist das Vektorprodukt auch linear im zweiten Argument.

4. Ist $x^\perp = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ und $y^\perp = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, so ist

$$x^\perp \times y^\perp = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

Beweis:

1. Durch die Vertauschung der Faktoren ändert sich der Flächeninhalt von $P(x^\perp, y^\perp)$, also der Betrag von $x^\perp \times y^\perp$ nicht, wohl aber die Richtung.
2. Der Flächeninhalt des Parallelogramms ist genau dann gleich 0, wenn x^\perp und y^\perp auf einer gemeinsamen Geraden liegen.
3. Die Formel mit λ ist klar. Die Formel mit “+” im ersten Argument ist zwar elementargeometrisch klar, aber etwas schwieriger.
4. erhält man aus 1) und 3) wie folgt:

$$x^\perp = x_1 e_1^\perp + x_2 e_2^\perp + x_3 e_3^\perp, \quad y^\perp = y_1 e_1^\perp + y_2 e_2^\perp + y_3 e_3^\perp$$

$$\text{mit } e_1^\perp = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2^\perp = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3^\perp = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dann gilt nämlich $e_1^\perp \times e_2^\perp = e_3^\perp$, $e_2^\perp \times e_3^\perp = e_1^\perp$ und $e_3^\perp \times e_1^\perp = e_2^\perp$, ferner $x^\perp \times y^\perp = \sum_{i,k=1}^3 x_i y_k e_i^\perp \times e_k^\perp$.

Berücksichtigung von $e_2^\perp \times e_1^\perp = -e_3^\perp$ usw. liefert die Behauptung.

Für das Vektorprodukt gilt kein Assoziativgesetz. Vielmehr hat man den folgenden Satz, der auch eine weitere Formel enthält.

Satz 7.15 1. $x^\perp \times (y^\perp \times z^\perp) = (x^\perp | z^\perp) y^\perp - (x^\perp | y^\perp) z^\perp$

$$2. (x^\perp \times y^\perp | u^\perp \times v^\perp) = (x^\perp | u^\perp)(y^\perp | v^\perp) - (x^\perp | v^\perp)(y^\perp | u^\perp)$$

Identität von Lagrange

Beweis:

1. $x^\perp \times (y^\perp \times z^\perp)$ steht insbesondere senkrecht auf $y^\perp \times z^\perp$, liegt also in der von y^\perp und z^\perp aufgespannten Ebene. Damit kann man die Formel durch den Ansatz $x^\perp \times (y^\perp \times z^\perp) = \lambda y^\perp + \mu z^\perp$ und weitere geometrische Überlegungen erhalten. Es muß dieser Vektor ja auch auf x^\perp senkrecht stehen, was $\lambda(x^\perp | y^\perp) + \mu(x^\perp | z^\perp) = 0$ ergibt. Man kann die Formel aber auch über die Koordinatenformel für das Vektorprodukt ausrechnen.
2. Man setzt $z^\perp = u^\perp \times v^\perp$ und erhält nach dem Vertauschungssatz (Satz 7.16)

$$\begin{aligned} (x^\perp \times y^\perp | z^\perp) &= (x^\perp | y^\perp \times z^\perp) = (x^\perp | y^\perp \times (u^\perp \times v^\perp)) \\ &= (x^\perp | (y^\perp | v^\perp)u^\perp - (y^\perp | u^\perp)v^\perp). \end{aligned}$$

Ausmultiplizieren liefert die Formel.

Der orientierte Rauminhalt des von den drei Vektoren $x^\perp, y^\perp, z^\perp$ im \mathbb{R}^3 aufgespannten Parallelepipeds oder Spats heißt **Spatprodukt**. Es ist Grundfläche mal Höhe. Die Formel ist

$$(x^\perp y^\perp z^\perp) = (x^\perp \times y^\perp | z^\perp)$$

Anleitung zum Applet “affine Transformationen” Schauen Sie sich das entsprechende Parallelepiped an:

$$a_1^\perp = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, a_2^\perp = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, a_3^\perp = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Schreiben Sie hierzu den ersten Vektor in die erste Spalte der “Matrix” auf der linken Seite (“Transformationen”), den zweiten in die zweite Spalte, den dritten in die dritte. Wählen Sie eigene Vektoren!

Für das Spatprodukt gilt der folgende Satz:

Satz 7.16 (Vertauschungssatz)

Es ist $(x^\perp y^\perp z^\perp) = (y^\perp z^\perp x^\perp) = (z^\perp x^\perp y^\perp)$.

Vertauscht man ein Argumentpaar, so wechselt das Spatprodukt das Vorzeichen.

Beweis: Es ist für die kanonischen Vektoren $e_1^\perp, e_2^\perp, e_3^\perp$ $(e_1^\perp e_2^\perp e_3^\perp) = 1$ (Inhalt des Würfels der Kantenlänge 1). Setzt man nun $x^\perp = x_1 e_1^\perp + x_2 e_2^\perp + x_3 e_3^\perp, y^\perp$ und z^\perp entsprechend, so folgt die Behauptung.

Falls Sie dreireihige Determinanten aus der Schule kennen, sehen Sie:

$$(x^\perp y^\perp z^\perp) = \det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix}.$$

Korollar 7.17 Es ist $(x^\perp y^\perp z^\perp)$ genau dann gleich 0 wenn $x^\perp, y^\perp, z^\perp$ linear abhängig sind.

Beweis: Nach der geometrischen Bedeutung ist $(x^\perp y^\perp z^\perp) = 0$ genau dann, wenn die drei Vektoren gar kein echtes Spat aufspannen, also in einer Ebene liegen.

Aufgaben:

1. Wie lang ist die Strecke zwischen $p^\perp = \begin{pmatrix} 2 \\ 5 \\ 11 \end{pmatrix}$ und $q^\perp = \begin{pmatrix} 10 \\ 3 \\ 18 \end{pmatrix}$?

2. Wo liegt der Punkt p_1^\perp der die Strecke p^\perp und q^\perp im Verhältnis 2 zu 3 teilt?

3. Bestimmen Sie den Schwerpunkt des Dreiecks mit den Spitzen

$$a_1^\perp = \begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix}, a_2^\perp = \begin{pmatrix} 11 \\ 3 \\ 8 \end{pmatrix}, a_3^\perp = \begin{pmatrix} 5 \\ 1 \\ 11 \end{pmatrix}.$$

4. Bestimmen Sie den Flächeninhalt des obigen Dreiecks.

Tip: Benutzen Sie das Vektorprodukt zweier Kantenvektoren.

5. Der Rauminhalt eines Tetraeders mit den Eckpunkten $p_1^\perp, p_2^\perp, p_3^\perp$ ist Grundfläche \times Höhe / 3. Bestimmen Sie den Rauminhalt für

$$p_1^\perp = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, p_2^\perp = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, p_3^\perp = \begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix}, p_4^\perp = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Tip: Verwenden Sie das Spatprodukt für die Kanten!

Geometrische Anwendungen

Wir hatten schon früher die Beschreibung einer Ebene durch drei Punkte $p_1^\perp, p_2^\perp, p_3^\perp$ behandelt:

$$E = \{p_1^\perp + s(p_2^\perp - p_1^\perp) + t(p_3^\perp - p_1^\perp) : s, t \in \mathbb{R}\}$$

Die Gleichung $z^\perp = p_1^\perp + s(p_2^\perp - p_1^\perp) + t(p_3^\perp - p_1^\perp) = z^\perp(s, t)$ heißt **Parameterdarstellung** der Ebene.

Wir setzen $e^\perp = \frac{(p_2^\perp - p_1^\perp) \times (p_3^\perp - p_1^\perp)}{\|(p_2^\perp - p_1^\perp) \times (p_3^\perp - p_1^\perp)\|}$. e^\perp heißt der **Stellungsvektor** oder **Normalenvektor der Ebene**.

Ebene. Nach Konstruktion steht e^\perp senkrecht auf $z^\perp(s, t) - p_1^\perp$ für alle s, t , also auf der in den Nullpunkt parallel verschobenen Ebene $E_0 = E - p_1^\perp$. Damit gilt $x^\perp \in E$ genau dann, wenn $0 = (e^\perp | x^\perp - p_1^\perp)$, also $x^\perp \in E \Leftrightarrow (e^\perp | x^\perp) = (e^\perp | p_1^\perp)$. Diese Darstellung der Ebene heißt **Hessesche Normalform**.

Appletanleitung Wählen Sie einen beliebigen Vektor der Norm 1, z. B. $e^\perp = (1, -1, 0)^t / \sqrt{2}$. Schauen Sie sich die Ebene durch den Nullpunkt senkrecht zu e^\perp an.

*** TO DO ***

Für einen beliebigen Punkt a^\perp berechnet sich der **Abstand von a^\perp zur Ebene E** zu $\text{Abstand}((a^\perp, E) = (a^\perp - p_1^\perp | e^\perp)$.

Der Abstand ist positiv, wenn a^\perp auf der Seite der Ebene liegt, in die e^\perp zeigt, negativ im anderen Fall.

Appletanleitung Wählen sie die Ebene aus der obigen Anschauung und den Vektor $a^\perp = (1, 1, 1)^t$. Schauen Sie sich den Abstandsvektor zur Ebene an.

*** TO DO ***

Schnittpunkt zwischen Gerade und Ebene

Wir denken uns die Ebene E durch $(x^\perp - p^\perp|e^\perp) = 0$ gegeben, die Gerade durch $y^\perp(t) = a^\perp + tb^\perp$. Dann gilt für den Durchstoßpunkt oder Schnittpunkt

$$(a^\perp + t_s b^\perp - p^\perp|e^\perp) = 0 = ((a^\perp - p^\perp)|e) + t_s(b^\perp|e).$$

Ist $(b^\perp|e^\perp) \neq 0$, also Gerade und Ebene **nicht** parallel, so gibt es genau einen Schnittpunkt: $t_s = \frac{(p^\perp - a^\perp|e^\perp)}{(b^\perp|e^\perp)}$.

Ist $(b^\perp|e^\perp) = 0$, so liegt die Gerade entweder in E oder schneidet E nie.

Anleitung zum Applet "Geraden und Ebenen im Raum" Wählen Sie $p^\perp = (1, 1, 0)^t$, e^\perp wie oben und wählen Sie eine bestimmte Gerade. Schauen Sie sich ihren Schnittpunkt mit der Ebene an.

Winkel zwischen zwei Ebenen

Seien E_1 und E_2 durch ihre Normalen f_1^\perp, f_2^\perp gegeben, also durch $(x^\perp - p_j^\perp|f_j^\perp) = 0$ ($j = 1, 2$), so ist der Winkel zwischen E_1 und E_2 derjenige zwischen f_1^\perp und f_2^\perp . Wegen $\|f_j\| = 1$ erhält man

$$\cos(\varphi) = (f_1^\perp|f_2^\perp), \quad \sin(\varphi) = \|f_1 \times f_2\|.$$

Projektion auf eine Ebene

Sei E durch $(y^\perp - p^\perp|e^\perp) = 0$ gegeben. Wir wollen einen Punkt x^\perp in Richtung ℓ^\perp auf E projizieren und den Bildpunkt \hat{x}^\perp berechnen. Es ist $\hat{x}^\perp = x^\perp + t\ell^\perp$ für ein t . Es muß $(e^\perp|\ell^\perp) \neq 0$ sein, weil sonst die Projektionsrichtung parallel zu e ist. Wegen $\hat{x}^\perp \in E$ erhält man $(x^\perp + t\ell^\perp - p^\perp|e^\perp) = 0$, also

$$t = \frac{(p^\perp - x^\perp|e^\perp)}{(e^\perp|\ell^\perp)} \quad \text{und damit}$$

$$\hat{x}^\perp = x^\perp + \frac{(p^\perp - x^\perp|e^\perp)}{(e^\perp|\ell^\perp)} \ell^\perp.$$

Appletanleitung Wählen Sie $e^\perp = (1, -1, 0)^t/\sqrt{2}$, $p^\perp = (1, 1, 1)^t$ und $\ell^\perp = (1, 2, -3)^t$. Schauen Sie sich die Projektion für verschiedene x^\perp an.

* * * TO DO * * *

Damit können wir **Spiegelungen an einer Ebene** berechnen. Hier handelt es sich zunächst um die senkrechte Projektion auf E , also $\ell^\perp = e^\perp$ und dann um Weitergehen um die gleiche Länge in Richtung e^\perp , also ist der gespiegelte Punkt

$$x_s^\perp = x^\perp + 2(p^\perp - x^\perp|e^\perp)e^\perp.$$

Sphären (Kugeln im klassischen Sinn)

Eine **Sphäre** ist der geometrische Ort aller Punkte, die von einem festen Punkt m^\downarrow , dem **Mittelpunkt** einen festen Abstand r ($r > 0$ oder Radius) haben.

Also $S = \{x^\downarrow : (x^\downarrow - m^\downarrow | x^\downarrow - m^\downarrow) = r^2\}$

Die Gleichung $(x^\downarrow - m^\downarrow | x^\downarrow - m^\downarrow) = r^2$ heißt Kugelgleichung. Wir schneiden S mit einer Ebene in Parameterform

$$z(s, t) = p^\downarrow + sa^\downarrow + tb^\downarrow \quad \text{mit} \quad \|a^\downarrow\| = \|b^\downarrow\| = 1, \quad (a^\downarrow | b^\downarrow) = 0.$$

Wir erhalten für die Schnittmenge die folgende Gleichung für s und t :

$$\|p^\downarrow + sa^\downarrow + tb^\downarrow - m^\downarrow\|^2 = r^2.$$

Es ergibt sich

$$\begin{aligned} \|p^\downarrow - m^\downarrow\|^2 + s^2 + t^2 + 2s(a^\downarrow | p^\downarrow - m^\downarrow) + 2t(b^\downarrow | p^\downarrow - m^\downarrow) &= r^2 \\ \text{oder} \quad s^2 + t^2 + 2s(a^\downarrow | p^\downarrow - m^\downarrow) + 2t(b^\downarrow | p^\downarrow - m^\downarrow) &= r^2 - \|p - m^\downarrow\|^2 \end{aligned}$$

Quadratische Ergänzung liefert

$$(s + (a^\downarrow | p^\downarrow - m^\downarrow))^2 + (t + (b^\downarrow | p^\downarrow - m^\downarrow))^2 = r^2 + (a^\downarrow | p^\downarrow - m^\downarrow)^2 + (b^\downarrow | p^\downarrow - m^\downarrow)^2 - \|p - m^\downarrow\|^2.$$

Da die Vektoren a^\downarrow und b^\downarrow die Länge 1 haben und senkrecht aufeinander stehen, ist die Schnittpunktmenge ein Kreis, ein Punkt oder leer, je nachdem, ob die rechte Seite $>$, $=$ oder < 0 ist.

Wiederholung:

- Spatprodukt
- Vektorprodukt
- zusammengesetzte Produkte
- Schnittpunkte zwischen Ebenen und Geraden, Ebenen und Ebenen, Ebenen und Kugeln

7.2 Lineare Abbildungen und Matrizen

Überblick: Der Begriff der linearen Abbildung ist einer der zentralen Begriffe für die gesamte Mathematik. Geometrisch gesprochen sind dies Abbildungen eines Vektorraumes V in einen anderen W , die den Nullpunkt in den Nullpunkt und Dreiecke in Dreiecke überführen (die auch entartet, also Strecken oder Punkte werden können). Wichtige Spezialfälle sind lineare Abbildungen des \mathbb{R}^p in sich, die die Länge von Vektoren erhalten. Hierunter fallen Drehungen und Spiegelungen. Ebenso wichtig sind gerade für die graphische Datenverarbeitung die Projektionen.

Lineare Abbildungen sind durch die Angabe des Bildes einer Basis vollständig bestimmt und können deshalb in Zahlentafeln (Matrizen) kodiert werden. Mit ihrer Hilfe lassen sich lineare Gleichungssysteme theoretisch elegant behandeln. Zentral ist daneben der Begriff der Determinante einer linearen Abbildung eines p -dimensionalen Vektorraumes in sich, der den Verzerrungsfaktor für das Bild des p -dimensionalen Einheitsvolumens angibt.

7.2.1 Lineare Abbildungen

Überblick: Ausgehend von Drehungen, Projektionen und Spiegelungen erklären wir den Begriff der linearen Abbildung, ihres Ranges, ihres Kernes und zeigen die berühmte Dimensionsformel.

Anleitung zum Applet “affine Transformationen”

1. Drehungen in der Ebene

Wir schauen uns eine Drehung einer Ebene um einen willkürlich festgelegten Nullpunkt um $45^\circ (= \pi/4)$ an. Wir führen zwei senkrecht aufeinander stehende Vektoren e_1^\perp und e_2^\perp der Länge 1 ein, so daß $\langle (e_1, e_2) = \pi/2$ ist. Dann ist das Bild von e_1^\perp unter der Drehung gerade $(e_1 + e_2)/\sqrt{2}$ und das Bild von e_2^\perp ist $(-e_1 + e_2)/\sqrt{2}$. Eine Drehung erhält Längen und Winkel und damit Dreiecke; genauer ist das Bilddreieck jeden Dreiecks kongruent zum (Urbild-) Dreieck. Bezeichnen wir die Drehung mit D , so ist also

$$D(x^\perp + y^\perp) = Dx^\perp + Dy^\perp$$

und

$$D(\lambda x^\perp) = \lambda D(x^\perp).$$

Abbildungen, die diese beiden Gleichungen erfüllen, heißen linear.

Damit können wir allein aus der Kenntnis der Bilder unserer Basis e_1^\perp und e_2^\perp die Bilder beliebiger Vektoren x^\perp berechnen. Denn es ist wegen der Linearität

$$\begin{aligned} D(x^\perp) &= D(x_1 e_1^\perp + x_2 e_2^\perp) = x_1 D(e_1^\perp) + x_2 D(e_2^\perp) \\ &= x_1 (e_1^\perp + e_2^\perp)/\sqrt{2} + x_2 (-e_1^\perp + e_2^\perp)/\sqrt{2} \\ &= (x_1 - x_2)/\sqrt{2} \cdot e_1^\perp + (x_1 + x_2)/\sqrt{2} \cdot e_2^\perp \end{aligned}$$

Sie können die Drehwinkel um eine bestimmte Achse direkt eingeben. Tragen sie in der linken “Matrix” (“Transformationen”) im Kästchen ganz rechts unten eine Null ein, dreht sich nur die x - y -Ebene. Versuchen Sie die “Transformation so zu ändern, daß sich nur die x - z - oder nur die y - z -Ebene dreht.

(Rechts sehen sie dann die “Matrix dieser Drehung”. Was das genau bedeutet wird im übernächsten Abschnitt behandelt)

2. Drehungen im Raum

Im Raum, den wir durch \mathbb{R}^3 beschreiben, hat eine Drehung immer eine Drehachse. Wir legen den 0-Punkt unseres Koordinatensystems so, daß die Drehachse durch ihn verläuft und wählen auf der Drehachse einen Vektor e_3^\perp der Länge 1 so, daß wenn der Daumen der rechten Faust in Richtung von e_3^\perp zeigt, die Finger der Faust in Drehrichtung zeigt. In der Ebene senkrecht zur Drehachse führen wir orthogonale Vektoren e_1^\perp und e_2^\perp so ein, daß $(e_1^\perp, e_2^\perp, e_3^\perp)$ ein Rechtssystem ist.

Es handle sich um eine Drehung um den Winkel $\varphi > 0$. Dann bleibt e_3^\perp fest, also $D(e_3^\perp) = e_3^\perp$ ($D =$ die Drehung), $D(e_1^\perp) = \cos \varphi e_1^\perp + \sin \varphi e_2^\perp$, $D(e_2^\perp) = (-\sin \varphi e_1^\perp + \cos \varphi e_2^\perp)$. Da wieder Länge und Winkel erhalten bleiben, werden beliebige Dreiecke in kongruente Dreiecke überführt. Insbesondere ist D linear und es gilt wieder

$$\begin{aligned} D(x^\perp) &= D(x_1 e_1^\perp + x_2 e_2^\perp + x_3 e_3^\perp) \\ &= x_1 D(e_1^\perp) + x_2 D(e_2^\perp) + x_3 D(e_3^\perp), \end{aligned}$$

also

$$D(x^\perp) = (x_1 \cos \varphi + x_2 \sin \varphi)e_1^\perp + (-x_1 \sin \varphi + x_2 \cos \varphi)e_2^\perp + x_3 e_3^\perp$$

Drehen Sie um die z -Achse. Versuchen Sie das gleich Ergebnis zu erzielen, indem Sie bei "Transformationen" Zahlen eingeben.

3. Projektionen auf eine Ebene

Wir denken uns die Koordinaten im Raum so eingeführt, daß die Ebene E von zwei orthogonalen Vektoren e_1^\perp, e_2^\perp der Norm 1 aufgespannt wird, und wählen $e_3^\perp = e_1^\perp \times e_2^\perp$. Sei b^\perp ein nicht in E liegender Vektor der Länge 1, also $b^\perp = b_1 e_1^\perp + b_2 e_2^\perp + b_3 e_3^\perp$ mit $b_3 \neq 0$. Der Normalenvektor von E ist gerade e_3^\perp , die Ebene E ist durch $e_3^\perp \cdot x^\perp = 0$ gegeben. Also ist der Bildpunkt von x^\perp unter der Projektion längs b^\perp durch

$$P(x^\perp) = x^\perp - \frac{(x^\perp | e_3^\perp)}{(b^\perp | e_3^\perp)} b^\perp$$

gegeben. Allein an dieser Formel sieht man: Dreiecke gehen wieder in Dreiecke über, die aber nicht mehr kongruent sind. Genauer: P ist wieder linear.

In Formeln erhält man zunächst

$$\begin{aligned} P(e_1^\perp) &= e_1^\perp, & P(e_2^\perp) &= e_2^\perp, \\ P(e_3^\perp) &= e_3^\perp - \frac{1}{b_3} b^\perp = -\frac{b_1}{b_3} e_1^\perp - \frac{b_2}{b_3} e_2^\perp \end{aligned}$$

Damit erhält man

$$\begin{aligned} P(x^\perp) &= P(x_1 e_1^\perp + x_2 e_2^\perp + x_3 e_3^\perp) = x_1 P(e_1^\perp) + x_2 P(e_2^\perp) + x_3 P(e_3^\perp) \\ &= \left(x_1 - \frac{b_1}{b_3} x_3\right) e_1^\perp + \left(x_2 - \frac{b_2}{b_3} x_3\right) e_2^\perp \end{aligned}$$

Wählen Sie $b^\perp = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Versuchen Sie durch Eingabe von Zahlen in "Transformationen" das durch die oben angegebenen Bilder der Basisvektoren definierte Bild des Würfels zu erhalten. Versuchen Sie das gleiche mit anderen Vektoren b^\perp .

Ausgehend von diesen vielen Beispielen erklären wir nun:

Definition 7.18 Eine Abbildung A vom Vektorraum V in den Vektorraum W heißt **linear**, wenn für alle $x, y \in V$ und alle $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} A(x + y) &= A(x) + A(y) \\ A(\lambda x) &= \lambda A(x) \end{aligned}$$

Im einzelnen gilt also:

- (i) Ist Δ ein Dreieck mit den Ecken $0, x$ und $x + y$ so wird es durch die lineare Abbildung A in das Dreieck (in W) mit den Ecken $0, A(x)$ und $A(x) + A(y)$ abgebildet.

- (ii) Dehnt man den Vektor x mit dem Faktor λ , so wird der so gedehnte Vektor abgebildet auf den Vektor, den man erhält, wenn man das Bild $A(x)$ von x um denselben Faktor dehnt.

Der Begriff der linearen Abbildung ist zentral für die gesamte Mathematik. Hier sind weitere Beispiele, die Sie sich zum Teil auch am Computer anschauen können.

Beispiele:

1. Spiegelung in \mathbb{R}^2 an einer Gerade durch den Nullpunkt. Sei $G = \{ta^\downarrow : t \in \mathbb{R}\}$, $a^\downarrow = \begin{pmatrix} a \\ b \end{pmatrix} \neq 0^\downarrow$. Sei $e^\downarrow = \frac{1}{\sqrt{a^2 + b^2}} \begin{pmatrix} b \\ -a \end{pmatrix}$. Dann ist $e^\downarrow \perp G$. Sei $x^\downarrow \in \mathbb{R}^2$ beliebig. Dann ist der an G gespiegelte Punkt $S(x^\downarrow) = x^\downarrow - 2(x^\downarrow | e^\downarrow) e^\downarrow$. Prüfen Sie nach, daß S linear ist.
2. Spiegelung in \mathbb{R}^3 an einer Ebene durch den Nullpunkt. Sei e^\downarrow der Normalenvektor der Ebene. Dann ist $S(x^\downarrow) = x^\downarrow - 2(x^\downarrow | e^\downarrow) e^\downarrow$ wieder der Spiegelungspunkt.
3. Die **Parametrisierung** einer Ebene in einem beliebigen Vektorraum V durch \mathbb{R}^2 ist eine lineare Abbildung. Genauer: Seien $a, b \in V$ linear unabhängig (wir lassen bei Vektoren eines beliebigen Vektorraums die Pfeile weg und bezeichnen reelle Zahlen mit griechischen Buchstaben). Sei $E = \{\xi a + \mu b : \xi, \mu \in \mathbb{R}\}$ die von a und b in V gespannte Ebene durch den Nullpunkt. Sei $P\left(\begin{pmatrix} \xi \\ \eta \end{pmatrix}\right) = \xi a + \eta b$. Dann ist P eine lineare Abbildung von \mathbb{R}^2 auf die Ebene E .
4. Sei $a^\downarrow = \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}$ ein fester Vektor aus \mathbb{R}^p . Dann ist die Abbildung $S_{a^\downarrow} : x^\downarrow \mapsto S_{a^\downarrow}(x^\downarrow) = (a^\downarrow | x^\downarrow) a^\downarrow$ linear.

Mit linearen Abbildungen kann man rechnen.

Satz 7.19 a) Seien V, W Vektorräume und $A, B : V \rightarrow W$ seien lineare Abbildungen. Dann ist auch

$$A + B : x \mapsto (A + B)(x) := A(x) + B(x)$$

eine lineare Abbildung. Ebenso ist für $\lambda \in \mathbb{R}$

$$\lambda A : x \mapsto (\lambda A)(x) := \lambda A(x)$$

eine lineare Abbildung.

b) Sei X ein weiterer Vektorraum und $A : V \rightarrow W$, $B : W \rightarrow X$ seien lineare Abbildungen. Dann ist auch die Hintereinanderausführung $B \cdot A : V \rightarrow X$ eine lineare Abbildung.

Korollar 7.20 Die Menge $L(V, W)$ aller linearen Abbildungen von V nach W bildet mit der oben angegebenen Addition und mit der oben angegebenen Multiplikation mit Skalaren einen Vektorraum.

Die folgenden Aufgaben liefern eine Vertiefung des Satzes und Beispiele.

Aufgaben:

1. Beweisen Sie bitte den vorausgegangenen Satz.
2. Zeigen Sie bitte: die Hintereinanderausführung zweier Drehungen in \mathbb{R}^2 ist eine Drehung.
3. Sei S eine Spiegelung des \mathbb{R}^3 an einer Ebene durch den Nullpunkt. Was ist $S^2 (= S \circ S)$?
4. Sei P eine Projektion des \mathbb{R}^3 auf eine Ebene durch den Nullpunkt längs einem Vektor $a^\perp (\neq 0^\perp)$. Was ist P^2 .

Wir haben uns bisher um das Bild $A(V)$ der linearen Abbildung A gekümmert und damit den Rang der Abbildung definiert. Bei einer Projektion werden nun eine Menge Vektoren "geschluckt", das heißt auf den Nullvektor abgebildet. Welche sind das (siehe Aufg. 4 oben)?

Definition 7.21 Sei A eine lineare Abbildung des Vektorraumes V in den Vektorraum W . Die Menge $\{x \in V : A(x) = 0\}$ heißt **Kern** $\ker(A)$ der Abbildung A .

Klar ist der folgende Satz:

Satz 7.22 Sei $A : V \rightarrow W$ eine lineare Abbildung.

- a) $\ker(A)$ ist ein linearer Teilraum von V .
- b) A ist genau dann injektiv, wenn $\ker(A) = \{0\}$, d.h. der Kern nur aus dem Nullvektor besteht.

Aufgabe: Beweisen Sie diesen Satz.

Seien V und W beliebige Vektorräume und $A : V \rightarrow W$ sei eine lineare Abbildung. Dann ist das Bild $A(V)$ ein Untervektorraum von W .

Aufgabe: Beweisen Sie diese Aussage.

Definition 7.23 Die Dimension des Bildes $A(V)$ einer linearen Abbildung $A : V \rightarrow W$ heißt der **Rang** $\text{rng}(A)$ von A .

Beispiele:

1. Der Rang einer Drehung des \mathbb{R}^2 ist 2
2. Der Rang einer Projektion auf eine Gerade ist 1
3. Der Rang einer Projektion auf eine Ebene ist 2
4. Der Rang der Abbildung $S_{a^\perp} : \mathbb{R}^p \rightarrow \mathbb{R}^p$ ($x^\perp \mapsto (a^\perp | x^\perp)$) ist 1 falls $a^\perp \neq 0$ und 0 sonst.

Lineare Abbildungen und Basen

Überblick: Wie die Drehungen sind alle linearen Abbildungen bereits vollkommen bestimmt durch die Angabe der Bilder einer Basis. Hieraus ergibt sich die Dimensionsformel, die es erlaubt, die Dimension des Bildes einer linearen Abbildung zu berechnen.

Satz 7.24 Seien V und W Vektorräume, V sei endlichdimensional und $B = \{b_1, \dots, b_p\}$ sei eine Basis von V .

- Sei $A : V \rightarrow W$ eine lineare Abbildung. Dann ist A bereits durch die Angabe der Bildvektoren $A(b_1), \dots, A(b_p)$ eindeutig festgelegt.
- Sei $\varphi : B \rightarrow W$ eine ganz beliebige Abbildung. Dann gibt es genau eine lineare Abbildung $A_\varphi : V \rightarrow W$ mit $A_\varphi(b_j) = \varphi(b_j)$ für $j = 1, \dots, p$.

Beweis:

- Es ist $x = \sum \xi_j b_j$ mit eindeutig bestimmten ξ_1, \dots, ξ_p . Wegen der Linearität ist $A(x) = \sum \xi_j A(b_j)$.
- Definiere $A_\varphi(x) = \sum \xi_j \varphi(b_j)$. Zeige selbst, daß A_φ wohldefiniert und linear ist.

Speziell für endlich-dimensionale Vektorräume erhalten wir die folgende Dimensionsbeziehung, die weitreichende Anwendungen hat (s.das Korollar).

Satz 7.25 Dimensionsformel

Seien V und W Vektorräume, V sei endlich-dimensional. Sei $A : V \rightarrow W$ eine lineare Abbildung. Dann gilt die Dimensionsformel

$$\begin{aligned} \dim(V) &= \dim(\ker(A)) + \text{rng}(A) \\ &= \dim(\ker(A)) + \dim(A(V)) \end{aligned}$$

Beweis: (Skizze)

Sei $\dim(V) = p$, $\dim(\ker(A)) = q$ und $\dim(A(V)) = r$. Wir wählen uns eine Basis $\{b_1, \dots, b_q\}$ des Kerns von A und ergänzen sie durch $\{b_{q+1}, \dots, b_p\}$ zu einer solchen von V . (Daß das möglich ist, ist anschaulich klar, aber nicht ganz einfach abstrakt zu beweisen). Wir zeigen

- $A(b_{q+1}), \dots, A(b_p)$ sind linear unabhängig
- $A(b_{q+1}), \dots, A(b_p)$ spannen das Bild $A(V)$ auf.

Wiederholung:

- Drehungen, Spiegelungen, Projektionen
- Lineare Abbildungen
- Rechnen mit linearen Abbildungen
- Kern und Bild einer linearen Abbildung
- Lineare Abbildungen und Basen
- Die Dimensionsformel

7.2.2 Lineare Abbildungen und Matrizen

Überblick: Eine lineare Abbildung liegt durch die Angabe der Bilder einer Basis fest. Ist auch der Bildraum endlichdimensional, so kann man die lineare Abbildung durch eine Zahlentafel - eine Matrix - kodieren. Die Rechenregeln für lineare Abbildungen spiegeln sich in den Rechenregeln für die Matrizen wieder. Dies erfordert eine spezielle Definition der Multiplikation von Matrizen. Auch Basiswechsel werden mit Matrizen kodiert.

Seien V und W endlichdimensionale Vektorräume. Wir zeichnen sowohl in V als auch in W zwei Basen aus. $B = \{b_1, \dots, b_p\}$ sei eine Basis in V , $C = \{c_1, \dots, c_q\}$ sei eine Basis in W .

Alles was folgt hängt von diesen beiden Basen ab!

Als Beispiele führen Sie sich vor Augen:

Beispiele:

$$1. V = \mathbb{R}^3, \quad b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad B = \{b_1, b_2, b_3\}$$

$$W = V \text{ und } C = B.$$

$$2. V \text{ wie in 1. } W = \mathbb{R}^2, \quad c_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad c_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad C = \{c_1, c_2\}$$

$$3. V = \mathbb{R}^2, \quad W = \mathbb{R}^3, \text{ Basis in } V \text{ sei } \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$$

$$\text{Basis in } W = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Sei A eine lineare Abbildung von V nach W . Nach Satz 7.24 kennen wir $A : V \mapsto W$, wenn wir die Bilder $A(b_1), \dots, A(b_p)$ der Basis B von V kennen. Wir schreiben

$$A(b_1) = a_{11}c_1 + a_{21}c_2 + \dots + a_{q1}c_q$$

a_{11}, \dots, a_{q1} sind eindeutig bestimmt, weil $C = \{c_1, \dots, c_q\}$ eine Basis in W ist.

Die Wahl der Indizes in dieser Form ist eine internationale stillschweigende Vereinbarung.

Entsprechend ist

$$\begin{aligned} A(b_2) &= a_{12}c_1 + a_{22}c_2 + a_{32}c_3 + \dots + a_{q2}c_q \\ &\vdots \\ A(b_p) &= a_{1p}c_1 + \dots + a_{qp}c_q. \end{aligned}$$

Wir fassen die Koeffizienten zu einer Zahlentafel - **Matrix** genannt - zusammen

$$\mathcal{A}_A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1p} \\ \vdots & & & & \vdots \\ a_{q1} & a_{q2} & a_{q3} & \cdots & a_{qp} \end{pmatrix}$$

Die k -te Spalte enthält gerade die Koordinaten des Bildes $A(b_k)$ des k -ten Basisvektors von V bezüglich der Basis C in W .

Definition 7.26 \mathcal{A}_A heißt die **Matrixdarstellung** der linearen Abbildung A **bezüglich der Basen** $B \subset V$ und $C \subset W$.

Beispiele:

1. $V = W = \mathbb{R}^2$, $B = C = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = \{e_1, e_2\}$

A sei eine Drehung um den Winkel φ . Dann ist

$$A(e_1) = \cos \varphi e_1 + \sin \varphi e_2, \quad A(e_2) = -\sin \varphi e_1 + \cos \varphi e_2,$$

$$\text{also } \mathcal{A}_A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

2. $V = W = \mathbb{R}^3$, $B = C = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} = \{e_1, e_2, e_3\}$.

A sei die Drehung um e_3 um den Winkel φ .

Dann ist $A(e_3) = e_3$, $A(e_1) = \cos \varphi e_1 + \sin \varphi e_2$

$A(e_2) = -\sin \varphi e_1 + \cos \varphi e_2$.

$$\text{Also ist } \mathcal{A}_A = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Sei $V = \mathbb{R}^3$ und b_3 der Normalenvektor einer von zwei Vektoren b_1 und b_2 aufgespannten Ebene durch den Nullpunkt. Hier wählen wir nicht die kanonische Basis wie in 2 sondern die problemangepaßte Basis $\{b_1, b_2, b_3\}$. Sei $\ell = \alpha b_1 + \beta b_2 + \gamma b_3$ mit $\gamma \neq 0$ eine vorgegebene Richtung, die *nicht* parallel zur Ebene verläuft. Die Projektionsformel lautet

$$P(x) = x - \frac{(b_3|x)}{(b_3|\ell)} \ell$$

also $P(b_1) = b_1$, $P(b_2) = b_2$. Weil $b_3 \perp b_1, b_2$ und $(b_3|b_3) = 1$ erhält man $P(b_3) = b_3 - \frac{1}{(b_3|\ell)} \ell$,

also $P(b_3) = -\frac{\alpha}{\gamma} b_1 - \frac{\beta}{\gamma} b_2$ und damit

$$A_P = \begin{pmatrix} 1 & 0 & -\frac{\alpha}{\gamma} \\ 0 & 1 & -\frac{\beta}{\gamma} \\ 0 & 0 & 0 \end{pmatrix}.$$

Beachten Sie bitte, daß A_P so einfach aussieht, weil wir die *Basis problemangepaßt gewählt* haben.

Ist insbesondere die Projektion *senkrecht* auf die Ebene, also $\ell = b_3$, so ist

$$A_P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

4. In $V = \mathbb{R}^2$ wählen wir eine beliebige orthogonale Basis $\{b_1, b_2\}$ mit $\|b_j\| = 1$. b_1 ist physikalisch der *Grundzustand*, b_2 der *angeregte Zustand* eines quantenmechanischen Teilchens.

A^- ist die lineare Abbildung, die das Teilchen im Grundzustand vernichtet, das Teilchen im angeregten Zustand in den Grundzustand überführt, also

$$A^-(b_1) = 0, \quad A^-(b_2) = b_1, \quad \mathcal{A}_{A^-} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

A^- heißt (Einteilchen-) **Vernichter**

Aufgabe: Finden Sie eine geeignete Basis und die Matrixdarstellung der Spiegelung an einer Ebene durch den Nullpunkt (im Raum).

Sind A und B gegeben, so lassen sich leicht \mathcal{A}_{A+B} und $\mathcal{A}_{\lambda B}$ aus $\mathcal{A}_A, \mathcal{A}_B$ berechnen.

Satz 7.27 Sei $\mathcal{A}_A = \begin{pmatrix} a_{11}, & \dots, & a_{1p} \\ \vdots & & \vdots \\ a_{q1}, & \dots, & a_{qp} \end{pmatrix}$ und $\mathcal{A}_B = \begin{pmatrix} b_{11}, & \dots, & b_{1p} \\ \vdots & & \vdots \\ b_{q1}, & \dots, & b_{qp} \end{pmatrix}$.

Dann ist

$$\mathcal{A}_{A+B} = \begin{pmatrix} a_{11} + b_{11}, & \dots, & a_{1p} + b_{1p} \\ \vdots & & \vdots \\ a_{q1} + b_{q1}, & \dots, & a_{qp} + b_{qp} \end{pmatrix},$$

das heißt man addiert koordinatenweise.

$$\mathcal{A}_{\lambda A} = \begin{pmatrix} \lambda a_{11}, & \dots, & \lambda a_{1p} \\ \vdots & & \vdots \\ \lambda a_{q1}, & \dots, & \lambda a_{qp} \end{pmatrix},$$

das heißt, man multipliziert jeden Eintrag mit λ .

Wie drückt sich die Hintereinanderausführung zweier linearer Abbildungen durch die entsprechenden Matrizen aus? Dazu berechnen wir zunächst, wie sich die Koordinaten des Bildes eines beliebigen Vektors x berechnen.

Seien V, W wie bisher endlichdimensionale Vektorräume mit Basen $B = \{b_1, \dots, b_p\} \subset V$ und $C = \{c_1, \dots, c_q\} \subset W$. Sei $A : V \rightarrow W$ eine lineare Abbildung von V nach W mit der Matrix

$$\mathcal{A}_A = \begin{pmatrix} a_{11}, & \dots, & a_{1p} \\ \vdots & & \vdots \\ a_{q1}, & \dots, & a_{qp} \end{pmatrix}.$$

Es ist also

$$A(b_j) = a_{1j} c_1 + \dots + a_{qj} c_q.$$

Sei $x = \xi_1 b_1 + \dots + \xi_p b_p$. Dann ist

$$\begin{aligned} A(x) &= \xi_1 A(b_1) + \dots + \xi_p A(b_p) \\ &= \xi_1 (a_{11} c_1 + a_{21} c_2 + \dots + a_{q1} c_q) + \dots + \xi_p (a_{1p} c_1 + \dots + a_{qp} c_q) \\ &= \sum_{j=1}^p \xi_j \left(\sum_{k=1}^q a_{kj} c_k \right) \\ &= \sum_{k=1}^q \left(\sum_{j=1}^p a_{kj} \xi_j \right) c_k \\ &= c_1 (a_{11} \xi_1 + a_{12} \xi_2 + \dots + a_{1p} \xi_p) + \dots + c_q (a_{q1} \xi_1 + \dots + a_{qp} \xi_p). \end{aligned}$$

Also ist die j -te Koordinate des Bildvektors $A(x) = \eta_1 c_1 + \dots + \eta_q c_q$ gerade

$$\eta_j = a_{j1} \xi_1 + a_{j2} \xi_2 + \dots + a_{jp} \xi_p.$$

Aufgabe: Berechnen Sie die Koordinaten eines Bildvektors $A(x)$, wo A wie die einzelnen Beispiele oben ist (Drehung in \mathbb{R}^2 , in \mathbb{R}^3 , Projektion auf eine Ebene und Spiegelung).

Wir führen den j -ten **Zeilenvektor** der Matrix \mathcal{A}_A ein: $\vec{a}_j = (a_{j1}, a_{j2}, \dots, a_{jp})$. Dann läßt sich \mathcal{A}_A schreiben als

$$\mathcal{A}_A = \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_q \end{pmatrix}.$$

Wir führen nun das **Produkt eines Zeilenvektors** $\vec{b} = (\beta_1, \dots, \beta_p)$ der Länge p mit einem gleich

langen Spaltenvektor $x^\downarrow = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix}$ ein:

$$\vec{b}x^\downarrow = \beta_1 \xi_1 + \beta_2 \xi_2 + \dots + \beta_p \xi_p.$$

Mit dieser Konvention erhalten wir für die Koordinaten des Bildvektors $A(x) = A(\sum_{j=1}^p \xi_j b_j)$:

$$\begin{aligned} \eta_1 &= \vec{a}_1 x^\downarrow \\ \eta_2 &= \vec{a}_2 x^\downarrow \\ &\vdots \\ \eta_q &= \vec{a}_q x^\downarrow \end{aligned}$$

Dies fassen wir zusammen zum Produkt der Matrix \mathcal{A}_A (mit q Zeilen und p Spalten) mit dem Spaltenvektor x^\downarrow der Länge p (= Anzahl der Spalten von \mathcal{A}_A).

$$\mathcal{A}_A x^\downarrow = \begin{pmatrix} \vec{a}_1 x^\downarrow \\ \vdots \\ \vec{a}_q x^\downarrow \end{pmatrix}.$$

Das Ergebnis ist ein Spaltenvektor der Länge q (= Zeilenzahl von \mathcal{A}_A).

Damit haben wir den Satz:

Satz 7.28 Seien V, W, B und C wie oben und $A : V \rightarrow W$ eine lineare Abbildung mit Matrix \mathcal{A}_A bezüglich der Basen B und C .

Sei $x \in V$ ein Vektor mit dem Koordinatenvektor

$$x^\downarrow = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix},$$

d.h. $x = \xi_1 b_1 + \dots + \xi_p b_p$. Dann besitzt das Bild $A(x)$ (bezüglich der Basis C) den Koordinatenvektor

$$y^\downarrow = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_q \end{pmatrix} = \mathcal{A}_A x^\downarrow,$$

d.h. $A(x) = \eta_1 c_1 + \dots + \eta_q c_q$ mit $\eta_j = \vec{a}_j x^\downarrow$.

Nachdem wir die Koordinaten des Bildvektors $A(x)$ aus denen des Urbildvektors berechnen können, ist es nicht mehr schwer, die Matrix der Hintereinanderausführung zweier linearer Abbildungen zu berechnen. Seien V, W und X endlichdimensionale Vektorräume mit Basen $B = \{b_1, \dots, b_p\}$, $C = \{c_1, \dots, c_q\}$ und $D = \{d_1, \dots, d_r\}$.

Seien $S : V \rightarrow W$ und $T : W \rightarrow X$ zwei lineare Abbildungen mit Matrizen $\mathcal{A}_S = \begin{pmatrix} s_{11}, \dots, s_{1p} \\ \vdots \\ s_{q1}, \dots, s_{qp} \end{pmatrix}$

und $\mathcal{A}_T = \begin{pmatrix} t_{11}, \dots, t_{1q} \\ \vdots \\ t_{r1}, \dots, t_{rq} \end{pmatrix}$. Das heißt, s_j^\downarrow ist der Koordinatenvektor von $S(b_j)$ bezüglich C , t_k^\downarrow ist der

Koordinatenvektor von $T(c_k)$ bezüglich D .

Nach dem Vorausgegangenen ist für

$(T \circ S)(b_j) = T(Sb_j)$ der Koordinatenvektor gerade $\mathcal{A}_T s_j^\downarrow$, also ergibt sich

$$\mathcal{A}_{T \circ S} = (\mathcal{A}_T S_1^\downarrow, \dots, \mathcal{A}_T S_p^\downarrow).$$

Setzen wir $\vec{t}_m = (t_{1m}, \dots, t_{qm})$ so erhalten wir

$$\begin{aligned} \mathcal{A}_{T \circ S} &= \begin{pmatrix} \vec{t}_1 s_1^\downarrow, \vec{t}_1 s_2^\downarrow, \dots, \vec{t}_1 s_p^\downarrow \\ \vdots \\ \vec{t}_r s_1^\downarrow, \vec{t}_r s_2^\downarrow, \dots, \vec{t}_r s_p^\downarrow \end{pmatrix} \\ &= \begin{pmatrix} \sum t_{1j} s_{j1}, \dots, \sum t_{1j} s_{jp} \\ \vdots \\ \sum t_{rj} s_{j1}, \dots, \sum t_{rj} s_{jp} \end{pmatrix}. \end{aligned}$$

Beispiel: Sei $V = W = X = \mathbb{R}^2$, $B = C = D = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ S die Drehung um den Winkel φ , T die um den Winkel ψ . Dann ist $\mathcal{A}_S = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$, $\mathcal{A}_T = \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix}$ und

$$\begin{aligned} \mathcal{A}_T \mathcal{A}_S &= \begin{pmatrix} \cos \psi \cos \varphi - \sin \psi \sin \varphi & -\cos \psi \cos \varphi - \sin \psi \sin \varphi \\ \sin \psi \cos \varphi + \cos \psi \sin \varphi & -\sin \psi \sin \varphi + \cos \psi \cos \varphi \end{pmatrix} \\ &= \begin{pmatrix} \cos(\varphi + \psi) & -\sin(\varphi + \psi) \\ \sin(\varphi + \psi) & \cos(\varphi + \psi) \end{pmatrix} \end{aligned}$$

und das ist tatsächlich die Matrix der Drehung um $\varphi + \psi$, also der Drehung $T \circ S$.

Gleichgültig ob die Matrizen \mathcal{A} und \mathcal{B} von linearen Abbildungen herrühren oder nicht, können wir ihr Produkt so erklären, daß es im Fall der Herkunft von linearen Abbildungen gerade mit der Matrix der

Hintereinanderausführung übereinstimmt. Seien $\mathcal{A} = \begin{pmatrix} a_{r1}, \dots, a_{1p} \\ \vdots \\ a_{q1}, \dots, a_{qp} \end{pmatrix}$ und $\mathcal{B} = \begin{pmatrix} b_{11}, \dots, b_{1q} \\ \vdots \\ t_{r1}, \dots, b_{rq} \end{pmatrix}$.

\mathcal{A} ist also eine Matrix mit q Zeilen und p Spalten, \mathcal{B} eine Matrix mit r Zeilen und q Spalten, also Zeilenzahl (\mathcal{A}) = Spaltenzahl (\mathcal{B}).

Dann (und nur dann) bilden wir das Matrizenprodukt

$$\begin{aligned} \mathcal{B}\mathcal{A} &= \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1q}a_{q1}, & \cdots & , b_{11}a_{1p} + \cdots + b_{1q}a_{qp} \\ & \vdots & \\ b_{r1}a_{11} + b_{r2}a_{21} + \cdots + b_{rq}a_{q1}, & \cdots & , b_{r1}a_{1p} + \cdots + b_{rq}a_{qp} \end{pmatrix} \\ &= \begin{pmatrix} \vec{b}_r a_1^\downarrow, & \cdots & , \vec{b}_r a_p^\downarrow \\ & \vdots & \\ \vec{b}_r a_1^\downarrow, & \cdots & , \vec{b}_r a_p^\downarrow \end{pmatrix}. \end{aligned}$$

Das Produkt ist dann eine $r \times p$ -Matrix kurz: *Das Produkt $\mathcal{B}\mathcal{A}$ einer $r \times q$ -Matrix \mathcal{B} mit einer $q \times p$ -Matrix \mathcal{A} ist eine $r \times p$ -Matrix.* q hebt sich weg.

Aufgaben: Trainieren Sie das Rechnen mit Matrizen.

1. $(1, 2, 3) \begin{pmatrix} 4 \\ 5 \\ 7 \end{pmatrix}$, also Produkt einer 1×3 -Matrix mit einer 3×1 -Matrix.

2. $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix}$, Produkt einer 2×2 -Matrix mit einer 2×1 -Matrix.

3. $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}$, Produkt zweier 2×2 -Matrizen.

4. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 7 & 10 & 13 \\ 8 & 11 & 14 \\ 9 & 12 & 15 \end{pmatrix}$, Produkt einer 2×3 -Matrix mit einer 3×3 -Matrix.

5. $\begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 7 & 9 \\ 8 & 10 \end{pmatrix}$ geht das?

6. Berechnen Sie im \mathbb{R}^3 die Matrix, die man erhält, indem man erst den Raum um $e_3^\downarrow = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ um

den Winkel φ dreht und dann auf die $\{e_1^\downarrow, e_2^\downarrow\}$ -Ebene längs $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ projiziert. Erhält man dasselbe, wenn man erst projiziert und dann dreht?

Für das Produkt haben wir folgende Rechenregeln

Satz 7.29 Seien \mathcal{B}_j zwei $r \times q$ -Matrix, \mathcal{A}_k zwei $q \times p$ -Matrizen und $\lambda, \in \mathbb{R}$. Dann gilt

$$\begin{aligned} (\mathcal{B}_1 + \mathcal{B}_2)\mathcal{A}_1 &= \mathcal{B}_1\mathcal{A}_1 + \mathcal{B}_2\mathcal{A}_1 \\ \mathcal{B}_1(\mathcal{A}_1 + \mathcal{A}_2) &= \mathcal{B}_1\mathcal{A}_1 + \mathcal{B}_1\mathcal{A}_2 \\ (\lambda\mathcal{B}_1)\mathcal{A}_1 &= \mathcal{B}_1(\lambda\mathcal{A}_1) = \lambda(\mathcal{B}_1\mathcal{A}_1) \end{aligned}$$

Wir merken noch die folgenden weiteren Regeln an:

Satz 7.30 a) Sei C eine $s \times r$ -Matrix, B eine $r \times q$ -Matrix und A eine $q \times p$ -Matrix. Dann ist $(CB)A = C(BA)$.

b) Seien B, A $p \times p$ -Matrizen (also quadratisch). Dann gilt im allgemeinen $BA \neq AB$ falls $p > 1$.

Aufgabe: Beweisen Sie bitte die beiden Sätze.

Um die folgenden Sätze besser zu verstehen, kehren wir noch einmal zu linearen Abbildungen zurück. Dabei bedenken wir, daß jede $q \times p$ -Matrix A eine lineare Abbildung A_A von \mathbb{R}^p in \mathbb{R}^q definiert durch $A_A(x^\downarrow) = Ax^\downarrow$, weil das Matrizenprodukt ja linear im 2. Argument ist.

Nach der Dimensionsformel $\dim(\ker(A)) + \dim(A(V)) = \dim V$ folgt, daß eine lineare Abbildung A von V nach W genau dann bijektiv ist, wenn $\dim(V) = \dim(W) = \dim(A(V))$ ist. Genau dann existiert die inverse Abbildung A^{-1} , die offensichtlich wiederum linear ist. Wir formulieren:

Satz 7.31 Sei $\dim(V) = \dim(W) = p$, $B = \{b_1, \dots, b_p\}$ sei eine Basis von V , $C = \{c_1, \dots, c_p\}$ sei eine solche von W . Sei $A: V \rightarrow W$. Dann gilt:

A ist genau dann bijektiv, wenn es eine $p \times p$ -Matrix B gibt mit

$$A_A B = B A_A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} = E_p,$$

wo E die $p \times p$ -Einheitsmatrix

$$E_p = (e_1^\downarrow, \dots, e_p^\downarrow)$$

ist. ($\{e_1^\downarrow, \dots, e_p^\downarrow\}$ die kanonische Basis in \mathbb{R}^p).

B ist dann eindeutig bestimmt und heißt die zu A_A inverse Matrix A_A^{-1} . Sie ist die zur inversen Abbildung A^{-1} gehörige Matrix, das heißt es gilt

$$A_A^{-1} = A_{A^{-1}}.$$

Beweis: Sei A bijektiv und $B = A_{A^{-1}}$.

Wegen $A^{-1}(A(b_j)) = b_j$ ist $A_{A^{-1}A} = E_p$, also $B A_A = E_p$. Ebenso ist $c_k = A(A^{-1}c_k)$, also ist $A_{AA^{-1}} = E_p$ und damit $A_A B = E_p$.

Mit der Matrix $B = (\beta_{ik})_{i,k=1 \dots p}$ erklärt man eine lineare Abbildung T von W nach V durch $T(c_k) = \sum_{\ell=1}^p \beta_{\ell k} b_\ell$ und erhält $A_T = B$ und wegen $A_A A_T = A_T A_A = E_p$ folgt $(AT)(c_k) = c_k$, also A bijektiv.

Natürlich nennt man allgemein eine $p \times p$ -Matrix B die Inverse zur $p \times p$ -Matrix A , wenn $BA = AB = E_p$ gilt. A muß nicht von einer bestimmten linearen Abbildung stammen.

Matrizen und Basiswechsel

Seien $B = \{b_1, \dots, b_p\}$ und $B' = \{b'_1, \dots, b'_p\}$ zwei Basen des Vektorraumes V . Sei $x \in V$. Dann ist $x = \sum \xi_j b_j = \sum_k \xi'_k b'_k$. Wie berechnen sich die Koordinaten ξ'_1, \dots, ξ'_p bezüglich der neuen Basis aus denen bezüglich der alten Basis?

Es ist $b'_j = \sum_{k=1}^p s_{kj} b_k$, wir bilden also einfach die Matrix $S_B = (s_1^\downarrow, \dots, s_p^\downarrow)$, wo $s_j^\downarrow = \begin{pmatrix} s_{1j} \\ \vdots \\ s_{pj} \end{pmatrix}$ der

Koordinatenvektor von b'_j bezüglich B ist. Genau so setzen wir

$$S_{B'} = (t_1^\downarrow, \dots, t_p^\downarrow) \text{ mit } t_k^\downarrow = \begin{pmatrix} t_{1k} \\ \vdots \\ t_{pk} \end{pmatrix}, \text{ wo } b_k = \sum_{\ell=1}^p t_{\ell k} b'_\ell \text{ ist.}$$

Wir führen nun zur Vereinfachung die folgende Abkürzung, das *Kronecker-Symbol* ein:

$$\delta_{k\ell} = \begin{cases} 1 & k = \ell \\ 0 & k \neq \ell \end{cases}$$

Damit erhalten wir für alle $k = 1, \dots, p$

$$b_k = \sum_{j=1}^p t_{jk} b'_j = \sum_{j=1}^p t_{jk} \sum_{\ell=1}^p s_{\ell j} b_\ell = \sum_{\ell=1}^p \left(\sum_{j=1}^p s_{\ell j} t_{jk} \right) b_\ell.$$

Aus dieser Gleichung folgt für alle $\ell = 1, \dots, p$ $\sum_{j=1}^p s_{\ell j} t_{jk} = \delta_{\ell k}$, denn die Koordinatendarstellung von b_k bezüglich der Basis $B = \{b_1, \dots, b_p\}$ ist nun einmal

$$b_k = 0 \cdot b_1 + 0 \cdot b_2 + \dots + 0 \cdot b_{k-1} + 1 \cdot b_k + 0 \cdot b_{k+1} + \dots + 0 \cdot b_p.$$

Das bedeutet aber $S_B S_{B'} = E_p$ und damit $S_{B'} = S_B^{-1}$.

Des weiteren erhalten wir aus

$$x = \sum_j \xi_j b_j = \sum_j \xi_j \sum_k t_{kj} b'_k = \sum_k \left(\sum_j t_{kj} \xi_j \right) b'_k = \sum_k \xi'_k b'_k$$

sofort

$$\begin{aligned} \xi^{\downarrow'} &= S_{B'} \xi^\downarrow \\ \xi^\downarrow &= S_{B'}^{-1} \xi^{\downarrow'} = S_B \xi^{\downarrow'}. \end{aligned}$$

Anders ausgedrückt: Um die *neuen* Koordinaten von x aus den alten zu berechnen, multipliziert man den alten Koordinatenvektor ξ^\downarrow mit der Matrix, die die Koordinaten der alten Basisvektoren bezüglich der neuen Basis haben.

Wir nennen S_B daher **Basiswechselmatrix**.

Beispiel: Sei $V = \mathbb{R}^2$, $B = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$, $B' = \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$.

Dann ist also $S_B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, $S_{B'} = S_B^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Machen Sie bitte die Probe!

Sei $x = b_1 + 2b_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Dann hat x bezüglich neuen, um 45 Grad gedrehten Basis die Koordinaten

$$\xi^{\downarrow'} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \text{ das heißt}$$

$$x = \frac{1}{2} \cdot 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2} \cdot 1 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Aufgaben: Berechnen Sie die Matrizen der folgenden Basiswechsel im \mathbb{R}^3 .

Dabei ist $B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

1. $B' = \left\{ \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. B' erhält man aus B durch Drehung um $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ um den Winkel φ .
2. $b'_1 = b_1, b'_2 = b_3, b'_3 = b_2$. Man erhält die neue Basis aus der alten durch Änderung des Drehsinus (Vertauschung zweier aufeinander folgender Vektoren).

Nun können wir leicht berechnen, wie sich die Matrixdarstellungen einer linearen Abbildung bei Basiswechsel verhalten:

Sei $A : V \rightarrow W$ eine lineare Abbildung, B, B' seien Basen in V mit S_B , wie oben, $C = \{c_1, \dots, c_q\}$, $C' = \{c'_1, \dots, c'_q\}$ seien Basen in W . S_C und $S_{C'}$ seien entsprechend gebildet.

Beantworten Sie bitte die Frage: Was enthalten S_B bzw. S_C als Spalten? Was enthalten S_C und $S_{C'}$ als Spalten? Sei $\mathcal{A}_A^{B,C}$ die Matrixdarstellung von A bezüglich der Basen B und C . Was sind die Spalten von $\mathcal{A}_A^{B',C'}$?

Satz 7.32 Die Matrixdarstellung $\mathcal{A}_A^{B',C'}$ von A bezüglich der neuen Basen berechnet sich aus $\mathcal{A}_A^{B,C}$ durch

$$\mathcal{A}_A^{B',C'} = S_B^{-1} \mathcal{A}_A^{B,C} S_C.$$

Beweis: Es ist

$$\begin{aligned} A(b'_j) &= A\left(\sum_{\ell} s_{\ell j} b_{\ell}\right) = \sum_{\ell} s_{\ell j} A(b_{\ell}) \\ &= \sum_{\ell} s_{\ell j} \sum_p a_{p\ell} c_p = \sum_{\ell} s_{\ell j} \sum_p a_{p\ell} \sum_r t'_{rp} c'_r \\ &= \sum_r \left(\sum_p t'_{rp} \sum_{\ell} a_{p\ell} s_{\ell j}\right) c'_r \\ &= \sum_r a'_{rj} c'_r. \end{aligned}$$

Beispiel: Wir nehmen $V = \mathbb{R}^2$, $B = \{e_1^{\downarrow}, e_2^{\downarrow}\} = C$ mit $e_1^{\downarrow} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2^{\downarrow} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$; $B' = \left\{ \frac{1}{\sqrt{2}}(e_1^{\downarrow} + e_2^{\downarrow}), \frac{1}{\sqrt{2}}(-e_1^{\downarrow} + e_2^{\downarrow}) \right\} = C'$.

Dann ist $S_B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Sei A die Spiegelung an der Hauptdiagonalen, also an der Geraden $G = \{\lambda(e_1^{\downarrow} + e_2^{\downarrow}) : \lambda \in \mathbb{R}\}$. Es gilt also $A(x^{\downarrow}) = x^{\downarrow} - (x^{\downarrow} - e_1^{\downarrow} + e_2^{\downarrow})(-e_1^{\downarrow} + e_2^{\downarrow})$. Das ergibt $A(e_1^{\downarrow}) = e_2^{\downarrow}$, $A(e_2^{\downarrow}) = e_1^{\downarrow}$, also $\mathcal{A}_A^{B,C} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Es ist $S_B^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Damit erhält man $\mathcal{A}_A^{B',C'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Wir haben also eine neue Basis gefunden, in der A eine besonders einfache Gestalt hat, nämlich *Diagonalgestalt*.

Man nennt eine quadratische Matrix A (Zeilenzahl = Spaltenzahl) **Diagonalmatrix** wenn sie die Form hat

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & \lambda_p \end{pmatrix} = (\lambda_1 e_1^{\downarrow}, \dots, \lambda_p e_p^{\downarrow}).$$

Wiederholung:

- Begriff der Matrix
- Lineare Abbildung und Matrix
- Addition und Multiplikation von Matrizen
- Matrizen und Basiswechsel

7.3 Determinanten und Lineare Gleichungssysteme

Überblick: Die Determinante einer $p \times p$ -Matrix ist das Volumen des von den Spaltenvektoren aufgespannten orientierten Parallel-Epipeds (p -dimensionales orientiertes Parallelogramm). Wir zeigen den Zusammenhang zwischen der Eigenschaft, daß die Determinante $\neq 0$ ist, der Invertierbarkeit der Matrix und der eindeutigen Lösbarkeit linearer Gleichungssysteme.

7.3.1 Determinanten

Wir haben oben das Spatprodukt $(a^\downarrow | b^\downarrow \times c^\downarrow)$ kennengelernt. Es gibt den Rauminhalt eines dreidimensionalen Parallelepipeds mit den Kanten $a^\downarrow, b^\downarrow, c^\downarrow$ an. Genauer informiert es darüber hinaus, ob $\{a^\downarrow, b^\downarrow, c^\downarrow\}$ in dieser Reihenfolge ein *Rechtssystem* ist, also die Vektoren wie Daumen, Zeigefinger und Mittelfinger der *rechten Hand* zeigen, in welchem Fall das Spatprodukt größer als Null ist, oder ob $\{a^\downarrow, b^\downarrow, c^\downarrow\}$ ein *Linkssystem* ist, also die Vektoren wie Daumen, Zeigefinger und Mittelfinger der *linken Hand* zeigen, in welchem Fall das Spatprodukt negativ ist. Genau dann ist schließlich das Spatprodukt 0, wenn $a^\downarrow, b^\downarrow, c^\downarrow$ linear abhängig sind, denn dann spannen sie gar kein Volumen auf sondern liegen in einer Ebene.

Man sagt, das Parallelepipid ist **positiv orientiert**, wenn die Kanten ein Rechtssystem bilden, **negativ orientiert**, wenn sie ein Linkssystem bilden und **entartet**, wenn sie in einer Ebene liegen. Ganz analoge Eigenschaften hat das nun eingeführte $*$ -Produkt zweier Vektoren $a^\downarrow, b^\downarrow$ in der Ebene:

$$a^\downarrow * b^\downarrow = \|a^\downarrow\| \|b^\downarrow\| \sin \angle(a^\downarrow, b^\downarrow).$$

Führen wir rechtwinklige Koordinaten $e_1^\downarrow, e_2^\downarrow$ ein, so erhalten wir für $a^\downarrow = a_1 e_1^\downarrow + a_2 e_2^\downarrow$ und $b^\downarrow = b_1 e_1^\downarrow + b_2 e_2^\downarrow$

$$a^\downarrow * b^\downarrow = a_1 b_2 - a_2 b_1.$$

Dann gilt auch hier, daß $a^\downarrow * b^\downarrow$ den *orientierten Flächeninhalt* des von a^\downarrow und b^\downarrow aufgespannten Parallelogramms angibt. Es ist positiv, wenn der Winkel von a^\downarrow zu b^\downarrow entgegengesetzt dem Uhrzeigersinn durchlaufen wird, negativ, wenn er im Uhrzeigersinn durchlaufen wird und 0 genau dann, wenn a^\downarrow und b^\downarrow auf einer Geraden liegen.

Beide Produkte haben die folgenden Eigenschaften:

1. Hält man alle Argumente bis auf eines fest, so ist das Produkt linear in diesem variablen Argument.
2. Vertauscht man 2 nebeneinander stehende Argumente, so dreht sich das Vorzeichen um.

3. Das Produkt aus normierten Vektoren einer positiv orientierten orthogonalen Basis ist gleich dem Inhalt des Einheitsquadrats bzw. des Einheitswürfels.

Dies verallgemeinern wir auf p Dimensionen.

Theorem 7.33 (Determinante)

Sei $V = \mathbb{R}^p$. Es gibt genau eine Funktion \det von $V^p = \underbrace{\mathbb{R}^p \times \mathbb{R}^p \times \dots \times \mathbb{R}^p}_{p \text{ mal}}$ in \mathbb{R} , **Determinante**

genannt, mit den folgenden Eigenschaften:

1. Es gilt

$$\begin{aligned} \det(x^\downarrow + y^\downarrow, a_2^\downarrow, \dots, a_p^\downarrow) &= \det(x^\downarrow, a_2^\downarrow, \dots, a_p^\downarrow) + \det(y^\downarrow, a_2^\downarrow, \dots, a_p^\downarrow), \\ \det(\lambda x^\downarrow, a_2^\downarrow, \dots, a_p^\downarrow) &= \lambda \det(x^\downarrow, a_2^\downarrow, \dots, a_p^\downarrow). \end{aligned}$$

Kurz: die Determinante ist linear im ersten Argument bei festgehaltenen Argumenten auf Platz 2 bis p .

2. $\det(a_1^\downarrow, a_2^\downarrow, \dots, a_{j+1}^\downarrow, a_j^\downarrow, a_{j+2}^\downarrow, \dots, a_p^\downarrow) = -\det(a_1^\downarrow, a_2^\downarrow, \dots, a_j^\downarrow, a_{j+1}^\downarrow, \dots, a_p^\downarrow)$.

Kurz: Vertauscht man 2 benachbarte Argumente, so dreht sich das Vorzeichen um.

3. Ist $e_1^\downarrow = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $e_2^\downarrow = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, \dots , $e_p^\downarrow = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$, so ist $\det(e_1^\downarrow, \dots, e_p^\downarrow) = 1$.

Den Beweis der Existenz und Eindeutigkeit dieser Funktion erhält man einfach, indem man die Funktion allein aufgrund der Eigenschaften 1) bis 3) auszurechnen versucht. Dies gelingt durch Induktion unter Beachtung des **Vorzeichens einer Permutation**, also einer Bijektion π von $\{1, \dots, p\}$ auf sich. Man setzt das **Vorzeichen (Signum)** fest durch

$$sg(\pi) = \frac{\prod_{i < j} (\pi(j) - \pi(i))}{\prod_{i < j} |\pi(j) - \pi(i)|}.$$

Dabei ist

$$\begin{aligned} \prod_{i < j} (\pi(j) - \pi(i)) &= (\pi(2) - \pi(1))(\pi(3) - \pi(1)) \cdots (\pi(p) - \pi(1)) \\ &\quad \cdot (\pi(3) - \pi(2)) \cdots (\pi(p) - \pi(2)) \\ &\quad \cdot (\pi(4) - \pi(3)) \cdots (\pi(p) - \pi(p-1)). \end{aligned}$$

Das Vorzeichen einer Permutation ist entweder $+1$ oder -1 .

Beispiele:

1. $p = 2$:

$$\pi(1) = 2, \pi(2) = 1, sg(\pi) = \frac{1-2}{|1-2|} = -1$$

2. $p = 3$

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad sg(\pi_1) = \frac{(3-2)(1-2)(1-3)}{|3-2||1-2||1-3|} = 1.$$

$$\pi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad sg(\pi_2) = \frac{(1-2)(3-2)(3-1)}{|1-2||3-2||3-1|} = -1$$

Aufgaben:

1. Berechnen Sie bitte das Vorzeichen der restlichen Permutationen der Menge $\{1, 2, 3\}$.
2. Zeigen Sie für die Permutationen der Menge $\{1, 2, 3\}$ $sg(\pi_1)sg(\pi_2) = sg(\pi_1\pi_2)$.
3. Schauen Sie in einem Lehrbuch der linearen Algebra, ob die Aussage von Aufgabe 2 auch für p statt 3 gilt. Versuchen Sie selbst einen Beweis durch Induktion nach p .

Wie oben schon ausgeführt, erhält man den Beweis des obigen Theorems durch den Beweis der folgenden Formel:

Satz 7.34 Sei $a_j^\downarrow = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{pj} \end{pmatrix}$. Dann ist die Determinante

$$\det(a_1^\downarrow, \dots, a_p^\downarrow) = \sum_{\pi \text{ Permut.}} sg(\pi) a_{\pi(1)1} a_{\pi(2)2} \cdots a_{\pi(p)p}$$

Den durch Induktion nach der Dimension p zu führenden Beweis lassen wir fort.

Aufgaben: Zeigen Sie bitte:

1. $\det\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right) = a_{11}a_{22} - a_{21}a_{12} = a_1^\downarrow * a_2^\downarrow$ (s.0.)

2. $\det\left(\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}\right) = a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{31} a_{22} a_{13} - a_{32} a_{23} a_{12} - a_{33} a_{21} a_{12}$

3. Für jede Permutation π der Menge $\{1, \dots, p\}$ ist $\det(a_{\pi(1)}^\downarrow, \dots, a_{\pi(p)}^\downarrow) = sg(\pi) \det(a_1^\downarrow, \dots, a_1^\downarrow)$.
Tipp: Jede Permutation π kann man als Produkt von zwei Vertauschungen schreiben. Verwenden Sie (ohne Beweis) $sg(\pi_1\pi_2) = sg(\pi_1)sg(\pi_2)$ (vergl. Aufgaben 2 und 3 oben) und Eigenschaft 2) der Determinante.

4. Die Determinante ist linear im j -ten Argument, wenn man alle anderen festhält.
Tipp: Benutzen Sie Aufgabe 3 und Eigenschaft 1) der Determinante.

Wir erhalten aufgrund der Eigenschaften der Determinante

Lemma 7.35 Sind zwei der Vektoren $a_1^\downarrow, \dots, a_p^\downarrow$ gleich, so ist $\det(a_1^\downarrow, \dots, a_p^\downarrow) = 0$.

Beweis: Angenommen es ist $a_1^\downarrow = a_2^\downarrow = b^\downarrow$. Nach E2 ist

$$\begin{aligned} \det(b^\downarrow, b^\downarrow, a_3^\downarrow, \dots, a_p^\downarrow) &= \det(a_2^\downarrow, a_1^\downarrow, a_3^\downarrow, \dots, a_p^\downarrow) \\ &= -\det(a_1^\downarrow, a_2^\downarrow, a_3^\downarrow, \dots, a_p^\downarrow) \\ &= -\det(b^\downarrow, b^\downarrow, a_3^\downarrow, \dots, a_p^\downarrow), \end{aligned}$$

also $\det(b^\downarrow, b^\downarrow, a_3^\downarrow, \dots, a_p^\downarrow) = 0$.

Ist nun allgemeiner $a_i^\downarrow = a_k^\downarrow$ mit $i < k$, so vertauscht man die Spalten unter Beachtung der Vorzeichen so lange, bis der i Vektor auf Platz 1 und der k -te auf Platz 2 steht. Aus dem ersten Teil des Beweises folgt die Behauptung.

Sei $\mathcal{A} = (a_1^\downarrow, \dots, a_p^\downarrow)$ eine $p \times p$ -Matrix. Dann setzen wir $\det(\mathcal{A}) = \det(a_1^\downarrow, \dots, a_p^\downarrow)$ und erhalten den **Multiplikationssatz**:

Satz 7.36 (Multiplikationssatz)

Für zwei $p \times p$ -Matrizen \mathcal{A} und \mathcal{B} gilt

$$\det(\mathcal{A}\mathcal{B}) = \det(\mathcal{A}) \det(\mathcal{B})$$

Dieser Satz folgt aus Aufgabe 4 und der Definition der Matrizenmultiplikation. Aus ihm erhalten wir die folgenden wichtigen Formeln.

Korollar 7.37 a) Eine $p \times p$ -Matrix \mathcal{A} ist genau dann invertierbar, wenn $\det(\mathcal{A}) \neq 0$ gilt. Dann ist

$$\det(\mathcal{A}^{-1}) = \frac{1}{\det(\mathcal{A})}$$

b) Sei A eine lineare Abbildung des p -dimensionalen Vektorraumes V mit Basis B in den Vektorraum W gleicher Dimension mit Basis C . A ist genau dann bijektiv, wenn für die zugehörige Matrix \mathcal{A}_A $\det(\mathcal{A}_A) \neq 0$ gilt.

Beweis: a) \mathcal{A} ist nach Definition und Satz 7.31 genau dann invertierbar, wenn es eine Matrix \mathcal{B} mit $\mathcal{A}\mathcal{B} = E_p$ gibt. Ist dies der Fall, so ist nach dem vorangehenden Satz $\det(\mathcal{A}) \det(\mathcal{B}) = \det(E_p) = 1$. Ist aber \mathcal{A} nicht invertierbar, so ist die durch \mathcal{A} auf \mathbb{R}^p gegebene lineare Abbildung $x^\downarrow \mapsto \mathcal{A}x^\downarrow$ nicht bijektiv. Damit ist nach der Dimensionsformel $\text{rng}(\mathcal{A}) < p$, also sind die Spalten $a_1^\downarrow, \dots, a_p^\downarrow$ von \mathcal{A} linear abhängig. Dann ist also für ein j $a_j^\downarrow = \sum_{k \neq j} \lambda_k a_k^\downarrow$ und damit nach Aufgabe 4 oben

$$\det(a_1^\downarrow, \dots, a_j^\downarrow, \dots, a_p^\downarrow) = \sum_{k \neq j} \lambda_k \det(a_1^\downarrow, \dots, \underbrace{a_k^\downarrow}_{j\text{-te Stelle}}, \dots, a_p^\downarrow)$$

Die rechts stehenden Determinanten haben alle a_k^\downarrow auf Platz $k \neq j$ und auf Platz j , sind also gleich 0, also ist $\det(a_1^\downarrow, \dots, a_j^\downarrow, \dots, a_p^\downarrow) = 0$.

Hieraus erhalten wir das folgende wirkungsvolle Kriterium für lineare Abhängigkeit.

Satz 7.38 p Vektoren $a_1^\downarrow, \dots, a_p^\downarrow$ in \mathbb{R}^p sind genau dann linear abhängig, wenn die Determinante

$$\det(a_1^\downarrow, \dots, a_p^\downarrow) = 0$$

ist.

Beweis: Sei $\mathcal{A} = (a_1^\downarrow, \dots, a_p^\downarrow)$. Dann ist die durch \mathcal{A} vermittelte lineare Abbildung $x^\downarrow \mapsto \mathcal{A}x^\downarrow$ des \mathbb{R}^p in sich nach der Dimensionsformel genau dann nicht bijektiv, wenn die Vektoren linear abhängig sind. Die Aussage folgt nun aus dem vorangegangenen Korollar.

Wiederholung:

- Definition der Determinante und anschauliche Bedeutung
- Rechenregeln für Determinanten
- Determinante einer linearen Abbildung
- Der Multiplikationssatz für Determinanten linearer Abbildungen
- Determinantenkriterium für die Invertierbarkeit einer linearen Abbildung

7.3.2 Lineare Gleichungssysteme

Überblick: Wir führen die Theorie linearer Gleichungssysteme auf diejenige linearer Abbildungen zurück. Für quadratische lineare Gleichungssysteme (Anzahl der Gleichungen = Anzahl der Unbekannten) erhalten wir ein Kriterium für die eindeutige Lösbarkeit mit Hilfe der Determinante der Koeffizientenmatrix. Allgemein behandeln wir das Gauß-Seidel-Verfahren zur Lösung beliebiger linearer Gleichungssysteme. Als Anwendung ergibt sich ein Verfahren zur Berechnung des Ranges einer beliebigen rechteckigen Matrix.

Allgemeine Theorie

Die ersten Gleichungssysteme kennen wir aus der Schule. Sie sind von der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Dabei sind die a_{ij} und b_k bekannte Zahlen und die x_1 und x_2 sind gesucht. Jedes Zahlenpaar (x_1, x_2) , das beide Gleichungen gleichzeitig löst, heißt Lösung des Gleichungssystems.

Zu Beginn dieses Kapitels haben wir auch andere Gleichungssysteme kennengelernt. Zum Beispiel kann eine Gerade G in der Ebene \mathbb{R}^2 durch eine Gleichung der Form $(a^\downarrow|x^\downarrow) = b$ beschrieben werden (s. Bsp 2 am Ende von Abschnitt 7.1.3). Hier haben wir *eine Gleichung* für *zwei Unbekannte*, denn die Gleichung lautet ja

$$a_1x_1 + a_2x_2 = b.$$

Ebenen im Raum werden ebenfalls durch eine Gleichung $(a^\downarrow|x^\downarrow) = b$ beschrieben. Hier jedoch sind $a^\downarrow, x^\downarrow \in \mathbb{R}^3$, die Gleichung lautet also

$$a_1x_1 + a_2x_2 + a_3x_3 = b.$$

Die Punkte der Schnittmenge zweier solcher Ebenen $(a_1^\downarrow|x_1^\downarrow) = b_1, (a_2^\downarrow|x_2^\downarrow) = b_2$ erfüllen also die *beiden* Gleichungen

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \end{aligned}$$

(hier sind $a_j^\downarrow = \begin{pmatrix} a_{j1} \\ a_{j2} \\ a_{j3} \end{pmatrix}$).

Die **allgemeine Form eines linearen Gleichungssystems (kurz: LGS)** ist also

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1q}x_q &= b_1 \\ &\vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pq}x_q &= b_p. \end{aligned}$$

Hier sind die a_{ij} und b_k bekannte Zahlen. Gesucht werden alle

q -Tupel $x^\downarrow = \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix}$, die alle Gleichungen gleichzeitig erfüllen.

Jedes solche x^\downarrow heißt **Lösung des Gleichungssystem**.

Die beiden entscheidenden Fragen sind

Existiert eine Lösung?

Ist die Lösung eindeutig bestimmt, d.h. existiert genau eine Lösung?

Wir setzen $a_k^\downarrow = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{pk} \end{pmatrix}$, $\vec{a}_\ell = (a_{\ell 1}, \dots, a_{\ell q})$

und $\mathcal{A} = (a_1^\downarrow, \dots, a_q^\downarrow) = \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_p \end{pmatrix}$, ferner $b^\downarrow = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$.

Dann lautet das Gleichungssystem $\mathcal{A}x^\downarrow = b^\downarrow$ (**Matrixform**)

oder $x_1a_1^\downarrow + \dots + x_qa_q^\downarrow = b^\downarrow$ (**Spaltenform**)

Aus der Spaltenform folgt sofort der folgende Satz.

Satz 7.39 (Existenz einer Lösung eines LGS)

Das gegebene Gleichungssystem hat genau dann mindestens eine Lösung, wenn der Rang von \mathcal{A} gleich dem Rang der erweiterten Matrix $(\mathcal{A}, b^\downarrow) = (a_1^\downarrow, \dots, a_q^\downarrow, b^\downarrow)$ ist.

Beweis: Ist $\begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix}$ eine Lösung, so ist b^\downarrow linear abhängig von $a_1^\downarrow, \dots, a_q^\downarrow$, der Rang erhöht sich also

nicht bei Hinzufügung von b^\downarrow . Erhöht sich umgekehrt der Rang nicht, so ist b^\downarrow linear abhängig von $a_1^\downarrow, \dots, a_q^\downarrow$.

Beispiele:

1. Betrachten Sie das Gleichungssystem.

$$5x + 4y = 1$$

$$4x + 5y = 10$$

Lösen Sie es wie in der Schule und zeigen Sie

$$\text{rng}\left(\begin{pmatrix} 5 & 4 \\ 4 & -5 \end{pmatrix}\right) = \text{rng}\left(\begin{pmatrix} 5 & 4 & 1 \\ 4 & 5 & 10 \end{pmatrix}\right).$$

2. Betrachten Sie das Gleichungssystem

$$2x - 4y = 6$$

$$2x - 4y = 0.$$

Es handelt sich um die Beschreibung zweier paralleler Geraden, die keinen Punkt gemeinsam haben, nämlich die Geraden beschrieben durch $y = \frac{1}{2}x$ und $y = \frac{1}{2}x - 3$. Zeigen Sie $\text{rng}\left(\begin{pmatrix} 2 & -4 \\ 2 & -4 \end{pmatrix}\right) \neq$

$$\text{rng}\left(\begin{pmatrix} 2 & -4 & 6 \\ 2 & -4 & 0 \end{pmatrix}\right).$$

Die Matrixform des Gleichungssystems läßt eine andere Interpretation zu:

$$\mathbb{R}^q \ni x^\downarrow \mapsto Ax^\downarrow \in \mathbb{R}^p$$

ist eine lineare Abbildung $A_{\mathcal{A}}$.

Ist $b^\downarrow = 0$, so heißt das Gleichungssystem **homogen**, sonst **inhomogen**.

Aus der Interpretation der Matrixform als linearer Abbildung erhalten wir

Theorem 7.40 a) Die Menge $L_0(\mathcal{A})$ der Lösungen des homogen Systems $Ax^\downarrow = 0^\downarrow$ bildet einen Untervektorraum der Dimension $q - \text{rng}(\mathcal{A})$ von \mathbb{R}^q .

b) Ist x_0^\downarrow eine willkürlich gewählte Lösung des inhomogenen Systems $Ax^\downarrow = b^\downarrow$, so erhält man alle Lösungen dieses Systems durch Addition von Lösungen des homogenen Systems, kurz:

$$L_{b^\downarrow}(\mathcal{A}) = \{x^\downarrow : Ax^\downarrow = b^\downarrow\} = \{x_0^\downarrow + y^\downarrow : Ay^\downarrow = 0^\downarrow\}$$

$L_0(\mathcal{A})$ ist nichts anderes als der Kern $\ker(A_{\mathcal{A}})$. Der Rest des Beweises ergibt sich sofort durch Matrizenrechnung. Damit haben wir die abstrakten Antworten auf unsere beiden Fragen nach der Lösung:

Korollar 7.41 (Existenz und Eindeutigkeit der Lösung eines LGS)

a) Die beiden folgenden Aussagen sind äquivalent:

- i) Für jede rechte Seite $b^\downarrow \in \mathbb{R}^p$ hat das Gleichungssystem $Ax^\downarrow = b^\downarrow$ mindestens eine Lösung
- ii) Der Rang $\text{rng}(\mathcal{A})$ von \mathcal{A} ist gleich p .

b) Die folgenden Aussagen sind äquivalent:

- i) Für jede rechte Seite $b^\downarrow \in \mathbb{R}^p$ hat das Gleichungssystem höchstens eine Lösung
- ii) Die homogene Gleichung $Ax^\downarrow = 0^\downarrow$ hat nur die Nulllösung.
- iii) $\text{rng}(\mathcal{A}) = q$.

c) Die beiden folgenden Aussagen sind äquivalent:

- i) Für jede rechte Seite $b^\downarrow \in \mathbb{R}^p$ hat das Gleichungssystem $Ax^\downarrow = b^\downarrow$ genau eine Lösung.
- ii) Es ist $p = q$ und die Determinante $\det(\mathcal{A}) \neq 0$.

Beweis:

a) i) gilt genau dann, wenn $A_{\mathcal{A}}$ surjektiv, also

$$\text{rng}(\mathcal{A}) = \dim(A_{\mathcal{A}}(\mathbb{R}^q)) = \dim(\mathbb{R}^p) = p$$

ist.

b) folgt unmittelbar aus der Aussage b) des vorausgegangenen Theorems.

c) i) gilt genau dann, wenn a) i) und b) i) gilt. Das ist genau dann der Fall, wenn $\text{rng}(\mathcal{A}) = p$ und $0 = \dim(L_0(\mathcal{A})) = \dim(\ker(A_{\mathcal{A}}))$ ist. Nach der Dimensionsformel gilt dies genau dann, wenn $q = p$ und $\text{rng}(\mathcal{A}) = p$, also $\det(\mathcal{A}) \neq 0$ ist.

Gauß-Seidel-Verfahren zur Lösung

Sei $\mathcal{A} = (a_{ik})$ eine $p \times q$ -Matrix und $b^{\downarrow} \in \mathbb{R}^p$ beliebig. Um das Gleichungssystem $\mathcal{A}x^{\downarrow} = b^{\downarrow}$ zu lösen, benutzen wir die folgenden Manipulationen, von denen die ersten drei die Lösungsmenge $L_{b^{\downarrow}}(\mathcal{A})$ nicht ändern. Genauer bedeutet dies: jede dieser ersten drei Manipulation liefert ein neues Gleichungssystem $\mathcal{A}'x^{\downarrow} = b'^{\downarrow}$ mit $L_{b'^{\downarrow}}(\mathcal{A}') = L_{b^{\downarrow}}(\mathcal{A})$:

1. Vertauschung von Zeilen
2. Multiplikation einer Zeile mit einer Zahl $\neq 0$
3. Addition eines Vielfachen einer Zeile zu einer anderen
4. Vertauschung zweier Spalten. Das bedeutet die Umnummerierung (Vertauschung) der entsprechenden Unbekannten, über die wir genau Buch führen müssen, um aus den Lösungen des veränderten Gleichungssystems die des ursprünglichen zu finden.

Beispiel:

$$\begin{aligned} x_2 - x_3 &= 1 \\ x_1 - x_2 + 2x_3 &= 4 \\ x_1 + x_2 &= -2. \end{aligned}$$

Hier ist $\mathcal{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 0 \end{pmatrix}$, also $x_1 a_1^{\downarrow} + x_2 a_2^{\downarrow} + x_3 a_3^{\downarrow} = \begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}$.

Vertauschen wir nun die zweite mit der ersten Spalte, setzen also $a_1^{\downarrow} = a_2^{\downarrow}$, $a_2^{\downarrow} = a_1^{\downarrow}$, so erhalten wir

$\mathcal{A}' = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix}$. $x'^{\downarrow} = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$ ist genau dann Lösung des Systems $\mathcal{A}'x^{\downarrow} = b^{\downarrow}$, wenn $x^{\downarrow} = \begin{pmatrix} x'_2 \\ x'_1 \\ x'_3 \end{pmatrix}$ Lösung des ursprünglichen Systems ist.

Nun führen wir schrittweise nach Bedarf die einzelnen Manipulationen durch:

Gauß-Seidel-Verfahren:

1. Schritt:

a) Ist $\mathcal{A} = 0_{pq}$ (die Matrix mit lauter Nullen), so ist $L_0(\mathcal{A}) = \mathbb{R}^q$, und für $b^\downarrow \neq 0$ ist das System nicht lösbar, d.h. $L_{b^\downarrow}(\mathcal{A}) = \emptyset$.

b) Sei $\mathcal{A} \neq 0$, etwa $a_{ij} \neq 0$. Dann vertauscht man zunächst die i -Zeile mit der ersten Zeile und dann die j -Spalte mit der ersten Spalte. Man erhält eine neue Matrix

$$\mathcal{A}' \text{ mit } a'_{11} = a_{ij} \neq 0.$$

In der Praxis sucht man das betragsmäßig größte a_{ij} , auch wenn $a_{11} \neq 0$ war (Pivot-Element) und führt die Vertauschung durch, damit a'_{11} betragsmäßig besonders groß ist, wegen des folgenden Schrittes:

Nun setzt man für $k \geq 2$

$$\vec{a}_k^{(1)} = \vec{a}'_k - \frac{a'_{k1}}{a_{k1}} \vec{a}'_1, \quad b_k^{(1)} = b'_k - \frac{a'_{k1}}{a_{k1}} b_k$$

und erhält $a_{k1}^{(1)} = 0$, insgesamt also

$$\mathcal{A}^{(1)} = \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1q} \\ 0 & a_{22}^{(1)} & \cdots & a_{2q}^{(1)} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & a_{p2}^{(1)} & \cdots & a_{pq}^{(1)} \end{pmatrix} \text{ und } \vec{b}^{(1)}.$$

Der Übersichtlichkeit wegen schreibt man $a'_{1k} = a_{1k}^{(1)}$, obwohl man ja hier nichts abgezogen hat.

2. Schritt:

Man betrachtet nun die Restmatrix $\mathcal{B} = \begin{pmatrix} a_{22}^{(1)} & \cdots & a_{2q}^{(1)} \\ \vdots & & \vdots \\ a_{p2}^{(1)} & \cdots & a_{pq}^{(1)} \end{pmatrix}$

a) Ist $\mathcal{B} = 0$, so bricht das Verfahren ab. Das neue Gleichungssystem $\mathcal{A}^{(1)} x^{\downarrow(1)} = b^\downarrow$ ist genau dann lösbar, wenn $b_2 = \cdots = b_p = 0$ ist. Dann ist $x_1^{(1)} = \frac{b_1}{a_{11}^{(1)}}$, $x_2^{(1)}, \dots, x_q^{(1)} \in \mathbb{R}$ beliebig.

b) Ist $\mathcal{B} \neq 0$, so gibt es ein $a_{ik}^{(1)} \neq 0$, wo $i, k \geq 2$. Man vertauscht um erst die 2 -te und die i -te Zeile von $\mathcal{A}^{(1)}$ (nicht von \mathcal{B}) und dann die k -te und 2 . Spalte von $\mathcal{A}^{(1)}$ und erhält eine neue Matrix $\mathcal{A}^{(1)}$ mit $a_{22}^{(1)} \neq 0$. Auch hier gilt: man sucht in der Praxis das betragsmäßig größte $a_{ik}^{(1)}$, selbst wenn $a_{22}^{(1)}$ schon ungleich 0 war.

Nun setzt man $\vec{a}_k^{(2)} = \vec{a}_k^{(1)'} - \frac{a_{k2}^{(1)'}}{a_{22}^{(1)'}} \vec{a}_2^{(1)'}$ für $k \geq 2$, ferner $b_k^{(2)} = b_k^{(1)} - \frac{a_{k2}^{(1)'}}{a_{22}^{(1)'}} b_2^{(1)'}$

und erhält

$$\mathcal{A}^{(2)} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1q}^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2q}^{(2)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3q}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{p3}^{(2)} & \cdots & a_{pq}^{(2)} \end{pmatrix}$$

sowie $b^{\downarrow(2)}$.

bf 3. bis p. Schritt:

Sie laufen analog zum 2. Schritt und zwar maximal bis zum p -Schritt, falls vorher kein Abbruch war. Bevor wir Schlußfolgerungen ziehen, demonstrieren wir das Verfahren an drei Beispielen.

Beispiele:

1.

$$\begin{aligned}2x_1 + x_2 &= 1 \\ x_1 - x_2 &= 4\end{aligned}$$

Matrix $\mathcal{A} = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$, $b^\downarrow = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ $a_n = 2$, man zieht das $\frac{1}{2}$ -fache der 1. Zeile von der zweiten ab: es ergibt sich:

$$\mathcal{A}^{(1)} = \begin{pmatrix} 2 & 1 \\ 0 & -\frac{3}{2} \end{pmatrix}, \quad b^{\downarrow(1)} = \begin{pmatrix} 1 \\ 7/2 \end{pmatrix}$$

Hieraus erhält man $x_2 = -7/3$, $x_1 = \frac{1}{2}(1 + \frac{7}{3}) = \frac{5}{3}$

2.

$$\begin{aligned}2x_1 + x_2 - x_3 &= 1 \\ x_1 - x_2 + x_3 &= 4\end{aligned}$$

Es ist $\mathcal{A} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$, $b^\downarrow = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$

Da $a_{11} = 2 \neq 0$, verfährt man wie oben und erhält

$$\mathcal{A}^{(1)} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & -\frac{3}{2} & \frac{3}{2} \end{pmatrix}, \quad b^{\downarrow(1)} = \begin{pmatrix} 1 \\ 7/2 \end{pmatrix}$$

und damit

$$\begin{aligned}x_2 &= -\frac{2}{3}\left(\frac{7}{2} - \frac{3}{2}x_3\right) = -\frac{7}{3} + x_3 \\ x_1 &= \frac{1}{2}\left(1 + \frac{7}{3} - x_3 + x_3\right) = \frac{5}{3}\end{aligned}$$

3.

$$\begin{aligned}2x_1 + x_2 - x_3 &= 1 \\ x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= 2\end{aligned}$$

$\mathcal{A} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}$, $b^\downarrow = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

Nach dem ersten Schritt ist

$$\mathcal{A}^{(1)} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad b^{\downarrow(1)} = \begin{pmatrix} 1 \\ 3/2 \end{pmatrix}$$

Das Gleichungssystem ist nicht lösbar. Hier wird $a_{22}^{(1)} = a_{23}^{(1)} = 0$.

4.

$$\begin{aligned}2x_1 + x_2 &= 1 \\ x_1 - x_2 &= 2 \\ x_1 + 2x_2 &= 3\end{aligned}$$

$$\mathcal{A} = \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 2 \end{pmatrix}, \quad b^\downarrow = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$a_{11} = 2, \text{ also } \mathcal{A}^{(1)} = \begin{pmatrix} 2 & 1 \\ 0 & -3/2 \\ 0 & 3/2 \end{pmatrix}, \quad b^{\downarrow(1)} = \begin{pmatrix} 1 \\ 3/2 \\ 5/2 \end{pmatrix}$$

$$\mathcal{A}^{(2)} = \begin{pmatrix} 2 & 1 \\ 0 & -3/2 \\ 0 & 0 \end{pmatrix}, \quad b^{\downarrow(2)} = \begin{pmatrix} 1 \\ 3/2 \\ 4 \end{pmatrix}$$

Dieses Gleichungssystem ist also nicht lösbar.

Wir kehren zum allgemeinen Fall zurück. Wir enden je nachdem ob $p \leq q$ oder $p > q$ ist:

1. Fall: $p \leq q$.

a) Das Verfahren bricht nicht vor p Schritten ab. Dann erhält man $\mathcal{A}^{(r)}x^{\downarrow(r)} = b^{\downarrow(r)}$ mit

$$\mathcal{A}^{(r)} = \begin{pmatrix} a_{11}^{(r)}, & a_{12}^{(r)}, & a_{13}^{(r)}, & \cdots, & a_{1p}^{(r)}, & \cdots, & a_{1q}^{(r)} \\ 0, & a_{22}^{(r)}, & a_{23}^{(r)}, & \cdots, & a_{2q}^{(r)}, & \cdots, & a_{2q}^{(r)} \\ 0, & 0, & a_{33}^{(r)}, & \cdots, & a_{3p}^{(r)}, & \cdots, & a_{3q}^{(r)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0, & 0, & 0, & 0, & a_{pp}^{(r)}, & \cdots, & a_{pq}^{(r)} \end{pmatrix}$$

mit $a_{kk}^{(r)} \neq 0$ für alle k . Dann kann man die Lösungen sukzessive berechnen

$$x_p^{(r)} = \frac{1}{a_{pp}^{(r)}}(b_p^{(r)} - \sum_{k=p+1}^q a_{pk}^{(r)}x_k)$$

$$x_{(p-1)}^{(r)} = \frac{1}{a_{p-1,p-1}^{(r)}}(b_{p-1}^{(r)} - \sum_{k=p}^q a_{p-1,k}^{(r)}x_k)$$

$$\vdots$$

b) Das Verfahren bricht vorzeitig ab, weil die Restmatrix 0 ist. Dann gibt es ein $t, 0 \leq t < p$ mit $a_{kk}^{(r)} \neq 0$ für $k \leq t$ ($t = 0$: \mathcal{A} ist 0-Matrix) und

$$\mathcal{A}^{(r)} = \begin{pmatrix} a_{11}^{(r)}, & a_{12}^{(r)}, & \cdots, & a_{1t}^{(r)}, & \cdots, & a_{1q}^{(r)} \\ 0, & a_{22}^{(r)}, & \cdots, & a_{2t}^{(r)}, & \cdots, & a_{2q}^{(r)} \\ 0, & 0, & & \vdots & & \\ 0, & 0, & \cdots, & a_{tt}^{(r)}, & \cdots, & a_{tq}^{(r)} \\ \vdots & \vdots & & 0, & & 0 \\ 0, & 0, & \cdots, & 0, & \cdots, & 0 \end{pmatrix}$$

und das Gleichungssystem ist genau dann lösbar, wenn $b_k^{(r)} = 0$ für $k \geq t + 1$. Dann kann man $x_k^{(r)}$ für $k \geq t + 1$ beliebig wählen und dann nach $x_t^{(r)}, x_{t-1}^{(r)}, \dots, x_1^{(r)}$ auflösen. 2. Fall: $p > q$.

Hier bricht das Verfahren spätestens nach q Schritten ab. Man erhält

$$\mathcal{A}^{(r)} = \begin{pmatrix} a_{11}^{(r)} & \cdots & a_{1q}^{(r)} \\ 0 & & \\ 0 & \cdots & a_{qq}^{(r)} \\ 0 & \cdots & 0 \\ \vdots & & \\ 0 & \cdots & 0 \end{pmatrix}$$

und die Lösbarkeit ist leicht zu entscheiden.

Aufgaben: Lösen Sie bitte die folgenden Gleichungssysteme:

1.

$$\begin{aligned} -x_1 + 2x_2 + x_3 &= -2 \\ 3x_1 - 8x_2 - 2x_3 &= 4 \\ x_1 + 4x_3 &= -2. \end{aligned}$$

2.

$$\begin{aligned} 5x_1 - 2x_2 + x_3 &= 7 \\ x_1 + x_2 - x_3 &= 6 \end{aligned}$$

Interpretieren Sie dieses Gleichungssystem geometrisch!

3.

$$\begin{aligned} \lambda x_1 - x_2 &= 0 \\ -x_1 + \lambda x_2 &= 0 \end{aligned}$$

Achtung: Für welche λ gibt es außer der 0-Lösung weitere Lösungen? Es handelt sich um ein sogenanntes Eigenwertproblem, siehe unten.

4. Schreiben Sie ein Programm für das Verfahren für $p = q = 3$.

Zeilen- und Spaltenrang einer Matrix

Sei \mathcal{A} eine $p \times q$ -Matrix und $1 \leq r \leq \min(p, q)$.

Suchen wir r Spalten aus, etwa

$a_{i_1}^\downarrow, \dots, a_{i_r}^\downarrow$ mit $i_1 < i_2 < \dots < i_r$, so erhalten wir eine $p \times r$ -Matrix. Wählen wir in dieser noch einmal r Zeilen aus etwa $j_1 < j_2 < \dots < j_r$, so erhalten wir eine $r \times r$ -Teilmatrix \mathcal{B} von \mathcal{A}

$$\mathcal{B} = \begin{pmatrix} a_{j_1 i_1} & \cdots & a_{j_1 i_r} \\ \vdots & & \\ a_{j_r i_1} & \cdots & a_{j_r i_r} \end{pmatrix}.$$

Aufgaben:

1. Sei $\mathcal{A} = \begin{pmatrix} 1 & 2 \\ 4 & 4 \\ 6 & 6 \end{pmatrix}$. Bestimmen Sie alle 2×2 -Teilmatrizen und berechnen Sie deren Determinante.
2. Wieviel 4×4 -Teilmatrizen hat eine 6×8 -Matrix?
3. Wieviel $r \times r$ -Teilmatrizen besitzt eine $p \times q$ -Matrix?

Sei $\mathcal{A} = (a_{ik})$ eine $p \times q$ -Matrix mit den Spalten $a^{\downarrow j} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{pj} \end{pmatrix}$ bzw. den Zeilen $\vec{a}_k = (a_{k1}, \dots, a_{kq})$.

Der Rang von A war definiert als die Maximalzahl linear unabhängiger *Spalten* ("Spaltenrang"). Wir schreiben für den Moment hierfür $rng_s(\mathcal{A})$. Ebenso kann man natürlich nach der Maximalzahl linear unabhängiger *Zeilen* fragen. Diese Zahl heißt **Zeilenrang** $rng_z(\mathcal{A})$. Schließlich sei $r_d(\mathcal{A})$ das Maximum aller Zahlen r mit der Eigenschaft: Es gibt eine $r \times r$ -Teilmatrix $\mathcal{A}_{r,r}$ mit $\det(\mathcal{A}_{r,r}) \neq 0$. Offensichtlich ist $0 \leq r_d(\mathcal{A}) \leq \min(p, q)$.

Beispiele:

1. $\mathcal{A} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$, $r_d(\mathcal{A}) = 0$
2. $\mathcal{A} = (5, 2, 0)$, $r_d(\mathcal{A}) = 1$
3. $\mathcal{A} = \begin{pmatrix} 5, & 2 \\ 7, & 6 \\ 0 & 1 \end{pmatrix}$, $r_d(\mathcal{A}) = 2$
4. $\mathcal{A} = \begin{pmatrix} 5, & 2, & 0 \\ 0, & 0, & 0 \\ 7, & 6, & 0 \end{pmatrix}$. Auch hier ist $r_d(\mathcal{A}) = 2$. Die 2×2 -Teilmatrix ist $\begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix}$.

Teilmatrizen müssen nicht aus aufeinander folgenden Zeilen und Spalten bestehen. Man kann ganz beliebig herausschneiden!

Wie reagieren die Zahlen $r_d(\mathcal{A})$ und $rng_s(\mathcal{A})$ auf die beim Gauß-Seidel-Verfahren angewendeten Schritte?

Lemma 7.42 *Weder $rng_s(\mathcal{A})$ noch $r_d(\mathcal{A})$ noch $rng_z(\mathcal{A})$ ändern sich bei den folgenden Operationen:*

1. *Vertauschung zweier Zeilen*
2. *Multiplikation einer Zeile mit einer Zahl $\neq 0$*
3. *Addition der k -ten Zeile zur j -ten Zeile*
4. *Vertauschung zweier Spalten.*

Der Beweis ist offensichtlich.

Damit können wir nun zeigen

Satz 7.43 *Für eine $p \times q$ -Matrix sind Spaltenrang und Zeilenrang gleich und gleich $r_d(\mathcal{A})$, kurz:*

$$rng_s(\mathcal{A}) = rng_z(\mathcal{A}) = r_d(\mathcal{A})$$

Beweis: Wir können \mathcal{A} durch die Umformungen auf die Gestalt $C = \begin{pmatrix} c_{11} & & * \\ 0 & & \\ 0 & \cdots & c_{tt} & \cdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$

bringen ohne $r_d(\mathcal{A})$ und $rng_s(\mathcal{A})$ zu ändern. Es ist aber $r_d(C) = t$ und $rng_s(C) = t$, also ist $r_d(\mathcal{A}) = r_d(C) = rng_s(C) = rng_s(\mathcal{A})$. Der Zeilenrang von C ist aber nichts anderes als der Spaltenrang von

$C^T = \begin{pmatrix} c_{11} & \cdots & c_{p1} \\ c_{12} & \cdots & c_{p2} \\ c_{1q} & \cdots & c_{pq} \end{pmatrix}$, das ist die Matrix, die man aus C erhält, indem man die Indizes vertauscht:

man erhält $C^T = \begin{pmatrix} c_{11} & 0, & \cdots & 0, & 0 & \cdots & 0 \\ * & & & & & & \\ & & & c_{tt}, \cdots & 0, & \cdots & 0 \end{pmatrix}$.

Daraus ergibt sich

$$rng_z(\mathcal{A}) = rng_z(C) = rng_s(C^T) = r_d(C^T) = r_d(C) = rng_s(\mathcal{A}).$$

7.4 Spektraltheorie

Einführung

Wir hatten am Anfang dieses Kapitels gesehen, daß eine Drehung im Raum immer um eine Achse senkrecht zur Drehebene erfolgt. Ist $D : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ diese Drehung, so gilt also für einen Vektor y^\perp senkrecht zur Drehebene stets $Dy^\perp = y^\perp$.

Bei Projektionen P auf eine Ebene werden alle Vektoren der Projektionsebene festgelassen, es gilt also $Px^\perp = x^\perp$ für alle x^\perp aus der Ebene.

Betrachten wir schließlich eine Spiegelung S an einer Ebene E , so gilt für jeden Vektor $x^\perp \perp E$ stets $Sx^\perp = -x^\perp$. Für $y^\perp \in E$ hingegen ist $Sy^\perp = y^\perp$. Für das Rechnen mit einer solchen linearen Abbildung ist natürlich klar, daß man eine Basis aus solchen *geeigneten Vektoren* sucht. Eine Spiegelung S an einer Ebene E zum Beispiel läßt sich dann so beschreiben: wähle $u^\perp, v^\perp \in E$, $u^\perp \perp v^\perp$ und setze $w^\perp = u^\perp \times v^\perp$. Dann ist $\{u^\perp, v^\perp, w^\perp\}$ eine Basis B und für die Matrixdarstellung von S bezüglich dieser

Basis gilt $\mathcal{A}_s = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

In der klassischen Mechanik werden verschiedene Größen als "Tensoren" dargestellt, das sind 3×3 -

Matrizen $\mathcal{A} = (a_{ik})$ mit $a_{ik} = a_{ki}$. Solche Matrizen heißen symmetrisch. Zum Beispiel ist $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$

symmetrisch.

Die Menge $\{x^\perp : (x^\perp | \mathcal{A}x^\perp) = 1\}$ nennt man eine **Quadrik**. Zu jeder symmetrischen Matrix \mathcal{A} gibt es nun eine Basis $\{b_1^\perp, b_2^\perp, b_3^\perp\}$ in \mathbb{R}^3 mit $(b_1^\perp | b_1^\perp \times b_2^\perp) = 1$, $b_3^\perp = b_1^\perp \times b_2^\perp$ schließlich $b_1^\perp \perp b_2^\perp$ und $\|b_j^\perp\| = 1$, sowie $\mathcal{A}b_j^\perp = \lambda_j b_j^\perp$ $\lambda_j \in \mathbb{R}$. Ist dann x^\perp in der neuen Basis dargestellt als $x^\perp = \xi_1 b_1^\perp + \xi_2 b_2^\perp + \xi_3 b_3^\perp$, so hat die Gleichung $(x^\perp | \mathcal{A}x^\perp) = 1$ die besonders einfache Form

$$\lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \lambda_3 \xi_3^2 = 1.$$

Die Vektoren $b_1^\perp, b_2^\perp, b_3^\perp$ nennt man auch **Hauptachsen des Tensors** bzw. der Quadrik. Für $\lambda_1 = \lambda_2 = \lambda_3 = 1$ erhält man die **Kugelfläche**, für $\lambda_j > 0$ für alle j ein **Ellipsoid**.

Appletanleitung Geben Sie verschiedene Werte $\lambda_1, \lambda_2, \lambda_3$ ein (auch negative!) und schauen Sie sich an, wie sich die Quadriken ändern in Abhängigkeit von $\lambda_1, \lambda_2, \lambda_3$.

* * * TO DO * * *

Die allgemeine Fragestellung

Sei im folgenden A eine lineare Abbildung des Vektorraumes V in sich.

Definition 7.44 Eine Zahl $\lambda \in \mathbb{R}$ heißt **Eigenwert** von A , wenn es ein $x \neq 0$ in V gibt mit $Ax = \lambda x$. x heißt dann ein zu λ gehöriger **Eigenvektor**. Die Menge $\sigma(A)$ aller Eigenwerte von A heißt **Spektrum** von A .

Beispiele:

1. $\mathcal{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ $Ax^\downarrow := \mathcal{A}x^\downarrow$. Eigenwerte sind 1 und -1 Eigenvektoren $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. (Es gibt mehr Eigenvektoren, s.u.)
2. $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $Ax^\downarrow = \mathcal{A}x^\downarrow$. Dies beschreibt die Spiegelung an der Hauptdiagonalen. Eigenwerte sind ± 1 (vergl. Aufgabe 19), mögliche Eigenvektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (zu $\lambda = 1$) und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ (zu $\lambda = -1$)
3. $\mathcal{A} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$, $Ax^\downarrow := \mathcal{A}x^\downarrow$. Das beschreibt die Drehung in der Ebene um den Winkel φ . Für welche φ gibt es Eigenwerte?

Mit I bezeichnen wir die identische Abbildung auf V , also $I(v) = v$. Damit ergibt sich leicht der folgende Satz:

Satz 7.45 Sei $A : V \rightarrow V$ eine lineare Abbildung. Dann gilt

- a) λ ist genau dann ein Eigenwert, wenn der Kern der linearen Abbildung $\lambda I - A$ nicht nur aus der 0 besteht, d.h. $\dim \ker(\lambda I - A) \geq 1$ gilt.
- b) $\ker(\lambda I - A) \setminus \{0\}$ ist die Menge aller zum Eigenwert λ gehörenden Eigenvektoren.

Ist λ ein Eigenwert, so nennt man $\ker(\lambda I - A)$ den zugehörigen **Eigenraum**.

Ist V endlich-dimensional, so können wir die Eigenwerte auf folgende Art bestimmen:

Wir führen eine Basis $B = \{b_1, \dots, b_p\}$ ein ($p = \dim(V)$). Dann ist $\mathcal{A}_I = E_p$ (die $p \times p$ -Einheitsmatrix) und wir erhalten:

Satz 7.46 λ ist genau dann Eigenwert von A , wenn $\det(\lambda E_p - \mathcal{A}_A) = 0$ gilt. Die Funktion $\lambda \mapsto \det(\lambda E_p - \mathcal{A}_A)$ ist ein Polynom von Grad p , das **charakteristische Polynom** von A . Es hängt nicht von der gewählten Basis ab.

Beweis: (I) λ ist genau dann Eigenwert, wenn es ein $x \neq 0$ mit $Ax = \lambda x$ gibt. Für den Koordinatenvektor x^\downarrow bezüglich der Basis B bedeutet dies $\mathcal{A}_A x^\downarrow = \lambda x^\downarrow$ oder $(\lambda E_p - \mathcal{A}_A)x^\downarrow = 0^\downarrow$, d.h. x ist im Kern der linearen Abbildung $\lambda I - A$. Das bedeutet $\dim L_0(\lambda E_p - \mathcal{A}_A) \geq 1$ und das ist nach Korollar 7.37 b) genau

dann der Fall, wenn $\det(\lambda E_p - \mathcal{A}_A) = 0$ gilt.

(II) Aufgrund der Entwicklungsformel für die Determinante folgt

$$0 = \det(\lambda E_p - \mathcal{A}_A) = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{pp}) + \sum_{id \neq \pi} \text{sg}(\pi) \prod_k b_{\pi(k)k}(\lambda)$$

wo $b_{ij}(\lambda)$ das (i, j) -Element der Matrix $(\lambda E_p - \mathcal{A}_A)$ ist, also $b_{ij}(\lambda) = \begin{cases} \lambda - a_{ij} & i = j \\ -a_{ij} & i \neq j \end{cases}$

und id die identische Abbildung auf $\{1, \dots, p\}$ ist.

Ist $id \neq \pi$ eine Permutation der Menge $\{1, \dots, p\}$, so kommt im Produkt $\prod b_{\pi(k)k}(\lambda)$ ein Faktor der Form $(\lambda - a_{ii})$ höchstens $p - 1$ mal vor, also ist das Produkt ein Polynom vom Grade $\leq p - 1$. Rechnet man $\prod_{i=1}^p (\lambda - a_{ii})$ aus so erhält man $\prod_{i=1}^p (\lambda - a_{ii}) = \lambda^p - \lambda^{p-1} \sum a_{ii} \pm \dots + (-1)^p a_{11} \cdots a_{pp}$. Also ist das Produkt ein Polynom vom Grad p (mit höchstem Koeffizienten 1).

(III) Wechseln wir die Basis, so erhalten wir eine neue Matrix-Darstellung \mathcal{A}'_A von A , und es gilt $\mathcal{A}'_A = S_B^{-1} \mathcal{A}_A S_B$ (s. Satz 7.32), also

$\lambda E_p - \mathcal{A}'_A = S_B^{-1} (\lambda E_p - \mathcal{A}_A) S_B$ und damit folgt aus dem Determinantensatz

$$\begin{aligned} \det(\lambda E_p - \mathcal{A}'_A) &= \det(S_B^{-1}) \det(\lambda E_p - \mathcal{A}_A) \det(S_B) \\ &= \det(S_B^{-1}) \det(S_B) \det(\lambda E_p - \mathcal{A}_A) \\ &= \det(S_B^{-1} S_B) \det(\lambda E_p - \mathcal{A}_A) \\ &= \det(\lambda E_p - \mathcal{A}_A). \end{aligned}$$

Leider gibt es Polynome mit reellen Koeffizienten, die in \mathbb{R} keine Nullstelle haben. Hier ist ein Beispiel für ein charakteristisches Polynom:

Beispiel: Sei D die Drehung um $\pi/2$ in der Ebene. D hat keinen Eigenvektor, alle Vektoren werden ja gedreht. Hier überzeugen wir uns auch rein formal davon: wir legen die kanonische Basis $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ zugrunde. Dann ist $\mathcal{A}_D = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\lambda E_2 - \mathcal{A}_D = \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix}$ und damit $\det(\lambda E_2 - \mathcal{A}_D) = \lambda^2 + 1$.

Aufgaben: Berechnen Sie die charakteristischen Polynome der folgenden Matrizen (aufgefaßt als lineare Abbildungen) und versuchen Sie, die Eigenwerte zu bestimmen!

1. $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

2. $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$

3. $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ -1 & \alpha \end{pmatrix}$. Bestimmen Sie die Eigenwerte als Funktion von α und diskutieren Sie die Funktion.

4. $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Bestimmen Sie hierzu auch die Eigenräume. Wieviele sind es? Welche Dimension haben sie?

5. $\mathcal{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Zeigen Sie, daß es 2 verschiedene oder eine doppelte reelle Nullstelle gibt.

Die Lösung in den komplexen Zahlen

In der ganzen bisherigen linearen Algebra haben wir (mit Ausnahme des Skalarprodukts) nie davon Gebrauch gemacht, daß wir speziell \mathbb{R} als Körper gewählt hatten. Sämtliche, wirklich sämtliche Definitionen kann man sinnvoll auch für beliebige Körper statt \mathbb{R} formulieren, insbesondere für \mathbb{C} . Sämtliche Sätze gelten dann auch für den entsprechenden Körper. Im folgenden benutzen wir den Körper \mathbb{C} . Starten wir mit einer reellen Matrix, etwa $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, so kann man natürlich auch die Multiplikation mit **komplexen Vektoren** betrachten.

Zum Beispiel ist $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1+i \\ 1-i \end{pmatrix} = \begin{pmatrix} 1-i \\ -(1+i) \end{pmatrix}$, denn jeder komplexe Vektor $z^\downarrow = \begin{pmatrix} z_1 \\ \vdots \\ z_p \end{pmatrix}$

ist Summe aus reellen Vektoren: $z^\downarrow = \begin{pmatrix} \Re z_1 & + & i\Im z_1 \\ \vdots & & \vdots \\ \Re z_p & + & i\Im z_p \end{pmatrix} = \Re(z^\downarrow) + i\Im(z^\downarrow)$ und damit ergibt sich

$$\mathcal{A}z^\downarrow = \mathcal{A}\Re(z^\downarrow) + i\mathcal{A}\Im(z^\downarrow).$$

In \mathbb{C} haben wir den Fundamentalsatz der Algebra zur Verfügung.

Theorem 7.47 (Fundamentalsatz der Algebra)

Sei $P(z) = z^n + a_1z^{n-1} + \dots + a_n$. Dann gibt es n komplexe Zahlen $\lambda_1, \dots, \lambda_n$ (die nicht notwendig verschieden sein müssen) mit

$$P(z) = (z - \lambda_1) \cdots (z - \lambda_n).$$

Korollar 7.48 Jede lineare Abbildung $A : \mathbb{C}^p \rightarrow \mathbb{C}^p$ hat p nicht notwendig verschiedene Eigenwerte.

Damit können wir unser Problem in \mathbb{C}^p lösen. Zunächst ein Beispiel.

Beispiel: Sei $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ die Drehmatrix um $\pi/2$.

Das charakteristische Polynom ist $\lambda^2 + 1$. Seine Nullstellen, also Eigenwerte von \mathcal{A} sind $\pm i$. Zu $+i$ gehört ein Eigenvektor $\begin{pmatrix} 1 \\ i \end{pmatrix}$ zu -1 gehört $\begin{pmatrix} 1 \\ -i \end{pmatrix}$. Sie sind linear unabhängig bilden also eine Basis.

Allgemein gilt:

Satz 7.49 Seien $\lambda_1, \dots, \lambda_r$ verschiedene Eigenwerte der linearen Abbildung $A : V \rightarrow V$ mit Eigenvektoren z_j (also $Az_j = \lambda_j z_j$).

Dann sind sie linear unabhängig.

Beweis: Für $r = 1$ ist nichts zu beweisen. Sei also $r > 1$. Seien z_1, \dots, z_s linear unabhängig und oBdA sei dies eine maximal linear unabhängige Menge.

Wäre $s < p$, so wäre $z_p = \sum_{j=1}^s \gamma_j z_j$, also

$$\lambda_p z_p = Az_p = \sum_{j=1}^s \gamma_j Az_j = \sum_{j=1}^s \lambda_j \gamma_j z_j.$$

Damit folgt $0 = \lambda_p z_p - \lambda_p z_p = \sum_{j=1}^s \gamma_j (\lambda_p - \lambda_j) z_j$. Nach Voraussetzung ist $(\lambda_p - \lambda_j) \neq 0$. Also muß jedes $\gamma_j = 0$ sein wegen der linearen Unabhängigkeit der z_j . Damit ist $0 = z_p$, ein Widerspruch.

Hat man eine Basis $C = \{c_1, \dots, c_p\}$ aus Eigenvektoren, so erhalten wir $\mathcal{A}_A^{c,c} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{pmatrix}$ also eine Diagonalmatrix. Nun gilt aber für die alte Matrix $\mathcal{A}_A^{B,B}$ bezüglich der Basis B $\mathcal{A}_A^{c,c} = S_B^{-1} \mathcal{A}_A^{B,B} S_B$. Matrizen der Form $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{pmatrix}$ heißen **Diagonalmatrizen**. Wir erhalten damit den Satz.

Satz 7.50 Sei A eine komplexe $p \times p$ -Matrix. \mathbb{C}^p besitze eine Basis aus Eigenvektoren von A . Dann gibt es eine Matrix S , und eine Diagonalmatrix D mit

$$A = SDS^{-1}$$

Man nennt dann A **diagonalisierbar**.

Wir erhalten sofort das folgende hinreichende Diagonalisierbarkeitskriterium:

Korollar 7.51 Die $p \times p$ -Matrix A besitze p verschiedene Eigenwerte. Dann ist sie diagonalisierbar.

Beweis: Seien b_1, \dots, b_p zu $\lambda_1, \dots, \lambda_2, \dots, \lambda_p$ gehörigen Eigenvektoren. Nach Satz 7.49 sind sie linear unabhängig. Wegen $\dim(\mathbb{C}^p) = p$ bilden sie eine Basis. Die Behauptung folgt nun aus dem Satz 7.50

Symmetrische und normale Matrizen

Überblick: Wir zeigen, daß eine bestimmte Klasse von Matrizen, die sog. normalen Matrizen, Orthonormalbasen aus Eigenvektoren besitzen.

Definition 7.52 Im folgenden sei $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} . Für das Skalarprodukt auf \mathbb{C}^p verweisen wir auf Aufgabe 9

Wir wiederholen und ergänzen: $x^\perp, y^\perp \in \mathbb{K}^p$ heißen **orthogonal**, in Zeichen: $x^\perp \perp y^\perp$, wenn $(x^\perp | y^\perp) = 0$ gilt.

Eine Menge $M = \{a^\perp_1, \dots, a^\perp_q\}$ heißt **Orthonormalsystem**, wenn $(a^\perp_i | a^\perp_j) = \delta_{ij}$ gilt, d.h. $\|a_i\|^2 = 1$, $(a^\perp_i | a^\perp_j) = 0$ für $i \neq j$. M heißt **Orthonormalbasis**, wenn M ein Orthonormalsystem und gleichzeitig eine Basis (also $p = q$) ist.

Im folgenden betrachten wir komplexe $p \times p$ -Matrizen $A = (a_{ik})$, das heißt, die Einträge sind aus \mathbb{C} .

Für $A = \begin{pmatrix} a_{11}, \dots, a_{1p} \\ \vdots \\ a_{p1}, \dots, a_{pp} \end{pmatrix}$ setzen wir die **transponierte Matrix** A^T , die wir durch Spiegelung an der Hauptdiagonale (links oben bis rechts unten) erhalten,

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{p1} \\ \vdots & & & \vdots \\ a_{1p} & a_{2p} & \dots & a_{pp} \end{pmatrix}$$

Die Zeilenvektoren von A^T sind also nichts anderes als die Spaltenvektoren von A .

Die **adjungierte Matrix** A^* ist nichts anderes als

$$A^* = \overline{A^T} = \begin{pmatrix} \overline{a_{11}} & \overline{a_{21}} & \dots & \overline{a_{p1}} \\ \vdots & & & \vdots \\ \overline{a_{1p}} & \overline{a_{2p}} & \dots & \overline{a_{pp}} \end{pmatrix}$$

Beispiele:

1. Ist A reell (also $a_{ij} \in \mathbb{R}$ wie früher), so ist $A^T = A^*$, also $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^* = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$
2. $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}^* = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$
3. $\begin{pmatrix} 1 & i & -i \\ 0 & 1 & i \\ i & 0 & i \end{pmatrix}^* = \begin{pmatrix} 1 & 0 & -i \\ -i & 1 & 0 \\ i & -i & -i \end{pmatrix} \neq \begin{pmatrix} 1 & i & -i \\ 0 & 1 & i \\ i & 0 & i \end{pmatrix}^T$

Definition 7.53 a) Eine $p \times p$ -Matrix A heißt **selbstadjungiert**, wenn $A^* = A$ ist.

b) Eine $p \times p$ -Matrix A heißt **unitär**, wenn sie invertierbar ist und $A^* = A^{-1}$ gilt.

c) Eine $p \times p$ -Matrix A heißt **normal**, wenn $A^*A = AA^*$ gilt.

Beispiele:

1. $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ist selbstadjungiert. Allgemeiner ist eine reelle Matrix A genau dann selbstadjungiert, wenn $A = A^T$ gilt. Solche Matrizen heißen **symmetrisch**.
2. Eine reelle Matrix A ist genau dann unitär, wenn ihre Spaltenvektoren $a^{\downarrow 1}, \dots, a^{\downarrow p}$ eine **Orthonormalbasis** bilden, d.h. wenn $(a^{\downarrow i}, | a^{\downarrow j}) = \delta_{ij}$ gilt.
Beweis?
reelle unitäre Matrizen heißen **orthogonal**. Typische Beispiele sind *2- und 3-dimensionale Drehmatrizen*.
3. Selbstadjungierte und unitäre Matrizen sind normal. Hier ist eine normale Matrix, die weder selbstadjungiert noch unitär ist: $A = \begin{pmatrix} 5 \cos \varphi & -5 \sin \varphi \\ 5 \sin \varphi & 5 \cos \varphi \end{pmatrix}$.

Satz 7.54 (Kennzeichnung selbstadjungierter unitärer Matrizen)

Sei $(x^{\downarrow} | y^{\downarrow}) = \sum_{j=1}^p \bar{x}_j y_j$ das komplexe Skalarprodukt (Aufgabe 9, Ü-Blatt 3) und A eine $p \times p$ -Matrix. Dann gilt:

- a) $(Ax^{\downarrow} | y^{\downarrow}) = (x^{\downarrow} | A^*y^{\downarrow})$ für alle $x^{\downarrow}, y^{\downarrow}$
- b) A ist genau dann selbstadjungiert, wenn $(Ax^{\downarrow} | y^{\downarrow}) = (x^{\downarrow} | Ay^{\downarrow})$ für alle $x^{\downarrow}, y^{\downarrow}$ gilt.
- c) A ist genau dann unitär, wenn $(Ax^{\downarrow} | Ay^{\downarrow}) = (x^{\downarrow} | y^{\downarrow})$ gilt.

Beweis: Sei $A = (a_{ik})_{i,k=1, \dots, p}$. Dann ist $a_{ik} = (e_i | Ae_k)$ für alle i, k . Die Behauptung folgt nun aus den Eigenschaften des Skalarprodukts.

Wir setzen für $M \subset \mathbb{C}^p$ einfach $M^{\perp} = \{y^{\downarrow} | (y^{\downarrow} | x^{\downarrow}) = 0 \text{ für alle } x^{\downarrow} \in M\}$. Wir benötigen folgendes einfache Lemma:

Lemma 7.55 Sei A eine $p \times p$ -Matrix und $V = \mathbb{C}^p$. Dann gilt:

a) $A^*x^\perp = 0^\perp \Rightarrow A(x^{\perp\perp}) \subset x^{\perp\perp}$

b) Ist A normal, so ist $\ker(A) = \ker(A^*)$.

Beweis: $x^\perp \in \ker(A) \Leftrightarrow$

$$\begin{aligned} 0 = \|Ax^\perp\|^2 &= (Ax^\perp | Ax^\perp) = (x^\perp | A^*Ax^\perp) \stackrel{\text{A normal}}{=} (x^\perp | AA^*x^\perp) \\ &= (A^*x^\perp | A^*x^\perp) = \|A^*x^\perp\|^2. \end{aligned}$$

Hieraus folgt die Behauptung.

Damit können wir den folgenden fundamentalen Satz beweisen:

Theorem 7.56 Jede normale Matrix A besitzt eine Orthonormalbasis aus Eigenvektoren.

Beweis: (Skizze) Man beweist dies durch Induktion nach der Dimension p .

(I) $p = 1$. $V = \mathbb{C}$, $A = (a)$ $Ax = ax$. Setze $e = 1$. Das ist die Orthonormalbasis für V .

(II) Sei der Satz richtig für $\dim V = p$. Sei A eine $(p+1) \times (p+1)$ -Matrix. Da das charakteristische Polynom $\det(\lambda E_{p+1} - A)$ nach dem Fundamentalsatz der Algebra mindestens eine Nullstelle λ_1 hat, gibt es einen Vektor b_1^\perp der Norm 1 mit $Ab_1^\perp = \lambda b_1^\perp$. Da mit A auch $A - \lambda_1 E_p$ normal ist, gilt nach Lemma 7.55

b) $A^*b_1^\perp = \bar{\lambda}b_1^\perp$. Daraus folgt nach Teil a) desselben Lemmas für $(b_1^\perp)^\perp = W$: $(A - \lambda_1 E)(W) \subset W$. Wegen $\dim W = p$ folgt nach Induktionsvoraussetzung: es gibt eine Orthonormalbasis $\{b_2^\perp, \dots, b_p^\perp\}$ von W aus Eigenvektoren von $(A - \lambda_1 E)|_W$, also von $A|_W$. Damit ist $\{b_1, \dots, b_{p+1}\}$ eine Orthonormalbasis aus Eigenvektoren von A .

Korollar 7.57 Jede reelle symmetrische $p \times p$ -Matrix A hat nur reelle Eigenwerte und besitzt eine Orthonormalbasis $\{b_1^\perp, \dots, b_p^\perp\}$ aus Eigenvektoren in \mathbb{R}^p .

Beweis: Sei A eine symmetrische reelle $p \times p$ -Matrix. Aufgefaßt als Matrix über \mathbb{C} ist A dann genau selbstadjungiert. Ist $Ax^\perp = \lambda x^\perp$ mit $x^\perp \neq 0$, so gilt

$$\begin{aligned} \lambda \|x^\perp\|^2 = \lambda (x^\perp | x^\perp) &= (x^\perp | \lambda x^\perp) = (x^\perp | Ax^\perp) \\ &\stackrel{\text{A s.a.}}{=} (Ax^\perp | x^\perp) = \overline{(x^\perp | Ax^\perp)} = \overline{(x^\perp | \lambda x^\perp)} = \bar{\lambda} \|x^\perp\|^2. \end{aligned}$$

Division durch $\|x^\perp\|^2$ liefert $\lambda \in \mathbb{R}$. Ist nun $x^\perp = u^\perp + iv^\perp$ mit $u^\perp, v^\perp \in \mathbb{R}^p$, so ist

$$Ax^\perp = Au^\perp + iAv^\perp = \lambda x^\perp = \lambda u^\perp + i\lambda v^\perp,$$

also ist u^\perp reeller Eigenvektor von A zu λ . Nun verläuft der Beweis analog zu dem des Theorems, jetzt aber in \mathbb{R}^p .

Anwendungen:

1. Sei A eine reelle symmetrische $p \times p$ -Matrix. Dann ist $(x^\perp | Ax^\perp) > 0$ für alle $0^\perp \neq x^\perp \in \mathbb{R}^p$ genau dann, wenn alle Eigenwerte von A größer als 0 sind. Solche Matrizen heißen **positiv definit**.

2. Sei $\varphi : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$ in beiden Argumenten linear; ferner gelte $\varphi(x^\perp, y^\perp) = \varphi(y^\perp, x^\perp)$. Solche Abbildungen nennt man **symmetrische Bilinearformen**. φ heißt **positiv definit**, wenn $0 < \varphi(x^\perp, x^\perp)$ für alle $x^\perp \neq 0$ gilt.

Sei $\{e_1^\perp, \dots, e_p^\perp\}$ die kanonische Basis und $a_{ik} = \varphi(e_i, e_k)$, ferner $A_\varphi = \begin{pmatrix} a_{11}, \dots, a_{1p} \\ \vdots \\ a_{p1}, \dots, a_{pp} \end{pmatrix}$. Dann

ist A reell, symmetrisch und es gilt

$$\varphi(x^\perp, y^\perp) = (x^\perp | Ay^\perp).$$

Damit ist φ genau dann positiv definit, wenn A dies ist.

Eine Orthonormalbasis $B = \{b_1^\perp, \dots, b_p^\perp\}$ aus Eigenvektoren von A erlaubt dann einen *Basiswechsel*,

die sog. **Hauptachsentransformation** von φ . Ist $x'^\perp = \begin{pmatrix} x_1 \\ \vdots \\ x_p' \end{pmatrix}$ der Koordinatenvektor von

x^\perp in der neuen Basis, so gilt

$$\varphi(x'^\perp, x'^\perp) = \lambda_1 x_1'^2 + \dots + \lambda_p x_p'^2.$$

Kapitel 8

Differentialrechnung mehrerer Veränderlicher

Überblick: Wir übertragen zunächst die Begriffe “Konvergenz von Folgen”, “Grenzwert” und “Stetigkeit” auf die mehrdimensionale Situation und führen dann den Begriff der partiellen und totalen Differenzierbarkeit ein. Wir geben Anwendungen auf skalare Funktionen, auf Kurven und auf krummlinige (d.h. nichtlineare) Koordinatentransformationen an.

8.1 Konvergenz, Grenzwert und Stetigkeit in \mathbb{R}^p

Überblick: Wie im eindimensionalen Fall studieren wir zunächst die Konvergenz von Folgen, bestimmte Klassen von Mengen und Grenzwerte von Funktionswerten. Die Stetigkeit ist ein Spezialfall der Grenzwerttheorie. Wir untersuchen die wichtigsten Eigenschaften stetiger Funktionen mehrerer Variablen.

8.1.1 Topologie im \mathbb{R}^p

Überblick: Wir führen verschiedene Entfernungsmessungen in \mathbb{R}^p bzw. im Raum linearer Abbildungen $L(\mathbb{R}^p, \mathbb{R}^q)$ von \mathbb{R}^p nach \mathbb{R}^q ein. Jede hat ihren spezifischen Sinn, aber Konvergenz, Grenzwert und Stetigkeit sind Begriffe, die für alle Entfernungen das Gleiche bedeuten.

Normen und Entfernung

Wir hatten in Abschnitt 7.1.3 bereits die **euklidische Norm** in \mathbb{R}^p eingeführt.

$$\text{Sei } x = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \quad \|x\|_2 = (x_1^2 + x_2^2 + \cdots + x_p^2)^{1/2}.$$

Für sie gilt

$$\|x\| = 0 \Leftrightarrow x = 0 \quad \text{Definitheit} \quad (8.1)$$

$$\|\lambda x\| = |\lambda| \|x\| \quad \text{Absolut-Homogenität} \quad (8.2)$$

$$\|x + y\| \leq \|x\| + \|y\| \quad \text{Dreiecksungleichung} \quad (8.3)$$

$$|\|x\| - \|y\|| \leq \|x - y\| \leq \|x\| + \|y\| \quad \text{Verallgemeinerte Dreiecksungleichung} \quad (8.4)$$

Der **Abstand zwischen 2 Vektoren** x und y wird einfach durch die Länge der Differenz, also durch $\|x - y\|$ erklärt. Für den Abstand nimmt die Dreiecksungleichung die folgende Form an:

$$\|x - y\| \leq \|x - z\| + \|z - y\|$$

Man kann sie daher auch **Umweg-Ungleichung** nennen.

Wichtig ist die folgende Abschätzung durch die Koordinaten

Satz 8.1 a) Auch die folgenden Funktionen erfüllen die Beziehungen (8.1) bis (8.4):

$$(i) \|x\|_\infty := \max(|x_1|, \dots, |x_p|)$$

$$(ii) \|x\|_1 := |x_1| + \dots + |x_p|$$

Sie heißen **Unendlich-Norm** bzw. **1-Norm**

b) Es gilt

$$\|x\|_\infty \leq \|x\| \leq \sqrt{p} \|x\|_\infty \leq \sqrt{p} \|x\|_1 \leq p \|x\| \leq p^{3/2} \|x\|_\infty \quad (8.5)$$

das heißt, man kann jeweils eine Norm durch die andere abschätzen.

Aufgabe: Zeigen Sie, daß die unter a) definierten Funktionen die genannten Beziehungen erfüllen.

Beweis: b) Es ist $|x_j| \leq \sqrt{\sum_{k=1}^p x_k^2} = \|x\|$, also auch $\|x\|_\infty = \max(|x_1|, \dots, |x_p|) \leq \|x\|$.

Ferner ist $\sum_{k=1}^p |x_k|^2 \leq p \cdot \max_k(|x_k|^2) = p \|x\|_\infty^2$.

Schließlich ist $\|x\|_\infty = \max(|x_1|, \dots, |x_p|) \leq \sum_{k=1}^p |x_k| = \|x\|_1$

$$\text{und } \|x\|_1 = \sum_{k=1}^p |x_k| = \left(\begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right) \cdot \left(\begin{array}{c} |x_1| \\ \vdots \\ |x_p| \end{array} \right) \leq \left\| \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right\| \cdot \|x\| = \sqrt{p} \|x\|.$$

Für Matrizen \mathcal{A} definieren wir

$$\text{Definition 8.2 } \|\mathcal{A}\| = \left\| \begin{pmatrix} a_{11}, \dots, a_{1p} \\ \vdots \\ a_{q1}, \dots, a_{qp} \end{pmatrix} \right\| = (\sum_{jk} |a_{jk}|^2)^{1/2} \text{ heißt Matrixnorm oder auch}$$

Hilbert-Schmidt-Norm

Bemerkung: Identifiziert man $\mathcal{A} = (a_{ik})$ mit dem pq -Vektor $\begin{pmatrix} a_{\downarrow 1} \\ \vdots \\ a_{\downarrow p} \end{pmatrix} =: x_{\mathcal{A}}$, so ist $\|\mathcal{A}\| = \|x_{\mathcal{A}}\|$, d.h.

auch für die Matrixnorm gelten die Abschätzungen aus Satz 8.1. Die Matrixnorm verhält sich gut bei Multiplikation.

Satz 8.3 a) Die Abbildung $\mathcal{A} \rightarrow \|\mathcal{A}\|$ erfüllt die Beziehungen (8.1) bis (8.4).

b) Ist \mathcal{A} eine $q \times p$ -Matrix, \mathcal{L} eine $r \times q$ -Matrix, so ist

$$\|\mathcal{L}\mathcal{A}\| \leq \|\mathcal{L}\| \|\mathcal{A}\| \quad \text{Submultiplikativität der Norm.}$$

Beweis:

a) folgt aus $\|\mathcal{A}\| = \|x_{\mathcal{A}}\|$.

b) Für eine Matrix \mathcal{A} ist $\|\mathcal{A}\|^2 = \sum_{k=1}^p \|a^{\perp_k}\|^2 = \sum_{j=1}^q \|\vec{a}_j^T\|^2$, wo $\vec{a}_j^T = \begin{pmatrix} a_{j1} \\ \vdots \\ a_{jp} \end{pmatrix}$ die transponierte

Matrix der $1 \times p$ -Matrix \vec{a}_j ist.

Damit ist wegen $\mathcal{L}\mathcal{A} = (\sum_{i=1}^r \vec{b}_i a^{\perp_k})_{\substack{i=1, \dots, r \\ k=1, \dots, p}}$ und $|\vec{b}_i a^{\perp_k}|^2 = (\vec{b}_i^T | a^{\perp_k})^2 \leq \|\vec{b}_i^T\|^2 \|a^{\perp_k}\|^2$

$$\begin{aligned} \|\mathcal{L}\mathcal{A}\|^2 &= \sum_{ik} |\vec{b}_i a^{\perp_k}|^2 \leq \left(\sum_{i=1}^r \|\vec{b}_i^T\|^2 \right) \left(\sum_{k=1}^p \|a^{\perp_k}\|^2 \right) \\ &= \|\mathcal{L}\|^2 \|\mathcal{A}\|^2 \end{aligned}$$

Spezielle Klassen von Mengen

Im Folgenden identifizieren wir $L(\mathbb{R}^p, \mathbb{R}^q)$ mit \mathbb{R}^{pq} , so daß alle Begriffe auch automatisch für $L(\mathbb{R}^p, \mathbb{R}^q)$ eingeführt sind.

Definition 8.4 Sei $A \subset \mathbb{R}^p$ eine Menge.

- a) A heißt **offen**, wenn zu jedem $x \in A$ ein $r(x) > 0$ existiert, so daß $B(x, r(x)) = \{y : \|y - x\| < r(x)\}$ ganz in A liegt. $B(x, r(x))$ selbst heißt **Kugel** mit Mittelpunkt x und Radius $r(x)$.
- b) A heißt **abgeschlossen**, wenn $\mathbb{R}^p \setminus A$ offen ist.
- c) A heißt **beschränkt**, wenn es ein M mit $A \subset B(0, M)$ gibt.
- d) A heißt **kompakt**, wenn A abgeschlossen und beschränkt ist.

Für $p = 1$ ist $B(x, r(x)) =]x - r(x), x + r(x)[$.

Für $p = 2$ ist $B(x, r(x))$ eine Kreisscheibe um x mit Radius $r(x)$.

Allgemein ist $B(x, r(x))$ offen. Denn ist $y \in B(x, r(x))$, also $\|y - x\| < r(x)$,

so liegt $B(y, \underbrace{r(x) - \|y - x\|}_{=r(y)})$ in $B(x, r(x))$. Machen Sie sich eine Skizze.

Aufgaben: Zeigen Sie bitte

1. Ist \mathcal{U} eine Menge von offenen Mengen, so ist $\cup_{U \in \mathcal{U}} U$ offen.
2. Ist U, V offen, so ist $U \cap V$ offen.

3. Ist \mathcal{A} eine Menge abgeschlossener Mengen, so ist $\bigcap_{A \in \mathcal{A}} A$ abgeschlossen.
4. Sind A, B abgeschlossen, so ist $A \cup B$ abgeschlossen.
5. \emptyset ist offen und abgeschlossen.

Der wichtigste Begriff neben dem der Kompaktheit ist der des Gebietes. Fast alle physikalischen Probleme sind auf Gebieten formuliert.

Definition 8.5 Eine Menge A heißt **offen zerlegbar**, wenn es zwei offene Mengen U und V mit $U \cap V = \emptyset$ und $A \cap U \neq \emptyset \neq A \cap V$ und $A \subset U \cup V$ gibt. $G \subset \mathbb{R}^p$ heißt **Gebiet**, wenn G offen, $\neq \emptyset$ und nicht offen zerlegbar ist.

Beispiele:

1. $p = 1$: Die offenen Intervalle sind die einzigen Gebiete in \mathbb{R} . Gebiet ist also die Verallgemeinerung offener Intervalle.
2. $p = 2$: Ein Kreisring $G = \{x : r_1 < \|x\| < r_2\}$ ist ein Gebiet.
3. Allgemein ist $B(x, r)$ ein Gebiet.

8.1.2 Folgen in \mathbb{R}^p

Im Grunde definieren wir alles wie in \mathbb{R} , ersetzen den Absolutbetrag einfach durch die Norm.

Definition 8.6 a) eine Abbildung $a : \mathbb{N} \rightarrow \mathbb{R}^p$, $n \mapsto a_n$ heißt **Folge** (in \mathbb{R}^p)

- b) Die Folge $a = (a_n)_n$ **konvergiert gegen** b , in Zeichen: $b = \lim_{n \rightarrow \infty} a_n$, wenn es zu jedem $\varepsilon > 0$ ein $n(\varepsilon)$ mit $\|b - a_n\| < \varepsilon$ für alle $n \geq n(\varepsilon)$ gibt.
- c) Die Folge $a = (a_n)_n$ heißt **Cauchy-Folge**, wenn es zu jedem $\varepsilon > 0$ ein $n(\varepsilon)$ gibt mit $\|a_m - a_n\| < \varepsilon$ für alle $m, n \geq n(\varepsilon)$.
- d) $a = (a_n)_n$ heißt **beschränkt**, wenn es ein $M > 0$ mit $\|a_n\| \leq M$ für alle n gibt.

Wir erhalten sofort:

Lemma 8.7 Genau dann konvergiert die Folge $(a_n)_n$ gegen b wenn die Folge $(\|b - a_n\|)_n$ der Entfernungen eine Nullfolge in \mathbb{R} ist.

Aufgabe: Beweisen Sie bitte das Lemma!

Wir erhalten den folgenden bequemen Satz, den Sie aus dem Anschauen der Bilder schon gewonnen haben:

Satz 8.8 Sei $a = (a_n)_n$ eine Folge in \mathbb{R}^p . Sei $a_n = \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{p,n} \end{pmatrix}$ und $(a_{j,n})_n$ die j . Koordinatenfolge. Dann gilt:

$a = (a_n)_n$ konvergiert genau dann gegen $b = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$, wenn jede Koordinatenfolge $(a_{j,n})_n$ in \mathbb{R} gegen b_j konvergiert.

Beweis: Nach Formel (8.5) gilt

$$\|b_j - a_{j,n}\| \leq \|b - a_n\| \leq \sqrt{p} \max_j \|b_j - a_{j,n}\|.$$

Sind alle $(\|b_j - a_{j,n}\|)_n$ Nullfolgen, so auch die rechte Seite, also auch $(\|b - a_n\|)_n$ und das impliziert wiederum, daß alle $(\|b_j - a_{j,n}\|)_n$ Nullfolgen sind.

Korollar 8.9 Jede Cauchyfolge konvergiert.

Beweis: Nach Formel (8.5) ist $(a_n)_n$ genau dann Cauchyfolge, wenn jede Koordinatenfolge Cauchyfolge in \mathbb{R} ist. In \mathbb{R} konvergiert aber jede Cauchyfolge.

Definition 8.10 b heißt Häufungspunkt der Folge $(a_n)_n$, wenn es eine gegen b konvergente Teilfolge gibt.

Theorem 8.11 (Bolzano-Weierstraß) Jede beschränkte Folge hat einen Häufungspunkt.

Beweis: Der Übersichtlichkeit führen wir den Beweis für $p = 2$: Sei $(a_n)_n$ beschränkt. Nach Formel (8.5) ist dann jede Koordinatenfolge beschränkt in \mathbb{R} . Da dort der Satz gilt, hat $(a_{1,n})_n$ eine konvergente Teilfolge $(a_{1,\varphi(n)})_n$. Dann ist $(a_{2,\varphi(n)})_n$ als Teilfolge von $(a_{2,n})_n$ beschränkt, hat also eine konvergente Teilfolge von $(a_{2,\psi(\varphi(n))})_n$. Da $(a_{1,\psi(\varphi(n))})_n$ Teilfolge von $(a_{1,\varphi(n)})_n$ ist, ist sie konvergent, d.h. beide Koordinatenfolgen der Teilfolge $(a_{\psi(\varphi(n))})_n$ konvergieren. Also konvergiert die Teilfolge.

Weil alle Rechenoperationen koordinatenweise erfolgen, erhalten wir leicht die folgenden Rechenregeln:

Satz 8.12 Bei den folgenden Gleichungen mögen die links stehenden Grenzwerte existieren (d.h. die Folgen konvergieren). Dann gelten die Gleichungen

- a) $\lim a_n + \lim b_n = \lim(a_n + b_n)$
- b) $\lim \lambda_n \cdot \lim a_n = \lim(\lambda_n a_n)$
- c) $(\lim a_n \mid \lim b_n) = \lim(a_n \mid b_n)$
- d) $\|\lim a_n\| = \lim \|a_n\|$

Korollar 8.13 Sei $(A_n)_n$ eine Folge von $q \times p$ -Matrizen, $(\mathcal{L}_n)_n$ eine Folge von $r \times q$ -Matrizen. Dann gilt

$$\lim \mathcal{L}_n \cdot \lim A_n = \lim \mathcal{L}_n A_n.$$

Als Anwendung zeigen wir, wie man Potenzreihen von Matrizen definieren kann.

Satz 8.14 Sei \mathcal{A} eine $p \times p$ -Matrix.

a) Sei $\|\mathcal{A}\| = q < 1$. Dann konvergiert die Reihe $\sum_{k=0}^{\infty} \mathcal{A}^k$ gegen $(E_p - \mathcal{A})^{-1}$

b) Die Reihe $\sum_{k=0}^{\infty} \frac{\mathcal{A}^k}{k!}$ konvergiert für jede $p \times p$ -Matrix \mathcal{A} . Der Grenzwert heißt Exponentialfunktion $\exp(\mathcal{A})$.

Bemerkung: b) ist zentral für die Theorie dynamischer Systeme.

Beweis: Wegen der Submultiplikativität der Matrixnorm ist $\|\mathcal{A}^2\| \leq \|\mathcal{A}\|^2$, also (Induktion) $\|\mathcal{A}^n\| \leq \|\mathcal{A}\|^n$.

a) Sei $S_n = \sum_{k=0}^n \mathcal{A}^k$. Dann ist nach der Dreiecksungleichung

$$\begin{aligned} \|S_{n+r} - S_n\| &\leq \sum_{k=n+1}^{n+r} \|\mathcal{A}^k\| \leq \sum_{k=n+1}^{n+p} \|\mathcal{A}\|^k \\ &= \sum_{k=n+1}^{n+r} q^k \leq \frac{q^{n+1}}{1-q} \end{aligned}$$

Wegen $0 \leq q < 1$ konvergiert die rechte Seite für $n \rightarrow \infty$ gegen 0, also ist $(S_n)_n$ eine Cauchyfolge und damit konvergent. Sei $S = \lim S_n$. Nach den Rechenregeln ist wegen

$$\begin{aligned} (E_p - \mathcal{A})S_n &= E_p - \mathcal{A}^{n+1} = S_n(E_p - \mathcal{A}) \\ (E_p - \mathcal{A})S &= E_p - \lim \mathcal{A}^{n+1} = S(E_p - \mathcal{A}). \end{aligned}$$

Nun ist $\|\mathcal{A}^{n+1}\| \leq q^{n+1}$, also $\lim \mathcal{A}^{n+1} = 0$ und a) folgt.

b) Man setzt $S_n = \sum_{k=0}^n \frac{\mathcal{A}^k}{k!}$ und erhält

$$\|S_{n+r} - S_n\| \leq \sum_{k=n+1}^{\infty} \frac{\|\mathcal{A}\|^k}{k!} = \exp(\|\mathcal{A}\|) - \sum_{k=0}^n \frac{\|\mathcal{A}\|^k}{k!} \rightarrow 0.$$

Daraus folgt, daß $(S_n)_n$ eine Cauchyfolge ist.

Aufgaben: Zeigen Sie bitte:

1. $\|\exp(\mathcal{A})\| \leq \exp(\|\mathcal{A}\|)$

Tipp: $\|S_n\| \leq \exp(\|\mathcal{A}\|)$, warum?

2. Sei \mathcal{A} invertierbar. Dann ist $\frac{1}{\|\mathcal{A}^{-1}\|} \leq \|\mathcal{A}\|$.

Tipp: $E_p = \mathcal{A}\mathcal{A}^{-1}$, Submultiplikativität der Norm

3. Sei \mathcal{A} invertierbar und $\|\mathcal{L} - \mathcal{A}\| < \frac{1}{\|\mathcal{A}^{-1}\|}$. Dann ist \mathcal{L} invertierbar und

$$\|\mathcal{L}^{-1} - \mathcal{A}^{-1}\| \leq \frac{\|\mathcal{A}^{-1}\|^2 \|\mathcal{A} - \mathcal{L}\|}{1 - \|\mathcal{A}^{-1}(\mathcal{A} - \mathcal{L})\|}.$$

Tipp: $\mathcal{L} = \mathcal{A} - (\mathcal{A} - \mathcal{L}) = \mathcal{A}(E_p - \mathcal{A}^{-1}(\mathcal{A} - \mathcal{L}))$ und $\|\mathcal{A}^{-1}(\mathcal{A} - \mathcal{L})\| < 1$, warum? Verwende nun den Teil a) des vorausgegangenen Satzes. Für den Rest benutze

$$\mathcal{L}^{-1} - \mathcal{A}^{-1} = ((E_p - \mathcal{A}^{-1}(\mathcal{A} - \mathcal{L}))^{-1} - E_p) \mathcal{A}^{-1}$$

8.1.3 Grenzwerte von Funktionswerten, Stetigkeit

Typen von Funktionen

skalare Funktionen:

Funktionen mit Werten in \mathbb{R} heißen **skalare** Funktionen. Typische Beispiele sind die Temperatur als Funktion des Ortes, die potentielle Energie eines Teilchens im Raum, oder die Populationsdichte einer Spezies als Funktion des Ortes.

Anleitung zum Applet "Parameterflächen im Raum" Schauen Sie sich eine skalare Funktion auf \mathbb{R}^2 an. Ihre Graphen sind meist wunderschöne Flächen.

Kurven im \mathbb{R}^p :

Eine Funktion φ eines Intervalls J in den \mathbb{R}^p erscheint als **Kurve**.

Appletanleitung Typische Kurven in \mathbb{R}^2 sind Kreise $\varphi(t) = R \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$. In \mathbb{R}^3 hat man Schraubenlinien, usw.

*** TO DO ***

Flächen in \mathbb{R}^p :

Man erhält sie durch Parameterdarstellungen, also Funktionen φ eines Gebietes $G \subset \mathbb{R}^2 \rightarrow \mathbb{R}^p$ ($p \geq 3$).

Appletanleitung $\varphi(u, v) = (u, v, u^2 - v^2)$ liefert eine Sattelfläche, $\psi(u, v) = \begin{pmatrix} \cos u \cos v \\ \sin u \cos v \\ \sin v \end{pmatrix}$ liefert die Einheitssphäre.

*** TO DO ***

Vektorfelder:

Funktionen des p -dimensionalen Raumes in sich können wir als Vektorfelder deuten.

Appletanleitung Wir veranschaulichen $\varphi : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$, indem wir an jeden Punkt x von D den Vektor $\varphi(x)$ anbringen. Die Vorstellung ist die eines Kraftfeldes.

Hier drei Beispiele:

a) $p = 2$, $f(x, y) = \begin{pmatrix} (x+y)/2 \\ (x-y)/2 \end{pmatrix}$. Auf der Hauptdiagonalen zeigt der Vektor $f(x, x)$ gerade in positiver Richtung der x -Achse, auf der Nebendiagonale in Richtung der negativen y -Achse.

b) $p = 2$. In Orts- und Impulskoordinaten unterliegt das physikalische Pendel dem Kraftfeld $f(x, y) = \begin{pmatrix} y \\ -\sin x \end{pmatrix}$.

c) $p = 3$. Das elektrische Feld einer geladenen Einheitssphäre wird durch $v(x, y, z) = \frac{-1}{\max(1, \|x^\dagger\|^3)} x^\dagger$ beschrieben.

*** TO DO ***

Koordinatentransformationen:

Eine Funktion $f : D \subset \mathbb{R}^p \rightarrow \mathbb{R}^p$ kann man auch als Koordinatentransformation auffassen. Hier drei Beispiele:

a) $f(x) = Ax$, wo $\det A \neq 0$. Dies haben wir als Basiswechsel interpretiert.

b) $f(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$. Dies sind Polarkoordinaten.

Genauer: hat man einen Punkt P mit rechtwinkligen Koordinaten $x = \begin{pmatrix} x \\ y \end{pmatrix}$ und setzt man $r = \|x\|$ und ist $x \neq 0$, so gibt es genau einen Winkel φ zwischen $-\pi$ und π ($-\pi \leq \varphi < \pi$) mit $x = r \cos \varphi$ und $y = r \sin \varphi$. r und φ sind die **Polarkoordinaten** von P .

c) Die **Kugelkoordinaten** in \mathbb{R}^3 ergeben sich als

$$x = r \cos(\varphi) \cos(\vartheta)$$

$$y = r \sin(\varphi) \cos(\vartheta)$$

$$z = r \sin(\vartheta)$$

mit $-\pi \leq \varphi < \pi$ und $-\pi/2 \leq \vartheta \leq \pi/2$.

r ist der Radius (Abstand vom Ursprung).

φ die geographische Länge, ϑ die geographische Breite.

Grenzwerte

Definition 8.15 Sei $\emptyset \neq D \subset \mathbb{R}^p$. $c \in \mathbb{R}^p$ heißt **Adhärenzpunkt** von D , wenn es eine Folge $(d_n) \subset D$ mit $c = \lim d_n$ gibt.

Aufgaben: Zeigen Sie bitte:

1. c ist genau dann Adhärenzpunkt von D , wenn für jedes $\varepsilon > 0$ stets $B(c, \varepsilon) \cap D \neq \emptyset$ gilt.
2. $A \subset \mathbb{R}^p$ ist genau dann abgeschlossen, wenn A alle seine Adhärenzpunkte enthält.
3. Sei $D \subset \mathbb{R}^p$ und $\overline{D} = \{x \in \mathbb{R}^p : x \text{ ist Adhärenzpunkt von } D\}$. \overline{D} ist abgeschlossen.
4. $\overline{B(x, r)} = \{y : \|y - x\| \leq r\}$
5. Sei $D \subset \mathbb{R}^p$. Ist A abgeschlossen und $D \subset A$, so ist auch $\overline{D} \subset A$. \overline{D} ist also die kleinste abgeschlossene Menge, die D enthält und heißt **abgeschlossene Hülle von D**

Ganz analog zum eindimensionalen Fall erklären wir den Grenzwert von Funktionswerten:

Definition 8.16 Sei c ein Adhärenzpunkt von $D \subset \mathbb{R}^p$ und $F : D \rightarrow \mathbb{R}^q$ eine Abbildung. $F(x)$ **konvergiert gegen v für x gegen c , in Zeichen:**

$$\lim_{x \rightarrow c} F(x) = v$$

wenn für jede gegen c konvergente Folge (d_n) aus D die Bildfolge $(F(d_n))_n$ gegen v konvergiert. v heißt dann **Grenzwert von $F(x)$ für x gegen c** .

Wie im eindimensionalen Fall ergibt sich:

Satz 8.17 Sei c ein Adhärenzpunkt von $D \subset \mathbb{R}^p$ und $F : D \rightarrow \mathbb{R}^q$ eine Abbildung. Die folgenden Aussagen sind äquivalent:

- a) $\lim_{x \rightarrow c} F(x) = v$.

- b) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle $x \in D$ mit $\|x - c\| < \delta$ stets $\|F(x) - v\| < \varepsilon$ ist.
- c) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $F(B(c, \delta) \cap D) \subset B(v, \varepsilon)$.

Der Beweis verläuft analog zu Theorem 3.6.

Beispiele:

- Sei $p = 1$, $q = 2$, $D = \mathbb{R}$ und $F(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$.
Dann gilt $\lim_{t \rightarrow c} F(t) = \begin{pmatrix} \cos(c) \\ \sin(c) \end{pmatrix}$.
- Seien p, q und F wie oben, $c \in \mathbb{R}$ und $D = \mathbb{R} \setminus \{c\}$ ferner $G(t) = \frac{F(t) - F(c)}{t - c}$. Dann gilt $\lim_{t \rightarrow c} G(t) = \begin{pmatrix} -\sin(c) \\ \cos(c) \end{pmatrix}$.
- Sei $D_0 = \mathbb{R}^2 \setminus \{0\}$ und $F(x, y) = \frac{2xy}{x^2 + y^2}$. Sei $c = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Dann existiert für $(x, y) \rightarrow c$ kein Grenzwert.
Setzen wir aber $D_1 = \{ \begin{pmatrix} t \\ t \end{pmatrix}, t \neq 0 \}$, so ist $\lim_{D_1 \ni x \rightarrow c} F(x) = 1$.

Aus den Rechenregeln für konvergente Folgen erhalten wir sofort den folgenden Satz:

Satz 8.18 Sei c ein Adhärenzpunkt von $D \subset \mathbb{R}^p$ und $F : D \rightarrow \mathbb{R}^q$, $F(x) = (F_1(x), \dots, F_q(x))^T$.

- a) Es gilt: $F(x)$ konvergiert genau dann gegen v für $x \rightarrow c$, wenn jede Koordinatenfunktion $F_j : D \rightarrow \mathbb{R}$ für $x \rightarrow c$ gegen v_j konvergiert. Dann ist

$$\lim_{x \rightarrow c} F(x) = \left(\lim_{x \rightarrow c} F_1(x), \dots, \lim_{x \rightarrow c} F_q(x) \right)^T.$$

- b) Sei zusätzlich $G : D \rightarrow \mathbb{R}^q$. Existieren die Grenzwerte auf der linken Seite, so gelten die folgenden Gleichungen:

$$\begin{aligned} \lim_{x \rightarrow c} F(x) + \lim_{x \rightarrow c} G(x) &= \lim_{x \rightarrow c} (F(x) + G(x)) \\ \left(\lim_{x \rightarrow c} F(x) \mid \lim_{x \rightarrow c} G(x) \right) &= \lim_{x \rightarrow c} (F(x) \mid G(x)) \\ \left\| \lim_{x \rightarrow c} F(x) \right\| &= \lim_{x \rightarrow c} \|F(x)\| \end{aligned}$$

- c) Ist $\varphi : D \rightarrow \mathbb{R}$ und existieren $\lim_{x \rightarrow c} \varphi(x)$ und $\lim_{x \rightarrow c} F(x)$, so auch $\lim_{x \rightarrow c} \varphi(x)F(x)$ und es gilt

$$\lim_{x \rightarrow c} \varphi(x) \lim_{x \rightarrow c} F(x) = \lim_{x \rightarrow c} \varphi(x)F(x).$$

Stetigkeit

Wie im eindimensionalen Fall (siehe Definition 3.9) erklären wir:

Definition 8.19 Sei $\emptyset \neq D \subset \mathbb{R}^p$ und $f : D \rightarrow \mathbb{R}^q$. f heißt **stetig** in $c \in D$, wenn $\lim_{x \rightarrow c} f(x) = f(c)$ gilt. f heißt **stetig**, wenn f in jedem Punkt von D stetig ist.

Aus Satz 8.18 erhalten wir sofort:

- Satz 8.20** a) $f : D \rightarrow \mathbb{R}^q$ ist genau dann stetig (in $c \in D$), wenn dies für jede Koordinatenfunktion f_j gilt ($f(x) = (f_1(x), \dots, f_q(x))^T$).
- b) Mit f und g sind auch $f + g$ und $x \rightarrow (f(x) \mid g(x))$ stetig (in $c \in D$).
- c) Ist $f : D \rightarrow \mathbb{R}^q$ und $\varphi : D \rightarrow \mathbb{R}$ stetig (in c), so auch $\varphi f : x \mapsto \varphi(x)f(x)$.

Aus der äquivalenten Form des Grenzwert-Kriteriums erhalten wir

Satz 8.21 Sei $f : D \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ eine Abbildung. Die folgenden Aussagen sind äquivalent:

- a) f ist stetig in c
- b) Für jede gegen c konvergente Folge $(d_n)_n$ aus D gilt

$$\lim_{n \rightarrow \infty} f(d_n) = f(c) = f\left(\lim_{n \rightarrow \infty} d_n\right).$$

- c) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $f(B(c, \delta) \cap D) \subset B(f(c), \varepsilon)$ oder äquivalent:

$$\forall x \in D \quad \text{gilt} \quad \|x - c\| < \delta \Rightarrow \|f(x) - f(c)\| < \varepsilon.$$

Bevor wir weitere Beispiele bringen, brauchen wir noch den folgenden Satz:

Satz 8.22 Sei $f : D \subset \mathbb{R}^p \rightarrow D' \subset \mathbb{R}^q$ stetig in c und $g : D' \rightarrow \mathbb{R}^r$ stetig in $d = f(c)$. Dann ist $g \circ f : D \rightarrow \mathbb{R}^r$ stetig in c .

Der Beweis läuft analog zum reellen Fall am schnellsten mit dem Folgenkriterium Satz 8.21 b).

Anleitung zum Applet "Parameterflächen im Raum" Schauen Sie sich das $\varepsilon - \delta$ -Kriterium für verschiedene Funktionen f an.

Gehen Sie wie folgt vor: Wählen Sie einen Punkt $x^{(0)} = (x_0, y_0)^T$. und lassen sie sich die Funktion f in einer Umgebung zeichnen. (Sie können den Definitionsbereich in "drawing options" angeben.) Bestimmen sie $f(x^{(0)})$ (rechnerisch oder mit dem Bild).

Wählen Sie ein ε und schränken Sie die z -Achse auf $[-\varepsilon, \varepsilon]$ ein.

Lassen Sie sich mit diesen Einstellungen $f - f(x^{(0)})$ zeichnen.

Finden sie wenn nötig eine kleinere Umgebung von $x^{(0)}$, sodaß f vollständig dargestellt wird, d.h. der Graph wird im *Innern* des Definitionsbereichs durch den eingeschränkten z -Wert nicht "abgeschnitten".

Beispiele: $x^{(0)} = 0$:

1. $f(x, y) = xy$.
2. $f(x, y) = \sin(xy)$.
3. $f(x, y) = \exp(-x^2 - y^2)$.
4. $f(x, y) = \exp(-x^2 + y^2)$.
5. $f(x, y) = \begin{cases} 0 & x = y = 0 \\ \frac{2xy}{x^2 + y^2} & \text{sonst} \end{cases}$

Beispiele:

1. Sei $A : \mathbb{R}^p \rightarrow \mathbb{R}^q$ eine lineare Abbildung, dargestellt durch \mathcal{A} bezüglich der kanonischen Basen. Wendet man Korollar 8.13 auf $q \times p$ und $1 \times q$ an, so erhält man für $\lim y_n = x$ einfach $\lim(\mathcal{A}(y_n) - \mathcal{A}(x)) = \lim \mathcal{A}(y_n - x) = 0$. Also ist jede lineare Abbildung stetig.
2. Jedes Polynom in p Variablen ist stetig. Dazu sei $\alpha \in \mathbb{N}_0^p$, also $\alpha = (\alpha_1, \dots, \alpha_p)$ mit $\alpha_j \in \mathbb{N}_0$. Für $x \in \mathbb{R}^p$ setzen wir $x^\alpha = x_1^{\alpha_1} \cdots x_p^{\alpha_p}$. Nach dem vorausgegangenen Beispiel ist die lineare Abbildung $x \rightarrow x_j$ stetig von $\mathbb{R}^p \rightarrow \mathbb{R}$. Auf \mathbb{R} ist die Potenzfunktion $t \rightarrow t^{\alpha_j}$ stetig, also ist die Hintereinanderausführung $x \mapsto x_j^{\alpha_j}$ stetig. Nach Satz 8.20 b) ist dann auch $x \mapsto x_1^{\alpha_1} x_2^{\alpha_2}$ stetig. Durch Induktion nach p erhält man daß $x \mapsto x^\alpha$ stetig ist. Setzt man $|\alpha| = \alpha_1 + \dots + \alpha_p$, so ist das Polynom $P(x) = \sum_{\ell=0}^n (\sum_{|\alpha|=\ell} a_\alpha x^\alpha)$ stetig nach Satz 8.20.
3. $(x, y) \rightarrow \sin(xy)$ ist stetig als Hintereinanderausführung von $(x, y) \rightarrow xy \rightarrow \sin(xy)$.

Aufgaben: Beweisen Sie die Stetigkeit der folgenden Funktionen:

1. $f(x, y) = x + y$.
2. $f(x, y) = \sin(x + y)$.
3. $f(x, y) = (\sin(x + y), \cos(xy))^T$.
4. $f(x, y) = x \begin{pmatrix} \cos(y) \\ \sin(y) \end{pmatrix}$.
5. $f(r, \varphi, \vartheta) = r \begin{pmatrix} \cos \varphi \cos \vartheta \\ \sin \varphi \cos \vartheta \\ \sin \vartheta \end{pmatrix}$
6. $f(r, \varphi, z) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix}$

Wir benötigen nun noch ein weiteres Mittel zur Konstruktion stetiger Funktionen. Wie im Fall $p = 1$ erklären wir

Definition 8.23 Sei $(f_n)_n$ eine Folge von Funktionen $f_n : D \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$. Die Folge konvergiert **gleichmäßig** gegen die Funktion $g : D \rightarrow \mathbb{R}^q$, wenn zu jedem $\varepsilon > 0$ ein $n(\varepsilon) \in \mathbb{N}$ existiert mit

$$\|f_n(x) - g(x)\| < \varepsilon$$

für alle $n \geq n(\varepsilon)$ und alle $x \in D$.

Mit derselben Beweisidee wie im Fall $p = q = 1$ erhält man:

Satz 8.24 Sei $(f_n)_n$ eine Folge stetiger Funktionen $f_n : D \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$, die gleichmäßig gegen die Funktion g konvergiert. Dann ist g stetig.

8.2 Differenzierbarkeit

8.2.1 Kurven im \mathbb{R}^p

Definition 8.25 Sei $J \subset \mathbb{R}^p$ ein Intervall und $x : J \rightarrow \mathbb{R}^p$, $t \mapsto x(t) = (x_1(t), \dots, x_p(t))^T$ eine Funktion.

- x heißt differenzierbar in t_0 , wenn der Grenzwert $\lim_{t \rightarrow t_0} (x(t) - x(t_0)) / (t - t_0) =: \dot{x}(t_0)$ existiert.
- x heißt differenzierbar, wenn x in jedem Punkt differenzierbar ist. Ist dann $t \rightarrow \dot{x}(t)$ auch noch stetig, so heißt x stetig differenzierbar.

Weil man Grenzwerte koordinatenweise bilden kann, ist x genau dann stetig differenzierbar, wenn jede Koordinatenfunktion dies ist.

Definition 8.26 Sei $x : J \rightarrow \mathbb{R}^p$ stetig differenzierbar und \dot{x} habe nur endlich viele Nullstellen. Dann nennt man das Bild $x(J)$ eine **Kurve** und $x : J \rightarrow x(J)$ heißt eine **Parametrisierung** dieser Kurve. Ist $\dot{x}(t) \neq 0$, so heißt $\dot{x}(t)$ Steigung der Tangente an $x(t)$.

Appletanleitung

Beispiele:

- $x(t) = r(\cos(t), \sin(t))^T$ ($r > 0$). Das ist eine Kreislinie. $\dot{x}(t) = r(-\sin(t), \cos(t))^T$ steht senkrecht auf $x(t)$.
- $x(t) = (r \cos(t), r \sin(t), \alpha t)^T$ ($\alpha \neq 0$). Dies ist eine Schraubenlinie.
 $\dot{x}(t) = (-r \sin(t), r \cos(t), \alpha)^T$.
- $x(t) = r \begin{pmatrix} \cos t \cos \vartheta \\ \sin t \cos \vartheta \\ \sin \vartheta \end{pmatrix}$. Dies ist ein Breitengrad auf einer Kugel.

* * * TO DO * * *

Wir können die Länge einer Kurve zwischen zwei Punkten berechnen:

Seien $a < b$, $a, b \in J$ und $C = x(J)$, wo $x : J \rightarrow \mathbb{R}^p$ stetig differenzierbar mit nicht verschwindender Ableitung ist.

Teilen wir das Intervall $[a, b]$ in n gleiche Teile ein, wo n sehr groß ist, so wird die Länge der Kurve zwischen $P = x(a)$ und $Q = x(b)$ ungefähr gleich $\sum_{j=1}^n \|x(a_k) - x(a_{k-1})\|$ sein wo $a_k = a + \frac{(b-a)}{n}k$ ist. Es ist $x_j(a_k) - x_j(a_{k-1}) \approx \dot{x}_j(a_{k-1})(a_k - a_{k-1})$. Also erhält man $\|x(a_k) - x(a_{k-1})\| \approx \|\dot{x}(a_{k-1})\|(a_k - a_{k-1})$ und damit $\sum_{j=1}^n \|x(a_k) - x(a_{k-1})\| \approx \int_a^b \|\dot{x}(t)\| dt$.

Also setzen wir

Definition 8.27 Sei $x : J \rightarrow \mathbb{R}^p$ stetig differenzierbar mit nicht verschwindender Ableitung und $a < b$, $a, b \in J$. Dann ist $\int_a^b \|\dot{x}(t)\| dt$ die **Länge der Kurve** $x(J)$ zwischen den Punkten $x(a)$ und $x(b)$. Hält man a fest, so ergibt sich $s(t) = \int_a^t \|\dot{x}(s)\| ds$, die **Bogenlänge** der Kurve als Funktion des Kurvenparameters.

Wegen $\dot{s}(t) = \|\dot{x}(t)\| \neq 0$ nach Voraussetzung kann man die Umkehrfunktion $t = t(s)$ bilden und erhält $y(s) = x(t(s))$. Diese Darstellung ist die **kanonische Parametrisierung** durch die **Bogenlänge**.

Aufgabe:

- a) Zeigen Sie bitte $\|y'(s)\| = 1$, d.h. $y'(s)$ ist die **normierte Tangentenrichtung**.
- b) Sei y 2 x stetig differenzierbar. Zeigen Sie bitte: $y''(s)$ steht senkrecht auf $y'(s)$.
Tip: Differenzieren Sie $1 = (y'(s) | y'(s))$.
- c) Berechnen Sie die Bogenlänge der Schraubenlinie $x(t) = (\cos t, \sin(t), t)^T$.

8.2.2 Differentiation von Funktionen mehrerer unabhängiger Variablen

Richtungsableitung, partielle Ableitungen

Die Frage ist: wie ändert sich eine Funktion $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ in der Nähe von x , wenn man ein Stück in Richtung $h \neq 0$ geht?

Appletanleitung Schauen Sie sich skalare Funktionen in einer vorgegebenen Richtung an. Geben Sie dazu eine skalare Funktion, die Stelle $x^{(0)}$ und die Richtung $h = (h_1, h_2)^T$ ein. Sie sehen eine Ebene senkrecht zur (x, y) -Ebene in Richtung von h . Sie schneidet aus der durch die Funktion gegebene Fläche eine Kurve aus. Die Richtung der Tangente an diese Kurve im Punkt $x^{(0)}$ ist die Richtungsableitung. Studieren Sie so die partiellen Ableitungen als Spezialfälle von Richtungsableitungen.

* * * TO DO * * *

Definition 8.28 Sei $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ und $x \in U$, $0 \neq h \in \mathbb{R}^p$. f heißt in x in Richtung h differenzierbar, wenn der Grenzwert

$$D_h(f)(x) = \lim_{t \rightarrow 0} \frac{1}{t} (f(x + th) - f(x))$$

existiert. $D_h(f)(x)$ heißt **Richtungsableitung in Richtung h** .

Für $h = e_j = (\delta_{jk})_{k=1, \dots, p}$ heißt $D_h(f)(x)$ **partielle Ableitung $\frac{\partial f(x)}{\partial x_j}$ oder $f_{x_j}(x)$ nach der Variablen x_j** .

Nach unseren Rechenregeln für Grenzwerte ist für $f = (f_1, \dots, f_q)^T$ stets

$$D_h f = (D_h f_1, \dots, D_h f_q)^T,$$

insbesondere

$$\frac{\partial f}{\partial x_j} = \left(\frac{\partial f_1}{\partial x_j}, \dots, \frac{\partial f_q}{\partial x_j} \right)^T.$$

Wir erhalten die partielle Ableitung nach x_j an der Stelle $x^{(0)} = (x_1^{(0)}, \dots, x_p^{(0)})$ indem wir alle Variablen außer x_j festhalten, also die Funktion $t \rightarrow f(x_1^{(0)}, \dots, x_{j-1}^{(0)}, t, x_{j+1}^{(0)}, \dots, x_p^{(0)})$ nach t differenzieren.

Beispiele:

- 1. $f(x, y) = xy, \quad \frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x$
- 2. $f(x, y) = \sin(xy), \quad f_x = y \cos(xy), \quad f_y = x \cos(xy).$

3. $f(x, y) = \begin{cases} \frac{2xy}{x^2 + y^2} & 0 \neq x, y \\ 0 & x = y = 0 \end{cases}$ $f_x(0, 0) = 0 = f_y(0, 0)$. Sei h beliebig ($\neq 0$). Dann ist $D_h f(0, 0) = 0$. Aber f ist in 0 nicht stetig!
4. $f(x_1, \dots, x_p) = x^\alpha = x_1^{\alpha_1} \cdots x_p^{\alpha_p}$. Ist $\alpha_j = 0$, so ist $f_{x_j} = 0$, andernfalls ist $f_{x_j} = \alpha_j x_1^{\alpha_1} \cdots x_{j-1}^{\alpha_{j-1}} \cdot x_j^{\alpha_j - 1} \cdots x_p^{\alpha_p}$.
5. $f(x) = \|x\| = \sqrt{x_1^2 + \cdots + x_p^2}$. Für $x_j \neq 0$ ist $f_{x_j}(x) = \frac{x_j}{\|x\|}$.

Wir führen gleich auch höhere partielle Ableitungen ein:

Definition 8.29 a) $f : U \rightarrow \mathbb{R}^q$ heißt **stetig differenzierbar**, wenn f überall nach allen Variablen differenzierbar ist und die Ableitungen $\frac{\partial f}{\partial x_j} : U \rightarrow \mathbb{R}^q$ stetig sind.

b) f heißt **2-mal stetig differenzierbar**, wenn f einmal stetig differenzierbar ist und die Ableitungen $\frac{\partial f}{\partial x_j}$ wiederum stetig differenzierbar sind; ihre partiellen Ableitungen sind dann $\frac{\partial^2 f}{\partial x_k \partial x_j}$.

c) Seien $\frac{\partial^n f}{\partial x_{k_1} \cdots \partial x_{k_n}}$ bereits erklärt und stetig differenzierbar. Dann heißt f $(n+1)$ mal stetig differenzierbar und die partiellen Ableitungen $(n+1)$. Ordnung sind $\frac{\partial^{n+1} f}{\partial x_\ell \partial x_{k_1} \cdots \partial x_{k_n}}$

Aufgaben: Zeigen Sie bitte, daß die folgenden Funktionen 2 mal stetig differenzierbar sind und berechnen Sie alle partiellen Ableitungen 1. und 2. Ordnung.

1. $f(x, y) = x^2 y^3$.
2. $f(x_1, \dots, x_p) = x_1^{\alpha_1} \cdots x_p^{\alpha_p}$.
3. $f(x) = \|x\|^2$.
4. $f(x) = \mathcal{A}x$, \mathcal{A} eine $q \times p$ -Matrix.
5. $f(u, v) = r \begin{pmatrix} \cos u \cos v \\ \sin u \cos v \\ \sin v \end{pmatrix}$ (Kugelsphäre mit Radius r).
6. $f(u, v) = \begin{pmatrix} r \cos u \\ r \sin u \\ v \end{pmatrix}$ (Zylindermantel).
7. $f(r, \varphi) = r \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$ (Polarkoordinaten).

Bei n mal stetig differenzierbaren Funktionen kommt es auf die Reihenfolge der Differentiation nicht an.

Satz 8.30 (H.A. Schwarz) Sei $f : U \rightarrow \mathbb{R}^q$ zweimal stetig differenzierbar. Dann ist stets

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Korollar 8.31 Sei $f : U \rightarrow \mathbb{R}^q$ n mal stetig differenzierbar. Dann kommt es bei der Berechnung der partiellen Ableitungen k -ter Ordnung ($k \geq n$) nicht auf die Reihenfolge der Ableitungen an.

Der Beweis des Satzes ist etwas technisch und wird hier ausgelassen.

Totale Ableitung

Anleitung zum Applet "Parameterflächen im Raum" Schauen Sie sich die Funktion

$$f(x, y) = \begin{cases} 0 & (x, y) = (0, 0) \\ \frac{2xy}{x^2 + y^2} & \text{sonst} \end{cases} \text{ in der Nähe des Nullpunktes an. (Wählen Sie den } x\text{- und } y\text{-}$$

Achsenabschnitt betragsmäßig klein.) Sie sieht dort wie ein Gebirge aus und besitzt anschaulich keine Tangentialebene im Nullpunkt, obwohl die partiellen Ableitungen existieren. (Variieren Sie den vertikalen Betrachtungswinkel. Bei 100 Grad sieht man eine scharfen Grad am Nullpunkt.)

Möchten wir in einem Punkt $x^{(0)} \in U$ an den Graphen der Funktion $f : U \rightarrow \mathbb{R}$ eine Tangentialebene anlegen, so reicht die partielle Differenzierbarkeit im allgemeinen nicht aus. Wir definieren daher:

Definition 8.32 Sei $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ eine Funktion. f heißt in $x^{(0)} \in U$ total differenzierbar, wenn es eine lineare Abbildung A mit

$$\lim_{h \rightarrow 0} \frac{f(x^{(0)} + h) - f(x^{(0)}) - Ah}{\|h\|} = 0.$$

A heißt dann totale Ableitung $f'(x^{(0)})$ an der Stelle $x^{(0)}$.

Den Zusammenhang zu partiellen Ableitungen klärt der folgende Satz:

Theorem 8.33 a) Sei $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ in $x^{(0)}$ total differenzierbar. Dann existiert jede Richtungsableitung $D_h f(x^{(0)})$ und es gilt

$$D_h f(x^{(0)}) = f'(x^{(0)})h.$$

Insbesondere hat $f'(x^{(0)})$ bezüglich der kanonischen Basen die Matrix

$$f'(x^{(0)}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^{(0)}), \dots, \frac{\partial f_1}{\partial x_p}(x^{(0)}) \\ \vdots \\ \frac{\partial f_q}{\partial x_1}(x^{(0)}), \dots, \frac{\partial f_q}{\partial x_p}(x^{(0)}) \end{pmatrix}. \text{ Diese Matrix heißt } \mathbf{Jacobi-Matrix}$$

b) Sei f stetig partiell differenzierbar auf U . Dann ist f in jedem Punkt $x \in U$ total differenzierbar und die Funktion $x \rightarrow f'(x)$ ist stetig von $U \rightarrow L(\mathbb{R}^p, \mathbb{R}^q)$.

Beweis: a) Wähle h_0 als feste Richtung und $(t_n) \subset \mathbb{R}$ mit $\lim t_n = 0$. Nach Voraussetzung ist

$$\lim_{n \rightarrow \infty} \frac{f(x + t_n h_0) - f(x) - t_n A h_0}{|t_n| \|h_0\|} = 0.$$

Ist $t_n > 0$ für alle n , so folgt $\frac{1}{\|h_0\|} A(h) = \lim_{n \rightarrow \infty} \frac{f(x^{(0)} + t_n h_0) - f(x)}{t_n \|h_0\|}$. Ist $t_n < 0$ für alle n , so folgt

$$-\frac{1}{\|h_0\|} A(h_0) = \lim_{n \rightarrow \infty} \frac{f(x^{(0)} + t_n h_0) - f(x)}{|t_n| \|h_0\|},$$

also

$$\frac{1}{\|h_0\|} A(h_0) = \lim_{n \rightarrow \infty} \frac{f(x^{(0)} + t_n h_0) - f(x)}{t_n \|h_0\|},$$

und damit

$$D_{h_0} f(x^{(0)}) = A h_0.$$

Insbesondere ergibt sich für $h_0 = e_j$ $\frac{\partial f}{\partial x_j} = A e_j$, woraus die Matrixgestalt folgt.

b) Es genügt, die Aussage für $f : U \rightarrow \mathbb{R}$ zu zeigen. Wir beschränken uns auf den Fall $p = 2$. Es ist $f(x+h) - f(x) = f(x+h) - f(x+h_1 e_1) + f(x+h_1 e_1) - f(x)$. Auf die erste Differenz wendet man den 1. Mittelwertsatz (Theorem 4.13) an und erhält

$$f(x+h_1 e_1 + h_2 e_2) - f(x+h_1 e_1) = f_{x_2}(x+h_1 e_1 + \vartheta_2 h_2 e_2) h_2.$$

Ebenso ist

$$f(x+h_1 e_1) - f(x) = f_{x_1}(x+\vartheta_1 h_1 e_1) h_1.$$

Sei $A(x) = (f_{x_1}(x), f_{x_2}(x))$. Dann ergibt sich

$$f(x+h) - f(x) - A(x)h = (f_{x_1}(x+h_1 e_1) - f_{x_1}(x), f_{x_2}(x+h_1 e_1 + \vartheta_2 h_2 e_2) - f_{x_2}(x)) \cdot h.$$

Also

$$\frac{f(x+h) - f(x) - A(x)h}{\|h\|} = (f_{x_1}(x+h_1 e_1) - f_{x_1}(x), f_{x_2}(x+h_1 e_1 + \vartheta_2 h_2 e_2) - f_{x_2}(x)) \cdot \frac{h}{\|h\|}.$$

Wegen der Stetigkeit der partiellen Ableitungen konvergiert die rechte Seite für h gegen 0 gegen 0.

Aufgaben: Zeigen Sie bitte:

1. Seien $f, g : U \rightarrow \mathbb{R}^q$ stetig differenzierbar. Dann ist $(f+g)'(x) = f'(x) + g'(x)$.
2. Seien f, g wie oben. Dann ist für $h(x) = (f(x) \mid g(x))$ $h'(x) = f'(x)^T g(x) + g'(x)^T f(x)$
3. Sei $f : U \rightarrow \mathbb{R}^q$ stetig differenzierbar und $\varphi : U \rightarrow \mathbb{R}$ ebenfalls. Dann ist $(\varphi f)'(x) = f(x)\varphi'(x) + \varphi(x)f'(x)$

Tip: Berechnen Sie die partiellen Ableitungen und benutzen Sie dabei die Ableitungsregeln für Funktionen einer unabhängigen Veränderlichen (siehe Satz 4.4).

Als wichtiges Instrument zur Berechnung von Ableitungen beweisen wir die Kettenregel, wofür wir einen Hilfssatz benötigen (vergleiche Satz 4.2 c):

Lemma 8.34 $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ ist genau dann in c total differenzierbar, wenn es eine in c stetige Funktion $S_f : U \rightarrow L(\mathbb{R}^p, \mathbb{R}^q)$ gibt mit

$$f(y) = f(c) + S_f(y)(y-c). \tag{8.6}$$

In diesem Fall ist $S_f(c) = f'(c)$.

Beweis: (I) Sei f in c total differenzierbar. Setze

$$R_f(y) = \begin{cases} \frac{f(y) - f(c) - f'(c)(y - c)}{\|y - c\|} & y \neq c \\ 0 & y = c \end{cases}$$

und

$$S_f(y) = \begin{cases} f'(c) & y = c \\ f'(c) + \frac{R_f(y)}{\|y - c\|} (y - c) & y \neq c \end{cases}.$$

$S_f(y)$ ordnet als jedem Vektor $v \in \mathbb{R}^p$ den Vektor

$$S_f(y)(v) = f'(c)(v) + \frac{R_f(y)}{\|y - c\|} (y - c | v) \in \mathbb{R}^q$$

zu. Dann folgt $\lim_{y \rightarrow c} S_f(y) = f'(c) = S_f(c)$ und $f(y) = f(c) + S_f(y)(y - c)$.

(II) Gilt die Formel (8.6), so ist offensichtlich die Bedingung für die totale Differenzierbarkeit erfüllt.

Theorem 8.35 (Kettenregel) Seien $f : U \subset \mathbb{R}^p \rightarrow V \subset \mathbb{R}^q$ und $g : V \rightarrow \mathbb{R}^r$ stetig differenzierbar. Dann ist $g \circ f : U \rightarrow \mathbb{R}^r$ stetig differenzierbar und es gilt $(g \circ f)'(x) = g'(f(x))f'(x)$.

Korollar 8.36 a) Sei $x : J \subset \mathbb{R} \rightarrow U \subset \mathbb{R}^p$ eine Kurvenparametrisierung und $f : U \rightarrow \mathbb{R}^q$ stetig differenzierbar. Dann ist $(f \circ x)'(t) = f'(x(t))x'(t)$ die Richtungsableitung in Richtungen der Tangente an $x(t)$.

b) Seien f, g wie im Theorem. Dann ist

$$\frac{\partial(g \circ f)(x)}{\partial x_j} = \sum_{k=1}^q \frac{\partial g(f(x))}{\partial y_k} \frac{\partial f_k(x)}{\partial x_j}$$

Teil a) des Korollars folgt aus dem Theorem für $p = 1$ und $f(t) = x(t)$.

Teil b) folgt daraus, daß $\frac{\partial(g \circ f)}{\partial x_j}(x)$ die j . Spalte von $(g \circ f)'(x)$ ist. Man erhält diese Spalte als $g'(f(x)) \frac{\partial f(x)}{\partial x_j}$ nach den Regeln der Matrizenmultiplikation.

Beweis: (Beweis des Theorems)

Es ist nach dem Lemma

$$g(f(y)) - g(f(x)) = S_g(f(y))(f(y) - f(x)) = S_g(f(y))S_f(y)(y - x).$$

Da die Funktion $S_{g \circ f}(y) := S_g(f(y))S_f(y)$ nach den Rechenregeln für stetige Funktionen stetig ist, folgt nach dem Lemma die Differenzierbarkeit und es gilt

$$(g \circ f)'(x) = S_{g \circ f}(x) = g'(f(x))f'(x).$$

Aufgaben: Berechnen Sie im folgenden die totale Ableitung an allen Stellen, an denen sie existiert.

1. $f(x_1, x_2, x_3) = x_1 x_2 + x_2^2$.

2. $f(x, y) = \sin(\sqrt{x^2 + y^2})$. Existiert die totale Ableitung in 0?
3. $f(x, y) = \exp(x \cos(y^2))$
4. $f(u, v, w) = (u, v^2, \exp(u^2 - w^2))^T$

Für skalare Funktionen $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$ ist die totale Ableitung ein Zeilenvektor:

$$f'(x_1, \dots, x_p) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_p} \right).$$

Definition 8.37 Der Spaltenvektor $f'(x)^T = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_p} \end{pmatrix}$ heißt **Gradient** $\text{grad}f(x)$.

Die Ableitung in Richtung h ist damit $D_h f(x) = f'(x)h = (\text{grad}f(x) \mid h)$, läßt sich also als Skalarprodukt schreiben. Damit erhalten wir folgende interessante geometrische Interpretation:

Satz 8.38 (Eigenschaften des Gradienten)

- a) Sei $J \ni t \rightarrow x(t)$ eine Kurvenparametrisierung mit $f(x(t)) = c$ ($x(J)$ heißt dann Höhenlinie). Dann ist $\text{grad}f(x(t)) \perp \dot{x}(t)$, der Gradient steht senkrecht auf Höhenlinien.
- b) Es ist $\text{grad}f(x)$ der Vektor in Richtung des steilsten Anstiegs von f . Das heißt

$$\|\text{grad}f(x)\| = \max\{|D_h f(x)| : \|h\| = 1\}$$

Beweis: a)

Nach der Kettenregel ist wegen $f(x(t)) = c$:

$$0 = \frac{d}{dt} f(x(t)) = f'(x(t))\dot{x}(t) = (\text{grad}f(x(t)) \mid \dot{x}(t)).$$

b) Für $\|h\| = 1$ ist nach der Cauchy-Schwarzschen Ungleichung

$$\begin{aligned} |D_h f(x)| &= |(\text{grad}f(x) \mid h)| \leq \|\text{grad}f(x)\| \cdot \|h\| \\ &= \|\text{grad}f(x)\| = |(\text{grad}f(x) \mid \frac{\text{grad}f(x)}{\|\text{grad}f(x)\|})| \\ &= |D_{h_o} f(x)| \quad \text{für } h_o = \frac{\text{grad}f(x)}{\|\text{grad}f(x)\|} \end{aligned}$$

Appletanleitung Bestimmen Sie die Höhenlinien, indem Sie $c = f(x_1, \dots, x_p)$ nach x_1 oder x_2, \dots, x_p auflösen. Überzeugen Sie sich vom Satz oben, indem Sie die Höhenlinien und den Gradienten bestimmen und anschauen.

1. $f(x, y) = x^2 + y^2$

2. $f(x, y) = x^2 - y^2$
3. $f(x, y) = 2xy$
4. $f(x, y, z) = z - (x^2 + y^2)$.

* * * T O D O * * *

8.3 Erste Anwendungen

Überblick: Wir erhalten auf konvexen Mengen einen Ersatz für den Mittelwertsatz in einer Veränderlichen. Außerdem entwickeln wir auch eine skalare Funktion in mehreren Veränderlichen nach Taylor und benutzen dies zur Charakterisierung von lokalen Extremstellen.

Definition 8.39 Eine Menge $\emptyset \neq C \subset \mathbb{R}^p$ heißt **konvex** wenn mit je zwei Punkten a und b aus C die ganze Verbindungsstrecke $[a, b] := \{a + t(b - a) : 0 \leq t \leq 1\}$ in C liegt.

Beispiele:

1. $B(x, r)$ ist konvex.
2. $\{(x, y) \in \mathbb{R}^2, 0 \leq y \leq x\}$ ist konvex.
3. Jedes Dreieck ist konvex.
4. Ellipsen sind konvex.
5. $M = \{(x, y) : 0 \leq y \leq x^2\}$ ist nicht konvex.

Wir erhalten den folgenden Satz, der unter anderem bei Iterationsverfahren eine Rolle spielt (siehe den Umkehrsatz).

Satz 8.40 (Schränkensatz)

Sei U offen in \mathbb{R}^p , $f : U \rightarrow \mathbb{R}^q$ sei stetig differenzierbar und $C \subset U$ sei eine kompakte konvexe Menge. Dann gilt für alle $x, y \in C$

$$\|f(y) - f(x)\| \leq \sup\{\|f'(z)\| : z \in C\} \cdot \|y - x\|.$$

Beweis: Sei erst einmal $q = 1$, f also eine skalare Funktion. Zunächst ist $z \rightarrow f'(z) \rightarrow \|f'(z)\|$ stetig auf der kompakten Menge C , das Supremum $L = \sup\{\|f'(z)\| : z \in C\}$ existiert (und wird sogar angenommen). Sei $x(t) = x + t(y - x)$. Die Strecke liegt in C . Dann ist nach dem ersten Mittelwertsatz

$$f(y) - f(x) = f(x(1)) - f(x(0)) = (f \circ x)'(\vartheta) \cdot (1 - 0)$$

für ein $\vartheta \in]0, 1[$. Nach der Kettenregel ist $(f \circ x)'(\vartheta) = f'(x(\vartheta)) \cdot x'(\vartheta)$ und $x'(\vartheta) = (y - x)$. Also ist

$$|f(y) - f(x)| \leq \|f'(x(\vartheta))\| \|y - x\| \leq L \|y - x\|.$$

Der allgemeine Fall $q \geq 1$ folgt nun durch sorgfältige Betrachtung der einzelnen Komponenten f_j . Mit dem gleichen Trick erhalten wir den Satz von Taylor. Zunächst definieren wir die Hessematrix.

Definition 8.41 Sei $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$ 2 mal stetig differenzierbar. Dann heißt die Ableitung $(\text{grad} f)' = (\frac{\partial^2 f}{\partial x_i \partial x_j})_{i,j=1,\dots,p} = H(f)$ **Hessematrix** von f .

Aufgaben: Berechnen Sie die Hessematrix der folgenden Funktionen und bestimmen Sie deren Eigenwerte.

1. $f(x, y) = x^2 + y^2 - 1$
2. $f(x, y) = x^2 - y^2 + 1$
3. $f(x, y) = \sin(x^2) + \sin(y^2)$
4. $f(x, y) = \exp(x + y)$
5. $f(x, y, z) = x^2 + y^2 + \exp(z)$

Wie im Fall $p = 1$ kann man glatte Funktionen durch Polynome annähern.

Theorem 8.42 (Satz von Taylor) Sei $\emptyset \neq U$ offen in \mathbb{R}^p und $f : U \rightarrow \mathbb{R}$ 2 mal stetig differenzierbar. Seien x und $y \neq x$ aus U und es liege die ganze Strecke $[x, y] = \{x + t(y - x) : 0 \leq t \leq 1\}$ in U . Dann gibt es einen Punkt $x_0 \in [x, y]$, so daß gilt

$$f(y) = f(x) + (\text{grad} f(x) \mid y - x) + \frac{1}{2}((y - x) \mid H(f)(x_0)(y - x)).$$

$(H(f)(x_0))$ ist die Hessematrix von f an der Stelle x_0 .

Beweis: Sei $g(t) = f(x + t(y - x))$. Wir wenden auf g den eindimensionalen Satz von Taylor an und erhalten

$$f(y) - f(x) = g(1) - g(0) = g'(0)(1 - 0) + \frac{1}{2}g''(\vartheta)(1 - 0)^2$$

mit $0 < \vartheta < 1$. Anwendung der Kettenregel liefert die Behauptung.

Wir wenden diesen Satz sofort auf Extremwertprobleme an. Dazu brauchen wir das folgende Lemma, das wir dann auf die Hessematrix anwenden, die nach dem Satz von Schwarz symmetrisch ist.

Lemma 8.43 Sei A eine symmetrische $p \times p$ -Matrix. Dann sind die folgenden Aussagen äquivalent:

- a) Für alle $x \neq 0$ gilt $(x \mid Ax) > 0$.
- b) A hat nur positive Eigenwerte.

Beweis: a) \Rightarrow b) Annahme b) gilt nicht. Dann gibt es einen Eigenwert $\lambda \leq 0$ (da A symmetrisch ist, sind alle Eigenwerte nach Korollar 7.64 reell). Sei $0 \neq x$ ein zugehöriger Eigenwert. Dann ist $(x \mid Ax) = \lambda(x \mid x) \leq 0$, ein Widerspruch zu a).

b) \Rightarrow a) Nach Korollar 7.64 hat A eine Orthonormalbasis $\{b_1, \dots, b_p\}$ aus Eigenvektoren. Also ist $x = \sum \xi_j b_j$ und

$$\begin{aligned} (x \mid Ax) &= \left(\sum_{j,k} \xi_j \xi_k (b_j \mid Ab_k) \right) \\ &= \sum_{j,k} \xi_j \xi_k (b_j \mid \lambda_k b_k) \\ &= \sum_{\ell} \xi_{\ell}^2 \lambda_{\ell} > 0, \end{aligned}$$

wobei die letzte Gleichheit aus $(b_j \mid b_k) = \delta_{jk}$ folgt.

Definition 8.44 Die symmetrische Matrix A heißt **positiv definit**, wenn sie eine der beiden äquivalenten Bedingungen aus dem Lemma erfüllt. Sie heißt **negativ definit**, wenn $-A$ positiv definit ist.

Wir definieren wie in Abschnitt 4 lokale Extremalstellen.

Definition 8.45 Sei $\emptyset \neq U \subset \mathbb{R}^p$ offen, $f_U \rightarrow \mathbb{R}$ eine skalare Funktion.

$x \in U$ heißt **lokale** $\left\{ \begin{array}{l} \text{Minimalstelle} \\ \text{Maximalstelle} \end{array} \right\}$, wenn es eine Kugel $B(x, r) \subset U$ gibt mit $f(y) \left\{ \begin{array}{l} \geq \\ \leq \end{array} \right\} f(x)$ für alle $y \in B(x, r)$.

$f(x)$ selbst heißt dann **lokales** $\left\{ \begin{array}{l} \text{Minimum} \\ \text{Maximum} \end{array} \right\}$.

Will man sich nicht festlegen, ob es ein Minimum oder Maximum ist, so spricht man von einem **lokalen Extremwert** bzw. einer **lokalen Extremstelle**.

Appletanleitung Schauen Sie sich die folgenden Funktionen an. Wo liegen lokale Minima, wo lokale Maxima? Bestimmen Sie dort (rechnerisch) den Gradienten und die Hessematrix.

1. $f(x, y) = x^2 + y^2$
2. $f(x, y) = x^2 - y^2$
3. $f(x, y) = \cos(x^2 + y^2)$
4. $f(x, y) = \sqrt{4 - x^2 - 2y^2}$

* * * TO DO * * *

An den Beispielen konnten Sie vielleicht folgenden Satz entdecken:

Satz 8.46 (lokale Extremwerte) Sei $\emptyset \neq U \subset \mathbb{R}^p$ offen und $f : U \rightarrow \mathbb{R}$ sei eine skalare Funktion.

a) Sei f einmal stetig differenzierbar. Ist x_0 eine lokale Extremalstelle, so ist $\text{grad}f(x_0) = 0$.

b) Sei f 2 mal stetig differenzierbar und $\text{grad}f(x_0) = 0$.

Ist die Hessematrix $H(f)(x_0) \left\{ \begin{array}{l} \text{positiv} \\ \text{negativ} \end{array} \right\}$ definit, so ist $f(x_0)$ ein lokales $\left\{ \begin{array}{l} \text{Minimum} \\ \text{Maximum} \end{array} \right\}$.

Beweis: a)

Sei $1 \leq j \leq p$ beliebig und x_0 eine lokale Extremalstelle von f . Dann ist 0 eine lokale Extremalstelle der Funktion einer Veränderlichen $g(s) = f(x_0 + se_j)$. Also ist $\frac{\partial f}{\partial x_j}(x_0) = g'(0) = 0$. Da j beliebig war, folgt die Behauptung.

b)

Wegen $f'(x_0) = 0$ gibt es nach dem Satz von Taylor ein x_1 auf der Strecken zwischen x_0 und x mit

$$f(x) = f(x_0) + \frac{1}{2}(x - x_0 | H(f)(x_1)(x - x_0)).$$

Sei $H(f)(x_0)$ positiv definit und x sehr dicht bei x_0 . Dann kann man zeigen, daß $H(f)$ auf der ganzen Strecke zwischen x_0 und x positiv definit ist. Also ist der 2. Summand der rechten Seite echt größer als 0. Damit ist $f(x) > f(x_0)$. Da x (nahe bei x_0) beliebig war, ist $f(x_0)$ ein lokales Minimum. Ist $H(f)(x_0)$ negativ definit, so betrachte $-f$.

Aufgabe: Untersuchen Sie, wo bei den Funktionen, die Sie sich vor Satz 1.57 angeschaut haben, lokale Extrema sind und bestimmen Sie, ob es sich um Minima oder Maxima handelt.

8.4 Der Umkehrsatz und seine Anwendungen

Überblick: Sei $f : I \subset \mathbb{R}^p \rightarrow \mathbb{R}^p$ stetig differenzierbar und $f'(x)$ sei bijektiv (invertierbar). Im Gegensatz zum Fall einer Veränderlichen kann man hieraus nicht schließen, daß man zu f eine inverse Abbildung f^{-1} hat. f läßt sich aber lokal umkehren. Als Folgerung erhält man den Satz über implizit (durch Gleichungen) gegebene Funktionen und den Satz über Extrema mit Nebenbedingungen.

8.4.1 Der Umkehrsatz

Beispiel: Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \rightarrow e^x \begin{pmatrix} \cos y \\ \sin y \end{pmatrix}$.

Die Ableitung $f'(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}$ hat die Determinante $e^{2x} \neq 0$. Trotzdem ist f nicht injektiv. Denn es ist $f(x, y) = f(x, y + 2k\pi) \quad (k \in \mathbb{Z})$. Aber man hat die Funktion f eingeschränkt auf $\mathcal{U} = \mathbb{R} \times]-\pi/2, \pi/2[$, so ist sie dort injektiv, $f(\mathcal{U}) = V$ ist offen und die Umkehrabbildung $f^{-1} : V \rightarrow \mathcal{U}$ ist stetig differenzierbar. $f^{-1}(u, v) = \begin{pmatrix} \ln(u) \\ \arctan(\frac{v}{u}) \end{pmatrix}$.

Dieser Sachverhalt gilt ganz allgemein.

Theorem 8.47 (Satz über die Umkehrfunktion)

Sei $f : \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}^p$ stetig differenzierbar. Sei für ein $x_0 \in \mathcal{U} \quad \det(f'(x_0)) \neq 0$. Dann gibt es eine offene zusammenhängende Menge $\mathcal{U}_0 \subset \mathcal{U}$ mit

1. $x_0 \in \mathcal{U}_0$ und $\det(f'(x)) \neq 0$ für alle $x \in \mathcal{U}_0$.
2. $f|_{\mathcal{U}_0}$ ist injektiv und $V = f(\mathcal{U}_0)$ ist offen.
3. Die Umkehrfunktion $g : V \rightarrow \mathcal{U}_0$ von $f|_{\mathcal{U}_0}$ ist stetig differenzierbar und es gilt

$$g'(f(x)) = (f'(x))^{-1}$$

Der Beweis ist nicht einfach (siehe Analysis Alive, S. 295 ff).

Aufgaben: Wenden Sie bitte den Satz auf die folgenden Abbildungen und angegebenen Stellen an, d.h. bestimmen Sie ein \mathcal{U}_0 , berechnen Sie V und $f^{-1}|_{\mathcal{U}_0} = g$, sowie g' . Sie können \mathcal{U}_0 klein und günstig wählen, es kommt auf das Prinzip an.

1. Sei A eine $p \times p$ -Matrix, $\det(A) \neq 0$, $f(x) = Ax$ und $x_0 = 0$.
2. Sei $f(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$; $\begin{pmatrix} r_0 \\ \varphi_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

$$3. \text{ Sei } f(r, \varphi, \vartheta) = r \begin{pmatrix} \cos \varphi & \cos \vartheta \\ \sin \varphi & \cos \vartheta \\ \sin \vartheta & \end{pmatrix}; \quad \begin{pmatrix} r_0 \\ \varphi_0 \\ \vartheta_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

8.4.2 Implize Funktionen

Die Kreisgleichung $x^2 + y^2 = 1$ kann man sowohl nach y als auch nach x stetig differenzierbar auflösen: $y(x) = \sqrt{1-x^2}$ oder $x(y) = \sqrt{1-y^2}$. Je nachdem, welchen der Halbbögen man möchte, muß man die positive oder negative Wurzel wählen. Ist man zum Beispiel interessiert daran, daß der Punkt $(0, 1)$ im Graphen der Funktion liegt und außerdem die Funktion differenzierbar ist, muß man $y(x) = +\sqrt{1-x^2}$ wählen. (Warum gehen die anderen Möglichkeiten nicht?)

Wir haben also eine Funktion $F: \mathbb{R}^2 \rightarrow \mathbb{R}^1$, nämlich $F(x, y) = x^2 + y^2 - 1$ und suchen eine Funktion $y: \mathbb{R}^1 \rightarrow \mathbb{R}^1$, $x \rightarrow \sqrt{1-x^2} = y(x)$ derart, daß sie die Gleichung $F(x, y(x)) = 0$ löst. Etwas komplizierter ausgedrückt: ihr Graph $G = \{(x, y(x)) : (x) < 1\}$ liegt in der Menge $M = \{(x, y) : x^2 + y^2 - 1 = 0\}$. Man sagt: y ist durch $F \equiv 0$ implizit definiert.

Bezeichnungen Sei $p = q + r$, wo $p, q, r \in \mathbb{N}$ und $1 \leq q < p$, also $r \neq 0$ gilt.

Im folgenden schreiben wir $z = (z_1, \dots, z_p) = (x, y)$ mit $x = (x_1, \dots, x_q)$, also $x_j = z_j$ für $j \leq q$ und $y = (y_1, \dots, y_r)$, also $y_k = z_{q+k}$. Damit haben wir $\mathbb{R}^p = \mathbb{R}^q \times \mathbb{R}^r$ und die Analogie zum einführenden Beispiel ($q = r = 1, p = 2$) wird deutlich.

Sei $F: \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}^r$ stetig differenzierbar. Dann setzen wir

$$F_x = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_q} \right), \quad F_y = \left(\frac{\partial F}{\partial y_1}, \dots, \frac{\partial F}{\partial y_r} \right).$$

F_y ist also eine $r \times r$ -Matrix und es ist die Jacobimatrix $F'(z) = (F_x, F_y)$ (bis auf die beiden überflüssigen Klammern), (in der Mitte zwischen F_x und F_y). Damit können wir schon den Satz über implizite Funktionen formulieren:

Theorem 8.48 (Satz über implizite Funktionen)

Sei $\mathcal{U} \subset \mathbb{R}^p$ offen, $F: \mathcal{U} \rightarrow \mathbb{R}^r$ sei stetig differenzierbar und es gebe ein $z_0 = (x_0, y_0)$ mit

1. $F(x_0, y_0) = 0$,
2. $\det(F_y(x_0, y_0)) \neq 0$.

Dann gibt es eine offene Menge $V \subset \mathbb{R}^q$ mit $x_0 \in V$ und eine stetig differenzierbare Funktion $\varphi: V \rightarrow \mathbb{R}^r$, mit folgenden zwei Eigenschaften:

- a) $\varphi(x_0) = y_0$
- b) Für alle $x \in V$ ist $(x, \varphi(x)) \in \mathcal{U}$ und $F(x, \varphi(x)) = 0$.

Darüber hinaus ist

$$\varphi'(x) = -F_y(x, \varphi(x))^{-1} \cdot F_x(x, \varphi(x)).$$

Beweis: (Skizze): Für die Abbildung $z \rightarrow g(z) = \begin{pmatrix} x \\ F(x, y) \end{pmatrix} \in \mathbb{R}^p$ gilt $g'(z) = \begin{pmatrix} E_p & 0 \\ F_x & F_y \end{pmatrix}$, also ist $\det g'(x_0, y_0) = \det(F_y(x_0, y_0)) \neq 0$. Nach dem Umkehrsatz kann man g in einer offenen, $g(x_0, y_0)$

enthaltenden Menge W umkehren. Wegen $g(x_0, y_0) = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$ ergibt sich, daß $\psi(x) = g^{-1}(x, 0) =: h(x)$ wohldefiniert und stetig differenzierbar ist. $\varphi(x) = (h_{q+1}(x), \dots, h_r(x))$ ist dann die gewünschte Funktion; denn es ist nach Definition von g stets $h_1(x) = x_1, \dots, h_q(x) = x_q$ und damit

$$\begin{pmatrix} x \\ F(x, \varphi(x)) \end{pmatrix} = \begin{pmatrix} x \\ \varphi(x) \end{pmatrix} = g(h(x)) = \begin{pmatrix} x \\ 0 \end{pmatrix}.$$

Mit ψ ist auch φ stetig differenzierbar, also auch $x \rightarrow \begin{pmatrix} x \\ \varphi(x) \end{pmatrix}$. Nach der Kettenregel folgt

$$0 = F(x, \varphi(x))' = F'(x, \varphi(x)) \begin{pmatrix} x \\ \varphi(x) \end{pmatrix}' = F_x(x, \varphi(x)) + F_y(x, \varphi(x))\varphi'(x).$$

Hieraus folgt die Behauptung.

Wir können den Satz verallgemeinern:

Korollar 8.49 Sei $F : \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}^r$ stetig differenzierbar und es gebe ein $z_0 \in \mathbb{R}^p$ mit

1. $F(z_0) = 0$
2. $\text{rng}(F'(z_0)) = r$

Dann gibt es eine Permutation π der Koordinaten, so daß für $F_\pi(z) = F(z_{\pi(1)}, \dots, z_{\pi(p)})$ das Theorem gilt.

Genauer:

Es gibt eine stetig differenzierbare Funktion $\varphi : (z_{\pi(1)}, \dots, z_{\pi(q)}) \rightarrow \mathbb{R}^r$, so daß

$$F(z_{\pi(1)}, \dots, z_{\pi(q)}, \varphi(z_{\pi(1)}, \dots, z_{\pi(q)})) = 0.$$

Beweis: Wegen $\text{rng}(F'(z_0)) = r$ kann man die Spalten von $F'(z)$ so permutieren, daß $\det F_y(z_0) \neq 0$ gilt.

8.4.3 Extrema unter Nebenbedingungen

Die klassische Aufgabe lautet: bestimme unter allen Rechtecken mit gegebenem Umfang $M (> 0)$ dasjenige mit dem größten Flächeninhalt. Wir skizzieren die Lösung: Seien x und y die Kantenlängen. Die "Nebenbedingung" des vorgegebenen Umfangs lautet $F(x, y) \equiv 2(x + y) - M = 0$. Der Flächeninhalt ist $f(x, y) = xy$. Aus der Nebenbedingung erhalten wir $y = (M - 2x)/2$. Dies in f eingesetzt ergibt $g(x) = f(x), (M - 2x)/2 = (Mx - 2x^2)/2$. Wir bestimmen die Extrema von g .

$g'(x) = \frac{M}{2} - 2x$, also $x_E = \frac{M}{4}$. Daraus ergibt sich $y_E = \frac{M - 2x_E}{2} = \frac{M}{4}$, also $x_E = y_E$. (x_E, y_E) erfüllen nach Konstruktion $F(x_E, y_E) = 0$. Außerdem ist $g(x_E)$ ein Maximum von g . Damit ist (x_E, y_E) ein Maximum von $f(xy) = xy$ unter der Nebenbedingung $F(x, y) = 0$.

Wir haben das Problem gelöst, indem wir die Nebenbedingung nach y aufgelöst haben, die so erhaltene Funktion in f eingesetzt und von der neuen Funktion die möglichen Extremalstellen gesucht haben. Führen wir dies etwas abstrakter durch, so sehen wir, daß wir die Funktion $y = \varphi(x)$ nicht explizit bestimmen müssen. Dazu präzisieren wir:

Definition 8.50 Sei \mathcal{U} offen in \mathbb{R}^p , seien $F : \mathcal{U} \rightarrow \mathbb{R}^r$ ($1 \leq r < p$) und $f : \mathcal{U} \rightarrow \mathbb{R}$ gegeben. Sei $M = \{z : F(z) = 0\}$. Ein Punkt $z_0 \in \mathcal{U}$ heißt $\left\{ \begin{array}{l} \text{Maximalstelle} \\ \text{Minimalstelle} \end{array} \right\}$ von f unter der Nebenbedingung $F(z) = 0$, wenn es ein $\varepsilon > 0$ gibt, so daß für alle $z \in B(z_0, \varepsilon) \cap M$ stets $\left\{ \begin{array}{l} f(z) \leq f(z_0) \\ f(z) \geq f(z_0) \end{array} \right\}$ gilt. z_0 heißt dann **Extremalstelle von f unter der Nebenbedingung $f(z) = 0$** .

Der Satz lautet nun:

Theorem 8.51 Sei $\mathcal{U} \subset \mathbb{R}^p$ offen, $F : \mathcal{U} \rightarrow \mathbb{R}^r$ ($1 \leq r < p$) sowie $f : \mathcal{U} \rightarrow \mathbb{R}$ seien stetig differenzierbar. Schließlich sei der Rang $\text{rang}(F'(z)) = r$ für alle $z \in \mathcal{U}$. Sei z_E eine Extremalstelle von f unter der Nebenbedingung $F(z) = 0$. Dann gibt es eindeutig bestimmte Zahlen $\lambda_1, \dots, \lambda_r$ mit

$$f'(z_E) = \lambda_1 F'_1(z_E) + \lambda_2 F'_2(z_E) + \dots + \lambda_r F'_r(z_E).$$

Die λ_j heißen **Lagrange-Multiplikatoren**.

Beweis: Sei $q = p - r$ und es gelte o.B.d.A. $\det\left(\left(\frac{\partial F}{\partial z_j}\right)(z_E)_{j=q+1, \dots, p}\right) \neq 0$ (sonst permutiere die Koordinaten). Wegen $F(z_E) = 0$ kann man die Gleichung nach $y = (z_{q+1}, \dots, z_{q+r})$ auflösen, man erhält $F(x, \varphi(x)) = 0$. Die Funktion $g(x) = f(x, \varphi(x))$ hat dann ein ganz gewöhnliches Maximum oder Minimum in x_E .

Also gilt

$$\begin{aligned} 0 = g'(x_E) &= f_x(x_E, y_E) \varphi'(x_E) \\ &= f_x(z_E) + f_y(z_E) \varphi'(x_E) \\ &= f_x(z_E) + f_y(z_E) (-F_y(z_E))^{-1} F_x(z_E). \end{aligned}$$

Daraus folgt mit $\vec{\lambda} = f_y(z_E) F_y(z_E)^{-1}$

$$f_x(z_E) = \vec{\lambda} F_x(z_E), \quad f_y(z_E) = \vec{\lambda} f_y(z_E),$$

insgesamt also $f'(z_E) = \vec{\lambda} F'(z_E)$.

Um lokale Extrema unter der Nebenbedingung $F(z) = 0$ zu finden, muß man das folgende (nichtlineare) Gleichungssystem für $p + r$ Unbekannte lösen:

$$\begin{aligned} F_1(z) &= 0, \dots, F_r(z) = 0 \\ \frac{\partial f}{\partial z_1}(z) &= \sum_{k=1}^r \lambda_k \frac{\partial F_k}{\partial z_1}(z) \\ &\vdots \\ \frac{\partial f}{\partial z_p}(z) &= \sum_{k=1}^r \lambda_k \frac{\partial F_k}{\partial z_p}(z) \end{aligned}$$

Es sind $p + r$ Gleichungen für die $p + r$ Unbekannten $z_1, \dots, z_p, \lambda_1, \dots, \lambda_r$.

Aufgaben:

1. Bestimmen Sie bitte unter allen Quadern mit Oberfläche $M (> 0)$ denjenigen maximalen Volumens.

2. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Bestimmen Sie bitte die Extrema der Funktion $(x \mid Ax)$ unter der Nebenbedingung $\|x\| = 1$.
Tipp: Die Nebenbedingung ist äquivalent zu $\|x\|^2 = 1$.
3. Wo nimmt die Funktion $f(x, y, z) = 8xyz$ unter der Nebenbedingung $x^2 + y^2 + z^2 = 1$ ihre Extremwerte an? Was bedeutet das geometrisch?

Kapitel 9

Mehrdimensionale Integration

Überblick: Wir erklären wie im eindimensionalen Fall zunächst das Integral für Treppenfunktionen und approximieren dann das untere Integral einer beliebigen Funktion f durch die Integrale von Treppenfunktionen, die unterhalb von f liegen. Analog wird das obere Integral definiert. Sind beide gleich, so heißt f integrierbar und der gemeinsame Wert das Integral von f . Stetige Funktionen sind integrierbar. Zur Berechnung des Integrals dient die iterierte Integration, die dem Cavalierischen Prinzip entspricht. Schließlich benötigt man noch das Verhalten des Integrals bei Wechsel des Koordinatensystems (etwa von rechtwinkligen zu Kugelkoordinaten).

9.1 Integral von Treppenfunktionen

Überblick: Das Integral von Treppenfunktionen basiert auf dem Prinzip Volumen = Grundfläche \times Höhe. Das Integral von Treppenfunktionen verhält sich so, wie der gesunde Menschenverstand es erwartet.

Der Bequemlichkeit wegen betrachten wir im Folgenden nur Intervalle der Form $]a, b] = \{x : a < x \leq b\}$ mit $-\infty < a < b < \infty$. Ein p -dimensionaler Quader Q ist das kartesische Produkt solcher Intervalle:

$$Q =]a_1, b_1] \times \cdots \times]a_p, b_p] = \{x : a_j < x_j \leq b_j \text{ für } 1 \leq j \leq p\}.$$

Sein p -dimensionales Volumen ist $V_p(Q) = (b_1 - a_1) \cdots (b_p - a_p)$, also das Produkt der Kantenlängen. Eine **Elementarfigur** E ist eine Vereinigung von endlich vielen paarweise disjunkten Quadern Q_j , $E = \cup_{j=1}^n Q_j$, $Q_i \cap Q_j = \emptyset$ für $i \neq j$. Wir setzen $V_p(E) = \sum_{j=1}^n V_p(Q_j)$ und überlegen uns, daß dieser Wert unabhängig von der speziellen Zerlegung von E ist. In der Tat: Sei $Q = \prod_{j=1}^p]a_j, b_j]$ und $a_1 < c_1 < b_1$. Dann ist

$$\begin{aligned} Q &=]a_1, c_1] \times \prod_{j=2}^p]a_j, b_j] \cup]c_1, b_1] \times \prod_{j=2}^p]a_j, b_j] \\ &= Q_1 \cup Q_2 \quad \text{mit} \quad Q_1 \cap Q_2 = \emptyset. \end{aligned}$$

Es ist

$$\begin{aligned} V_p(Q_1) + V_p(Q_2) &= (c_1 - a_1)(b_2 - a_2) \cdots (b_p - a_p) + (b_1 - c_1) \cdots (b_p - a_p) \\ &= ((c_1 - a_1) + (b_1 - c_1))(b_2 - a_2) \cdots (b_p - a_p) \\ &= V_p(Q). \end{aligned}$$

Durch Induktion nach der Zahl der zusätzlichen Einteilungspunkte erhält man die behauptete Unabhängigkeit von der Zerlegung von E .

Definition 9.1 Eine Treppenfunktion f ist eine Funktion von $\mathbb{R}^p \rightarrow \mathbb{R}$, die nur endlich viele Werte annimmt und für die $f^{-1}(\alpha)$ eine Elementarfigur ist, falls $\alpha \neq 0$ ist.

Da die Vereinigung zweier Quader $Q \cup R$ eine Elementarfigur ist, ist auch die Vereinigung von endlich vielen Elementarfiguren eine Elementarfigur. Damit erhalten wir

Satz 9.2 Summe, skalares Vielfaches, Produkt und Absolutbetrag von Treppenfunktionen sind Treppenfunktionen.

Wir können nun das Integral definieren:

Definition 9.3 Sei $f : \mathbb{R}^p \rightarrow \mathbb{R}$ eine Treppenfunktion.

Dann heißt $\sum_{0 \neq \alpha \in f(\mathbb{R}^p)} \alpha V_p(f^{-1}(\alpha)) = \int f d^p x$ Integral von f (über \mathbb{R}^p).

Satz 9.4 Das Integral hat folgende Eigenschaften:

- a) $\int (\alpha f + \beta g) d^p x = \alpha \int f d^p x + \beta \int g d^p x$.
- b) $f \leq g \Rightarrow \int f d^p x \leq \int g d^p x$.
- c) $|\int f d^p x| \leq \int |f| d^p x$.
- d) Sei $E = \{x : f(x) \neq 0\}$. Dann gilt $\min(f(\mathbb{R}^p)) V_p(E) \leq \int f d^p x \leq \max(f(\mathbb{R}^p)) V_p(E)$.

Beweis: a) Wir zerlegen die Elementarfigur $\{x : f(x) \neq 0\} \cup \{x : g(x) \neq 0\}$ in lauter disjunkte Quader Q_j und erhalten $f = \sum \alpha_j 1_{Q_j}$, $g = \sum \beta_j 1_{Q_j}$, also $\alpha f + \beta g = \sum (\alpha \alpha_j + \beta \beta_j) 1_{Q_j}$ und damit

$$\begin{aligned} \int (\alpha f + \beta g) d^p x &= \sum (\alpha \alpha_j + \beta \beta_j) V_p(Q_j) \\ &= \alpha \sum \alpha_j V_p(Q_j) + \beta \sum \beta_j V_p(Q_j) \\ &= \alpha \int f d^p x + \beta \int g d^p x. \end{aligned}$$

b) Ist $0 \leq h \Rightarrow 0 \leq \int h d^p x$. Setze nun im Fall $f \leq g$ einfach $h = g - f$. Dann ist nach a)

$$0 \leq \int h d^p x = \int g d^p x - \int f d^p x.$$

c) Dies folgt aus $\pm f \leq |f|$.

d) $\min(f(\mathbb{R}^p)) 1_E \leq f \leq \max(f(\mathbb{R}^p)) 1_E$ und b) liefert die Behauptung.

9.2 Fortsetzung des Integrals

Überblick: Wir definieren analog zum eindimensionalen Fall das untere und das obere Integral. Eine Funktion heißt integrierbar, wenn beide übereinstimmen.

Hinter der Definition des allgemeine Integrals steckt wie im eindimensionalen Fall der Ausschöpfungsge-
danke von Archimedes.

Appletanleitung Sie können sich einige Funktionen (also ihre Graphen) anschauen und sehen, wie sich die Volumina der Mengen

$$\{(x, y, z) : a \leq x \leq b, c \leq y \leq d, 0 \leq z \leq f(x, y)\}$$

durch Quader von unten und oben ausschöpfen lassen. Vergewissern Sie sich, daß dies bedeutet, daß unterhalb und oberhalb von f geeignete Treppenfunktionen liegen, deren Graphen Sie sehen können. Das bedeutet gerade das Ausschöpfen von unten und oben.

* * * T O D O * * *

Definition 9.5 Sei f beschränkt und verschwinde außerhalb eines genügend großen Quaders Q . Dann heißt

$$\overline{\int} f dx^p = \inf \left\{ \int h dx^p : f \leq h, h \text{ Treppenfunktion} \right\}$$

oberes Integral und

$$\underline{\int} f dx^p = \sup \left\{ \int g dx^p : g \leq f, g \text{ Treppenfunktion} \right\}$$

heißt **unteres Integral** von f . Ist $\overline{\int} f dx^p = \underline{\int} f dx^p$, so heißt f integrierbar und der gemeinsame Wert heißt Integral

$$\int f dx^p = \overline{\int} f dx^p = \underline{\int} f dx^p.$$

Aus den Rechenregeln für das Supremum und das Infimum erhält man wie im Falle $p = 1$ aus Satz 9.4 den folgenden Satz:

Satz 9.6 Seien f und g integrierbar. Dann gilt

a) $\int (\alpha f + \beta g) dx^p = \alpha \int f dx^p + \beta \int g dx^p$

b) $f \leq g \Rightarrow \int f dx^p \leq \int g dx^p$

c) $|\int f dx^p| \leq \int |f| dx^p$

d) fg ist integrierbar.

Beweis: Nach der Bemerkung vor dem Satz muß nur noch d) bewiesen werden. Es ist

$$fg = \frac{1}{4}((f+g)^2 - (f-g)^2),$$

also muß nur bewiesen werden: Ist f integrierbar, so auch f^2 und das beweist man analog zum entsprechenden Satz im Wintersemester.

Definition 9.7 Sei $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$, wo U offen ist. Dann heißt der Durchschnitt von U mit dem Abschluß \overline{A}_f der Menge $A_f = \{x : f(x) \neq 0\}$ der **Träger** von f .

Satz 9.8 Jede stetige Funktion mit kompaktem Träger ist integrierbar.

Beweis: Der Träger A_f liege im Würfel $W = \prod_1^p]-a, a[=]-a, a]^p$. Da A_f kompakt und f außerhalb von A_f gleich 0 ist, ist f gleichmäßig stetig. Zu $\varepsilon > 0$ gibt es also ein $\delta > 0$ mit $|f(x) - f(y)| < \varepsilon$ für alle x, y mit $\|x - y\|_\infty < \delta$ (beachte die Formel 8.6).

Wir wählen n mit $\frac{2a}{n} < \delta$ und setzen $a_k = -a + \frac{2ka}{n}$. Dann ist $|f(x) - f(y)| < \varepsilon$, sobald x und y in einem Würfel der Form $\prod_{j=1}^p]a_{k(j)}, a_{k(j)+1}[$ liegt. Sei $\vec{k} = (k(1), \dots, k(p))$ mit $0 \leq k(j) < n$ und $W_{\vec{k}} = \prod_{j=1}^p]a_{k(j)}, a_{k(j)+1}[$. Dann ist $W = \bigcup_{\vec{k}} W_{\vec{k}}$ und $W_{\vec{k}} \cap W_{\vec{\ell}} = \emptyset$ für $\vec{k} \neq \vec{\ell}$. Sei $\underline{f}_{\vec{k}} = \inf\{f(x) : x \in W_{\vec{k}}\}$ und $\overline{f}_{\vec{k}} = \sup\{f(x) : x \in W_{\vec{k}}\}$. Dann ist

$$0 \leq \overline{f}_{\vec{k}} - \underline{f}_{\vec{k}} \leq \varepsilon. \quad \text{Sei } g = \sum \underline{f}_{\vec{k}} 1_{W_{\vec{k}}} \quad \text{und } h = \sum \overline{f}_{\vec{k}} 1_{W_{\vec{k}}}.$$

Dann ist $g \leq f \leq h$ und $\int (h - g) d^p x \leq \varepsilon V_p(W)$.

Also ist $0 \leq \int f d^p x - \int g d^p x \leq \int h d^p x - \int g d^p x \leq \varepsilon V_p(W)$. Da ε beliebig war folgt die Behauptung.

Wir definieren nun das Integral über eine Menge.

Definition 9.9 Sei C eine beliebige Menge, für die 1_C integrierbar ist. Sei f integrierbar. Dann ist $1_C f$ integrierbar und wir setzen

$$\int_C f d^p x = \int 1_C f d^p x$$

9.3 Iterierte Integration

Das wichtigste Hilfsmittel zur Integration ist die **iterierte Integration**. Für $z \in \mathbb{R}^p = \mathbb{R}^q \times \mathbb{R}^r$ schreiben wir wie beim Satz über implizite Funktionen

$$z = (x, y) \quad \text{mit } x = (x_1, \dots, x_q) = (z_1, \dots, z_q), \quad y = (y_1, \dots, y_r) = (z_{q+1}, \dots, z_p).$$

Appletanleitung Die iterierte Integration bedeutet im wesentlichen, daß wir "scheibchenweise" integrieren: Sei wieder

$$K = \{(x, y, z) : a \leq x \leq b, c \leq y \leq d, 0 \leq z \leq f(x, y)\}.$$

Wir zerlegen das x -Intervall $[a, b]$ in kleine Stücke $a_k = a + k\delta x$ mit $\delta x = (b - a)/n$. Durch jeden Punkt $(a_k, 0, 0)$ legen wir eine Ebene parallel zur (y, z) -Ebene, die uns aus der Menge eine Grundfläche der Größe $F(a_k) = \int_c^d f(a_k, y) dy$ herauschneidet. Das Volumen von K , also das Integral $\int_{[a,b] \times [c,d]} f(x, y) d^2(x, y)$ ist dann ungefähr gleich der Summe der Zylinder mit dieser Grundfläche und der Höhe δx also gleich $\sum_{k=0}^{n-1} F(a_k) \delta x$, also hat man

$$\int_{[a,b] \times [c,d]} f(x, y) d^2(x, y) \approx \sum_{k=0}^{n-1} \left(\int_c^d f(a_k, y) dy \right) \delta x \approx \int_a^b \left(\int_c^d f(x, y) dy \right) dx.$$

* * * TO DO * * *

Wir legitimieren nun das Vorgehen, das für Treppenfunktionen fast selbstverständlich ist.

Lemma 9.10 Sei f eine Treppenfunktion in $\mathbb{R}^p = \mathbb{R}^q \times \mathbb{R}^r$. Für jedes $x \in \mathbb{R}^q$ ist ${}_x f : y \rightarrow f(x, y)$ als Funktion von y eine Treppenfunktion, ferner ist $x \mapsto h(x) = \int {}_x f d^r y$ als Funktion von x eine Treppenfunktion und es gilt

$$\int f(x, y) d^p(x, y) = \int {}_x \left(\int f d^r y \right) d^q x = \int (f(x, y) d^r y) d^q x.$$

Beweis: Aufgrund der Rechenregeln für das Integral von Treppenfunktionen genügt es, den Satz für Funktionen 1_Q zu zeigen mit $Q = \prod_{j=1}^p]a_j, b_j]$. Setzt man $Q = Q_q \times Q_r$ mit $Q_q = \prod_{j=1}^q]a_j, b_j]$, $Q_r = \prod_{k=q+1}^p]a_k, b_k]$, so erhält man ${}_x 1_Q = \begin{cases} 1_{Q_r} & x \in Q_q \\ 0 & \text{sonst} \end{cases}$ und die Behauptung ist offensichtlich.

Mit der Konstruktion des Integrals ergibt sich:

Theorem 9.11 Sei $f : U \rightarrow \mathbb{R}$ integrierbar. Dann ist

$$\int f(z) d^p z = \int \left(\int f(x, y) d^r(y) \right) d^q x = \int \left(\int f(x, y) d^q x \right) d^p y.$$

Beispiele:

1.

$$\begin{aligned} \int_{[0, a] \times [0, b]} \sin(x + y) d^2(x, y) &= \int 1_{[0, a] \times [0, b]}(x, y) \sin(x + y) d^2(x, y) \\ &= \int_0^a \left(\int_0^b \sin(x + y) dy \right) dx = \int_0^a (\cos(x) - \cos(x + b)) dx \\ &= \sin(a) - \sin(a + b) + \sin(b) \\ &= \sin(a) + \sin(b) - \sin(a + b). \end{aligned}$$

2. Sei K der obere Halbkreis, $K = \{(x, y) : -1 \leq x \leq 1, 0 \leq y \leq \sqrt{1 - x^2}\}$. Dann ist

$$\begin{aligned} \int 1_K(x, y) d^2(x, y) &= \int_{-1}^1 \left(\int_0^{\sqrt{1-x^2}} dy \right) dx \\ &= \int_{-1}^1 \sqrt{1-x^2} dx \\ &= \frac{1}{2} (x\sqrt{1-x^2} + \arcsin(x)) \Big|_{-1}^1 = \pi/2. \end{aligned}$$

9.4 Transformation von Integralen

Das letzte Integral war sehr schwer, weil wir K "falsch" parametrisiert haben. In Polarkoordinaten ist

$$K = \{(r \cos \varphi, r \sin \varphi) : 0 \leq r \leq 1, 0 \leq \varphi \leq \pi\}$$

Aber wie berechnet man das Integral nur unter Benutzung der Polarkoordinaten?

Es ist $dx = x(r + dr, \varphi + d\varphi) - x(r, \varphi) \approx \frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \varphi} d\varphi$

und analog $dy = \frac{\partial y}{\partial r} dr + \frac{\partial y}{\partial \varphi} d\varphi$, also

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} \begin{pmatrix} dr \\ d\varphi \end{pmatrix} = J(r, \varphi) \begin{pmatrix} dr \\ d\varphi \end{pmatrix}.$$

Der Flächeninhalt wird aber unter einer linearen Abbildung A durch $\det(A)$ verzerrt, genauer: Das Bild des Einheitsquadrates ist ein Parallelogramm mit dem Flächeninhalt $\det(A)$. Also ist

$$dxdy \approx \det(J(r, \varphi)) dr d\varphi$$

und damit

$$\int f(x, y) d^2(x, y) = \int \int f(x, y) dxdy = \int \int f(x(r, \varphi), y(r, \varphi)) \cdot r dr d\varphi,$$

ferner

$$1_K(x(r, \varphi), y(r, \varphi)) = 1_{[0,1]}(r) 1_{[0,\pi]}(\varphi)$$

und damit

$$\int 1_K(x, y) dxdy = \int_0^\pi \left(\int_0^1 r dr \right) d\varphi = \frac{1}{2} \pi.$$

Unser eben betrachtete Spezialfall läßt sich verallgemeinern:

Theorem 9.12 (Transformationsatz)

Seien $\emptyset \neq U, V$ offene Teilmengen von \mathbb{R}^p und $f : U \subset \mathbb{R}^p \rightarrow V$ stetig. Ferner sei $\Phi : V \rightarrow U$ eine stetig differenzierbare bijektive Abbildung. Dann gilt für jede kompakte integrierbare Menge $C \subset U$

$$\int_C f(x) d^p x = \int_{\Phi^{-1}(C)} f(\Phi(y)) |\det \Phi'(y)| dy.$$

Der Beweis beruht auf der Formel

$$dx_1 \cdots dx_p \approx |\det \Phi'| dy_1 \cdots dy_p,$$

die wir oben plausibel gemacht haben.

Beispiele:

- Wir schauen noch einmal Rotationskörper an. Sei $\varphi : [0, a] \rightarrow \mathbb{R}_+$ stetig und

$$K = \{(x, y, z) : 0 \leq x \leq a, 0 \leq y^2 + z^2 \leq \varphi(x)^2\}$$

Hier bieten sich die folgenden Zylinderkoordinaten an: $x = x$, $y = r \cos \psi$, $z = r \sin \psi$ mit $-\pi \leq \psi \leq \pi$ und $0 < r$, also

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = f \begin{pmatrix} x \\ r \\ \psi \end{pmatrix}, f' \begin{pmatrix} x \\ r \\ \psi \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & r(-\sin \psi) \\ 0 & \sin \psi & r \cos \psi \end{pmatrix}, \det f' \begin{pmatrix} x \\ r \\ \psi \end{pmatrix} = r.$$

Damit ergibt sich

$$\begin{aligned} V(K) &= \int_K d^3(x, y, z) = \int_{-\pi}^{\pi} \left(\int_0^a \left(\int_0^{\varphi(x)} r dr \right) dx \right) d\psi \\ &= \pi \int_0^a \varphi(x)^2 dx, \end{aligned}$$

ein Ergebnis, das wir auch schon anschaulich im Wintersemester erhalten haben.

2. Sei $K = \{(x, y) : 0 \leq x \leq R, 0 \leq y \leq \sqrt{R^2 - x^2}\}$ und $f(x, y) = \exp(-(x^2 + y^2))$.

Hier bieten sich Polarkoordinaten an.

$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} = \Phi(r, \varphi)$. Dann ist $1_K(x, y) = 1_{[0, R]}(r)1_{[0, \pi/2]}(\varphi)$ und $f(x, y) = \exp(-r^2)$.

Also erhält man wegen $\det(\Phi') = r$ mit der Substitutionsregel (in bezug auf r):

$$\begin{aligned} \int_K f(x, y) d^2(x, y) &= \int_0^{\pi/2} \int_0^R r \exp(-r^2) dr d\varphi \\ &= \pi/2 \cdot \frac{1}{2} (1 - \exp(-R^2)) \end{aligned}$$

Ende der Vorlesung

Literaturverzeichnis

[Wolff] M. Wolff, O. Gloor, C. Richard *Analysis Alive* Birkhäuser, 1991

Index

- $(A + B)(x)$, 115
 $(\mathbb{K}, +, \cdot)$, 13
 $(\lambda A)(x)$, 115
 $(a_n)_{n \geq k}$, 21
 $(f_n)_{n \in \mathbb{N}}$, 48
 $(x^\perp | y^\perp)$, 105
 $(x^\perp y^\perp z^\perp)$, 109
 *-Produkt, 127
 1_A , 37
 $<$, 14
 $< a, b >$, 36
 $>$, 14
 $A(\varepsilon)$, 22
 A^* , 144
 A^T , 144
 A_f , 174
 $B(x, r(x))$, 149
 $D_h(f)(x)$, 159
 E_p , 124
 $F(A)$, 6
 G_f , 5
 $H(f)$, 165
 K_f , 93
 $L(V, W)$, 115
 $L_0(\mathcal{A})$, 132
 $L_{b^\perp}(\mathcal{A})$, 132
 M^c , 4
 M^n , 5
 $M_1 \times M_2$, 5
 $N \setminus M$, 4
 $O(K_f)$, 93
 S_B , 124
 $S_Z(f)$, 76
 $U(c, \varepsilon)$, 22
 $V(K_f)$, 93
 V_p , 172
 $[a, b[$, 35
 $[a, b]$, 35
 $\mathcal{A}_A^{B,C}$, 125
 \mathcal{A}_A , 118
 \mathcal{A}_A^{-1} , 124
 $\mathcal{A}_{T \circ S}$, 122
 $\mathcal{B}(M, N)$, 9
 $\Im(z)$, 13
 \Leftrightarrow , 3
 $\Re(z)$, 13
 \Rightarrow , 2
 $\mathcal{T}(J)$, 77
 $\mathcal{T}(Z)$, 75
 $\mathcal{T}([a, b])$, 77
 arcosh, 73
 arccos, 72
 arcsin, 72
 arctan, 72
 arsinh, 73
 artanh, 73
 $\bigcap_{i \in I} A_i$, 7
 $\bigcup_{i \in I} A_i$, 7
 $\binom{x}{y}$, 98
 $\binom{\alpha}{k}$, 69
 $\binom{n}{k}$, 10
 $\frac{\partial f(x)}{\partial x_j}$, 159
 $\frac{\partial^n f}{\partial x_{k_1} \dots \partial x_{k_n}}$, 159
 $\mathcal{P}(M)$, 7
 \cap , 4
 \mathbb{C} , 4
 cos, 71
 cosh, 72
 \cup , 4
 δ_{kl} , 124
 $\det(\mathcal{A})$, 129
 $\det(a_1^\perp, \dots, a_p^\perp)$, 129
 \det , 127
 $\dot{f}(c)$, 53
 $\dot{x}(t)$, 157
 \emptyset , 4
 $\epsilon - \delta$ -Kriterium, 41
 \equiv , 3

$\frac{1}{x}$, 12
 $\frac{d^n f}{dx^n}$, 60
 $\frac{df}{dx}(c)$, 53
 $\frac{d}{dx}$, 12
 \geq , 14
 \in , 4
 $\inf(M)$, 17
 $\int f dx^p$, 173, 174
 $\int_a^b f(x) dx$, 77
 $\int_C f dx^p$, 175
 \int_a^b , 78
 \int_a^b , 78
 $\ker(A)$, 116
 \mathbb{K} , 13
 \leq , 14
 \liminf , 23
 \limsup , 23
 $\lim_{n \rightarrow \infty} a_n$, 24
 $\lim_{x \rightarrow \infty} f(x) = \infty$, 41
 $\lim_{x \rightarrow \infty} f(x)$, 41
 $\lim_{x \rightarrow c} f(x)$, 39
 $\lim_{x \rightarrow c} cf(x) = \infty$, 41
 $\ln(x)$, 46
 $\log_a(x)$, 46
 $\mathcal{H}(D)$, 38
 $\max(a, b)$, 15
 $\operatorname{grad} f(x)$, 163
 $\operatorname{span}(a_1, \dots, a_p)$, 103
 $\min(a, b)$, 15
 \neg , 1
 \mathbb{N} , 4
 \mathbb{N}_0 , 4
 \notin , 4
 \overline{D} , 38
 $\int f dx^p$, 174
 \overline{z} , 13
 \prod_k^n , 5
 $\prod_{k=0}^n$, 9
 $\dim(V)$, 103
 \mathbb{Q} , 4
 \mathbb{R} , 4
 \sharp , 9
 $\sigma(A)$, 140
 \sin , 71
 \sinh , 72
 \subset , 4
 $\sum_{k=0}^n$, 9
 $\sup(M)$, 17
 $a^\perp * b^\perp$, 127
 $\int f dx^p$, 174
 \overline{a}_j , 120
 \vee , 2
 \wedge , 1
 \mathbb{Z} , 4
 $]a, b[$, 35
 $]a, b]$, 35
 $+\sqrt{a}$, 17
 xf , 175
 $a \mid b$, 4
 a^+ , 16
 a^- , 16
 a^x , 46
 $d(a, b)$, 16
 $f'(c)$, 53
 $f(x)$, 5
 $f: M \mapsto N$, 6
 $f(n)$, 60
 $f^{-1}(B)$, 6
 $f_{x_j}(x)$, 159
 $g \circ f$, 6
 id_M , 7
 $\operatorname{rng}(A)$, 116
 rng_z , 138
 r_d , 138
 $\operatorname{sg}(\pi)$, 128
 x^{-1} , 12
 $x^\perp \times y^\perp$, 107
 x^\perp , 99
 $z + M$, 18
 zM , 18
 \mathbb{K}_2 , 14
 $\text{\AA}quivalenz$, 3
 \"uberall stetig , 42
 $n!$, 9
 1-Norm , 148
 $A(P, Q)$, 104
 Abbildung , 5
 identische, 36
 lineare, 114
 abgeschlossen , 149
 $\text{abgeschlossene H\"ulle}$, 154
 $\text{abgeschlossenes Intervall}$, 35
 ableitbar , 53
 Ableitung , 53
 $n\text{-te}$, 60
 der Umkehrfunktion, 56
 erste, 60

- partielle, 159
- totale, 161
- absolut konvergent, 29
- Absolutbetrag, 16
- absolute Homogenität
der Norm, 106
- Abstand
zu einer Ebene, 110
- Addition
von Vektoren, 99
- Adhärenzpunkt, 38, 154
- adjungierte Matrix, 144
- angeregter Zustand, 119
- Antisymmetrie, 14
- Archimedes' Axiom, 14
- Area Cosinus Hyperbolicus, 73
- Area Sinus Hyperbolicus, 73
- Area Tangens Hyperbolicus, 73
- arithmetischer Vektorraum, 100
- Arkuscosinus, 72
- Arkussinus, 72
- Arkustangens, 72
- assoziativ, 6
- Assoziativgesetz
der Addition, 11
der Multiplikation, 12
- Basis, 103
- Basiswechsel, 125
- Basiswechselmatrix, 125
- Bernoullische Ungleichung, 15
- beschränkt, 17, 149
(bei Folgen in \mathbb{R}^p), 150
nach oben, 17
nach unten, 17
- bestimmte Divergenz, 41
- Beweis
indirekter, 3
- Bijektion, 6
- bijektiv, 6
- Bild, 5
- Bildmenge, 6
- Bilinearform
symmetrische, 146
- Binomialkoeffizienten, 10
- Bisektionsverfahren, 43
- Bogenlänge, 158
einer Kurve, 93
- Bolzano-Weierstraß, Satz von, 22
- Boolesche Zahlen, 14

- Cauchy-Folge
in \mathbb{R}^p , 150
- Cauchy-Schwarzsche Ungleichung, 105
- Cauchyprodukt, 30
- Cauchys Konvergenzkriterium, 26
- charakteristisches Polynom, 141
- Cosinus, 71
- Cosinus Hyperbolicus, 72
- Cosinus-Reihe, 32
- De L'Hospital
1. Regel, 64
2. Regel, 64
- Definitheit
der Norm, 106
des Skalarproduktes, 105
- Definitheit (des Abstandes), 16
- Definitionsbereich, 5
- Dehnungsinvarianz (der Ordnung), 14
- Determinante, 127
- Diagonalgestalt, 126
- diagonalisierbar, 143
- Diagonalmatrix, 126
- Differentialgleichung
des mathematischen Pendels, 70
- Differentialquotient, 53
- Differentiation
von Potenzreihen, 57
- differenzierbar, 53
beliebig oft, 60
in einem Punkt, 53
in Richtung h , 159
unendlich oft, 60
- Differenzierbarkeit
komplexer Funktionen, 94
von Kurven, 157
- Dimension, 103
- Dimensionsformel, 117
- Distributivgesetz, 12
- Divergenz
bestimmte, 41
- Dreiecksungleichung, 16, 106
allgemeine, 16
verallgemeinerte, 16
- Durchschnitt, 4
beliebiger, 7
- Ebene, 98, 100
- Eigenraum, 141
- Eigenschaften

- der Determinanten, 127
 - des Gradienten, 164
 - des Integrals, 173
- Eigenwert, 140
- Eindeutigkeit
 - der Lösung eines LGS, 133
- Einheitsmatrix, 124
- Elementarfigur, 172
- Elementarfläche, 78
- Ellipsoid, 140
- entartet, 127
- Entwickelbarkeit (in eine Taylorreihe), 68
- Entwicklungspunkt, 66
- Entwicklungssatz (von Taylor), 66
- erste Ableitung, 60
- erzeugter Untervektorraum, 103
- Existenz
 - der Lösung eines LGS, 133
- Exponentialreihe, 32
- Extremstelle, 66
 - unter der Nebenbedingung, 170
- Extremum
 - lokales, 66
- Extremwert, 167
- Extremwert, 66
- Fakultät, 9
- feinere Zerlegung, 76
- Fläche, 153
- Flächeninhalt
 - einer Kreisscheibe, 92
- Folge
 - in \mathbb{R}^p , 150
 - von Funktionen, 48
- Folge der Teilsummen, 28
- Folge, beschränkte, 22
- Folge, komplexe, 32
- Folge, konvergente, 24
- Folge, nach oben beschränkte, 22
- Folge, nach unten beschränkte, 22
- Fourier-Reihe, 94
- Fourierkoeffizienten, 95
- Fourierreihe, 51
- Fundamentalsatz der Algebra, 142
- Funktion
 - (streng) monoton fallende, 45
 - (streng) monoton wachsende, 45
 - konstante, 36
 - periodische, 94
 - rationale, 36
 - reelle, 35
- Funktionenfolge, 48
- Gauß-Seidel-Verfahren, 134
- Gebiet, 150
- gemeinsame Verfeinerung, 76
- geometrische Reihe, 29, 32
- Gerade, 52, 100
- Geradengleichung, 52
- gewöhnliche Differentialgleichung, 70
- Gleichheit (von Abbildungen), 6
- gleichmäßig
 - konvergent, 157
- gleichmäßig stetig, 47
- gleichmäßige Konvergenz, 49
 - und Stetigkeit, 49
- Gleichungssystem
 - allgemeine Form, 131
 - lineares, 131
- Grad
 - eines Polynoms, 36
- Gradient, 163
- Graph, 5
- Grenze
 - obere, 17
 - untere, 17
- Grenzwert, 24
 - bei vektorwertigen Funktionen, 154
 - von $f(x)$ für x gegen c , 39
- Grundzustand, 119
- Häufungspunkt, 22
 - einer Menge, 38
- Höhenlinie, 164
- höhere Ableitung, 60
- Hauptachsen
 - eines Tensors, 140
- Hauptachsentransformation, 146
- Hauptsatz
 - der Differential- und Integralrechnung, 85
- Hessematrix, 165
- Hessesche Normalenform, 110
- Hilber-Schmidt-Norm, 148
- Hintereinanderausführung, 6, 37
- homogen, 132
- Homogenität, absolute, 16
- identische Abbildung, 7, 36
- Identität von Lagrange, 109
- Imaginärteil, 13

Implikation, 2
 implizit definiert, 169
 Indikatorfunktion, 37
 Induktion, vollständige, 9
 Induktionsanfang, 9
 Induktionsaxiom, 8
 Induktionsschritt, 9
 Inegral, 78
 Inegration
 rationaler Funktionen, 88
 Infimum, 17
 inhomogen, 132
 injektiv, 6
 Integral, 173
 über eine Menge, 175
 oberes, 78
 unteres, 78
 von Treppenfunktionen, 77
 Integration
 durch Substitution, 88
 iterierte, 175
 komplexer Funktionen, 94
 partielle, 87
 von Potenzreihen, 89
 Intervall
 abgeschlossenes, 35
 links abgeschlossen, rechts offen, 35
 links offen, rechts abgeschlossen, 35
 offenes, 35
 Intervalladditivität, 78
 Integral
 oberes, 174
 unteres, 174
 inverse Matrix, 124
 inverses Element,
 additiv, 11
 multiplikatives, 12
 iterierte Integration, 175

 Jakobi-Matrix, 161

 Körper, 13
 kanonische Parametrisierung, 158
 kartesisches Produkt, 5
 Kern, 116
 Kettenregel, 55, 163
 Kommutativgesetz
 der Addition, 11
 der Multiplikation, 12
 kompakt, 149

 Komplement, 4
 komplexe Folge, 32
 komplexe Reihe, 32
 komplexe Zahlen, 13
 komplexer Vektor, 142
 konjugiert komplexe Zahl, 13
 konstante Funktion, 36
 Konvergenz, 24
 gleichmäßige, 49
 punktweise, 49
 vektorwertiger Funktionen, 154
 von Folgen in \mathbb{R}^p , 150
 von Fourierreihen, 96
 Konvergenz komplexer Zahlen, 32
 Konvergenzkriterium von Cauchy, 26
 Konvergenzradius, 31
 konvex, 165
 Koordinaten, 98
 Koordinatentransformation, 153
 Kreisumfang, 93
 Kronecker-Symbol, 124
 Kugel, 149
 Kugelkoordinaten, 154
 Kurven (im \mathbb{R}^p), 153

 Länge
 einer Kurve, 158
 Lösung
 einer Differentialgleichung, 70
 eines LGS, 131
 Lagrange, Identität von, 109
 Lagrange-Multiplikatoren, 170
 leere Menge, 4
 LGS, 131
 Limes, 24
 Limes Inferior, 23
 Limes Superior, 23
 linear, 114
 linear abhängig, 102
 linear unabhängig, 102
 lineare Evolution, 70
 lineare Evolutionsgleichung, 70
 linearer Teilraum, 103
 lineares Gleichungssystem, 131
 Linearität
 der Ableitung, 55
 des Integrals, 80
 Linearität
 des Skalarproduktes, 105
 Linearkombination, 103

- Linkssystem, 127
- Logarithmus naturalis, 46
- logisches Oder, 2
- logisches Und, 1
- lokale
 - Maximalstelle, 167
 - Minimalstelle, 167
- lokales
 - Maximum, 167
 - Minimum, 167
- lokales Extremum, 66
- lokales Maximum, 66
- lokales Minimum, 66
- Mantelfläche
 - eines Rotationskörpers, 93
- mathematisches Pendel, 70
- Matrix, 118
 - adjungierte, 144
 - einer Hintereinanderausführung, 122
 - quadratische, 126
 - transponierte, 144
- Matrixdarstellung, 118
- Matrixform
 - eines LGS, 132
- Matrixnorm, 148
- Matrizenprodukt, 122
- Maximalstelle, 167
 - unter der Nebenbedingung, 170
- Maximum, 15, 167
 - lokales, 66
- Menge, 3
- Minimalstelle, 167
 - unter der Nebenbedingung, 170
- Minimum, 15, 167
 - lokales, 66
- Mittelwertsatz
 - der Integralrechnung, 85
 - erster, 62
 - zweiter, 63
- monoton, 27
- monoton fallend, 27
- monoton wachsend, 27
- Monotonie
 - bei Funktionen, 45
- Multiplikation
 - mit Skalaren, 99
- Multiplikationssatz, 129
- negativ
 - orientiert, 127
- negativ definit, 166
- neutrales Element
 - der Addition, 11
 - der Multiplikation, 12
- Norm, 104
- normal, 144
- Normalenvektor, 110
- Normalverteilung, 89
- Nullfolge, 24
- Nullstellensatz
 - für stetige Funktionen, 43
- oberes Integral, 78, 174
- Oberfläche
 - eines Rotationskörpers, 93
- offen, 149
 - zerlegbar, 150
- offenes Intervall, 35
- Ordnung, 14
- orthogonal, 106
- orthogonale Matrix, 145
- Orthonormalbasis, 144
- Orthonormalsystem, 144
- parallel, 102
- Parameterdarstellung, 110
- Parametrisierung, 158
 - durch die Bogenlänge, 158
 - kanonische, 158
- partielle Ableitung, 159
- partielle Integration, 87
- Pendel, 70
- Periode, 94
- periodische Funktion, 94
- Permutation, 10
- Pivot-Element, 134
- Polarkoordinaten, 154
- Polynom, 36
 - charakteristisches, 141
 - trigonometrisches, 94
- Polynomgrad, 36
- positiv definit, 146, 166
- Positivität
 - des Integrals, 80
- positiv
 - orientiert, 127
- Potenz, 9
- Potenzmenge, 7
- Potenzreihe, 31

Potenzreihen, 31
 Produkt, 9
 eines Zeilen- m.e. Spaltenvektor, 120
 von Funktionen, 36
 zweier Matrizen, 122
 Produktregel
 Leibnizsche, 55
 Produktreihe, 31

 Quader
 p -dimensionaler, 172
 quadratische Matrix, 126
 Quadrik, 140
 Quotient
 von Funktionen, 36
 Quotientenkriterium, 30
 Quotientenregel, 56

 Rang, 116
 rationale Funktion, 36
 Realteil, 13
 Rechenregeln
 für das Integral, 80
 für das Matrizenprodukt, 123
 für die Differentiation, 55
 für Stetigkeit, 43
 Rechtssystem, 107, 127
 reelle Funktion, 35
 Reflexivität, 14
 Reihe, 28
 trigonometrische, 50, 94
 Reihe, geometrische, 29
 Reihe, komplexe, 32
 Rekursion, 9
 Rekursionsschritt, 9
 Restglied, 66
 Richtungsableitung, 159
 Rolle, Satz von, 62
 Rotationskörper, 93

 Satz
 über die Umkehrfunktion, 168
 über implizite Funktionen, 169
 über lokale Extremwerte, 167
 von Bolzano-Weierstraß, 151
 von Schwarz, 160
 von Taylor, 166
 Satz von Rolle, 62
 Schranke
 obere, 17
 untere, 17
 Schrankensatz, 165
 Schwingungsgleichung, 70
 Sekante, 52
 selbstadjungiert, 144
 Signum, 128
 Sinus, 71
 Sinus Hyperbolicus, 72
 Sinus-Reihe, 32
 Skalar, 99
 skalares Vielfaches, 36
 Skalarmultiplikation, 99
 Skalarprodukt, 105
 skalare Funktion, 153
 Spalte
 einer Matrix, 118
 Spaltenform
 eines LGS, 132
 Spaltenrang, 138
 Spektrum, 140
 Sphäre, 111
 Spiegelung
 an einer Ebene, 111
 Stammfunktion, 85
 Steigung, 52
 Stellungsvektor, 110
 stetig
 überall, 42
 differenzierbar, 159
 Stetigkeit, 42
 an der Stelle x_0 , 42
 auf abgeschlossenen beschränkten Intervallen, 47
 der Umkehrfunktion, 45
 gleichmäßige, 47
 komplexer Funktionen, 94
 vektorwertiger Funktionen, 155
 von Potenzreihen, 50
 Streckung, 99
 streng monoton fallend, 27
 streng monoton wachsend, 27
 Submultiplikativität der Norm, 149
 Summe, 9
 von Funktionen, 36
 Supremum, 17
 Supremumsnorm, 68
 surjektiv, 6
 Symmetrie
 des Skalarproduktes, 105

symmetrische Matrix, 145
 Tangens, 58, 72
 Tangente, 52
 Tautologie, 3
 Taylor
 Satz von, 66
 Taylorpolynom, 66
 Taylorreihe, 68
 Teilfolge, 28
 Teilmatrix, 138
 Teilmenge, 4
 Teilraum
 linearer, 103
 Tensor, 140
 total differenzierbar, 161
 totale Ableitung, 161
 Totalität (der Ordnung), 14
 Träger, 174
 Transformationsatz, 177
 Transitivität, 14
 Translationsinvarianz (der Ordnung), 14
 transponierte Matrix, 144
 Treppenfunktion, 77, 173
 Treppenfunktionen, 37
 trigonometrische Reihe, 50, 94
 trigonometrisches Polynom, 94

 Umkehrabbildung, 7
 Unendlich-Norm, 148
 unendlichdimensional, 103
 Unendliche Reihe, 28
 unitär, 144
 unteres Integral, 78, 174
 Untervektorraum
 erzeugter, 103
 Untervektorraum, 100, 103
 eindimensionaler, 100
 zweidimensionaler, 100
 Urbild, 5, 6

 Vektor, 98
 komplexer, 142
 Vektoraddition, 99
 Vektorfeld, 153
 Vektorprodukt, 107
 Vektorraum, 99
 arithmetischer, 100
 Vereinigung, 4
 beliebige, 7

 Verfeinerung
 gemeinsame, 76
 Verneinung, 1
 Vernichter, 119
 Vertauschbarkeit
 von Differentiation und Konvergenz, 90
 von Limes und Integral, 83
 Vertauschungssatz, 109
 Vielfaches
 skalares, 36
 Vollständigkeitsaxiom, 18
 Volumen
 p -dimensionales, 172
 eines Rotationskörpers, 93
 Vorzeichen
 einer Permutation, 128

 Wurzel, positive, 17
 Wurzelkriterium, 30

 Zeilenrang, 138
 Zeilenvektor
 einer Matrix, 120
 Zerlegung, 75
 feinere, 76
 Zustand
 angeregter, 119