Mathematik

für BiophysikerInnen
I und II

Vorlesung

gehalten von

Werner Römisch

Humboldt-Universität Berlin
Institut für Mathematik

Beginn: Wintersemester 2005/2006

Inhaltsverzeichnis Seite

0 Einleitung

Gegenstand des Gebietes „Mathematik“:

- klassische Grundlagen wie Mengen, Zahlen, Strukturen usw.;
- klassische lineare Algebra;
- klassische Differential- und Integralrechnung;
- Einführung in die gewöhnlichen Differentialgleichungen und Anfangsgründe partieller Differentialgleichungen;
- Grundlagen der Numerischen Mathematik.

Zielstellung:

- Vermittlung von Methoden und wesentlichen Ergebnissen der Mathematik;
- Wechselwirkung: praktisches Modell - mathematische Theorie - numerische, computergestützte Lösung.

Historie der Mathematik:

- antike Blüte der Mathematik: Pythagoras, Euklid, Archimedes;
- erst etwa im 15. Jh. entstand durch Probleme der Anwendung wieder das Bedürfnis nach Mathematik (Navigation, Kriegswesen, Astronomie, Optik);
- wichtige Mathematiker des 16./17. und 18. Jh.: Kepler, Newton, Leibniz, Fermat, Euler, Lagrange;
- 19. Jh.: Bolzano, Chauchy, Weierstraß, Cantor, Dedekind (exakte Begründung der Mathematik).

Anwendungsbeispiele (auf die wir später in der Vorlesung zurückkommen):

a) Populationsdynamik:
Es bezeichne $p(t)$ die Population einer gegebenen biologischen Art zum Zeitpunkt t.
Gesetz (Malthus, Anfang 19. Jh.): Die Populationsgeschwindigkeit ist proportional zur Population!

In mathematischen Termini:

$$\frac{dp}{dt} = \dot{p}(t) = a\ p(t) \qquad (a = \text{ Konstante})$$

Für die biologische Art „Mensch“: $a = 0.02$.
Das Gesetz ist demnach eine Differentialgleichung für die Funktion $p(t)$. Diese kann explizit gelöst werden:

$$p(t) = p(t_0)e^{0.02(t-t_0)},\ p(1961) = 3.06 \cdot 10^9.$$

Mit dieser Formel konnte die Populationsentwicklung der Menschheit in der Periode 1700-1961 überraschend genau widergespiegelt werden. Die Formel bedeutet die Verdopplung der Menscheit alle 34.6 Jahre. Sie ergibt aber in der Zukunft unrealistische Zahlen, z. B.:

2510: 2.84 m^2 für jeden Menschen, wobei auch die gesamte Wasseroberfläche der Erde genutzt wird,

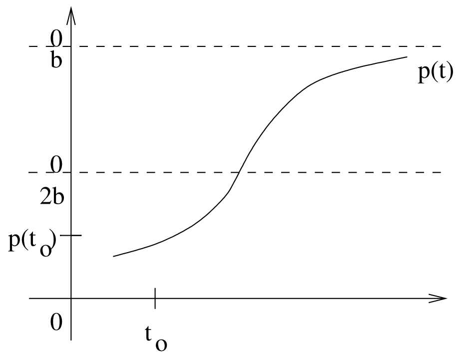
2670: $3.6{\cdot}10^{18}$ Menschen (jeder Mensch trägt einen weiteren auf der Schulter).

Ursache dieser unrealistischen Ergebnisse ist, daß menschliche Konflikte unberücksichtigt sind.

Einführung eines Konkurrenz-Terms, der die begrenzten Ressourcen und die Grenzen des Lebensraumes ausdrückt:

$\dot{p}(t) = ap(t) - bp(t)^2 \quad (b = \text{Konstante})$, Mensch: $a = 0.029, \quad b = 2.941 \cdot 10^{-12}$

Man kann zeigen: $p(2000) = 5.74 \cdot 10^9 p(t) \underset{t\to\infty}{\longrightarrow} \frac{a}{b} = 9.86 \cdot 10^9$



Wie wir wissen, ist der Wert $p(2000)$ nicht mehr korrekt. Also müssen a und b neu angepaßt werden.

b) „Tacoma Bridge“-Katastrophe:

- Eröffnung der Brücke am 1. 7. 1940 (Tacoma, Washington)
- von Beginn an (kleine) vertikale Schwingungen
- am 7. 11. 40, 7-10 Uhr, wellenförmige Bewegung, danach wilde Oszillation: 10.30 Uhr beginnt die Brücke zu krachen und 11.10 Uhr stürzt sie zusammen.
- Ursache: Aerodynamisches Phänomen
    - Luftstrom - Hindernis - Wirbel hinter dem Hindernis mit einer gewissen Periode, die von der Struktur des Hindernisses und der Geschwindigkeit des Luftstromes abhängt;
    - periodische Kraft senkrecht zum Luftstrom von der Größe $F_0 \cos \omega t$ ($\omega-$ Frequenz);

- Resonanz-Effekt mit der Frequenz der Struktur, d. h., Schwingung mit wachsender Amplitude!
- mathematischer Hintergrund: Lineare Differentialgleichung 2. Ordnung mit periodischer rechter Seite.

Literatur: (weitere Literatur in den Kapiteln)

H. Heuser:	Lehrbuch der Analysis, Teil 1 und 2, Teubner, Stuttgart 1990 (7. Auflage).
M. Braun:	Differentialgleichungen und ihre Anwendungen, Springer, Berlin 1994 (3. Auflage).
G. Hämmerlin, K.-H. Hoffmann:	Numerische Mathematik, Springer, Berlin 1994 (4. Auflage).

1 Grundlagen der Mathematik

1.1 Mathematische Aussagen und Beweise

Eine mathematische Aussage A ist eine Aussage, die entweder wahr (w) oder falsch (f) ist. Schreibweise: $A = w$ oder $A = f$

Zu jeder Aussage A verstehen wir unter ihrer Negation $\neg A$ („nicht A") die Aussage

$$\neg A = \begin{cases} w, & \text{falls } A = f \\ f, & \text{falls } A = w \end{cases}$$

Beispiele:

a) Sei a eine relle Zahl.
A sei die Aussage, daß a größer als 0 ist. Oder formalisiert: $A : a > 0$.
Dann gilt $\neg A : a \leq 0$.

b) A: für alle rellen Zahlen a gilt $a^2 \geq 0$.
Offenbar gilt $A = w$.
$\neg A$: es existiert eine relle Zahl a mit $a^2 < 0$.

c) A : es existieren natürliche Zahlen a, b, c, so daß $a^2 + b^2 = c^2$.
Es gilt: $A = w$.

d) A : Für keine natürliche Zahl $n > 2$ existieren natürliche Zahlen a, b, c, so daß $a^n + b^n = c^n$. Es gilt: $A = w$ (Fermat's Theorem, Beweis A. Wiles 1995)

Verknüpfungen von Aussagen

Name	Symbol	Schreibweise	Sprechweise	Bedeutung
Konjunktion	$\wedge$	$A \wedge B$	A und B	sowohl A als auch B
Disjunktion	$\vee$	$A \vee B$	A oder B	nicht ausschließendes oder
Implikation	$\Rightarrow$	$A \Rightarrow B$	wenn A, so B	aus A folgt B
Äquivalenz	$\Longleftrightarrow$	$A \Longleftrightarrow B$	A äquivalent B	A ist genau dann wahr (falsch), wenn B es ist.

Diese Verknüpfungen sind durch die folgenden Wertetabellen definiert:

A	B	¬ A	¬ B	A ∧ B	A ∨ B	A ⇒ B	A ⇔ B
w	w	f	f	w	w	w	w
w	f	f	w	f	w	f	f
f	w	w	f	f	w	w	f
f	f	w	w	f	f	w	w

Entscheidend für mathematische Beweismethoden sind die Negation, die Implikation und die Äquivalenz. Besonders die Implikation ist wichtig, da sie die Grundlage für direkte und indirekte Beweise ist:

- eine wahre Aussage impliziert immer eine wahre Aussage;
- eine falsche Aussage kann nur aus einer falschen Aussage gefolgert werden.

Für die Verknüpfungen von Aussagen gelten folgende Eigenschaften:

Eigenschaften:

1. $[A \Leftrightarrow B] \Leftrightarrow [(A \Rightarrow B) \wedge (B \Rightarrow A)]$

   Beweis mit Hilfe einer Wertetabelle für $C = [A \Leftrightarrow B], D = [A \Rightarrow B]$, $E = [B \Rightarrow A]$ und $D \wedge E$:

A	B	C	D	E	$D \wedge E$
w	w	w	w	w	w
w	f	f	f	w	f
f	w	f	w	f	f
f	f	w	w	w	w

   Also gilt $C = D \Leftrightarrow E$.

2. $\neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B)$

A	B	$\neg A$	$\neg B$	$A \wedge B$	$\neg(A \wedge B)$	$\neg A \vee \neg B$
w	w	f	f	w	f	f
w	f	f	w	f	w	w
f	w	w	f	f	w	w
f	f	w	w	f	w	w

3. $\neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B), (A \Leftrightarrow B) \Leftrightarrow (\neg A \Leftrightarrow \neg B)$ (Übung)

4. $(A \Rightarrow B) \Leftrightarrow (B \vee \neg A) \Leftrightarrow (\neg B \Rightarrow \neg A)$

A	B	$\neg A$	$A \Rightarrow B$	$B \vee \neg A$	$\neg B$	$\neg B \Rightarrow \neg A$
w	w	f	w	w	f	w
w	f	f	f	f	w	f
f	w	w	w	w	f	w
f	f	w	w	w	w	w

5. Kommutativitätsgesetze: (Übung)

   $(A \wedge B) \Leftrightarrow (B \wedge A), (A \vee B) \Leftrightarrow (B \vee A), (A \Leftrightarrow B) \Leftrightarrow (B \Leftrightarrow A)$

   Frage: $(A \Rightarrow B) \Leftrightarrow (B \Rightarrow A)$?
   Antwort: Nein $(A = w, B = f)$

6. Distributivgesetze:
   $[A \wedge (B \vee C)] \Leftrightarrow [(A \wedge B) \vee (A \wedge C)]$
   $[A \vee (B \wedge C)] \Leftrightarrow [(A \vee B) \wedge (A \vee C)]$ (Übung)

A	B	C	$D = (B \vee C)$	$E = (A \wedge B)$	$F = (A \wedge C)$	$A \wedge D$	$E \vee F$
w	w	w	w	w	w	w	w
w	w	f	w	w	f	w	w
w	f	w	w	f	w	w	w
f	w	w	w	f	f	f	f
w	f	f	f	f	f	f	f
f	w	f	w	f	f	f	f
f	f	w	w	f	f	f	f
f	f	f	f	f	f	f	f

<u>Mathematische Sätze und Beweise</u>
Die Mathematik besteht wesentlich in der Formulierung und im Beweisen von Sätzen. Zur Formulierung von Sätzen (math. Aussagen) benötigt man Begriffe, die definiert werden müssen. Gewisse Grundbegriffe muß man dabei als Grundlage akzeptieren.
Um Sätze zu beweisen, benötigt man immer bereits bewiesene Aussagen, aus denen die Behauptung des Satzes hergeleitet wird. Dabei muß man gewisse grundlegende Aussagen, die Axiome einer Theorie, als gültig annehmen.

Aufbau einer mathematischen Theorie:

$$\begin{tikzcd}
\text{Grundbegriffe} \arrow[r, "\text{Definition}"] \arrow[d] & \text{Begriffe} \arrow[d] \\
\text{Axiome} \arrow[r, "\text{Beweise}"] & \text{Sätze}
\end{tikzcd}$$

Mathematische Sätze bestehen stets aus einer Voraussetzung V und einer Behauptung B.
Der Satz sagt dann $V \Rightarrow B$ oder $V \Leftrightarrow B$.

Im Fall $V \Rightarrow B$ sagt man,	V ist hinreichend für B bzw. B ist notwendig für V.
Im Fall $V \Leftrightarrow B$ sagt man folglich,	daß V notwendig und hinreichend für B ist.

<u>Beispiele:</u>
$V: a, b$ relle Zahlen, $a = b$
$B: a^2 = b^2$
Es gilt: $V \Rightarrow B$ aber $B \not\Rightarrow V$.

<u>Satz:</u>
$V: a, b$ reell, $ab \geq 0$
$B: \frac{1}{2}(a+b) \geq \sqrt{ab}$ (Ungleichung zwischen arithmetischem und geometrischem Mittel)

„Beweis“:
Nach Quadrieren der Ungleichung in B erhalten wir
$\Rightarrow \frac{1}{4}(a^2 + 2ab + b^2) \geq ab \Rightarrow (a-b)^2 = a^2 - 2ab + b^2 \geq 0$,
und die letztere Aussage ist richtig.

<u>Frage:</u>
Ist der Beweis richtig und ist die Aussage richtig?

<u>Antwort:</u>
Der Beweis ist falsch, da man auch aus einer falschen Aussage eine richtige herleiten kann.
Der Beweis kann korrigiert werden, wenn „$\Rightarrow$“ durch „$\Leftrightarrow$“ ersetzt werden kann. Dazu benötigt man allerdings die stärkere Voraussetzung $V^*: a, b$ reell, $a \geq 0, b \geq 0$.

<u>Direkter</u> Beweis: Aus $V = w$ folgert man $V \Rightarrow B$.
<u>Indirekter</u> Beweis: Aus $V = w$ folgert man $\neg B \Rightarrow \neg V = f$.

<u>Beispiel:</u>
Im korrigierten obigen Satz gilt $\neg B: \frac{1}{2}(a+b) < \sqrt{ab}$.
Daraus leitet man $(a-b)^2 < 0$ her!

Für die rationelle Formulierung mathematischer Sachverhalte sind die folgenden „Operatoren“ nützlich:
$\forall$: „für alle“
$\exists$: „es existiert (mindestens) ein“.
$\exists!$: „es existiert genau ein“.
$\nexists$: „es existiert kein“.
Die Negation von $\forall(\exists)$ ist $\exists(\forall)$ usw. !

1.2 Mengen

Der Mengenbegriff ist ein Grundbegriff; er wird also nicht definiert, sondern zur gegenseitigen Verständigung erklärt.

Unter einer Menge M versteht man die Zusammenfassung wohlunterscheidbarer Objekte zu einem Ganzen. Diese Objekte heißen Elemente der Menge M.

Beschreibungsmöglichkeiten einer Menge M:

$$M = \{x : x \text{ besitzt die Eigenschaft } E\},$$

$$M = \{x_1, x_2, x_3, x_4\}, \quad M = \{x_1, \ldots, x_n\}, \quad M = \{x_1, x_2, \ldots\}.$$

Beispiele:

$I\!N = \{1, 2, 3, \ldots\}$ - Menge der natürlichen Zahlen (vgl. Kap. 1.5),

$\mathbb{Z} = \{x : x = 0 \vee x \in I\!N \vee -x \in I\!N\}$ - Menge der ganzen Zahlen,

$\mathbb{Q} = \{x : \exists p \in \mathbb{Z},\ q \in \mathbb{Z},\ q \neq 0, \text{ so daß } x = \frac{p}{q}\}$ - Menge der rationalen Zahlen.

$\mathbb{R}$ - Menge der reellen Zahlen.

Bezeichnungen und Definitionen
Es seien M und N Mengen.

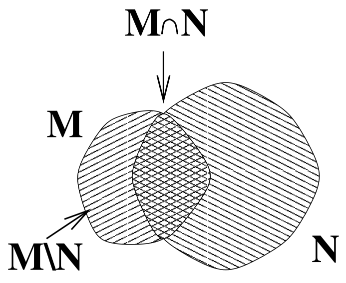
a) „$x \in M$“ bedeutet „x ist ein Element von M“ („x gehört zu M“),
„$x \notin M$“ bedeutet „x ist kein Element von M“ („x gehört nicht zu M“),

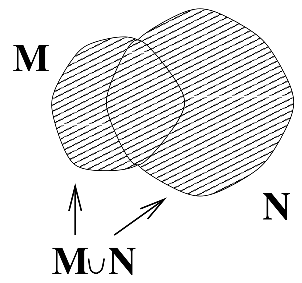
b) $\emptyset$ bezeichnet die leere Menge, d. h. die Menge, die keine Elemente enthält.

c) $[M = N] \Leftrightarrow [(x \in M) \Leftrightarrow (x \in N)]$
(Zwei Mengen sind gleich genau dann, wenn sie die gleichen Elemente enthalten.)
$[M \subseteq N] \Leftrightarrow [(x \in M) \Rightarrow (x \in N)]$ (Inklusion)
(Eine Menge M ist in einer Menge N enthalten genau dann, wenn alle Elemente von M auch Elemente von N sind.)
$[M \subset N] \Leftrightarrow [(M \subseteq N) \wedge (M \neq N)]$
(Eine Menge M ist in einer Menge N echt enthalten genau dann, wenn M in N enthalten ist und (mindestens) ein Element von N existiert, das nicht zu M gehört.)

d) $M \cup N = \{x : (x \in M) \vee (x \in N)\}$ (Vereinigung)
(Die Vereinigung $M \cup N$ besteht aus allen Elementen, die zu M oder zu N (oder zu beiden) gehören.)

e) $M \cap N = \{x :\in M) \wedge (x \in N)\}$ (Durchschnitt)
(Der Durchschnitt $M \cap N$ besteht aus allen Elementen, die sowohl zu M als auch zu N gehören.)





f) $M \backslash N = \{x : (x \in M) \wedge (x \notin N)\}$ (Differenz)
(Die Differenz $M \backslash N$ besteht aus allen Elementen, die zwar zu M aber nicht zu N gehören.)

g) $C_M(N) = M \backslash N$ (Komplement von $N \subseteq M$ bzgl. M)

Eigenschaften:
Es seien M, N und P Mengen.

1) $M \subseteq M, [(M \subseteq N) \wedge (N \subseteq P)] \Rightarrow [M \subseteq P]$,
$[(M \subseteq N) \wedge (N \subseteq M)] \Rightarrow [M = N]$;

2) $M \cap N \subseteq M \subseteq M \cup N, M \backslash N \subseteq M$;

3) $\emptyset \subseteq M, \; M \backslash M = \emptyset$;

4) $M \cup N = N \cup M, \quad M \cap N = N \cap M$ (Kommutativgesetze)
$(M \cup N) \cup P = M \cup (N \cup P), \; (M \cap N) \cap P = M \cap (N \cap P)$ (Assoziativgesetze)
$M \cap (N \cup P) = (M \cap N) \cup (M \cap P), \; (M \cup (N \cap P) = (M \cup N) \cap (M \cup P)$
(Distributivgesetze)

5) $M \cap N = M \backslash (M \backslash N)$;

6) Es seien $N \subseteq M$ und $P \subseteq M$. Dann gilt:
(i) $C_M(N \cup P) = C_M(N) \cap C_M(P)$
(ii) $C_M(N \cap P) = C_M(N) \cup C_M(P)$
(Morgan'sche Regeln)

Beweis:
1)-3) ist klar und bei 4) beweisen wir nur das erste Distributivgesetz. Hier und später zeigen wir jeweils zwei Inklusionen und verwenden dann die Gleichheitscharakterisierung aus 1):

$$\begin{aligned} \text{Sei } x \in M \cap (N \cup P) &\rightsquigarrow x \in M \text{ und } x \in N \cup P \\ &\rightsquigarrow x \in M \cap N \text{ oder } x \in M \cap P \\ &\rightsquigarrow x \in (M \cap N) \cup (M \cap P) \end{aligned}$$

$$\begin{aligned} \text{Sei } x \in (M \cap N) \cup (M \cap P) &\rightsquigarrow x \in M \text{ und } x \in N \text{ oder } x \in P \\ &\rightsquigarrow x \in M \cap (N \cup P) \end{aligned}$$

$$\text{Insgesamt:} \quad M \cap (N \cup P) \subseteq (M \cap N) \cup (M \cap P) \subseteq M \cap (N \cup P)$$

5) $x \in M \cap N \quad \rightsquigarrow \quad x \in M$ und $x \in N \rightsquigarrow x \notin M \backslash N \rightsquigarrow x \in M \backslash (M \backslash N)$

$\rightsquigarrow \quad M \cap N \subseteq M \backslash (M \backslash N)$

$x \in M \backslash (\backslash N) \quad \rightsquigarrow \quad x \in M$ und $x \notin (M \backslash N) \rightsquigarrow x \in M$ und $x \in N$

$\rightsquigarrow \quad x \in M \cap N \rightsquigarrow M \backslash (M \backslash N) \subseteq M \cap N$

6) (ii) (Übung)

(*i*) $x \in C_M(N \cup P) = M \backslash (N \cup P) \rightsquigarrow x \in M$ und $x \notin N \cup P$
$\rightsquigarrow x \notin N$ und $x \notin P \rightsquigarrow x \in C_M(N)$ und $x \in C_M(P)$
$\rightsquigarrow x \in C_M(N) \cap C_M(N) \rightsquigarrow C_M(N \cup P) \subseteq C_M(N) \cap C_M(P)$

$x \in C_M(N) \cap C_M(P) \rightsquigarrow x \in M$ und $x \notin N$ und $x \notin P$
$\rightsquigarrow x \in M \backslash (N \cup P) = C_M(N \cup P)$.

$\square$

Bemerkung:
Es besteht eine enge Verwandtschaft von Mengenoperationen $(\cup, \cap)$ und Aussagenverknüpfungen $(\vee, \wedge)$.

Definition:

(i) Es sei $I \neq \emptyset$ eine Menge. Die Zuordnung, die jedem $i \in I$ ein Objekt x_i zuordnet, heißt Familie. Für eine solche Familie verwendet man die Schreibweise $(x_i)_{i \in I}$ und nennt I den Indexbereich oder die Indexmenge dieser Familie.

(ii) Zwei Familien $(x_i)_{i \in I}$ und $(y_j)_{j \in J}$ heißen gleich, falls $I = J$ und $x_i = y_i$, $\forall i \in I$.

(iii) Eine Familie $(X_i)_{i \in I}$ heißt Mengenfamilie, falls für jedes $i \in I$ das Objekt X_i eine Menge ist.

(iv) Die Familien der Gestalt $(x_i)_{i \in I\!N}$ (Indexbereich $I = I\!N$) heißen Folgen. Schreibweise: $(x_i)_{i \in I\!N} = (x_1, x_2, \ldots)$. Eine Folge $(x_i)_{i \in I\!N}$ mit $x_i \in \mathbb{R}$ ($\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{C}$) für alle $i \in I\!N$ heißt Folge von rellen (ganzen, rationalen, komplexen) Zahlen.

Bezeichnungen:
Es sei $\mathcal{M}$ ein (endliches oder unendliches) System von Mengen (bei einem System sind evtl. nicht alle Elemente unterscheidbar; wir behandeln es im folgenden trotzdem wie eine Menge).

$$\bigcup_{M \in \mathcal{M}} M = \{x : \exists M \in \mathcal{M} \text{ mit } x \in M\}$$
$$\bigcap_{M \in \mathcal{M}} M = \{x : \forall M \in \mathcal{M} \text{ gilt } x \in M\}$$

In spezielleren Fällen verwenden wir auch folgende Schreibweise:

$$\bigcup_{i=1}^{n} M_i \text{ für } \bigcup_{M \in \mathcal{M}} M \text{ falls } \mathcal{M} = \{M_1, \ldots, M_n\},$$

$$\bigcap_{i \in I\!N} M_i = \bigcap_{i=1}^{\infty} M_i \text{ für } \bigcap_{M \in \mathcal{M}} M \text{ falls } \mathcal{M} = \{M_i : i \in I\!N\}.$$

Beispiel:
Sei $M_i = \{1, \ldots, i\}, i \in I\!N \rightsquigarrow \bigcup\limits_{i \in I\!N} M_i = I\!N, \bigcap\limits_{i \in I\!N} M_i = \{1\}.$

Übung: (erweiterte Morgan'sche Regel)
Es gelte $N \subseteq M$ für alle $N \in \mathcal{N}$($\mathcal{N}$ Mengensystem). Dann gilt:

$$C_M\left(\bigcup_{N \in \mathcal{N}} N\right) = \bigcap_{N \in \mathcal{N}} C_M(N),\ C_M\left(\bigcap_{N \in \mathcal{N}} N\right) = \bigcup_{N \in \mathcal{N}} C_M(N).$$

Definition:
Eine Menge mit $2, 3, \ldots, n$ $(n \in I\!N)$ Elementen, bei der ein erstes, ein zweites usw. ein n-tes Element festgelegt ist, heißt (geordnetes) Paar, Tripel, $\cdots$, n-Tupel.
Mögliche Bezeichnungen: (x, y) (Paar), , $(x_1, \ldots, x_n)$ (n-Tupel).

Für beliebige Mengen M und N heißt die Menge $M \times N = \{(x, y) : x \in M \wedge y \in N\}$, d. h. aller Paare von Elementen aus M und N, Produkt von M und N.
Allgemeiner definiert man das Produkt von n Mengen $M_1, \ldots, M_n$ als die Menge aller n-Tupel $(x_1, \ldots, x_n)$ mit $x_i \in M_i, i = 1, \ldots, n$, d. h.

$$\times_{i=1}^n M_i = M_1 \times M_2 \times \ldots \times M_n = \{(x_1, \ldots, x_n) : x_i \in M_i, i = 1, \ldots, n\}.$$

Beispiele:

$$\begin{aligned} I\!N \times I\!N = \{(m, n) : m \in I\!N, n \in I\!N\} &= \{(1,1), (1,2), (2,1), (1,3), (3,1), \\ &\quad (2,2), (2,3), (3,2), (3,3), \ldots\} \end{aligned}$$

$$M \times \emptyset = \emptyset,\ I\!N \times \{1\} = \{(n, 1) : n \in I\!N\}.$$

1.3 Abbildungen

Wir führen nun den für viele Gebiete der Mathematik zentralen Begriff der Abbildung von einer Menge X in eine Menge Y ein. Anschaulich ordnet eine Abbildung gewissen Elementen von X gewisse Elemente von Y zu, d. h. , eine Abbildung bestimmt Paare von einander zugehörigen Elementen.

Definition 1.1 *Es seien X und Y Mengen.*
Eine Menge $F \subseteq X \times Y$ heißt Abbildung aus X in Y.
Für jedes $x \in X$ heißt die Menge $F(x) := \{y \in Y : (x, y) \in F\}$ Wert von F in x.
$D(F) := \{x \in X : F(x) \neq \emptyset\} = \{x \in X : \exists y \in Y \text{ mit } (x, y) \in F\}$ heißt Definitionsbereich der Abbildung F.
$R(F) := \bigcup\limits_{x \in X} F(x) = \{y \in Y : \exists x \in X \text{ mit } (x, y) \in F\}$ heißt Wertebereich der Abbildung F.
Ist $(x, y) \in F$, so heißt x Urbild von y bzgl. F und y heißt Bild von x bzgl. F.
$F \subseteq X \times Y$ heißt Abbildung aus/von X, falls $D(F) \subset X / D(F) = X$,
in/auf Y, falls $R(F) \subset Y / R(F) = Y$.
$F \subseteq X \times Y$ heißt eindeutige Abbildung, falls für jedes $x \in X$ die Menge $F(x)$ höchstens ein Element besitzt. Sonst heißt sie mehrdeutig oder mengenwertig.

Bemerkung 1.2 *Für eindeutige Abbildungen $F \subseteq X \times Y$ existiert also zu jedem $x \in D(F)$ genau ein $y \in Y$ mit $(x, y) \in F$.*
In diesem Fall verwenden wir häufig folgende Schreibweisen:

$$
\begin{aligned}
&F : D(F) \longrightarrow Y \text{ und} \\
&F(x) = y \text{ oder } Fx = y \text{ für } (x, y) \in F
\end{aligned}
$$

(Wir identifizieren also die einelementige Menge $F(x)$ mit dem Element. Im Rahmen dieser Vorlesung werden fast ausschließlich eindeutige Abbildungen betrachtet.)
Bei eindeutigen Abbildungen sprechen wir im folgenden auch oft kurz von Abbildungen, Funktionen oder (seltener) Operatoren.

Beispiel 1.3

a) *Sei* $r \in Q_+ = Q \cap \mathbb{R}_+$. $F = \{(x, x^r) \in \mathbb{R} \times \mathbb{R} : x \geq 0\}$ *oder* $F(x) = x^r, \forall x \geq 0$, *heißt Potenzfunktion.(Beispiele für* $r = n, n \in I\!N$, *und* $r = \frac{1}{2}$*)*

b) $F : I\!N \longrightarrow I\!N,\ F(n) := n! = 1 \cdot 2 \cdot 3 \cdots n = \prod_{i=1}^{n} i, \forall n \in I\!N$ *(Fakultätsfunktion).*

c) *Ist* $\bar{y} \in Y$ *fixiert, so heißt* $F = X \times \{\bar{y}\}$ *die konstante Abbildung von* X *in* Y *mit Wert* $\bar{y}$.

d) *Ist* $X = Y$, *so ist* $F = \{(x, x) : x \in X\}$ *die identische Abbildung von* X *in* X. Bezeichnung: $F = I_X$

e) *Es seien* X *und* Y *beliebige Mengen. Dann heißen die Abbildungen*
$pr_1 : X \times Y \longrightarrow X,\ pr_1(x, y) = x$,
$pr_2 : X \times Y \longrightarrow Y,\ pr_2(x, y) = y$,
die erste bzw. zweite Projektion von $X \times Y$.

Definition 1.4 *Es sei* F *eine Abbildung aus* X *in* Y.
Für $A \subseteq X$ *heißt* $F(A) := \{y \in Y : \exists x \in A \text{ mit } (x, y) \in F\}$ *das Bild von* A *bzgl.* F.

$F^{-1} := \{(y, x) \in Y \times X : (x, y) \in F\}$ *heißt zu* F *inverse Abbildung.*

Die Abbildung F *heißt* *surjektiv, falls* $R(F) = Y$;
injektiv, falls F *und* F^{-1} *eindeutig sind;*
bijektiv, falls F *injektiv und surjektiv ist.*

Die Abbildung $F|_A := \{(x, y) \in F : x \in A\}$ *heißt Einschränkung von* F *auf* $A \subseteq X$.
Eine Abbildung $\hat{F} \subseteq X \times Y$ *heißt Fortsetzung von* F, *falls* $D(F) \subseteq D(\hat{F})$ *und* $F(x) = \hat{F}(x), \forall x \in D(F)$.

Beispiel 1.5

a) *Für jede nichtleere Menge* X *ist* I_X *bijektiv mit* $I_X^{-1} = I_X$.

b) *Die Projektionen* pr_1 *und* pr_2 *sind eindeutig und surjektiv. Es gilt z.B.* $pr_1^{-1} = \{(x, (x, y)) : y \in Y\}$ *oder* $pr_1^{-1}(x) = \{x\} \times Y$, $\forall x \in X$. *D.h.* pr_1^{-1} *ist nicht eindeutig.*

c) *Für* $r \in \mathbb{Q}_+$ *ist die Potenzfunktion* $F : \{x \in \mathbb{R} : x \geq 0\} \longrightarrow \mathbb{R}$, $F(x) = x^r$ *injektiv und es gilt* $F^{-1}(y) = y^{\frac{1}{r}}, \forall y \in R(F)$.

Eigenschaften 1.6
Es sei F *eine Abbildung von* X *in* Y; *es seien* $A, \tilde{A} \subseteq X$ *und* $B, \tilde{B} \subseteq Y$.

1) $A \neq \emptyset \Rightarrow F(A) \neq \emptyset$;
2) $A \subseteq \tilde{A} \Rightarrow F(A) \subseteq F(\tilde{A})$;
3) $F(A \cap \tilde{A}) \subseteq F(A) \cap F(\tilde{A})$, *wobei die Gleichheit gilt, falls* F^{-1} *eindeutig;*

4) $F(A \cup \tilde{A}) = F(A) \cup F(\tilde{A})$*;*

5) $F(A) = pr_2(F \cap (A \times Y))$ *und* $F^{-1}(B) = pr_1(F \cap (X \times B))$*;*

6) $D(F) = R(F^{-1})$ *und* $D(F^{-1}) = R(F)$*;*

7) $F^{-1}(B \cap \tilde{B}) \subseteq F^{-1}(B) \cap F^{-1}(\tilde{B})$*, wobei Gleichheit gilt, falls* F *eindeutig;*

8) $F^{-1}(B \cup \tilde{B}) = F^{-1}(B) \cup F^{-1}(\tilde{B})$*;*

9) $F(F^{-1}(B)) = B \cap R(F)$*, falls* F *eindeutig ist;*

10) $F^{-1}(F(A)) = A$*, falls* F^{-1} *eindeutig ist;*

11) $Z \subseteq pr_1(Z) \times pr_2(Z), \ \forall Z \subseteq X \times Y$.

Beweis:

1) und 2) sind klar.

3) $y \in F(A \cap \tilde{A}) \rightsquigarrow \exists x \in A \cap \tilde{A} : (x, y) \in F$
$\rightsquigarrow x \in F(A)$ und $x \in F(\tilde{A}) \rightsquigarrow x \in F(A) \cap F(\tilde{A})$.
Damit ist die Inklusion gezeigt. Es sei nun F^{-1} eindeutig.
Es sei $y \in F(A) \cap F(\tilde{A})$, d. h., $y \in F(A)$ und $y \in F(\tilde{A})$.
$\rightsquigarrow \ \exists x \in A : \ (x, y) \in F$ und $\exists \tilde{x} \in \tilde{A} : \ (\tilde{x}, y) \in F$
$\rightsquigarrow \ (y, x) \in F^{-1}$ und $(y, \tilde{x}) \in F^{-1}$
Da F^{-1} eindeutig ist, ist $F^{-1}(y)$ höchstens einelementig.
Wegen $x \in F^{-1}(y)$ und $\tilde{x} \in F^{-1}(y)$ muß $x = \tilde{x}$ gelten.
$\rightsquigarrow x = \tilde{x} \in A \cap \tilde{A} \rightsquigarrow y \in F(A \cap \tilde{A})$.

4) Übung (analog zum Beginn des Beweises von 3)).

5) Wir beweisen stellvertretend die erste Beziehung.
Es sei $y \in F(A) \rightsquigarrow \exists x \in A : (x, y) \in F$
oder $\exists x \in X : (x, y) \in F \cap (A \times Y)$
$\rightsquigarrow y \in pr_2(F \cap (A \times Y))$
Es sei $y \in pr_2(F \cap (A \times Y))$
$\rightsquigarrow \exists x \in X : (x, y) \in F \cap (A \times Y)$
$\rightsquigarrow x \in A$ und $(x, y) \in F \rightsquigarrow y \in F(A)$.

6) Übung (Anwendung der Definitionen)

7) Übung (Beweis analog zu 3))

8) Es sei $x \in F^{-1}(B \cup \tilde{B}) \rightsquigarrow \exists y \in B \cup \tilde{B} : (y, x) \in F^{-1}$
$\rightsquigarrow y \in B$ oder $y \in \tilde{B}$
$\rightsquigarrow x \in F^{-1}(B) \cup F^{-1}(B)$
Da jeder Schluß umkehrbar ist, folgt die Aussage.

9) Es sei F eindeutig und $y \in F(F^{-1}(B))$.
$\rightsquigarrow \exists x \in F^{-1}(B) \quad : \quad (x,y) \in F$
$\rightsquigarrow \exists \tilde{y} \in B \quad : \quad (\tilde{y}, x) \in F^{-1}$ bzw. $(x, \tilde{y}) \in F$
Aus der Eindeutigkeit von F folgt: $y = \tilde{y}$
$\rightsquigarrow y \in B$ und $y \in R(F)$
Es sei $y \in B \cap R(F) \rightsquigarrow \exists x \in X : (x,y) \in F$ bzw. $(y,x) \in F^{-1}$
$\rightsquigarrow x \in F^{-1}(B) \rightsquigarrow y \in F(F^{-1}(B))$.

10) Übung (ähnlich zu 9))

11) Sei $Z \subseteq X \times Y$ und $z = (x,y) \in Z$.
$\rightsquigarrow pr_1(z) = x$ und $pr_2(z) = y$
$\rightsquigarrow x \in pr_1(Z)$ und $y \in pr_2(Z) \rightsquigarrow z = (x,y) \in pr_1(Z) \times pr_2(Z)$.

□

Definition 1.7 *Es seien X, Y und Z Mengen und $F \subseteq X \times Y$, $G \subseteq Y \times Z$ seien Abbildungen. Die Abbildung*
$G \circ F := \{(x,z) \in X \times Z : \exists y \in Y$ mit $(x,y) \in F$ und $(y,z) \in G\}$
aus X in Z heißt Hintereinanderausführung oder Zusammensetzung der Abbildungen F und G.

Eigenschaften 1.8 *Es seien X, Y, Z, W Mengen und $F \subseteq X \times Y$, $G \subseteq Y \times Z$, $H \subseteq Z \times W$ Abbildungen. Dann gilt:*

1) $G \circ F(A) = G(F(A))$ für jede Teilmenge A von X;

2) $D(G \circ F) \subseteq D(F)$ und es gilt Gleichheit, falls $R(F) \subseteq D(G)$;

3) $H \circ (G \circ F) = (H \circ G) \circ F$;

4) $(G \circ F)^{-1} = F^{-1} \circ G^{-1}$.

Beweis:

1) $G \circ F(A) = \{z \in Z : \exists x \in A$ mit $(x,z) \in G \circ F\}$
$= \{z \in Z : \exists x \in A \exists y \in Y$ mit $(x,y) \in F$ und $(y,z) \in G\}$
$= \{z \in Z : \exists y \in F(A)$ und $(y,z) \in G\}$
$= G(F(A))$

2) Die Inklusion folgt aus der Definition.
Es gelte nun $R(F) \subseteq D(G)$ und es sei $x \in D(F)$.
$\rightsquigarrow \exists y \in Y : (x,y) \in F \rightsquigarrow y \in R(F) \subseteq D(G)$
$\rightsquigarrow \exists z \in Z : (y,z) \in G \rightsquigarrow (x,z) \in G \circ F \rightsquigarrow x \in D(G \circ F)$.

3) Übung

4) Sei $(z,x) \in (G \circ F)^{-1} \subseteq Z \times X \rightsquigarrow (x,z) \in G \circ F$
$\rightsquigarrow \exists y \in Y : (x,y) \in F$ und $(y,z) \in G$
$\rightsquigarrow \qquad (y,x) \in F^{-1}$ und $(z,y) \in G^{-1}$
$\rightsquigarrow (z,x) \in F^{-1} \circ G^{-1}$
Die andere Inklusion folgt aus der Umkehr aller Schlüsse.

□

Definition 1.9 *Es seien L und X zwei nichtleere Mengen.*
Unter einer Familie von Elementen von X mit Indexmenge L versteht man eine eindeutige Abbildung von L in X.
Schreibweise: $(x_\lambda)_{\lambda \in L}$
Ist L eine endliche oder unendliche Teilmenge von $\mathbb{N} \cup \{0\}$, so heißt $(x_\lambda)_{\lambda \in L}$ eine (endliche oder unendliche) Folge in X. Man nennt dann x_λ Folgenglied (Folgenelement) mit dem (Lauf-) Index λ.
Ist $L' \subseteq L$, so heißt die Einschränkung der Abbildung von L in X auf L' Teilfamilie von $(x_\lambda)_{\lambda \in L}$. Bei Folgen spricht man auch von Teilfolgen.

1.4 Algebraische Strukturen, Operationen und Relationen

Definition 1.10 *Es sei X eine nichtleere Menge.*
Eine eindeutige Abbildung $\circ$ von $X \times X$ in X heißt Operation oder Verknüpfung auf X. Wir nennen X eine algebraische Struktur, falls auf X eine oder mehrere Operationen mit „Rechenregeln" definiert sind.
Schreibweise: $(X, \circ), X, \circ, *)$ *usw.*

Definition 1.11 $(X, \circ)$ *heißt Gruppe, falls $X \neq \emptyset$ und $\circ$ eine Operation auf X ist, die die folgenden Eigenschaften besitzt:*

(i) $x \circ (y \circ z) = (x \circ y) \circ z$ *für alle $x, y, z \in X$ (Assoziativgesetz);*

(ii) Es existiert ein Element $e \in X$ („neutrales Element"), so daß $x \circ e = e \circ x = x$, *für alle $x \in X$;*

(iii) Zu jedem $x \in X$ existiert ein $x^* \in X$ („zu x inverses Element"), so dass $x \circ x^* = x^* \circ x = e$.

Eine Gruppe $(X, \circ)$ heißt Abelsch oder kommutativ, falls $x \circ y = y \circ x$ gilt, für alle $x, y \in X$.

Beispiel 1.12 $(\mathbb{Z}, +)$ *und* $(\mathbb{Q}, +)$ *sind Abelsche Gruppen,* $(\mathbb{N}, +)$ *ist keine Gruppe,* $(\mathbb{Q}, \cdot)$ *ist keine Gruppe.*

Definition 1.13 $(X, \circ, *)$ *heißt Körper, falls $X \neq \emptyset$ und $\circ$ bzw. $*$ Operationen auf X sind, die die folgenden Eigenschaften besitzen:*

(i) $(X, \circ)$ *ist eine Abelsche Gruppe (mit neutralem Element e);*

(ii) $x * (y * z) = (x * y) * z, \ \forall x, y, z \in X$ *(Assoziativität von $*$); es existiert ein Element $\mathbb{1} \in X$, $\mathbb{1} \neq e$ („Einselement"), so daß $\mathbb{1} * x = x * \mathbb{1} = x$ für alle $x \in X$; zu jedem $x \in X, x \neq e$, existiert ein $\bar{x} \in X$, so daß $\bar{x} * x = x * \bar{x} = \mathbb{1}$;*

(iii) $x * (y \circ z) = (x * y) \circ (x * z)$, $(x \circ y) * z = (x * z) \circ (y * z)$, *für alle $x, y, z \in X$ (Distributivgesetze).*

Ein Körper $(X, \circ, *)$ heißt Abelsch oder kommutativ, falls zusätzlich $x * y = y * x$ für alle $x, y \in X$ gilt.

Beispiel 1.14 $(\mathbb{Q}, +, \cdot)$ *ist ein Abelscher Körper mit neutralem Element* $e = 0$ *und Einselement* $\mathbb{1} = 1$.

Eigenschaften 1.15

(1) Es sei $(X, \circ)$ *eine Gruppe. Dann sind das neutrale Element und sämtliche inversen Elemente eindeutig bestimmt.*

(2) Es sei $(X, \circ, *)$ *ein Körper. Dann gelten die folgenden Rechenregeln:*
$x * e = e * x = e$ *für alle* $x \in X$*;*
$x * y = e$ *genau dann, wenn* $x = e$ *oder* $y = e$
(hierbei ist e *das neutrale Element);*
für alle $a, b \in X$ *mit* $a \neq e$ *existiert genau ein* $x \in X$ *mit* $a * x = b$ *(Lösbarkeit von Gleichungen).*

Beweis:

(1) Wir zeigen die Eindeutigkeit des neutralen Elements. Dazu nehmen wir an, daß es zwei neutrale Elemente e_1 und e_2 gibt. Dann gilt wegen Regel (ii) in Def. 1.11:

$$e_2 \circ e_1 = e_2 = e_1 \circ e_2 = e_1 \text{ d. h. } e_1 = e_2.$$

Weiterhin nehmen wir an, daß es zu einem $x \in X$ zwei inverse Elemente x_1^* und x_2^* gibt. Dann folgt aus den Regeln (i)-(iii):

$$x_2^* = e \circ x_2^* = (x_1^* \circ x) \circ x_2^* = x_1^* \circ (x \circ x_2^*) = x_1^* \circ e = x_1^*.$$

(2) Sei $x \in X$ beliebig. Dann folgt durch Anwendung von (iii) in Def. 1.13:

$$x * e = x * (e \circ e) = (x * e) \circ (x * e).$$

Also ist $x * e$ neutrales Element, d. h., $x * e = e$.
Analog zeigt man $e * x = e$.
Die zweite Aussage ist eine Übung (Hinweis: Man verwende das zu x inverse Element bzgl. der Operation $*$).
Es seien $a, b \in X$ mit $a \neq e$. Ist $\bar{a} \in X$ so gewählt, daß $a * \bar{a} = \mathbb{1}$, so ist $x = \bar{a} * b$ eine Lösung der Gleichung $a * x = b$.
Annahme: Es existieren zwei Lösungen x_1 und x_2 in X, d. h.,

$$a * x_1 = a * x_2 = b.$$

$$\rightsquigarrow \quad \bar{a} * (a * x_1) = \bar{a} * (a * x_2) \rightsquigarrow \mathbb{1} * x_1 = \mathbb{1} * x_2 \rightsquigarrow x_1 = x_2.$$

□

Beispiel 1.16 *Es sei $m \in I\!N$ und $X(m) := \{0, 1, 2, \dots m-1\}$. Wir definieren die folgenden Operationen für bel. $x, y \in X(m)$:*

$$x \circ y := \begin{cases} x+y \quad , & x+y \leq m-1 \\ x+y-m, & x+y \geq m \end{cases}$$

$$x * y := r, \quad \text{wobei } r \in \{0, 1, \dots, m-1\} \text{ definiert ist durch } r := x \cdot y - km$$
$$\text{mit einem geeigneten } k \in I\!N \cup \{0\} \text{ („Rest bei Division durch } m\text{“)}$$

Neutrales Element: $e = 0$; Einselement: $1\!\!1 = 1$

Inverses Element zu $x \in X(m)$ bzgl. $\circ$: $x^* = \begin{cases} m-x & , \ x \neq 0 \\ 0 & , \ x = 0 \end{cases}$

Es gilt: $(X(m), \circ)$ ist für jedes $m \in I\!N$ eine Abelsche Gruppe;

$(X(m), \circ, *)$ ist ein Abelscher Körper, falls $m = p$ eine Primzahl ist (Übung).

Ein weiterer Grundbegriff ist der Begriff einer <u>Relation</u> R in einer Menge X. Man versteht darunter eine Teilmenge von $X \times X$.
<u>Schreibweise</u>: $xRy \Leftrightarrow (x, y) \in R$

Definition 1.17 *Es sei X eine Menge und R eine Relation in X.*
R heißt eine <u>Äquivalenzrelation</u> in X, wenn R die folgenden Eigenschaften besitzt:

(i) xRx (Reflexivität);

(ii) $xRy \Leftrightarrow yRx$ (Symmetrie);

(iii) $xRy \wedge yRz \Rightarrow xRz$ (Transitivität).

<u>Schreibweise:</u> $x \sim y$ anstelle von xRy („x ist äquivalent zu y“).

Beispiel 1.18 *Es sei X eine nichtleere Menge und $(A_\lambda)_{\lambda \in L}$ eine Familie von Teilmengen von X mit den folgenden Eigenschaften*

$$A_\lambda \cap A_\mu = \emptyset, \forall \lambda, \mu \in L \text{ mit } \lambda \neq \mu; \quad \bigcup_{\lambda \in L} A_\lambda = X.$$

(Man spricht dann auch von einer Zerlegung von X.)
Wir definieren nun eine Relation R in X durch xRy für je 2 Elemente $x, y \in X$, falls $\exists \lambda \in L$ mit $x \in A_\lambda$ und $y \in A_\lambda$. Dann ist R eine Äquivalenzrelation.

<u>Übung:</u>
Man zeige, daß umgekehrt auch jede Äquivalenzrelation in X eine Zerlegung von X definiert.

Definition 1.19 *Eine Relation R in einer nichtleeren Menge X heißt Ordnungsrelation, falls sie die Eigenschaften (i) und (iii) der Reflexivität und Transitivität aus Definition 1.17 besitzt und überdies (ii)* erfüllt.*
(ii)* $xRy \wedge yRx \Rightarrow x = y$ (Antisymmetrie).
Schreibweise: $x \leq y$ anstelle von xRy („x ist kleiner gleich y“).
$(X, \leq)$ heißt dann geordnete Menge.

Beispiel 1.20 *Jede Menge von Mengen ist mit „$\subseteq$“ eine geordnete Menge.*
Die Menge $(\mathbb{Q}, \leq)$, wobei „$\leq$“ die übliche „kleiner oder gleich“-Beziehung zwischen Zahlen ist, ist geordnet.

Definition 1.21 *Es sei $(X, \leq)$ eine geordnete Menge und es sei $A \subseteq X$.*
A heißt „nach oben beschränkt“ falls ein $c \in X$ existiert, so dass $x \leq c$ für alle $x \in A$. c heißt dann obere Schranke von A.
Es sei nun A nach oben beschränkt und $S := \{c \in X : x \leq c, \ \forall x \in A\}$ die Menge aller oberen Schranken. Falls ein $s \in S$ existiert mit $s \leq c$ für alle $c \in S$, so heißt s die kleinste obere Schranke von A oder das Supremum von A.
Bezeichnung: $\sup A := s$
Analog definiert man „nach unten beschränkt“, „größte untere Schranke von A“= „Infimum von A“.

Beispiel 1.22 *Wir betrachten die geordnete Menge $(\mathbb{Q}, \leq)$.*
$A := \{x \in \mathbb{Q} : x < 1\}$ ist nach oben beschränkt. Jede rationale Zahl $c \geq 1$ ist obere Schranke und es gilt $1 = \sup A \notin A$.
$A := \{x \in \mathbb{Q} : x > 1\}$ ist nicht nach oben beschränkt!

2 Zahlen und Räume

2.1 Reelle Zahlen und Zahlbereiche

Im folgenden werden zunächst die reellen Zahlen axiomatisch („per Postulat“) eingeführt und aus diesen relativ wenigen Axiomen die Eigenschaften der reellen Zahlen hergeleitet.

Definition 2.1 *Eine Menge $\mathbb{R}$ heißt Menge der reellen Zahlen, falls zwei Operationen „$+$“ und „$\cdot$“ (Addition und Multiplikation) und eine Ordnungsrelation „$\leq$“ in $\mathbb{R}$ definiert sind, die die folgenden Eigenschaften haben:*

(I) $(\mathbb{R}, +, \cdot)$ *ist ein Abelscher Körper;*

(II) $(\mathbb{R}, \leq)$ *hat die (zusätzlichen) Eigenschaften:*
(II_1) $\forall x, y \in \mathbb{R}$ *gilt* $x \leq y$ *oder* $y \leq x$,
(II_2) $\forall x, y, z \in \mathbb{R}$ *mit* $x \leq y$ *gilt* $x + z \leq y + z$,
(II_3) *aus* $x, y \in \mathbb{R}$ *mit* $x \geq 0$ *und* $y \geq 0$ *folgt* $xy \geq 0$;

(III) für alle $A \subseteq \mathbb{R}, A \neq \emptyset$ *und A nach oben beschränkt, existiert das Supremum von A in* $\mathbb{R}$.

Bemerkung 2.2 *Die <u>Existenz</u> einer nichtleeren Menge $\mathbb{R}$ mit den in Def. 2.1 angeführten Eigenschaften wird als Grundlage für den gesamten weiteren Aufbau postuliert! $\mathbb{R}$ ist durch Def. 2.1 nicht eindeutig bestimmt, jedoch unterscheiden sich verschiedene Modelle von $\mathbb{R}$ nur in Eigenschaften, die für den weiteren mathematischen Aufbau uninteressant sind.*
All dies ist kaum problematisch, da die Axiome der anschaulichen Vorstellung der „Zahlen" entsprechen. Interessant ist dabei vielleicht (III), das die Vorstellung von der Lückenlosigkeit von $\mathbb{R}$ sichert!

Bezeichnungen 2.3
0 bzw. 1 bezeichnen das neutrale Element bzw. das Einselement in $\mathbb{R}$,
für $x \in \mathbb{R}$ bezeichnet $-x$ das inverse Element zu x bzgl. „+",
für $x \in \mathbb{R}\backslash\{0\}$ bezeichnet $x^{-1} = \frac{1}{x}$ das inverse Element bzgl. „·",
für den Sachverhalt $x \leq y$ und $x \neq y$ für $x, y \in \mathbb{R}$ schreiben wir $x < y$,
für $a, b \in \mathbb{R}$ mit $a < b$ bezeichne
$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$ *(abgeschlossenes Intervall)*
$]a, b[:= \{x \in \mathbb{R} : a < x < b\}$ *(offenes Intervall)*
$[a, b[:= \{x \in \mathbb{R} : a \leq x < b\}$ *(halboffenes Intervall)*
$]a, b] := \{x \in \mathbb{R} : a < x \leq b\}, \quad \mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$,
für $x \in \mathbb{R}$ heißt $|x| := \begin{cases} x, & \text{falls } x \geq 0 \\ -x, & \text{falls } x < 0 \end{cases}$ *der absolute Betrag von x.*

Eigenschaften 2.4 *(Rechenregeln in $\mathbb{R}$)*
Es seien $a, b \in \mathbb{R}$. Dann gilt:

1) $ab > 0 \Leftrightarrow [(a > 0) \wedge (b > 0)] \vee [(a < 0) \wedge (b < 0)]$;
   $aa > 0 \Leftrightarrow a \neq 0$; *insbesondere gilt* $1 > 0$ *und* $a > 0 \Leftrightarrow \frac{1}{a} > 0$;

2) $a < b \Rightarrow a < \frac{1}{2}(a + b) < b$ *(hierbei ist $2 := 1 + 1$);*

3) $|ab| = |a||b|, \quad |a + b| \leq |a| + |b|$ *(Dreiecksungleichung) und*
   $||a| - |b|| \leq |a - b|$.

<u>Beweis:</u>

1) Es gelte $ab > 0$. *Annahme:* $a > 0$ und $b < 0$.
   $$(II)_2 \rightsquigarrow -b > 0 \rightsquigarrow a(-b) = a(-1)b = -ab > 0 \quad (\text{aus } (II_3) \wedge (I))$$
   $$\rightsquigarrow ab < 0 \rightsquigarrow \text{Widerspruch.}$$
   Damit ist die Richtung „$\Rightarrow$" gezeigt.
   Zum Beweis der Umkehrung schlußfolgern wir aus (II_3) und $a > 0, b > 0$ bzw. $a < 0, b < 0$, daß $ab \geq 0$ ist.
   (im zweiten Fall gilt $(-a)(-b) = (-1)(-1)ab = ab \geq 0$)
   Annahme: $ab = 0$
   Dies ist aber wegen 2) in 1.15 unmöglich!
   Insbesondere gilt $a \cdot a > 0 \Leftrightarrow a \neq 0$; sowie insbesondere $1 > 0$, da $1 \neq 0$, und $a \cdot \frac{1}{a} = 1 > 0 \rightsquigarrow a$ und $\frac{1}{a}$ sind entweder beide > 0 oder beide < 0.

2) Es gelte $a < b$. Mit $2 := 1 + 1$ schlußfolgern wir aus (II_2):
$a + a = (1 + 1)a = 2a < a + b < b + b = 2b$
$\rightsquigarrow 0 < (a + b) - 2a \rightsquigarrow 0 < \frac{1}{2}[(a + b) - 2a]$
$\rightsquigarrow 0 < \frac{1}{2}(a + b) - a$
$\rightsquigarrow a < \frac{1}{2}(a + b)$
Analog folgt die rechte Seite der Ungleichung.

3) *Übung:* $|ab| = |a||b|$
Wir beweisen die Dreiecksungleichung. Nach Definition gilt:

$$\left.\begin{array}{ll} a \leq |a|, & -a \leq |a| \\ b \leq |b|, & -b \leq |b| \end{array}\right\} \underset{(II_2)}{\Longrightarrow} \quad \begin{array}{rl} a + b & \leq |a| + |b| \\ -a + (-b) & \leq |a| + |b| \end{array}$$

$$\begin{array}{lrl} \rightsquigarrow & -(|a| + |b|) \leq & a + b \leq |a| + |b| \\ \rightsquigarrow & & |a + b| \leq |a| + |b| (\text{ nach Definition von } | \cdot |) \end{array}$$

Zum Beweis der letzten Aussage folgern wir aus der Dreieckungleichung:

$$\begin{array}{rcllr} |a| & \leq & |a - b| + |b| & \text{und} & |b| \leq |b - a| + |a| \\ \rightsquigarrow |a| - |b| & \leq & |a - b| & \text{und} & |b| - |a| \leq |a - b|. \end{array}$$

□

Ähnlich wie die Rechenregeln in 2.4 können nun alle bekannten Rechenregeln mit reellen Zahlen aus den Axiomen in Def. 2.1 hergeleitet werden. Wir verzichten darauf und arbeiten mit reellen Zahlen wie gewohnt.
Im folgenden definieren wir nun die Zahlbereiche $I\!N, \mathbb{Z}$ und $\mathbb{Q}$ als Teilmengen von $\mathbb{R}$ mit gewissen Eigenschaften.

Definition 2.5

a) *Eine Menge $M \subseteq \mathbb{R}$ heißt* <u>*induktiv*</u>*, falls*
(i) $1 \in M$ und (ii) $x + 1 \in M, \forall x \in M$.
$\mathcal{M}_i := \{M : M \subseteq \mathbb{R}$ und M ist induktiv$\}$ bezeichnet die Menge aller induktiven Mengen.

b) *$I\!N := \bigcap\limits_{M \in \mathcal{M}_i} M$ heißt* <u>*Menge der natürlichen Zahlen,*</u>
$I\!N_0 := I\!N \cup \{0\}$.

c) *$\mathbb{Z} := \{x \in \mathbb{R} : x \in I\!N_0$ oder $-x \in I\!N\}$ heißt* <u>*Menge der ganzen Zahlen.*</u>

d) *$\mathbb{Q} := \{x \in \mathbb{R} : \exists p \in \mathbb{Z} \exists q \in I\!N$ mit $x = \frac{p}{q}\}$ heißt* <u>*Menge der rationalen Zahlen.*</u>
$\mathbb{R} \setminus \mathbb{Q}$ heißt <u>*Menge der irrationalen Zahlen.*</u>

Bemerkung 2.6 *Es existieren viele induktive Mengen, z. B.* $\mathbb{R}, \{x \in \mathbb{R} : x \geq 1\}$ *und* $\{x \in \mathbb{R} : x = 1, x \geq 2\}$ *usw. Nach Definition ist die Menge* $I\!N$ *der natürlichen Zahlen die kleinste aller induktiven Mengen!*
Diese Eigenschaft von $I\!N$ *begründet das sog. Induktionsprinzip:*
Ist $M \subseteq N$ *und zeigt man, daß* M *induktiv ist, so gilt* $M = I\!N$*!*
Dieses Prinzip ist von großer Bedeutung für die Beweistechnik („vollständige Induktion“) bzw. zur Definition von Größen.

(i) Prinzip der vollständigen Induktion:
Um zu zeigen, daß eine Aussage $A(n)$ *für alle* $n \in I\!N$ *richtig ist, genügt es zu zeigen:* $\{n \in I\!N : A(n) \text{ ist wahr}\}$ *ist induktiv.*
D.h. man geht wie folgt wor:
1. Man zeigt: $A(1)$ *ist wahr,*
2. Man zeigt für jedes $n \in I\!N : [A(n) \text{ ist wahr} \Rightarrow A(n+1) \text{ ist wahr}]$.

(ii) Induktives Definieren:
Um eine Größe $G(n)$ *für jedes* $n \in I\!N$ *zu definieren, geht man wie folgt vor:*
1. $G(1)$ *definieren (oft trivial),*
2. für jedes $n \in I\!N$ *wird* $G(n+1)$ *in Abhängigkeit von* $G(n)$ *definiert.*
Beispiele:
$n! := (n-1)! \cdot n, \quad \forall n \in I\!N, \quad 0! := 1;$
$a^n := a^{n-1} \cdot a, \quad \forall n \in I\!N, \quad a^0 := 1 \quad (\forall a \in \mathbb{R});$
$\sum\limits_{i=1}^{n} a_i := \sum\limits_{i=1}^{n-1} a_i + a_n, \quad \forall n \in I\!N, \quad \sum\limits_{i=1}^{0} a_i := 0;$
$\prod\limits_{i=1}^{n} a_i := \left(\prod\limits_{i=1}^{n-1} a_i\right) a_n, \quad \forall n \in I\!N, \quad \prod\limits_{i=1}^{0} a_i := 1 \quad (\forall a_i \in \mathbb{R}, i = 1, \ldots, n).$

Wir befassen uns nun mit Eigenschaften von $I\!N$ und der „Lage “von $I\!N$ und $\mathbb{Q}$ innerhalb von $\mathbb{R}$.

Satz 2.7
1) $I\!N$ *ist nicht nach oben beschränkt in* $\mathbb{R}$*. (Archimedes)*
2) $\forall \varepsilon > 0 \exists n \in I\!N : \frac{1}{n} < \varepsilon$*. (Eudoxos)*

Beweis:

1) *Annahme:* $I\!N$ ist nach oben beschränkt.
Aus (III) von Def. 2.1 folgt dann: $s := \sup I\!N \in \mathbb{R}$
$\rightsquigarrow s - 1$ ist keine obere Schranke von $I\!N$
$\rightsquigarrow \exists n \in I\!N : s - 1 < n \rightsquigarrow s < n + 1 \in I\!N$
$\rightsquigarrow s$ ist keine obere Schranke von $I\!N$
$\rightsquigarrow$ Widerspruch zur Annahme!

2) *Annahme:* $\exists \varepsilon_0 > 0 : \frac{1}{n} \geq \varepsilon_0, \forall n \in I\!N$
$\rightsquigarrow n \leq \frac{1}{\varepsilon_0}, \forall n \in I\!N \rightsquigarrow I\!N$ ist nach oben beschränkt
$\rightsquigarrow$ Widerspruch! $\square$

Satz 2.8 *(Wohlordnungssatz)*
Jede nichtleere Teilmenge von $I\!N$ besitzt ein kleinstes Element.

Beweis:
Es sei $M \neq \emptyset$ eine Teilmenge von $I\!N$.
Annahme: M besitzt kein kleinstes Element.
Es sei $M_u := \{n \in I\!N : n < m, \forall m \in M\}$ die Teilmenge von $I\!N$, die sämtliche unteren Schranken von M enthält, die zu $I\!N \setminus M$ gehören.
Beh.: M_u ist induktiv.
Bew.: $1 \in M_u$, da sonst $1 \in M$ kleinstes Element von M ist.
Es sei $n \in M_u$ und wir zeigen $n+1 \in M_u$.
Annahme: $n+1 \notin M_u$
$\rightsquigarrow \exists m_1 \in M : m_1 \leq n+1$
$\rightsquigarrow \exists m_2 \in M : m_2 < m_1$, da sonst m_1 kleinstes Element von M wäre.
$\rightsquigarrow m_2 \leq n$, da sonst $n < m_2 < n+1$ gelten würde
$\rightsquigarrow n \notin M_u \rightsquigarrow$ Widerspruch!
Also ist M_u induktiv, d. h., $M_u = I\!N \rightsquigarrow M = \emptyset \rightsquigarrow$ Widerspruch! $\square$

Satz 2.9 $\forall x \in \mathbb{R}\ \forall \varepsilon > 0 \exists r \in \mathbb{Q} \cap]x-\varepsilon, x+\varepsilon[$. *(„$\mathbb{Q}$ ist dicht in $\mathbb{R}$“)*

Beweis:
Es sei $x \in \mathbb{R}$ und zunächst $x \geq 0$, und es sei $\varepsilon > 0$ bel. gewählt.
Nach Satz 2.7 existiert ein $n \in I\!N$ mit $\frac{1}{n} < \varepsilon$.
Es sei $M := \{m \in I\!N : m > nx\}$. Nach Satz 2.7 gilt $M \neq \emptyset$.
Nach Satz 2.8 besitzt M ein kleinstes Element $m \in M$.
$\rightsquigarrow m - 1 \leq nx < m$.
Dann gilt für $r := \frac{m}{n} \in \mathbb{Q} : r - \varepsilon < r - \frac{1}{n} = \frac{m-1}{n} \leq x < r$
$\rightsquigarrow x - \varepsilon < x < r < x + \varepsilon$, d. h., $r \in]x-\varepsilon, x+\varepsilon[$
Es sei nun $x < 0$. Dann existiert nach dem Beweis von eben ein
$r \in \mathbb{Q} : -x-\varepsilon < r < -x+\varepsilon \rightsquigarrow x - \varepsilon < -r < x + \varepsilon$ und $-r \in \mathbb{Q}$. $\square$

Als nächstes beantworten wir die Frage, ob $\mathbb{Q}$ und $\mathbb{R}$ wirklich verschieden sind. Dazu benötigen wir noch einige Vorbereitungen.

Satz 2.10

1) *Für alle $a, b \in \mathbb{R}$ und $n \in I\!N$ gilt:*
   $(a+b)^n = \sum_{i=0}^{n} \begin{pmatrix} n \\ i \end{pmatrix} a^{n-i} b^i$ *(„Binomischer Lehrsatz“).*
   wobei $\begin{pmatrix} n \\ i \end{pmatrix} := \frac{n!}{i!(n-i)!}, \forall i = 0, \ldots, n$ *(„Binomialkoeffizienten“).*

2) *Für alle $a \in \mathbb{R}, a > -1$, und alle $n \in I\!N$ gilt:* $(1+a)^n \geq 1 + na$
   (Bernoullische Ungleichung).

3) *Für alle $a \in [0,1]$ und alle $n \in I\!N$ gilt:* $(1+a)^n \leq 1 + (2^n - 1)a$.

Beweis:

1) Es seien $a, b \in \mathbb{R}$ beliebig und wir beweisen die Aussage mit dem Induktionsprinzip.
Für $n = 1$ ist sie richtig wegen $(a+b)^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} a^1 b^0 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} a^0 b^1 = a + b$.
Die Aussage sei für n richtig und wir beweisen sie für $n+1$.

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n(a+b) = \left(\sum_{i=0}^{n} \begin{pmatrix} n \\ i \end{pmatrix} a^{n-i} b^i\right)(a+b) \\
&= \sum_{i=0}^{n} \begin{pmatrix} n \\ i \end{pmatrix} a^{n+1-i} b^i + \sum_{i=0}^{n} \begin{pmatrix} n \\ i \end{pmatrix} a^{n-i} b^{i+1} \\
&= a^{n+1} + \sum_{i=1}^{n} \left[\begin{pmatrix} n \\ i \end{pmatrix} + \begin{pmatrix} n \\ i-1 \end{pmatrix}\right] a^{n+1-i} b^i + b^{n+1} \\
&= \sum_{i=0}^{n+1} \begin{pmatrix} n+1 \\ i \end{pmatrix} a^{n+1-i} b^i
\end{aligned}$$

wegen $\begin{pmatrix} n \\ i \end{pmatrix} + \begin{pmatrix} n \\ i-1 \end{pmatrix} = \dfrac{(n+1)!}{i!(n-i+1)!} = \begin{pmatrix} n+1 \\ i \end{pmatrix}, \forall i = 1, \ldots, n.$

2) Es sei $a \in \mathbb{R}, a > -1$ und $M := \{n \in I\!N : (1+a)^n \geq 1 + na\}$.
Wir zeigen: M ist induktiv.
Es gilt natürlich $1 \in M$. Es sei $n \in M$ und wir zeigen $n+1 \in M$.

$$\begin{aligned}
(1+a)^{n+1} &= (1+a)^n(1+a) \geq (1+na)(1+a) = 1 + (n+1)a + na^2 \\
&\geq 1 + (n+1)a \rightsquigarrow n+1 \in M
\end{aligned}$$

3) Übung. □

Satz 2.11 *Es seien $a > 0$ und $n \in I\!N$. Dann existiert genau ein $x \in \mathbb{R}$ mit $x > 0$ und $x^n = a$.*

Beweis:
Wir betrachten die Menge $M := \{y \in \mathbb{R} : y > 0 \text{ und } y^n \leq a\}$.
Beh.: M ist nichtleer und nach oben beschränkt.
Bew.: Es gilt $1 \in M$, falls $a \geq 1$, und $a \in M$, falls $a < 1$. In jedem Fall haben wir also $M \neq \emptyset$, da $\min\{1, a\} \in M$. Ferner gilt für jedes $y \in M : y^n \leq a < 1+a \leq (1+a)^n$, d.h. $y < 1+a$ (da $y \geq 1+a$ impliziert $y^n \geq (1+a)^n$.) Also ist M nach oben berschränkt.
Deshalb gilt nach Axiom (III) in Def. 2.1: $x := \sup M \in \mathbb{R}$ und $x > 0$, wegen $\min\{1, a\} \leq x$.
Beh.: $x^n = a$.
Bew.: *Annahme:* $x^n < a$.
Wir zeigen: $\exists \varepsilon > 0$ mit $x + \varepsilon \in M (\rightsquigarrow x < \sup M)$.
Es gilt: $(x+\varepsilon)^n = x^n(1 + \frac{\varepsilon}{x})^n \leq x^n \left(1 + (2^n - 1)\frac{\varepsilon}{x}\right)$ (2.12.!)

$\rightsquigarrow (x+\varepsilon)^n \leq x^n + (2^n - 1)\, x^{n-1}\varepsilon$
Wir wählen nun $\varepsilon > 0$ so klein, daß $(2^n-1)x^{n-1}\varepsilon \leq a - x^n$, und erhalten $(x+\varepsilon)^n \leq a$, d. h., $x + \varepsilon \in M$.
Annahme: $x^n > a$
Wir zeigen: $\exists \varepsilon > 0$ mit $\sup M \leq x - \varepsilon < x$.
Es gilt: $(x-\varepsilon)^n = x^n \left(1 - \frac{\varepsilon}{x}\right)^n \geq x^n \left(1 - \frac{\varepsilon}{x} n\right)$, falls $\varepsilon < x$ (nach 2.12).
Wir wählen nun $\varepsilon \in]0, x[$ so klein, daß $\frac{\varepsilon n}{x} \leq \frac{x^n - a}{x^n}$.
$\rightsquigarrow (x-\varepsilon)^n \geq x^n \left(1 - \frac{x^n-a}{x^n}\right) = a \rightsquigarrow y^n \leq a \leq (x-\varepsilon)^n,\ \forall y \in M$
$\rightsquigarrow \sup M \leq x - \varepsilon$.
In beiden Fällen ergibt sich also ein Widerspruch zu $x = \sup M$.
Also folgt $x^n = a$.
Schließlich zeigen wir noch die Eindeutigkeit von x.
Annahme: $\exists x_1, x_2 \in \mathbb{R}$ mit $x_1 \neq x_2$ und $x_1^n = a = x_2^n$.
Nach (II_1) in Def. 2.1 gilt $x_1 < x_2$ oder $x_1 > x_2$. Es sei o.B.d.A. $x_1 < x_2$.
Nach (II_3) folgt $x_1^n < x_2^n \rightsquigarrow$ Widerspruch. $\square$

Bemerkung 2.12
Ein Element $x \in \mathbb{R}$ mit der Eigenschaft von Satz 2.11 heißt n-te Wurzel aus a.
Bezeichnung: $\sqrt[n]{a},\ a^{\frac{1}{n}}$
Nun lassen sich beliebige Potenzen positiver reeller Zahlen a mit rationalen Exponenten definieren:

$$a^\circ := 1,\ a^n,\ \forall n \in \mathbb{N}\ \text{vgl. 2.6},\ a^{-n} := \frac{1}{a^n},\ \forall n \in \mathbb{N},$$

$$a^{\frac{p}{q}} := \sqrt[q]{a^p}, \quad \forall p \in \mathbb{Z}\ \forall q \in \mathbb{N}.$$

Überdies gelten die bekannten Rechenregeln für die Potenzrechnung!

Satz 2.13 *Es gilt $\sqrt{2} \notin \mathbb{Q}$, d.h., $\sqrt{2}$ ist eine irrationale Zahl.*

Beweis:
Annahme: $\sqrt{2} \in \mathbb{Q}$, d.h., $\exists n, m \in \mathbb{N} : \sqrt{2} = \frac{m}{n}$ (da $\sqrt{2} > 1$).
Wir betrachten $M := \{n \in \mathbb{N} : n\sqrt{2} \in \mathbb{N}\}$. Da M nichtleer ist, besitzt M nach Satz 2.8 ein kleinstes Element $n_0 \in M$.
$\rightsquigarrow n_0\sqrt{2} \in \mathbb{N}$ und $m_0 := n_0\sqrt{2} - n_0 \in \mathbb{N}$
$\rightsquigarrow m_0\sqrt{2} = 2n_0 - n_0\sqrt{2} \in \mathbb{N} \rightsquigarrow m_0 \in M$ und $m_0 < n_0$
$\rightsquigarrow$ Widerspruch! $\square$

Folgerung 2.14

1) *Zwischen je zwei voneinander verschiedenen (rationalen) reellen Zahlen liegt eine (rationale) reelle Zahl.*

2) *Zwischen je zwei voneinander verschiedenen (rationalen) reellen Zahlen liegt eine (irrationale) rationale Zahl.*

Beweis:

1) folgt aus Eigenschaften 2.4 ($a < \frac{1}{2}(a+b) < b, \ \forall a,b \in \mathbb{Q}$ bzw. $\in \mathbb{R}$).

2) Seien $a,b \in \mathbb{R},\ a \neq b$. Dann liefert Satz 2.9 die Existenz einer rationalen Zahl, die zwischen a und b liegt.
Es seien nun $a,b \in \mathbb{Q}$ mit $a < b$. Dann gilt $\frac{a}{\sqrt{2}} < \frac{b}{\sqrt{2}}$.
Dann folgt aus Satz 2.9: $\exists r \in \mathbb{Q}$ mit $\frac{a}{\sqrt{2}} < r < \frac{b}{\sqrt{2}}$ und $r \neq 0$.
$\rightsquigarrow a < \sqrt{2}r < b$ und $\sqrt{2}r \notin \mathbb{Q}$
(letzteres folgt daraus, daß $\mathbb{Q}$ ein Körper ist und $\sqrt{2} \notin \mathbb{Q}$). $\square$

Als nächstes leiten wir mögliche Darstellungen reeller Zahlen her.

Satz 2.15 *Es seien $m \in \mathbb{N}, m \geq 2$, und $x \in \mathbb{R}$ mit $x \geq 0$.
Dann existieren eindeutig bestimmte Zahlen $k \in \mathbb{N}_0$ und $z_i \in \{n \in \mathbb{N}_0 : n \leq m-1\}$, so daß x die m-Darstellung*

$$x = z_1 z_2 \cdots z_k \cdot z_{k+1} z_{k+2} \cdots$$

besitzt. Dabei gilt

$$k := \min\{n \in \mathbb{N}_0 : x < m^n\} \ \text{ und } \ z_j \leq \frac{1}{m^{k-j}}\left(x - \sum_{i=1}^{j-1} z_i m^{k-i}\right) < z_j + 1 \ \ (j \in \mathbb{N}).$$

<u>Beweis:</u>
Die Zahl $k := \min\{n \in \mathbb{N}_0 : x < m^n\}$ ist nach Satz 2.8) wohldefiniert und wir zeigen, daß wir sukzessive die Zahlen $z_j \in \{n \in \mathbb{N}_0 : n \leq m-1\},\ j \in \mathbb{N}$,
so wählen können, daß $z_j m^{k-j} \leq x - \sum\limits_{i=1}^{j-1} z_i m^{k-i} < (z_j+1)m^{k-j}$
oder $z_j \leq \frac{1}{m^{k-j}}\left(x - \sum\limits_{i=1}^{j-1} z_i m^{k-i}\right) < z_j + 1 \qquad (*)$
oder $\sum\limits_{i=1}^{j} z_i m^{k-i} \leq x < \sum\limits_{i=1}^{j} z_i m^{k-i} + m^{k-j} \qquad (\forall j \in \mathbb{N})$.

Ist also die obige Konstruktion der $z_j,\ \forall j \in \mathbb{N}$, möglich, so ist die Aussage bewiesen, da die z_j dann eindeutig bestimmt wären. Wir zeigen die Möglichkeit der Konstruktion der z_j mit dem Induktionsprinzip. Zunächst gilt nach Definition von k, daß

$$m^{k-1} \leq x < m^k.$$

Für $j = 1$ bedeutet nun (*) $\ z_1 \leq \frac{x}{m^{k-1}} < z_1 + 1$. Wegen $\frac{x}{m^{k-1}} \in [0, m[$ kann aber $z_1 \in \{0, \cdots, m-1\}$ so gewählt werden, daß dies erfüllt ist.
Es gelte nun $\frac{1}{m^{k-j}}\left(x - \sum\limits^{j-1} i = 1 z_i m^{k-i}\right) \in [0,m[$ und $z_j \in \{0, \cdots, m-1\}$ sei wie in (*) gewählt. Dann gilt für $j+1$, daß

$$\frac{1}{m^{k-j-1}}\left(x - \sum_{i=1}^{j} z_i m^{k-i}\right) = m\left[\underbrace{\frac{1}{m^{k-j}}\left(x - \sum_{i=1}^{j-1} z_i m^{k-i}\right)}_{\in[0,m[} - z_j\right] \in [0, m[$$

und z_{j+1} ist definiert. $\square$

Bemerkung 2.16
Die Darstellung reeller Zahlen in Satz 2.15 heißt m-Darstellung. Für $m = 2$ spricht man von Dual-Darstellung und für $m = 10$ von Dezimal-Darstellung.
Die Zahlen $z_i \in \{n \in \mathbb{N}_0 : n \leq m-1\}$ nennt man auch Ziffern und die Zahl $k \in \mathbb{N}_0$ entspricht der Anzahl der Ziffern vor dem „Komma“.

Beispiel 2.17

$x = \frac{1}{2}$ *und* $m = 10: \;\rightsquigarrow k = 0,\; z_1 = 5 \rightsquigarrow x = .500\ldots$ *(Dezimal-Darstellung)*
$x = \frac{1}{2}$ *und* $m = 2: \;\rightsquigarrow k = 0,\; z_1 = 1 \rightsquigarrow x = .100\ldots$ *(Dual-Darstellung)*

Wir wollen nun untersuchen, wie umfangreich die Mengen $\mathbb{N}, \mathbb{Q}, \mathbb{R}, \mathbb{R}\backslash\mathbb{Q}$ sind. Dazu benötigen wir noch einige Vorbereitungen allgemeiner Art.

Definition 2.18 *Es seien X und Y Mengen.*

a) X und Y heißen gleichmächtig, falls eine bijektive Abbildung von X auf Y existiert. Schreibweise: $X \sim Y$

b) X heißt endlich, falls $X = \emptyset$ oder $\exists m \in \mathbb{N}$ mit $X \sim \{n \in \mathbb{N} : n \leq m\}$.

c) X heißt abzählbar, falls $X \sim \mathbb{N}$.

d) X heißt höchstens abzählbar, falls X endlich oder abzählbar ist.

c) X heißt überabzählbar, falls X nicht höchstens abzählbar ist.

Bemerkung 2.19
Die Relation $\sim$ zwischen Mengen stellt eine Äquivalenzrelation dar.
Es gilt $X \sim X$, die Symmetrie $X \sim Y \Leftrightarrow Y \sim X$ und die Transitivität (Übung).
Die Endlichkeit oder Abzählbarkeit einer Menge stellt eine Numerierungsmöglichkeit der Elemente dar: $X = \{x_1, \ldots, x_m\}$ oder $X = \{x_1, x_2, \ldots\}$.

Satz 2.20
1) Jede Teilmenge von $\mathbb{N}$ ist höchstens abzählbar.
2) Die Menge $\mathbb{N} \times \mathbb{N}$ ist abzählbar.

Beweis:

1) Es sei M eine Teilmenge von $\mathbb{N}$. Es genügt, den Fall zu betrachten, daß M nicht endlich ist. Wir zeigen: Dann ist M abzählbar.
Dazu betrachten wir die folgenden, induktiv definierten Elemente von M.
$x_1 := \min M, x_n := \min M\backslash\{x_1, \ldots, x_{n-1}\}, n \geq 2$ (gemäß Satz 2.8); und wir definieren die Abbildung

$$f : \mathbb{N} \longrightarrow M,\; f(n) := x_n,\; \forall n \in \mathbb{N}.$$

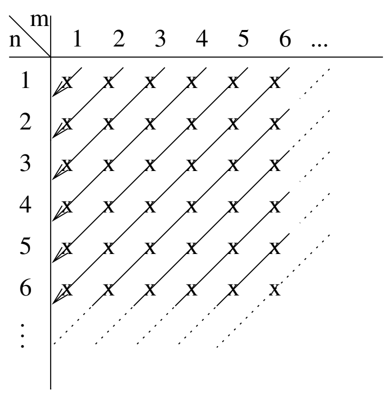
Da M nicht endlich ist, ist f eine Abbildung von $\mathbb{N}$ in M.
Wir zeigen: f ist bijektiv (und folglich $\mathbb{N} \sim M$).

Nach Definition gilt $x_i < x_n$, falls $i < n$, und $n \leq x_n$, $\forall n \in I\!N$.
Deshalb ist zunächst f injektiv.
Um zu zeigen, daß f auch surjektiv ist, sei nun $x \in I\!N$ beliebig gewählt. O.B.d.A. gelte $x > x_1$, da sonst $f(1) = x$.
Wir definieren $m := \min\{n \in I\!N : x \leq x_n\}$ (nach Satz 2.8).
Dann gilt nach Konstruktion $x_1 < x_2 < x_{m-1} < x < x_{m+1}$ und damit $f(m) = x$, d. h., $x \in f(I\!N) \rightsquigarrow M = f(I\!N)$ und M ist abzählbar.

2) Wir betrachten die folgende Abbildung $f : I\!N \times I\!N \longrightarrow I\!N$,
$f(m,n) := n + \frac{1}{2}(m+n)(m+n+1),\ \forall m,n \in I\!N$.
Wenn diese Abbildung f injektiv wäre, wäre $I\!N \times I\!N$ das bijektive Bild der Teilmenge $f(I\!N \times I\!N)$ von $I\!N$. Nach Teil 1) ist aber $f(I\!N \times I\!N)$ und damit $I\!N \times I\!N$ höchstens abzählbar. Da aber $I\!N \times I\!N$ nicht endlich ist, wäre dann $I\!N \times I\!N$ abzählbar.
Es genügt also zu zeigen, daß f injektiv ist.
Es seien dazu $(m_1, n_1), (m_2, n_2) \in I\!N \times I\!N$ mit $f(m_1, n_1) = f(m_2, n_2)$ gewählt.
Wir müssen zeigen: $(m_1, n_1) = (m_2, n_2)$.
1. Fall: $s_1 := m_1 + n_1 = m_2 + n_2 := s_2$
$\rightsquigarrow f(m_1, n_1) - f(m_2, n_2) = n_1 - n_2 \rightsquigarrow n_1 = n_2 \rightsquigarrow m_1 = m_2$.
2. Fall: $s_1 < s_2$ (o.B.d.A.)
$\rightsquigarrow f(m_1, n_1) = n_1 + \frac{1}{2}s_1(s_1+1) < s_1 + \frac{1}{2}s_1(s_1+1) = \frac{1}{2}(s_1+1)(s_1+2) - 1.$
$\leq \frac{1}{2}s_2(s_2+1) - 1 < n_2 + \frac{1}{2}s_2(s_2+1) = f(m_2, n_2).$
$\rightsquigarrow$ Widerspruch!
Also muß $s_1 = s_2$ gelten und die Aussage ist bewiesen.

$\square$

Bemerkung 2.21 *Die in Teil 2) des obigen Beweises definierte Abbildung f kann als Durchnumerierung von $I\!N \times I\!N$ nach dem sog. Cantor'schen Diagonalprinzip interpretiert werden. Sortiert man nämlich die Werte von f entsprechend der Ordnungsrelation in $\mathbb{R}$ und numeriert diese neu durch, so entspricht dies dem Abzählverfahren entlang der Pfeile beginnend von links oben in der Abbildung. Die eingezeichneten Kreuze ent-*



sprechen dabei jeweils einem Paar $(m,n) \in I\!N \times I\!N$. Z. B. gilt $f(1,1) = 4$, $f(2,1) = 7$, $f(1,2) = 8$, $f(3,1) = 11$, $f(2,2) = 12$, $\ldots$.

Satz 2.22

1) Es seien X und Y Mengen, X sei abzählbar und $f : X \longrightarrow Y$ surjektiv. Dann ist Y höchstens abzählbar.

2) Es seien L eine Indexmenge und $X_\lambda, \lambda \in L$, Mengen. Alle diese Mengen seien höchstens abzählbar. Dann ist auch die Menge $\bigcup\limits_{\lambda \in L} X_\lambda$ höchstens abzählbar.

Beweis:

1) Es sei $X = \{x_1, x_2, \ldots, x_n, \ldots\}$ und wir betrachten die Abbildung $g : I\!N \longrightarrow Y,\ g(n) := f(x_n),\ \forall n \in I\!N$.

   Da nach Voraussetzung g surjektiv ist, läßt sich auch die folgende Abbildung definieren

$$h : Y \longrightarrow I\!N,\ h(y) := \min\{n \in I\!N : g(n) = y\},\ \forall y \in Y.$$

   (Dies ist möglich, da wegen der Surjektivität von g, die Menge $\{n \in I\!N : g(n) = y\}$ für jedes $y \in Y$ nichtleer und Satz 2.8 anwendbar ist.)
   Nach Konstruktion ist h eindeutig und wegen $g(h(y)) = y$ auch eineindeutig, d. h., injektiv. Also ist h eine bijektive Abbildung von Y auf $h(Y) \subseteq I\!N$. Da $h(Y)$ nach Satz 2.20 höchstens abzählbar ist, gilt dies nach Definition auch für Y.

2) Nach Voraussetzung existieren surjektive Abbildungen $n \longrightarrow \lambda_n$ (von $I\!N$ auf L) und $m \longrightarrow n_\lambda^{(m)}$ (von $I\!N$ auf X_λ) für jedes $\lambda \in L$.
   Wir definieren nun die folgende Abbildung

$$f : I\!N \times I\!N \longrightarrow \bigcup_{\lambda \in L} X_\lambda,\ \ f(m,n) := x_{\lambda_n}^{(m)},\ \ \forall (m,n) \in I\!N \times I\!N.$$

   Diese Abbildung ist nach Konstruktion surjektiv und die Menge $I\!N \times I\!N$ ist nach Satz 2.20 abzählbar. Also folgt aus Teil 1), daß die Menge $\bigcup_{\lambda \in L} X_\lambda$ höchstens abzählbar ist. $\square$

Beispiel 2.23 *(Hilbert's Hotel oder das Paradoxon der Abzählbarkeit)*
Ein Hotel besitze abzählbar viele Zimmer und sie seien alle belegt. G_n bezeichne den Gast in Zimmer $n \in I\!N$.
Plötzlich erscheinen abzählbar viele neue Gäste $g_n, n \in I\!N$, und fragen nach Zimmern. Da nach Satz 2.22 die Menge $\{G_n : n \in I\!N\} \cup \{g_n : n \in I\!N\}$ wieder abzählbar ist, können die alten und die neuen Gäste im Hotel untergebracht werden.
Wie kann das in einfacher Weise erfolgen? Antwort:
Der Portier verteilt die Zimmer wie folgt neu: G_n wohnt in Zimmer $2n$ und g_n in $2n-1$, $\forall n \in I\!N$. Allerdings muß dann jeder Gast umziehen.

Wir kommen nun zurück zur Untersuchung von Teilmengen von $\mathbb{R}$ und ihrer Mächtigkeiten.

Satz 2.24
Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar.

Beweis:
Wir schreiben $\mathbb{Q}$ in der folgenden Form:

$$\mathbb{Q} = \{r \in \mathbb{Q} : r > 0\} \cup \{0\} \cup \{r \in \mathbb{Q} : r < 0\}$$

Nach Satz 2.22 genügt es zu zeigen, daß die beiden Mengen $\{r \in \mathbb{Q} : r > 0\}$ und $\{r \in \mathbb{Q} : r < 0\}$ abzählbar sind. Da die zweite Menge genau die negativen Zahlen der ersten Menge enthält, muß nur die Abzählbarkeit von

$$\{r \in \mathbb{Q} : r > 0\} = \{r = \frac{m}{n} : m, n \in I\!N\}$$

gezeigt werden. Dazu definieren wir die Abbildung $f : I\!N \times I\!N \longrightarrow \{r \in \mathbb{Q} : r > 0\}$ durch $f(m,n) = \frac{m}{n}$, $\forall m, n \in I\!N$.
Nach Definition ist f surjektiv und nach Satz 2.20 ist $I\!N \times I\!N$ abzählbar. Also liefert Satz 2.22, 1): $\{r \in \mathbb{Q} : r > 0\}$ ist höchstens abzählbar. Wegen $I\!N \subset \{r \in \mathbb{Q} : r > 0\}$ kann $\{r \in \mathbb{Q} : r > 0\}$ aber nicht endlich und muß abzählbar sein. □
Unser nächstes Ziel besteht darin, zu zeigen, daß $\mathbb{R}$ überabzählbar ist.

Satz 2.25
Die Menge $]0,1[\subset \mathbb{R}$ *ist überabzählbar.*

Beweis:
Da die Zahlen $\frac{1}{n}$ für jedes $n \in I\!N$ zu $]0,1[$ gehören, ist die Menge $]0,1[$ nicht endlich.
Annahme: $]0,1[$ ist abzählbar.
Wir bezeichnen mit $\{x_i : i \in I\!N\}$ sämtliche Elemente von $]0,1[$. Nach Satz 2.15 läßt sich jedes x_i in Dezimal-Darstellung schreiben: $x_i = .z_{1i}z_{2i}z_{3i}\cdots$, wobei $z_{ji} \in \{0,1,\ldots,9\}$, $\forall i, j \in I\!N$, eindeutig bestimmt ist.
Wir wählen nun für alle $i \in I\!N$ eine „Ziffer“ $z_i \in \{1,2,\ldots,8\}\backslash\{z_{ii}\}$ und betrachten nun die relle Zahl

$$x := \sup\left\{\sum_{i=1}^{n} z_i 10^{-i} : n \in I\!N\right\}.$$

Dieses Element x hat die Dezimal-Darstellung $x = .z_1z_2z_3\cdots z_i\cdots$ und es gilt $x \in]0,1[$.
Überdies gilt nach Konstruktion $x \neq x_i$, $\forall i \in I\!N$. Dies ist aber ein Widerspruch zur Annahme. Also kann $]0,1[$ nicht höchstens abzählbar sein. □

Satz 2.26
Die Mengen $\mathbb{R}$ *und* $\mathbb{R}\backslash\mathbb{Q}$ *sind überabzählbar und gleichmächtig.*

Beweis:
Wäre $\mathbb{R}$ höchstens abzählbar, so würde aus Satz 2.23 folgen, daß $]0,1[\subset \mathbb{R}$ ebenfalls höchstens abzählbar ist. Dies ist aber nach Satz 2.25 unmöglich. Also ist $\mathbb{R}$ überabzählbar. Wäre $\mathbb{R}\backslash\mathbb{Q}$ höchstens abzählbar, so auch $\mathbb{R} = (\mathbb{R}\backslash\mathbb{Q}) \cup \mathbb{Q}$ nach Satz 2.22, da $\mathbb{Q}$ abzählbar nach Satz 2.27 ist. Also ist auch $\mathbb{R}\backslash\mathbb{Q}$ überabzählbar.
Um die Gleichmächtigkeit von $\mathbb{R}$ und $\mathbb{R}\backslash\mathbb{Q}$ zu zeigen, müssen wir eine bijektive Abbildung $f : \mathbb{R} \longrightarrow \mathbb{R}\backslash\mathbb{Q}$ konstruieren.
Dazu wählen wir zunächst eine abzählbare Menge X wie folgt aus $\mathbb{R}\backslash\mathbb{Q}$ aus:

$x_1 \in \mathbb{R}\backslash\mathbb{Q},\ x_{n+1} \in (\mathbb{R}\backslash\mathbb{Q})\backslash\{x_1, \ldots, x_n\}$, $n \geq 2$, und $X := \{x_n : n \in I\!N\}$.
Dieser Auswahlprozeß kann nicht abbrechen und die Menge $Y := (\mathbb{R}\backslash\mathbb{Q})\backslash X$ ist nichtleer, da $\mathbb{R}\backslash\mathbb{Q}$ nicht höchstens abzählbar ist.
Nach Satz 2.22 ist die Menge $X \cup \mathbb{Q}$ abzählbar. Deshalb existiert eine bijektive Abbildung $g : X \cup \mathbb{Q} \longrightarrow X$.
Wir definieren nun die gewünschte Abbildung wie folgt:

$$f : \mathbb{R} \longrightarrow \mathbb{R}\backslash\mathbb{Q}\,,\ f(x) := \begin{cases} x, & \forall x \in Y = \mathbb{R}\backslash(X \cup \mathbb{Q}), \\ g(x), & \forall x \in X \cup \mathbb{Q}. \end{cases}$$

Damit ist f surjektiv und auch injektiv, da sowohl die identische Abbildung von Y also auch g injektiv sind. $\square$

Wir sehen also, daß es vielmehr irrationale Zahlen als rationale gibt. Im folgenden führen wir nun weitere Mengen neben $\mathbb{R}$ als Grundstrukturen der Mathematik ein. Sie bauen aber alle auf $\mathbb{R}$ auf.

2.2 Komplexe Zahlen

Die historische Motivation für die Erweiterung der reellen Zahlen $\mathbb{R}$ war, die Potenzen a^r mit $r \in \mathbb{Q}$ auch für $a < 0$ zu definieren (vgl. 2.11). Unser *Ausgangspunkt* ist die folgende Fragestellung:
Kann man die Menge aller Paare $\{(x, y) : x, y \in \mathbb{R}\} = \mathbb{R} \times \mathbb{R}$ reeller Zahlen durch Einführung geeigneter Operationen zu einem Körper machen ?
Wir definieren dazu die folgenden *Operationen* auf $\{(x, y) : x, y \in \mathbb{R}\}$:

$$(x, y) + (u, v) := (x + u, y + v), (x, y) \cdot (u, v) := (xu - yv, xv + yu) (\forall x, y, u, v \in \mathbb{R})$$

Satz 2.27 $\mathbb{C} := (\{(x, y) : x, y \in \mathbb{R}\}, +, \cdot)$ *ist ein Abelscher Körper mit Nullelement* $(0, 0)$ *und Einselement* $(1, 0)$.
Bezeichnung: $\mathbb{C}$ *heißt Körper der komplexen Zahlen.*

Beweis:
Beide Operationen „$+$“ und „$\cdot$“ sind nach Definition kommutativ, da es die entsprechenden Operationen auf $\mathbb{R}$ sind.
Klar ist auch, daß $(\{(x, y) : x, y \in \mathbb{R}\}, +)$ eine Gruppe ist. Die Assoziativität folgt aus der in $\mathbb{R}$ und das inverse Element zu (x, y) ist (natürlich) $(-x, -y)$.
Zum Nachweis der Körperaxiome (vgl. 1.13) wenden wir uns jetzt der Operation „$\cdot$“ zu. Zunächst gilt:

$$(x, y) \cdot (1, 0) = (x - 0, 0 + y) = (x, y),\ \ \forall x, y \in \mathbb{R},$$

d. h. $(1, 0)$ ist das Einselement. Es sei nun $(x, y) \neq (0, 0)$ und wir betrachten

$$(x, y) \cdot \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2}\right) = \left(\frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2}, \frac{-xy}{x^2 + y^2} + \frac{xy}{x^2 + y^2}\right) = (1, 0)$$

Also ist $\left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2}\right)$ das inverse Element zu (x, y).
Nachzuweisen bleiben nun nur noch das Assoziativgesetz und das Distributivgesetz.

Stellvertretend untersuchen wir die Gültigkeit des Distributivgesetzes.
Es seien dazu $x, y, u, v, w, z \in \mathbb{R}$ beliebig gewählt. dann gilt:

$$\begin{aligned}(x,y)\cdot((u,v)+(w,z)) &= (x,y)\cdot(u+w,v+z)\\ &= (x(u+w)-y(v+z),x(v+z)+y(u+w))\\ &= (xu-yv,xv+yu)+(xw-yz,xz+yw)\\ &= (x,y)\cdot(u,v)+(x,y)\cdot(w,z)\end{aligned}$$

□

Bemerkung 2.28

a) *Alle Rechenregeln mit reellen Zahlen, die allein aus den Körper-Eigenschaften von* $(\mathbb{R},+,\cdot)$ *resultieren, lassen sich auf* $\mathbb{C}$ *übertragen. Vorsicht aber bei allem, was aus der Ordnungsrelation hergeleitet wurde (z. B. die Wurzel in 2.11).*

b) *Wir betrachten die folgende Teilmenge* $\{(x,0) : x \in \mathbb{R}\}$ *von* $\mathbb{C}$.
Nach Definition gilt:

$$(x,0)+(y,0)=(x+y,0),(x,0)\cdot(y,0)=(xy,0),\forall x,y\in\mathbb{R}.$$

Die Operationen führen also aus dieser Teilmenge nicht heraus und, gesehen vom Standpunkt algebraischer Strukturen, besteht kein Unterschied zwischen $(x,0)\in\mathbb{C}$ *und* $x\in\mathbb{R}$. *Deshalb kann man* $\mathbb{R}$ *mit der Menge* $\{(x,0) : x\in\mathbb{R}\}$ *in folgendem Sinn identifizieren: Es sei* pr_1 *die Projektion von* $\mathbb{R}\times\mathbb{R}$ *auf die erste Komponente (vgl. 1.3). Dann ist die Abbildung*

$$pr_1 : \{(x,0) : x\in\mathbb{R}\} \longrightarrow \mathbb{R}, \quad pr_1(x,0)=x,\ \forall x\in\mathbb{R},$$

bijektiv. Da es also hierbei keine Verwechslung bei der Zuordnung von Urbild und Bild geben kann, werden wir einfach schreiben:

$$(x,0)=x \text{ und insbesondere } (0,0)=0,\ (1,0)=1.$$

In diesem Sinne ist $\mathbb{C}$ *eine Erweiterung von* $\mathbb{R}$*!*

c) *Für jedes* $z=(x,y)\in\mathbb{C}$ *gilt die Darstellung*

$$z=(x,0)+(0,y)=(x,0)+(y,0)\cdot(0,1).$$

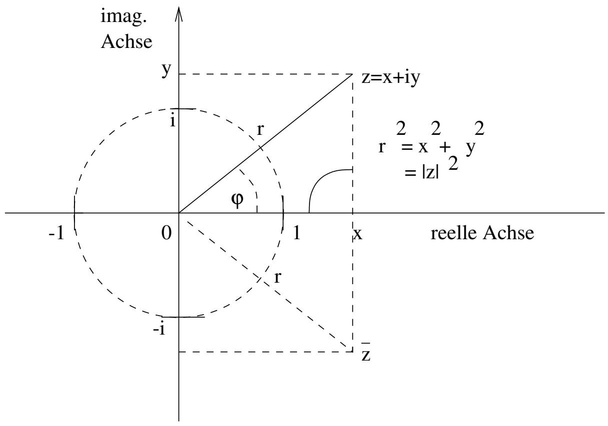
Bezeichnung: $i := (0,1)$ *(Euler 1777; „imaginäre Einheit")*

$$\rightsquigarrow z=(x,0)+(y,0)i=x+iy$$

wobei das zweite Gleichheitszeichen wie in b) zu verstehen ist.
Dann heißt x *Realteil von* $z\in\mathbb{C}$ *und*
y *Imaginärteil von* $z\in\mathbb{C}$.
Schreibweise: $\mathrm{Re}\ z := x,\ \mathrm{Im}\ z := y.\ \rightsquigarrow z = Re\ z + i\ Im\ \ z,\ \forall z\in\mathbb{C}.$
$|z|=|(x,y)|=(x^2+y^2)^{\frac{1}{2}}=(\mathrm{Re}\ z)^2+(\mathrm{Im}\ z)^2)^{\frac{1}{2}}$ *heißt absoluter Betrag von* $z\in\mathbb{C}$.
$\bar{z} := \mathrm{Re}\ z - i\,\mathrm{Im}\ z$ *heißt die zu* $z\in\mathbb{C}$ *konjugiert komplexe Zahl.*

d) Die komplexen Zahlen können wir als Punkte in der <u>*komplexen*</u> *(oder* <u>*Gaußschen)*</u> <u>*Zahlenebene*</u> *auffassen:*

$$\begin{aligned} z &= x+iy = |z|\left(\tfrac{x}{(x^2+y^2)^{\frac{1}{2}}} + \tfrac{y}{(x^2+y^2)^{\frac{1}{2}}}\right) \\ &= |z|(\cos\varphi + i\sin\varphi) = r(\cos\varphi + i\sin\varphi) \end{aligned}$$



Letzteres nennt man auch die Darstellung von z in <u>*Polarkoordinaten*</u> (r, φ).

Eigenschaften 2.29

1) $i^2 = -1, \quad i^3 = -i, \quad i^4 = 1.$

2) Es gilt: $\mathrm{Re}\ z = \frac{1}{2}(z+\bar{z})$ *und* $\mathrm{Im}\ z = \frac{1}{2i}(z-\bar{z}),\ \forall z \in \mathbb{C}$.
Insbesondere ist $z \in \mathbb{R}$ *gdw.* $z = \bar{z}$.

3) $\overline{z_1+z_2} = \bar{z}_1 + \bar{z}_2, \overline{z_1 z_2} = \bar{z}_1\bar{z}_2, z = \bar{\bar{z}},\ \forall z, z_1, z_2 \in \mathbb{C}.$

4) $|z| \geq 0,\ \forall z \in \mathbb{C}.$ *und* $|z| = 0$ *gdw.* $z = 0; |z|^2 = z\bar{z},\ \forall z \in \mathbb{C}$
$|z_1 z_2| = |z_1||z_2|; |z_1+z_2| \leq |z_1| + |z_2|,\ \forall z_1, z_2 \in \mathbb{C}.$

5) $\mathrm{Re}\left(\frac{z_1}{z_2}\right) = (\mathrm{Re}\ z_1 \mathrm{Re}\ z_2 + \mathrm{Im}\ z_1 \mathrm{Im}\ z_2)/((\mathrm{Re}\ z_2)^2 + (\mathrm{Im} z_2)^2),$
Im $\left(\frac{z_1}{z_2}\right) = (\mathrm{Re}\ z_2 \mathrm{Im}\ z_1 - \mathrm{Re}\ z_1 \mathrm{Im}\ z_2)/((\mathrm{Re}\ z_2)^2 + (\mathrm{Im}\ z_2)^2),\ \forall z_1, z_2 \in \mathbb{C}$

<u>Beweis:</u>

1) $i^2 = i \cdot i = (0,1)\cdot(0,1) = (0-1, 0+0) = (-1,0) = -1$ usw.

2) Gilt wegen $z + \bar{z} = 2\mathrm{Re}\ z$ und $z - \bar{z} = 2i\ \mathrm{Im}\ z$. Deshalb ist $z \in \mathbb{C}$ reell gdw. $\mathrm{Im}\ z = 0$ gdw. $z = \bar{z}$.

3) Wir beweisen als Beispiel $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \ \forall z_1, z_2 \in \mathbb{C}$.
Es sei $z_i = (x_i, y_i), i = 1, 2$.
$\rightsquigarrow \overline{z_1 z_2} = \overline{(x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1)} = (x_1 x_2 - y_1 y_2, -x_1 y_2 - x_2 y_1)$
$\bar{z}_1 \bar{z}_2 = (x_1, -y_1)(x_2, -y_2) = (x_1 x_2 - y_1 y_2, -x_1 y_2 - x_2 y_1)$.

4) Der erste Teil folgt aus der Definition des absoluten Betrages. Ferner gilt: $z\bar{z} = ((\text{Re } z)^2 + (\text{Im } z)^2, 0), \ \forall z \in \mathbb{C}$. Wir beweisen als Beispiel außerdem die Dreiecksungleichung. Es seien $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in \mathbb{C}$ und $\alpha, \beta \in \mathbb{R}$ bel. gewählt. Dann gilt:

$$\begin{aligned} 0 &\leq |\alpha z_1 + \beta z_2|^2 = |(\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2)|^2 \\ &= (\alpha x_1 + \beta x_2)^2 + (\alpha y_1 + \beta y_2)^2 \\ &= \alpha^2 (x_1^2 + y_1^2) + 2\alpha\beta(x_1 x_2 + y_1 y_2) + \beta^2 (x_2^2 + y_2^2) \\ &= \alpha^2 |z_1|^2 + 2\alpha\beta(x_1 x_2 + y_1 y_2) + \beta^2 |z_2|^2 \end{aligned}$$

Für $\alpha = \beta = 1$ sieht man daraus, daß die Dreiecksungleichung richtig ist, falls:

$$2(x_1 x_2 + y_1 y_2) \leq |z_1||z_2|.$$

Diese Ungleichung folgt aber auch aus der obigen Ungleichung, indem man

$$\alpha := -\frac{x_1 x_2 + y_1 y_2}{|z_1|}, \beta := |z_1| \text{ (o. B. d. A. } z_1 \neq 0)$$

setzt. Dann ergibt sich nämlich

$$0 \leq (x_1 x_2 + y_1 y_2)^2 - 2(x_1 x_2 + y_1 y_2)^2 + |z_1|^2 |z_2|^2.$$

5) Es gilt

$$\frac{z_1}{z_2} = \frac{z_1 \bar{z}_2}{|z_2|^2} = \frac{z_1 \bar{z}_2}{((\text{Re } z_2)^2 + (\text{Im } z_2)^2} \text{ usw.}$$

□

Bemerkung 2.30
Wegen 2.29 1) kann man nun Potenzen a^r für $a < 0$ und $r \in \mathbb{Q}$ wie folgt definieren:

$$a^{\frac{m}{n}} := \sqrt[n]{(-a)^m} \sqrt[n]{(-1)^m} = \sqrt[n]{(-a)^m} \cdot \left\{ \begin{array}{ll} 1 & , \ m \text{ gerade} \\ -1 & , \ m \text{ ungerade}, n \text{ ungerade} \\ i^{\frac{n}{2}} & , \ m \text{ ungerade}, n \text{ gerade} \end{array} \right\}$$

2.3 Der m-dimensionale Euklidische Raum

Für jede Zahl $m \in I\!N$ betrachten wir die Menge

$$\mathbb{R}^m := \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{m} = \{(x_1, \ldots, x_m) : \ x_i \in \mathbb{R}, \ i = 1, \ldots, m\}$$

und definieren für Elemente dieser Menge eine Addition „+“ und eine Multiplikation „·“ mit reellen Zahlen wie folgt:
Es seien $x = (x_1, \ldots, x_m) \in \mathbb{R}^m, \ y = (y_1, \ldots, y_m) \in \mathbb{R}^m, \alpha \in \mathbb{R}$.

$$x + y := (x_1 + y_1, \ldots, x_m + y_m), \quad \alpha \cdot x := (\alpha x_1, \ldots, \alpha x_m)$$

Eigenschaften 2.31
1) $(\mathbb{R}^m, +)$ *ist eine Abelsche Gruppe mit Nullelement* $\Theta = (0, \ldots, 0)$;
2) $1 \cdot x = x, \forall x \in \mathbb{R}^m$;
3) $(\alpha + \beta)x = \alpha x + \beta y, \alpha(x + y) = \alpha x + \alpha y, \alpha(\beta x) = (\alpha\beta)x, \ \forall x, y \in \mathbb{R}^m, \ \forall \alpha, \beta \in \mathbb{R}$.

Definition 2.32

a) $(\mathbb{R}^m; +, \cdot)$ *heißt m-dimensionaler linearer Raum (oder: Vektorraum)*
Kurzbezeichnung: $\mathbb{R}^m$

b) Die Abbildung $\langle \cdot, \cdot \rangle : \mathbb{R}^m \times \mathbb{R}^m \longrightarrow \mathbb{R}$ *definiert durch*

$$\langle x, y \rangle := \sum_{i=1}^{m} x_i y_i, \ \forall x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m) \in \mathbb{R}^m,$$

heißt Skalarprodukt auf $\mathbb{R}^m$.
Die Abbildung $\| \cdot \| : \mathbb{R}^m \longrightarrow \mathbb{R}$ *definiert durch* $\|x\| := \langle x, x \rangle^{\frac{1}{2}}, \forall x \in \mathbb{R}^m$ *heißt Euklidische Norm auf* $\mathbb{R}^m$.

c) $(\mathbb{R}^m, \| \cdot \|)$ *heißt m-dimensionaler Euklidischer Raum.*

Bemerkung 2.33 *Später werden wir jede Abelsche Gruppe* $(X, +)$*, für die eine Operation („Multiplikation“) mit Elementen eines Körpers* K *erklärt ist, die die Eigenschaften 2.34 erfüllen, einen linearen Raum (über* K*) nennen.*
Die Zahl $\|x - y\|$ *heißt Euklidischer Abstand von* $x, y \in \mathbb{R}^m$.
Die Elemente $e_i := (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^m$ *(i-te Komponente gleich 1),* $i = 1, \ldots, m$*, heißen Einheitsvektoren von* $\mathbb{R}^m$*. Es gilt:*

$$x = \sum_{i=1}^{m} x_i e_i, \quad \forall x = (x_1, \ldots, x_m) \in \mathbb{R}^m.$$

Satz 2.34 *Für alle* $x, y \in \mathbb{R}^m$ *gilt die Cauchy-Schwarzsche Ungleichung:*

$$|\langle x, y \rangle| \leq \|x\| \|y\| \ .$$

Für zwei Elemente $x, y \in \mathbb{R}^m$ *gilt in der Ungleichung die Gleichheit gdw.* $\exists\, \alpha, \beta \in \mathbb{R}$ *mit* $(\alpha, \beta) \neq (0, 0)$ *und* $\alpha x + \beta y = \Theta$.

Beweis:
Im Fall $\|x\| = 0$ ist die Ungleichung richtig. Es gelte nun $\|x\| \neq 0$.
Dann gilt für beliebige $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned} 0 \leq \|\alpha x + \beta y\|^2 &= \sum_{i=1}^{m} (\alpha x_i + \beta y_i)^2 \\ &= \alpha^2 \sum_{i=1}^{m} x_i^2 + 2\alpha\beta \sum_{i=1}^{m} x_i y_i + \beta^2 \sum_{i=1}^{m} y_i^2 \\ &= \alpha^2 \|x\|^2 + 2\alpha\beta \langle x, y\rangle + \beta^2 \|y\|^2 \end{aligned}$$

Wir wählen nun speziell

$$\alpha = -\frac{\langle x, y\rangle}{\|x\|} \text{ und } \beta := \|x\|$$

und erhalten $0 \leq \langle x, y\rangle^2 - 2\langle x, y\rangle^2 + \|x\|^2\|y\|^2$, d.h., die behauptete Ungleichung.
Es seien nun $x, y \in \mathbb{R}^m$, so daß $\langle x, y\rangle^2 = \|x\|^2\|y\|^2$.
Eine erste Möglichkeit ist $x = \Theta$, d. h., $x + 0y = \Theta$.
Im Fall $x \neq \Theta$ schlußfolgern wir aus unseren obigen Überlegungen, daß

$$0 = \left\| -\frac{\langle x, y\rangle}{\|x\|} x + \|x\| y \right\|^2 \quad \text{oder} \quad \Theta = -\frac{\langle x, y\rangle}{\|x\|} x + \|x\| y$$

Also: $\alpha x + \beta y = \Theta$ mit $\alpha := -\dfrac{\langle x, y\rangle}{\|x\|} = -\|y\|$ und $\beta := \|x\|$. □

Satz 2.35 *Für alle $x, y \in \mathbb{R}^m$ und $\alpha \in \mathbb{R}$ gilt:*

(i) $\|x\| \geq 0$ *und* $\|x\| = 0$ *gwd.* $x = \Theta$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$;

(iii) $\|x + y\| \leq \|x\| + \|y\|$ *(Dreiecksungleichung);*

(iv) $|\|x\| - \|y\|| \leq \|x - y\|$.

Beweis:
(i) und (ii) sind klar nach Definition von $\|x\| = \left(\sum_{i=1}^{m} x_i^2\right)^{\frac{1}{2}}$.
Für beliebige $x, y \in \mathbb{R}^m$ gilt ferner (vgl. Beweis von 2.34)

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \quad (2.34!) \end{aligned}$$

d. h. die Aussage (iii). Aussage (iv) folgt aus den Ungleichungen

$$\|x\| \leq \|x - y\| + \|y\| \text{ und } \|y\| \leq \|y - x\| + \|x\|. \qquad \square$$

Wir beschäftigen uns nun mit der inneren Geometrie von $\mathbb{R}^m$.

Definition 2.36

a) *Für alle* $x \in \mathbb{R}^m$ *und* $r > 0$ *heißt* $B(x,r) := \{y \in \mathbb{R}^m : \|x-y\| \le r\}$ *(abgeschlossene)* *Kugel* *um* x *mit Radius* r.

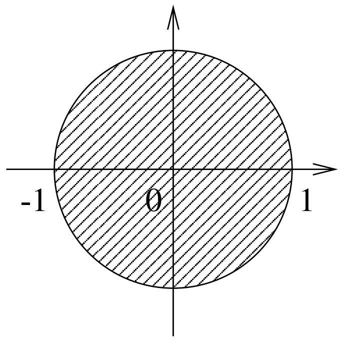
b) *Für jedes* $x \in \mathbb{R}^m$ *und jede Menge* $\emptyset \neq A \subseteq \mathbb{R}^m$ *heißt die reelle Zahl* $d(x,A) := \inf\limits_{y \in A} \|x-y\|$ *Abstand* *von* x *zu* A.

c) $A \subseteq \mathbb{R}^m$ *heißt* *beschränkt*, *falls eine Konstante* $C > 0$ *existiert, so daß* $\|x-y\| \le C$, $\forall x, y \in A$. *In diesem Fall heißt* $\operatorname{diam} A := \sup\limits_{x,y \in A} \|x-y\|$ *Durchmesser* *von* A.

Beispiel 2.37

a) *Für* $m = 1$ *gilt:*
$B(x,r) = [x-r, x+r]$ *und* $B\left(\frac{1}{2}(a+b), \frac{1}{2}(b-a)\right) = [a,b]$.

Für $m = 2$ *gilt:*
$B((0,0),1) = \{(y_1,y_2) \in \mathbb{R}^2 : y_1^2 + y_2^2 \le 1\}$ *(„Einheitskreisscheibe“).*



b) *Es gilt* $\operatorname{diam} B(x,r) = 2r, \quad \forall x \in \mathbb{R}^m \forall r > 0.$

c) *Für jede Teilmenge* $A \neq \emptyset$ *gilt:*
$|d(x,A) - d(y,A)| \le \|x-y\|$, $\forall\, x, y \in \mathbb{R}^m$ *(Übung).*

Definition 2.38 $A \subseteq \mathbb{R}^m$ *heißt* *offen*, *wenn zu jedem* $x \in A$ *ein* $\varepsilon > 0$ *ecxistiert mit* $B(x,\varepsilon) \subseteq A$. $A \subseteq \mathbb{R}^m$ *heißt* *abgeschlossen*, *wenn* $\mathbb{R}^m \backslash A$ *offen ist.*

Beispiel 2.39

a) $\emptyset$ *und* $\mathbb{R}^m$ *sind sowohl offen als auch abgeschlossen.*

b) *Die „offene“ Kugel* $\overset{\circ}{B}(x,r) := \{y \in \mathbb{R}^m : \|x-y\| < r\}$ *ist offen,* $\forall x \in \mathbb{R}^m \forall r > 0$.
Beweis: *Es sei* $y \in \overset{\circ}{B}(x,r)$ *und* $\varepsilon := \frac{1}{2}(r - \|x-y\|)$.
Wir zeigen: $B(y,\varepsilon) \subseteq \overset{\circ}{B}(x,r)$. *Sei*

$$\begin{aligned} z \in B(y,\varepsilon) \sim \|x-z\| &\le \|x-y\| + \|y-z\| \le \|x-y\| + \varepsilon \\ &= \tfrac{1}{2}(r + \|x-y\|) < r. \end{aligned}$$

$\square$

c) *Die „abgeschlossene“ Kugel* $B(x,r)$ *ist abgeschlossen,* $\forall\, x \in \mathbb{R}^m \forall\, r > 0$.
Beweis:
Wir zeigen: $\mathbb{R}^m \backslash B(x,r) = \{y \in \mathbb{R}^m : \|x-y\| > r\}$ *ist offen.*
Es sei $y \in \mathbb{R}^m \backslash B(x,r)$, *d. h.,* $\|x-y\| > r$.
Wir setzen $\varepsilon := \frac{1}{2}(\|x-y\| - r)$ *und zeigen* $B(y,\varepsilon) \subseteq \mathbb{R}^m \backslash B(x,r)$.
Sei $z \in B(y,\varepsilon)$. *Dann gilt* $\|x-y\| \leq \|x-z\| + \|z-y\|$
$\rightsquigarrow \|x-z\| \geq \|x-y\| - \|y-z\| \geq \|x-y\| - \varepsilon > \|x-y\| - 2\varepsilon = r$
$\rightsquigarrow z \in \mathbb{R}^m \backslash B(x,r)$. □

d) *Ist* $A \subseteq \mathbb{R}^m$ *offen, so gilt* $A = \bigcup_{x \in A} B(x, \varepsilon(x))$ *mit einem geeignet gewähltem* $\varepsilon(x) > 0$ *für jedes* $x \in A$.
Beweis:
Es sei $x \in A$ *bel. Dann existiert ein* $\varepsilon(x) > 0$ *mit der Eigenschaft* $B(x,\varepsilon(x)) \subseteq A$ *(da* A *offen). Also gilt* $\bigcup_{x \in A} B(x,\varepsilon(x)) \subseteq A$. *Die andere Inklusion ist trivial.* □

e) *Für* $m = 1$ *ist die Menge* $A := [0,1[$ *weder offen noch abgeschlossen (Übung).*

Satz 2.40 *Es sei* $(A_\lambda)_{\lambda \in L}$ *eine Familie von Teilmengen von* $\mathbb{R}^m$.

a) *Sind die Mengen* A_λ, $\forall \lambda \in L$, *offen bzw. abgeschlossen, so ist* $\bigcup_{\lambda \in L} A_\lambda$ *offen bzw.* $\bigcap_{\lambda \in L} A_\lambda$ *abgeschlossen.*

b) *Die Indexmenge* L *sei endlich. Sind die Mengen* A_λ, $\forall\, \lambda \in L$, *offen bzw. abgeschlossen, so ist auch* $\bigcap_{\lambda \in L} A_\lambda$ *offen bzw.* $\bigcup_{\lambda \in L} A_\lambda$ *abgeschlossen.*

Beweis:
Wir betrachten zunächst den Fall offener Mengen A_λ, $\forall \lambda \in L$.

a) Es sei $x \in \bigcup_{\lambda \in L} A_\lambda$. Dann existiert ein $\lambda_0 \in L$ mit $x \in A_{\lambda_0}$.
Da A_{λ_0} offen ist, existiert ein $\varepsilon > 0$ mit $B(x,\varepsilon) \subseteq A_{\lambda_0} \subseteq \bigcup_{\lambda \in L} A_\lambda$.
Also ist auch $\bigcup_{\lambda \in L} A_\lambda$ offen.

b) Es sei $x \in \bigcap_{\lambda \in L} A_\lambda \rightsquigarrow x \in A_\lambda,\ \forall\, \lambda \in L$.
Da alle A_λ offen sind, existiert für jedes $\lambda \in L$ ein $\varepsilon_\lambda > 0$ mit $B(x,\varepsilon_\lambda)$.
Wir definieren $\varepsilon := \min\{\varepsilon_\lambda : \lambda \in L\}$.
Da L endlich ist, gilt $\varepsilon > 0$ und $B(x,\varepsilon) \subseteq A_\lambda,\ \forall\, \lambda \in L$.
Also gilt auch $B(x,\varepsilon) \subseteq \bigcup_{\lambda \in L} A_\lambda$.

Im Fall, daß alle Mengen A_λ, $\forall\, \lambda \in L$, abgeschlossen sind, verwendet man die Morgan'sche Regel (vgl. Kap. 1.2) und z. B. in Teil b) die Aussage

$$\mathbb{R}^m \backslash \left(\bigcup_{\lambda \in L} A_\lambda \right) = \bigcap_{\lambda \in L} (\mathbb{R}^m \backslash A_\lambda) \qquad \text{(Übung)}.$$

□

Beispiel 2.41 *Wir zeigen, daß Satz 2.40b) i. a. nicht für unendliche Indexmengen L richtig ist. Es gilt:*

$$\bigcap_{n \in I\!N} \{x \in \mathbb{R} : x > -\frac{1}{n}\} = \{x \in \mathbb{R} : x \geq 0\}$$

Die Mengen auf der linken Seite der Gleichung sind offene Mengen und die Menge rechts ist abgeschlossen.

Definition 2.42 *Es sei A eine Teilmenge von $\mathbb{R}^m$.*

a) *$x \in \mathbb{R}^m$ heißt Häufungspunkt von A, falls für jedes $\varepsilon > 0$ gilt:*

$$A \cap B(x, \varepsilon) \backslash \{x\} \neq \emptyset.$$

$x \in A$ heißt isolierter Punkt von A, falls x kein Häufungspunkt von A ist.

b) *$x \in \mathbb{R}^m$ heißt innerer (bzw. äußerer) Punkt von $A \subseteq \mathbb{R}^m$, falls ein $\varepsilon > 0$ existiert, so daß $B(x, \varepsilon) \subseteq A$ (bzw. $B(x, \varepsilon) \cap A = \emptyset$).*
$x \in \mathbb{R}^m$ heißt Randpunkt von A, falls für jedes $\varepsilon > 0$ gilt:

$$B(x, \varepsilon) \cap A \neq \emptyset \quad \text{und} \quad B(x, \varepsilon) \cap (\mathbb{R}^m \backslash A) \neq \emptyset.$$

Beispiel 2.43

a) *$m = 1, A := \{\frac{1}{n} : n \in I\!N\}$. Nach Satz 2.7 gilt:*
$B(0, \varepsilon) \cap A \backslash \{0\} = [-\varepsilon, \varepsilon] \cap A \backslash \{0\} \neq \emptyset, \ \forall \varepsilon > 0$.
Also ist 0 Häufungspunkt von A.

b) *$m = 1, A := \mathbb{Q}$. Nach Satz 2.9 gilt: $\forall x \in \mathbb{R} \forall \varepsilon > 0 \exists r \in \mathbb{Q} : r \in]x - \varepsilon, x + \varepsilon[$.*
Insbesondere bedeutet das $B(x, \varepsilon) \cap \mathbb{Q} \backslash \{x\} \neq \emptyset, \ \forall \varepsilon > 0$.
Also ist jede relle Zahl Häufungspunkt von $\mathbb{Q}$.

c) *Endliche Teilmengen A von $\mathbb{R}^m$ enthalten nur isolierte Punkte.*
Beweis:
Sei $A = \{x_1, \ldots, x_n\}$ und $\varepsilon := \frac{1}{2} \min_{i,j=1,\ldots,n} \|x_i - x_j\| > 0$.
Dann gilt $B(x_i, \varepsilon) \cap A \backslash \{x_i\} = \emptyset), \forall i = 1, \ldots, n$, d. h., x_i ist kein Häufungspunkt von A, $\forall \, i = 1, \ldots, n$. □

d) *Es sei $x \notin A$. x ist Häufungspunkt von A gdw. x ist Randpunkt.*
Beweis:
Es sei $x \notin A \subseteq \mathbb{R}^m$ und es sei x Häufungspunkt von A. Dann:
$A \cap B(x, \varepsilon) \backslash \{x\} = A \cap B(x, \varepsilon) \neq \emptyset$ und $x \in B(x, \varepsilon) \cap (\mathbb{R}^m \backslash A)$.
Deshalb ist x Randpunkt von A.
Es sei nun x ein Randpunkt von A und es sei $\varepsilon > 0$ bel.
$\rightsquigarrow B(x, \varepsilon) \cap A \backslash \{x\} = B(x, \varepsilon) \cap A \neq \emptyset$
$\rightsquigarrow$ x ist Häufungspunkt von A. □

Definition 2.44 *Es sei A eine Teilmenge von $\mathbb{R}^m$.*

a) $A' := \{x \in \mathbb{R}^m : x \text{ ist Häufungspunkt von } A\}$,
$\overset{\circ}{A} := \mathrm{int}\,(A) := \{x \in A : x \text{ ist innerer Punkt von } A\}$.
$\bar{A} := \mathrm{cl}\,(A) := \bigcap\limits_{\substack{A \subseteq D \subseteq \mathbb{R}^m \\ D \text{ abgeschlossen}}} D$ *heißt* *Abschließung* *von A.*
$\mathrm{bd}(A) := \{x \in \mathbb{R}^m : x \text{ ist Randpunkt von } A\}$ *heißt* *Randmenge* *(Rand) von A.*
Man nennt $\overset{\circ}{A}$ *auch „Inneres" von A und es bedeutet* int=*interior*, cl=*closure*, bd=*boundary.*

b) *A heißt dicht bzgl. einer Menge* $B \subseteq \mathbb{R}^m$ *falls* $B \subseteq \bar{A}$.
A heißt (überall) dicht, falls $\bar{A} = \mathbb{R}^m$.

Satz 2.45 *Es sei A eine Teilmenge von* $\mathbb{R}^m$*. Dann gilt:*

a) $\bar{A} = A \cup A'$.

b) *A ist abgeschlossen (bzw. offen) gdw.* $A = \bar{A}$ *(bzw.* $A = \overset{\circ}{A}$*).*

Beweis:

a) Zunächst ist klar, daß $A \subseteq \bar{A}$ nach Definition 2.44, da $\bar{A}$ nach Satz 2.40 abgeschlossen ist. Wir zeigen: $A' \subseteq \bar{A}$.
Annahme: $\exists x \in A'$ mit $x \notin \bar{A}$.
Da $\mathbb{R}^m \backslash \bar{A}$ offen ist, existiert $\varepsilon > 0$ mit $B(x,\varepsilon) \subseteq \mathbb{R}^m \setminus \bar{A}$.
$\rightsquigarrow \bar{A} \cap B(x,\varepsilon) \setminus \{x\} = \bar{A} \cap B(x,\varepsilon) = \emptyset \ \rightsquigarrow x \notin A' \rightsquigarrow$ Widerspruch.
Also gilt: $A \cup A' \subseteq \bar{A}$.
Für den Beweis der umgekehrten Inklusion sei $x \in \bar{A} \setminus A$.
Annahme: x ist kein Häufungspunkt von A.
Dann existiert ein $\varepsilon_0 > 0$, so daß $A \cap B(x,\varepsilon_0) \setminus \{x\} = \emptyset$.
$\rightsquigarrow A \cap \overset{\circ}{B}(x,\varepsilon_0) = \emptyset$ (da $x \notin A$), d.h., $A \subseteq \bar{A} \setminus \overset{\circ}{B}(x,\varepsilon_0)$.
Da die Menge $\bar{A} \setminus \overset{\circ}{B}(x,\varepsilon_0)$ abgeschlossen ist, gilt $\bar{A} \subseteq \bar{A} \setminus \overset{\circ}{B}(x,\varepsilon_0)$.
$\rightsquigarrow$ Widerspruch zu $x \in \bar{A}$. Also muß $x \in A'$ gelten.

b) Ist A abgeschlossen so gilt $\bar{A} \subseteq A$. Die Inklusion $A \subseteq \bar{A}$ ist aber klar. Gilt $A = \bar{A}$, so ist A abgeschlossen, da $\bar{A}$ abgeschlossen ist.
Ist A offen, so gilt $A \subseteq \overset{\circ}{A}$, andererseits ist die Inklusion $\overset{\circ}{A} \subseteq A$ klar. Gilt umgekehrt $A = \overset{\circ}{A}$, so ist A offen, da $\overset{\circ}{A}$ offen ist. □

Beispiel 2.46

a) $m = 1, A := \left\{\frac{1}{n} : n \in I\!N\right\}$
$\rightsquigarrow A' = \{0\}$, $\bar{A} = \{0, \frac{1}{n} : n \in I\!N\}$, $\overset{\circ}{A} = \emptyset$, $\mathrm{bd}(A) = \{0, \frac{1}{n} : n \in I\!N\}$ *(vgl. 2.43).*

b) $m = 1, A :=]0,1[\rightsquigarrow A' = [0,1], \bar{A} = [0,1], \overset{\circ}{A} = A, \mathrm{bd}(A) = \{0,1\}$

c) $m = 1, A := \mathbb{Q} \rightsquigarrow \mathbb{Q}' = \mathbb{R} = \bar{\mathbb{Q}}$, *d. h.,* $\mathbb{Q}$ *ist dicht,* $\overset{\circ}{\mathbb{Q}} = \emptyset$ *und* $\mathrm{bd}(\mathbb{Q}) = \mathbb{R}$.

d) $\mathrm{cl}\overset{\circ}{B}(x,r) = B(x,r), \forall x \in \mathbb{R}^m \forall r > 0.$
(„Die Abschließung der offenen Kugel ist gerade die abgeschlossene Kugel“.)
Beweis:
Die Inklusion „$\subseteq$ “ ist klar nach Definition der Abschließung und da $B(x,r)$ abgeschlossen ist (vgl. 2.39).
Es sei nun $y \in B(x,r) \setminus \overset{\circ}{B}(x,r)$, d.h., $\|y\| = r$.
Wir zeigen: y ist Randpunkt von $B(x,r)$.
($\rightsquigarrow$ y ist Häufungspunkt von $\overset{\circ}{B}(x,r)$ (2.46) $\rightsquigarrow$ $y \in \mathrm{cl}\,\overset{\circ}{B}(x,r)$ (2.45)).
Sei $\varepsilon > 0$ bel. $\rightsquigarrow$ $B(y,\varepsilon) \cap \overset{\circ}{B}(x,r) \neq \emptyset$ und

$$B(y,\varepsilon) \cap (\mathbb{R}^m \setminus \overset{\circ}{B}(x,r)) \neq \emptyset,$$

(durch Betrachtung der Geraden durch x und y!) □

- *Es gilt:* $\mathrm{bd}(A) = (A' \setminus \overset{\circ}{A}) \cup \{x \in A : x \text{ ist isolierter Punkt von } A\}$ *(Übung)*

Alle Überlegungen zur „inneren Geometrie“ des $\mathbb{R}^m$ seit Definition 2.36 beruhten ausschließlich auf dem Begriff der Kugel und damit dem Begriff des Euklidischen Abstandes (vgl. 2.33). Wir kommen nun zum für die Mathematik zentralen Begriff der *Konvergenz einer Folge* (im $\mathbb{R}^m$).

Definition 2.47 *Es seien $(x_n)_{n\in\mathbb{N}}$ eine Folge von Elementen aus $\mathbb{R}^m$ und $x \in \mathbb{R}^m$. $(x_n)_{n\in\mathbb{N}}$ heißt konvergent gegen x, falls*

$$\forall \varepsilon > 0 \exists n_0(\varepsilon) \in \mathbb{N} : \|x_n - x\| < \varepsilon, \ \forall n \geq n_0(\varepsilon).$$

x heißt Grenzwert oder Limes der Folge $(x_n)_{n\in\mathbb{N}}$. $(x_n)_{n\in\mathbb{N}}$ heißt konvergent, falls ein $x \in \mathbb{R}^m$ existiert, so daß $(x_n)_{n\in\mathbb{N}}$ gegen x konvergent ist. Schreibweise:

$$x = \lim_{n\to\infty} x_n, \quad x_n \xrightarrow[n\to\infty]{} x$$

Beispiel 2.48

a) $m := 1, x_n := \frac{1}{n}, \ \forall n \in \mathbb{N}.$
Nach Satz 2.9 gilt: $\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} : \frac{1}{n_0} < \varepsilon$. Daraus folgt:

$$\frac{1}{n} \leq \frac{1}{n_0} < \varepsilon, \quad \forall n \geq n_0.$$

Deshalb konvergiert $(\frac{1}{n})_{n\in\mathbb{N}}$ gegen Null.

b) $m := 1, r \in]-1,1[$. *Dann gilt:* $\lim\limits_{n\to\infty} r^n = 0.$
Beweis: Für $r = 0$ ist die Aussage klar. Es sei $r \neq 0$ und es sei $\varepsilon > 0$ beliebig gewählt. Wir setzen $a := \frac{1}{|r|}(1-|r|) > 0$ und wählen $n_0(\varepsilon) \in \mathbb{N}$, so daß $\frac{1}{a\varepsilon} < n_0(\varepsilon)$ (Satz 2.7). Dann gilt $\forall n \geq n_0(\varepsilon)$ mit Hilfe von Satz 2.10:

$$|r^n - 0| = |r^n| = |r|^n = \left(\frac{1}{1+a}\right)^n = \frac{1}{(1+a)^n} \leq \frac{1}{1+na} \leq \frac{1}{na} \leq \frac{1}{n_0(\varepsilon)a} < \varepsilon. \square$$

c) $m := 1$. Die beiden folgenden Folgen sind <u>*nicht konvergent*</u> *(in $\mathbb{R}$):*

$$(n)_{n\in I\!N} \text{ und } (x_n)_{n\in I\!N} \text{ mit } x_n := \begin{cases} 0, & n \text{ gerade} \\ 1, & n \text{ ungerade} \end{cases}$$

<u>*Beweis:*</u>
$(x_n)_{n\in I\!N}$ ist nicht konvergent bedeutet:

$$\forall\, x \in \mathbb{R} \exists \varepsilon_0 > 0 \forall\, n \in I\!N \exists k(n) \geq n : \|x_{k(n)} - x\| \geq \varepsilon_0.$$

Wir betrachten zunächst die Folge $(n)_{n\in I\!N}$.
Es sei $x \in \mathbb{R}, \varepsilon_0 := 1$ und $k(n) := \max\{n, n_0 + 1\}$, wobei
$n_0 \in I\!N$ so gewählt ist, daß $|x| \leq n_0$.
$\rightsquigarrow \varepsilon_0 = 1 \leq k(n) - |x| \leq |k(n) - x|,\ \forall\, n \in I\!N$.
Für die zweite Folge wählen wir $\varepsilon_0 := \max\{|x|, |1 - x|\}$ und

$$k(n) \quad := \quad \begin{cases} 2n+1 & , \quad \text{falls } \varepsilon_0 = |1-x| \\ 2n & , \quad \text{falls } \varepsilon_0 = |x| \end{cases} \quad , \forall\, n \in I\!N.$$

$$\rightsquigarrow \varepsilon_0 \quad = \quad |x_{k(n)} - x|,\ \forall\, n \in I\!N. \qquad \square$$

d) Ist $(x_n)_{n\in I\!N}$ konvergent gegen $x \in \mathbb{R}^m$, so konvergiert die Folge $(\|x_n\|)_{n\in I\!N}$ der Normen der Folgenelemente in $\mathbb{R}$ gegen $\|x\|$.
(Der Beweis folgt aus der Ungleichung $|\|x_n\| - \|x\|| \leq \|x_n - x\|$ (vgl. Satz 2.35).)

Lemma 2.49
Jede in $\mathbb{R}^m$ konvergente Folge besitzt einen eindeutig bestimmten Grenzwert.

<u>Beweis:</u>
Annahme: Es existiert eine Folge $(x_n)_{n\in I\!N}$ in $\mathbb{R}^m$ und Elemente $x, \tilde{x} \in \mathbb{R}^m$ mit $x \neq \tilde{x}$ und $x = \lim\limits_{n\to\infty} x_n, \tilde{x} = \lim\limits_{n\to\infty} x_n$.
Wir wählen $\varepsilon > 0$ nun so, daß $2\varepsilon < \|x - \tilde{x}\|$ gilt. Wegen obiger Annahme existieren Zahlen $n_0(\varepsilon) \in I\!N$ und $\tilde{n}_0(\varepsilon) \in I\!N$, so daß

$$\|x - x_n\| < \varepsilon, \forall\, n \geq n_0(\varepsilon), \text{ und } \|\tilde{x} - x_n\| < \varepsilon, \forall\, n \geq \tilde{n}_0(\varepsilon).$$

Dann gilt für jedes $n \geq \max\{n_0(\varepsilon), \tilde{n}_0(\varepsilon)\}$:

$$2\varepsilon < \|x - \tilde{x}\| \leq \|x - x_n\| + \|x_n - \tilde{x}\| < \varepsilon + \varepsilon = 2\varepsilon \rightsquigarrow \text{ Widerspruch!} \quad \square$$

Der Konvergenzbegriff erlaubt nun eine neue Charakterisierung der Menge aller Häufungspunkte und der Abschließung einer Menge.

Satz 2.50 *Es sei A eine Teilmenge von $\mathbb{R}^m$.*

a) Für jedes $x \in A'$ existiert eine Folge $(x_n)_{n\in I\!N}$ von Elementen aus A, die gegen x konvergiert .

b) Ist $(x_n)_{n\in I\!N}$ eine konvergente Folge aus A, so gehört ihr Grenzwert zu $\bar{A}$.

Beweis:

a) Es sei $x \in A'$ und wir betrachten die Mengen

$$A_n := (A \cap B(x, \frac{1}{n})) \setminus \{x\} \neq \emptyset, \forall\, n \in I\!N.$$

Wir wählen sukzessive $x_1 \in A_1, x_n \in A_n, \forall\, n \geq 2$, und erhalten

$$\|x - x_n\| < \frac{1}{n} \text{ und } x_n \in A,\ \forall\, n \in I\!N.$$

Aus Beispiel 2.48 a) folgt dann die Konvergenz von $(x_n)_{n \in I\!N}$ gegen x.

b) Es sei $(x_n)_{n \in I\!N}$ eine Folge von Elementen aus A, die gegen ein $x \in \mathbb{R}^m$ konvergiert.
Annahme: $x \notin \bar{A}$.
Da $\bar{A}$ abgeschlossen ist, ist $\mathbb{R}^m \backslash \bar{A}$ offen. Deshalb existiert ein $\varepsilon_0 > 0$, so daß $B(x, \varepsilon_0) \subseteq \mathbb{R}^m \backslash \bar{A}$.
$\rightsquigarrow x_n \notin B(x, \varepsilon_0)$ oder $\|x - x_n\| > \varepsilon_0$, $\forall\, n \in I\!N$.
$\rightsquigarrow (x_n)_{n \in I\!N}$ ist nicht gegen x konvergent $\rightsquigarrow$ Widerspruch! □

Satz 2.50 besagt, daß die Abschließung $\bar{A}$ die Menge A und alle Grenzwerte von konvergenten Folgen aus A enthält.

Bemerkung 2.51 *In Definition 1.9 wurde allgemein der Begriff einer Folge $(x_n)_{n \in I\!N}$ und einer Teilfolge eingeführt. Wir nehmen im folgenden stets an, daß Teilfolgen die Form haben: $(x_n)_{n \in \tilde{I\!N}}$, $\tilde{I\!N} \subseteq I\!N$ und $\tilde{I\!N}$ unendlich.*
Da $I\!N$ und $\tilde{I\!N}$ beide abzählbar sind, existiert eine bijektive Abbildung $k : I\!N \longrightarrow \tilde{I\!N}$, die die folgende Schreibweise einer Teilfolge ermöglicht:

$$(x_{k(n)})_{n \in I\!N}$$

Dabei nehmen wir o.B.d.A. an, daß $k(1) < k(2) < \cdots < k(n+1) < \cdots$ gilt.
Eine Teilfolge $(x_n)_{n \in \tilde{I\!N}}$ von $(x_n)_{n \in I\!N}$ ist dementsprechend konvergent gegen ein Element $x \in \mathbb{R}^m$, falls
$\forall \varepsilon > 0 \exists n_0(\varepsilon) I\!N\ \forall\, n \geq n_0(\varepsilon) : \|x - x_{k(n)}\| < \varepsilon$ *oder*
$\forall \varepsilon > 0 \exists n_0(\varepsilon) I\!N\ \forall\, n \in \tilde{I\!N}, n \geq n_0(\varepsilon) : \|x - x_n\| < \varepsilon.$

Schreibweise: $x_{k(n)} \xrightarrow[n \to \infty]{} x$.

Satz 2.52 *Eine Folge $(x_n)_{n \in I\!N}$ in $\mathbb{R}^m$ ist gegen $x \in \mathbb{R}^m$ konvergent gdw. jede ihrer Teilfolgen gegen x konvergiert.*

Beweis:

($\Rightarrow$) Ist (x_n) gegen x konvergent, so gilt:
$\forall \varepsilon > 0 \exists n_0(\varepsilon) \in I\!N \forall n \geq n_0(\varepsilon) : \|x - x_n\| < \varepsilon$.
Ist nun $\tilde{I\!N}$ eine bel. unendliche Teilmenge von $I\!N$, so folgt erst recht, daß
$\forall \varepsilon > 0 \exists n_0(\varepsilon) \in I\!N \forall n \in \tilde{I\!N}, n \geq n_0(\varepsilon) : \|x - x_n\| < \varepsilon$.
Deshalb ist die Teilfolge $(x_n)_{n \in \tilde{I\!N}}$ konvergent.

($\Leftarrow$) Annahme: Die Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert nicht gegen x, d. h.,
$\exists\varepsilon_0 > 0 \forall n \in \mathbb{N} \exists k(n) \geq n : \|x - x_{k(n)}\| \geq \varepsilon_0$.
Dies bedeutet: Die Teilfolge $(x_{k(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ konvergiert nicht gegen x.
$\rightsquigarrow$ Widerspruch! $\square$

Satz 2.52 motiviert uns nun, nach Bedingungen zu suchen, die die Konvergenz von Teilfolgen einer Folge nach sich ziehen.

Definition 2.53
Es sei (x_n) eine Folge in $\mathbb{R}$. Diese heißt

- *(streng) monoton wachsend, falls $x_{n+1} \geq x_n$ $(x_{n+1} > x_n)$, $\forall n \in \mathbb{N}$;*
- *(streng) monoton fallend, falls $x_{n+1} \leq x_n$ $(x_{n+1} < x_n)$, $\forall n \in \mathbb{N}$;*
- *monoton, falls sie monoton wachsend oder monoton fallend ist;*
- *Nullfolge, falls $\lim\limits_{n\to\infty} x_n = 0$.*

Satz 2.54
Jede Folge in $\mathbb{R}$ enthält eine monotone Teilfolge.

Beweis:
Wir definieren für eine gegebene Folge (x_n) in $\mathbb{R}$ die Menge

$$M := \{m \in \mathbb{N} : x_n \geq x_m, \forall n > m\}.$$

(Beispiele: $(\frac{1}{n})_{n\in\mathbb{N}} \;\Rightarrow\; M = \emptyset$, $(1,2,1,4,1,6,1,8,1,\ldots) \;\Rightarrow\; M = \{1,3,5,7,\ldots\}$)
Wir unterscheiden nun die folgenden beiden Fälle:

(i) M ist nicht endlich:
Es seien $m(n) \in M$, $\forall n \in \mathbb{N}$, mit $m(1) < m(2) < m(3) < \ldots$
Dann gilt nach Definition von M:

$$x_{m(1)} \leq x_{m(2)} \leq x_{m(3)} \leq \ldots$$

Also ist die Teilfolge $(x_{m(n)})_{n\in\mathbb{N}}$ monoton wachsend.

(ii) M ist endlich:
Ist $M = \emptyset$, setzen wir formal $\max M = 0$. Andernfalls ist $\max M \in \mathbb{N}$.
$\rightsquigarrow m \notin M$, $\forall m > \max M$.
($m \notin M$ bedeutet: $\exists n(m) \in \mathbb{N}$, $n(m) > m$, so daß $x_{n(m)} < x_m$)
Wir konstruieren nun induktiv:

$$\begin{array}{lll} n(1) \in \mathbb{N}, & n(1) > \max M & \text{beliebig;} \\ n(k) \in \mathbb{N}, & n(k) > n(k-1), & \forall k \in \mathbb{N},\ k \geq 2, \text{ so daß} \\ & x_{n(k)} < x_{n(k-1)}. & \end{array}$$

Dann ist die Folge $(x_{n(k)})_{k\in\mathbb{N}}$ streng monoton fallend.
In beiden Fällen gelingt es also, eine monotone Teilfolge auszuwählen. $\square$

Satz 2.55
Eine monotone Folge (x_n) *in* $\mathbb{R}$ *ist konvergent* <u>*gdw.*</u> $\{x_n : n \in I\!N\}$ *beschränkt ist.*
Ist (x_n) *monoton wachsend, so gilt* $\lim\limits_{n\to\infty} x_n = \sup\{x_n : n \in I\!N\}$,
ist (x_n) *monoton fallend, so gilt* $\lim\limits_{n\to\infty} x_n = \inf\{x_n : n \in I\!N\}$.

<u>Beweis:</u>
Die eine Richtung ist trivial, da konvergente Folgen beschränkt sind (vgl. Satz 2.61).Für die andere Richtung sei (x_n) beschränkt und o.B.d.A. monoton wachsend.
Wir zeigen: (x_n) ist konvergent.
Nach Voraussetzung gilt $s := \sup\{x_n : n \in I\!N\} \in \mathbb{R}$. Sei $\varepsilon > 0$ beliebig gewählt. Nach Definition des Supremums existiert nun ein $n_o(\varepsilon) \in I\!N$, so daß $s - \varepsilon < x_{n_o} \leq s$.
Wegen der Monotonie der Folge gilt dann aber

$$s - \varepsilon < x_{n_o} \leq x_n \leq s, \quad \forall n \in I\!N, n \geq n_o,$$

$$\rightsquigarrow |s - x_n| < \varepsilon, \quad \forall n \geq n_o.$$

Also ist s Grenzwert der Folge (x_n). Für monoton fallende Folgen argumentiert man analog. □

Folgerung 2.56
Jede in $\mathbb{R}$ *beschränkte Folge* (x_n) *(d.h.* $\{x_n : n \in I\!N\}$ *ist beschränkt) besitzt eine konvergente Teilfolge.*

<u>Beweis:</u>
Die nach Satz 2.54 existierende monotone Teilfolge von (x_n) ist nach Satz 2.55 konvergent. □

Lemma 2.57
Eine Folge $(x_n)_{n\in I\!N}$ *in* $\mathbb{R}^m$ *ist konvergent* <u>*gdw.*</u> *ihre Komponentenfolgen* $(x_n^{(i)})$, $\forall i = 1, \ldots, m$, *konvergent sind.*
(Hierbei ist also $x_n = (x_n^{(1)}, \ldots, x_n^{(m)})$, $\forall n \in I\!N$.*)*

<u>Beweis:</u>
Eine Folge (x_n) konvergiert in $\mathbb{R}^m$ gegen $x \in \mathbb{R}^m$ <u>gdw.</u>

$$\forall \varepsilon > 0 \exists n_0(\varepsilon) \in I\!N \; \forall n \geq n_0(\varepsilon) : \; \|x_n - x\| = \sqrt{\sum_{i=1}^{m} |x_n^{(i)} - x^{(i)}|^2} < \varepsilon$$

Gilt $\|x_n - x\| < \varepsilon$, so auch $|x_n^{(i)} - x^{(i)}| < \varepsilon$, $\forall i = 1, \ldots, m$, d.h. $(x_n^{(i)})$ konvergiert gegen $x^{(i)}$, $\forall i = 1, \ldots, m$. Ist umgekehrt ein $\varepsilon > 0$ beliebig vorgegeben, so existiert für jedes $i \in \{1, \ldots, m\}$ ein $n_0^{(i)} \in I\!N$, so daß

$$|x_n^{(i)} - x^{(i)}| < \frac{\varepsilon}{\sqrt{m}}, \quad \forall n \geq n_0^{(i)}.$$

Folglich gilt für $n \geq \max\{n_0^{(i)} : i = 1, \ldots, m\}$, daß

$$\|x_n - x\|^2 = \sum_{i=1}^{m} |x_n^{(i)} - x^{(i)}|^2 < \sum_{i=1}^{m} \frac{\varepsilon^2}{m} = \varepsilon^2 \qquad \square$$

Definition 2.58 *Eine Menge $A \subseteq \mathbb{R}^m$ heißt relativ kompakt, falls jede Folge (x_n) in A eine konvergente Teilfolge besitzt.*
$A \subseteq \mathbb{R}^m$ heißt kompakt, falls A relativ kompakt und abgeschlossen ist.

Satz 2.59 *(Bolzano)*
Für eine Menge $A \subseteq \mathbb{R}^m$ sind folgende Aussagen äquivalent:

a) A ist beschränkt.

b) A ist relativ kompakt.

Beweis:

b) $\Rightarrow$ a) : $A \subseteq \mathbb{R}^m$ sei relativ kompakt und es sei $\varepsilon > 0$ bel. gewählt.
Wir wählen $x_1 \in A$ (o.B.d.A. $A \neq \emptyset$). Dann sind 2 Fälle möglich:
1. Fall: $A \subseteq B(x_1, \varepsilon) \rightsquigarrow n(\varepsilon) := 1$.
2. Fall: $\exists x_2 \in A$ mit $\|x_2 - x_1\| > \varepsilon$.
Jetzt gibt es wiederum 2 Fälle:
1. Fall: $A \subseteq B(x_1, \varepsilon) \cup B(x_2, \varepsilon) \rightsquigarrow n(\varepsilon) := 2$.
2. Fall: $\exists x_3 \in A \backslash (B(x_1, \varepsilon) \cup B(x_2, \varepsilon))$, d. h., $\|x_1 - x_3\| > \varepsilon$, $\|x_2 - x_3\| > \varepsilon$.
Diesen Prozeß setzen wir fort und erhalten im n-ten Schritt

$$x_n \in A \backslash \bigcup_{i=1}^{n-1} B(x_i, \varepsilon).$$

Annahme: Dieser Prozeß bricht nicht nach endlich vielen Schritten ab.
Dann haben wir eine Folge $(x_n)_{n \in \mathbb{N}}$ in A konstruiert, die die Eigenschaft besitzt

$$\|x_n - x_m\| > \varepsilon, \ \forall n, m \in \mathbb{N}, n \neq m.$$

Diese Folge kann keine konvergente Teilfolge besitzen. Dies ist aber ein Widerspruch zur relativen Kompaktheit von A! Also muß dieser Prozeß nach einer endlichen Anzahl $n(\varepsilon)$ von Schritten abbrechen!
Damit zeigen wir jetzt, daß $\sup\limits_{x,y \in A} \|x - y\| \in \mathbb{R}$ gilt. Es seien $x, y \in A$ bel. gewählt und wir setzen $\varepsilon := 1$. Nach unserer obigen Betrachtung existieren Punkte $\{x_1, \ldots, x_{n(\varepsilon)}\}$ und $i, j \in \{1, \ldots, n(\varepsilon)\}$ so daß $\|x - x_i\| \leq 1$ und $\|y - x_j\| \leq 1$.
Damit erhalten wir
$\|x - y\| \leq \|x - x_i\| + \|x_i - x_j\| + \|x_j - y\| \leq 2 + \max\limits_{i,j \in \{1,\ldots,n(\varepsilon)\}} \|x_i - x_j\|$.
Auf der rechten Seite steht aber eine fixe obere Schranke für $\|x - y\|$.

a) $\Rightarrow$ b): Zu zeigen: $A \subseteq \mathbb{R}^m$ ist beschränkt$\Rightarrow$ A ist relativ kompakt.
Es sei (x_n) eine Folge in A. Dann ist die Menge $\{x_n : n \in \mathbb{N}\}$ beschränkt in $\mathbb{R}^m$. Insbesondere existiert eine Konstante $C > 0$, daß $|x_n^{(i)}| \leq \|x_n\|_\infty \leq C$, $\forall n \in \mathbb{N}$, $\forall i = 1, \ldots, m$. Deshalb liegen auch die Folgen $(x_n^{(i)})$ für jedes $i = 1, \ldots, m$ in einer beschränkten Teilmenge von $\mathbb{R}$.
Wir wenden nun Folgerung 2.56 zunächst auf $(x_{n(1)})$ an und schlußfolgern, daß

$(x_n^{(1)})$ eine konvergente Teilfolge $(x_{k_1(n)}^{(1)})$ besitzt. Wir betrachten nun die Teilfolge $(x_{k_1(n)}^{(2)})$ von $(x_n^{(2)})$ und wenden darauf Folgerung 2.56 an. Wieder schlußfolgern wir die Existenz einer konvergenten Teilfolge $(x_{k_2(n)}^{(2)})$. Danach betrachten wir $(x_{k_2(n)}^{(3)})$ usw. Schließlich erhalten wir eine konvergente Teilfolge $(x_{k_m(n)}^{(m)})$. Nun ist aber die Folge $(x_{k_m(n)}^{(i)})$ eine konvergente Teilfolge von $(x_n^{(i)})$ für jedes $i = 1, \ldots, m$. Deshalb konvergiert nach Lemma 2.57 die Folge $(x_{k_m(n)})_{n\in I\!N}$ in $\mathbb{R}^m$ und (x_n) besitzt eine konvergente Teilfolge. □

Definition 2.60 *Eine Folge $(x_n)_{n\in I\!N}$ heißt Fundamentalfolge (oder Cauchy-Folge), falls* $\forall\, \varepsilon > 0 \exists n_0(\varepsilon) \in I\!N : \|x_n - x_m\| < \varepsilon,\ \forall\, n, m \geq n_0(\varepsilon)$.

Satz 2.61
Eine Folge $(x_n)_{n\in I\!N}$ in $\mathbb{R}^m$ ist konvergent gdw. sie eine Fundamentalfolge ist.
Überdies ist jede Fundamentalfolge beschränkt.

Beweis:

($\Rightarrow$) Es sei (x_n) konvergent gegen $x \in \mathbb{R}^m$ und $\varepsilon > 0$ sei bel. gewählt. Dann existiert ein $n_0(\frac{\varepsilon}{2}) \in I\!N$ mit $\|x_n - x\| < \frac{\varepsilon}{2},\ \forall\, n \geq n_0(\frac{\varepsilon}{2})$.
Daraus folgt für alle $m, n \geq n_0(\frac{\varepsilon}{2})$:

$$\|x_n - x_m\| \leq \|x_n - x\| + \|x - x_m\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

$\rightsquigarrow$ (x_n) ist eine Fundamentalfolge.

($\Leftarrow$) Es sei (x_n) eine Fundamentalfolge und wir setzen $\varepsilon := 1$.
Dann existiert ein $n_0 \in I\!N$ mit $\|x_n - x_m\| < \varepsilon,\ \forall\, n, m \geq n_0$.
Wir definieren

$$r := \max\left\{\varepsilon, \max_{i\in\{1,\ldots,n_0-1\}} \|x_i - x_{n_0}\|\right\}$$

und erhalten $x_n \in B(x_{n_0}, r),\ \forall\, n \in I\!N$. Also gehört die Folge (x_n) zu einer beschränkten Menge. Nach Satz 2.57 besitzt diese Folge deshalb eine konvergente Teilfolge $(x_n)_{n\in \tilde{I\!N}}, \tilde{I\!N} \subseteq I\!N$. Es bezeichne $x \in \mathbb{R}^m$ ihren Grenzwert.
Wir zeigen: Die gesamte Folge (x_n) konvergiert gegen x.
Es sei also $\varepsilon > 0$ bel. gewählt.
Dann existiert ein $n_0(\frac{\varepsilon}{2}) \in I\!N$ mit $\|x_n - x_m\| < \frac{\varepsilon}{2}, \forall\, n, m \geq n_0(\frac{\varepsilon}{2})$ und ein $\tilde{n}_0(\frac{\varepsilon}{2}) \in \tilde{I\!N}$ mit $\|x_n - x\| < \frac{\varepsilon}{2},\ \forall\, n \geq \tilde{n}_0(\frac{\varepsilon}{2}), n \in \tilde{I\!N}$.

Es seien nun $n, m \geq \max\left\{n_0\left(\frac{\varepsilon}{2}\right), \tilde{n}_0\left(\frac{\varepsilon}{2}\right)\right\}, n \in I\!N, m \in \tilde{I\!N}$. Dann gilt:

$$\|x_n - x\| \leq \|x_n - x_m\| + \|x_m - x\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Also ist $(x_n)_{n\in I\!N}$ gegen x konvergent. □

Bemerkung 2.62 *Auf $\mathbb{R}^m$ kann man noch viele weitere „Normen", d. h. Abbildungen von $\mathbb{R}^m$ in $\mathbb{R}_+$ mit den Eigenschaften (i)-(iii) von Satz 2.38 angeben.*
Wichtige Beispiele:

$$\|x\|_p := \left(\sum_{i=1}^{m} |x_i|^p\right)^{\frac{1}{p}}, p \geq 1,\ \|x\|_\infty := \max_{i=1,\ldots,m} |x_i|, \forall\, x = (x_1, \ldots, x_m) \in \mathbb{R}^m.$$

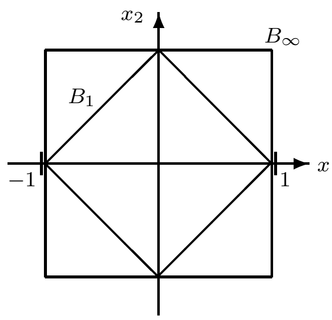
Die bisher betrachtete Euklidische Norm ist damit $\|\cdot\|_2$.
Die Eigenschaften (i) und (ii) sind für alle diese Normen einfach einzusehen, auch die Dreiecksungleichung für $\|\cdot\|_\infty$.
Übung: Man beweise die Dreiecksungleichung für $\|\cdot\|_p, p \geq 1$.

Definition 2.63 *Zwei Normen $\|\cdot\|$ und $\|\cdot\|_*$ auf $\mathbb{R}^m$ heißen äquivalent, falls Konstanten $C_1 > 0, C_2 > 0$ existieren, so daß*

$$C_1\|x\| \leq \|x\|_* \leq C_2\|x\|,\ \forall\, x \in \mathbb{R}^m.$$

Beispiel 2.64 *Es gilt $|x_j| \leq \|x\|_p,\ \forall\, j = 1, \ldots, m,\ \forall\, x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ und*

$$\|x\|_p \leq \left(m \cdot \max_{j=1,\ldots,m} |x_j|^p\right)^{\frac{1}{p}} = m^{\frac{1}{p}}\|x\|_\infty .$$



Also sind $\|\cdot\|_p$ und $\|\cdot\|_\infty$ äquivalent.
Es gilt: $\|x\|_1 = \sum_{i=1}^{m} |x_i| \leq m^{\frac{1}{2}}\|x\|_2$ (nach 2.37).
Was sind die Einheitskugeln $B_\infty((0,0),1)$ und $B_1((0,0),1)$ bzgl. $\|\cdot\|_\infty$ bzw. $\|\cdot\|_1$?
Antwort: siehe Abbildung

Bemerkung 2.65 *Für äquivalente Normen gilt, daß Folgen, die bzgl. der einen Norm konvergent sind, auch bzgl. der anderen gegen denselben Grenzwert konvergieren. Außerdem bleiben auch sämtliche Begriffe und Zusammenhänge der inneren Geometrie (wie offen, abgeschlossen, beschränkt, kompakt usw.) beim Übergang zwischen äquivalenten Normen erhalten !*
Wir vergewissern uns dessen im Fall der „Offenheit" von Mengen. Es seien also $\|\cdot\|$ und $\|\cdot\|_*$ zwei äquivalente Normen in $\mathbb{R}^m$.
Aussage: Eine Menge $A \subseteq \mathbb{R}^m$ ist offen bzgl. $\|\cdot\|$ gdw. A offen bzgl. $\|\cdot\|_*$ ist.
Beweis:

($\Rightarrow$) *Es sei A offen bzgl. $\|\cdot\|$ und es sei $x \in A$ beliebig gewählt.*
$\rightsquigarrow \exists \varepsilon > 0 : \{y : \|x-y\| \leq \varepsilon\} \subseteq A$
$\rightsquigarrow \{y : \frac{1}{C_1}\|x-y\|_* \leq \varepsilon\} \subseteq \{y : \|x-y\| \leq \varepsilon\} \subseteq A$
$\rightsquigarrow \{y : \|x-y\|_* \leq C_1\varepsilon\} \subseteq A$ *und A ist offen bzgl. $\|\cdot\|_*$.*

($\Leftarrow$) *Analog nutzt man die Inklusion $\{y : \|x-y\| \leq \frac{\varepsilon}{C_2}\} \subseteq \{y : \|x-y\|_* \leq \varepsilon\}$.* $\square$

Im Raum $\mathbb{R}^m$ gilt die folgende bemerkenswerte Aussage.

Satz 2.66 *Im Raum $\mathbb{R}^m$ sind je zwei beliebige Normen äquivalent.*

Beweis:
Es genügt zu zeigen, daß jede Norm auf $\mathbb{R}^m$ zur Euklidischen Norm $\|\cdot\|_2$ äquivalent ist. Es ist also $\|\cdot\|$ eine Norm auf $\mathbb{R}^m$. Für jedes $x \in \mathbb{R}^m$ gilt mit den kanonischen Einheitsvektoren e_j (vgl. 2.36) die Darstellung $x = \sum_{j=1}^m x_j e_j$. Es folgt:

$$\begin{aligned} \|x\| &\leq \sum_{j=1}^m \|x_j e_j\| = \sum_{j=1}^m |x_j|\, \|e_j\| \leq \left(\sum_{j=1}^m |x_j|^2\right)^{\frac{1}{2}} \left(\sum_{j=1}^m \|e_j\|^2\right)^{\frac{1}{2}} \\ &= \left(\sum_{j=1}^m \|e_j\|^2\right)^{\frac{1}{2}} \|x\|_2 \end{aligned}$$

Also gilt mit $C_2 := \left(\sum_{j=1}^m \|e_j\|^2\right)^{\frac{1}{2}}$, daß $\|x\| \leq C_2 \|x\|_2,\ \forall\, x \in \mathbb{R}^m$.
Jetzt betrachten wir die „Einheitssphäre“ $S := \{x \in \mathbb{R}^m : \|x\|_2 = 1\}$. Diese Menge ist beschränkt und abgeschlossen, also auch kompakt bzgl. $\|\cdot\|_2$.
Wir betrachten nun die Zahl $C_1 := \inf_{x \in S} \|x\| \geq 0$ und wir wählen eine Folge (x_n) in S mit $\|x_n\| \xrightarrow[n\to\infty]{} \inf_{x\in S} \|x\|$. Da S kompakt ist, existiert eine bzgl. $\|\cdot\|_2$ konvergente Teilfolge $(x_{k(n)})$ von (x_n). D.h. es existiert ein $x_* \in S$ mit

$$0 \leq \|x_{k(n)} - x_*\| \leq C_2 \|x_{k(n)} - x_*\|_2 \xrightarrow[n\to\infty]{} 0.$$

Deshalb gilt auch $\|x_{k(n)}\| \xrightarrow[n\to\infty]{} \|x_*\|$ und damit $C_1 = \|x_*\|$.
Folglich gilt $C_1 > 0$, da $x_* \neq \Theta$.
Ist nun $x \in \mathbb{R}^m \backslash \{\Theta\}$ bel. gewählt, so gilt $\frac{x}{\|x\|_2} \in S$ und folglich

$$\|\frac{x}{\|x\|_2}\| = \frac{1}{\|x\|_2}\|x\| \geq C_1, \text{ d. h., } C_1\|x\|_2 \leq \|x\|.$$

Für $x = \Theta$ gilt natürlich die letztere Ungleichung trivialerweise. □

Folglich sind die innere Geometrie und der Konvergenzbegriff in $\mathbb{R}^m$ unabhängig von der Wahl der konkreten Norm.
Schließlich gehen wir nochmals auf Folgen in $\mathbb{R}$ ein.

Satz 2.67 *(Rechenregeln für Grenzwerte)*
Es seien (x_n) und (y_n) in $\mathbb{R}$ konvergente Folgen (gegen $x \in \mathbb{R}$ bzw. $y \in \mathbb{R}$). Dann gilt:

a) *Für beliebige $\alpha, \beta \in \mathbb{R}$ konvergiert die Folge $(\alpha x_n + \beta y_n)$ gegen $\alpha x + \beta y$.*

b) *Die Folge $(x_n y_n)$ konvergiert gegen xy.*

c) *Falls $y \neq 0$, so existiert ein $n_o \in \mathbb{N}$ mit $y_n \neq 0$, $\forall n \geq n_o$, und die Folge* $\left(\frac{x_n}{y_n}\right)_{n \geq n_o}$ *konvergiert gegen* $\frac{x}{y}$.

d) $(|x_n|)$ *ist konvergent gegen* $|x|$.

<u>Beweis:</u>

a) Es gilt für alle $n \in \mathbb{N}$:

$$|\alpha x_n + \beta y_n - (\alpha x + \beta y)| \leq |\alpha||x_n - x| + |\beta||y_n - y|$$

Sei $\varepsilon > 0$ beliebig vorgegeben.
Nach Vor. existieren $n_o(\varepsilon) \in \mathbb{N}$ bzw. $\tilde{n}_o(\varepsilon) \in \mathbb{N}$, so daß

$$\begin{aligned} |x_n - x| &\leq \frac{\varepsilon}{1 + |\alpha| + |\beta|}, \quad \forall n \geq n_o(\varepsilon), \\ |y_n - y| &\leq \frac{\varepsilon}{1 + |\alpha| + |\beta|}, \quad \forall n \geq \tilde{n}_o(\varepsilon). \end{aligned}$$

Deshalb folgt für alle $n \geq \max\{n_o(\varepsilon), \tilde{n}_o(\varepsilon)\}$:

$$|\alpha x_n + \beta y_n - (\alpha x + \beta y)| \leq \frac{|\alpha| + |\beta|}{1 + |\alpha| + |\beta|}\varepsilon < \varepsilon.$$

b) Es gilt für alle $n \in \mathbb{N}$:

$$\begin{aligned} |x_n y_n - xy| &\leq |x_n y_n - x_n y| + |x_n y - xy| \\ &\leq |x_n||y_n - y| + |y||x_n - x| \end{aligned}$$

Da (x_n) konvergent ist, ist $\{x_n : n \in \mathbb{N}\} \subseteq \mathbb{R}$ beschränkt (Satz 2.59).
$\rightsquigarrow \exists C > 0 : |x_n| \leq C$, $\forall n \in \mathbb{N}$.
Sei $\varepsilon > 0$ beliebig vorgegeben.
Nach Vor. existieren $n_0(\varepsilon) \in \mathbb{N}$ bzw. $\tilde{n}_0(\varepsilon) \in \mathbb{N}$, so daß

$$\begin{aligned} |x_n - x| &< \frac{\varepsilon}{2C} \quad , \forall n \geq n_0(\varepsilon) \quad , \\ |y_n - y| &< \frac{\varepsilon}{2(1 + |y|)} \quad , \forall n \geq \tilde{n}_0(\varepsilon) \quad . \end{aligned}$$

$$\rightsquigarrow |x_n y_n - xy| < C\frac{\varepsilon}{2C} + |y|\frac{\varepsilon}{2(1 + |y|)} < \varepsilon, \ \forall n \geq \max\{n_0(\varepsilon), \tilde{n}_0(\varepsilon)\}.$$

c) Sei $y \neq 0$. Dann existiert ein $n_0 \in I\!\!N$ mit $|y| - |y_n| \leq |y - y_n| \leq \frac{|y|}{2}$ und folglich $\frac{|y|}{2} \leq |y_n|$ für alle $n \geq n_0$. Außerdem gilt für $n \geq n_0$:

$$\begin{aligned}
\left|\frac{x_n}{y_n} - \frac{x}{y}\right| = \left|\frac{x_n}{y_n} - \frac{x}{y_n} + \frac{x}{y_n} - \frac{x}{y}\right| &\leq \left|\frac{1}{y_n}\right| |x_n - x| + |x| \left|\frac{1}{y_n} - \frac{1}{y}\right| \\
&= \left|\frac{1}{y_n}\right| |x_n - x| + \frac{|x|}{|y_n||y|} \left|\frac{1}{y_n} - \frac{1}{y}\right| \\
&\leq \left|\frac{1}{y_n}\right| (|x_n - x| + \frac{|x|}{|y|}|y_n - y|) \\
&\leq \left|\frac{2}{y}\right| (|x_n - x| + \frac{|x|}{|y|}|y_n - y|)
\end{aligned}$$

Schließlich führt man den Beweis analog zu Teil a) zu Ende. □

Definition 2.68
Es sei (x_n) eine Folge in $\mathbb{R}$.

a) *(x_n) heißt divergent, wenn (x_n) nicht in $\mathbb{R}$ konvergiert.*

b) *(x_n) heißt bestimmt divergent (gegen $+\infty$ bzw. $-\infty$), falls für jedes $r \in \mathbb{R}$ ein $n_o(r) \in I\!\!N$ existiert, so daß $x_n \geq r$ bzw. $x_n \leq r$, $\forall n \geq n_o(r)$.*

Schreibweise: $\lim_{n\to\infty} x_n = +\infty$ *bzw.* $\lim_{n\to\infty} x_n = -\infty$,
$x_n \underset{n\to\infty}{\longrightarrow} +\infty$ *bzw.* $x_n \underset{n\to\infty}{\longrightarrow} -\infty$.

c) *Die Menge $\mathcal{L}((x_n))$ aller Grenzwerte von konvergenten Teilfolgen von (x_n) heißt Limesmenge von (x_n).*

d) *Die Limesmenge $\mathcal{L} := \mathcal{L}((x_n))$ der Folge (x_n) sei nichtleer. Ist (x_n) von oben (bzw. unten) beschränkt, so heißt $\sup \mathcal{L}$ bzw. $\inf \mathcal{L}$ limes superior bzw. limes inferior der Folge (x_n). Schreibweise:*

$$\begin{aligned}
\overline{\lim_{n\to\infty}}\, x_n &:= \limsup_{n\to\infty} x_n := \sup \mathcal{L} \\
\underline{\lim_{n\to\infty}}\, x_n &:= \liminf_{n\to\infty} x_n := \inf \mathcal{L}
\end{aligned}$$

Ist (x_n) von oben (bzw. unten) nicht beschränkt, so schreibt man auch
$\limsup_{n\to\infty} x_n = +\infty$ *bzw.* $\liminf_{n\to\infty} x_n = -\infty$.

Bemerkung 2.69
Ist (x_n) von oben (bzw. unten) beschränkt, so ist auch $\mathcal{L}((x_n))$ nach oben (bzw. unten) beschränkt und limes superior (bzw. inferior) sind folglich definiert in $\mathbb{R}$!
Satz 3.2 und Satz 3.3 ergeben, daß $\sup \mathcal{L}$ bzw. $\inf \mathcal{L}$ Grenzwerte gewisser monotoner Teilfolgen von (x_n) sind.
Aus den Definitionen ergibt sich ferner, daß eine Folge (x_n) konvergent ist in $\mathbb{R}$ gdw. $\overline{\lim_{n\to\infty}}\, x_n = \underline{\lim_{n\to\infty}}\, x_n$.

(Bew.: $\inf \mathcal{L} = \sup \mathcal{L}$ bedeutet: $\mathcal{L} \neq \emptyset$ und $\mathcal{L}$ einelementig.
Ferner muß (x_n) beschränkt sein und, folglich, jede Teilfolge eine konvergente Teilfolge enthalten!)

Beispiel 2.70

a) *Sei $r \in \mathbb{R}$. Dann gilt für*

$|r| < 1$: $\lim\limits_{n\to\infty} r^n = 0$ *(vgl. Bsp. 2.51 b)),*

$r = 1$: $\lim\limits_{n\to\infty} r^n = 1$,

$r \leq -1$: (r^n) *ist divergent,*

$r > 1$: (r^n) *ist bestimmt divergent gegen* $+\infty$, *da* $\left(\frac{1}{r^n}\right)$ *Nullfolge ist.*

b) *Sei $k \in I\!N$, $r \in \mathbb{R}$ mit $|r| < 1$. Dann gilt: $\lim\limits_{n\to\infty} n^k r^n = 0$.*

<u>*Bew.:*</u> *Sei o.B.d.A. $r \neq 0$; $a := \frac{1}{|r|} - 1 > 0$.*

$\rightsquigarrow |r| = (1+a)^{-1}$.

Sei $n \geq k+1$. Dann gilt (Binomischer Lehrsatz (2.12)):

$|n^k r^n| = n^k \frac{1}{(1+a)^n} \leq \frac{n^k}{\binom{n}{k+1} a^{k+1}}$ *(Satz 2.12)*

$$\rightsquigarrow |n^k r^n| \leq \frac{n^k (k+1)!}{\prod\limits_{i=0}^{k} (n-i) a^{k+1}} = \frac{1}{n} \left(\prod_{i=1}^{k} (1 - \tfrac{i}{n}) \right)^{-1} \frac{(k+1)!}{a^{k+1}} \underset{n\to\infty}{\longrightarrow} 0.$$

(letzteres folgt aus 3.7 b)) □

c) *Die Folge $((1+\frac{1}{n})^n)_{n\in I\!N}$ ist monoton wachsend und konvergent gegen* e= 2.7182818284.

<u>*Bew.:*</u> *Wir zeigen zunächst, daß die Folge monoton wachsend ist. Es gilt:*

$$\begin{aligned}
\frac{(1+\frac{1}{n+1})^{n+1}}{(1+\frac{1}{n})^{n+1}} &= \left(\frac{(n+2)n}{(n+1)^2}\right)^{n+1} = \left(\frac{(n+1)^2 + n - (n+1)}{(n+1)^2}\right)^{n+1} \\
&= \left(1 - \frac{1}{(n+1)^2}\right)^{n+1} \geq 1 - (n+1)\frac{1}{(n+1)^2} \text{ (2.12 2))} \\
&\geq \frac{n}{n+1} = \frac{1}{1+\frac{1}{n}}
\end{aligned}$$

Also gilt: $\left(1+\frac{1}{n+1}\right)^{n+1} \leq \left(1+\frac{1}{n}\right)^n$, $\forall n \in I\!N$.

Zum Beweis der Konvergenz der Folge verwenden wir Satz 3.3 und zeigen dazu ihre Beschränktheit. Aus dem Binomischen Lehrsatz erhalten wir für beliebiges $n \in I\!N$, $n \geq 2$:

$$\begin{aligned}
\left(1+\frac{1}{n}\right)^n &= 1 + \sum_{i=1}^{n} \binom{n}{i} \frac{1}{n^i} = 2 + \sum_{i=2}^{n} \binom{n}{i} \frac{1}{n^i} \\
&\leq 2 + \sum_{i=2}^{n} \frac{n(n-1)\cdots(n-i+1)}{i! \quad n^i} < 2 + \sum_{i=2}^{n} \frac{1}{i!} \\
&\leq 2 + \sum_{i=2}^{n} \frac{1}{2^{i-1}} = 2 + (1 - \frac{1}{2^n}) < 3
\end{aligned}$$

Also gilt: $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \sup\{\left(1+\frac{1}{n}\right)^n : n \in I\!N\} =: e.$

Die Dezimaldarstellung von e beweisen wir hier nicht ! $\square$

d) $\forall r \in \mathbb{R}$ *gilt:* $\lim\limits_{n\to\infty} \frac{r^n}{n!} = 0.$

<u>Bew.:</u> *Wir wählen* $n_o \in I\!N$, *so daß* $|r| \le n_o$. *Dann gilt für* $n > n_o$*:*

$$\left|\frac{r^n}{n!}\right| = \prod_{i=1}^{n} \frac{|r|}{i} = \prod_{i=1}^{n_o} \frac{|r|}{i} \prod_{i=n_o+1}^{n} \frac{|r|}{i} \le \prod_{i=1}^{n_o} \frac{|r|}{i} \left(\frac{|r|}{(n_o+1)}\right)^{n-n_o-1}$$

$$\le C(n_o, r)\left(\frac{|r|}{(n_o+1)}\right)^n \underset{n\to\infty}{\longrightarrow} 0 \text{ (wegen a))}$$

$$\text{wobei } C(n_o, r) := \frac{|r|^{n_o}}{n_o!}\frac{(n_o+1)^{n_o+1}}{|r|^{n_o+1}} = \frac{1}{|r|}\frac{(n_o+1)^{n_o+1}}{n_o!}.$$

Bemerkung 2.71

Im Falle $x_n \to \pm\infty$ *sagt man,* $\pm\infty$ *sei der <u>uneigentliche Grenzwert</u> der Folge* (x_n)*. Für die bestimmte Divergenz gelten die folgenden Eigenschaften:*

(i) (x_n) *bestimmt divergent,* (y_n) *beschränkt*
$\rightsquigarrow (x_n + y_n)$ *bestimmt divergent.*

(ii) $x_n \to +\infty,\ y_n \to +\infty \rightsquigarrow x_n + y_n \to +\infty,\ x_n y_n \to +\infty$

Die Symbole $+\infty$, $-\infty$ *sind <u>keine Zahlen</u>. Jedoch kann man die eben dargestellten Eigenschaften (und ähnliche) wie folgt formalisieren:*

$$\begin{aligned} (+\infty) + (+\infty) &= +\infty, \quad (\pm\infty) + r = \pm\infty, \quad \forall r \in \mathbb{R}, \\ (-\infty) + (-\infty) &= -\infty, \quad r(\pm\infty) = \pm\infty, \quad \forall r > 0, \\ (+\infty)(+\infty) &= +\infty, \quad (-\infty)(-\infty) = +\infty, \quad (-\infty)(+\infty) = -\infty. \end{aligned}$$

Kein bestimmter Sinn kann allerdings den folgenden Ausdrücken zugeschrieben werden:
$(+\infty) - (+\infty)$, $\frac{\pm\infty}{\pm\infty}$, $0 \cdot \pm\infty$

<u>Vereinbarung:</u> $-\infty < +\infty$, $-\infty < r < +\infty$, $\forall r \in \mathbb{R}$.

3 Reihen

3.1 Reihen in $\mathbb{C}$

Definition 3.1
Es sei (a_n) eine Folge in $\mathbb{C}$ (bzw. speziell in $\mathbb{R}$). Für jedes $n \in \mathbb{N}$ betrachten wir

$$s_n := \sum_{k=1}^{n} a_k.$$

Die Folge (s_n) heißt Folge der Partialsummen von (a_n) oder unendliche Reihe (mit den Gliedern a_n) und wird mit

$$\sum_{k=1}^{\infty} a_k \quad \textit{bezeichnet.}$$

Definition 3.2
Es sei (a_n) eine Folge in $\mathbb{C}$, und $a \in \mathbb{C}$.
Die unendliche Reihe $\sum\limits_{k=1}^{\infty} a_k$ heißt konvergent (gegen a), falls die Folge der Partialsummen (s_n) (gegen a) konvergiert. Dann heißt a Summe der unendlichen Reihe.
Schreibweise: $\sum\limits_{k=1}^{\infty} a_k = a$ (Doppelbedeutung der linken Seite!)
Anderenfalls heißt die unendliche Reihe divergent.
Die unendliche Reihe $\sum\limits_{k=1}^{\infty} a_k$ heißt absolut konvergent, falls $\sum\limits_{k=1}^{\infty} |a_k|$ konvergent ist.

Beispiel 3.3

a) *Geometrische Reihe: $\sum\limits_{k=0}^{\infty} a^k, \quad a \in \mathbb{C}$.*

Es gilt: $s_n = \sum\limits_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}, \quad \forall n \in \mathbb{N}_0, a \neq 1$.
Also gilt: $\left|s_n - \frac{1}{1-a}\right| = \frac{|a|^{n+1}}{|1-a|}, \quad \forall n \in \mathbb{N}_0$.
D.h. für $|a| < 1$ ist die geometrische Reihe konvergent und es gilt:

$$\sum_{k=1}^{\infty} a^k = \frac{1}{1-a}$$

Die Reihe ist nach Def. 3.2 dann sogar absolut konvergent. Für $|a| \geq 1$ ist die geometrische Reihe divergent!

b) *m-Darstellung reeller Zahlen:*

Für $m \in \mathbb{N}$, $m \geq 2$, erlaubt jedes $x \in \mathbb{R}$, $x \geq 0$, die Darstellung

$$x = \sup\{\sum_{i=1}^{n} z_i m^{k-i} : n \in \mathbb{N}\} \quad \textit{(Bem. 2.18)}$$

wobei $k \in I\!N_0$, $z_i \in \{n \in I\!N_0 : n \leq m-1\}$, $i \in I\!N$.
Da die Folge (s_n), $s_n := \sum\limits_{i=1}^{n} z_i m^{k-i}$, $\forall n \in I\!N$, *monoton wachsend und beschränkt ist, gilt nach Satz 3.3:* $x = \lim\limits_{n\to\infty} s_n = \sum\limits_{i=1}^{\infty} z_i m^{k-i}$.
Dies ist also eine neue Interpretation der m-Darstellung einer reellen Zahl als unendliche Reihe!

Satz 3.4
Es sei (a_n) *eine Folge komplexer Zahlen.*

a) *Die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *ist konvergent gdw.* $\forall \varepsilon > 0 \exists n_o(\varepsilon) \in I\!N$ *so daß*

$$\left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon, \quad \forall m > n \geq n_o(\varepsilon).$$

b) *Ist die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *absolut konvergent, so ist sie auch konvergent.*

c) *Ist die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *konvergent, so ist* (a_n) *eine Nullfolge.*

d) *Gilt* $a_n \in \mathbb{R}$, $a_n \geq 0$, $\forall n \in I\!N$, *so konvergiert die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *genau dann, wenn die Folge ihrer Partialsummen* (s_n) *beschränkt ist.*

e) *Sind außerdem* $\alpha, \beta \in \mathbb{C}$ *und ist* (b_n) *eine weitere Folge komplexer Zahlen, so daß die Reihen*
$\sum\limits_{k=1}^{\infty} a_k$ *und* $\sum\limits_{k=1}^{\infty} b_k$ *konvergent sind, so ist auch die Reihe*
$\sum\limits_{k=1}^{\infty} (\alpha a_k + \beta b_k)$ *konvergent und es gilt*
$\sum\limits_{k=1}^{\infty} (\alpha a_k + \beta b_k) = \alpha \sum\limits_{k=1}^{\infty} a_k + \beta \sum\limits_{k=1}^{\infty} b_k$.

Beweis:

a) Nach Satz 2.64 ist die Reihe $\sum\limits_{k=1}^{\infty} a_k$ konvergent gdw. die Folge der Partialsummen (s_n) eine Fundamentalfolge in $(\mathbb{C}, |\cdot|)$ ist. Wegen der Beziehung

$$|s_m - s_n| = |\sum_{k=n+1}^{m} a_k| \text{ für } m > n, \text{folgt daraus die Aussage.}$$

b) Die Aussage folgt wegen $|\sum\limits_{k=n+1}^{m} a_k| \leq \sum\limits_{k=n+1}^{m} |a_k| = s_m - s_n < \varepsilon \quad \forall m > n \geq n_o(\varepsilon)$, aus der vorausgesetzten absoluten Konvergenz der Reihe und a).

c) Sei $\sum_{k=1}^{\infty} a_k$ konvergent.

Aus a) folgt speziell für $m := n+1 : \forall \varepsilon > 0 \exists n_o(\varepsilon) \in I\!N$:

$$|a_{n+1}| < \varepsilon, \quad \forall n \geq n_o(\varepsilon).$$

D.h. (a_n) ist eine Nullfolge.

d) folgt sofort aus Satz 2.58, da die Folge der Partialsummen (s_n) monoton wachsend ist.

e) folgt wegen $\sum_{k=1}^{n}(\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^{n} a_k + \beta \sum_{k=1} n b_k$ aus Satz 2.70a) angewendet auf die Folgen der Partialsummen beider Reihen. □

Beispiel 3.5

Harmonische Reihe: $\sum_{k=1}^{\infty} \frac{1}{k}$

Nach Satz 3.4d) ist diese Reihe konvergent gdw. die Folge $\left(\sum_{k=1}^{n} \frac{1}{k}\right)_{n \in I\!N}$ *der Partialsummen beschränkt ist.*

Wir zeigen: Die letztere Folge ist nicht beschränkt!

Behauptung: $\frac{n}{2} + 1 \leq s_{2^n} \leq n + \frac{1}{2}$, $\forall n \in I\!N$ *(wobei* $s_n := \sum_{k=1}^{n} \frac{1}{k}$, $\forall n \in I\!N$*).*

Beweis:

Für $n = 1$ *ist die Behauptung richtig. Die Aussage sei für* $n-1$ *richtig, d.h., es gelte* $\frac{n}{2} + \frac{1}{2} \leq s_{2^{n-1}} \leq n - \frac{1}{2}$. *Dann erhalten wir*

$$\begin{aligned} s_{2^n} - s_{2^{n-1}} = \sum_{k=2^{n-1}+1}^{2^n} \frac{1}{k} &\geq 2^{n-1}\frac{1}{2^n} = \frac{1}{2} \\ &\leq 2^{n-1}\frac{1}{2^{n-1}} = 1 \end{aligned}$$

$\rightsquigarrow s_{2^n} \geq s_{2^{n-1}} + \frac{1}{2} \geq \frac{n}{2} + 1$

$\rightsquigarrow s_{2^n} \leq s_{2^{n-1}} + 1 \leq n + \frac{1}{2}$. □

Also ist die Folge (s_n) *nicht beschränkt und deshalb die harmonische Reihe divergent! (obwohl* (s_n) *sehr langsam wächst; es gilt: z.B.* $s_{5000} = 9.0945$, $s_{10000} = 9.7876$, $s_{100000} = 12.090146$.*)*

Im folgenden beschäftigen wir uns nun mit Konvergenz– und Divergenz–Kriterien für unendliche Reihen, d.h. hinreichenden Bedingungen für Konvergenz bzw. Divergenz, die einfach zu überprüfen sind.

Satz 3.6 *(Leibniz–Kriterium für alternierende Reihen)*

Es sei (a_n) *eine monoton fallende Nullfolge mit* $a_n \geq 0$, $\forall n \in I\!N$. *Dann ist die "alternierende" Reihe* $\sum_{k=1}^{\infty}(-1)^k a_k$ *konvergent und für ihre Summe a gilt:*

$$|a - \sum_{k=1}^{n}(-1)^k a_k| \leq a_{n+1}, \quad \forall n \in I\!N.$$

Beweis:
Von der Folge $(s_n) = (\sum\limits_{k=1}^{n}(-1)^k a_k)$ betrachten wir die Teilfolgen $(s_{2m})_{m\in\mathbb{N}}$ und $(s_{2m+1})_{m\in\mathbb{N}_0}$. Es gilt:

$$\begin{aligned} s_{2m+2} &= s_{2m} - a_{2m+1} + a_{2m+2} \leq s_{2m}, \quad \forall m \in \mathbb{N}, \\ s_{2m+3} &= s_{2m+1} + a_{2m+2} - a_{2m+3} \geq s_{2m+1}, \quad \forall m \in \mathbb{N}_0. \end{aligned}$$

$\rightsquigarrow$ (s_{2m}) ist monoton fallend und (s_{2m+1}) monoton wachsend.
Ferner gilt: $-a_1 \leq s_n \leq a_2, \quad \forall n \in \mathbb{N}$.
Deshalb ist (s_n) beschränkt und beide Teilfolgen sind konvergent!
Wegen $s_{2m+1} - s_{2m} = a_{2m+1}$ und $a_{2m+1} \underset{m\to\infty}{\longrightarrow} 0$ haben aber beide Teilfolgen denselben Grenzwert a.
Wegen der Wahl der Teilfolgen (sie "zerlegen" die gesamte Folge (s_n)) konvergiert deshalb (s_n) gegen a.
Wir betrachten nun $n, m \in \mathbb{N}$ mit $m > n$:

$$\rightsquigarrow |s_m - s_n| = |\sum_{j=n+1}^{m}(-1)^j a_j| = |a_{n+1} - a_{n+2} + a_{n+3} \ldots| \leq a_{n+1}$$

$$\rightsquigarrow |a - s_n| = \lim_{m\to\infty} |s_m - s_n| \leq a_{n+1} \qquad \square$$

Beispiel 3.7
Leibnizsche Reihe: $\sum\limits_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}$.
Diese Reihe ist nach Satz 3.6 konvergent und für ihre Summe a gilt: $|a - s_n| \leq \frac{1}{n+1}$, $\forall n \in \mathbb{N}$.
Es gilt: $s_{1000} = 0.692647$, $s_{10000} = 0.693097$
$a = 0.693147$,
Allerdings ist die Leibnizsche Reihe nicht absolut konvergent! (nach Bsp. 3.5)

Satz 3.8 *(Majoranten– bzw. Minoranten–Kriterium)*
Es seien (a_n) *und* (b_n) *Folgen in* $\mathbb{R}$.

a) *Existiert ein* $n_o \in \mathbb{N}$, *so daß* $|a_n| \leq b_n$, $\forall n \geq n_o$, *und ist die Reihe* $\sum\limits_{k=1}^{\infty} b_k$ *konvergent, so ist die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *absolut konvergent.*

b) *Existiert ein* $n_o \in \mathbb{N}$, *so daß* $0 \leq b_n \leq a_n$, $\forall n \geq n_o$, *und ist die Reihe* $\sum\limits_{k=1}^{\infty} b_k$ *divergent, so ist auch* $\sum\limits_{k=1}^{\infty} a_k$ *divergent.*

Beweis:

a) Für alle $m > n \geq n_o$ gilt: $\sum\limits_{k=n+1}^{m} |a_k| \leq \sum\limits_{k=n+1}^{m} b_k$.

Die Behauptung folgt nun aus der vorausgesetzten Konvergenz der Reihe $\sum\limits_{k=1}^{\infty} b_k$ und aus Satz 3.4a).

b) Für die Partialsummen der Reihen gilt für $n > n_o$:

$$0 \leq s_n - s_{n_o} = \sum_{k=n_o+1}^{n} b_k \leq \sum_{k=n_o+1}^{n} a_k =: \tilde{s}_n - \tilde{s}_{n_o}.$$

Nach Voraussetzung muß die monoton wachsende Folge $(s_n - s_{n_o})$ unbeschränkt wachsen, da sie sonst konvergent wäre.
Dasselbe gilt dann wegen $s_n - s_{n_o} \leq \tilde{s}_n - \tilde{s}_{n_o}$ auch für $(\tilde{s}_n)$! □

Satz 3.9 *(Wurzelkriterium)*
Es sei (a_n) eine Folge in $\mathbb{C}$ und $\alpha := \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} \in [0, +\infty]$.

a) *Ist $\alpha < 1$, so ist die Reihe $\sum_{k=1}^{\infty} a_k$ absolut konvergent.*

b) *Ist $\alpha > 1$, so ist diese Reihe divergent.*

Beweis:

a) Es sei $\alpha < 1$. Wegen $\alpha < \frac{\alpha+1}{2} < 1$ existiert ein $n_o \in I\!N$, so daß $\sqrt[n]{|a_n|} \leq \frac{\alpha+1}{2}$, $\forall n \geq n_o$.
$\rightsquigarrow$ $|a_n| \leq \left(\frac{\alpha+1}{2}\right)^n$, $\forall n \geq n_o$, die geometrische Reihe
$\sum_{k=1}^{\infty} \left(\frac{\alpha+1}{2}\right)^k$
ist aber konvergent.
$\rightsquigarrow$ die Behauptung folgt aus Satz 3.8a).

b) Ist $\alpha > 1$, so existiert $\tilde{I\!N} := \{n_1 < n_2 < \ldots < n_k < \ldots\} \subseteq I\!N$, so daß $\sqrt[n]{|a_n|} \geq 1 \rightsquigarrow |a_n| \geq 1$, $\forall n \in \tilde{I\!N}$.
Folglich ist (a_n) keine Nullfolge und die Reihe $\sum_{k=1}^{\infty} a_k$ nach Satz 3.4c) nicht konvergent. □

Satz 3.10 *(Quotientenkriterium)*
Es sei (a_n) eine Folge in $\mathbb{C}$ und $\beta := \overline{\lim_{n\to\infty}} \left|\frac{a_{n+1}}{a_n}\right|$, $\gamma := \underline{\lim_{n\to\infty}} \left|\frac{a_{n+1}}{a_n}\right| \in [0, +\infty]$ ($a_n \neq 0$, $\forall n$ hinreichend groß).

a) *Ist $\beta < 1$, so ist die Reihe $\sum_{k=1}^{\infty} a_k$ absolut konvergent.*

b) *Ist $\gamma > 1$, so ist diese Reihe divergent.*

Beweis:

a) Ist $\beta < 1 \leadsto \beta < \frac{\beta+1}{2} < 1$, so existiert ein $n_o \in I\!N$, so daß $\left|\frac{a_{n+1}}{a_n}\right| \leq \frac{\beta+1}{2} < 1$, für alle $n \geq n_o$. Dann gilt für $n > n_o$:

$$\begin{aligned} \left|\frac{a_n}{a_{n_o}}\right| &= \prod_{i=n_o}^{n-1} \left|\frac{a_{i+1}}{a_i}\right| \leq \prod_{i=n_o}^{n-1} \frac{\beta+1}{2} = \left(\frac{\beta+1}{2}\right)^{n-n_o} \\ \leadsto |a_n| &\leq \left[|a_{n_o}| \left(\frac{2}{\beta+1}\right)^{n_o}\right] \left(\frac{\beta+1}{2}\right)^n, \quad \forall n > n_o. \end{aligned}$$

Die Aussage folgt nun wieder aus dem Majorantenkriterium (3.8a)).

b) Ist $\gamma > 1$ so existiert ein $n_o \in I\!N$, so daß $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$, $\forall n \geq n_o$. Also gilt $\forall n > n_o$:

$$\left|\frac{a_n}{a_{n_o}}\right| = \prod_{i=n_o}^{n-1} \left|\frac{a_{i+1}}{a_i}\right| \geq 1, \quad \text{d.h.} \quad |a_n| \geq |a_{n_o}| > 0$$

Also ist (a_n) keine Nullfolge und damit $\sum\limits_{k=1}^{\infty} a_k$ divergent. □

Bemerkung 3.11
Für die in den Sätzen 3.9 und 3.10 definierten Zahlen α und β gilt die folgende Beziehung:

$\alpha \leq \beta$ *(Übung: vgl. Heuser Bd. 1, S. 182)*

Dies bedeutet, daß das Wurzelkriterium etwas leistungsfähiger (aber auch schwerer zu handhaben) ist als das Quotientenkriterium.

Beispiel 3.12
Exponentialreihe: $\sum\limits_{k=0}^{\infty} \frac{a^k}{k!} \quad (a \in \mathbb{C})$.
Mit $a_n := \frac{a^n}{n!}$, $\forall n \in I\!N$, *gilt* $\left|\frac{a_{n+1}}{a_n}\right| = \frac{|a|}{n+1}$, $\forall n \in I\!N$.
Also folgt $\beta := \overline{\lim\limits_{n\to\infty}} \left|\frac{a_{n+1}}{a_n}\right| = 0$ *und die Exponentialreihe ist absolut konvergent nach Satz 3.10 a).*

Definition 3.13
Es sei $\sum\limits_{k=1}^{\infty} a_k$ *eine gegebene unendliche Reihe.*
Ist $f : I\!N \to I\!N$ *bijektiv, so heißt die Reihe* $\sum\limits_{k=1}^{\infty} a_{f(k)}$ *eine <u>Umordnung</u> der Reihe.*

$\sum\limits_{k=1}^{\infty} a_k$ *heißt <u>unbedingt konvergent</u>, wenn jede Umordnung der gegebenen Reihe konvergent ist und dieselbe Summe hat.*

$\sum\limits_{k=1}^{\infty} a_k$ *heißt <u>bedingt konvergent</u>, wenn sie konvergent, aber nicht unbedingt konvergent ist.*

Satz 3.14
Jede absolut konvergente Reihe $\sum\limits_{k=1}^{\infty} a_k$ *in* $\mathbb{C}$ *ist unbedingt konvergent.*

Beweis:
Es sei $f : I\!N \to I\!N$ bijektiv und $b_n := a_{f(n)}$, $\forall n \in I\!N$.
Sei $\varepsilon > 0$ beliebig gewählt.
Nach Vor. $\exists n_o \in I\!N : \sum\limits_{k=n_o+1}^{m} |a_k| < \frac{\varepsilon}{2}$, $\forall m > n_o$ (Satz 3.16 a)).
Wir definieren $r_o := \max\{f^{-1}(n) : n = 1, \ldots, n_o\} \quad (r_o \geq n_o)$.

$$\rightsquigarrow \quad \{1, \ldots, n_o\} \subseteq \{f(1), \ldots, f(r_o)\}$$

$$\rightsquigarrow \quad \sum_{k=1}^{r_o} b_k = \sum_{k=1}^{n_o} a_k + \sum_{\substack{k=1 \\ f(k)>n_o}}^{r_o} b_k$$

$$\rightsquigarrow \left| \sum_{k=1}^{r} b_k - \sum_{k=1}^{n_o} a_k \right| = \left| \sum_{\substack{k=1 \\ f(k)>n_o}}^{r_o} a_{f(k)} \right| \leq \sum_{\substack{k=1 \\ f(k)>n_o}}^{r_o} |a_{f(k)}| < \frac{\varepsilon}{2}, \forall r > r_o.$$

Damit gilt für alle $r > r_o$ und $n > n_o$:

$$\left| \sum_{k=1}^{r} b_k - \sum_{k=1}^{n} a_k \right| \leq \left| \sum_{k=1}^{r} b_k - \sum_{k=1}^{n_o} a_k \right| + \sum_{k=n_o+1}^{n} |a_k| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Also konvergiert auch die Folge der Partialsummen $\left(\sum\limits_{k=1}^{r} b_k \right)_{r \in I\!N}$ gegen den Grenzwert $\sum\limits_{k=1}^{\infty} a_k$. Da die Abbildung f beliebig gewählt war, ist die Aussage bewiesen. $\square$

Satz 3.15 *(Riemannscher Umordnungssatz)*
Eine konvergente, aber nicht absolut konvergente, Reihe $\sum\limits_{k=1}^{\infty} a_k$ *in* $\mathbb{C}$ *ist bedingt konvergent. Für jede bedingt konvergente Reihe in* $\mathbb{R}$ *und jede Zahl* $a \in \mathbb{R}$ *existiert eine Umordnung, die gegen* a *konvergiert.*

Beweisidee: (vgl. Heuser, Bd. 1, S. 197–199)
Es sei $a \in \mathbb{R}$ beliebig gewählt und wir definieren $a_k^+ := \max\{0, a_k\}$ und $a_k^- := \max\{0, -a_k\}$, d.h. es gilt $a_k = a_k^+ - a_k^-$ und $|a_k| = a_k^+ + a_k^-$ für jedes $k \in I\!N$. Nach Voraussetzung müssen beide Reihen $\sum_{k=1}^{\infty} a_k^+$ und $\sum_{k=1}^{\infty} a_k^-$ divergent sein, d.h., unbeschränkte Partialsummenfolgen besitzen.
O.B.d.A. sei $a > 0$. Es sei nun $n_1 \in I\!N$ der kleinste Index, für den $\sum_{k=1}^{n_1} a_k^+ > a$ gilt. Als nächstes wählen wir $n_2 \in I\!N$ als kleinsten Index für den die Ungleichung $\sum_{k=1}^{n_1} a_k^+ - \sum_{k=0}^{n_2} a_k^- < a$ gültig ist. Dann entsprechend $n_3 \in I\!N$ mit $\sum_{k=1}^{n_1} a_k^+ - \sum_{k=0}^{n_2} a_k^- + \sum_{n_1+1}^{n_3} a_k^+ > a$ usw. Da die Indizes n_i jeweils minimal gewählt waren, konvergiert der Abstand von a und der jeweiligen Summe gegen 0. Damit haben wir eine Umordnung der Reihe konstruiert, die gegen a konvergiert. $\square$

Bemerkung 3.16
Satz 3.15 erklärt die Bedeutung der bedingten und der absoluten Konvergenz von Reihen. Nach diesem Satz und Bsp. 3.5 ist also die Leibnizsche Reihe bedingt konvergent. Die Problematik der Umordnung von Reihen ist sofort gegenwärtig, wenn man die Konvergenz der Produkte von Reihen untersucht: Es seien (a_n) und (b_m) Folgen in $\mathbb{C}$ und wir betrachten die Folge $\left(\sum_{k=1}^{n}\sum_{l=1}^{m} a_k b_l\right)_{n,m\in I\!N}$. Legt man nun eine Durchnumerierung von $I\!N \times I\!N$ fest, so entsteht eine sog. <u>Produktreihe</u>.
Nun erhebt sich die Frage, ob diese Produktreihe oder gar alle Varianten (sprich: Umordnungen) solcher Produktreihen konvergieren ? Satz 3.15 macht klar, daß im Fall nur bedingter Konvergenz hierbei große Probleme entstehen.

Wir geben nun eine positive Antwort auf diese Frage:

Satz 3.17 *(Multiplikationssatz)*
Es seien $(a_n)_{n\in I\!N_0}$ und $(b_n)_{n\in I\!N_0}$ Folgen in $\mathbb{C}$ und wir betrachten die Folge $(c_n)_{n\in I\!N_0}$, wobei $c_n := \sum_{i=0}^{n} a_i b_{n-i}$, $\forall n \in I\!N_0$.
Sind die Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergent, so ist ihr sog. "Cauchy–Produkt" $\sum_{k=0}^{\infty} c_k$ ebenfalls absolut konvergent mit der Summe $\left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{k=0}^{\infty} b_k\right)$.

<u>Beweis:</u>
Es bezeichne für $n \in I\!N_0 : \tilde{c}_n := \sum_{i=0}^{n} |a_i||b_{n-i}|$,

$$s_n := \sum_{i=0}^{n} |a_i|, \tilde{s}_n := \sum_{i=0}^{n} |b_i|.$$

Wir zeigen im folgenden, daß die Reihe $\sum_{k=0}^{\infty} \tilde{c}_k$ konvergent ist. Wegen $|c_n| = |\sum_{i=0}^{n} a_i b_{n-i}| \leq \sum_{i=0}^{n} |a_i||b_{n-i}| = \tilde{c}_n$, $\forall n \in I\!N_0$, folgt dann die Aussage aus dem Majoranten–Kriterium Satz 3.8. Es sei $n \in I\!N$ bel. Dann gilt für hinreichend großes $m \in I\!N$:

$$\sum_{k=0}^{n} \tilde{c}_k = \sum_{k=0}^{n}\sum_{i=0}^{n} |a_i||b_{n-i}| \leq \left(\sum_{k=0}^{m} |a_k|\right)\left(\sum_{i=0}^{m} |b_i|\right) \leq \left(\sum_{k=0}^{\infty} |a_k|\right)\left(\sum_{k=0}^{\infty} |b_k|\right)$$

Also ist die Folge der Partialsummen $\left(\sum_{k=0}^{n} \tilde{c}_k\right)_{n\in I\!N_0}$ beschränkt und deshalb konvergent (3.4d)). Da das Cauchy-Produkt eine bestimmte Numerierung des Produktes der Partialsummen beider Reihen darstellt, muß die Summe des Cauchy-Produktes gerade das Produkt der beiden Summen der Ausgangsreihen sein. □

Eine analoge Aussage gilt für jede beliebige Durchnumerierung von $I\!N_0 \times I\!N_0$. Die Durchnumerierung der Produkte $\{a_k b_l : k, l \in I\!N_0\}$ entspricht dabei dem Cauchy–Cantorschen Diagonalverfahren (vgl. Beweis von Satz 2.23).

3.2 Potenzreihen und Elementarfunktionen

Definition 3.18

Es sei $(a_n)_{n\in \mathbb{N}_0}$ eine Folge in $\mathbb{C}$, $a \in \mathbb{C}$ und $z \in \mathbb{C}$.
Die unendliche Reihe $\sum\limits_{k=0}^{\infty} a_k(z-a)^k$ heißt *Potenzreihe* *(in z) mit der Koeffizientenfolge (a_n) und der Entwicklungsstelle a (Vereinbarung: $(z-a)^0 = 1$).*

Frage: *Für welche $z \in \mathbb{C}$ konvergiert eine Potenzreihe (bei gegebenen a, a_n, $n \in \mathbb{N}_0$)? Auf jeden Fall für $z = a$!*

Beispiel 3.19

a) $\sum\limits_{k=0}^{\infty} \frac{1}{k!}(z-a)^k$

Nach Beispiel 3.12 ist diese Reihe $\forall z \in \mathbb{C}$ absolut konvergent!

b) $\sum\limits_{k=0}^{\infty} k!(z-a)^k$

Wir wenden das Quotientenkriterium (Satz 3.10) für $z \neq a$ an:

$$\lim_{n\to\infty} \left| \frac{(n+1)!(z-a)^{n+1}}{n!\,(z-a)^n} \right| = \lim_{n\to\infty} (n+1)|z-a| = +\infty$$

Deshalb ist die Potenzreihe für $z \neq a$ divergent (3.10 b))!

Satz 3.20

Es sei $(a_n)_{n\in \mathbb{N}_0}$ eine Folge in $\mathbb{C}$, $\bar{\alpha} := \overline{\lim\limits_{n\to\infty}} \sqrt[n]{|a_n|} \in [0, +\infty]$ und $a \in \mathbb{C}$.

a) *Ist $\bar{\alpha} \in]0, +\infty[$, so konvergiert die Potenzreihe $\sum\limits_{k=0}^{\infty} a_k(z-a)^k$ absolut, falls $|z-a| < \frac{1}{\bar{\alpha}}$, und divergiert falls $|z-a| > \frac{1}{\bar{\alpha}}$.*

b) *Ist $\bar{\alpha} = 0$, so konvergiert die Potenzreihe absolut $\forall z \in \mathbb{C}$.*

c) *Ist $\bar{\alpha} = +\infty$, so konvergiert die Potenzreihe nur für $z = a$.*

Beweis:
Wir wenden das Wurzelkriterium (Satz 3.9) an und definieren

$$\alpha(z) := \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n(z-a)^n|} = \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n||z-a|} = \bar{\alpha}|z-a|$$

a) Es gilt: $|z-a| < \bar{\alpha}^{-1} \Longleftrightarrow \alpha(z) < 1$,
d.h. die Aussage folgt aus Satz 3.9a).
Es gilt: $|z-a| > \bar{\alpha}^{-1} \Longleftrightarrow \alpha(z) > 1$
d.h. der zweite Teil der Aussage folgt aus Satz 3.9b).

b) Falls $\bar{\alpha} = 0$ so gilt $\alpha(z) = 0$, $\forall z \in \mathbb{C}$, und 3.9a) kann angewendet werden.

c) Falls $\bar{\alpha} = +\infty$, so gilt $\alpha(z) = +\infty$, $\forall z \neq a$, und die Aussage folgt wiederum aus 3.9b). $\square$

Definition 3.21
Es sei $a \in \mathbb{C}$, $(a_n)_{n \in \mathbb{N}_0}$ eine Folge in $\mathbb{C}$ und $\bar{\alpha} := \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} \in [0, +\infty]$.
Dann heißt $\rho := \begin{cases} +\infty, & \bar{\alpha} = 0 \\ \bar{\alpha}^{-1}, & \bar{\alpha} \in]0, +\infty[, \\ 0, & \bar{\alpha} = +\infty \end{cases}$ Konvergenzradius der Potenzreihe $\sum_{k=0}^{\infty} a_k(z - a)^k$. Die Menge $\{z \in \mathbb{C} : |z - a| < \rho\}$ heißt Konvergenz–Kreisscheibe.

Bemerkung 3.22
Satz 3.20 rechtfertigt diese Definition des Konvergenzradius einer Potenzreihe und besagt, daß eine Potenzreihe stets auf der (offenen) Kreisscheibe $\{z \in \mathbb{C} : |z - a| < \rho\}$ absolut konvergiert und in $\mathbb{C} \backslash \{z \in \mathbb{C} : |z - a| \leq \rho\}$ divergiert. Zum Verhalten auf dem Rand $\{z \in \mathbb{C} : |z - a| = \rho\}$ des "Konvergenzkreises" macht der Satz keine Aussage.
Handelt es sich um eine reelle Potenzreihe, d.h. $a_n \in \mathbb{R}$, $\forall n \in \mathbb{N}_0$, und $a \in \mathbb{R}$, $\sum_{k=0}^{\infty} a_k(x - a)^k$ $(x \in \mathbb{R})$, so sprechen wir analog vom "Konvergenzintervall" $]a - \rho, a + \rho[$.
Als Schlußfolgerung aus den Sätzen 3.4 e) und 3.17 könnten nun auch Aussagen über die Addition und Multiplikation von Potenzreihen formuliert werden. Wir verzichten hier darauf, kommen aber später auf Potenzreihen zurück.

Beispiel 3.23
Die Potenzreihe $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ konvergiert für alle $z \in \mathbb{C}$ absolut (3.12).
Es gilt: $\bar{\alpha} := \overline{\lim_{n\to\infty}} \sqrt[n]{\frac{1}{n!}} = 0$, d.h. $\rho = +\infty$.
Beweis: Sei $\varepsilon > 0$ beliebig gewählt. Da die Folge $\left(\frac{\frac{1}{\varepsilon}^n}{n!}\right)$ eine Nullfolge ist, existiert ein $n_0 \in \mathbb{N}$, so daß $\frac{\frac{1}{\varepsilon}^n}{n!} < 1$, $\forall n \geq n_0$. Deshalb gilt für alle $n \geq n_0$, daß $\frac{\frac{1}{\varepsilon}}{\sqrt[n]{n!}} < 1$ oder $\frac{1}{\sqrt[n]{n!}} < \varepsilon$. $\square$
Es gilt sogar die Abschätzung: $e\left(\frac{n}{e}\right)^n \leq n! \leq en\left(\frac{n}{e}\right)^n$, $\forall n \in \mathbb{N}$ (Heuser, Bd. 1).

Dies rechtfertigt die folgende

Definition 3.24
Die Abbildung $\exp : \mathbb{C} \to \mathbb{C}$, $\exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!}$, $\forall z \in \mathbb{C}$, heißt (komplexe) Exponentialfunktion. Die Einschränkung $\exp|_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ heißt (reelle) Exponentialfunktion.

Satz 3.25
Die Exponentialfunktion hat folgende Eigenschaften:

a) $\exp(z_1 + z_2) = \exp(z_1)\exp(z_2)$, $\forall z_1, z_2 \in \mathbb{C}$ *(Additionstheorem);*

b) $\exp(0) = 1$, $\exp(x) > 1$, $\forall x > 0$, $\quad 0 < \exp(x) < 1$, $\forall x < 0$;

c) $\exp(1) = e$ *und e ist irrational;*

d) $\exp(x) = e^x$, *$\forall x \in \mathbb{Q}$;*

e) $\exp : \mathbb{R} \to \mathbb{R}$ *ist injektiv.*

<u>Beweis:</u>

a) Für beliebige $z_1, z_2 \in \mathbb{C}$ gilt nach Satz 3.17:

$$\begin{aligned}\exp(z_1)\exp(z_2) &= \left(\sum_{k=0}^{\infty}\frac{z_1^k}{k!}\right)\left(\sum_{k=0}^{\infty}\frac{z_2^k}{k!}\right) = \sum_{k=0}^{\infty}\sum_{i=0}^{k}\frac{z_1^i}{i!}\frac{z_2^{k-i}}{(k-i)!} \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\sum_{i=0}^{k}\binom{k}{i}z_1^i z_2^{k-i} = \sum_{k=0}^{\infty}\frac{1}{k!}(z_1+z_2)^k \text{ (2.10!)} \\ &= \exp(z_1+z_2)\end{aligned}$$

b) $\exp(0) = 1$ ist klar nach Definition 3.24.
Aus a) folgt ferner $\exp(x)\exp(-x) = 1$, $\forall x \in \mathbb{R}$.
$\forall x > 0$ gilt:

$$\exp(x) = 1 + \sum_{k=1}^{\infty}\frac{x^k}{k!} > 1$$

$\forall x < 0$: $\exp(-x) > 1 \leadsto \exp(x) = \frac{1}{\exp(-x)} \in]0,1[$.

c) Wir zeigen zunächst: $\exp(1) = \lim\limits_{n\to\infty}(1+\frac{1}{n})^n$.
$\forall n \in I\!N$ gilt:

$$\begin{aligned}&1 + \tfrac{1}{n} < \sum_{k=0}^{\infty}\tfrac{1}{k!}\left(\tfrac{1}{n}\right)^k = \exp\left(\tfrac{1}{n}\right),\ \exp(1) = \exp\left(\tfrac{1}{n}\cdot n\right) = \left(\exp\left(\tfrac{1}{n}\right)\right)^n \\ \leadsto\ &1+\tfrac{1}{n} < (\exp(1))^{1/n} \leadsto (1+\tfrac{1}{n})^n < \exp(1) \\ \leadsto\ &\lim_{n\to\infty}(1+\tfrac{1}{n})^n \le \exp(1)\end{aligned}$$

Andererseits gilt für $n \ge 2$:
$(\exp(1))^{1/n} = \exp\left(\frac{1}{n}\right) = \sum\limits_{k=0}^{\infty}\frac{1}{k!n^k} < \sum\limits_{k=0}^{\infty}\frac{1}{n^k} = \frac{1}{1-\frac{1}{n}} = 1 + \frac{1}{n-1}$ (vgl. 3.15 a))
$\leadsto \exp(1)\left(1+\frac{1}{n-1}\right)^{-1} < \left(1+\frac{1}{n-1}\right)^{n-1}$
$\leadsto$ durch Grenzübergang $n \to \infty$ erhalten wir

$$\exp(1) \le \lim_{n\to\infty}(1+\frac{1}{n})^n.$$

Deshalb gilt nach Beispiel 2.70c): $\exp(1) = e$.
Wir zeigen jetzt: e ist irrational.
<u>*Annahme:*</u> e ist rational, d.h. $\exists m, n \in I\!N : e = \frac{m}{n}$.

$$\begin{aligned}\leadsto e &= \frac{m}{n} = \sum_{k=0}^{\infty}\frac{1}{k!} > \sum_{k=0}^{n}\frac{1}{k!} \\ \leadsto 0 &< m(n-1)! - \sum_{k=0}^{n}\frac{n!}{k!} \in I\!N \quad (\star)\end{aligned}$$

Ferner: $e - \sum\limits_{k=0}^{n} \frac{1}{k!} = \sum\limits_{k=n+1}^{\infty} \frac{1}{k!} < \frac{1}{(n+1)!} \sum\limits_{k=0}^{\infty} \frac{1}{(n+1)^k} = \frac{1}{(n+1)!} \frac{1}{1-\frac{1}{n+1}} = \frac{1}{n!}\frac{1}{n}$

$\rightsquigarrow m(n-1)! - \sum\limits_{k=0}^{n} \frac{n!}{k!} = n! \left(e - \sum\limits_{k=0}^{n} \frac{1}{k!} \right) < \frac{1}{n}$

Wegen $e \in]2,3[$, d.h. $n \neq 1$, ist dies ein Widerspruch zu $(\star)$.

d) Für alle $x \in \mathbb{R}$, $n \in I\!N$ gilt zunächst:

$$\exp(nx) = (\exp(x))^n \text{ (vgl. a)); } \exp\left(\frac{1}{n}\right) = \exp(1)^{\frac{1}{n}} = e^{\frac{1}{n}} \text{ (vgl. c))}$$

Also gilt für alle $m, n \in I\!N$:

$$\exp\left(\frac{m}{n}\right) = \left(\exp\left(\frac{1}{n}\right)\right)^m = (e^{1/n})^m = e^{\frac{m}{n}}$$

Daraus und aus $\exp(-x) = (\exp(x))^{-1}$, $\forall x \in \mathbb{R}$, folgt die Aussage.

e) Wir zeigen: $(\exp|_{\mathbb{R}})^{-1} : R(exp) \to \mathbb{R}$ ist eindeutig.

Seien (y, x_1), $(y, x_2) \in (\exp|_{\mathbb{R}})^{-1}$, d.h. $(x_i, y) \in \exp|_{\mathbb{R}}$, $i = 1, 2$.

$\rightsquigarrow \exp(x_1) = \exp(x_2) \rightsquigarrow \exp(x_1 - x_2) = 1 \quad$ (nach a))

$\rightsquigarrow$ nach b): $x_1 - x_2 = 0 \rightsquigarrow x_1 = x_2$

$\rightsquigarrow \exp|_{\mathbb{R}}$ ist injektiv. □

Definition 3.26

a) $e^x := \exp(x)$, $\forall x \in \mathbb{R} \setminus \mathbb{Q}$;

b) *Die zu* $\exp : \mathbb{R} \to \mathbb{R}$ *inverse Abbildung* $\ln : R(\exp) \to \mathbb{R}$ *heißt* *natürlicher Logarithmus*, *es gilt also:*

$$\exp(\ln x) = x, \quad \forall x \in R(\exp),$$
$$\ln(\exp x) = x, \qquad \forall x \in \mathbb{R}.$$

c) *Für* $a \in R(\exp)$ *wird die Exponentialfunktion zur Basis a als Abbildung von* $\mathbb{R}$ *in* $\mathbb{R}$ *definiert durch:*

$$a^x := \exp(x \cdot \ln(a)) = e^{x \ln a}, \forall x \in \mathbb{R}.$$

Bemerkung 3.27

Definition 3.26 a) wird durch Satz 3.25 d) motiviert. Später erfolgt die entgültige Rechtfertigung durch den Nachweis der Stetigkeit von exp.

Die inverse Abbildung ln *zu* exp *ist eindeutig wegen 3.25 e). Später zeigen wir noch, daß* $R(\exp) = \exp(\mathbb{R}) =]0, +\infty[$ *gilt. Klar ist bisher:* $R(\exp) \subseteq]0, +\infty[$ *(Satz 3.25 b)).*

Für die stärkere Aussage reichen unsere gegenwärtig verfügbaren mathematischen Hilfsmittel noch nicht aus!

Analog zum Beweis von d) zeigt man für beliebige $m, n \in I\!N$*:*

$$\exp\left(\frac{m}{n} \ln a\right) = (\exp(\ln a))^{\frac{m}{n}} = a^{\frac{m}{n}}$$

D.h. für alle $x \in \mathbb{Q}$ *stimmt die in 3.26 c) gegebene Definition von* a^x *mit der früheren aus Bem. 2.12 überein.*

Folgerung 3.28
Die Logarithmusfunktion $\ln$ *hat folgende Eigenschaften:*

a) $\ln(xy) = \ln x + \ln y, \quad \forall x, y \in R(\exp),$

b) $\ln(a^x) = x \ln a, \quad \forall a \in R(\exp), \forall x \in \mathbb{R},$

c) $\ln x < 0$, *falls* $x < 1$ $(x \in R(\exp))$
$\ln x > 0$, *falls* $x > 1$
$\ln x = 0$, *falls* $x = 1$.

Beweis: folgt sofort aus Satz 3.25 und Def. 3.26. □

Definition 3.29
Wir definieren die folgenden Elementarfunktionen als Abbildungen von $\mathbb{R}$ *in* $\mathbb{R}$*:*

$$\begin{array}{rll} \cos x & := \operatorname{Re}(\exp(ix)) & \textit{Kosinusfunktion} \\ \sin x & := \operatorname{Im}(\exp(ix)) & \textit{Sinusfunktion} \\ \cosh x & := \frac{1}{2}(\exp(x) + \exp(-x)) & \textit{Cosinus hyperbolicus} \\ \sinh x & := \frac{1}{2}(\exp(x) - \exp(-x)) & \textit{Sinus hyperbolicus} \\ & (\forall x \in \mathbb{R}) & \end{array}$$

Folgerung 3.30
Für diese Elementarfunktionen gilt die folgende Darstellung:

$$\begin{array}{llll} a) & \cos x = \sum\limits_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, & c) & \cosh x = \sum\limits_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \\ b) & \sin x = \sum\limits_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, & d) & \sinh x = \sum\limits_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \end{array} \qquad (\forall x \in \mathbb{R})$$

Alle diese Reihen sind für jedes $x \in \mathbb{R}$ *absolut konvergent.*

Beweis:
Für alle $x \in \mathbb{R}$ gilt: $\exp(ix) = \sum\limits_{k=0}^{\infty} \frac{i^k x^k}{k!}$. Deshalb folgt aus Lemma 2.57:

$$\cos x = \operatorname{Re}(\exp(ix)) = \sum_{k=0}^{\infty} \operatorname{Re}\left(\frac{i^k x^k}{k!}\right) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

Analog resultiert die Formel für $\sin x$.
Wir beweisen nun noch c) (analog folgt d)):

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

Aus der absoluten Konvergenz der Exponentialreihe folgt die absolute Konvergenz der Reihen der Real– bzw. Imaginärteile und insgesamt die absolute Konvergenz aller Reihen 3.30 a) bis d). □

Bemerkung 3.31
Man könnte jetzt weitere Elementarfunktionen, abgeleitet aus exp, ln, cos, sin, cosh, sinh *definieren und die entsprechenden Eigenschaften diskutieren, wie z.B.* tan, cot *usw. Für die in 3.29 definierten reellen Funktionen ergeben sich nun aus Satz 3.25 und Folgerung 3.30 vielfältige Eigenschaften, wie z.B.*

$$\begin{array}{rcl} \cos x = \cos(-x) & , & \sin(-x) = -\sin x, \\ (\cos x)^2 + (\sin x)^2 & = & |\exp(ix)|^2 = \exp(ix) \cdot \overline{\exp(ix)} \\ & = & \exp(ix)\exp(-ix) = 1 \end{array}$$

(vgl. Bem. 2.28d): komplexe Zahlenebene)

$$\begin{array}{rcl} \sin(x+y) & = & \mathrm{Im}(\exp(i(x+y))) = \mathrm{Im}(\exp(ix)\exp(iy)) \\ & = & \mathrm{Re}(\exp(ix))\mathrm{Im}(\exp(iy)) + \mathrm{Im}(\exp(ix))\mathrm{Re}(\exp(iy)) \\ & = & \cos(x)\sin(y) + \cos(y)\sin(x) \end{array}$$

Weitere Additionstheoreme der Winkelfunktionen u.v.a.m.

4 Lineare Algebra

4.1 Lineare Räume

In Verallgemeinerung der Definition des m-dimensionalen linearen Raums $\mathbb{R}^m$ (vgl. Kapitel 2.3) führen wir einen (allgemeinen) linearen Raum ein.

Definition 4.1 *(Axiomensystem des linearen Raumes)*
Eine Menge X mit einer Abbildung von $X \times X$ in X, $(x,y) \to x+y$ (genannt Addition) und einer Abbildung von $\mathbb{R} \times X$ in X, $(\lambda, x) \to \lambda x$ (genannt Multiplikation mit Skalaren) heißt linearer Raum, wenn folgende Eigenschaften erfüllt sind:
(I) $(X,+)$ *ist eine Abelsche Gruppe.*
(II) Für alle $\lambda, \mu \in \mathbb{R}$, $x \in X$ gilt $(\lambda\mu)x = \lambda(\mu x)$ und $(\lambda+\mu)x = \lambda x + \mu x$.
(III) Für alle $\lambda \in \mathbb{R}$ und $x, y \in X$ gilt $\lambda(x+y) = \lambda x + \lambda y$.
(IV) Für alle $x \in X$ gilt $1 \cdot x = x$.

Bemerkung 4.2 *Ein linearer Raum wird häufig auch als "abstrakter" Vektorraum bezeichnet, wobei die Elemente als Vektoren bezeichnet werden. Das Symbol + wird analog zum $\mathbb{R}^m$ in zweierlei Sinn benutzt, nämlich als Addition zwischen Zahlen und zwischen Elementen des linearen Raumes.*

Beispiel 4.3 *(Produkt zweier Vektorräume)*
Wenn X und Y zwei lineare Räume sind, so ist auch ihr kartesisches Produkt

$$X \times Y = \{(x,y) : \ x \in X, y \in Y\}$$

ein linearer Raum mit den "komponentenweise" definierten Operationen

$$\begin{array}{rcl} (x_1,y_1) + (x_2,y_2) & = & (x_1+x_2, y_1+y_2), \\ \lambda(x,y) & = & (\lambda x, \lambda y). \end{array}$$

Analog macht man mit Hilfe der algebraischen Operationen in X das m-fache kartesische Produkt $X \times \cdots \times X$ (m Faktoren) zu einem linearen Raum, und diesen bezeichnet man mit X^m. Offenbar ist die Bezeichnung $\mathbb{R}^m$ ein Spezialfall dessen.

Beispiel 4.4 *(ein Funktionenraum)*
Die Menge $\mathcal{F}(\mathbb{R})$ aller Funktionen $f : \mathbb{R} \to \mathbb{R}$ ist ein linearer Raum mit den "punktweise" definierten Operationen

$$\left.\begin{array}{rcl} (f+g)(x) &=& f(x)+g(x), \\ (\lambda f)(x) &=& \lambda f(x), \end{array}\right\} \text{ für alle } x \in \mathbb{R}.$$

Hier stehen wieder links von den Gleichheitszeichen die Operationen "Addition von Funktionen" und "Multiplikation von Zahlen mit Funktionen", die zu definieren sind, und rechts stehen die bekannten Summen und Produkte von Zahlen. Das Nullelement ist die Funktion, die identisch gleich Null ist.

Beispiel 4.5 *(der lineare Raum aller Polynome)*
$\mathcal{P}(\mathbb{R})$ bezeichne die Menge aller Polynome mit reellen Koeffizienten, d.h.,

$$\mathcal{P}(\mathbb{R}) := \{f \in \mathcal{F}(\mathbb{R}) : \exists m \in \mathbb{N} \exists (a_0, \ldots, a_m) \in \mathbb{R}^{m+1} : f(x) = \sum_{i=0}^{m} a_i x^i, \forall x \in \mathbb{R}\}.$$

$\mathcal{P}(\mathbb{R})$ ist ein linearer Raum mit den Operationen aus $\mathcal{F}(\mathbb{R})$ und es gilt

$$\begin{array}{rcl} \displaystyle\sum_{i=0}^{m} a_i x^i + \sum_{j=0}^{n} b_j x^j &=& \displaystyle\sum_{i=0}^{m} (a_i + b_i) x^i + \sum_{j=m+1}^{n} b_j x^j \\ \lambda(\displaystyle\sum_{i=0}^{m} a_i x^i) &=& \displaystyle\sum_{i=0}^{m} (\lambda a_i) x^i \end{array}$$

wobei ohne Beschränkung der Allgemeinheit $m \leq n$ betrachtet wurde. Bei der Addition von Polynomen bzw. Multiplikation von Polynomen mit Skalaren addieren sich einfach die Koeffizienten bzw. multipliziert sich der Skalar mit den Koeffizienten. Das Nullelement ist das Polynom, dessen Koeffizienten alle Null sind.

Definition 4.6 *(Unterräume)*
Eine Teilmenge U eines linearen Raumes X heißt <u>Unterraum</u> (oder <u>Teilraum</u>) von X, wenn für alle $x, y \in U$ und $\lambda \in \mathbb{R}$ gilt $x + y \in U$ und $\lambda x \in U$. Man kommt also nicht aus U heraus, indem man mit Elementen von U die Operationen "Addition" und "Multiplikation mit Skalaren" ausführt. Insbesondere ist U mit den von X "geerbten" Operationen wieder ein linearer Raum.

Definition 4.7 *(Affine Unterräume)*
Eine Untermenge A eines linearen Raumes X heißt <u>affiner Unterraum</u> von X, wenn ein $x \in X$ und ein Unterraum U von X existieren, so daß gilt $A = \{x + u : u \in U\}$. Dann schreibt man auch $A = x + U$.

Bemerkung 4.8

(i) Ein Unterraum ist auch ein affiner Unterraum, aber ein affiner Unterraum ist im allgemeinen kein Unterraum.

(ii) In der Darstellung eines affinen Unterraumes A mit Hilfe eines Elementes x und eines Unterraumes U ist U eindeutig bestimmt, nicht aber x. Es gilt nämlich $x_1 + U_1 = x_2 + U_2$ genau dann, wenn die Unterräume U_1 und U_2 gleich sind und wenn $x_1 - x_2$ in diesem Unterraum liegt. In diesem Sinne sagt man, daß U "der" Unterraum zu dem affinen Unterraum $A = x + U$ ist, während x "eine" Translation ist, durch die A aus U hervorgeht.

Beispiel 4.9 *(Unterräume und affine Unterräume in $\mathbb{R}^2$)*
In $\mathbb{R}^2$ existieren genau folgende Unterräume: Die trivialen Unterräume $\{(0,0)\}$ und $\mathbb{R}^2$ sowie die Geraden durch den Nullpunkt. Eine Untermenge des $\mathbb{R}^2$ ist genau dann affiner Unterraum, wenn sie ein Punkt oder eine Gerade im $\mathbb{R}^2$ ist.

Beispiel 4.10 *(Unterräume und affine Unterräume im Funktionenraum)*
Der Raum $\mathcal{P}(\mathbb{R})$ der Polynome bildet einen Unterraum des Funktionenraums $\mathcal{F}(\mathbb{R})$ von Funktion $f : \mathbb{R} \to \mathbb{R}$. Die Menge der Polynome, deren Grad nicht größer als eine fixierte natürliche Zahl m ist, ist ein Unterraum von $\mathcal{P}(\mathbb{R})$. Dieser Unterraum wird mit $\mathcal{P}_m(\mathbb{R})$ bezeichnet. Im linearen Raum $\mathcal{F}(\mathbb{R})$ ist ferner die Untermenge $\{f \in \mathcal{F}(\mathbb{R}) : f(0) = 0\}$ ein Unterraum. Die Untermenge $\{f \in \mathcal{F}(\mathbb{R}) : f(0) = 1\}$ ist kein Unterraum, aber ein affiner Unterraum. Bezeichnet man die Funktion, die identisch Eins ist, mit 1, so gilt

$$\{f \in \mathcal{F}(\mathbb{R}) : f(0) = 1\} = 1 + \{f \in \mathcal{F}(\mathbb{R}) : f(x) = 0\}.$$

Definition 4.11 *(Linearkombinationen und lineare Hülle)*
Es seien $x_1, \ldots, x_m$ Elemente eines linearen Raumes X und $\lambda_1, \cdots, \lambda_m$ Zahlen. Dann nennt man die Summe $\sum\limits_{i=1}^{m} \lambda_i x_i$ eine <u>Linearkombination</u> der Elemente $x_1, \cdots, x_m$.
Die Menge aller Linearkombinationen ist ein Unterraum von X, heißt <u>lineare Hülle</u> der Elemente $x_1, \ldots, x_m$ und wird mit $\operatorname{span}\{x_1, \ldots, x_m\}$ bezeichnet, d.h.

$$\operatorname{span}\{x_1, \ldots, x_m\} = \left\{ \sum_{i=1}^{m} \lambda_i x_i : \lambda_i \in \mathbb{R}, i = 1, \ldots, m \right\}.$$

Satz 4.12 *(Durchschnitt und Summe von Unterräumen)*
Wenn U_1 und U_2 zwei Unterräume eines linearen Raumes X sind, so ist auch ihr Durchschnitt $U_1 \cap U_2$ ein Unterraum von X. Ferner ist auch ihre (Minkowski-) Summe

$$U_1 + U_2 = \{u_1 + u_2 : u_1 \in U_1, u_2 \in U_2\}$$

ein Unterraum von X. Dagegen ist die Vereinigung $U_1 \cup U_2$ i.a. kein Unterraum.

<u>Beweis:</u> Für den Durchschnitt ist die Aussage klar. Wir betrachten die Minkowski-Summe $U_1 + U_2$. Seien $x, y \in U_1 + U_2$, $\lambda, \mu \in \mathbb{R}$ und $x_i, y_i \in U_i$ für $i = 1, 2$, so daß $x = x_1 + x_2$ und $y = y_1 + y_2$. Dann gilt $\lambda x_i + \mu y_i \in U_i$ für $i = 1, 2$ und folglich

$$\lambda x + \mu y = (\lambda x_1 + \mu y_1) + (\lambda x_2 + \mu y_2) \in U_1 + U_2.$$

Deshalb ist $U_1 + U_2$ ein Unterraum von X. Als Gegenbeispiel im Fall der Vereinigung zweier Unterräume betrachten wir $X := \mathbb{R}^2$, $U_k := \{(x, kx) \in \mathbb{R}^2 : x \in \mathbb{R}\}$, $k = 1, 2$. Dann enthält $U_1 \cup U_2$ die Punkte $(1, 1)$ und $(1, 2)$, aber nicht $(1, 1) + (1, 2) = (2, 3)$. $\square$

Nach Satz 4.12 gilt insbesondere für Elemente $x_1, \ldots, x_m$ eines linearen Raumes X, daß $\operatorname{span}\{x_1, \ldots, x_m\} = \sum_{i=1}^{m} \operatorname{span}\{x_i\}$.

Folgerung 4.13 *Es seien U_1 und U_2 zwei Unterräume eines Vektorraumes X. Dann gilt $U_1 \cap U_2 = \{0\}$ gdw. für alle $x \in U_1 + U_2$ existieren je genau ein $x_1 \in U_1$ und $x_2 \in U_2$, so daß $x = x_1 + x_2$.*

Beweis: Es sei $x \in U_1 + U_2$ beliebig gewählt.
Annahme: Es existieren $x_1, \tilde{x}_1 \in U_1$ und $x_2, \tilde{x}_2 \in U_2$, so daß $x = x_1 + x_2 = \tilde{x}_1 + \tilde{x}_2$. Es folgt $x_1 - \tilde{x}_1 = \tilde{x}_2 - x_2 \in U_1 \cap U_2 = \{0\}$ und damit $x_1 = \tilde{x}_1$, $x_2 = \tilde{x}_2$.
Die Umkehrung gilt, da jedes $x \in U_1 \cap U_2$, $x \neq 0$, auf zwei verschiedene Weisen darstellbar wäre. $\square$

Definition 4.14 *(Komplement, direkte Summe und Projektor)*
Es seien U_1 und U_2 zwei Unterräume eines Vektorraumes X. Falls $U_1 \cap U_2 = \{0\}$, so heißt $U_1 + U_2$ direkte Summe von U_1 und U_2. Man schreibt dann $U_1 \oplus U_2$. Die Abbildungen von $U_1 \oplus U_2$ in U_1 bzw. in U_2, die durch Folgerung 4.13 definiert sind, heißen Projektoren auf U_1 entlang U_2 bzw. auf U_2 entlang U_1.
Wenn $U_1 \oplus U_2 = X$ ist, so heißen U_1 bzw. U_2 Komplement zu U_2 bzw. U_1 in X.

Beispiel 4.15

a) Wenn U_1 und U_2 zwei Ebenen in $\mathbb{R}^3$ sind, die den Nullpunkt enthalten und die nicht gleich sind, so gilt $U_1 + U_2 = \mathbb{R}^3$, aber $\mathbb{R}^3$ ist nicht die direkte Summe von U_1 und U_2.

b) Wenn U_1 eine Gerade in $\mathbb{R}^2$ durch den Nullpunkt ist, so ist U_2 ein Komplement zu U_1 in $\mathbb{R}^2$ gdw. U_2 eine andere Gerade durch den Nullpunkt ist.

Definition 4.16 *(lineare Abhängigkeit)*
Elemente $x_1, \ldots, x_m$ eines linearen Raumes X heißen linear abhängig (bzw. linear unabhängig), wenn $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ existieren, so daß $\sum_{i=1}^{m} \lambda_i^2 > 0$ und $\sum_{i=1}^{m} \lambda_i x_i = 0$ gilt (bzw. wenn solche Zahlen nicht existieren).

Elemente $x_1, \ldots, x_m$ sind also genau dann linear abhängig, wenn ein $j \in \{1, \ldots, m\}$ existiert, so daß gilt

$$x_j \in \operatorname{span}\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m\}.$$

Beispiel 4.17 *Zwei Vektoren in $\mathbb{R}^2$ sind linear abhängig genau dann, wenn sie auf einer gemeinsamen Geraden, die den Nullpunkt enthält, liegen.*

Satz 4.18 *(Basen und Koordinaten)*
Für Elemente $b_1, \ldots, b_m$ eines linearen Raumes X sind (i) und (ii) äquivalent:

(i) Die Elemente $b_1, \ldots, b_m$ sind linear unabhängig, und gemeinsam mit jedem weiteren Element aus X sind sie linear abhängig.

(ii) Für jedes $x \in X$ existiert genau ein $(\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$, so daß $x = \sum\limits_{i=1}^{m} \lambda_i b_i$ gilt.
In diesem Fall sagt man, daß die Elemente $b_1, \ldots, b_m$ eine Basis in X bilden, und die Zahlen $\lambda_1, \ldots, \lambda_m$ heißen Koordinaten des Elements x bezüglich dieser Basis.

Falls $a_1, \ldots, a_m$ und $b_1, \ldots, b_n$ Basen von X sind, so gilt $m = n$.

Beweis: Aus (ii) folgt sofort (i). Es sei nun (i) erfüllt und $x \in X$ beliebig gewählt. Dann ist $x, b_1, \ldots, b_m$ linear abhängig. Also existieren $\mu_i \in \mathbb{R}$, $i = 1, \ldots, m+1$, so daß $\mu_1 x + \sum\limits_{i=2}^{m+1} \mu_i b_{i-1} = 0$ und $\mu_1 \neq 0$. Mit $\lambda_i := -\frac{\mu_{i+1}}{\mu_1}$, $i = 1, \ldots, m$, ist die Existenz der λ_i, $i = 1, \ldots, m$, gezeigt. Um die Eindeutigkeit zu zeigen, nehmen wir an, daß überdies Zahlen $\overline{\lambda}_i$, $i = 1, \ldots, m$, existieren, so daß gilt

$$x = \sum_{i=1}^{m} \lambda_i b_i = \sum_{i=1}^{m} \overline{\lambda}_i b_i.$$

Daraus folgt aber $\sum\limits_{i=1}^{m} (\lambda_i - \overline{\lambda}_i) b_i = 0$ und deshalb $\lambda_i = \overline{\lambda}_i$, $i = 1, \ldots, m$.
Um die letzte Aussage zu beweisen, machen wir die Annahme, daß o.B.d.A. $n < m$ gilt. Um zu zeigen, daß bereits $\operatorname{span}\{a_1, \ldots, a_n\} = \operatorname{span}\{b_1, \ldots, b_n\} = X$ gilt und damit $a_1, \ldots, a_n, a_{n+1}$ linear abhängig wären, beweisen wir mit Induktion die folgende Aussage: Für jedes $k \in \mathbb{N}_0$ ist $a_1, \ldots, a_k, b_{k+1}, \ldots, b_n$ eine Basis in X.
Für $k = 0$ ist die Aussage richtig, da nach Voraussetzung $b_1, \ldots, b_n$ eine Basis von X ist. Die Aussage sei jetzt für k richtig und wir zeigen sie für $k+1$.
Nach Induktionsvoraussetzung gilt, daß $\lambda_j \in \mathbb{R}$, $j = 1, \ldots, n$ existieren, so daß

$$a_{k+1} = \sum_{i=1}^{k} \lambda_i a_i + \sum_{j=k+1}^{n} \lambda_j b_j$$

gilt. Dabei sei o.B.d.A. $\lambda_{k+1} \neq 0$, ansonsten sind $b_{k+1}, \ldots, b_n$ entsprechend umzunumerieren. Dann gilt $b_{k+1} \in \operatorname{span}\{a_1, \ldots, a_k, a_{k+1}, b_{k+2}, \ldots, b_n\}$ und damit

$$X = \operatorname{span}\{a_1, \ldots, a_k, b_{k+1}, \ldots, b_n\} \subseteq \operatorname{span}\{a_1, \ldots, a_k, a_{k+1}, b_{k+2}, \ldots, b_n\} \subseteq X.$$

Nun ist noch die lineare Unabhängigkeit von $a_1, \ldots, a_k, a_{k+1}, b_{k+2}, \ldots, b_n$ zu zeigen. Dazu betrachten wir die Gleichung

$$0 = \sum_{i=1}^{k} \mu_i a_i + \sum_{j=k+2}^{n} \mu_j b_j + \mu a_{k+1} = \sum_{i=1}^{k} (\mu_i + \mu\lambda_i) a_i + \sum_{j=k+2} (\mu_j + \mu\lambda_j) b_j + \mu\lambda_{k+1} b_{k+1}$$

Daraus folgt sofort $\mu = 0$ (wegen der Induktionsvoraussetzung und $\lambda_{k+1} \neq 0$) und $\mu_i + \mu\lambda_i = \mu_i = 0$, $i = 1, \ldots, k, k+2, \ldots, n$. Damit ist alles gezeigt. □

Beispiel 4.19 *Die Polynome* $1, x, x^2, \dots, x^m$ *bilden eine Basis in* $\mathcal{P}_m(\mathbb{R})$ *(vgl. 4.10). Die Koordinaten eines Polynoms* $\sum_{i=0}^{m} \alpha_i x^i$ *bezüglich dieser Basis sind die reellen Zahlen* $\alpha_0, \dots, \alpha_m$. *Eine andere Basis in* $\mathcal{P}_m(\mathbb{R})$ *bilden z.B. die Polynome* $(x+1)^m, (x+2)^m, \dots, (x+m+1)^m$ *(Übung).*

Definition 4.20 *(Dimension)*
Ist $b_1, \dots, b_m$ *eine Basis in einem linearen Raum* X, *so heißt* m *die Dimension von* X *und wird mit* $\dim X$ *bezeichnet. Konsistent damit definieren wir* $\dim\{0\} = 0$. *Ist* $X \neq \{0\}$ *ein linearer Raum und existieren* $m \in I\!N$ *und Elemente* $b_1, \dots, b_m$ *in* X, *die eine Basis in* X *bilden, so heißt* X *endlich-dimensional, anderenfalls unendlich-dimensional und man schreibt* $\dim X = \infty$.
Unter der Dimension eines affinen Unterraums versteht man die Dimension des entsprechenden Unterraums, d.h. $\dim(x + U) = \dim U$.

Beispiel 4.21 *Es gilt* $\dim \mathcal{P}_m(\mathbb{R}) = m + 1$ *und* $\dim \mathcal{P}(\mathbb{R}) = \infty$, *da die Polynome* $1, x, x^2, \dots, x^m$ *für jedes* $m \in I\!N$ *linear unabhängig sind (nach dem Fundamentalsatz der Algebra über die Anzahl der Nullstellen von Polynomen).*

Folgerung 4.22 *(Dimension des kartesischen Produkts)*
Wenn X *und* Y *zwei endlich-dimensionale lineare Räume sind, so gilt* $\dim(X \times Y) = \dim X + \dim Y$.

Beweis: Sind $a_1, \dots, a_m$ und $b_1, \dots, b_n$ Basen von X bzw. Y, so sind die Elemente $(a_i, 0)$, $(0, b_j)$, $i = 1, \dots, m$, $j = 1, \dots, n$, eine Basis von $X \times Y$. $\square$

Satz 4.23 *(Dimensionsformel für Unterräume)*
Für zwei endlich-dimensionale Unterräume U_1 *und* U_2 *eines linearen Raumes gilt stets* $\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$. *Falls insbesondere* $U_1 \cap U_2 = \{0\}$ *ist, so gilt* $\dim U_1 + \dim U_2 = \dim(U_1 \oplus U_2)$.

Beweis: Es sei $a_1, \dots, a_m$ eine Basis von U_1 und $b_1, \dots, b_n$ eine Basis von U_2. Wir überprüfen, ob die Elemente $a_1, \dots, a_m, b_1$ linear unabhängig sind. Wenn ja, betrachten wir die linear unabhängigen Elemente $a_1, \dots, a_m, b_1$ und stellen die Frage ob auch die Elemente $a_1, \dots, a_m, b_1, b_2$ linear unabhängig sind. Wenn ja setzen wir diesen Prozeß solange fort, bis die Antwort nein ist. Im letzteren Fall fügen wir das entsprechende b_j nicht zu den linear unabhängigen Elementen hinzu und untersuchen b_{j+1}.
Wir nehmen jetzt o.B.d.A. an, daß für die ersten k Elemente $b_1, \dots, b_k$ die Antwort jeweils ja ist, also $a_1, \dots, a_m, b_1, \dots, b_k$ linear unabhängig sind, und daß für jedes weitere b_j, $j = k+1, \dots, n$, die $m+k+1$ Elemente $a_1, \dots, a_m, b_1, \dots, b_k, b_j$ linear abhängig sind. Folglich gilt $\dim(U_1 + U_2) = m + k = \dim U_1 + \dim U_2 - (n-k)$ und es genügt zu zeigen, daß $\dim(U_1 \cap U_2) = n - k$ gilt. Es sei $x \in U_1 \cap U_2$. Dann gilt auch $x \in U_1 + U_2$ und folglich existieren Koeffizienten λ_i, μ_j, $i = 1, \dots, m$, $j = 1, \dots, k$, so daß $x = \sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{k} \mu_j b_j$. Da andererseits x zu U_1 gehört, muß $\mu_j = 0$, $j = 1, \dots, k$, gelten. Da x auch zu U_2 gehört, muß sich x als Linearkombination der Elemente $b_{k+1}, \dots, b_n$ darstellen lassen. Dies sind aber gerade $n - k$ linear unabhängige Elemente, die folglich eine Basis von $U_1 \cap U_2$ darstellen. Der Rest folgt aus Satz 4.12. $\square$

Bemerkung 4.24 *(Reelle und komplexe lineare Räume)*
Alle bisherigen Resultate gelten unverändert, wenn überall $\mathbb{R}$ durch $\mathbb{C}$ ersetzt wird. Man spricht dann von komplexen linearen Räumen (oder linearen Räumen über $\mathbb{C}$) bzw., bei linearen Räumen wie bisher, von reellen linearen Räumen (oder linearen Räumen über $\mathbb{R}$). Allgemeiner spricht man von linearen Räumen über einem Körper $\mathbb{K}$, wobei $\mathbb{K}$ ein beliebiger Körper sein kann. Zur Definition eines linearen Raumes in 4.1 werden nämlich nur die Körper-Eigenschaften von $\mathbb{R}$ verwendet. Im folgenden bezeichne $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$.

Definition 4.25 *(Norm und Skalarprodukt)*
Es sei X ein linearer Raum. Eine Norm in X ist eine Abbildung $\|\cdot\| : X \to \mathbb{R}$, so daß für alle $x, y \in X$ und $\lambda \in \mathbb{K}$ gilt:

$$\begin{aligned} \|x\| &\geq 0 \quad \textit{und} \quad \|x\| = 0 \textit{ gdw. } x = 0, \\ \|\lambda x\| &= |\lambda|\|x\|, \\ \|x + y\| &\leq \|x\| + \|y\| \quad (\textit{Dreiecksungleichung}). \end{aligned}$$

Ein linearer Raum X über $\mathbb{K}$ mit Norm heißt normierter Raum.
Ein Skalarprodukt in X ist eine Abbildung $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$, so daß für alle $x, y, z \in X$ und $\lambda, \mu \in \mathbb{K}$ gilt:

$$\begin{aligned} \langle \lambda x + \mu y, z\rangle &= \lambda\langle x, z\rangle + \mu\langle y, z\rangle, \\ \langle x, y\rangle &= \overline{\langle y, x\rangle}, \\ \langle x, x\rangle &> 0, \text{ falls } x \neq 0. \end{aligned}$$

(Hierbei ist $\overline{\langle y, x\rangle}$ die zu $\langle y, x\rangle$ konjugiert komplexe Zahl.)
Ein linearer Raum über $\mathbb{K}$ mit Skalarprodukt heißt unitärer oder Euklidischer Raum. Zwei Elemente $x, y \in X$ heißen orthogonal, wenn $\langle x, y\rangle = 0$ gilt.

Satz 4.26 *Es sei X ein unitärer Raum mit Skalarprodukt $\langle\cdot,\cdot\rangle$.*
Dann ist $\|x\| := \sqrt{\langle x, x\rangle}$, $\forall x \in X$, eine Norm in X und es gilt die Cauchy-Schwarzsche Ungleichung

$$|\langle x, y\rangle| \leq \|x\|\|y\| \quad \forall x, y \in X.$$

In der Cauchy-Schwarzsche Ungleichung gilt die Gleichheit gdw. die Elemente x und y linear abhängig sind. Überdies gilt für beliebige $x \in X$, daß $\langle x, 0\rangle = \langle 0, x\rangle = 0$.

Beweis: Aus den Eigenschaften des Skalarproduktes folgt für jedes $x \in X$, daß $\langle x, 0\rangle = \langle 0, x\rangle = 0$. Dies impliziert auch die erste Normeigenschaft. Die zweite Normeigenschaft folgt sofort aus der Definition eines Skalarproduktes. Die Dreiecksungleichung beweist man wie in Satz 2.35. Dort wurde die konkrete Struktur des Euklidischen Skalarproduktes nicht benutzt.
Zum Beweis der Cauchy-Schwarzschen Ungleichung gehen wir analog zum Beweis von Satz 2.34 vor. Für beliebige $x, y \in X$ und $\alpha, \beta \in \mathbb{K}$ folgt aus den ersten beiden Skalarprodukt-Eigenschaften:

$$0 \leq \|\alpha x + \beta y\|^2 = |\alpha|^2\|x\|^2 + \alpha\overline{\beta}\langle x, y\rangle + \overline{\alpha}\beta\overline{\langle x, y\rangle} + |\beta|^2\|y\|^2.$$

In dieser Identität setzen wir $\alpha := -\frac{\overline{\langle x,y\rangle}}{\|x\|}$ und $\beta := \|x\|$ und erhalten

$$0 \leq |\langle x, y\rangle|^2 - 2|\langle x, y\rangle|^2 + \|x\|^2\|y\|^2,$$

d.h. die behauptete Ungleichung. Es seien nun $x, y \in X$ derart, daß die Gleichheit $|\langle x, y\rangle| = \|x\|\|y\|$ gilt. Man zeigt dann $-\|y\|x + \|x\|y = 0$ analog zum Beweis von 2.34. □

Beispiel 4.27 *(Das Hermitesche Skalarprodukt)*
Das sogenannte Hermitesche Skalarprodukt in $\mathbb{C}^m$ ist definiert durch

$$\langle(x_1, \ldots, x_m), (y_1, \ldots, y_m)\rangle := \sum_{j=1}^{m} x_j\overline{y_j}.$$

Die Einschränkung dieses Skalarproduktes auf $\mathbb{R}^m$ ergibt gerade das Euklidische Skalarprodukt. Wenn im folgenden ein Skalarprodukt in $\mathbb{C}^m$ benutzt wird, so ist stets das Hermitesche Skalarprodukt gemeint.

Folgerung 4.28 *(Winkel)*
Es sei X ein unitärer Raum über $\mathbb{K} = \mathbb{R}$. Für alle $x, y \in X$ mit $x \neq 0$ und $y \neq 0$ existiert genau ein $\varphi \in [0, \pi]$ mit

$$\cos\varphi = \frac{\langle x, y\rangle}{\|x\|\|y\|},$$

der sogenannte Winkel zwischen den Vektoren x und y. Überdies gilt

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|\cos\varphi.$$

Beweis: Nach Satz 4.26 gilt für alle von 0 verschiedenen $x, y \in X$, daß $\frac{|\langle x,y\rangle|}{\|x\|\|y\|} \leq 1$. Deshalb existiert genau ein $\varphi \in [0, \pi]$, so daß

$$\cos\varphi = \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

Der Rest folgt durch Ausrechnen von $\|x + y\|^2$. □

Definition 4.29 *(Orthonormalsysteme)*
Es sei X ein unitärer Raum und $I \subseteq \mathbb{N}$ höchstens abzählbar. Eine Familie $(e_i)_{i\in I}$ von Elementen in X heißt Orthonormalsystem, falls $\|e_i\| = 1$ und $\langle e_i, e_j\rangle = 0$ für alle $i, j = 1, \ldots, m$ und $i \neq j$.

Natürlich ist die kanonische Basis der Einheitsvektoren im $\mathbb{R}^m$ eine Orthonormalsystem (vgl. Bem. 2.33).

Satz 4.30 *(Orthonormalisierung)*
Ist $(e_i)_{i\in I}$ ein Orthonormalsystem und $\tilde I$ eine endliche Teilmenge von I, so sind die Elemente $\{e_i : i \in \tilde I\}$ linear unabhängig. Ist I endlich und gilt $\operatorname{span}\{e_i : i \in I\} = X$, *so gilt für alle $x \in X$:*

$$x = \sum_{j=1}^{m} \langle x, e_j\rangle e_j \quad \textit{und} \quad \|x\|^2 = \sum_{j=1}^{m} |\langle x, e_j\rangle|^2.$$

Sind $b_1, \ldots, b_m$ linear unabhängig in X, so entsteht durch

$$e_i := \frac{b_i - \sum\limits_{k=1}^{i-1} \langle b_i, e_k\rangle e_k}{\|b_i - \sum\limits_{k=1}^{i-1} \langle b_i, e_k\rangle e_k\|}, \; i = 1, \ldots, m,$$

ein Orthonormalsystem $e_1, \ldots, e_m$ (Orthonormalisierungsverfahren nach Schmidt).

Beweis: Sei $\tilde I$ eine endliche Teilmenge von I. Wir betrachten die Gleichung

$$\sum_{i\in \tilde I} \lambda_i e_i = 0$$

mit $\lambda_i \in \mathbb{K}$, $i \in I$, und multiplizieren sie skalar mit e_j, wobei $j \in \tilde I$. Dann gilt

$$0 = \sum_{i\in\tilde I} \lambda_i \langle e_i, e_j\rangle = \lambda_j \langle e_j, e_j\rangle = \lambda_j \,.$$

Wiederholen wir das für jedes $j \in \tilde I$, so folgt, daß alle Koeffizienten λ_j verschwinden. Damit ist die erste Aussage gezeigt.
Es sei jetzt I endlich und es gelte $\operatorname{span}\{e_i : i \in I\} = X$. Dann ist $\{e_i : i \in I\}$ eine Basis in X und es gilt für jedes $x \in X$, daß $\mu_i \in \mathbb{K}$, $i \in I$, existieren, so daß $x = \sum\limits_{i\in I} \mu_i e_i$. Durch skalare Multiplikation mit e_j im Sinne von entsteht $\mu_j = \langle x, e_j\rangle$ für jedes $j \in I$. Weiterhin gilt:

$$\|x\|^2 = \langle \sum_{i\in I} \langle x, e_i\rangle e_i, \sum_{j\in I} \langle x, e_j\rangle e_j\rangle = \sum_{i,j\in I} \langle x, e_i\rangle \overline{\langle x, e_j\rangle} \langle e_i, e_j\rangle = \sum_{i\in I} |\langle x, e_i\rangle|^2.$$

Es sei jetzt $b_1, \ldots, b_m$ linear unabhängig in X und wir betrachten $\hat e_i \in X$ definiert durch

$$\hat e_i := b_i - \sum_{k=1}^{i-1} \langle b_i, e_k\rangle e_k = b_i - \sum_{k=1}^{i-1} \frac{1}{\|\hat e_k\|^2} \langle b_i, \hat e_k\rangle \hat e_k$$

für $i = 1, \ldots, m$. Es gilt also $e_i := \frac{\hat e_i}{\|\hat e_i\|}$ für $i = 1, \ldots, m$.
Um zu zeigen, daß die Elemente $e_1, \ldots, e_m$ orthonormal sind, genügt es, $\langle \hat e_i, \hat e_j\rangle = 0$ für $j = 1, \ldots, i-1$ und $i = 1, \ldots, m$ zu beweisen. Dazu führen wir eine Induktion über i durch. Für $i = 1$ ist die Aussage trivial. Es gelte jetzt $\langle \hat e_i, \hat e_j\rangle = 0$ für $j = 1, \ldots, i-1$ und wir betrachten für $j \in \{1, \ldots, i\}$

$$\langle \hat e_{i+1}, \hat e_j\rangle = \langle b_{i+1}, \hat e_j\rangle - \sum_{k=1}^{i} \frac{1}{\|\hat e_k\|^2} \langle b_{i+1}, \hat e_k\rangle \langle \hat e_k, \hat e_j\rangle = \langle b_{i+1}, \hat e_j\rangle - \frac{1}{\|\hat e_j\|^2} \langle b_{i+1}, \hat e_j\rangle \langle \hat e_j, \hat e_j\rangle = 0.$$

Damit ist alles gezeigt. $\square$

Satz 4.31 *(Isometrie mit $\mathbb{K}^m$)*
Es sei X ein normierter Raum über $\mathbb{K}$ mit Norm $\|\cdot\|$ und $\dim X = m$. Dann existieren eine Norm $\|\cdot\|_*$ auf $\mathbb{K}^m$ und eine bijektive Abbildung $F : X \to \mathbb{K}^m$, so daß für alle $x \in X$ die "Isometrie" $\|F(x)\|_* = \|x\|$ gilt.

Beweis: Es sei $b_1, \ldots, b_m$ eine Basis in X. Wir definieren eine Abbildung $F : X \to \mathbb{R}^m$ durch $F(\sum\limits_{i=1}^{m} \lambda_i b_i) := (\lambda_1, \ldots, \lambda_m)$. Diese Abbildung ist eindeutig nach Satz 4.18 und surjektiv wegen $X = \mathrm{span}\{b_1, \ldots, b_m\}$. Die inverse Abbildung ist ebenfalls eindeutig, d.h. insgesamt ist F bijektiv. Wir definieren jetzt die Norm $\|\cdot\|_*$ auf $\mathbb{K}^m$ durch $\|(\lambda_1, \ldots, \lambda_m)\|_* := \|\sum\limits_{i=1}^{m} \lambda_i b_i\|$ für alle $(\lambda_1, \ldots, \lambda_m) \in \mathbb{K}^m$. Da sich sämtliche Normeigenschaften von $\|\cdot\|_*$ sofort aus denen von $\|\cdot\|$ ergeben, ist alles gezeigt. $\square$

Satz 4.31 bedeutet, daß wir Eigenschaften von Elementen eines normierten Raumes X, die unter Verwendung der Norm definiert sind, sofort als entsprechende Eigenschaften der Koeffizienten der Elemente bzgl. der Basis $b_1, \ldots, b_m$ interpretieren und auf die Ergebnisse in Kapitel 2 zurückgreifen können. So ist eine beschränkte Menge in X relativ kompakt durch Anwendung von Satz 2.59.

Lemma 4.32 *Es seien U, U_1 und U_2 Unterräume eines unitären Raums X.*

(i) Die Menge $U^\perp = \{x \in X : \langle x, u\rangle = 0, \forall u \in U\}$ ist ein Unterraum in X und es gilt $U \cap U^\perp = \{0\}$.

(ii) $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$.

Beweis: (i) Die Menge $U^\perp$ ist ein linearer Raum und damit ein Unterraum von X. Es sei nun $x \in U \cap U^\perp$. Nach Definition von $U^\perp$ muß deshalb $\langle x, x\rangle = \|x\|^2 = 0$ gelten, d.h., $x = 0$.
(ii) Beide Inklusionen folgen sofort aus den Definitionen (Übung). $\square$

Satz 4.33 *(orthogonaler Projektor und orthogonales Komplement)*
Es sei U ein endlich-dimensionaler Unterraum eines unitären Raumes X.
Dann existiert für jedes $x \in X$ genau ein $\bar{u} \in U$, so daß gilt

$$\|x - \bar{u}\| = \min\{\|x - u\| : u \in U\}.$$

Überdies gilt $x - \bar{u} \in U^\perp$ und $X = U \oplus U^\perp$.
(Das Element $\bar{u}$ heißt orthogonale Projektion von x auf U der Projektor von X auf U entlang $U^\perp$ heißt orthogonaler Projektor. $U^\perp$ heißt orthogonales Komplement zu U und $U \oplus U^\perp$ orthogonale Summe.)
Ist $e_1, \cdots, e_m$ eine Orthonormalbasis in U, so gilt $\bar{u} = \sum\limits_{j=1}^{m} \langle x, e_j\rangle e_j$.

Beweis: Es sei $x \in X$ und (u_n) eine Folge in U mit der Eigenschaft

$$\lim_{n\to\infty} \|x - u_n\| = \inf\{\|x - u\| : u \in U\}.$$

Diese Folge ist beschränkt. Da U endlichdimensional ist, läßt sich Satz 2.59 anwenden. Deshalb besitzt (u_n) eine konvergente Teilfolge (u_{n_k}). Für deren Grenzwert $\bar{u}$ gilt:

$$\|x - \bar{u}\| = \lim_{k\to\infty} \|x - u_{n_k}\| = \inf\{\|x - u\| : u \in U\}.$$

Wir zeigen jetzt, daß $x - \bar{u} \in U^{\perp}$. Für jedes $\alpha \in \mathbb{C}$ und jedes $u \in U$ gilt:

$$\|x - \bar{u}\|^2 \leq \|x - (\bar{u} + \alpha u)\|^2 = \|x - \bar{u}\|^2 - \overline{\alpha}\langle x - \bar{u}, u\rangle - \alpha\overline{\langle x - \bar{u}, u\rangle} + |\alpha|^2\|u\|^2.$$

Wir setzen $\alpha := t\langle x - \bar{u}, u\rangle$ mit $t > 0$ und erhalten aus der obigen Ungleichung

$$0 \leq -2t|\langle x - \bar{u}, u\rangle|^2 - t^2|\langle x - \bar{u}, u\rangle|^2\|u\|^2.$$

Mittels Division durch t und anschließendem Grenzübergang für $t \to 0$ folgt daraus $|\langle x - \bar{u}, u\rangle| = 0$. Da dies für beliebiges $u \in U$ gilt, folgt $x - \bar{u} \in U^{\perp}$. Wegen $x = \bar{u} + (x - \bar{u})$ folgt $X = U + U^{\perp}$ und damit $X = U \oplus U^{\perp}$ wegen Lemma 4.31. Da die Eigenschaft $x - \bar{u} \in U^{\perp}$ nur aus der Tatsache hergeleitet wurde, daß $\bar{u}$ eine Lösung von $\min\{\|x - u\| : u \in U\}$ ist, ist nach Folgerung 4.13 auch die Eindeutigkeit von $\bar{u}$ klar. Ist schließlich $e_1, \cdots, e_m$ eine Orthonormalbasis in U, so folgt $\langle x - \bar{u}, e_j\rangle = 0$ für jedes $j = 1, \ldots, m$ und damit wegen Satz 4.30 auch die letzte Aussage. □

Definition 4.34 *Es sei $(e_i)_{i\in\mathbb{N}}$ ein Orthonormalsystem im unitären Raum X. Dann heißt die Reihe*

$$\sum_{k=1}^{\infty}\langle x, e_k\rangle e_k$$

Fourierreihe von $x \in X$ und $\langle x, e_k\rangle \in \mathbb{K}$ heißt k-ter Fourierkoeffizient. Das Orthonormalsystem heißt vollständig, falls der kleinste abgeschlossene Teilraum U von X, der $(e_i)_{i\in\mathbb{N}}$ enthält, gleich X ist.

Dabei definieren wir mit der Norm $\|\cdot\|$ in X wie in Kap. 2 die Konvergenz von Folgen in X, die Offenheit, Abgeschlossenheit usw. von Teilmengen von X.

Satz 4.35 *(Fourierreihen)*
Es sei $(e_i)_{i\in\mathbb{N}}$ ein Orthonormalsystem im unitären Raum X.
Für jedes $x \in X$ ist die Folge $(\sum_{k=1}^{n}\langle x, e_k\rangle e_k)_{n\in\mathbb{N}}$ der Partialsummen der Fourierreihe eine Fundamentalfolge in X. Ist das Orthonormalsystem vollständig, so konvergiert die Fourierreihe gegen x, d.h.

$$x = \sum_{k=1}^{\infty}\langle x, e_k\rangle e_k\,.$$

Überdies gilt: $\|x\|^2 = \sum\limits_{k=1}^{\infty} |\langle x, e_k\rangle|^2$ (Parsevalsche Gleichung).

Beweis: Sei $x \in X$. Wir definieren $s_n := \sum_{k=1}^n \langle x, e_k\rangle e_k$, $\forall n \in I\!N$, und erhalten zunächst wie in Satz 4.30

$$\begin{aligned} \|x - s_n\|^2 &= \langle x - s_n, x - s_n\rangle = \|x\|^2 - \langle x, s_n\rangle - \langle s_n, x\rangle + \|s_n\|^2 \\ &= \|x\|^2 - 2\sum_{k=1}^n |\langle x, e_k\rangle|^2 + \sum_{k=1}^n |\langle x, e_k\rangle|^2 \\ &= \|x\|^2 - \sum_{k=1}^n |\langle x, e_k\rangle|^2 = \|x\|^2 - \|s_n\|^2 \,. \end{aligned}$$

Insbesondere gilt $\|s_n\|^2 = sum_{k=1}^n |\langle x, e_k\rangle|^2 \le \|x\|^2$. Für $m > n$ erhalten wir nun

$$\|s_m - s_n\|^2 = \| \sum_{k=n+1}^m \langle x, e_k\rangle e_k\|^2 = \sum_{k=n+1}^m |\langle x, e_k\rangle|^2 \,.$$

Da die Folge $(\sum_{k=1}^n |\langle x, e_k\rangle|^2)_{n\in I\!N}$ monoton wachsend und beschränkt ist, ist sie konvergent. Deshalb wird $\sum_{k=n+1}^m |\langle x, e_k\rangle|^2$ beliebig klein, wenn nur $n < m$ hinreichend groß gewählt wird. Also ist $(s_n)_{n\in I\!N}$ eine Fundamentalfolge in X. Schließlich betrachten wir die Unterräume $U_n := \text{span}\{e_1, \ldots, e_n\}$ für jedes $n \in I\!N$. Diese sind monoton wachsend bzgl. $\subseteq$. Ist nun das Orthonormalsystem vollständig, so ist die Abschließung der Vereinigung $\bigcup_{n\in I\!N} U_n$ gleich X. Deshalb existiert eine Folge (x_n) in $x_n \in U_n$, $n \in I\!N$, mit $x_n \to x$. Da nach Satz 4.32 s_n gerade die orthogonale Projektion von x auf U_n ist, gilt: $\|x - s_n\| \le \|x - x_n\|$, $\forall n \in I\!N$. Deshalb konvergiert (s_n) gegen x. Die Parsevalsche Gleichung folgt schließlich aus der Identidät $\|x - s_n\|^2 = \|x\|^2 - \|s_n\|^2$ für jedes $n \in I\!N$. Damit ist alles gezeigt. □

4.2 Lineare Abbildungen

Definition 4.36 *Es seien X und Y zwei lineare Räume über (demselben Körper) $\mathbb{K}$. Eine Abbildung $F : X \to Y$ heißt linear, wenn gilt*

$$F(\lambda x + \mu y) = \lambda F(x) + \mu F(y) \quad \text{für alle } \lambda, \mu \in \mathbb{K} \text{ und } x, y \in X.$$

Die Menge aller linearen Abbildungen von X nach Y wird mit $\mathcal{L}(X;Y)$ bezeichnet. Anstelle von $\mathcal{L}(X;X)$ schreibt man auch $\mathcal{L}(X)$. Mit $I \in \mathcal{L}(X)$ bezeichnet man die identische Abbildung in X.

Beispiel 4.37

a) *$F : \mathbb{R} \to \mathbb{R}^m$ ist linear, falls F die Gestalt $F(x) = xa$, $\forall x \in \mathbb{R}$, mit $a \in \mathbb{R}^m$ hat.*

b) *$F : \mathbb{R}^m \to \mathbb{R}$ ist linear, falls F die Form $F(x) = \langle x, a\rangle$, $\forall x \in \mathbb{R}^m$, mit $a \in \mathbb{R}^m$ hat.*

Eigenschaften 4.38

(i) $\mathcal{L}(X;Y)$ *ist ein linearer Raum über* $\mathbb{K}$ *mit den Operationen*

$$\left.\begin{array}{rcl}(F+G)(x) & = & F(x)+G(x),\\ (\lambda F)(x) & = & \lambda F(x).\end{array}\right\} \text{ für alle } x \in X, \lambda \in \mathbb{K}.$$

Das Nullelement in $\mathcal{L}(X;Y)$ *ist die Abbildung, die alle* $x \in X$ *auf das Nullelement in* Y *abbildet. Wenn* X *und* Y *endlich-dimensional sind, so gilt*

$$\dim \mathcal{L}(X;Y) = \dim X \cdot \dim Y.$$

(ii) *Wenn* $F : X \to Y$ *und* $G : Y \to Z$ *zwei lineare Abbildungen sind, so ist auch ihre Verkettung* $G \circ F : X \to Z$ *linear.*

(iii) *Wenn* $F : X \to Y$ *linear ist, so sind die Mengen* $\{x \in X : F(x) = 0\}$ *und* $R(A) = \{F(x) : x \in X\}$ *Unterräume von* X *bzw.* Y.

(iv) *Wenn* $F : X \to Y$ *linear und bijektiv ist, so ist auch* $F^{-1} : Y \to X$ *linear.*

(v) *Für fixiertes* $G \in \mathcal{L}(X)$ *ist die Abbildung* $F \mapsto F \circ G$ *von* $\mathcal{L}(X)$ *in* $\mathcal{L}(X)$ *linear. Dagegen ist die Abbildung* $(F, G) \mapsto F \circ G$ *von* $\mathcal{L}(X) \times \mathcal{L}(X)$ *in* $\mathcal{L}(X)$ *i.a. nicht linear.*

Beweis:

(i) Die Axiome eines linearen Raumes sind für $\mathcal{L}(X;Y)$ erfüllt. Es gelte nun $\dim X = m$ und $\dim Y = n$. Sei $F \in \mathcal{L}(X;Y)$ und $a_1, \ldots, a_m$ bzw. $b_1, \ldots, b_n$ Basen aus X bzw. Y. Wir definieren $\lambda_i(x)$ als den Koeffizienten von a_i in der Basisdarstellung von $x \in X$. Diese Koeffizienten $\lambda_i(\cdot)$ stellen lineare Abbildungen von X in $\mathbb{K}$ dar. Wir definieren nun lineare Abbildungen $F_{ij} \in \mathcal{L}(X;Y)$ durch $F_{ij}(x) := \lambda_i(x) b_j$ für $i = 1, \ldots, m$, $j = 1, \ldots, n$. Diese Abbildungen bilden eine Basis von $\mathcal{L}(X;Y)$ (Übung).

(iv) Es seien $y, \tilde{y} \in Y$ und $\lambda, \mu \in \mathbb{K}$. Ferner seien $x, \tilde{x} \in X$ so gewählt, daß $F(x) = y$ und $F(\tilde{x}) = \tilde{y}$. Dann gilt $F(\lambda x + \mu \tilde{x}) = \lambda F(x) + \mu F(\tilde{x}) = \lambda y + \mu \tilde{y}$ und folglich $F^{-1}(\lambda y + \mu \tilde{y}) = \lambda x + \mu \tilde{x} = \lambda F^{-1}(y) + \mu F^{-1}(\tilde{y})$.

(ii) , *(iii)*, *(v)* (Übung). □

Definition 4.39 *(Nullraum, Defekt und Rang linearer Abbildungen)*
Für $F \in \mathcal{L}(X;Y)$ *heißt der Unterraum* $N(F) = \{x \in X : F(x) = 0\}$ *Nullraum bzw. Kern von* F. *Die natürlichen Zahlen* $d(F) := \dim N(F)$ *und* $r(F) := \dim R(F)$ *heißen Defekt bzw. Rang von* F.

Satz 4.40 *(Dimensionsformel für lineare Abbildungen)*
Es seien $F \in \mathcal{L}(X;Y)$ *und* X *endlichdimensional. Dann ist*

$$\dim N(F) + \dim R(F) = \dim X \quad \textit{bzw.} \quad d(F) + r(F) = \dim X.$$

Wenn $\dim X = \dim Y$ *ist, so gilt insbesondere, daß* F *genau dann injektiv (d.h.* $N(F) = \{0\}$*) ist, wenn* F *surjektiv (d.h.* $R(F) = Y$*) ist.*
Mit anderen Worten, ist $\dim X = \dim Y$, *so gilt die Alternative: Entweder die homogene Gleichung* $F(x) = 0$ *besitzt eine Lösung* $x \neq 0$ *oder die inhomogene Gleichung* $F(x) = y$ *besitzt für jedes* $y \in Y$ *genau eine Lösung* $x \in X$.

Beweis: Es sei $\dim X = m$ und $\dim N(F) = n \leq m$. Die Elemente $b_1, \dots, b_n$ seien eine Basis für $N(F)$ und $a_1, \dots, a_m$ eine Basis für X. Wir gehen jetzt vor wie im Beweis von Satz 4.23 und überprüfen, ob $b_1, \dots, b_n, a_1$ linear unabhängig sind. Ist dies der Fall setzen wir mit $b_1, \dots, b_n, a_1$ fort und fragen, ob $b_1, \dots, b_n, a_1, a_2$ linear unabhängig sind. Ist dies nicht der Fall, so untersuchen wir als nächstes die lineare Unabhängigkeit von $b_1, \dots, b_n, a_2$. Diesen Prozeß setzen wir fort und erhalten die linear unabhängigen Elemente $b_1, \dots, b_n, a_{i_1}, \dots, a_{i_k}$, deren lineare Unabhängigkeit bei Hinzunahme eines a_j mit $j \in \{1, \dots, m\} \setminus \{i_1, \dots, i_k\}$ verloren geht. Nach Konstruktion muß deshalb $k = m-n$ gelten und $b_1, \dots, b_n, a_{i_1}, \dots, a_{i_k}$ eine Basis von X sein. Wir betrachten die Bilder $F(a_{i_j}) \in R(F)$ für $j = 1, \dots, m-n$ und untersuchen die Gleichung $\sum\limits_{j=1}^{m-n} \lambda_j F(a_{i_j}) = 0$ mit $\lambda_j \in \mathbb{K}$, $j = 1, \dots, m-n$. Wegen der Linearität von F gilt

$$F(\sum_{j=1}^{m-n} \lambda_j a_{i_j}) = 0 \quad \text{d.h.} \quad x := \sum_{j=1}^{m-n} \lambda_j a_{i_j} \in N(F).$$

Stellt man x als Linearkombination der Basis $b_1, \dots, b_n$ von $N(F)$ dar, so wird klar, daß dies der linearen Unabhängigkeit von $b_1, \dots, b_n, a_{i_1}, \dots, a_{i_k}$ widerspricht, falls gilt $(\lambda_1, \dots, \lambda_{m-n}) \neq (0, \dots, 0)$. Folglich muß $\lambda_j = 0$, $j = 1, \dots, m-n$ gelten und die $F(a_{i_1}), \dots, F(a_{i_{m-n}})$ sind linear unabhängig. Es sei nun $y \in R(F)$ beliebig gewählt und es sei $x \in X$ mit $F(x) = y$. Dann läßt sich x als Linearkombination der Basiselemente mit Koeffizienten μ_i, $\imath = 1, \dots, m$, darstellen und es gilt

$$y = \sum_{i=1}^{n} \mu_i F(b_i) + \sum_{j=1}^{m-n} \mu_{n+j} F(a_{i_j}) = \sum_{j=1}^{m-n} \mu_{n+j} F(a_{i_j}).$$

Also ist $F(a_{i_1}), \dots, F(a_{i_{m-n}})$ eine Basis in $R(F)$ und es gilt $\dim R(F) = m-n$. Der letzte Teil der Aussage folgt daraus sofort. □

Folgerung 4.41 *(Lösungsmengen inhomogener linearer Gleichungen)*
Wenn X endlichdimensional, $F \in \mathcal{L}(X;Y)$ und $y \in Y$, so ist die Menge aller Lösungen $x \in X$ der Gleichung $F(x) = y$ ein affiner Unterraum der Dimension $\dim X - r(F)$. *Es gilt nämlich* $\quad \{x \in X : F(x) = y\} = x_0 + N(F) \quad$ *gdw.* $\quad F(x_0) = y$.

Beweis: $(\Longrightarrow)$ Gilt $\{x \in X : F(x) = y\} = x_0 + N(F)$, so muß $F(x_0) = y$ gelten.
$(\Longleftarrow)$ Gilt $F(x_0) = y$, so auch $x_0 + N(F) \subseteq \{x \in X : F(x) = y\}$. Ist x eine Lösung von $F(x) = y$, so gilt $F(x - x_0) = 0$, d.h. $x - x_0 \in N(F)$ oder $x \in x_0 + N(F)$.
Die Dimension von $N(F)$ ist nach Satz 4.38 aber gerade $\dim X - r(F)$. □

Bemerkung 4.42 *Den Sachverhalt in Folgerung 4.39 formuliert man bisweilen auch wie folgt: Die allgemeine Lösung der inhomogenen Gleichung, nämlich $\{x \in X : F(x) = y\}$, ist gleich der Summe aus einer partikulären Lösung der inhomogenen Gleichung, nämlich x_0, und der allgemeinen Lösung der homogenen Gleichung, nämlich $N(F)$.*

Folgerung 4.43 *(Projektoren)*
Es seien X ein linearer Raum und U_1, U_2 Unterräume von X mit $X = U_1 \oplus U_2$.

(i) Ist P der Projektor auf U_1 entlang U_2, so ist P linear, es gilt $P \circ P = P$, und $I - P$ ist der Projektor auf U_2 entlang U_1.

(ii) Erfüllt $P \in \mathcal{L}(X)$ die Gleichung $P \circ P = P$, so ist P der Projektor auf $R(P)$ entlang $N(P)$.

Beweis:

(i) Seien $x, y \in X$ und $\lambda, \mu \in \mathbb{K}$. Dann existieren eindeutig bestimmte $x_1, y_1 \in U_1$ und $x_2, y_2 \in U_2$, so daß $x = x_1 + x_2$ und $y = y_1 + y_2$, und es gilt $P(x) = x_1$ und $P(y) = y_1$ sowie $(I - P)(x) = x - x_1 = x_2$. Also ist $I - P$ der Projektor auf U_2 entlang U_1. Überdies folgt $\lambda x_1 + \mu y_1 \in U_1$ und $P(\lambda x + \mu y) = \lambda x_1 + \mu y_1 = \lambda P(x_1) + \mu P(y_1)$. Damit ist P linear. Weiterhin gilt $P(u) = u$ für alle $u \in U_1$ und damit $P \circ P(x) = P(P(x)) = P(x)$ für alle $x \in X$.

(ii) Alles ist gezeigt, wenn $X = R(P) \oplus N(P)$ gilt. Für $x \in X$ betrachten wir die Darstellung $x = P(x) + (x - P(x))$. Wegen $P(x) \in R(P)$, bleibt $x - P(x) \in N(P)$ zu zeigen. Da P linear ist und $P \circ P = P$ gilt, folgt $P(x - P(x)) = P(x) - P(P(x)) = 0$. Die Darstellung $x = P(x) + (x - P(x))$ für $x \in X$ ist eindeutig, da $R(P) \cap N(P) = \emptyset$ gilt (vgl. 4.14). □

4.3 Matrizen

Definition 4.44 (Matrix)
Eine $m \times n$-Matrix über $\mathbb{K}$ ist eine Anordnung

$$A = [a_{ij}] = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

mit $a_{ij} \in \mathbb{K}$. Die Zahlen a_{ij} heißen Koeffizienten von A. Die m-Tupel bzw. n-Tupel

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad \text{bzw.} \quad \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{bmatrix}$$

heißen Spalten bzw. Zeilen von A. Wenn $a_1, a_2, \dots, a_n$ die Spalten der Matrix A sind, so schreibt man auch

$$A = [a_1\, a_2\, \dots a_n].$$

Die Menge aller $m \times n$-Matrizen über $\mathbb{K}$ wird mit $\mathbb{K}^{m \times n}$ bezeichnet. Elemente von $\mathbb{K}^{m \times 1}$ bzw. $\mathbb{K}^{1 \times n}$ werden mit Spalten- bzw. Zeilenvektoren aus $\mathbb{K}^m$ bzw. $\mathbb{K}^n$ identifiziert. Matrizen mit gleicher Spalten- bzw. Zeilenzahl heißen quadratisch. Unter der Hauptdiagonalen einer quadratischen Matrix versteht man die Elemente $a_{11}, a_{22}, \dots, a_{nn}$. Quadratische Matrizen, bei denen alle Elemente außerhalb der Hauptdiagonalen gleich Null sind, heißen Diagonalmatrizen. Die Diagonalmatrix mit Einsen auf der Diagonale heißt Einheitsmatrix und wird mit I bezeichnet. Häufig schreibt man auch $I = [\delta_{ij}]$ (mit dem Kronecker-Symbol δ_{ij}).

Eigenschaften 4.45 *(Addition und Multiplikation mit Zahlen)*
$\mathbb{K}^{m\times n}$ ist ein linearer Raum über $\mathbb{K}$ mit den Operationen

$$\begin{aligned} [a_{ij}] + [b_{ij}] &= [a_{ij} + b_{ij}], \\ \lambda[a_{ij}] &= [\lambda a_{ij}]. \end{aligned}$$

Das Nullelement in $\mathbb{K}^{m\times n}$ ist die Matrix, deren Koeffizienten alle Null sind, und sie wird auch mit dem Symbol 0 bezeichnet. Ferner gilt $\dim \mathbb{K}^{m\times n} = m \cdot n$.

Beweis: Dies ist ein Analogon zu 4.36 (i) (vgl. auch Satz 4.50). Eine Basis in $\mathbb{K}^{m\times n}$ sind die mn Matrizen E_{ij}, deren Koeffizienten mit Ausnahme einer 1 in der i-ten Zeile und j-ten Spalte gleich Null sind. □

Definition 4.46 *(Multiplikation von Matrizen)*
Es seien $A = [a_{ij}] \in \mathbb{K}^{\ell\times m}$ und $B = [b_{jk}] \in \mathbb{K}^{m\times n}$. Dann heißt die Matrix

$$C = [c_{ik}] := AB \in \mathbb{K}^{\ell\times n}, \quad \textit{wobei} \quad c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}, i = 1, \ldots, \ell, k = 1, \ldots, n.$$

Produkt der Matrizen A und B. Wenn insbesondere B ein Vektor $b \in \mathbb{K}^m$ ist, so ist auch das Produkt $c = Ab \in \mathbb{K}^\ell$ ein Vektor, wobei

$$c_i = \sum_{j=1}^{m} a_{ij} b_j \quad \text{mit} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \quad \text{und} \quad c = \begin{bmatrix} c_1 \\ \vdots \\ c_\ell \end{bmatrix}.$$

Definition 4.47 *(inverse Matrix)*
Wenn zu einem gegebenem $A \in \mathbb{K}^{m\times m}$ ein $B \in \mathbb{K}^{m\times m}$ existiert, so daß gilt $AB = I$ und $BA = I$, so heißt die Matrix A regulär oder invertierbar, B heißt inverse Matrix zu A, und man schreibt $B := A^{-1}$. Anderenfalls heißt A singulär.

Definition 4.48 *(transponierte und adjungierte Matrix)*

(i) Für $A \in \mathbb{K}^{m\times n}$ bezeichnet man mit $A^T \in \mathbb{K}^{n\times m}$ die Matrix, deren $i$te Zeile (bzw. $j$te Spalte) gleich der $i$ten Spalte (bzw. $j$ten Zeile) von A ist, also

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{nm} \end{bmatrix},$$

und nennt sie die transponierte Matrix von A.

(ii) Für $A \in \mathbb{C}^{m\times n}$ ist $A^* \in \mathbb{C}^{n\times m}$ die Matrix deren Koeffizienten die komplex-konjugierten Zahlen der entsprechenden Koeffizienten von A^T sind, also

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \ldots & \overline{a_{m1}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \ldots & \overline{a_{nm}} \end{bmatrix},$$

und nennt sie die adjungierte oder konjugiert-transponierte Matrix von A. Insbesondere ist $A^T = A^*$, wenn alle Koeffizienten von A reell sind.

Eigenschaften 4.49 *(Rechenregeln)*

(i) Für beliebige $A \in \mathbb{K}^{m\times n}$, $x \in \mathbb{K}^n$ *und* $y \in \mathbb{K}^m$ *gilt* $\langle Ax, y\rangle = \langle x, A^*y\rangle$.

(ii) Es gelten die folgenden Rechenregeln:

$$(\lambda A + \mu B)C = \lambda(AC) + \mu(BC), \;\; C(\lambda A + \mu B) = \lambda(CA) + \mu(CB),$$
$$A(BC) = (AB)C, \;\; 0A = B0 = 0, \;\; IA = AI = A,$$
$$(AB)^* = B^*A^*, \;\; (A^{-1})^* = (A^*)^{-1}, \;\; (AB)^{-1} = B^{-1}A^{-1}.$$

Dabei sind A, B *und* C *bzw.* λ *und* μ *beliebige Matrizen bzw. Zahlen derart, daß die Operationen ausgeführt werden können.*

Beweis

(i) $\langle Ax, y\rangle = \sum\limits_{i=1}^{m}\left(\sum\limits_{j=1}^{n} a_{ij}x_j\right)\overline{y_i} = \sum\limits_{j=1}^{n} x_j\left(\overline{\sum\limits_{i=1}^{m}\overline{a_{ij}}y_i}\right) = \langle x, A^*y\rangle$, $\forall x \in \mathbb{K}^n, y \in \mathbb{K}^m$.

(ii) Wir zeigen als Beispiele die drei letzten Eigenschaften:
Es seien $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times k}$. Dann gilt für alle $x \in \mathbb{K}^k$ und $y \in \mathbb{K}^m$ nach (i): $\langle x, (AB)^*y\rangle = \langle ABx, y\rangle = \langle Bx, A^*y\rangle = \langle x, B^*A^*y\rangle$. Deshalb muß $(AB)^*y = B^*A^*y$, $\forall y \in \mathbb{K}^m$ gelten.
Es sei $A \in \mathbb{K}^{n\times n}$ regulär. Dann gilt $(A^{-1})^*A^* = (AA^{-1})^* = I^* = I$ und $A^*(A^{-1})^* = (A^{-1}A)^* = I^* = I$. Es folgt $(A^*)^{-1} = (A^{-1})^*$.
Es seien $A, B \in \mathbb{K}^{n\times n}$ regulär. Dann gilt $B^{-1}A^{-1}AB = B^{-1}(A^{-1}A)B = B^{-1}B = I$ und $ABB^{-1}A^{-1} = A(BB^{-1})A^{-1} = AA^{-1} = I$, also $(AB)^{-1} = B^{-1}A^{-1}$. □

Beispiel 4.50 *Die Rechnung*

$$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

zeigt, daß im allgemeinen in Matrizenprodukten die Faktoren nicht vertauscht werden dürfen. Das Matrizen-Produkt ist also <u>*nicht kommutativ*</u>.
Aus 4.47 folgt, daß die Menge aller regulären Matrizen in $\mathbb{K}^{m\times m}$ *eine* <u>*nicht-kommutative*</u> *Gruppe bzgl. der Multiplikation bildet.*

Satz 4.51 *(Basiswechsel-Matrix)*

(i) Es seien $b_1, \cdots, b_m$ *und* $b'_1, \cdots, b'_m$ *zwei Basen in einem linearen Raum* X *und* $C = [c_{ij}] \in \mathbb{K}^{m\times m}$ *eine Matrix. Dann gilt*

$$b_j = \sum_{i=1}^{m} c_{ij}b'_i\,, \;\forall\, j = 1, \ldots, m \quad \underline{\textit{gdw.}} \quad \sum_{j=1}^{m}\lambda_j b_j = \sum_{i=1}^{m}\mu_i b'_i, \;\textit{wobei}\; \mu = C\lambda.$$

Dabei bedeutet $\mu = C\lambda$, *daß* $\mu_i = \sum_{j=1}^{m} c_{ij}\lambda_j$ $(i = 1, \ldots, m)$.
Die Matrix C *heißt dann Matrix des* <u>*Basiswechsels*</u> *von* $b_1, \ldots, b_m$ *zu* $b'_1, \ldots, b'_m$.
Die j*-te Spalte von* C *sind die Koordinaten des* j*-ten alten Basisvektors bzgl. der neuen Basis.*

(ii) Die Matrix C ist regulär, und C^{-1} ist die Matrix des Basiswechsels von der Basis $b'_1, \dots, b'_m$ zu der Basis $b_1, \dots, b_m$.

(iii) Wenn $b'_1, \dots, b'_m$ eine Orthonormalbasis bzgl. eines Skalarprodukts in X ist, so gilt $c_{ij} = \langle b_j, b'_i \rangle$ für alle $i, j = 1, \dots, m$. Wenn ferner $b_1, \dots, b_m$ ebenfalls eine Orthonormalbasis ist, so gilt $C^{-1} = C^*$.

Beweis:

(i) Für alle $i, j \in \{1, \dots, m\}$ seien $c_{ij} \in \mathbb{K}$ so gewählt, daß $b_j = \sum\limits_{i=1}^m c_{ij} b'_i$ gilt. Daraus folgt

$$\sum_{j=1}^m \lambda_j b_j = \sum_{j=1}^m \lambda_j \left(\sum_{i=1}^m c_{ij} b'_i \right) = \sum_{i=1}^m \left(\sum_{j=1}^m c_{ij} \lambda_j \right) b'_i.$$

Gilt umgekehrt die Gleichung $\sum\limits_{j=1}^m \lambda_j b_j = \sum\limits_{i=1}^m \mu_i b'_i$, so folgt die behauptete Darstellung für die μ_i, $i = 1, \dots, m$, aus der Eindeutigkeit der Basisdarstellung.

(ii) Es seien $C = (c_{ij})$ und $\tilde{C} = (\tilde{c}_{ij})$ die Matrizen des Basiswechsels von $b_1, \dots, b_m$ zu $b'_1, \dots, b'_m$ bzw. von $b'_1, \dots, b'_m$ zu $b_1, \dots, b_m$. Es gilt also

$$b_j = \sum_{i=1}^m c_{ij} b'_i \ (\forall j = 1, \dots, m) \quad \text{und} \quad b'_i = \sum_{j=1}^m \tilde{c}_{ji} b_j \ (\forall i = 1, \dots, m).$$

Daraus folgt durch Einsetzen

$$b_k = \sum_{i=1}^m c_{ik} \sum_{j=1}^m \tilde{c}_{ji} b_j = \sum_{j=1}^m (\sum_{i=1}^m \tilde{c}_{ji} c_{ik}) b_j \ (\forall k = 1, \dots, m).$$

Wegen der Eindeutigkeit der Basisdarstellung folgt aber daraus, dass

$$\sum_{i=1}^m \tilde{c}_{ji} c_{ik} = \delta_{jk} \quad (\forall j, k = 1, \dots, m) \quad \text{oder} \quad \tilde{C} C = I.$$

Durch umgekehrtes Einsetzen und Koeffizientenvergleich bzgl. der Basis $b'_1, \dots, b'_m$ erhält man analog $C \tilde{C} = I$. Also ist die Matrix C regulär und $\tilde{C}$ ist die inverse Matrix.

(iii) Ist $b'_1, \dots, b'_m$ eine Orthonormalbasis, so folgt aus Satz 4.30, daß $b_j = \sum\limits_{i=1}^m \langle b_j, b'_i \rangle b'_i$ für jedes $j = 1, \dots, m$ und folglich $c_{ij} = \langle b_j, b'_i \rangle$, $i, j = 1, \dots, m$. Ist auch $b_1, \dots, b_m$ ein Orthonormalsystem, so gilt $b'_i = \sum\limits_{j=1}^m \langle b'_j, b_j \rangle b_j = \sum\limits_{j=1}^m \tilde{c}_{ij} b_j$ für $i = 1, \dots, m$. Aus (ii) folgt schließlich die Aussage. $\square$

Satz 4.52 *(Matrix-Darstellungen linearer Abbildungen)*
Es seien X und Y lineare Räume über $\mathbb{K}$ mit den Basen $b_1, \dots, b_m$ in X und $b'_1, \dots, b'_n$ in Y.

(i) Es sei $F \in \mathcal{L}(X;Y)$ und $A_F = [a_{ij}] \in \mathbb{K}^{m\times n}$ eine Matrix. Dann gilt

$$F(b_i) = \sum_{j=1}^{n} a_{ij} b_j^{'}\,,\ \forall i = 1,\dots,m \ \ \underline{\textit{gdw.}}\ \ F(\sum_{i=1}^{m} \lambda_i b_i) = \sum_{j=1}^{n} \mu_j b_j^{'}\,,\ \textit{wobei}\ \mu = A^T\lambda\,.$$

Dabei bedeutet $\mu = A^T\lambda$, daß $\mu_j = \sum_{i=1}^m a_{ij}\lambda_i \quad (\forall j = 1,\dots,n)$.
Die Matrix A_F heißt dann Matrixdarstellung von F bzgl. der Basen $b_1,\dots,b_m$ und $b_1^{'},\dots,b_n^{'}$.

(ii) Für beliebige $\lambda,\mu \in \mathbb{K}$ und $F,G \in \mathcal{L}(X;Y)$ gilt $A_{\lambda F+\mu G} = \lambda A_F + \mu A_G$.

(iii) Für $G \in \mathcal{L}(X;Y)$ und $F \in \mathcal{L}(Y;Z)$ gilt $A_{F\circ G} = A_G A_F$.

(iv) Wenn $b_1^{'},\dots,b_n^{'}$ eine Orthonormalbasis bzgl. eines Skalarprodukts $\langle\cdot,\cdot\rangle$ in Y ist, so gilt $a_{ij} = \langle F(b_i), b_j^{'}\rangle$ für alle $i = 1,\dots,m$, $j = 1,\dots,n$.

Beweis:

(i) Analog zum Beweis von Satz 4.49 (i) erhält man die erste Aussage aus

$$F(\sum_{i=1}^{m} \lambda_i b_i) = \sum_{i=1}^{m} \lambda_i F(b_i) = \sum_{i=1}^{m} \lambda_i \sum_{j=1}^{n} a_{ij} b_j^{'} = \sum_{j=1}^{n}(\sum_{i=1}^{m} a_{ij}\lambda_i) b_j^{'}.$$

(ii) (Übung)

(iii) Ist $b_1^{''},\dots,b_l^{''}$ eine Basis in Z und gilt $G(b_i) = \sum_{j=1}^{n} a_{ij} b_j^{'}$ $(\forall i = 1,\dots,m)$ sowie $F(b_j^{'}) = \sum_{k=1}^{l} \tilde{a}_{jk} b_k^{''}$ $(\forall j = 1,\dots,n)$, so folgt

$$F\circ G(b_i) = \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{l} \tilde{a}_{jk} b_k^{''} = \sum_{k=1}^{l}(\sum_{j=1}^{n} a_{ij}\tilde{a}_{jk}) b_k^{''} \quad (i = 1,\dots,m)$$

und damit die Aussage wegen $A_{F\circ G} = (\sum_{j=1}^n a_{ij}\tilde{a}_{jk}) = A_G A_F$.

(iv) folgt analog zu Satz 4.49 (iii) aus Satz 4.30. □

Definition 4.53 *(ähnliche, symmetrische und orthogonale Matrizen)*

(i) Zwei Matrizen $A, B \in \mathbb{K}^{m\times m}$ heißen ähnlich, wenn eine reguläre Matrix $C \in \mathbb{K}^{m\times m}$ existiert, so daß gilt $B = CAC^{-1}$.

(ii) Eine Matrix $A \in \mathbb{K}^{m\times m}$ heißt selbstadjungiert bzw. symmetrisch, wenn $A^* = A$ bzw. $A^T = A$ gilt.

(iii) Eine Matrix $A \in \mathbb{K}^{m\times m}$ heißt orthogonal, wenn $A^{-1} = A^*$ gilt.

4.4 Determinanten

Definition 4.54 *(induktive Determinantendefinition)*
Es sei $[a_{ij}] \in \mathbb{K}^{m \times m}$ *eine quadratische Matrix.*
Ist $m = 1$*, so ist* $\det(A) := a_{11}$ *die* <u>*Determinante*</u> *von* A*.*
Ist $m > 1$*, so bezeichne* A_{ik} *die Matrix aus* $\mathbb{K}^{(m-1)\times(m-1)}$*, die aus* A *durch Streichen der* i*-ten Zeile und* k*-ten Spalte entsteht.*
Dann heißt $\det(A) := \sum\limits_{i=1}^{m} a_{i1}(-1)^{i+1}\det(A_{i1})$ *die* <u>*Determinante*</u> *von* A*.*

Beispiel 4.55
a) $m = 2$*:*

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}(-1)^2 a_{22} + a_{21}(-1)^3 a_{12} = a_{11}a_{22} - a_{12}a_{21}$$

b) $m = 3$*:*

$$\begin{aligned}
\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{23} \end{bmatrix} &= a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{21} \det \begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} \\
&\quad + a_{31} \det \begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix} \\
&= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} \\
&\quad + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} \quad \textit{(\underline{Sarus'sche Regel}).}
\end{aligned}$$

c) $m \in \mathbb{N}$*,* $A := \operatorname{diag}(d_1, \ldots, d_m)$*, d.h.,* A *ist die Diagonalmatrix mit Diagonalelementen* $d_1, \ldots, d_m$*. Es gilt:*

$$\det(\operatorname{diag}(d_1, \ldots, d_m)) = \det \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_m \end{pmatrix} = d_1 \det(\operatorname{diag}(d_2, \ldots, d_m)) = \prod_{i=1}^{m} d_i$$

Insbesondere gilt für die Determinante der Einheitsmatrix $I \in \mathbb{K}^{m \times m}$*:* $\det(I) = 1$*.*

Satz 4.56 *(Laplacescher Entwicklungssatz)*
Es sei $A = [a_{ij}] \in \mathbb{K}^{m \times m}$ *mit* $m > 2$*. Dann gilt für jedes* $k = 1, \ldots, m$*:*

$$\det(A) = \sum_{i=1}^{m} a_{ik}(-1)^{i+k}\det(A_{ik}).$$

Die Zahl $(-1)^{i+k}\det(A_{ik})$ *heißt die* <u>*Adjunkte*</u> *zu* a_{ik} *für* $i, k = 1, \ldots, m$*.*

<u>Beweis:</u> durch vollständige Induktion über m unter Ausnutzung der Richtigkeit der Aussage für $m = 2$ (vgl Beispiel 4.53a)) und von 4.52, wonach die Aussage für $k = 1$ und jedes $m \in \mathbb{N}$ richtig ist.
(Literatur: S. Brehmer, H. Belkner: Einführung in die analytische Geometrie und lineare Algebra, Verlag der Wissenschaften, Berlin 1966, Kap. 3.4.3.)

Eigenschaften 4.57

a) Ist $t \in \mathbb{K}$, $A \in \mathbb{K}^{m\times m}$, so gilt für die Determinante der Matrix, die aus A dadurch entsteht, daß die k-te Spalte mit t multipliziert wird, daß gilt:

$$\det \begin{pmatrix} a_{11} & \dots & ta_{1k} & \dots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & ta_{mk} & \dots & a_{mm} \end{pmatrix} = t \det(A).$$

Insbesondere gilt $\det(tA) = t^m \det(A)$.

b) Vertauscht man in einer quadratischen Matrix A zwei Spalten, so wechselt die Determinante von A das Vorzeichen.

c) Haben die Elemente in der k-ten Spalte einer quadratischen Matrix die Form $a_{ik} + a'_{ik}$, $k = 1, \dots, m$, so gilt für die Determinante

$$\begin{aligned} \det \begin{pmatrix} a_{11} & \dots & a_{1k} + a'_{1k} & \dots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mk} + a'_{mk} & \dots & a_{mm} \end{pmatrix} &= \det \begin{pmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mk} & \dots & a_{mm} \end{pmatrix} \\ &+ \det \begin{pmatrix} a_{11} & \dots & a'_{1k} & \dots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a'_{mk} & \dots & a_{mm} \end{pmatrix} \end{aligned}$$

d) Sind in einer quadratischen Matrix A zwei Spalten gleich, so gilt $\det(A) = 0$.

e) Addiert man zu einer Spalte einer quadratischen Matrix ein beliebiges Vielfaches einer anderen Spalte, so ändert sich die Determinante nicht.

<u>Beweis:</u>
a) Wir wenden Satz 4.54 an und erhalten

$$\det \begin{pmatrix} a_{11} & \dots & ta_{1k} & \dots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & ta_{mk} & \dots & a_{mm} \end{pmatrix} = \sum_{i=1}^{m} ta_{ik}(-1)^{i+k} \det(A_{ik}) = t \det(A).$$

b) Die Aussage ist für $m = 1$ und $m = 2$ richtig. Wir beweisen sie für jedes m durch vollständige Induktion und nehmen an, daß sie für $m-1$ richtig ist. Es sei A' diejenige Matrix, die aus A entsteht, indem deren k-te und l-te Spalte vertauscht wird. Wir wählen $p \in \{1, \dots, m\}$ mit $p \neq k$ und $p \neq l$. Nach Satz 4.54 gilt:

$$\begin{aligned} \det(A) &= \sum_{i=1}^{m} a_{ip}(-1)^{i+p} \det(A_{ip}), \\ \det(A') &= \sum_{i=1}^{m} a'_{ip}(-1)^{i+p} \det(A'_{ip}). \end{aligned}$$

Nach Konstruktion gilt aber $a_{ip} = a'_{ip}$ und nach Induktionsvoraussetzung $\det(A_{ip}) = -\det(A'_{ip})$ jeweils für $i = 1, \ldots, m$. Also folgt: $\det(A) = -\det(A')$.
c) folgt sofort ebenfalls aus Satz 4.54.
d) Bei Vertauschung der beiden gleichen Spalten ändert sich die Matrix nicht. Andererseits ändert sich nach b) das Vorzeichen der Matrix. Deshalb muß $\det(A) = 0$ gelten.
e) folgt aus den Eigenschaften c), a) und d) (in dieser Reihenfolge). $\square$

Faßt man die Determinante einer quadratischen Matrix als Funktion auf, die vom Raum aller m-Tupel von Spalten in $\mathbb{R}$ abbildet, so besagen die Eigenschaften a) und c), daß die Determinante als Funktion einer beliebigen Spalte (bei fixierten anderen $m-1$ Spalten) linear ist. Solche Abbildungen nennt man auch <u>m-multilinear</u>.

Satz 4.58 *Es seien $A, B \in \mathbb{K}^{m\times m}$. Dann gilt:*

(a) $\det(A^T) = \det(A)$*,*

(b) $\det(A\,B) = \det(A)\det(B)$*.*
Insbesondere gilt für jede reguläre Matrix A: $\det(A^{-1}) = (\det(A))^{-1}$.

<u>Beweis:</u>
b) Wir beginnen damit, die Aussage für den Fall zu zeigen, daß eine der Matrizen, sagen wir B, eine Diagonalmatrix $D = \mathrm{diag}(d_1, \ldots, d_n)$ ist. Es gilt

$$\det(A\,D) = \det(d_1 a_{i1}, \ldots, d_m a_{im}) = \prod_{i=1}^{m} d_i \det(A) = \det(D)\det(A).$$

Dabei folgt das zweite Gleichheitszeichen durch Matrix-Multiplikation ($(d_k a_{ik})$ bezeichnet dabei die k-te Spalte von AD), das dritte folgt aus 4.55a) und das vierte aus Beispiel 4.53c). Im folgenden entwickeln wir nun eine Methodik, die den allgemeinen Fall auf die eben betrachtete Situation zurückführt.
Wir betrachten die folgenden speziellen (Transformations-) Matrizen

$$L_{ij}(t) := I + t\begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} = I + te^i(e^j)^T \ (t \in \mathbb{R}, i \neq j),$$

d.h., die Matrizen enstehen dadurch, daß zur Einheitsmatrix $I \in \mathbb{K}^{m\times m}$ eine Matrix addiert wird, die mit Ausnahme des Elementes t in der i-ten Zeile und j-ten Spalte nur Nullen enthält. Die Matrix-Multiplikationen $L_{ij}(t)\,A$ bzw. $A\,L_{ij}(t)$ bewirken, daß zur i-Zeile bzw. j-ten Spalte das t-fache der j-ten Zeile bzw. i-ten Spalte addiert wird. Durch Multiplikationen von A (und B) von rechts und links mit geeigneten Matrizen der Form $L_{ij}(t)$ gelingt es, beide Matrizen A und B auf Diagonalform zu transformieren. Wir demonstrieren das im Fall der Matrix $A = [a_{ij}]$. Ist $a_{11} \neq 0$, so multiplizieren wir A von rechts nacheinander mit $L_{1k}(-\frac{a_{1k}}{a_{11}})$, $k = 2, \ldots, m$. Dadurch entstehen in der ersten Zeile ab der zweiten Spalte Nullen. Ist $a_{11} = 0$, aber $a_{1l} \neq 0$ für ein $l \in$

$\{1, \ldots, m\}$, so multiplizieren wir zuerst A von rechts mit $L_{l1}(1)$ und erreichen dadurch, dass das Element in der ersten Zeile und Spalte verschieden von 0 ist. Sind in der ersten Zeile alle Elemente gleich 0, so gehen wir zur zweiten Zeile über und erreichen durch Multiplizieren mit Matrizen von rechts, daß alle Elemente ab der dritten Spalte zu 0 werden. So setzen wir fort und erhalten eine untere Dreiecksmatrix. Anschließend multiplizieren wir analog von links mit solchen Transformations-Matrizen und erhalten am Schluß eine Diagonalmatrix D_A. Es existieren also Matrizen $L_A^{(l)}$ und $L_A^{(r)}$, die jeweils Produkte solcher Transformations-Matrizen sind, für die gilt

$$L_A^{(l)} \, A \, L_A^{(r)} = D_A \, .$$

Da überdies für die Inversen der Transformations-Matrizen gilt $(L_{ij}(t))^{-1} = L_{ij}(-t)$, kann die letztere Relation auch in der Form

$$A = \hat{L}_A^{(l)} \, D_A \, \hat{L}_A^{(r)}$$

geschrieben werden, wobei die dort auftretenden Matrizen $\hat{L}_A^{(l)}$ und $\hat{L}_A^{(r)}$ vom selben Typ sind wie vorher. Damit gilt

$$\det(A) = \det(\hat{L}_A^{(l)} \, D_A \, \hat{L}_A^{(r)}) = \det(\hat{L}_A^{(l)} \, D_A) = \det(\hat{L}_A^{(l)}) \det(D_A) = \det(D_A),$$

wobei das zweite Gleichheitszeichen sukzessive aus 4.55e), das dritte aus unserer Anfangsbetrachtung und das vierte daraus folgt, daß wieder sukzessive 4.55e) und abschließend $\det(L_{ij}(t)) = 1 \; \forall t \in \mathbb{R}$ verwendet wird.
Betrachtet man nun auch eine entsprechende Transformation

$$B = \hat{L}_B^{(l)} \, D_B \, \hat{L}_B^{(r)}$$

der Matrix B, so kann man wie folgt vorgehen:

$$\begin{aligned}
\det(A\,B) &= \det(\hat{L}_B^{(l)} \, D_B \, \hat{L}_B^{(r)} \hat{L}_A^{(l)} \, D_A \, \hat{L}_A^{(r)}) \\
&= \det(\hat{L}_B^{(l)} \, D_B \, \hat{L}_B^{(r)} \hat{L}_A^{(l)} \, D_A) \\
&= \det(\hat{L}_B^{(l)} \, D_B \, \hat{L}_B^{(r)} \hat{L}_A^{(l)}) \det(D_A) = \det(\hat{L}_B^{(l)} \, D_B) \det(D_A) \\
&= \det(\hat{L}_B^{(l)}) \det(D_B) \det(D_A) = \det(D_B) \det(D_A) \\
&= \det(A) \det(B).
\end{aligned}$$

Dabei werden wieder dieselben Schlüsse verwendet wie vorhin.
a) Da natürlich $A^T = (\hat{L}_A^{(r)})^T D_A (\hat{L}_A^{(l)})^T$ und $(L_{ij}(t))^T = L_{ji}(t)$ gelten, folgt analog zu oben auch

$$\det(A^T) = \det(D_A) = \det(A) \, .$$

Damit ist alles gezeigt. □

Satz 4.56(a) impliziert, daß Satz 4.54 und die Eigenschaften 4.55 richtig bleiben, wenn man jeweils "Spalte" durch "Zeile" ersetzt.
Die folgende Aussage verwendete Leibniz zur (alternativen) Determinantendefinition.

Satz 4.59 *(Leibniz)*
Es sei $A = [a_{ij}] \in \mathbb{K}^{m\times m}$. Dann gilt:

$$\det(A) = \sum_{j_1,\dots,j_m=1}^{m} \det([e^{j_1},\dots,e^{j_m}])a_{j_1,1}\cdots a_{j_m,m},$$

wobei e^j der j-te kanonische Einheitsvektor im $\mathbb{R}^m$ ist, d.h., die Spalte mit einer 1 in der j-ten Zeile und Nullen in allen anderen Zeilen. Es gilt

$$\det([e^{j_1},\dots,e^{j_m}]) = \begin{cases} 1 & \textit{falls die Anzahl der Spaltenvertauschungen von} \\ & [e^{j_1},\dots,e^{j_m}] \textit{zur Einheitsmatrix } I \textit{ gerade ist,} \\ -1 & \textit{falls die Anzahl der Spaltenvertauschungen von} \\ & [e^{j_1},\dots,e^{j_m}] \textit{zur Einheitsmatrix } I \textit{ ungerade ist,} \\ 0 & \textit{falls } \{j_1,\dots,j_m\} \textit{ zwei gleiche Elemente enthält.} \end{cases}$$

(Man nennt das m-Tupel $(j_1,\dots,j_m)$ auch eine <u>Permutation</u> von $(1,\dots,m)$, falls $j_i \neq j_k$ für $i \neq k$ gilt. Die Zahl $\det([e^{j_1},\dots,e^{j_m}])$ nennt man dann das <u>Vorzeichen</u> der Permutation $(j_1,\dots,j_m)$.)
<u>Beweis:</u> Es seien a_k, $k = 1,\dots,m$, die Spaltenvektoren von A. Dann gilt:

$$a_k = \sum_{j=1}^{m} a_{jk}e^j = \sum_{j_k=1}^{m} a_{j_k,k}e^{j_k} \quad (k = 1,\dots,m).$$

Unter Ausnutzung der Rechenregeln 4.55c) und a) für Determinanten erhalten wir

$$\begin{aligned} \det(A) &= \det([\sum_{j_1=1}^{m} a_{j_1,1}e^{j_1},\dots,\sum_{j_m=1}^{m} a_{j_m,m}e^{j_m}]) \\ &= \sum_{j_1,\dots,j_m=1}^{m} \det([a_{j_1,1}e^{j_1},\dots,a_{j_m,m}e^{j_m}]) \\ &= \sum_{j_1,\dots,j_m=1}^{m} \det([e^{j_1},\dots,e^{j_m}])a_{j_1,1}\cdots a_{j_m,m}. \end{aligned}$$

Die Formel für $\det([e^{j_1},\dots,e^{j_m}])$ folgt aus 4.55b). $\square$

Satz 4.60 *Es sei $A = [a_{ij}] \in \mathbb{K}^{m\times m}$.*
Dann ist A regulär <u>gdw.</u> $\det(A) \neq 0$ und

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} (-1)^{1+1}\det(A_{11}) & (-1)^{2+1}\det(A_{21}) & \cdots & (-1)^{m+1}\det(A_{m1}) \\ (-1)^{1+2}\det(A_{12}) & (-1)^{2+2}\det(A_{22}) & \cdots & (-1)^{m+2}\det(A_{m2}) \\ \vdots & \vdots & \vdots & \\ (-1)^{1+m}\det(A_{1m}) & (-1)^{2+m}\det(A_{2m}) & \cdots & (-1)^{m+m}\det(A_{mm}) \end{pmatrix}.$$

Beweis: Nach Satz 4.54 gilt: $\det(A) = \sum_{i=1}^{m} a_{ik}(-1)^{i+k}\det(A_{ik}) \quad (k = 1, \ldots, m)$.
Würde in die k-te Spalte von A die j-te Spalte von A geschrieben werden, so wären in der entstandenen Matrix 2 Spalten gleich und folglich deren Determinante gleich 0 nach 4.55d). Also würde gelten

$$\sum_{i=1}^{m} a_{ij}(-1)^{i+k}\det(A_{ik}) = 0 \quad (j \neq k, j, k = 1, \ldots, m).$$

Wir betrachten jetzt die Matrix $\tilde{A}$ der Adjunkten (zu den Elementen von A), d.h.,

$$\tilde{A} = \begin{pmatrix} (-1)^{1+1}\det(A_{11}) & \cdots & (-1)^{1+m}\det(A_{1m}) \\ \vdots & \vdots & \vdots \\ (-1)^{m+1}\det(A_{m1}) & \cdots & (-1)^{m+m}\det(A_{mm}) \end{pmatrix}.$$

Dann gilt nach den obigen Überlegungen

$$\begin{aligned} \tilde{A}^T A &= \Big[\sum_{i=1}^{m} a_{ij}(-1)^{i+k}\det(A_{ik})\Big]_{j,k=1,\ldots,m} \\ &= \operatorname{diag}(\det(A), \ldots, \det(A)) = \det(A)\, I. \end{aligned}$$

($\Longrightarrow$) Ist A regulär, so gilt $\det(A) \neq 0$ nach 4.56b), und damit $\frac{1}{\det(A)}\tilde{A}^T A = I$.
($\Longleftarrow$) folgt aus der letzten Identität. □

Beispiel 4.61
Wir betrachten $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ *mit* $\det(A) = -1$. *Nach Satz 4.58 gilt:*

$$A^{-1} = -\begin{pmatrix} \det(A_{11}) & -\det(A_{12}) & \det(A_{13}) \\ -\det(A_{21}) & \det(A_{22}) & -\det(A_{23}) \\ \det(A_{31}) & -\det(A_{32}) & \det(A_{33}) \end{pmatrix} = -\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} = A.$$

4.5 Lineare Gleichungssysteme

Definition 4.62 *Es sei $A \in \mathbb{K}^{m\times n}$ und $b \in \mathbb{K}^m$ gegeben. Dann bezeichnet man die Gleichung $Ax = b$ als lineares Gleichungssystem mit dem Vektor $x \in \mathbb{K}^n$ der Unbekannten. Ist der Vektor b der rechten Seiten gleich Null (bzw. ungleich Null), so heißt das lineare Gleichungssystem homogen (bzw. inhomogen).*

Es sei eine Matrix $A = (a_1, \ldots, a_n) \in \mathbb{K}^{m\times n}$ mit den Spalten $a_i \in \mathbb{K}^m$, $i = 1, \ldots, n$, gegeben. Da die Matrix A eine lineare Abbildung von $\mathbb{K}^n$ in $\mathbb{K}^m$ definiert, bezeichnen wir analog zu Definition 4.37 mit $\ker(A) = \{x \in \mathbb{K}^n : Ax = 0\}$ den Kern von A und mit $\mathrm{r}(A) = \dim \operatorname{span}\{a_1, \ldots, a_n\}$ den Rang von A. Nach Satz 4.38 gilt dann

$$\mathrm{r}(A) + \dim\ker(A) = n.$$

Satz 4.63 *(Lösbarkeit und Lösungsmenge)*
Es seien $A \in \mathbb{K}^{m\times n}$ und $b \in \mathbb{K}^m$.
Das lineare Gleichungssystem $Ax = b$ besitzt eine Lösung gdw. $\mathrm{r}\,[A\ b] = \mathrm{r}\,(A)$.
Ist $x_0 \in \mathbb{K}^n$ eine beliebige Lösung, so gilt $\{x \in \mathbb{K}^n : Ax = b\} = x_0 + \ker(A)$.
Falls $m = n$ und $\det(A) \neq 0$, *so besitzt das lineare Gleichungssystem für alle rechten Seiten $b \in \mathbb{K}^n$ eine eindeutige Lösung $\bar{x} = A^{-1}b$. Für diese gilt:*

$$\bar{x}_k = \frac{\sum\limits_{i=1}^{m} b_i(-1)^{i+k}\det(A_{ik})}{\det(A)}, k = 1, \ldots, m. \quad \textit{(Cramersche Regel)}$$

Beweis: Die Rang-Bedingung $\mathrm{r}\,[A\ b] = \mathrm{r}\,(A)$ für die um die Spalte b erweiterte Matrix A zu $[A\ b]$ ist gleichbedeutend damit, daß b zum Wertebereich $\mathrm{span}\{a_1, \ldots, a_n\}$ von A gehört. Dies bedeutet aber gerade Lösbarkeit von $Ax = b$.
Die zweite Aussage ist eine Konsequenz von Folgerung 4.39. Die dritte Aussage folgt aus Satz 4.58 und der Formel $\bar{x} = \frac{1}{\det(A)}\tilde{A}^T b$, wobei $\tilde{A}$ die Matrix der Adjunkten darstellt. □

Für die numerische Lösung linearer Gleichungssysteme mit "großen" $m = n$ ist aber die Cramersche Regel nicht geeignet, da die Berechnung von Determinanten nach Definition 4.52 bzw. Satz 4.57 eine Anzahl von Operationen (Additionen und Multiplikationen) erfordert, die mit $m!$ schnell anwächst. Deshalb erinnern wir uns an die Beweis-Methodik von Satz 4.56 mittels Transformations-Matrizen. Die dortige Methodik wenden wir aber nur zur Transformation von A auf Dreiecksgestalt an. Ausgangspunkt sind die Beobachtungen, dass die folgenden Operationen die Lösungsmenge nicht verändern:

- Multiplikation einer Zeile von $Ax = b$ mit einem Faktor ungleich Null,
- Addition einer Zeile von $Ax = b$ zu einer anderen Zeile.

Dies bringt uns auf die Idee zu folgendem Algorithmus, der auf die algemeine Situation (rechteckiger) linearer Gleichungssysteme anwendbar ist.

Algorithmus 4.64 *(Gaußscher Algorithmus)*
Gegeben sei ein lineares Gleichungssystem $Ax = b$ mit $A \in \mathbb{K}^{m\times n}$ und $b \in \mathbb{K}^m$.

- $A^{(1)} := A,\ b^{(1)} := b;$
- *für* $k = 1, \ldots, n-1$:
  - *finde einen Index $s(k) \in \{k,\ k+1, \ldots, m\}$ mit $a^{(k)}_{s(k),k} \neq 0$ und vertausche die Zeilen k und $s(k)$ in $A^{(k)}$ und $b^{(k)}$ und bezeichne die Elemente wie vorher;*
  - $a^{(k+1)}_{ij} := \begin{cases} 0 & , \quad i = k+1, \ldots, m, j = k \\ a^{(k)}_{ij} - \frac{a^{(k)}_{ik}}{a^{(k)}_{kk}} a^{(k)}_{kj} & , \quad i = k+1, \ldots, m, j = k+1, \ldots, n \\ a^{(k)}_{ij} & \quad \textit{sonst} \end{cases}$
  - $b^{(k+1)}_i := \begin{cases} b^{(k)}_i - \frac{a^{(k)}_{ik}}{a^{(k)}_{kk}} b^{(k)}_k & , \quad i = k+1, \ldots, m \\ b^{(k)}_i & , \quad \textit{sonst} \end{cases}$

- $R := A^{(n)}, \; c := b^{(n)}$.

Satz 4.65 *Die Lösungsmenge von $Rx = c$ ist gleich der Lösungsmenge von $Ax = b$. Die Matrix R besitzt obere Dreiecksgestalt, wobei die i-ten Zeilen für $\mathrm{r}(A) < i \leq m$ gleich Null sind. Ist c_i mit $\mathrm{r}(A) < i \leq m$ ungleich Null, so ist das Gleichungssystem nicht lösbar. Gilt $m = n$, so kann die Lösungsmenge von $Rx = c$ durch die sogenannte Rückwärtselimination sukzessive bestimmt werden:*

$$x_m := \frac{c_m}{r_{mm}}, \quad x_i = c_i - \sum_{j=i+1}^{m} r_{ij} x_j / r_{ii}, \quad i = m-1, \ldots, 1\,.$$

4.6 Eigenwerte und Eigenvektoren

Definition 4.66 *(Eigenwerte, Eigenvektoren)*
Es seien $A \in \mathbb{C}^{m \times m}$, $\lambda \in \mathbb{C}$ und $x \in \mathbb{C}^m$ mit

$$Ax = \lambda x, \;\; x \neq 0.$$

Dann heißen λ bzw. x Eigenwert bzw. Eigenvektor von A. Die Menge aller Eigenwerte von A heißt Spektrum von A und wird mit $\mathrm{spec}(A)$ bezeichnet.

Die gleiche Terminologie kann man für lineare Abbildungen $F : X \to X$ auf einem beliebigen linearen Raum X einführen. Falls X endlich-dimensional ist, gelten dann Analoga zu den nachfolgenden Resultaten.

Satz 4.67 *(charakteristisches Polynom)*
Eine Zahl $\lambda \in \mathbb{C}$ ist Eigenwert von $A \in \mathbb{C}^{m \times m}$ gdw. sie Nullstelle des charakteristischen Polynoms von A ist, d.h.,

$$\varphi_A(\lambda) := \det(A - \lambda I) = 0.$$

Die Menge $\mathrm{spec}(A) \subset \mathbb{C}$ ist nichtleer und besteht aus höchstens m Elementen.

Beweis: Ein Eigenvektor x zum Eigenwert λ von A ist nach Definition 4.64 nichttriviale Lösung des homogenen linearen Gleichungssystems

$$(A - \lambda I)x = 0\,.$$

Da $x_0 = 0$ Lösung dieses linearen Gleicungssystems ist, ist seine Lösungsmenge gerade der Unterraum $\ker(A - \lambda I)$. Das homogene lineare Gleichungssystem hat also genau dann eine nichttriviale Lösung, wenn $\ker(A - \lambda I) \neq \{0\}$ gilt. Also folgt:
x ist Eigenvektor zu λ gdw. $\ker(A - \lambda I) \neq \{0\}$ gdw. $A - \lambda I$ ist nicht regulär gdw. $\det(A - \lambda I) = 0$ (nach Satz 4.58).
Nach Satz 4.57 muß aber $\varphi_A(\lambda) := \det(A - \lambda I)$ ein Polynom m-ten Grades mit komplexen Koeffizienten sein (der Term mit der höchsten λ-Potenz ist gerade $(-1)^m \lambda^m$). Nach dem Fundamentalsatz der Algebra hat aber ein solches Polynom mindestens eine Nullstelle und höchstens m verschiedene Nullstellen in $\mathbb{C}$. $\square$

Definition 4.68 *(Vielfachheit und Eigenraum)*
Es sei $\tilde{\lambda}$ Eigenwert von $A \in \mathbb{C}^{m\times m}$.

(i) Der Unterraum $\ker(A-\tilde{\lambda}I)$ *aller Eigenvektoren zu $\tilde{\lambda}$ heißt Eigenraum zu $\tilde{\lambda}$, und seine Dimension heißt geometrische Vielfachheit von $\tilde{\lambda}$.*

(ii) Existiert ein $\alpha \in \{1,\ldots,m\}$, so daß

$$\varphi_A(\lambda) = (\tilde{\lambda}-\lambda)^\alpha \psi(\lambda) \quad \textit{und} \quad \psi(\tilde{\lambda}) \neq 0,$$

so heißt α algebraische Vielfachheit von $\tilde{\lambda}$ und der Unterraum $\ker(A-\tilde{\lambda}I)^\alpha$ *heißt verallgemeinerter Eigenraum von $\tilde{\lambda}$.*

Das nachfolgende Beispiel zeigt, daß die geometrische und algebraische Vielfachheit eines Eigenwerts i.a. verschieden sind.

Beispiel 4.69
Wir betrachten die Matrix $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$. *Es gilt* $\det(A-\lambda I) = (\lambda-2)^2$ *und* $\ker(A-2I) = \ker\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. *Der Eigenwert $\tilde{\lambda}=2$ hat also algebraische Vielfachheit $\alpha=2$ und geometrische Vielfachheit $\gamma=1$.*

Ehe wir zeigen, daß die algebraische Vielfachheit stets größer oder gleich der geometrischen Vielfachheit eines Eigenwerts ist, benötigen wir eine wichtige Vorbereitung zu Eigenwerten ähnlicher Matrizen.

Satz 4.70 *(Eigenwerte ähnlicher Matrizen)*
Es seien A, B und T in $\mathbb{C}^{m\times m}$, T sei regulär und es gelte $B = T^{-1}AT$, d.h., A und B seien ähnlich. Dann gilt:

$$\varphi_A(\lambda) = \varphi_B(\lambda) \quad \textit{und} \quad \dim\ker(A-\lambda I) = \dim\ker(B-\lambda I) \quad \forall \lambda \in \mathbb{C}.$$

Insbesondere besitzen also A und B diesselben Eigenwerte mit denselben geometrischen und algebraischen Vielfachheiten.

Beweis: Für jedes $\lambda \in \mathbb{C}$ gilt:

$$\begin{aligned}\det(B-\lambda I) &= \det(T^{-1}AT-\lambda I) = \det(T^{-1}(A-\lambda I)T) \\ &= \det(T^{-1})\det(A-\lambda I)\det(T) = \det(A-\lambda I).\end{aligned}$$

Dabei wurde insbesondere Satz 4.56 verwendet. Überdies gilt:

$$By = \lambda y \quad \Leftrightarrow \quad A(Ty) = \lambda(Ty) \quad \Leftrightarrow \quad (A-\lambda I)(Ty) = 0.$$

Also ist y Eigenvektor von B zu λ genau dann, wenn Ty Eigenvektor von A zu λ ist und es gilt $\ker(A-\lambda I) = T\ker(B-\lambda I)$. Da T regulär ist, stimmen aber die Dimensionen von $\ker(A-\lambda I)$ und $\ker(B-\lambda I)$ überein. $\square$

Satz 4.71
Es seien $\lambda_1, \dots, \lambda_p$ sämtliche verschiedenen Eigenwerte einer Matrix $A \in \mathbb{C}^{m\times m}$ mit den algebraischen bzw. geometrischen Vielfachheiten $\alpha_1, \dots, \alpha_p$ bzw. $\gamma_1, \dots, \gamma_p$. Dann gilt

$$\varphi_A(\lambda) = \prod_{i=1}^{p}(\lambda_i - \lambda)^{\alpha_i}, \quad \sum_{i=1}^{p}\alpha_i = m \quad \textit{und} \quad 1 \le \gamma_i \le \alpha_i \le m,\ i = 1, \dots, p.$$

Beweis: Da die $\lambda_1, \dots, \lambda_p$ sämtliche verschiedenen Eigenwerte von A muß nach Definition 4.66 $\varphi_A(\lambda) = \prod_{i=1}^{p}(\lambda_i - \lambda)^{\alpha_i}\psi(\lambda)$ gelten, wobei das Polynom $\psi(\cdot)$ keine weiteren Nullstellen besitzen darf, also $\psi(\lambda) \equiv C$ =const gelten muß. Dann folgt aber $\sum_{i=1}^{p}\alpha_i = m$ und $C = 1$, da $\varphi_A(\cdot)$ ein Polynom m-ten Grades mit höchstem Koeffizienten $(-1)^m$ ist.
Klar ist nach Definition 4.66 auch $1 \le \gamma_i, \alpha_i \le m$ für $i = 1, \dots, p$. Um die verbliebene Ungleichung zu zeigen, betrachten wir den Eigenwert λ_i mit $i \in \{1, \dots, p\}$. Es seien $x_{i1}, \dots, x_{i\gamma_i}$ eine Basis des Eigenraums $\ker(A - \lambda_i I)$. Wir ergänzen diese γ_i linear unabhängigen Elemente zu einer Basis von $\mathbb{C}^m$ und bilden eine Matrix $T \in \mathbb{C}^{m\times m}$, indem wir beginnend mit $x_{i1}, \dots, x_{i\gamma_i}$ diese linear unabhängigen Vektoren in ihre Spalten eintragen. Dann gilt $Te_j = x_{ij}$ für $j = 1, \dots, \gamma_i$ und die Elemente e_j der kanonischen Basis. Folglich gilt $T^{-1}ATe_j = T^{-1}Ax_{ij} = \lambda_i e_j$, $j = 1, \dots, \gamma_i$, bzw. $T^{-1}AT = \begin{pmatrix} \lambda_i I & B \\ 0 & C \end{pmatrix}$, wobei I die Einheitsmatrix im $\mathbb{C}^{\gamma_i\times\gamma_i}$ ist sowie B und C Matrizen passender Dimension sind. Durch sukzessive Entwicklung nach der ersten Spalte ergibt sich nach Satz 4.68

$$\varphi_A(\lambda) = \det(T^{-1}AT - \lambda I) = (\lambda_i - \lambda)^{\gamma_i}\det(C - \lambda I),$$

wobei mit I jeweils die Einheitsmatrix passender Dimension bezeichnet wird. Also muß nach Definition 4.66 $\gamma_i \le \alpha_i$ gelten. $\square$

Satz 4.72
Es seien $\lambda_1, \dots, \lambda_p$ verschiedene Eigenwerte einer Matrix $A \in \mathbb{C}^{m\times m}$ mit den Eigenvektoren $x_1, \dots, x_p$. Dann sind die Vektoren $x_1, \dots, x_p$ linear unabhängig.

Beweis: Wir wissen, daß gilt: $Ax_i = \lambda_i x_i$, $\lambda_i \ne \lambda_k$, $i, k = 1, \dots, p$, $i \ne k$.
Wir betrachten die Gleichung $\sum\limits_{i=1}^{p} c_i x_i = 0$ mit Zahlen $c_i \in \mathbb{C}$. Durch Anwendung der Matrizenprodukte $(A - \lambda_1 I)\cdots(A - \lambda_j I)$, $j = 1, \dots, p-1$, auf diese Gleichung entstehen die neuen Gleichungen $\sum\limits_{i=j+1}^{p} c_i(\lambda_i - \lambda_{i-1})\cdots(\lambda_i - \lambda_1)x_i = 0$, $j = 1, \dots, p-1$.
Aus der Gleichung für $j = p-1$ ergibt sich sofort $c_p = 0$. Daraus dann aus der Gleichung für $j = p-2$ auch $c_{p-1} = 0$ usw. Man erhält $c_1 = \dots = c_p = 0$ und damit die gewünschte lineare Unabhängigkeit. $\square$

Sind nun $\lambda_1, \dots, \lambda_p$ sämtliche verschiedenen Eigenwerte von A mit den geometrischen Vielfachheiten $\gamma_1, \dots, \gamma_p$, so wählen wir Basen $x_{i1}, \dots, x_{i\gamma_i}$ der Eigenräume $\ker(A - \lambda_i I)$ für $i = 1, \dots, p$. Nach Satz 4.70 ist dann das System

$$x_{11}, \dots, x_{1\gamma_1}, x_{21}, \dots, x_{2\gamma_2}, \dots, x_{p1}, \dots, x_{p\gamma_p},$$

linear unabhängig in $\mathbb{C}^m$. Allerdings gilt i.a. nach Satz 4.69 nur

$$p \leq \sum_{i=1}^{p} \gamma_i \leq \sum_{i=1}^{p} \alpha_i = m,$$

wobei α_i die algebraische Vielfachheit von λ_i, $i = 1, \ldots, p$, ist. Wir nennen nun die Matrix $A \in \mathbb{C}^{m \times m}$ nichtdefektiv, falls $\sum_{i=1}^{p} \gamma_i = m$.

Satz 4.73 *(Diagonalähnlichkeit nichtdefektiver Matrizen)*
Ist $A \in \mathbb{C}^{m \times m}$ nichtdefektiv, so existiert eine reguläre Matrix $X \in \mathbb{C}^{m \times m}$ mit der Eigenschaft

$$X^{-1}AX = \Lambda := \begin{pmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & 0 & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & 0 & & & & \lambda_p & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \lambda_p \end{pmatrix}.$$

Hierbei sind $\lambda_1, \ldots, \lambda_p$ sämtliche verschiedenen Eigenwerte von A, die in der Hauptdiagonalen von Λ entspechend ihrer Vielfachheiten $\gamma_1, \ldots, \gamma_p$ auftreten.
(Matrizen mit der obigen Eigenschaft heißen auch diagonalähnlich. Matrizen sind diagonalähnlich gdw. sie nichtdefektiv sind.)

Beweis: Da A nichtdefektiv ist, ist die Anzahl der linear unabhängigen Eigenvektoren

$$x_{11}, \ldots, x_{1\gamma_1}, x_{21}, \ldots, x_{2\gamma_2}, \ldots, x_{p1}, \ldots, x_{p\gamma_p}$$

in $\mathbb{C}^m$ gleich m. Wir bilden aus diesen Vektoren eine Matrix $X \in \mathbb{C}^{m \times m}$. Wir wissen, daß X regulär ist und nach Konstruktion gilt:

$$AX = X\Lambda \quad \text{oder} \quad X^{-1}AX = \Lambda. \qquad \square$$

Ein wesentlicher Spezialfall diagonalähnlicher Matrizen sind Matrizen mit sämtlich verschiedenen (algebraisch) einfachen Eigenwerten $\lambda_1, \ldots, \lambda_m$ mit $\gamma_i = \alpha_i = 1$, $i = 1, \ldots, m$. Ein anderer wichtiger Spezialfall sind selbstadjungierte Matrizen.

Satz 4.74 *(selbstadjungierte Matrizen)*
Die Matrix $A \in \mathbb{C}^{m \times m}$ sei selbstadjungiert.
Dann sind sämtliche Eigenwerte von A reell, A ist diagonalähnlich und besitzt eine Orthonormalbasis von Eigenvektoren in $\mathbb{C}^m$. Die zugehörige Matrix X (vgl. 4.71) ist orthogonal. Überdies gilt

$$\lambda_{\min}\|x\|^2 \leq \langle Ax, x\rangle \leq \lambda_{\max}\|x\|^2 \quad \forall x \in \mathbb{C}^m,$$

wobei $\lambda_{\min}$ bzw. $\lambda_{\max}$ der kleinste bzw. der größte Eigenwert von A ist.

Beweis: Wir zeigen zunächst, daß die adjungierte Matrix A^* den Eigenwert $\bar{\lambda}$ genau dann besitzt, wenn λ Eigenwert von A ist. Dies folgt aus der Gleichungskette

$$0 = \overline{\det(A - \lambda I)} = \overline{\det((A - \lambda I)^T)} = \overline{\det(A^T - \lambda I)} = \det(\overline{(A^T - \lambda I)}) = \det(A^* - \bar{\lambda} I).$$

Da A selbstadjungiert ist, d.h. $A = A^*$ gilt, muß also jeder Eigenwert von A reell sein. Wir zeigen nun, daß jede selbstadjungierte Matrix A diagonalähnlich mit einer orthogonalen (Transformations-) Matrix X ist. Den Beweis führen wir mit vollständiger Induktion über die Dimension m. Für $m = 1$ ist die Aussage trivial. Sie sei nun für $m - 1$ richtig und $A \in \mathbb{C}^{m\times m}$ sei eine selbstadjungierte Matrix (der Dimension m).
Es sei λ_1 ein Eigenwert von A, d.h. $\lambda_1 \in \mathbb{R}$, und $x_1 \in \mathbb{C}^m$ ein zugehöriger Eigenvektor, d.h. $Ax_1 = \lambda_1 x_1$, mit $\|x_1\|_2 = 1$. Wir ergänzen x_1 durch $x_2, \ldots, x_m \in \mathbb{C}^m$ zu einer Orthonormalbasis von $\mathbb{C}^m$. Wir betrachten die Matrix $X_0 \in \mathbb{C}^{m\times m}$, deren Spalten gerade die Basiselemente $x_1, \ldots, x_m$ sind. Dann gilt:

$$X_0^* X_0 = I \quad \text{und} \quad X_0^* A X_0 e_1 = X_0^* A x_1 = \lambda_1 X_0^* x_1 = \lambda_1 e_1 \, .$$

Deshalb hat $X_0^* A X_0$ die Gestalt $\begin{pmatrix} \lambda_1 & a \\ 0 & A_1 \end{pmatrix}$ mit $A_1 \in \mathbb{C}^{(m-1)\times(m-1)}$ und $a^T \in \mathbb{C}^{m-1}$.
Da auch $X_0^* A X_0$ selbstadjungiert ist, muß $a^T = 0$ gelten.
Nach Induktionsannahme existiert eine orthogonale Matrix $X_1 \in \mathbb{C}^{(m-1)\times(m-1)}$, so daß $X_1^* A_1 X_1$ eine Diagonalmatrix mit den Eigenwerten $\lambda_2, \ldots, \lambda_m$ von A_1 in der Hauptdiagonale ist. Wir definieren nun die Matrix $X := X_0 \begin{pmatrix} 1 & 0 \\ 0 & X_1 \end{pmatrix}$ und berechnen $X^* A X$.

$$\begin{aligned} X^* A X &= \begin{pmatrix} 1 & 0 \\ 0 & X_1^* \end{pmatrix} X_0^* A X_0 \begin{pmatrix} 1 & 0 \\ 0 & X_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & X_1^* \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & X_1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & X_1^* A_1 X_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{pmatrix} \end{aligned}$$

Die Zahlen $\lambda_1, \ldots, \lambda_m$ sind aber gerade die (nicht notwendig verschiedenen reellen) Eigenwerte von A. Ferner gilt $\langle x, x\rangle = \sum_{i=1}^m |y_i|^2$ für alle $x \in \mathbb{C}^m$, $y = X^* x$ und

$$\langle Ax, x\rangle = \langle AX(X^*x), X(X^*x)\rangle = \langle X^*AX(X^*x), (X^*x)\rangle = \sum_{i=1}^m \lambda_i |y_i|^2.$$

Hieraus folgt abschließend die gewünschte beidseitige Abschätzung von $\frac{\langle Ax,x\rangle}{\|x\|^2}$. □

Für den allgemeinen Fall geben wir die Ergebnisse ohne Beweis an.

Lemma 4.75 *(Jordan-Kette)*
Es sei λ Eigenwert von $A \in \mathbb{C}^{m\times m}$ mit geometrischer Vielfachheit γ und algebraischer Vielfachheit α. Ferner sei $x_{11}, \ldots, x_{1\gamma}$ eine Basis im zu λ gehörenden Eigenraum $\ker(A - \lambda I)$. *Gilt $\gamma < \alpha$, so kann diese Basis zu einer Menge mit α linear unabhängigen Vektoren*

$$x_{11}, x_{21}, \ldots, x_{\ell_1 1}, x_{12}, \ldots, x_{\ell_2 2}, \ldots, x_{1\gamma}, \ldots, x_{\ell_\gamma \gamma}$$

ergänzt werden. Hierbei ist $\sum_{j=1}^{\gamma} \ell_j = \alpha$ *und für jedes* $j = 1, \ldots, \gamma$ *gilt* $Ax_{1j} = \lambda x_{1j}$ *und* x_{kj}, $k = 2, \ldots, \ell_j$ *ist Lösung von* $Ax_{kj} = \lambda x_{kj} + x_{k-1,j}$ *(solange diese existieren).*
Die Vektoren x_{kj}, $k = 2, \ldots, \ell_j$, *heißen* *Hauptvektoren* *zu* λ, $x_{1j}, x_{2j}, \ldots, x_{\ell_j j}$ *heißt* *Jordan-Kette* *zum Eigenvektor* x_{1j} *und die Matrix*

$$J_{\ell_j}(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \ddots & & \\ \vdots & & \ddots & \ddots & & 0 \\ \vdots & & \cdots & 0 & \lambda & 1 \\ 0 & & \cdots & 0 & 0 & \lambda \end{pmatrix} \in \mathbb{C}^{\ell_j \times \ell_j},$$

das zugehörige *Jordan-Kästchen*.

Satz 4.76 *(Jordansche Normalform einer Matrix)*
Es seien $\lambda_1, \ldots, \lambda_p$ *sämtliche verschiedene Eigenwerte von* $A \in \mathbb{C}^{m \times m}$ *mit den geometrischen Vielfachheiten* γ_i *und den algebraischen Vielfachheiten* a_i, $i = 1, \ldots, p$. *Ausgehend von Basen* $x_{i1}, \ldots, x_{i\gamma_i}$ *der Eigenräume* $\ker(A - \lambda_i I)$ *bildet man die zugehörigen Jordan-Ketten und Jordan-Kästchen* $J_{\ell_{ij}}(\lambda_i)$, $j = 1, \ldots, \gamma_i$, $i = 1, \ldots, p$.
Bezeichnet X *die Matrix, die sämtliche Vektoren der Jordanketten in den Spalten enthält, so ist* X *regulär und es gilt*

$$X^{-1}AX = J := \begin{pmatrix} J_{\ell_{11}}(\lambda_1) & 0 & & & \cdots & & 0 \\ 0 & \ddots & & & & \cdots & \\ & & J_{\ell_{1\gamma_1}}(\lambda_1) & & 0 & \cdots & \\ \vdots & & \ddots & & & & \vdots \\ \vdots & & & & J_{\ell_{p1}}(\lambda_p) & & \\ & & & & & \ddots & 0 \\ 0 & & & & & 0 & J_{\ell_{p\gamma_p}}(\lambda_p) \end{pmatrix}.$$

In der Hauptdiagonalen von J *stehen also die Eigenwerte von* A *entsprechend ihrer algebraischen Vielfachheit, die Anzahl der Einsen oberhalb der Hauptdiagonalen in jedem Jordan-Kästchen entspricht der jeweiligen Anzahl der Hauptvektoren in der Jordan-Kette, die Anzahl der Jordan-Kästchen ist* $\sum_{i=1}^{p} \gamma_i$.
Die Matrix J *heißt* *Jordansche Normalform* *von* A.

Beispiel 4.77 *(vgl. Beispiel 4.67)*
Wir betrachten wieder die Matrix $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$. *Wir wissen, daß* $\lambda_1 = 2$, $\alpha_1 = 2$ *und* $\gamma_1 = 1$ *gilt. Zunächst bestimmen wir einen Eigenvektor* x_1 *zu* λ_1:
$(A - \lambda_1 I)x_1 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} x_1 = 0 \rightsquigarrow x_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.
Als nächstes berechnen wir die Jordankette zu $x_1 = x_{11}$:
$(A - \lambda_1 I)x_{12} = x_{11} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} x_{12} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rightsquigarrow x_{12} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

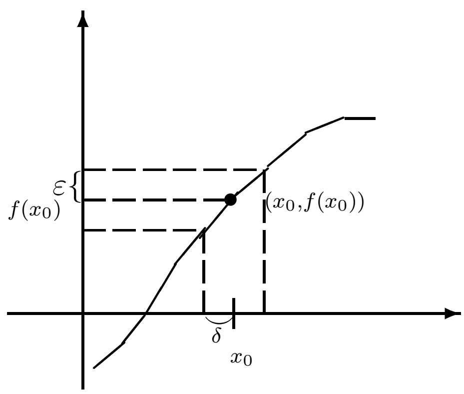
Damit gilt für das einzige zu $\lambda_1 = 2$ gehörige Jordan-Kästchen $J_{11}(2) = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$ und für die Jordansche Normalform $J = J_{11}(2)$.

5 Stetige Funktionen

Wir kehren nun zur Analysis zurück und wenden uns verstärkt der Untersuchung von Abbildungen zu. Der erste zentrale Begriff ist dabei die Stetigkeit.
Wir betrachten eine Funktion $f : D(f) \to \mathbb{R}^k$ mit $D(f) \subseteq \mathbb{R}^m$. Dabei können $\mathbb{R}^m$ und $\mathbb{R}^k$ mit Normen versehen sein, die möglicherweise verschieden sind, die wir aber beide mit $\|\cdot\|$ bezeichnen.

Definition 5.1
$f : D(f) \to \mathbb{R}^k$ heißt stetig in $x_0 \in D(f)$, falls für jedes $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert, so daß aus $x \in D(f)$ und $\|x - x_0\| < \delta$ folgt, daß $\|f(x) - f(x_0)\| < \varepsilon$.
(kurz: $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D(f) \cap \overset{o}{B}(x_0, \delta) : f(x) \in \overset{o}{B}(f(x_0), \varepsilon)$
oder: $\forall \varepsilon > 0 \exists \delta > 0 : f(D(f) \cap \overset{o}{B}(x_0, \delta)) \subseteq \overset{o}{B}(f(x_0), \varepsilon))$
f heißt stetig, falls f in jedem $x_0 \in D(f)$ stetig ist.

Bemerkung 5.2
In Def 5.1 kann anstelle von „$<$" auch „$\leq$" geschrieben werden. In der Kurzschreibweise bedeutet dies, daß man auch schreiben kann:
$\forall \varepsilon > 0 \exists \delta > 0 :$
$f(D(f) \cap B(x_0, \delta)) \subseteq B(f(x_0), \varepsilon)$.
Anschaulich bedeutet Stetigkeit in $x_0 \in D(f)$: So klein man auch eine Kugel um $f(x_0)$ wählt, es existiert immer eine (evtl. „sehr kleine") Kugel um x_0, deren Durchschnitt mit $D(f)$ durch f in die Kugel um $f(x_0)$ abgebildet wird.



Satz 5.3
f ist stetig in $x_0 \in D(f)$ gdw. für alle Folgen (x_n) in $D(f)$ mit $\|x_n - x_0\| \xrightarrow{n \to \infty} 0$ gilt $\|f(x_n) - f(x_0)\| \xrightarrow{n \to \infty} 0$.

Beweis:

($\Rightarrow$) Es sei (x_n) eine Folge in $D(f)$ mit $\|x_n - x_0\| \xrightarrow{n \to \infty} 0$. Es sei $\varepsilon > 0$ beliebig gewählt. Dann existiert ein $\delta > 0$, so daß

$$f(D(f) \cap B(x_0, \delta)) \subseteq B(f(x_0), \varepsilon).$$

$\rightsquigarrow$ $\exists n_0 \in I\!N : \|x_n - x_0\| \leq \delta, \forall n \geq n_0$.
$\rightsquigarrow$ $\|f(x_n) - f(x_0)\| \leq \varepsilon$, $\forall n \geq n_0$.
$\rightsquigarrow$ $(f(x_0))$ konvergiert gegen $f(x_0)$.

($\Leftarrow$) Es sei $\varepsilon > 0$ beliebig gewählt.

Annahme: $\forall \delta > 0 : f(D(f) \cap B(x_0, \delta)) \not\subseteq B(f(x_0), \varepsilon)$

$\rightsquigarrow \quad \forall n \in \mathbb{N} \exists y_n \in f(D(f) \cap B(x_0, \frac{1}{n})) \backslash B(f(x_0), \varepsilon)$

$\rightsquigarrow \quad \forall n \in \mathbb{N} \exists x_n \in D(f) : \|x_n - x_0\| \leq \frac{1}{n}$ und $f(x_n) = y_n$.

Nach Voraussetzung konvergiert $(f(x_n)) = (y_n)$ gegen $f(x_0)$.

$\rightsquigarrow \quad \exists n_0 \in \mathbb{N} : f(x_n) = y_n \in B(f(x_0), \varepsilon), \forall n \geq n_0$.

$\rightsquigarrow$ Widerspruch! $\square$

Bemerkung 5.4

Die ε–δ–Charakterisierung der Stetigkeit einer Funktion f geht auf Weierstraß zurück. Das dort auftretende δ hängt im allgemeinen von ε und x_0 ab, d.h. $\delta = \delta(\varepsilon, x_0)$! Die Limes–Charakterisierung der Stetigkeit in Satz 5.3 erweist sich häufig als sehr bequem zum Nachweis der Stetigkeit von Funktionen, während sich die ε–δ–Charakterisierung für Negativ–Aussagen als günstig erweist.

Beispiele 5.5

Wir betrachten durchweg Funktionen $f : D(f) \longrightarrow \mathbb{R}$ mit $D(f) \subseteq \mathbb{R}$. Für solche Funktionen f bedeutet Stetigkeit in $x_0 \in D(f)$: $\forall \varepsilon > 0 \exists \delta > 0 : \forall x \in D(f)$ mit $|x - x_0| < \delta$ gilt $|f(x) - f(x_0)| < \varepsilon$.

a) *Polynome: $f(x) := \sum_{j=0}^{\ell} a_j x^j, \quad \forall x \in \mathbb{R} \qquad (a_j \in \mathbb{R},\ j = 0, \ldots, \ell)$.*
   Sei $x_0 \in \mathbb{R}$ beliebig und (x_n) eine Folge in $\mathbb{R}$ mit $x_n \longrightarrow x_0$. Aus Satz 3.6 folgt dann

$$f(x_n) = \sum_{j=0}^{\ell} a_j x_n^j \longrightarrow \sum_{j=0}^{\ell} a_j x_0^j = f(x_0).$$

   Also ist f stetig.

b) $f(x) := \begin{cases} 1, & x > x_0 \\ 0, & x \leq x_0 \end{cases}, \qquad (x_0 \in \mathbb{R})$.
   f ist nicht stetig in x_0, wegen $f(B(x_0, \delta)) \not\subseteq B(f(x_0), \frac{1}{2}), \forall \delta > 0$.

c) $f(x) := \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{R} \backslash \mathbb{Q} \end{cases}$ *ist in keinem $x_0 \in \mathbb{R}$ stetig (in jeder Kugel um ein beliebiges $x_0 \in \mathbb{R}$ befinden sich rationale und irrationale Zahlen. Deshalb gilt $f(B(x_0, \delta)) = \{0, 1\}$ für jedes $\delta > 0$.)*

d) *$f : \mathbb{N} \longrightarrow \mathbb{R},\ f(2\ell) := 1,\ f(2\ell - 1) := 0,\ \forall \ell \in \mathbb{N}$.*
   f ist stetig, da konvergente Folgen in $\mathbb{N}$ ab einem gewissen Index konstant sein müssen.

Wir kommen nun zu globalen Eigenschaften stetiger Funktionen.

Satz 5.6

Für eine Funktion $f : D(f) \longrightarrow \mathbb{R}^k$ mit $D(f) \subseteq \mathbb{R}^m$ sind folgende Aussagen äquivalent:

a) *f ist stetig.*

b) *Für jede offene Menge $A \subseteq \mathbb{R}^k$ existiert eine offene Menge $B \subseteq \mathbb{R}^m$, so daß $f^{-1}(A) = D(f) \cap B$.*

c) *Für jede abgeschlossene Menge $A \subseteq \mathbb{R}^k$ existiert eine abgeschlossene Menge $B \subseteq \mathbb{R}^m$, so daß $f^{-1}(A) = D(f) \cap B$.*

Ist $D(f) = \mathbb{R}^m$, so sind also stetige Urbilder offener bzw. abgeschlossener Mengen offen bzw. abgeschlossen.

Beweis: Wir beweisen zunächst die Aussage b) $\Longrightarrow$ a):
Es seien $x_0 \in D(f)$ und $\varepsilon > 0$ beliebig gewählt. Wir wissen dann,
$f^{-1}(\overset{o}{B}(f(x_0), \varepsilon)) = D(f) \cap B$ und $B \subseteq \mathbb{R}^m$ ist offen.
Außerdem gilt $x_0 \in f^{-1}(\overset{o}{B}(f(x_0), \varepsilon))$ und damit $x_0 \in B$. Deshalb existiert ein $\delta > 0$ mit $\overset{o}{B}(x_0, \delta) \subseteq B$ und damit

$$f(D(f) \cap \overset{o}{B}(x_0, \delta)) \subseteq \overset{o}{B}(f(x_0), \varepsilon).$$

Als nächstes beweisen wir noch c) $\Longrightarrow$ b) und lassen a) $\Longrightarrow$ c) als Übung.
Es sei $A \subseteq \mathbb{R}^k$ offen. Folglich ist $\mathbb{R}^k \setminus A$ abgeschlossen und wir wissen, daß $f^{-1}(\mathbb{R}^k \setminus A) = D(f) \cap B$ mit einer abgeschlossenen Menge $B \subseteq \mathbb{R}^m$ gilt. Außerdem gilt

$$D(f) = f^{-1}(A) \cup f^{-1}(\mathbb{R}^k \setminus A) = f^{-1}(A) \cup (D(f) \cap B),$$

woraus $f^{-1}(A) = D(f) \cap (\mathbb{R}^m \setminus B)$ folgt. □

Bemerkung 5.7
Achtung: Stetige Funktionen bilden im allgemeinen offene bzw. abgeschlossene Mengen nicht in offene bzw. abgeschlossene Mengen ab. Satz 5.6 gilt für Urbilder, aber nicht für Bilder!
Beispiele: $f : \mathbb{R} \longrightarrow \mathbb{R},\ f(x) := x^2 \rightsquigarrow f(\,]-1,1[\,) = [0,1[,$
$f :]0,+\infty[\longrightarrow \mathbb{R},\ f(x) := \frac{1}{x} \rightsquigarrow f(\,[1,\infty[\,) =]0,1].$

Die nächste Aussage zeigt nun, daß eine andere wichtige Mengeneigenschaft durch stetige Abbildungen unverändert bleibt.

Satz 5.8
Es sei $f : D(f) \longrightarrow \mathbb{R}^k$ stetig und $A \subseteq \bar{A} \subseteq D(f) \subseteq \mathbb{R}^m$. Ist A (relativ) kompakt, so ist $f(A)$ (relativ) kompakt.

Beweis: Es sei A relativ kompakt und es sei (y_n) eine Folge in $f(A)$.
Wir zeigen: y_n besitzt eine konvergente Teilfolge.
Zunächst existiert ein $x_n \in A$ mit $y_n = f(x_n)$, $\forall n \in \mathbb{N}$. Da A relativ kompakt ist, existiert eine Teilfolge $(x_{k(n)})$ von(x_n), die gegen ein $x \in \mathbb{R}^m$ konvergiert. Da nach Satz 2.53 das Element x zu $\bar{A}$ gehört, folgt $x \in D(f)$. Deshalb gilt:

$$y_{k(n)} = f(x_{k(n)}) \overrightarrow{n \to \infty}\ f(x)\ .$$

Ist A kompakt (d.h. relativ kompakt und abgeschlossen), so folgt aus dem ersten Teil zunächst: $f(A)$ ist relativ kompakt.
Es bleibt zu zeigen: Jeder Häufungspunkt y von $f(A)$ ist in $f(A)$ (Satz 2.48).
$\rightsquigarrow$ $\exists(y_n)$ in $f(A)$ mit $y_n \to y \rightsquigarrow \exists(x_n)$ in A mit $f(x_n) = y_n$
$\rightsquigarrow$ wieder über eine Teilfolge von (x_n) argumentiert, folgt:
$\exists x \in A$ mit $y_n \longrightarrow f(x) \rightsquigarrow y = f(x)$, d. h., $y \in f(A)$. $\square$

Eine wesentliche Schlußfolgerung aus Satz 5.8 ist der Hauptsatz der mathematischen Optimierung.

Satz 5.9 *(Weierstraß)*
Es sei $\emptyset \neq K \subseteq \mathbb{R}^m$ kompakt und $f : K \longrightarrow \mathbb{R}$ sei stetig.
Dann existieren $\underline{x}$, $\overline{x} \in K$, so daß

$$f(\underline{x}) = \inf\{f(x) : x \in K\} \quad \text{und} \quad f(\overline{x}) = \sup\{f(x) : x \in K\}.$$

Beweis: Nach Satz 5.8 ist $f(K)$ kompakt in $\mathbb{R}$, d. h., beschränkt und abgeschlossen (Satz 2.57). Damit sind $a := \inf f(K)$ und $b := \sup f(K)$ in $\mathbb{R}$ und gehören zu $f(K)$ oder sind Häufungspunkte von $f(K)$. Da $f(K)$ abgeschlossen ist, gehören sie auch im letzteren Fall zu $f(K)$. $\square$

Jede reellwertige stetige Funktion nimmt also auf einer kompakten Menge ihr Minimum und Maximum an. Stetige injektive Funktionen erlauben überdies die folgende Schlußfolgerung.

Satz 5.10
Es sei $K \subseteq \mathbb{R}^m$ kompakt, $f : K \longrightarrow \mathbb{R}^k$ sei stetig und injektiv.
Dann ist $f^{-1} : f(K) \longrightarrow \mathbb{R}^m$ stetig.

Beweis: Es sei $A \subseteq \mathbb{R}^m$ abgeschlossen. Zu zeigen: $(f^{-1})^{-1}(A) = f(A \cap K)$ ist abgeschlossen (Satz 5.6!).
Nach Voraussetzung ist $A \cap K$ kompakt und deshalb nach Satz 5.8 auch $f(A \cap K)$.
Kompakte Mengen sind aber abgeschlossen. $\square$

Beispiel 5.11 *(Exponentialfunktion)*
Wir zeigen: $\exp : \mathbb{C} \longrightarrow \mathbb{C}$ *ist stetig. ($\mathbb{C}$ mit $|\cdot|$ enspricht $\mathbb{R}^2$ mit Euklidischer Norm.)*
Beweis:
Es sei $z_0 \in \mathbb{C}$ und (z_n) eine Folge in $\mathbb{C}$ mit $z_n \to z_0$.
Nach Satz 3.25 gilt für alle $n \in \mathbb{N}$:

$$\exp(z_n) - \exp(z_0) = \exp(z_0)(\exp(z_n - z_0) - 1).$$

Das Additionstheorem von exp *reduziert also die Stetigkeit auf $\mathbb{C}$ auf die Stetigkeit im Punkt* 0 *und es genügt zu zeigen:*
Ist $(\tilde{z}_n)$ eine Folge in $\mathbb{C}$ mit $\tilde{z}_n \to 0$, so gilt $\exp(\tilde{z}_n) \to 1$.
Nach Definition gilt für jedes $z \in \mathbb{C}$ mit $|z| \leq \frac{1}{2}$

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

$$\rightsquigarrow |\exp(z)-1| \le |z| \sum_{k=0}^{\infty} \frac{|z|^k}{(k+1)!} \le |z| \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = 2|z|.$$

Also gilt für hinreichend großes $n \in I\!N$:

$$|\exp(\tilde{z}_n)-1| \le 2|\tilde{z}_n| \xrightarrow[n\to\infty]{} 0. \qquad \square$$

Nach Satz 3.25 ist aber $\exp : \mathbb{R} \to \mathbb{R}$ *injektiv und die inverse Abbildung* $\ln : \exp(\mathbb{R}) \to \mathbb{R}$ *ist definiert (vgl. Def. 3.26).*

Wir zeigen: $\ln : \exp(\mathbb{R}) \to \mathbb{R}$ *ist stetig.*
Beweis: Es sei $y \in \exp(\mathbb{R})$ beliebig gewählt. Wir wählen $a, b \in \mathbb{R}$ mit $a < b$ und $\exp(a) < y < \exp(b)$. Das Intervall $[a, b]$ ist kompakt, $\exp$ *ist stetig und injektiv auf $[a, b]$. Nach Satz 5.10 ist* $\ln$ *stetig auf* $[\exp(a), \exp(b)]$*, also speziell in y.* $\square$

Direkt aus der Definition der Exponentialfunktion folgt auch, daß sin, cos, sinh *und* cosh *als Funktionen von $\mathbb{R}$ in $\mathbb{R}$ stetig sind (Übung!).*

Schließlich kommen wir noch zum Zwischenwertsatz, mit dem auch die noch offene Frage $\exp(\mathbb{R}) =]0, +\infty[$ geklärt werden kann.

Satz 5.12 *(Zwischenwertsatz)*
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ stetig. Dann ist $f([a, b])$ ein Intervall.

Beweis:
Zu zeigen: Aus $x, \tilde{x} \in [a, b]$ und $f(x) \le y \le f(\tilde{x})$ folgt $y \in f([a, b])$.
Wir betrachten $A := \{x \in [a, b] : f(x) \le y\}$ und $\bar{x} := \sup A$, was möglich ist, da A nichtleer und beschränkt ist. Da A abgeschlossen ist, folgt $\bar{x} \in A$ und $f(\bar{x}) \le y$.
Annahme: $f(\bar{x}) < y$, d.h. $y \notin f([a, b])$.
Nach Satz 5.6 existieren offene Mengen B, $\tilde{B}$ in $\mathbb{R}$, so daß $f^{-1}(]-\infty, y[) = [a, b] \cap B$ und $f^{-1}(]y, \infty[) = [a, b] \cap \tilde{B}$. Dann gilt: $B \cap \tilde{B} = \emptyset$ und $[a, b] = [a, b] \cap (B \cup \tilde{B})$.
Es seien nun $x \in B$ und $\tilde{x} \in \tilde{B}$. Wegen $x \ne \tilde{x}$ sei o.B.d.A. $x < \tilde{x}$.
Wir definieren $z := \sup B \cap [x, \tilde{x}]$. Dann gilt $z \in [x, \tilde{x}] \subseteq [a, b]$ z muß entweder zu B oder zu $\tilde{B}$ gehören.
1. Fall: $z \in B$. $\rightsquigarrow \exists \varepsilon > 0$ mit $[z, z+\varepsilon] \subseteq B \cap [x, \tilde{x}]$ (wegen B offen und $z < \tilde{x}$).
$\rightsquigarrow$ Widerspruch zur Definition von z.
2. Fall: $z \in \tilde{B}$.$\rightsquigarrow \exists \varepsilon > 0$ mit $[z-\varepsilon, z] \subseteq \tilde{B} \cap [x, \tilde{x}]$ (wegen $\tilde{B}$ offen und $x < z$). $\rightsquigarrow z - \varepsilon \notin B \rightsquigarrow \sup B \cap [x, \tilde{x}] \le z - \varepsilon \rightsquigarrow$ Widerspruch zur Definition von z.
Also muß die Annahme falsch sein und es gilt $f(\bar{x}) = y$. $\square$

Übung 5.13 *Man zeige* $\exp(\mathbb{R}) =]0, +\infty[$ *mit Hilfe von Satz 5.12.*
(Hinweis: Man verwende, daß für jedes $y > 0$ Punkte x und $\tilde{x}$ in $\mathbb{R}$ existieren, so daß $\exp(x) \le y \le \exp(\tilde{x})$.)

Definition 5.14

(a) $f : D(f) \longrightarrow \mathbb{R}^k$ mit $D(f) \subseteq \mathbb{R}^m$ heißt gleichmäßig stetig, falls
$\forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) > 0 : \forall x, \tilde{x} \in D(f)$ *mit* $\|x - \tilde{x}\| < \delta$ *gilt* $\|f(x) - f(\tilde{x})\| < \varepsilon$.

(b) $f : D(f) \longrightarrow \mathbb{R}^k$ *heißt Lipschitzstetig, falls ein* $L > 0$ *existiert, so daß für alle* $x, \tilde{x} \in D(f) : \|f(x) - f(\tilde{x})\| \leq L\|x - \tilde{x}\|$.

Bemerkung 5.15
Für eine Abbildung $f : D(f) \to \mathbb{R}^k$ *gilt nach 5.1 bzw. 5.14:*
f *ist Lipschitzstetig* $\Longrightarrow$ f *ist gleichmäßig stetig (mit* $\delta(\varepsilon) := \frac{\varepsilon}{L}$*)* $\Longrightarrow$ f *ist stetig.*
Ein Vergleich des Begriffs ,,gleichmäßige Stetigkeit" mit dem ε*–*δ*–Kriterium in Definition 5.1 zeigt, daß dort* δ *von* ε *und* x_0 *abhängt und daß in 5.14* $\delta = \delta(\varepsilon)$ *gleichmäßig bzgl. der Elemente aus* $D(f)$ *gewählt werden kann.*

Satz 5.16
Es sei $f : K \to \mathbb{R}^k$ *stetig und* $K \subseteq \mathbb{R}^m$ *kompakt. Dann ist* f *gleichmäßig stetig.*

Beweis: Annahme: f ist nicht gleichmäßig stetig auf K. Dann gilt:
$\exists \varepsilon_0 > 0 \forall n \in \mathbb{N} \exists x_n, y_n \in K : \|x_n - y_n\| \leq \frac{1}{n}$ und $\|f(x_n) - f(y_n)\| \geq \varepsilon_0$.
Da K kompakt ist, existieren Teilfolgen von $(x_n), (y_n)$, die gegen $x, y \in K$ konvergieren.
Durch Grenzübergang: $x = y$ und $0 = \|f(x) - f(y)\| \geq \varepsilon_0 > 0 \rightsquigarrow$ Widerspruch! $\square$

Beispiele 5.17 a) $f : \mathbb{R} \to \mathbb{R}$, $f(x) := \begin{cases} x \cdot \sin(\frac{1}{x}), & x \neq 0 \\ 0, & x = 0 \end{cases}$ *, ist stetig.*

Beweis: Es sei zunächst $x_0 \neq 0$. *Dann ist die Funktion* $g(x) := \frac{1}{x}$ *stetig in* x_0 *und nach Beispiel 5.11 (Stetigkeit von* $\sin$*) ist auch* $\sin(\frac{1}{x})$ *stetig in* x_0.
Hieraus folgt schließlich die Stetigkeit von f *in* x_0.
Sei nun $x_0 = 0$. *Dann gilt für beliebige* $x \neq 0$*:*

$$|f(0) - f(x)| \leq |x||\sin\frac{1}{x}| \leq |x| \quad \text{da nach Definition} \quad |\sin x| \leq 1,\ \forall x \in \mathbb{R}.$$

Ist also $\varepsilon > 0$ *beliebig vorgegeben, so gilt mit* $\delta := \varepsilon$*:*

$$|f(0) - f(x)| < \varepsilon, \quad \text{falls} \quad |x| < \delta,$$

d.h. f *ist auch stetig in* $x_0 = 0$. $\square$

b) $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^2$, $\forall x \in \mathbb{R}$, *ist stetig, aber nicht gleichmäßig stetig (auf* $\mathbb{R}$*).*
Ursache:

$$|x^2 - y^2| = |x + y||x - y|, \quad \forall x, y \in \mathbb{R},$$

und $|x + y|$ *kann beliebig groß werden!*
Jedoch ist f *sogar Lipschitzstetig auf jeder beschränkten Menge!*

c) $f : [0, +\infty[\to \mathbb{R}$, $f(x) := x^{\frac{1}{2}}$, $\forall x \in [0, +\infty[$, *ist gleichmäßig stetig, aber nicht Lipschitzstetig (auch nicht auf* $[0, 1]$*).*
Beweis:
Da $g : [0, +\infty[\to \mathbb{R}$, $g(y) := y^2$, $\forall y \in \mathbb{R}$, *stetig und injektiv ist, folgt aus Satz*

5.10, daß $f = g^{-1}$ ebenfalls stetig ist.
Ist nun für $x, y \in [0, +\infty[$, $x \geq 1$ oder $y \geq 1$, so gilt:

$$|x^{\frac{1}{2}} - y^{\frac{1}{2}}| = \frac{|x-y|}{x^{\frac{1}{2}} + y^{\frac{1}{2}}} \leq |x-y|,$$

d.h. für solche $x, y \in [0, +\infty[$ ist f gleichmäßig stetig. Auf $[0,1]$ resultiert die gleichmäßige Stetigkeit von f aus Satz 5.17 (da f stetig und $[0,1]$ kompakt).
Also ist insgesamt f gleichmäßig stetig auf $[0, +\infty[$.
Wäre f Lipschitzstetig, so müßte insbesondere die Menge

$$\left\{ \frac{|f(x) - f(0)|}{|x|} : x > 0 \right\} \text{ beschränkt sein.}$$

Für $x_n := \frac{1}{n}$ $(n \in \mathbb{N})$ gilt aber: $\frac{|f(x_n) - f(0)|}{|x_n|} = \sqrt{n} \underset{n\to\infty}{\longrightarrow} +\infty.$ $\square$

d) $\|\cdot\| : \mathbb{R}^m \longrightarrow \mathbb{R}$ *ist Lipschitzstetig.*
($|\|x\| - \|y\|| \leq \|x - y\|$, $\forall x, y \in \mathbb{R}^m$ *vgl. auch 2.38.)*

e) $f : \mathbb{R} \to \mathbb{C}$, $f(x) := \exp(ix)$, $\forall x \in \mathbb{R}$, *ist Lipschitzstetig (mit Konstante $L := 1$).*
Eine wichtige Konsequenz dieser Aussage ist, daß auch die Funktionen sin *und* cos *Lipschitzstetig mit Konstante $L := 1$ sind.*
<u>Beweis:</u>
Zunächst folgt für alle $x, y \in \mathbb{R}$ aus Satz 3.25:

$$\begin{aligned} f(x+y) - f(x-y) &= \exp(ix)(\exp(iy) - \exp(-iy)) \\ &= \exp(ix)(\exp(iy) - \overline{\exp(iy)}) \\ &= 2i\exp(ix)\,\mathrm{Im}(\exp(iy)) = 2i\exp(ix)\sin y \end{aligned}$$

$\rightsquigarrow |f(x+y) - f(x-y)| = 2|\sin y|$
$\rightsquigarrow |f(x) - f(\tilde{x})| = 2|\sin(\frac{x-\tilde{x}}{2})|, \quad \forall x, \tilde{x} \in \mathbb{R}.$
Wir zeigen schließlich $|\sin y| \leq |y|$, $\forall y \in \mathbb{R}$, woraus dann die Aussage folgt.
Zunächst ist wegen $|\sin(-y)| = |-\sin y| = |\sin y|$ klar, daß man sich auf den Fall $y \geq 0$ beschränken kann. Für $y \geq 1$ gilt offenbar

$$|\sin y| \leq |\exp(iy)| = 1 \leq y.$$

Es sei nun $y \in [0, 1[$. Dafür gilt

$$\sin y - y = \sum_{k=1}^{\infty} (-1)^k \frac{y^{2k+1}}{(2k+1)!} \leq 0,$$

d.h. $\sin y \leq y$. $\square$

Satz 5.18
Es seien $f : D(f) \to \mathbb{R}^k$ und $g : D(g) \to \mathbb{R}^l$ Funktionen mit $D(f) \subseteq \mathbb{R}^m$ und $R(f) \subseteq D(g) \subseteq \mathbb{R}^k$.
Ist f stetig in $x_0 \in D(f)$ und g stetig in $f(x_0) \in \mathbb{R}^k$, so ist die Funktion $h := g \circ f : D(f) \to \mathbb{R}^l$ stetig in x_0. Sind f und g stetig, so ist auch h stetig.

Beweis: Da der zweite Teil der Aussage sofort aus dem ersten Teil folgt, beweisen wir diesen. Es sei $\varepsilon > 0$ beliebig gewählt und wir betrachten die Kugel $B(h(x_0), \varepsilon) = B(g(f(x_0)), \varepsilon)$. Da g in $f(x_0)$ stetig ist, existiert ein $\eta > 0$, so daß $g(B(f(x_0), \eta)) \subseteq B(h(x_0), \varepsilon)$. Da auch f in x_0 stetig ist, existiert ein $\delta > 0$, so daß $f(B(x_0, \delta)) \subseteq B(f(x_0), \eta)$. Insgesamt folgt:

$$h(B(x_0, \delta)) = g(f(B(x_0, \delta))) \subseteq g(B(f(x_0), \eta)) \subseteq B(h(x_0), \varepsilon)$$

und h ist stetig in x_0. $\square$

6 Differentialrechnung

6.1 Differentialrechnung reeller Funktionen einer reellen Veränderlichen

Es sei $f : D(f) \to \mathbb{R}$, $D(f) \subseteq \mathbb{R}$ und $x_0 \in \mathrm{int}(D(f))$.
Dann heißt die Funktion $\varphi : \{h \in \mathbb{R}\backslash\{0\} : x_0 + h \in D(f)\} \to \mathbb{R}$, definiert durch

$$\varphi(h) := \frac{f(x_0 + h) - f(x_0)}{h}, \quad \forall h \in D(\varphi),$$

Differenzenquotient von f in x_0.

Definition 6.1
Falls der Grenzwert von φ in $h = 0$ existiert, d.h.

$$\lim_{h \to 0} \varphi(h) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} =: f'(x_0) =: \frac{df}{dx}(x_0),$$

so heißt $f'(x_0)$ die Ableitung (oder: der Differentialquotient) von f in x_0. Man sagt dann, daß f differenzierbar in x_0 ist.
Ist $D(f)$ offen, so heißt f differenzierbar, falls f in jedem $x_0 \in D(f)$ differenzierbar ist. In diesem Fall heißt die Funktion $f' : D(f) \to \mathbb{R}$, wobei $f'(x_0)$ für jedes $x_0 \in D(f)$ wie oben definiert ist, die Ableitung von f.

Bemerkung 6.2
Da nach Voraussetzung gilt, daß $x_0 \in int\ (D(f))$, existiert $\varepsilon > 0$, so daß $[x_0 - \varepsilon, x_0 + \varepsilon] \subseteq D(f) \rightsquigarrow [-\varepsilon, \varepsilon]\backslash\{0\} \subseteq D(\varphi)$. Also ist 0 Häufungspunkt von $D(\varphi)$ und der Grenzwert $\lim_{h \to 0} \varphi(h)$ ist definiert gemäß Def. 5.19.
Folglich ist also f differenzierbar in x_0 mit Ableitung $f'(x_0)$, falls für jede Folge (h_n) mit $h_n \to 0$, $h_n \neq 0$ und $x_0 + h_n \in D(f)$, $\forall n \in \mathbb{N}$, gilt:

$$\lim_{n \to \infty} \frac{f(x_0 + h_n) - f(x_0)}{h_n} = f'(x_0).$$

Diese Definition geht auf Cauchy (1821) zurück. Der Differentialkalkül war jedoch bereits von Newton und Leibniz entwickelt worden, allerdings nicht auf exakter Grundlage, da ohne Limesbegriff. Ist $D(f)$ ein Intervall und x_0 ein Randpunkt von $D(f)$, so gilt $D(\varphi) \subseteq]0, +\infty[$ bzw. $D(\varphi) \subseteq]-\infty, 0[$, und es lassen sich einseitige Ableitungen von f in x_0 als einseitige Grenzwerte von φ in $h = 0$, d.h.,

$$\lim_{h\to 0-} \varphi(h) = f'(x_0-), \quad \lim_{h\to 0+} \varphi(h) = f'(x_0+),$$

im Sinne von Def. 5.25 definieren.
Ist $D(f)$ ein abgeschlossenes Intervall, so werden wir f differenzierbar auf $D(f)$ nennen, falls f in allen inneren Punkten differenzierbar ist und in den Randpunkten die einseitigen Ableitungen existieren. Natürlich lassen sich einseitige Ableitungen auch in inneren Punkten von $D(f)$ definieren.

Bemerkung 6.3
Geometrische und analytische Interpretation der Ableitung:
Zu gegebenem $f : D(f) \to \mathbb{R}$ und $x_0 \in int\ (D(f)) \subseteq \mathbb{R}$ betrachten wir die beiden folgenden Funktionen:
$S(x) := f(x_0) + \frac{f(x_0+h)-f(x_0)}{h}(x - x_0)$,
$T(x) := f(x_0) + f'(x_0)(x - x_0)$ *(falls $f'(x_0)$ existiert).*
S ist die "Sekante" und T die "Tangente" an f (in x_0 und $x_0 + h$ bzw. in x_0), d.h. $f'(x_0)$ ist der Anstieg der Tangente, $\frac{f(x_0+h)-f(x_0)}{h}$ der der Sekante.
Analytisch bedeutet die Differenzierbarkeit von f:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \omega(f, x - x_0) \ \text{("Restglied")}$$

d.h. die Funktion f läßt sich lokal durch eine lineare Funktion approximieren (die "Tangente"), d.h. $\exists a \in \mathbb{R}$:

$$f(x) = f(x_0) + a(x - x_0) + \omega(f; x - x_0), \quad \forall x \in D(f)$$

wobei $\lim_{x\to x_0} \frac{\omega(f;x-x_0)}{x-x_0} = 0$.
Diese Formulierung der Definition der Differenzierbarkeit ist verallgemeinerbar auf den mehrdimensionalen Fall ! (Kap. 6.2)

Bemerkung 6.4
Die Definition der Ableitung in 6.1 kann natürlich auf Funktionen $f : D(f) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}$, verallgemeinert werden, indem der Grenzwert in $\mathbb{R}^m$ (was gleichbedeutend ist mit "komponentenweise", vgl. Kap. 3.1) gebildet wird, d.h. $f'(x_0) \in \mathbb{R}^m$.
Eine direkte Verallgemeinerung auf den Fall $D(f) \subseteq \mathbb{R}^k$, $k > 1$, ist wegen der nicht erklärten Division aber unmöglich (Ausnahme $D(f) \subseteq \mathbb{C}$)! Jedoch zeigt sich, daß die "Linearisierungs"–Idee aus Bem. 6.3 zur Definition einer Ableitung in diesem Fall geeignet ist (vgl. Kap. 6.2).

Satz 6.5
Es seien $f : D(f) \to \mathbb{R}$, $D(f) \subseteq \mathbb{R}$, f sei in x_0 differenzierbar.

a) *Dann ist f in x_0 stetig.*

b) *Sei $g : D(g) \to \mathbb{R}$, $D(g) \subseteq \mathbb{R}$, in $x_0 \in \operatorname{int} D(f) \cap D(g)$ differenzierbar. Dann sind auch die Funktionen (von $D(f) \cap D(g)$ in $\mathbb{R}$) $\alpha f + \beta g$, fg und $\frac{f}{g}$ (falls $g(x_0) \neq 0$) in x_0 differenzierbar, und es gelten die Regeln:*

$$\begin{aligned}
(\alpha f + \beta g)'(x_0) &= \alpha f'(x_0) + \beta g'(x_0) \quad (\alpha, \beta \in \mathbb{R}),\\
(fg)'(x_0) &= f(x_0)g'(x_0) + f'(x_0)g(x_0) \quad \text{("Produktregel")},\\
(\frac{f}{g})'(x_0) &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2} \quad \text{("Quotientenregel")}.
\end{aligned}$$

Beweis:

a) Sei $\varepsilon > 0$ beliebig. Es gilt nach Voraussetzung:

$$\lim_{h \to 0} \frac{1}{|h|}|f(x_0 + h) - f(x_0)| = |f'(x_0)|,$$

$\rightsquigarrow \exists \delta_0 > 0 : \frac{1}{|h|}|f(x_0 + h) - f(x_0)| \leq |f'(x_0)| + 1$, falls $|h| < \delta_0$,
$\rightsquigarrow |f(x_0 + h) - f(x_0)| \leq (|f'(x_0)| + 1)|h|$, falls $|h| < \delta_0$,
Wir wählen $0 < \delta \leq \delta_0$ so, daß $(|f'(x_0)| + 1)\delta < \varepsilon$
$\rightsquigarrow \forall x \in D(f)$ mit $|x - x_0| < \delta$ ist $|f(x) - f(x_0)| \leq (|f'(x_0)| + 1)|h| < \varepsilon$
$\rightsquigarrow f$ ist stetig in x_0.

b) Mit $h \neq 0$, so daß $x_0 + h \in D(f) \cap D(g)$, erhält man aus den Rechenregeln für Grenzwerte (Satz 3.6):

$$\frac{\alpha f(x_0 + h) + \beta g(x_0 + h) - \alpha f(x_0) - \beta g(x_0)}{h} =$$

$$= \alpha \frac{f(x_0 + h) - f(x_0)}{h} + \beta \frac{g(x_0 + h) - g(x_0)}{h} \underset{h \to 0}{\longrightarrow} \alpha f'(x_0) + \beta g'(x_0),$$

$$\frac{f(x_0 + h)g(x_0 + h) - f(x_0)g(x_0)}{h} =$$

$$= f(x_0 + h)\frac{g(x_0 + h) - g(x_0)}{h} + g(x_0)\frac{f(x_0 + h) - f(x_0)}{h}$$

$$\underset{h \to 0}{\longrightarrow} f(x_0)g'(x_0) + g(x_0)f'(x_0) \quad \text{(nach Vor. bzw. aus a)).}$$

$$\begin{aligned}
\frac{\frac{f(x_0+h)}{g(x_0+h)} - \frac{f(x_0)}{g(x_0)}}{h} &= \frac{1}{h}\frac{f(x_0 + h)g(x_0) - f(x_0)g(x_0 + h)}{g(x_0 + h)g(x_0)}\\
&= \frac{1}{h}\frac{(f(x_0 + h) - f(x_0))g(x_0) - f(x_0)(g(x_0 + h) - g(x_0))}{g(x_0 + h)g(x_0)}\\
&= \frac{\frac{f(x_0+h) - f(x_0)}{h}g(x_0) - \frac{g(x_0+h) - g(x_0)}{h}f(x_0)}{g(x_0 + h)g(x_0)}\\
&\underset{h \to 0}{\longrightarrow} \frac{g(x_0)f'(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}
\end{aligned}$$

(nach Vor. bzw. aus a), wobei noch angemerkt werden muß, daß wegen $g(x_0) \neq 0$ und der Stetigkeit von g in x_0, auch $g(x_0 + h) \neq 0$ für hinreichend kleines $|h|$ gilt). $\square$

Satz 6.6 *(Ableitung der inversen Funktion)*
Sei $f : D(f) \to \mathbb{R}$ injektiv, stetig und in $x_0 \in \operatorname{int}(D(f))$ differenzierbar. Dann ist die inverse Funktion $f^{-1} : R(f) \to \mathbb{R}$ in $y_0 = f(x_0)$ differenzierbar, wenn $f'(x_0) \neq 0$ ist, und es gilt

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

Beweis:
Für f muß auf $[x_0 - \varepsilon, x_0 + \varepsilon] \subseteq D(f)$ mit einem geeignet gewählten $\varepsilon > 0$ gelten, daß aus $x, y \in [x_0 - \varepsilon, x_0 + \varepsilon]$ und $x < y$ folgt $f(x) < f(y)$ oder $f(y) < f(x)$, d.h., f ist streng monoton (o.B.d.A. wachsend) und f^{-1} ist auf $f([x_0 - \varepsilon, x_0 + \varepsilon])$ stetig (vgl. Satz 5.10). Nach Satz 5.12 ist $f([x_0 - \varepsilon, x_0 + \varepsilon])$ ein Intervall. Da f streng monoton wachsend ist, muß gelten

$$y_0 \in]f(x_0 - \varepsilon), f(x_0 + \varepsilon)[\subseteq R(f),$$

d.h. y_0 ist innerer Punkt von $R(f)$.
Es sei nun (y_n) eine Folge in $R(f)$ mit $y_n \neq y_0$, $y_n \to y_0$.
$\rightsquigarrow x_n := f^{-1}(y_n) \underset{n\to\infty}{\longrightarrow} f^{-1}(y_0) = x_0$ und $x_n \neq x_0$.

$$\rightsquigarrow \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{x_n - x_0}{f(x_n) - f(x_0)} = \frac{1}{\frac{f(x_n)-f(x_0)}{x_n - x_0}} \underset{n\to\infty}{\longrightarrow} \frac{1}{f'(x_0)}$$

$\rightsquigarrow f^{-1}$ ist in y_0 differenzierbar, und die angegebene Formel gilt. $\square$

Beispiel 6.7

a) *$f(x) := x^k$, $\forall x \in \mathbb{R}$ $(k \in \mathbb{N})$. Dann gilt für $x_0 \in \mathbb{R}$, $h \neq 0$ bel.:*

$$\begin{aligned}\frac{(x_0+h)^k - x_0^k}{h} &= \frac{1}{h}\left(\sum_{i=0}^{k}\binom{k}{i}x_0^{k-i}h^i - x_0^k\right) = \sum_{i=1}^{k}\binom{k}{i}x_0^{k-i}h^{i-1} \quad \textit{(1.11 !)}\\ &\underset{h\to 0}{\longrightarrow} \binom{k}{1}x_0^{k-1} = k\,x_0^{k-1}\end{aligned}$$

Also ist f differenzierbar in x_0 mit $f'(x_0) = kx_0^{k-1}$.
Als Konsequenz aus Satz 6.5 b) sind alle Polynome auf $\mathbb{R}$ differenzierbar.
Aus demselben Resultat folgt für $g(x) := x^{-k}$, $\forall x \in \mathbb{R}\backslash\{0\}$ $(k \in \mathbb{N})$:

$$g'(x_0) = (\frac{1}{f})'(x_0) = -\frac{f'(x_0)}{(f(x_0))^2} = -k\frac{x_0^{k-1}}{x_0^{2k}} = -kx_0^{-k-1}.$$

Als Konsequenz aus Satz 6.5 a) sind alle rationalen Funktionen, d.h. Quotienten zweier Polynome, in jedem Punkt ihres Definitionsbereiches differenzierbar.

b) $f(x) := \exp(x)$, $\forall x \in \mathbb{R}$. *Seien* $x_0 \in \mathbb{R}$, $h \neq 0$ *bel.*

$$\begin{aligned}\rightsquigarrow \frac{1}{h}(\exp(x_0+h) - \exp(x_0)) &= \exp(x_0)\left[\frac{1}{h}(\exp(h)-1)\right] \quad \textit{(3.47 !)} \\ &= \exp(x_0)\left(\sum_{k=1}^{\infty} \frac{h^{k-1}}{k!}\right)\end{aligned}$$

und $\lim_{h\to 0}\sum_{k=1}^{\infty}\frac{h^{k-1}}{k!} = 1$ *(analog zu Bsp. 5.11)!).*
$\rightsquigarrow f'(x_0) = \exp(x_0)$.
Als Konsequenz aus Satz 6.6 ergibt sich, da exp *injektiv, stetig und differenzierbar ist (vgl. 3.47, 4.5):*

$$(f^{-1})'(y_0) = \ln'(y_0) = \frac{1}{\exp(\ln(y_0))} = \frac{1}{y_0}, \quad \forall y_0 > 0,$$

d.h. $\ln :]0, +\infty[\to \mathbb{R}$ *ist ebenfalls differenzierbar.*

Wir betrachten jetzt $g(x) := \exp(ix)$, $\forall x \in \mathbb{R}$.
Seien $x_0 \in \mathbb{R}$, $h \neq 0$ *bel. Analog zu oben erhält man:*

$$\begin{aligned}\frac{1}{h}(g(x_0+h) - g(x_0)) &= g(x_0)\left[\frac{1}{h}\exp(ih) - 1\right] \\ &= g(x_0)\left(\sum_{k=1}^{\infty}\frac{i^k h^{k-1}}{k!}\right) \underset{h\to 0}{\longrightarrow} g(x_0)\cdot i.\end{aligned}$$

Also gilt: $g'(x_0) = i\exp(ix_0)$.
Da ferner gilt $g(x) = \cos x + i \sin x$, *erhalten wir daraus:*

$$\begin{aligned}(\cos)'(x_0) &= \frac{d}{dx}(\mathrm{Re}(\exp(i\cdot)))(x_0) = \mathrm{Re}\,(i\exp(ix_0)) \\ &= \mathrm{Re}\,(i(\cos x_0 + i\sin x_0)) = -\sin x_0\end{aligned}$$

und analog $(\sin)'(x_0) = \cos x_0$.

c) $f(x) := |x|$, $\forall x \in \mathbb{R}$.
f *ist nicht differenzierbar in* $x_0 = 0$, *da gilt:*

$$\frac{f(\frac{1}{n}) - f(0)}{\frac{1}{n} - 0} = 1, \quad \frac{f(-\frac{1}{n}) - f(0)}{-\frac{1}{n} - 0} = -1, \quad \forall n \in I\!N.$$

Es gilt: $f'(0+) = 1$, $f'(0-) = -1$ *(Die einseitigen Ableitungen existieren also).*

d) *Wir konstruieren eine Funktion $f : \mathbb{R} \to \mathbb{R}$, die stetig, aber nirgends differenzierbar ist. Dazu betrachten wir zunächst die folgende "Sägezahn"–Funktion:*

$$g : \mathbb{R} \to \mathbb{R} \quad , \quad g(x) := |x|, \forall x \in \left[-\frac{1}{2}, \frac{1}{2}\right], \text{ und}$$
$$g(x+1) = g(x), \forall x \in \mathbb{R}.$$

g ist nach c) nicht differenzierbar in den Stellen $\frac{j}{2}$, $\forall j \in \mathbb{Z}$, aber stetig.
"Verdichtete" Sägezahn–Funktionen:

$$g_k : \mathbb{R} \to \mathbb{R}, \quad g_k(x) := \frac{1}{2^k} g(2^k x), \quad \forall x \in \mathbb{R}, \forall k \in \mathbb{N}_0.$$

g_k ist stetig und nicht differenzierbar an den Stellen $\frac{j}{2^{k+1}}$, $\forall j \in \mathbb{Z}$.
Wir definieren nun $f(x) := \sum_{k=0}^{\infty} g_k(x), \quad \forall x \in \mathbb{R}$.
Wegen $0 \leq g_k(x) \leq \frac{1}{2^{k+1}}$, $\forall x \in \mathbb{R}\ \forall k \in \mathbb{N}$, ist die Funktionenreihe im folgenden Sinn gleichmßig konvergent:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} : \quad \max_{x \in \mathbb{R}} \left| \sum_{k=n}^{m} g_k(x) \right| < \varepsilon, \quad \forall n_0 \leq n < m.$$

Da alle g_k ($k \in \mathbb{N}$) stetig sind, ist dann auch f stetig.
Man kann nun zeigen, daß f in keinem $x_0 \in \mathbb{R}$ differenzierbar ist (Barner/Flohr, Analysis I, de Gruyter, Kap. 9.1, S. 263).
Beweisskizze: Wir zeigen zunächst, daß f nirgends differenzierbar ist. Es sei $a \in \mathbb{R}$ beliebig gewählt und wir wählen $k_n \in \mathbb{N}$ so, daß $x_n := \frac{k_n}{2^n} \leq a \leq \frac{k_n+1}{s^n} =: y_n$, $\forall n \in \mathbb{N}$, gilt. Dann folgt

$$\frac{f(y_n) - f(x_n)}{y_n - x_n} = \sum_{k=0}^{\infty} \frac{g_k(y_n) - g_k(x_n)}{y_n - x_n} = \sum_{k=0}^{n-1} \frac{g_k(y_n) - g_k(x_n)}{y_n - x_n}$$

und jeder der n Summanden ist entweder gleich 1 oder gleich -1. Deshalb ist der Quotient auf der linken Seite gerade (ungerade), falls n gerade (ungerade) ist. Der Quotient müßte aber gegen $f'(a)$ konvergieren, falls f differenzierbar ist (Übung !).
Die Funktion f ist aber stetig in a, wie man aus folgender Abschätzung erkennt:

$$\begin{aligned} |f(x) - f(a)| &\leq \sum_{k=0}^{p} |g_k(x) - g_k(a)| + 2 \max_{y \in \mathbb{R}} \sum_{k=p+1}^{\infty} |g_k(x)| \\ &\leq \sum_{k=0}^{p} |g_k(x) - g_k(a)| + 2 \sum_{k=p+1}^{\infty} \frac{1}{2^{k+1}}. \end{aligned}$$

Die Stetigkeit von f folgt daraus, daß zu beliebig vorgegebenem $\varepsilon > 0$ p so gewählt werden kann, daß der zweite Summand kleiner als $\frac{\varepsilon}{2}$ ist und wegen der Stetigkeit der Funktionen g_k, $k = 1, \ldots, p$, der erste Summand beliebig klein wird, falls x nahe an a liegt. □

Satz 6.8
Es sei $f : D(f) \to \mathbb{R}$ in $x_0 \in \operatorname{int}(D(f))$ differenzierbar, und f besitze in x_0 ein lokales Extremum, d.h. es existiert eine Umgebung U von x_0, so daß

$$\forall x \in U \cap D(f) : f(x) \leq f(x_0) \quad \textit{bzw.} \quad f(x_0) \leq f(x).$$

Dann gilt $f'(x_0) = 0$.

Beweis:
Wir können U so klein wählen, daß $x_0 \in U \subseteq D(f)$ und o.B.d.A.

$$f(x) \leq f(x_0), \quad \forall x \in U.$$

Deshalb gilt:

$$\begin{aligned} \frac{f(x) - f(x_0)}{x - x_0} &\leq 0, \quad \text{falls} \quad x > x_0, \quad \text{und} \\ &\geq 0, \quad \text{falls} \quad x < x_0. \end{aligned}$$

Da aber f in x_0 differenzierbar ist, existiert der Grenzwert

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) \quad \text{und muß gleich 0 sein.} \qquad \square$$

Satz 6.9 *(Rolle)*
Es sei $f : [a, b] \to \mathbb{R}$ stetig, f sei auf $]a, b[$ differenzierbar und es gelte $f(a) = f(b)$. Dann existiert $\tilde{x} \in]a, b[$ mit $f'(\tilde{x}) = 0$.

Beweis:
Ist f eine konstante Funktion, so ist die Aussage trivial.
Es sei also f nicht konstant. Dann existiert ein $x_0 \in]a, b[$ mit o.B.d.A. $f(x_0) > f(a)$.
Nach Satz 5.13 existiert $\tilde{x} \in [a, b]$ mit $f(\tilde{x}) = \max\limits_{x \in [a,b]} f(x)$.
$\rightsquigarrow f(\tilde{x}) \geq f(x_0) > f(a) = f(b) \rightsquigarrow \tilde{x} \in]a, b[$ und $f(x) \leq f(\tilde{x}), \quad \forall x \in [a, b]$
$\rightsquigarrow f'(\tilde{x}) = 0$ nach Satz 6.8. $\square$

Satz 6.10 *(Mittelwertsatz)*
Es sei $f : [a, b] \to \mathbb{R}$ stetig und auf $]a, b[$ differenzierbar. Dann existiert ein $\tilde{x} \in]a, b[$ mit

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

Beweis:
Wir betrachten die Funktion $g(x) := f(x) - \frac{f(b)-f(a)}{b-a}(x - a), \forall x \in [a, b]$.
$\rightsquigarrow g$ ist stetig auf $[a, b]$, differenzierbar auf $]a, b[$ und es gilt

$$g(a) = f(a) = g(b).$$

$\rightsquigarrow$ aus dem Satz von Rolle folgt: $\exists \tilde{x} \in]a, b[: g'(\tilde{x}) = 0$.
Wegen $g'(\tilde{x}) = f'(\tilde{x}) - \frac{f(b)-f(a)}{b-a}$ ist damit alles gezeigt. $\square$

Folgerung 6.11
Es sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ stetig und auf $\overset{\circ}{I}$ differenzierbar.

a) *Gilt $L := \sup_{x \in I} |f'(x)| < +\infty$, so ist f Lipschitzstetig mit Konstante L.*

b) *f ist [streng] monoton wachsend (fallend), wenn*
$\forall x \in \overset{\circ}{I}$: $\quad f'(x) \geq 0 \quad (f'(x) \leq 0)\ [f'(x) > 0 \quad (f'(x) < 0)]$.

Beweis:

a) Es seien $x, y \in I$, $x < y$ bel. gewählt. Dann sind die Vor. von Satz 5.10 auf $[x, y]$ erfüllt, und es existiert $\tilde{x} \in]x, y[$, so daß

$$f(y) - f(x) = f'(\tilde{x})(y - x),$$

$\rightsquigarrow |f(x) - f(y)| \leq L|x - y|$.

b) Alles folgt wie in a) sofort aus der Gleichung

$$f(y) - f(x) = f'(\tilde{x})(y - x)$$

für bel. $x, y \in I$, $x < y$ ($\rightsquigarrow y - x > 0$)
z.B. $f'(\tilde{x}) \geq 0$, $\forall \tilde{x} \in \overset{\circ}{I} \rightsquigarrow f(y) - f(x) \geq 0$, d.h. $f(x) \leq f(y) \rightsquigarrow f$ ist monoton wachsend. In den anderen Fällen schlußfolgert man analog. □

Satz 6.12 *(verallgem. Mittelwertsatz)*
Seien $f, g : [a, b] \to \mathbb{R}$ stetig und auf $]a, b[$ differenzierbar. Dann existiert $\tilde{x} \in]a, b[$ mit

$$[f(b) - f(a)]g'(\tilde{x}) = [g(b) - g(a)]f'(\tilde{x}).$$

Gilt überdies $g'(x) \neq 0$, $\forall x \in]a, b[$, so gilt auch:

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\tilde{x})}{g'(\tilde{x})}.$$

Beweis:
Analog zum Beweis von Satz 6.10 wendet man Satz 6.9 an, jetzt aber auf die Funktion:

$$h(x) := [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x), \quad \forall x \in [a, b].$$

Es gilt $h(a) = f(a)g(b) - f(b)g(a) = h(b)$ und es folgt aus 6.9
$\rightsquigarrow \exists \tilde{x} \in]a, b[: h'(\tilde{x}) = 0 = [g(b) - g(a)]f'(\tilde{x}) - [f(b) - f(a)]g(\tilde{x})$
Ist noch $g'(x) \neq 0$, $\forall x \in]a, b[$, so muß $g(a) \neq g(b)$ gelten (6.9!), und die letztere Aussage folgt durch Division. □

In Kap. 6.2 werden wir die Mittelwertsätze auf den mehrdimensionalen Fall verallgemeinern und sehen, daß eine direkte Übertragung nicht möglich ist.

Satz 6.13 *(Regel von de l'Hospital)*
Die Funktionen f und g seien differenzierbar auf dem Intervall $]a,b[$ mit $a,b \in \mathbb{R} \cup \{-\infty,+\infty\}$, $a < b$ und es gelte $g'(x) \neq 0$, $\forall x \in]a,b[$. Ferner sei eine der folgenden Voraussetzungen erfüllt:
(V1) $\lim_{x\to a+} f(x) = \lim_{x\to a+} g(x) = 0$,
(V2) $\lim_{x\to a+} g(x) = +\infty$ *bzw.* $= -\infty$.
Dann gilt

$$\lim_{x\to a+} \frac{f(x)}{g(x)} = \lim_{x\to a+} \frac{f'(x)}{g'(x)},$$

falls der Limes auf der rechten Seite im eigentlichen oder uneigentlichen Sinn existiert. Ein analoges Resultat gilt für den Fall $x \to b-$.

Beweis:

(i) Es sei zunächst $y := \lim_{x\to a+} \frac{f'(x)}{g'(x)} \in [-\infty,+\infty[$ und wir wählen $\tilde{y} > y$ beliebig. Ferner sei $y_1 \in \mathbb{R}$ so gewählt, daß $y < y_1 < \tilde{y}$.
Nach Voraussetzung existiert dann ein $x_1 \in]a,b[$, so daß

$$\frac{f'(x)}{g'(x)} < y_1, \quad \forall x \in]a,x_1[.$$

Aus Satz 6.12 schlußfolgern wir nun, daß für alle $x,u \in]a,x_1[$, $u < x$ ein $\tilde{x}$ zwischen x und u existiert, so daß

$$\frac{f(x)-f(u)}{g(x)-g(u)} = \frac{f'(\tilde{x})}{g'(\tilde{x})} < y_1 < \tilde{y}$$

Im Fall (V1) folgt daraus durch Grenzübergang $u \to a+$:

$$\frac{f(x)}{g(x)} \leq y_1 < \tilde{y}, \quad \forall x \in]a,x_1[.$$

Im Fall (V2) gehen wir wie folgt vor: Zu fixiertem $u \in]a,x_1[$ existiert $x_2 \in]a,u[$, so daß für alle $x \in]a,x_2[$ gilt:

$$g(x) > \max\{0,g(u)\} \quad \text{bzw.} \quad g(x) < \min\{0,g(u)\},$$

$$\rightsquigarrow \frac{g(x)-g(u)}{g(x)} > 0, \quad \forall x \in]a,x_2[,$$

$$\rightsquigarrow \frac{f(x)-f(u)}{g(x)} < y_1 \frac{g(x)-g(u)}{g(x)}, \quad \forall x \in]a,x_2],$$

$$\rightsquigarrow \frac{f(x)}{g(x)} < y_1 - y_1\frac{g(u)}{g(x)} + \frac{f(u)}{g(x)}, \quad \forall x \in]a,x_2].$$

Strebt nun $x \to a$, so konvergiert wegen (V2) die rechte Seite der Ungleichung gegen y_1. Wegen $y_1 < \tilde{y}$ existiert deshalb $x_3 \in]a, x_2[$, so daß

$$\frac{f(x)}{g(x)} < \tilde{y}, \quad \forall x \in]a, x_3[.$$

Wir erhalten also insgesamt in beiden Fällen:
$\forall \tilde{y} \in \mathbb{R}$ mit $\tilde{y} > \lim_{x\to a+} \frac{f'(x)}{g'(x)} = y \exists \bar{x} \in]a, b[$:

$$\frac{f(x)}{g(x)} < \tilde{y}, \quad \forall x \in]a, \bar{x}[.$$

(ii) Es sei nun $y \in]-\infty, +\infty]$. Dann schlußfolgert man analog zu Teil (i) des Beweises, indem man jetzt alles nach unten abschätzt, daß zu jedem $\tilde{\tilde{y}} < y$ ein $\bar{\bar{x}} \in]a, b[$ existiert, so daß

$$\tilde{\tilde{y}} < \frac{f(x)}{g(x)}, \quad \forall x \in]a, \bar{x}[.$$

(iii) Man unterscheidet jetzt folgende Fälle:

1. Fall: $y = -\infty : \rightsquigarrow \forall \tilde{y} \in \mathbb{R} \exists \tilde{x} \in]a, b[: -\infty < \frac{f(x)}{g(x)} < \tilde{y}$
$\rightsquigarrow \exists \lim_{x\to a+} \frac{f(x)}{g(x)} = -\infty = y$

2. Fall: $y = +\infty :$ analog $\exists \lim_{x\to a+} \frac{f(x)}{g(x)} = +\infty = y$

3. Fall: y endlich: durch Kombination von i) und ii) folgt:
$\forall \tilde{y}, \tilde{\tilde{y}} \in \mathbb{R}, \tilde{y} > y > \tilde{\tilde{y}} \; \exists \tilde{x} \in]a, b[: \tilde{\tilde{y}} < \frac{f(x)}{g(x)} < \tilde{y} \, \forall x \in]a, \tilde{x}[$
$\rightsquigarrow \lim_{x\to a+} \frac{f(x)}{g(x)} = y = \lim_{x\to a+} \frac{f'(x)}{g'(x)}$. $\square$

Beispiel 6.14

a) $\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = 1$ *(Satz 5.13 mit $a = 0$, $b = +\infty$. bzw. $a = -\infty$, $b = 0$ und (V1); Anwendung von Satz 4.52)*

$\lim_{x\to+\infty} x \sin \frac{1}{x} = \lim_{x\to+\infty} \frac{\sin \frac{1}{x}}{\frac{1}{x}} = \lim_{x\to+\infty} \cos \frac{1}{x} = 1.$

b) $\lim_{x\to+\infty} \frac{\exp(x)}{x^k} = \lim_{x\to+\infty} \frac{\exp(x)}{kx^{k-1}} = \ldots = \lim_{x\to+\infty} \frac{\exp(x)}{k!} = +\infty$
(wiederholte Anwendung von Satz 5.13; vgl. auch 4.60 e)).

c) $\lim_{x\to+\infty} \frac{\ln x}{x} = \lim_{x\to+\infty} \frac{\frac{1}{x}}{1} = 0$ *(5.13 mit $a = 0$, $b = +\infty$),*
$\lim_{x\to 0+} x \ln x = \lim_{x\to 0+} \frac{\ln x}{\frac{1}{x}} = \lim_{x\to 0+} \frac{x^{-1}}{-x^{-2}} = \lim_{x\to 0+}(-x) = 0.$

d) $\lim_{x\to 0+} x^x = \lim_{x\to 0+} \exp(x \ln(x)) = \exp(\lim_{x\to 0+} x \ln(x)) = 1$
(wegen der Stetigkeit von $\exp$ *und c)).*

Definition 6.15
Es sei $D(f) \subseteq \mathbb{R}$ offen und $f : D(f) \to \mathbb{R}$ differenzierbar mit der Ableitung $g := f' : D(f) \to \mathbb{R}$.
Ist g in $x_0 \in D(f)$ differenzierbar, so heißt $g'(x_0)$ die <u>zweite Ableitung</u> von f in x_0.
Bezeichnung: $f''(x_0) := g'(x_0)$.
Induktiv definiert man $f^{(n)}(x_0) := (f^{(n-1)})'(x_0)$, die <u>n–te Ableitung</u> von f in x_0. Man sagt dann, daß f <u>n–mal differenzierbar</u> in x_0 ist. Schreibweise: $f^{(0)} := f$, $f^{(1)} = f'$ *etc.*

Wir kommen nun zu einer Verallgemeinerung des Mittelwertsatzes:

Satz 6.16 *(Taylor)*
Es sei $D(f)$ ein offenes Intervall, $f : D(f) \to \mathbb{R}$ sei $(n+1)$–mal differenzierbar für ein $n \in \mathbb{N}_0$. Dann existiert für alle $x, x_0 \in D(f)$ ein $\Theta \in]0,1[$, so daß

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(x_0 + \Theta(x-x_0))}{(n+1)!}(x-x_0)^{n+1}.$$

<u>**Beweis:**</u>
Wir betrachten die beiden folgenden Funktionen $g, h : D(f) \to \mathbb{R}$:

$$\begin{array}{ll} g(x) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k, & \quad (\forall x \in D(f); x_0 \in D(f) \text{ fest}) \\ h(x) := (x-x_0)^{n+1}, & \end{array}$$

g und h sind $(n+1)$–mal differenzierbar (g nach Vor.!) und es gilt
$\forall x \in D(f)$, $j = 0, \ldots, n+1$:

$$\begin{array}{l} g^{(j)}(x) = f^{(j)}(x) - \sum_{k=j}^{n} \frac{f^{(k)}(x_0)}{(k-j)!}(x-x_0)^{k-j}, \\ h^{(j)}(x) = (n+1)\cdots(n+2-j)(x-x_0)^{n+1-j}, \end{array}$$

Es sei nun $x \in D(f)$ ebenfalls fixiert, und wir betrachten das Intervall
$I := [\min\{x, x_0\}, \max\{x, x_0\}]$. Aus 6.12 folgt die Existenz von $\tilde{x}_1 \in \mathrm{int}(I)$ mit

$$\frac{g(x) - g(x_0)}{h(x) - h(x_0)} = \frac{g'(\tilde{x}_1)}{h'(\tilde{x}_1)}$$

Allgemein existiert nach Satz 6.12 für jedes $j = 1, \ldots, n$ ein $\tilde{x}_j \in I$, so daß

$$\frac{g^{(j+1)}(\tilde{x}_{j+1})}{h^{(j+1)}(\tilde{x}_{j+1})} = \frac{g^{(j)}(\tilde{x}_j) - g^{(j)}(x_0)}{h^{(j)}(\tilde{x}_j) - h^{(j)}(x_0)}, \ \forall j = 0, \ldots, n, \ \tilde{x}_0 := x.$$

(Dies erfolgt sukzessive und ist möglich da $h^{(j)}(x) \neq 0$, $\forall x \neq x_0$ und $\forall j = 1, \ldots, n+1$)
Insgesamt ergibt sich daraus wegen $g^{(j)}(x_0) = h^{(j)}(x_0) = 0$ für alle $j = 0, \ldots, n$:

$$\frac{g(x)}{h(x)} = \frac{g'(\tilde{x}_1)}{h'(\tilde{x}_1)} = \frac{g^{(n+1)}(\tilde{x}_{n+1})}{h^{(n+1)}(\tilde{x}_{n+1})} = \frac{f^{(n+1)}(\tilde{x}_{n+1})}{(n+1)!}.$$

$$\rightsquigarrow g(x) = \frac{f^{(n+1)}(\tilde{x}_{n+1})}{(n+1)!}(x-x_0)^{n+1},$$

und es gilt $\frac{\tilde{x}_{n+1} - x_0}{x - x_0} =: \Theta \in]0,1[$. Damit ist alles bewiesen. □

Bemerkung 6.17
Ein Spezialfall von Satz 6.16 für $n = 0$ ist der Mittelwertsatz 6.10 auf I.
Das Polynom

$$T_n(x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \quad (x \in \mathbb{R})$$

heißt das *n–te Taylor–Polynom* *von f an der Stelle x_0, und der Term*

$$R_n(x) := f(x) - T_n(x) \quad (x \in D(f))$$

heißt *Taylorsches Restglied.*
Satz 6.16 besagt dazu folgendes: $\forall x \in D(f)$ existiert $\Theta \in]0, 1[$, so daß

$$R_n(x) = \frac{f^{(n+1)}(x_0 + \Theta(x - x_0))}{(n + 1)!}(x - x_0)^{n+1}.$$

Dabei ist klar, daß Θ von x, x_0 und n abhängt! (vgl. Beweis von 6.16)

Definition 6.18
Es sei $f : D(f) \to \mathbb{R}$ in $x_0 \in \operatorname{int}(D(f))$ beliebig oft differenzierbar.
Dann heißt die Potenzreihe $\sum\limits_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k$ Taylorreihe von f in x_0.
Man sagt, f ist um x_0 in eine Taylorreihe entwickelbar, wenn ein $\rho > 0$ existiert, so daß $f(x) = \sum\limits_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k, \quad \forall x \in B(x_0, \rho)$.

Folgerung 6.19
Es seien $a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$, $a < b$, und $f :]a, b[\to \mathbb{R}$ sei beliebig oft differenzierbar. Für jedes $n \in \mathbb{N}$ existiere M_n, so daß

$$|f^{(n)}(x)| \leq M_n, \quad \forall x \in]a, b[.$$

Dann ist f um jedes $x_0 \in]a, b[$ in eine Taylorreihe entwickelbar (in $B(x_0, \rho) \subset]a, b[$), wenn

$$\lim_{n \to \infty} M_n \frac{\rho^n}{n!} = 0.$$

Falls $M_n = \alpha C^n$, $\forall n \in \mathbb{N}$, mit geeigneten $\alpha > 0$ und $C > 0$, so ist die letztere Bedingung erfüllt.

Beweis:
Aus der Ungleichung in Bem. 6.17 erhalten wir für alle $x \in B(x_0, \rho)$:

$$|f(x) - T_n(x)| = |f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k| \leq M_{n+1} \frac{\rho^{n+1}}{(n + 1)!} \underset{n \to \infty}{\longrightarrow} 0.$$

Mit $M_n = \alpha C^n$ gilt $\alpha \frac{(C\rho)^n}{n!} \underset{n \to \infty}{\longrightarrow} 0$ (vgl. Bsp. 3.10 f)). $\square$

Beispiel 6.20

a) *Taylorreihe von* $f(x) := \sin x$ $(x \in \mathbb{R})$ *um* $x_0 = 0$.
Die Voraussetzungen von Folg. 6.19 sind erfüllt mit $M_n := 1, \forall n \in I\!N$, $a := -\infty$, $b := +\infty$ *und es folgt für bel.* $\rho > 0$, *d.h.* $\forall x \in \mathbb{R}$:

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

(da $\sin^{(2k)} = (-1)^k \sin$ *und* $\sin^{(2k+1)} = (-1)^k \cos$*) Man erhält also gerade die Reihendarstellung aus Folg. 3.52.*
(Analoges gilt für $\exp$, $\cos$, $\sinh$, $\cosh$.*)*

b) $\ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}$, $\quad \forall x \in]-1, 1]$ *(Übung).*

c) *Taylorreihe von* $f(x) := \begin{cases} \exp(-\frac{1}{x^2}) & , \; x \neq 0, \\ 0 & , \; x = 0, \end{cases}$ *um* $x_0 = 0$.
f *ist auf* $\mathbb{R}$ *beliebig oft differenzierbar und* $\forall n \in I\!N$ *hat* $f^{(n)}(x)$ *für* $x \neq 0$ *die Darstellung:*

$$f^{(n)}(x) = p_n(\frac{1}{x}) \exp(-\frac{1}{x^2}), \quad f^{(n)}(0) = 0,$$

wobei p_n *ein geeignetes Polynom ist.*
Man kann nun zeigen, daß die Taylorreihe von f *in* $x_0 = 0$ *auf* $\mathbb{R}$ *konvergent ist (alle Partialsummen sind* 0 $\forall x \in \mathbb{R}$*), aber mit Ausnahme von* $f(0)$ *in* $x = 0$ *nicht mit* $f(x)$ *übereinstimmt (vgl. Übungen).*

6.2 Differentialrechnung im $\mathbb{R}^m$

Definition 6.21
Sei $f : D(f) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}^k$ *und* $x_0 \in \operatorname{int}(D(f))$.
f *heißt* <u>*differenzierbar*</u> *(oder: Fréchet–differenzierbar) in* x_0, *wenn es eine* $m \times k$*–Matrix* A *gibt, so daß*

$$f(x_0 + h) = f(x_0) + Ah + r(x_0, h)$$

für alle $h \in \mathbb{R}^k$ *mit* $x_0 + h \in D(f)$ *und*

$$\lim_{\|h\| \to 0} \frac{r(x_0, h)}{\|h\|} = 0.$$

$f'(x_0) := A$ *heißt* <u>*Ableitung*</u> *(oder: Fréchet–Ableitung) von* f *in* x_0;
$f'(x_0)h$ *heißt* <u>*Differential*</u> *von* f *in* x_0 *bez.* $h \in \mathbb{R}^k$.
Ist $D(f)$ *offen, so heißt* f *differenzierbar (auf* $D(f)$*), falls* f *in jedem* $x_0 \in D(f)$ *differenzierbar ist.*

Bemerkung 6.22
Differenzierbarkeit von f *in* x_0 *bedeutet, daß* f *in einer Umgebung von* x_0 *"lokal linearisierbar" ist. Für* $k = m = 1$ *entsteht der Ableitungsbegriff aus Kap. 6.1!*

Satz 6.23
Es sei $f : D(f) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}^k$, *differenzierbar in* $x_0 \in \operatorname{int}(D(f))$.

a) *Die Fréchet–Ableitung von* f *in* x_0 *ist eindeutig bestimmt.*

b) f *ist in* x_0 *stetig.*

c) *Ist* $g : D(g) \to \mathbb{R}^m$, $D(g) \subseteq \mathbb{R}^k$, *ebenfalls differenzierbar in* $x_0 \in \operatorname{int}(D(f) \cap D(g))$, *so gilt:* $(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0) \quad (\alpha, \beta \in \mathbb{R})$.

Beweis:

a) Seien A_1 und A_2 Fréchet–Ableitungen von f in x_0.

$$\rightsquigarrow \lim_{\|h\| \to 0} \frac{1}{\|h\|} \|f(x_0 + h) - f(x_0) - A_i h\| = 0, \quad i = 1, 2$$

Dann gilt für beliebiges $x \in \mathbb{R}^k$ und $t \in \mathbb{R}$, so daß $x_0 + tx \in D(f)$:

$$\begin{aligned}
\|A_1 x - A_2 x\| &= \|\frac{1}{t}(A_1(tx) - A_2(tx)\| = \frac{1}{|t|\|x\|}\|A_1(tx) - A_2(tx)\|\|x\| \\
&= \frac{1}{\|tx\|}\|A_1(tx) - A_2(tx)\|\|x\| \\
&= \frac{1}{\|tx\|}\|A_1(tx) - (f(x_0 + tx) - f(x_0)) + \\
&\qquad (f(x_0 + tx) - f(x_0)) - A_2(tx)\|\|x\| \\
&\le \frac{1}{\|tx\|}\Big[\|A_1(tx) - (f(x_0 + tx) - f(x_0))\| + \\
&\qquad \|(f(x_0 + tx) - f(x_0)) - A_2(tx)\|]\|x\|]\Big] \\
&= \Big[\frac{1}{\|tx\|}\|f(x_0 + tx) - f(x_0) - A_1(tx)\| + \\
&\qquad \frac{1}{\|tx\|}\|f(x_0 + tx) - f(x_0) - A_2(tx)\|\Big]\|x\| \underset{t \to 0}{\longrightarrow} 0.
\end{aligned}$$

$\rightsquigarrow A_1 x = A_2 x$, $\forall x \in \mathbb{R}^k$.
$\rightsquigarrow$ beide $m \times k$–Matrizen stimmen überein.

b) Für bel. $x \in D(f)$ mit $x \neq x_0$ gilt:
$\|f(x) - f(x_0)\| \le \|f(x) - f(x_0) - f'(x_0)(x - x_0)\| + \|f'(x_0)(x - x_0)\|$.
Nach Vor. existiert $\delta > 0$, so daß

$$\frac{1}{\|x - x_0\|}\|f(x) - f(x_0) - f'(x_0)(x - x_0)\| \le 1, \quad \forall x \in B(x_0, \delta).$$

$$\rightsquigarrow \|f(x) - f(x_0)\| \le \|x - x_0\| + \|f'(x_0)(x - x_0)\| \quad , \forall x \in B(x_0, \delta).$$

$$\underset{x \to x_0}{\longrightarrow} 0.$$

c) Seien $\alpha, \beta \in \mathbb{R}$ bel., $h \in \mathbb{R}^k : x_0 + h \in D(f), h \neq 0$
$\rightsquigarrow \frac{1}{\|h\|}\|\alpha f(x_0+h) + \beta g(x_0+h) - \alpha f(x_0) - \beta g(x_0) - (\alpha f'(x_0) + \beta g'(x_0))h\|$

$$\begin{aligned}
\leq \ & |\alpha| \underbrace{\frac{1}{\|h\|}\|f(x_0+h) - f(x_0) - f'(x_0)h\|}_{\xrightarrow[\|h\|\to 0]{} 0} \\
+ \ & |\beta| \underbrace{\frac{1}{\|h\|}\|g(x_0+h) - g(x_0) - g'(x_0)h\|}_{\xrightarrow[\|h\|\to 0]{} 0}
\end{aligned}$$

$\rightsquigarrow \alpha f + \beta g$ ist differenzierbar in x_0, und es gilt

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0). \qquad \square$$

Beispiel 6.24

a) *$f : \mathbb{R}^k \to \mathbb{R}^m$, $f(x) := Ax + b$, $\forall x \in \mathbb{R}^k$; A $m \times k$–Matrix, $b \in \mathbb{R}^m$.*
$\rightsquigarrow f(x_0+h) - f(x_0) = A(x_0+h) - Ax_0 = Ah + 0$, $\forall h \in \mathbb{R}^k$.
$\rightsquigarrow f$ ist differenzierbar in jedem $x_0 \in \mathbb{R}^k$, und $f'(x_0) = A$.

b) *$f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) := x_1^2 + x_2^2 + x_1x_2$, $\forall x = (x_1, x_2) \in \mathbb{R}^2$.*

$$\begin{aligned}
\rightsquigarrow f(x+h) - f(x) &= (x_1+h_1)^2 + (x_2+h_2)^2 + (x_1+h_1)(x_2+h_2) \\
&\quad -x_1^2 - x_2^2 - x_1x_2 \\
&= 2x_1h_1 + h_1^2 + 2x_2h_2 + h_2^2 + x_1h_2 + x_2h_1 + h_1h_2 \\
&= (2x_1+x_2)h_1 + (2x_2+x_1)h_2 + (h_1^2 + h_1h_2 + h_2^2) \\
&= \underbrace{(2x_1+x_2, 2x_2+x_1)}_{=:A} \underbrace{\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}}_{=h} + r(h),
\end{aligned}$$

und es gilt:

$$\frac{|r(h)|}{\|h\|} = \frac{|h_1^2 + h_1h_2 + h_2^2|}{(h_1^2 + h_2^2)^{1/2}} \leq |h_1| + |h_1| + |h_2| \xrightarrow[\|h\|\to 0]{} 0.$$

$\rightsquigarrow f$ ist differenzierbar und $f'(x) = (2x_1 + x_2, 2x_2 + x_1)$, $\forall x \in \mathbb{R}^2$.

Definition 6.25
Es sei $f : D(f) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}^k$, $x_0 \in \operatorname{int}(D(f))$.

a) *Existiert der Grenzwert*

$$f'(x_0; h) := \lim_{t\to 0} \frac{1}{t}(f(x_0+th) - f(x_0)) \quad (\textit{in } \mathbb{R}^m),$$

so heißt er Richtungsableitung von f in x_0 in Richtung $h \in \mathbb{R}^k$.

b) *Es bezeichne e_j, $j = 1, \ldots, k$, den j–ten kanonischen Einheitsvektor in $\mathbb{R}^k$. Existiert nun die Richtungsableitung von f in x_0 in Richtung e_j, so heißt*

$$\frac{\partial f}{\partial x_j}(x_0) := f'(x_0; e_j) \in \mathbb{R}^m \quad (j = 1, \ldots, k)$$

die partielle Ableitung von f nach x_j in $x_0 \in \mathbb{R}^k$.

c) *Falls alle partiellen Ableitungen $\frac{\partial f}{\partial x_j}(x_0) \quad (j = 1, \ldots, k)$ von f in x_0 existieren, so heißt die $m \times k$–Matrix*

$$J_f(x_0) := \left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_k}(x_0) \right)$$

Jacobi–Matrix (oder: Funktionalmatrix) von f in x_0.
Für $m = 1$ heißt $\nabla f(x_0) := J_f(x_0)$ Gradient von f in x_0.

Bemerkung 6.26
Im Unterschied zum Differential $f'(x_0)h$ in Def. 6.21 werden zur Definition der Richtungsableitung $f'(x_0; h)$ nur gegen x_0 konvergierende Folgen auf der Geraden $\{x_0 + th : t \in \mathbb{R}\}$ "zur Konkurrenz zugelassen". Für $h := e_j$ bedeutet dies, daß nur gegen x_0 konvergierende Folgen zugelassen werden, bei denen sich nur die j–te Komponente verändert. Alle anderen Komponenten bleiben fest!

$$\begin{aligned}
\frac{\partial f}{\partial x_j}(x_0) &= \lim_{t \to 0}(f(x_0 + te_j) - f(x_0)) \quad \left(x_0 = \begin{pmatrix} x_{01} \\ \vdots \\ x_{0k} \end{pmatrix} \in \operatorname{int}(D(f))\right) \\
&= \lim_{t \to 0} \frac{1}{t}(f(x_{01}, \ldots, x_{0j} + t, \ldots, x_{0k}) - f(x_{01}, \ldots, x_{0j}, \ldots, x_{0k})).
\end{aligned}$$

Deshalb also der Begriff "partielle Ableitung" von f nach der j–ten Komponente x_j!
Schreibt man die Funktion $f : D(f) \to \mathbb{R}^m$ in der Form
$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$, *$f_i : D(f) \to \mathbb{R}$, $(i = 1, \ldots, m)$ sind also die Komponenten von f,*
so hat die Jacobi–Matrix von f in x die Form:

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \ldots & \frac{\partial f_1}{\partial x_k}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \ldots & \frac{\partial f_m}{\partial x_k}(x) \end{pmatrix}.$$

Satz 6.27
Es sei $f : D(f) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}^k$, in $x_0 \in \operatorname{int}(D(f))$ differenzierbar. Dann existieren alle partiellen Ableitungen $\frac{\partial f}{\partial x_j}(x_0)$, $j = 1, \ldots, k$, und es gilt

$$f'(x_0) = J_f(x_0).$$

Allgemeiner existieren auch alle Richtungsableitungen $f'(x_0; h)$, $\forall h \in \mathbb{R}^k$, und es gilt $f'(x_0)h = f'(x_0; h)$, $\forall h \in \mathbb{R}^k$.

Beweis:
Es sei $h \in \mathbb{R}^k$ bel. gewählt. Dann gilt:

$$\begin{aligned}\|\frac{1}{t}(f(x_0+th)-f(x_0))-f'(x_0)h\| &= \frac{1}{\|th\|}\|f(x_0+th)-f(x_0)-f'(x_0)(th)\|\|h\| \\ &\underset{t\to 0}{\to} 0\end{aligned}$$

und $f'(x_0;h) = f'(x_0)h$.
Speziell existieren alle Richtungsableitungen $f'(x_0;e_j)$, $j=1,\ldots,k$, und es gilt:

$$\frac{\partial f}{\partial x_j}(x_0) = f'(x_0;e_j) = f'(x_0)e_j, \quad j=1,\ldots,k.$$

Es sei nun $x \in \mathbb{R}^k$ bel. $\rightsquigarrow x = \sum\limits_{j=1}^{k} x_j e_j$ (vgl. Bem. 1.38).

$$\rightsquigarrow f'(x_0)x = \sum_{j=1}^{k} x_j f'(x_0)e_j = \sum_{j=1}^{k} x_j \tfrac{\partial f}{\partial x_j}(x_0) = J_f(x_0)\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = J_f(x_0)x.$$

$\rightsquigarrow f'(x_0) = J_f(x_0).$ □

Die Fréchet–Ableitung einer Funktion erweist sich also gerade als ihre Jacobi–Matrix. Diese Erkenntnis stellt gleichzeitig einen neuen Einzigkeitsbeweis für die Ableitung dar.

Beispiel 6.28

a) $f(x_1,x_2) := x_1^2 + x_2^2 + x_1x_2$, $\forall (x_1,x_2) \in \mathbb{R}^2$ *(vgl. Bsp. 6.24 b)).*

$$\begin{aligned}\rightsquigarrow \frac{\partial f}{\partial x_1}(x_1,x_2) &= \lim_{t\to 0}\frac{1}{t}((x_1+t)^2 + x_2^2 + (x_1+t)x_2 - x_1^2 - x_2^2 - x_1x_2) \\ &= \lim_{t\to 0}\frac{1}{t}(2x_1t + t^2 + x_2t) = 2x_1 + x_2, \\ \frac{\partial f}{\partial x_2}(x_1,x_2) &= 2x_2 + x_1. \\ \rightsquigarrow J_f(x_1,x_2) &= (2x_1+x_2, 2x_2+x_1).\end{aligned}$$

b) *Dieses Beispiel zeigt, daß aus der Existenz aller partiellen Ableitungen i.a. <u>nicht</u> die Existenz aller Richtungsableitungen folgt!*

$$f:\mathbb{R}^2 \to \mathbb{R}, \quad f(x_1,x_2) := \begin{cases} x_1 & , \textit{ falls } x_2 = 0, \\ x_2 & , \textit{ falls } x_1 = 0, \\ 1 & \textit{ sonst.} \end{cases}$$

f besitzt in $(0,0)$ alle partiellen Ableitungen, denn:

$$\begin{aligned}\frac{\partial f}{\partial x_1}(0,0) &= \lim_{t\to 0}\frac{1}{t}(f(t,0)-f(0,0)) = \lim_{t\to 0}\frac{t}{t} = 1, \\ \frac{\partial f}{\partial x_2}(0,0) &= \lim_{t\to 0}\frac{t}{t} = 1.\end{aligned}$$

Aber: *f ist nicht stetig in $(0,0)$, und es existiert auch nicht die Richtungsableitung von f in $(0,0)$ in Richtung $h := (1,1)$.*

Beweis: *$\frac{1}{t}(f(t,t) - f(0,0)) = \frac{1}{t}$, d.h. der zugehörige Limes existiert nicht! Ebenso sieht man, daß f nicht stetig in $(0,0)$ ist, da $f(\frac{1}{n}, \frac{1}{n}) = 1 \not\to 0 = f(0,0)$.* $\square$

c) *Das folgende Beispiel zeigt, daß selbst aus der Existenz aller Richtungsableitungen in einem Punkt nicht die Stetigkeit in diesem Punkt, also erst recht nicht die Fréchet–Differenzierbarkeit, folgt!*

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x_1, x_2) := \begin{cases} 0 & x_1 = x_2 = 0 \\ \frac{x_1^4 x_2^2}{x_1^8 + x_2^4} & \text{sonst} \end{cases}.$$

Es sei $h \in \mathbb{R}^2$, $h \neq (0,0)$, bel. gewählt. Dann gilt:

$$\lim_{t \to 0} \frac{1}{t}(f(0 + th_1, 0 + th_2) - f(0,0)) = \lim_{t \to 0} \frac{1}{t} \frac{t^6 h_1^4 h_2^2}{t^8 h_1^8 + t^4 h_2^4}$$
$$= \lim_{t \to 0} t \frac{h_1^4 h_2^2}{t^4 h_1^8 + h_2^4} = 0, \quad \text{d.h. } \exists f'((0,0); h) = 0,$$

also existieren alle Richtungsableitungen von f in $(0,0)$.
Aber: f ist nicht stetig in $(0,0)$.

Bew.: *Wir betrachten die Folge $((\frac{1}{n}, \frac{1}{n^2}))_{n \in \mathbb{N}}$. Es gilt: $f(\frac{1}{n}, \frac{1}{n^2}) = \frac{1}{2} \not\to 0 = f(0,0)$.* $\square$

(Ursache: Bei Richtungsableitungen werden nur Folgen auf Geraden zugelassen; bei Stetigkeiten alle!)

Problem: Unter welcher zusätzlichen Voraussetzung folgt nun aus der Existenz der partiellen Ableitungen die Differenzierbarkeit?

Definition 6.29
Es sei $f : D(f) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}^k$, $x_0 \in \operatorname{int}(D(f))$.
f heißt stetig differenzierbar in x_0, wenn ein $\delta > 0$ existiert, so daß alle partiellen Ableitungen von f in jedem $x \in B(x_0, \delta) \subseteq D(f)$ existieren und in x_0 stetig sind.

Satz 6.30
Sei $f : D(f) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}^k$, in $x_0 \in \operatorname{int}(D(f))$ stetig differenzierbar. Dann ist f in x_0 Fréchet–differenzierbar.

Beweis:
Es genügt, die Fréchet–Differenzierbarkeit von f in x_0 zu beweisen. Der Rest folgt aus Satz 6.27. Nach Def. 6.21 genügt es, den Fall $m = 1$ zu betrachten (sonst argumentiert man komponentenweise). Wir führen den Beweis außerdem nur für $k = 2$, da er für $k > 2$ bis auf mehr Schreibarbeit völlig analog verläuft.

Es sei nun $\varepsilon > 0$ beliebig gewählt. Nach Def. 6.21 genügt es nun zu zeigen, daß ein $\delta = \delta(\varepsilon) > 0$ existiert, so daß

$$\frac{|f(x) - f(x_0) - J_f(x_0)(x - x_0)|}{\|x - x_0\|} < \varepsilon, \quad \forall x \in B(x_0, \delta).$$

(Hierbei ist $\|\cdot\|$ die Euklidische Norm im $\mathbb{R}^2$.)
Nach Voraussetzung existiert zunächst ein $\delta > 0$, so daß $B(x_0, \delta) \subseteq D(f)$, in $B(x_0, \delta)$ die partiellen Ableitungen von f existieren und daß

$$\left|\frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(x_0)\right| < \frac{\varepsilon}{2}, \quad \forall x \in B(x_0, \delta), \ \forall j = 1, 2.$$

Es sei nun $x_0 = (x_{o1}, x_{o2})$, und wir betrachten die Funktion
$f(\cdot, x_{o2}) :]x_{o1} - \delta, x_{o1} + \delta[\to \mathbb{R}$.
Diese Funktion ist auf dem angegebenen Intervall differenzierbar mit der Ableitung $\frac{\partial f}{\partial x_1}(\cdot, x_{o2})$.
Ist $x = (x_1, x_2) \in B(x_0, \delta)$ bel. gewählt, so folgt aus dem Mittelwertsatz 6.10, daß ein $\Theta_1 \in]0, 1[$ existiert, so daß

$$\frac{f(x_1, x_{o2}) - f(x_{o2}, x_{o2})}{x_1 - x_{o1}} = \frac{\partial f}{\partial x_1}(x_{o1} + \Theta_1(x_1 - x_{o1}), x_{o2}).$$

Analog erhält man: $\exists \Theta_2 \in]0, 1[$ mit

$$\frac{f(x_1, x_2) - f(x_1, x_{o2})}{x_2 - x_{o2}} = \frac{\partial f}{\partial x_2}(x_1, x_{o2} + \Theta_2(x_2 - x_{o2})).$$

Daraus erhält man insgesamt:

$$\begin{aligned}
f(x) - f(x_0) &= f(x_1, x_2) - f(x_{o1}, x_{o2}) \\
&= f(x_1, x_2) - f(x_1, x_{o2}) + f(x_1, x_{o2}) - f(x_{o1}, x_{o2}) \\
&= \frac{\partial f}{\partial x_2}(x_1, x_{o2} + \Theta_2(x_2 - x_{o2}))(x_2 - x_{o2}) + \\
&\quad\ \frac{\partial f}{\partial x_1}(x_{o1} + \Theta_1(x_1 - x_{o1}), x_{o2})(x_1 - x_{o1}), \\
\rightsquigarrow f(x) - f(x_0) &= \left(\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0)\right) \begin{pmatrix} x_1 - x_{o1} \\ x_2 - x_{o2} \end{pmatrix} + r_f(x, x_0) \\
&\quad\ (\forall x \in B(x_0, \delta)),
\end{aligned}$$

wobei

$$\begin{aligned}
r_f(x, x_0) := {} & [\frac{\partial f}{\partial x_1}(x_{o1} + \Theta_1(x_1 - x_{o1}), x_{o2}) - \frac{\partial f}{\partial x_1}(x_{o1}, x_{o2})](x_1 - x_{o1}) + \\
& [\frac{\partial f}{\partial x_2}(x_1, x_{o2} + \Theta_2(x_2 - x_{o2})) - \frac{\partial f}{\partial x_2}(x_{o1}, x_{o2})](x_2 - x_{o2}).
\end{aligned}$$

$\rightsquigarrow |r_f(x, x_0)| \le \frac{\varepsilon}{2}|x_1 - x_{o1}| + \frac{\varepsilon}{2}|x_2 - x_{o2}| \le \varepsilon\|x - x_0\|, \ \forall x \in B(x_0, \delta)$.
(wegen $\left.\begin{matrix} (x_{o1} + \Theta_1(x_1 - x_{o1}), x_{o2}) \in B(x_0, \delta) \\ (x_1, x_{o2} + \Theta_2(x_2 - x_{o2})) \in B(x_0, \delta) \end{matrix}\right\}. \ \forall x \in B(x_0, \delta))$
Damit ist alles bewiesen! □

Wir kommen nun zur Kettenregel als der fundamentalen Differentiationsregel.

Satz 6.31 *(Kettenregel)*
Es seien $f : D(f) \to \mathbb{R}^l$, $g : D(g) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}^k$, $D(g) \subseteq \mathbb{R}^l$, und wir betrachten die Abbildung

$$F := g \circ f : D(g \circ f) \to \mathbb{R}^m, \quad D(g \circ f) \subseteq D(f) \subseteq \mathbb{R}^k.$$

Es seien $x_0 \in \operatorname{int}(D(g \circ f))$ und $f(x_0) \in \operatorname{int}(D(g))$, f sei in x_0 und g sei in $f(x_0)$ Fréchet–differenzierbar.
Dann ist $F = g \circ f$ in x_0 Fréchet–differenzierbar, und es gilt

$$F'(x_0) = g'(f(x_0))f'(x_0).$$

Beweis:
Nach Vor. und Def. 5.22 bzw. Bem. 5.23 gilt:

$$f(x_0 + h) = f(x_0) + f'(x_0)h + r_f(x_0, h), \quad \forall h \in \mathbb{R}^k \text{ mit } x_0 + h \in D(f),$$

wobei $\lim\limits_{\|h\| \to 0} \dfrac{\|r_f(x_0, h)\|}{\|h\|} = 0$, und

$g(f(x_0) + y) = g(f(x_0)) + g'(f(x_0))y + r_g(f(x_0), y), \quad \forall y \in \mathbb{R}^l$ mit $f(x_0) + y \in D(g)$, wobei $\lim\limits_{\|y\| \to 0} \frac{\|r_g(f(x_0), y)\|}{\|y\|} = 0.$
Daraus folgt für alle $h \in \mathbb{R}^k$ mit $x_0 + h \in D(F) = D(g \circ f)$:

$$\begin{aligned} F(x_0 + h) - F(x_0) &= g(f(x_0) + (f(x_0 + h) - f(x_0))) - g(f(x_0)) \\ &= g'(f(x_0))(f(x_0 + h) - f(x_0)) + r_g(f(x_0), f(x_0 + h) - f(x_0)) \\ &= g'(f(x_0))f'(x_0)h + r_F(x_0, h), \end{aligned}$$

wobei $r_F(x_0, h) := g'(f(x_0))r_f(x_0, h) + r_g(f(x_0), f(x_0 + h) - f(x_0))$.
Wir zeigen: $\lim\limits_{\|h\| \to 0} \frac{\|r_F(x_0,h)\|}{\|h\|} = 0.$
Da $\frac{\|g'(f(x_0))r_g(x_0;h)\|}{\|h\|} \leq \frac{\|g'(f(x_0))\| \|r_g(x_0;h)\|}{\|h\|} \underset{\|h\| \to \infty}{\longrightarrow} 0$,
genügt es, zu zeigen: $\lim\limits_{\|h\| \to 0} \frac{\|r_g(f(x_0), f(x_0 + h) - f(x_0))\|}{\|h\|} = 0.$
Es sei $\varepsilon > 0$ bel. gewählt. Dann existiert zunächst ein $\delta_1(\varepsilon) > 0$ so, daß

$$\frac{\|r_g(f(x_0), y)\|}{\|y\|} < \frac{\varepsilon}{\|f'(x_0)\| + 1}, \quad \forall y \in \mathbb{R}^l \text{ mit } 0 \neq \|y\| < \delta_1(\varepsilon).$$

(Hierbei ist $\|f'(x_0)\|$ die Norm der (l, k)–Matrix $f'(x_0)$ d.h. $\|f'(x_0)\| = \max_{\|x\|=1} \|f'(x_0)x\|$.)
$\rightsquigarrow \exists \delta(\varepsilon) > 0$ so, daß für alle $h \in \mathbb{R}^k$ mit $0 \neq \|h\| < \delta(\varepsilon)$ gilt:

$$\|f(x_0 + h) - f(x_0)\| < \delta_1(\varepsilon) \quad \text{(vgl. 5.24 b)) und}$$
$$\frac{\|r_f(x_0, h)\|}{\|h\|} \leq 1.$$

Dann gilt für alle $h \in \mathbb{R}^k$ mit $0 \neq \|h\| < \delta(\varepsilon)$:

$$\begin{aligned} \frac{\|r_g(f(x_0), f(x_0 + h) - f(x_0))\|}{\|h\|} &= \frac{\|r_g(f(x_0), f(x_0 + h) - f(x_0))\|}{\|f(x_0 + h) - f(x_0)\|} \frac{\|f'(x_0)h + r_f(x_0, h)\|}{\|h\|} \\ &< \frac{\varepsilon}{\|f'(x_0)\| + 1}(\|f'(x_0)\frac{h}{\|h\|}\| + \frac{\|r_f(x_0, h)\|}{\|h\|}) \\ &\leq \frac{\varepsilon}{\|f'(x_0)\| + 1}(\|f'(x_0)\| + 1) = \varepsilon. \qquad \square \end{aligned}$$

Folgerung 6.32
Es seien alle Voraussetzungen von Satz 6.31 erfüllt mit der Ausnahme, daß f in x_0 lediglich eine Richtungsableitung in Richtung $h \in \mathbb{R}^k$ besitze.
Dann besitzt auch $F := g \circ f$ in x_0 eine Richtungsableitung in Richtung $h \in \mathbb{R}^k$ und es gilt:

$$F'(x_0; h) = g'(f(x_0))f'(x_0; h).$$

<u>Beweis:</u>
Man wiederholt den prinzipiellen Ablauf des Beweises von 6.31 und ersieht aus der Gleichung

$$\begin{aligned}\frac{1}{t}(F(x_0 + th) - F(x_0)) &= g'(f(x_0))\frac{1}{t}(f(x_0 + th) - f(x_0)) \\ &\quad + \frac{1}{t} r_g(f(x_0), f(x_0 + th) - f(x_0))\end{aligned}$$

und indem man analog zu 6.31 beweist $\lim\limits_{t \to 0} \frac{1}{t} r_g(f(x_0), f(x_0 + th) - f(x_0)) = 0$, daß die behauptete Formel für die Richtungsableitungen gültig ist. □

Bemerkung 6.33
Unter den Voraussetzungen von Satz 6.31/Folg. 6.32 ergeben sich aus Satz 6.27 folgende Formeln für die Jacobi–Matrizen bzw. Richtungsableitungen:

$$\begin{aligned} J_F(x_0) &= J_g(f(x_0))J_f(x_0) \quad \textit{(nach 6.31)}, \\ F'(x_0; h) &= J_g(f(x_0))f'(x_0; h) \quad \textit{(nach 6.32)}. \end{aligned}$$

Für den Spezialfall $h := e_j$ $(j = 1, \ldots, k)$ erhalten wir daraus:

$$\frac{\partial F}{\partial x_j}(x_0) = J_g(f(x_0))\frac{\partial f}{\partial x_j}(x_0) \quad (j = 1, \ldots, k).$$

Haben ferner F, f bzw. g die folgende Gestalt

$$F(x) = \begin{pmatrix} F_1(x) \\ \vdots \\ F_m(x) \end{pmatrix}, \quad f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}, \quad g(y) = \begin{pmatrix} g_1(y) \\ \vdots \\ g_m(y) \end{pmatrix},$$

$(\forall x \in D(F))$ $\qquad$ $(\forall x \in D(f))$ $\qquad$ $(\forall y \in D(g))$

so ergeben sich folgende <u>Kettenregeln</u> für partielle Ableitungen:

$$\frac{\partial F_r}{\partial x_j}(x_0) = \sum_{i=1}^{l} \frac{\partial g_r}{\partial y_i}(f(x_0))\frac{\partial f_i}{\partial x_j}(x_0) \quad (r = 1, \ldots, m; j = 1, \ldots, k).$$

<u>*Spezialfälle:*</u>

$$\begin{aligned} m = 1 : \tfrac{\partial F}{\partial x_i}(x_0) &= \sum_{i=1}^{l} \frac{\partial g}{\partial y_i}(f(x_0))\frac{\partial f_i}{\partial x_j}(x_0) \quad (j = 1, \ldots, k), \\ l = 1 : \tfrac{\partial F_r}{\partial x_j}(x_0) &= \frac{\partial g_r}{\partial y}(f(x_0))\frac{\partial f}{\partial x_j}(x_0) \quad \begin{matrix}(r = 1, \ldots, m; \\ j = 1, \ldots, k),\end{matrix} \\ k = l = m = 1 : F'(x_0) &= g'(f(x_0))f'(x_0). \end{aligned}$$

Ausdrücklich vermerkt werden muß aber, daß die angegebenen Beziehungen für die partiellen Ableitungen von $F := g \circ f$ nur unter den Voraussetzungen von 6.31 bzw. 6.32 abgeleitet werden können! Insbesondere ist die Fréchet–Differenzierbarkeit von g in $f(x_0)$ unverzichtbar! (hinreichende Bedingungen dafür siehe Satz 6.30)

Beispiel 6.34 *a) Wir betrachten die folgenden beiden Funktionen:*

$$\begin{aligned} f : \mathbb{R}^2 \to \mathbb{R}, \quad & f(x_1, x_2) := x_1^2 + x_2^2 + x_1 x_2, \quad && \forall (x_1, x_2) \in \mathbb{R}^2, \\ g : \mathbb{R} \to \mathbb{R}, \quad & g(y) := \exp(y), && \forall y \in \mathbb{R}. \end{aligned}$$

Nach Bsp. 6.24 b) bzw. Satz 6.30 sind beide Funktionen Fréchet–differenzierbar auf ihren Definitionsbereichen.
Nach Satz 6.31 ist deshalb $g \circ f$ Fréchet–differenzierbar auf $\mathbb{R}^2$ und es gilt für die partiellen Ableitungen nach Bem. 6.33:

$$\begin{aligned} \frac{\partial (g \circ f)}{\partial x_1}(x_1, x_2) &= g'(f(x_1, x_2)) \frac{\partial f}{\partial x_1}(x_1, x_2) = (2x_1 + x_2) \exp(x_1^2 + x_2^2 + x_1 x_2), \\ \frac{\partial (g \circ f)}{\partial x_2}(x_1, x_2) &= (2x_2 + x_1) \exp(x_1^2 + x_2^2 + x_1 x_2), \end{aligned}$$

und damit $(g \circ f)' = \exp(x_1^2 + x_2^2 + x_1 x_2)(2x_1 + x_2, x_1 + 2x_2)$.

b) Wir zeigen, wie die Differentiationsregel für Produkte mit Hilfe der Kettenregel hergeleitet werden kann. Dazu definieren wir

$$\begin{aligned} & G : D(f) \cap D(g) \to \mathbb{R}^2, G(x) := \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}, \\ & H : \mathbb{R}^2 \to \mathbb{R}, H(y_1, y_2) := y_1 y_2, \forall (y_1, y_2) \in \mathbb{R}^2. \end{aligned}$$

und erhalten $f \cdot g = H \circ G$. G ist Fréchet–differenzierbar in x_0, da dies für f und g erfüllt ist. H ist Fréchet–differenzierbar auf $\mathbb{R}^2$, da H stetig differenzierbar ist und nach Satz 6.30 ! Ferner gilt:

$$H'(y_1, y_2) = \left(\frac{\partial H}{\partial y_1}(y_1, y_2), \frac{\partial H}{\partial y_2}(y_1, y_2) \right) = (y_2, y_1).$$

$$\begin{aligned} \textit{Satz 6.36} \rightsquigarrow (f \cdot g)'(x_0) &= (H \circ G)'(x_0) = H'(G(x_0)) G'(x_0) \\ &= (g(x_0), f(x_0)) \begin{pmatrix} f'(x_0) \\ g'(x_0) \end{pmatrix} \\ &= g(x_0) f'(x_0) + f(x_0) g'(x_0). \end{aligned}$$

Beispiel 6.35
Für $m > 1$ und $k > 1$ gilt der Mittelwertsatz in der Form 6.10 i.a. nicht mehr!
Sei $k := m := 2$, $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x_1, x_2) := (x_1^3, x_2^2)$, $\forall (x_1, x_2) \in \mathbb{R}^2$.
$\rightsquigarrow$ f ist stetig differenzierbar auf $\mathbb{R}^2$;
$\rightsquigarrow$ f ist Fréchet–differenzierbar auf $\mathbb{R}^2$ nach Satz 6.31, und es gilt:

$$f'(x_1, x_2) = \begin{pmatrix} 3x_1^2 & 0 \\ 0 & 2x_2 \end{pmatrix}, \quad \forall (x_1, x_2) \in \mathbb{R}^2.$$

Annahme: $\exists \Theta \in]0,1[$: $f(1,1) - f(0,0) = f'((0,0)) + \Theta(1,1)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\rightsquigarrow f(1,1) = f'(\Theta,\Theta)\begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3\Theta^2 \\ 2\Theta \end{pmatrix}$

$\rightsquigarrow$ *Widerspruch!*

Problem: *Existieren (wenigstens) Abschätzungen für $\|f(x) - f(x_0)\|$ mit Hilfe der Fréchet–Ableitung f'?*

Definition 6.36

Für beliebige $x, y \in \mathbb{R}^k$ bezeichne $[x,y] := \{x + t(y-x) : t \in [0,1]\}$ ein k–dimensionales Intervall. $C \subseteq \mathbb{R}^k$ heißt konvex, falls für alle $x, y \in C$ gilt: $[x,y] \subseteq C$.

Satz 6.37 *(verallgemeinerter Mittelwertsatz)*

Es sei $D(f) \subseteq \mathbb{R}^k$ offen und konvex, $f : D(f) \to \mathbb{R}^m$ sei Fréchet–differenzierbar. Dann gilt für alle $x, x_0 \in D(f)$:

$$\|f(x) - f(x_0)\| \leq \sup_{f \in [0,1]} \|f'(x_0 + t(x - x_0))\| \|x - x_0\|.$$

wobei $\|A\| = \max\limits_{\|x\|=1} \|Ax\|$ für alle $m \times k$-Matrizen A.

Beweisidee: Für die Funktion $t \mapsto g_i(t) := f_i(x_0 + t(x - x_0))$ ist der Mittelwertsatz 6.10 anwendbar, d.h. für jedes $i = 1, \ldots, n$ existiert ein $\theta_i \in]0,1[$, so daß

$$|g_i(1) - g_i(0)| = |f_i(x) - f_i(x_0)| = |g_i'(\theta_i)| \leq \sup_{t \in [0,1]} |g_i'(t)| = \sup_{t \in [0,1]} |f_i'(x_0 + t(x - x_0))(x - x_0)|.$$

Daraus folgt die gewünschte Aussage für die Norm $\|\cdot\|_\infty$ in $\mathbb{R}^k$ und $\mathbb{R}^m$. □

Wir kommen schließlich zu partiellen Ableitungen höherer Ordnung.

Definition 6.38

Es sei $f : D(f) \to \mathbb{R}^m$, und $D(f) \subseteq \mathbb{R}^k$ sei offen. f besitze auf $D(f)$ die partielle Ableitung $\frac{\partial f}{\partial x_j} : D(f) \to \mathbb{R}^m$ $(j \in \{1, \ldots, k\})$.

Existiert nun die partielle Ableitung $\frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j})(x_0)$, für $x_0 \in D(f)$, so heißt $\frac{\partial^2 f}{\partial x_i \partial x_j})(x_0) := \frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j}(x_0)$ partielle Ableitung zweiter Ordnung von f in x_0 nach x_j und x_i $(i, j \in \{1, \ldots, k\})$. Im Fall $i = j$ verwenden wir die Schreibweise: $\frac{\partial^2 f}{\partial x_j^2}(x_0)$.

Ist $\alpha := (\alpha_1, \ldots, a_k)$ ein sog. Multi–Index und ist $|\alpha| := \sum\limits_{i=1}^{k} \alpha_i$, $\alpha_i \in \mathbb{N}_o$, $\forall i = 1, \ldots, k$, so definiert man auf die obige Weise sukzessive eine partielle Ableitung $|\alpha|$–ter Ordnung

$$\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_k^{\alpha_k}}(x_0) \text{ von } f \text{ in } x_0.$$

(Dabei bedeutet $\alpha_i = 0$, daß nach der Variablen x_i nicht abgeleitet wird.)

Für eine offene Menge $G \subseteq \mathbb{R}^k$ bezeichnet $C_m^{|\alpha|}(G)$ die Menge aller Funktionen (mit Werten in $\mathbb{R}^m$), für die alle partiellen Ableitungen bis zur $|\alpha|$-ten Ordnung existieren und stetig sind.

Beispiel 6.39

a) $f(x_1,x_2) := (x_1^2+x_2^2)\exp(x_1x_2), \quad \forall (x_1,x_2)\in\mathbb{R}^2.$

$$\begin{aligned}\frac{\partial f}{\partial x_1}(x_1,x_2) &= 2x_1\exp(x_1x_2)+(x_1^2+x_2^2)x_2\exp(x_1x_2),\\ &= (x_1^2x_2+2x_1+x_2^3)\exp(x_1x_2),\end{aligned}$$

$$\frac{\partial f}{\partial x_2}(x_1,x_2) = (x_1^3+x_1x_2^2+2x_2)\exp(x_1x_2),$$

$$\begin{aligned}\frac{\partial^2 f}{\partial x_2\partial x_1}(x_1,x_2) &= (x_1^2+3x_2^2)\exp(x_1x_2)+(x_1^2x_2+2x_1+x_2^3)x_1\exp(x_1x_2)\\ &= (3x_1^2+3x_2^2+x_1^3x_2+x_1x_2^3)\exp(x_1x_2),\\ \frac{\partial^2 f}{\partial x_1\partial x_2}(x_1,x_2) &= (3x_1^2+x_2^2)\exp(x_1x_2)+(x_1^3+x_1x_2^2+2x_2)x_2\exp(x_1x_2)\\ &= \frac{\partial^2 f}{\partial x_2\partial x_1}(x_1,x_2),\end{aligned}$$

d.h. bei dieser Funktion kann die Reihenfolge der Ableitungen vertauscht werden.

b) $f(x_1,x_2) := \begin{cases} x_1x_2\frac{x_1^2-x_2^2}{x_1^2+x_2^2}, & (x_1,x_2)\neq(0,0),\\ 0 & \textit{sonst.}\end{cases}$

$$\begin{aligned}\frac{\partial f}{\partial x_1}(x_1,x_2) &= x_2\frac{x_1^2-x_2^2}{x_1^2+x_2^2}+x_1x_2\frac{2x_1(x_1^2+x_2^2)-(x_1^2-x_2^2)2x_1}{(x_1^2+x_2^2)^2}\\ &= \frac{x_2(x_1^4-x_2^4)+2x_1^2x_2^3}{(x_1^2+x_2^2)^2}.\end{aligned}$$

$$\rightsquigarrow \frac{\partial f}{\partial x_1}(0,x_2) = -x_2 \textit{ und analog:} \frac{\partial f}{\partial x_2}(x_1,0)=x_1.$$

$$\rightsquigarrow \frac{\partial^2 f}{\partial x_2\partial x_1}(0,0) = -1 \textit{ und } \frac{\partial^2 f}{\partial x_1\partial x_2}(0,0)=1.$$

In diesem Fall läßt sich also die Reihenfolge der Ableitungen nicht vertauschen !

Wann sind die partiellen Ableitungen einer Funktion unabhängig von der Reihenfolge der Differentiationen?

Satz 6.40
Es sei $f : D(f) \to \mathbb{R}^m$, $D(f) \subseteq \mathbb{R}^k$, $x_0 \in \operatorname{int}(D(f))$. Existieren die partiellen Ableitungen

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \quad \textit{und} \quad \frac{\partial^2 f}{\partial x_j \partial x_i}$$

für gewisse $i, j \in \{1, \ldots, k\}$, $i \neq j$, auf einer Umgebung U von x_0 und sind in x_0 stetig, so sind sie in x_0 gleich.

(ohne Beweis)

Abschließend behandeln wir nun eine Verallgemeinerung des Satzes von Taylor (6.16) auf den mehrdimensionalen Fall. Jedoch beschränken wir uns der Einfachheit halber auf eine Entwicklung nur bis einschließlich der zweiten partiellen Ableitungen.
Für eine zweimal stetig differenzierbare Funktion $f : G \mapsto \mathbb{R}$, d.h. $f \in C^2(G)$, mit $G \subseteq \mathbb{R}^k$ offen, wird für $x_0 \in G$ definiert:

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_k}(x_0) \right) \text{(Gradient von } f \text{ in } x_0)$$

und die sog. Hesse–Matrix $\nabla^2 f(x_0)$ von f in x_0

$$\nabla^2 f(x_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_k \partial x_1}(x_0) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_k}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_k^2}(x_0) \end{pmatrix}$$

(die also nach 6.40 eine symmetrische Matrix in $\mathbb{R}^{k \times k}$ ist)

Satz 6.41 *(Taylor)*
Sei $G \subseteq \mathbb{R}^k$ offen und konvex, und es gelte $f \in C^2(G)$.
Dann gilt für alle $x_0 \in G$ und $h \in \mathbb{R}^k$ derart, daß $x_0 + h \in G$:

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle \nabla^2 f(x_0) h, h \rangle + \|h\|^2 \rho(h)$$

wobei $\lim_{\|h\| \to 0} \rho(h) = 0.$
Dabei bezeichnet $\langle \cdot, \cdot \rangle$ das Skalarprodukt auf $\mathbb{R}^k$ (vgl. 1.37), und $\| \cdot \|$ die Euklidische Norm.

(ohne Beweis)

Mit dieser mehrdimensionalen Variante der Taylorformel läßt sich nun die folgende notwendige und hinreichende Bedingung für das Vorliegen eines lokalen Minimums einer Funktion f in einem Punkt x_0 herleiten.

Satz 6.42 *(notwendige und hinreichende Optimalitätsbedingung)*
(a) Seien $f \in C^1(\mathbb{R}^m)$, $C \subseteq \mathbb{R}^m$ abgeschlossen und konvex, und $x_0 \in C$. Besitzt f in x_0 ein lokales Minimum auf C, d.h. existiert ein $\varepsilon > 0$, so daß $f(x_0) \leq f(x)$ für jedes $x \in C \cap B(x_0, \varepsilon)$, so gilt

$$\langle \nabla f(x_0), x - x_0 \rangle \geq 0, \, \forall x \in C.$$

Gilt $x_0 \in \operatorname{int}(C)$, so folgt sogar $\nabla f(x_0) = 0$.
(b) Sei $f \in C^2(\mathbb{R}^m)$ und es gelte für $x_0 \in \mathbb{R}^m$, daß $\nabla f(x_0) = 0$ und $\langle \nabla^2 f(x_0)h, h \rangle > 0$, $\forall h \in \mathbb{R}^m \backslash \{0\}$. Dann ist x_0 lokales Minimum von f auf $\mathbb{R}^m$.

Die Bedingung an die Hesse-Matrix $\nabla^2 f(x_0)$ in Satz 6.42 ist dabei gleichbedeutend damit, daß ihre sämtlichen Eigenwerte positiv sind.

7 Integralrechnung

Für den Integralbegriff gibt es zwei wesentliche Motivationen:

(1) Berechnung von Flächen- bzw. Rauminhalten von Körpern, und

(2) die "Umkehrung" der Differentiation, d.h. die Rekonstruktion einer Funktion aus ihrer Ableitung.

Beide Motivationen hängen eng zusammen. Wir beginnen mit (1) und der Definition des sog. Riemann-Integrals. Später zeigen wir, daß es enge Zusammenhänge zwischen dem Riemann-Integral und der Umkehrung der Differentiation gibt. Anschließend gehen wir noch auf eine Erweiterung der bis dahin dargelegten Integralrechnung ein: das uneigentliche Integral.

7.1 Das Riemann-Integral im $\mathbb{R}^m$

Wir betrachten Funktionen $f : D(f) \to \mathbb{R}$, $D(f) \subseteq \mathbb{R}^m$, und definieren (schrittweise) Riemann-Integrale über immer komplizierteren Mengen, beginnend mit m-dimensionalen kompakten Intervallen $I \subseteq D(f)$ (oder "Quadern")

Definition 7.1

a) *Für beliebige a_i, $b_i \in \mathbb{R}$, $a_i < b_i$, $i = 1, \ldots, m$, heißt die Menge $I := \times_{i=1}^m \, [a_i, b_i]$ $\underline{m\text{-dimensionales kompaktes Intervall}}$.*
$\mu(I) := \prod_{i=1}^{m} (b_i - a_i)$ heißt $\underline{\text{Maß}}$ (bzw. Inhalt) von I.

b) *Ist $I := \times_{i=1}^m [a_i, b_i]$ ein m-dimensionales kompaktes Intervall und sind $Z_i := \{x_{i0}, x_{i1}, \ldots, x_{i,n(i)}\}$ Zerlegungen von $[a_i, b_i]$ $(i = 1, \ldots, m)$, d.h. $a_i = x_{i0} < x_{i1} < \ldots < x_{i,n(i)} = b_i$, so heißt $Z := \times_{i=1}^m Z_i$ $\underline{\text{Zerlegung}}$ von I.*
$\mathcal{Z}(I)$ bezeichne die Menge aller Zerlegungen von I.

Bemerkung 7.2
Ist $Z = \times_{i=1}^m Z_i \in \mathcal{Z}(I)$*, so kann man alle möglichen kartesischen Produkte von* m *Intervallen, die aus jeweils zwei aufeinanderfolgenden Punkten der Zerlegungen* $Z_1, \ldots, Z_m$ *gebildet werden, erzeugen. Man erhält dadurch eine gewisse Anzahl* M *(kleinerer)* m*-dimensionaler Intervalle* $I_1, \ldots, I_M$*, d.h.*

$$I_j = \times_{i=1}^m [x_{i,k_i}, x_{i,k_i+1}],\ k_i = k_i(j) \in \{0, \ldots, n(i)-1\}, i = 1, \ldots, m, j = 1, \ldots, M.$$

Für diese Intervalle gilt: $I = \bigcup_{j=1}^M I_j,\ \overset{\circ}{I}_j \cap \overset{\circ}{I}_k = \emptyset,\ k \neq j,$ *und*

$$\mu(I) = \sum_{j=1}^M \mu(I_j).$$

Eine Zerlegung $\tilde{Z} \in \mathcal{Z}(I)$ *heißt* *<u>Verfeinerung</u>* *von* $Z = \times_{i=1}^m Z_i$*, falls* $\tilde{Z}$ *die Gestalt* $\tilde{Z} = \times_{i=1}^m \tilde{Z}_i$ *hat und* $\tilde{Z}_i \supseteq Z_i$, $\forall i = 1, \ldots, m$*, gilt.*

Definition 7.3
Es sei I *ein* m*-dimensionaler kompaktes Intervall und* $f : I \to \mathbb{R}$ *beschränkt, d.h.* $f \in B(I, \mathbb{R})$*. Für jede Zerlegung* $Z \in \mathcal{Z}(I)$ *mit den "Teilintervallen"* I_j, $j = 1, \ldots, M$ *(vgl. 7.2) heißen*

$$s(Z; f) := \sum_{j=1}^M \inf_{x \in I_j} f(x)\mu(I_j) \quad bzw. \quad S(Z; f) := \sum_{j=1}^M \sup_{x \in I_j} f(x)\mu(I_j)$$

die <u>Darboux'sche Untersumme</u> bzw. <u>Obersumme</u> bez. f *und* Z*.*
Wir nennen weiterhin

$$\mathcal{I}_u := \sup_{Z \in \mathcal{Z}(I)} s(Z; f) \qquad bzw. \quad \mathcal{I}_o := \inf_{Z \in \mathcal{Z}(I)} S(Z; f)$$

das <u>Unter-Integral</u> bzw. <u>Ober-Integral von</u> f *<u>über</u>* I*.*

Lemma 7.4
Es sei I *ein* m*-dimensionales kompaktes Interval und* $f \in B(I, \mathbb{R})$*. Dann gilt für jedes* $Z \in \mathcal{Z}(I)$*:*
$-\infty < \inf_{x \in I} f(x) \cdot \mu(I) \leq s(Z; f) \leq \mathcal{I}_u \leq \mathcal{I}_o \leq S(Z; f) \leq \sup_{x \in I} f(x)\mu(I) < +\infty.$

<u>Beweis:</u>
Nach Definition und Konstruktion sind die folgenden Ungleichungen klar:

$$S(Z; f) \leq \mathcal{I}_u, \quad \mathcal{I}_o \geq S(Z; f), \quad s(Z; f) \leq S(Z; f)$$

$$-\infty < \inf_{x \in I} f(x) \cdot \mu(I) \leq \inf_{x \in I} f(x) \sum_{j=1}^M \mu(I_j) \leq \sum_{j=1}^M \inf_{x \in I_j} f(x)\mu(Ij) = s(Z; f)$$

und analog die beiden rechten Ungleichungszeichen in der Kette.
Es bleibt zu zeigen: $\mathcal{I}_u \leq \mathcal{I}_o$.
Sind zunächst Z und $\tilde{Z}$ Zerlegungen von I und ist $\tilde{Z}$ eine Verfeinerung von Z. Dann

setzt sich jedes Teilintervall I_j von A evtl. aus mehreren Teilintervallen $\tilde{I}j$ von $\tilde{Z}$ zusammen, d.h. es gilt z.B.

$$\inf_{x\in I_j} f(x)\mu(I_j) \leq \sum_i \inf_{x\in\tilde{I}_{j_i}} f(x)\mu(\tilde{I}_{j_i}),$$

was (durch aufsummieren über j) zur Ungleichung

$$s(Z;f) \leq s(\tilde{Z},f) \leq S(\tilde{Z};f) \leq S(Z;f) \qquad \text{führt.}$$

Sind ferner Z und $\bar{Z}$ zwei beliebige Zerlegungen von I und ist $\tilde{Z}$ eine "gemeinsame Verfeinerung" von Z und $\bar{Z}$, d.h. eine Zerlegung von I, die sowohl die Zerlegungspunkte zon Z als auch die von $\bar{Z}$ enthält. Dann gilt:

$$s(Z;f) \leq s(\tilde{Z};f) \leq S(\tilde{Z};f) \leq S(\bar{Z};f).$$

Daraus folgt: $\mathcal{I}_u = \sup\limits_{Z\in\mathcal{Z}(I)} s(Z;f) \leq S(\tilde{Z};f)$, $\mathcal{I}_u \leq \inf\limits_{Z\in\mathcal{Z}(I)} S(\tilde{Z};f) = \mathcal{I}_o$

Also gilt: $\mathcal{I}_u \leq \mathcal{I}_o$. $\square$

Definition 7.5

Es sei $I \subseteq \mathbb{R}^m$ ein m-dimensionales Intervall und $f \in B(I,\mathbb{R})$. $\mathcal{I}_u$ und $\mathcal{I}_o$ seien das Unter-Integral bzw. das Ober-Integral von f über I. Falls $\mathcal{I}_u = \mathcal{I}_o$, so heißt f über I Riemann-integrierbar und $\int\limits_I f(x)dx := \mathcal{I}_u = \mathcal{I}_o$ heißt Riemann-Integral von f über I.

Ehe wir zu Eigenschaften und zur Berechnung des Riemann-Integrals kommen, wenden wir uns der Frage nach der Charakterisierung Riemann-integrierbarer Funktionen zu. Unser erstes Beispiel zeigt, daß nicht alle beschränkten Funktionen Riemann-integrierbar sind. Anschließend kommen wir zu einer ersten positiven Antwort.

Beispiel 7.6

Wir setzen $m := 1$, $I := [0,1]$ und $f(x) := \begin{cases} 0, & x \in [0,1]\cap\mathbb{Q} \\ 1, & x \in [0,1]\backslash\mathbb{Q} \end{cases}$, d.h. f ist die Dirichlet-Funktion auf $[0,1]$.

Wir wissen: f ist nirgends stetig (vgl. Bsp. 5.5d)).

Ist nun Z eine beliebige Zerlegung von $[0,1]$, so folgt nach Definition:

$$s(Z;f) = 0 \quad , S(Z;f) = 1.$$

$$\rightsquigarrow \quad \mathcal{I}_u = 0 \quad \text{und} \quad \mathcal{I}_o = 1$$

$\rightsquigarrow$ *f ist nicht Riemann-integrierbar über I!*

Satz 7.7 *Es sei $I \subseteq \mathbb{R}^m$ ein m-dimensionales kompaktes Intervall. Dann ist jede stetige Funktion $f : I \to \mathbb{R}$, d.h. $f \in C(I)$, Riemann-integrierbar.*

Beweis:

Wir wählen $\varepsilon > 0$ beliebig und zeigen: $0 \leq \mathcal{I}_o - \mathcal{I}_u < \varepsilon$.

Nach Voraussetzung ist f auf I gleichmäßig stetig (Satz 5.17). Folglich existiert $\delta > 0$, so daß

$$|f(x) - f(\tilde{x})| < \frac{\varepsilon}{\mu(I)}, \quad \text{falls } x, \tilde{x} \in I \text{ mit } \|x - \tilde{x}\| < \delta.$$

Wir wählen nun $Z \in \mathcal{Z}(I)$ so, daß $\max\limits_{j=1,\ldots,M} \operatorname{diam}(I_j) < \delta$. Weiter gilt $\forall j \in \{1,\ldots,M\} \exists \bar{x}_j, \tilde{x}_j \in I_j$, so daß

$$f(\bar{x}_j) = \inf_{x \in I_j} f(x), \quad f(\tilde{x}_j) = \sup_{x \in I_j} f(x) \quad \text{(Satz 4.9).}$$

$$\begin{aligned}
\rightsquigarrow 0 \le \mathcal{I}_o - \mathcal{I}_u &\le S(Z;f) - s(Z;f) = \sum_{j=1}^{M} (\sup_{x \in I_j} f(x) - \inf_{x \in I_j} f(x)) \mu(I_j) \\
&= \sum_{j=1}^{M} (f(\tilde{x}_j) - f(\bar{x}_j)) \mu(I_j) < \frac{\varepsilon}{\mu(I)} \sum_{j=1}^{M} \mu(I_j) = \varepsilon
\end{aligned}$$

(wegen $\|\bar{x}_j - \tilde{x}_j\| \le \operatorname{diam}(I_j) < \delta$, $\forall j = 1,\ldots,M$). □

Unser nächstes Ziel besteht nun darin, die Klasse aller Riemann-integrierbaren Funktionen genau zu charakterisieren. Es wird sich zeigen, daß Riemann-integrierbare Funktionen "im wesentlichen" stetig sind. Dazu benötigen wir aber noch einige Vorbereitungen.

Definition 7.8
$A \subseteq \mathbb{R}^m$ heißt Nullmenge, wenn $\forall \varepsilon > 0$ eine höchstens abzählbare Menge $\{I_j : j \in \mathcal{J}\}$ von m-dimensionalen kompakten Intervallen existiert, so daß

$$A \subseteq \bigcup_{j \in \mathcal{J}} I_j \quad \textit{und} \quad \sum_{j \in \mathcal{J}} \mu(I_j) < \varepsilon.$$

Lemma 7.9

a) *Jede Teilmenge einer Nullmenge ist eine Nullmenge.*

b) *Endliche oder abzählbare Teilmengen von $\mathbb{R}^m$ sind Nullmengen.*

c) *Die Vereinigung höchstens abzählbar vieler Nullmengen ist wieder eine Nullmenge.*

Beweis:

a) ist offensichtlich nach Def. 7.8.

b) Es sei $\{x_j : j \in \mathcal{J}\}$ eine höchstens abzählbare Teilmenge des $\mathbb{R}^m$ und es sei $\varepsilon > 0$ beliebig gewählt. Dabei o.B.d.A. $\mathcal{J} \subseteq \mathbb{N}$.
Wir definieren $I_j := \{x \in \mathbb{R}^m : \|x - x_j\|_\infty \le \frac{1}{2} \left(\frac{\varepsilon}{2^{j+1}}\right)^{\frac{1}{m}}\}$, $\forall j \in \mathcal{J}$.

$$\rightsquigarrow \{x_j : j \in \mathcal{J}\} \subseteq \bigcup_{j \in \mathcal{J}} I_j \quad \text{und} \mu(I_j) = \prod_{i=1}^{m} (\frac{\varepsilon}{2^{j+1}})^{\frac{1}{m}} = \frac{\varepsilon}{2^{j+1}}$$

$$\rightsquigarrow \sum_{j \in \mathcal{J}} \mu(I_j) = \sum_{j \in \mathcal{J}} \frac{\varepsilon}{2^{j+1}} \le \varepsilon \sum_{j=1}^{\infty} \frac{1}{2^{j+1}} = \frac{\varepsilon}{2}.$$

c) Seien $\{A_i : i \in \mathcal{I}\}$ höchstens abzählbar viele Nullmengen, wobei o.B.d.A. $\mathcal{I} \subseteq I\!N$.
Wir setzen $A := \bigcup_{i\in\mathcal{I}} A_i$ und es sei $\varepsilon > 0$ beliebig gewählt.

$\rightsquigarrow$ $\exists$ höchstens abzählbare Mengen $\{I_{ij} : j \in \mathcal{J}_i\}$ (o.B.d.A. $\mathcal{J}_i \subseteq I\!N$) von m-dimensionalen kompakten Intervallen, so daß

$$A_i \subseteq \bigcup_{j\in\mathcal{J})_i} I_{ij} \quad \text{und} \quad \sum_{j\in\mathcal{J}_i} \mu(I_{ij}) < \tfrac{\varepsilon}{2^i}, \quad \forall i \in \mathcal{I}.$$

$\rightsquigarrow$ $\{I_{ij} : j \in \mathcal{J}_i, i \in \mathcal{I}\}$ ist höchstens abzählbar und es gilt

$$A \subseteq \bigcup_{i\in\mathcal{I}} \bigcup_{j\in\mathcal{J}_i} I_{ij} \quad \text{und} \quad \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}_i} \mu(I_{ij}) < \varepsilon \sum_{i\in\mathcal{I}} \tfrac{1}{2^i} \le \varepsilon.$$

$\rightsquigarrow$ A ist Nullmenge. $\square$

Beispiel 7.10
Jede Hyperebene $H := \{(x_1, \ldots, x_m) \in \mathbb{R}^m : x_s = c\}$ ($s \in \{1, \ldots, m\}$ und $c \in \mathbb{R}$ fixiert) ist eine Nullmenge.

Bew.: *Sei $\varepsilon > 0$ beliebig gewählt und wir definieren $I_j := \times_{i=1}^{s}[a_{ij}, b_{ij}]$ wobei $a_{ij} := -j$, $b_{ij} := j$, $\forall i \ne s$, $\forall j \in I\!N$,*

$$a_{sj} := c - \tfrac{\varepsilon}{2^{j+2}(2j)^{m-1}}, \; b_{sj} := c + \tfrac{\varepsilon}{2^{j+2}(2j)^{m-1}}.$$

$$\rightsquigarrow H \subseteq \bigcup_{j=1}^{\infty} I_j \quad \textit{und} \quad \sum_{j=1}^{\infty} \mu(I_j) = \sum_{j=1}^{\infty} (2j)^{m-1} \frac{2\varepsilon}{2^{j+2}(2j)^{m-1}} = \frac{\varepsilon}{2}. \quad \square$$

Satz 7.11 *(Lebesgue)*
Es sei $I \subseteq \mathbb{R}^m$ ein m-dimensionales kompaktes Intervall und $f : B(I, \mathbb{R})$. f ist Riemann-integrierbar über I gdw. eine Nullmenge $A \subseteq I$ existiert, so daß f auf $I \setminus A$ stetig ist.

(ohne Beweis)

Folgerung 7.12
Ist $f : [a, b] \to \mathbb{R}$ ($a, b \in \mathbb{R}, a < b$) beschränkt und hat höchstens abzählbar viele Unstetigkeitsstellen, so ist f Riemann-integrierbar.

Beweis:
folgt aus Satz 7.11 und Lemma 7.9b). $\square$

Die Lebesgue'sche Charakterisierung der Riemann-Integrierbarkeit in Satz 7.11 wird uns im folgenden bei der Herleitung von Eigenschaften des Riemann-Integrals bzw. beim weiteren Aufbau der Integralrechnung wertvolle Dienste leisten.

Satz 7.13
Es sei $I \subseteq \mathbb{R}^m$ ein m-dimensionales kompaktes Intervall und $f, g \in B(I, \mathbb{R})$ seien Riemann-integrierbar. Dann gilt:

a) *Für alle $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g$ Riemann-integrierbar und es gilt*

$$\int_I (\alpha f + \beta g)(x)dx = \alpha \int_I f(x)dx + \beta \int_I g(x)dx.$$

b) *Falls* $f(x) \geq 0$, $\forall x \in I$, *so gilt* $\int\limits_I f(x)dx \geq 0$.

c) $|f|$ *ist Riemann-integrierbar und es gilt* $|\int\limits_I f(x)dx| \leq \int\limits_I |f(x)|dx$.

Beweis:
Die Riemann-Integrierbarkeit von $\alpha f + \beta g$ und von $|f|$ folgt jeweils sofort aus Satz 7.11.

a) Es sei zunächst $\alpha = \beta = 1$ und $\varepsilon > 0$ beliebig gewählt. Dann existieren nach Voraussetzung Zerlegungen $Z, \tilde{Z} \in \mathcal{Z}(I)$ mit

$$0 \leq S(Z;f) - s(Z;f) < \frac{\varepsilon}{2}, \quad 0 \leq S(\tilde{Z};g) - s(\tilde{Z};g) < \frac{\varepsilon}{2}.$$

$$\rightsquigarrow -\frac{\varepsilon}{2} < s(Z;f) - S(Z;f) < s(Z;f) - \int\limits_I f(x)dx$$
$$\rightsquigarrow -\frac{\varepsilon}{2} < s(\tilde{Z};g) - S(\tilde{Z};g) < s(\tilde{Z};g) - \int\limits_I g(x)dx$$

Es sei nun Z^* eine gemeinsame Verfeinerung von Z und $\tilde{Z}$, d.h. $Z^* \supseteq Z \cup \tilde{Z}$, mit den Teilintervallen I_j, $j = 1, \ldots, M$. Dann gilt zunächst:

$$\begin{aligned} -\frac{\varepsilon}{2} < s(Z;f) - \int\limits_I f(x)dx &\leq s(Z^*;f) - \int\limits_I f(x)dx \quad \text{(Lemma 7.4)} \\ \leq S(Z^*;f) - \int\limits_I f(x)dx &\leq S(Z;f) - \int\limits_I f(x)dx < \frac{\varepsilon}{2} \end{aligned}$$

(und analoges für $\tilde{Z}$ bzw. g anstelle von Z bzw. f), und

$$\begin{aligned} s(Z^*; f+g) &= \sum_{j=1}^{M} \inf_{x \in I_j} (f+g)(x)\mu(I_j) \geq s(Z^*;f) + s(Z^*;g) \\ S(Z^*; f+g) &= \sum_{j=1}^{M} \sup_{x \in I_j} (f+g)(x)\mu(I_j) \leq S(Z^*;f) + S(Z^*;g). \end{aligned}$$

Daraus folgt insgesamt:

$$\begin{aligned} -\varepsilon &< s(Z;f) + s(\tilde{Z};f) - (\int\limits_I f(x)dx + \int\limits_I g(x)dx) \\ &\leq s(Z^*; f+g) - (\int\limits_I f(x)dx + \int\limits_I g(x)dx) \end{aligned}$$

$$\begin{aligned}
&\le S(Z^*;f+g)-(\int\limits_I f(x)dx+\int\limits_I g(x)dx)\\
&\le (S(Z;f)-\int\limits_I f(x)dx)+(S(\tilde{Z};g)-\int\limits_I g(x)dx)<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.
\end{aligned}$$

Also gilt: $-\varepsilon < \int\limits_I (f+g)(x)dx - (\int\limits_I f(x)dx + \int\limits_I g(g)dx) < \varepsilon$

wegen $\quad s(Z^*;f+g) \le \int\limits_I (f+g)(x)dx \le S(Z^*;f+g)$.

Wir haben nun die Aussage für $\alpha=\beta=1$ gezeigt!
Es bleibt zu zeigen: $\int\limits_I (\alpha f)(x)dx = \alpha \int\limits_I f(x)dx$, $\forall \alpha \in \mathbb{R}$.

Sei zunächst $\alpha \ge 0$. Dann gilt:

$$\int\limits_I (\alpha f)(x)dx = \sup_{Z\in\mathcal{Z}(I)} s(Z;\alpha f) = \sup_{Z\in\mathcal{Z}(I)} \alpha s(Z;f) = \alpha \int\limits_I f(x)dx.$$

Ferner erhalten wir:

$$\begin{aligned}
\int\limits_I (-f)(x)dx &= \sup_{Z\in\mathcal{Z}(I)} s(Z;-f) = \sup_{z\in\mathcal{Z}(I)} -S(Z;f)\\
&= -\inf_{Z\in\mathcal{Z}(I)} S(Z;f) = -\int\limits_I f(x)dx.
\end{aligned}$$

Damit ist Teil a) vollständig bewiesen!

b) Es gelte $f(x)\ge 0$, $\forall x\in I$. Dann folgt:

$$\int\limits_I f(x)dx = \sup_{Z\in\mathcal{Z}(I)} s(Z;f) \ge \sum_{j=1}^{M} \inf_{x\in I_j} f(x)\mu(I_j) \ge 0.$$

c) Es gilt: $|f|\pm f\ge 0$. Folglich ergibt sich aus a) und b):

$$\int\limits_I |f(x)|dx \pm \int\limits_I f(x)dx = \int\limits_I (|f(x)|\pm f(x))dx \ge 0.$$

Also: $-\int\limits_I |f(x)|dx \le \int\limits_I f(x)dx \le \int\limits_I |f(x)|dx$ und die Behauptung folgt. □

Satz 7.14
Es seien I, $I^{(1)}$ und $I^{(2)}$ m-dimensionale kompakte Intervalle und es gelte $I=I^{(1)}\cup I^{(2)}$ und $\operatorname{int}(I^{(1)})\cap\operatorname{int}(I^{(2)})=\emptyset$.
Es sei $f\in B(I,\mathbb{R})$ Riemann-integrierbar über I.
Dann ist f auch Riemann-integrierbar über $I^{(1)}$ und $I^{(2)}$ und es gilt

$$\int\limits_I f(x)dx = \int\limits_{I^{(1)}} f(x)dx + \int\limits_{I^{(2)}} f(x)dx.$$

Beweis:
Die Riemann-Integrierbarkeit von f über $I^{(1)}$ und $I^{(2)}$ folgt sofort aus der Voraussetzung und aus Satz 7.11. Deshalb existieren für $j = 1, 2$ und zu beliebig vorgegebenem $\varepsilon > 0$ Zerlegungen $Z_j \in \mathcal{Z}(I^{(j)})$, so daß

$$S(Z_j; f) - s(Z_j; f) < \frac{\varepsilon}{2}.$$

Es sei $Z \in \mathcal{Z}(I)$ derart gewählt, daß $Z \cap I^{(j)}$ eine Verfeinerung von Z_j für $j = 1, 2$ darstellt. Dann gilt:

$$\begin{aligned}
&- \varepsilon + \int\limits_{I^{(1)}} f(x)dx + \int\limits_{I^{(2)}} f(x)dx < s(Z_1; f) + s(Z_2; f) \\
&\leq s(Z \cap I^{(1)}; f) + s(Z \cap I^{(2)}; f) = s(Z; f) \\
&\leq \int\limits_I f(x)dx \leq S(Z; f) \leq S(Z \cap I^{(1)}; f) + S(Z \cap I^{(2)}; f) \\
&\leq S(Z_1; f) + S(Z_2; f) < \int\limits_{I^{(1)}} f(x)dx + \int\limits_{I^{(2)}} f(x)dx + \varepsilon. \qquad \square
\end{aligned}$$

Satz 7.15 *(Mittelwertsatz)*
Es seien $I \subseteq \mathbb{R}^m$ ein m-dimensionales kompaktes Intervall und $f \in C(I)$.
Dann existiert ein $\bar{x} \in I$ mit $\int\limits_I f(x)dx = f(\bar{x})\mu(I)$.

Beweis: Nach Satz 7.7 (bzw. 7.11) ist f Riemann-integrierbar über I. Nach Lemma 7.4 gilt für alle $Z \in \mathcal{Z}(I)$:

$$\inf_{x \in I} f(x)\mu(I) \leq s(Z; f) \leq \int\limits_I f(x)dx \leq S(Z; f) \leq \sup_{x \in I} f(x)\mu(I)$$

$$\rightsquigarrow \qquad \inf_{x \in I} f(x) \leq \frac{1}{\mu(I)} \int\limits_I f(x)dx \leq \sup_{x \in I} f(x).$$

Da f auf I stetig ist, existieren nach Satz 5.9 (Satz von Weierstraß) x_*, $x^* \in I$, so daß

$$f(x_*) = \inf_{x \in I} f(x) \quad \text{und} \quad f(x^*) = \sup_{x \in I} f(x).$$

$$\rightsquigarrow \qquad f(x_*) \leq \frac{1}{\mu(I)} \int\limits_I f(x)dx \leq f(x^*).$$

Durch komponentenweise Anwendung der Beweisidee des eindimensionalen Zwischenwertsatz 5.12 zeigt man, daß $[f(x_*), f(x^*)] \subseteq f(I)$ gilt. Insbesondere ergibt sich

$$\frac{1}{\mu(I)} \int\limits_I f(x)dx \in f(I).$$

d.h. $\exists \bar{x} \in I$ mit $f(\bar{x}) = \frac{1}{\mu(I)} \int\limits_I f(x)dx$.

Satz 7.16
Es sei $I \subseteq \mathbb{R}^m$ ein m-dimensionales kompaktes Intervall.
Es sei $f : I \mapsto \mathbb{R}$ stetig mit $f(x) \geq 0$, $\forall x \in I$, und $\int\limits_I f(x)dx = 0$. Dann gilt: $f(x) = 0$, $\forall x \in I$.

<u>Beweis:</u>
<u>Annahme:</u> $\exists \tilde{x} \in I : f(\tilde{x}) > 0$.
Wegen der Stetigkeit von f auf I folgt:
$\exists \delta > 0 : f(x) > 0$, $\forall x \in \bar{B}_\infty(\tilde{x}, \delta) := \{y \in I : \|y - \tilde{x}\|_\infty \leq \delta\}$. Wir zerlegen nun I in endlich viele Teilintervalle, zu denen auch $\tilde{I} := \bar{B}_\infty(\tilde{x}, \delta)$ gehört. Dann folgt aus den Sätzen 7.13b) und 7.14:

$$\int\limits_I f(x)dx \geq \int\limits_{\tilde{I}} f(x)dx.$$

7.15 $\rightsquigarrow \exists \bar{x} \in \tilde{I} : \int\limits_{\tilde{I}} f(x)dx = f(\bar{x})\mu(I)$.
$\rightsquigarrow \int\limits_I f(x)dx \geq f(\bar{x})\mu(\tilde{I}) > 0$.
$\rightsquigarrow$ Widerspruch zur Annahme! $\square$

Wir kommen nun zur Definition des Riemann-Integrals über allgemeineren Mengen $B \subseteq \mathbb{R}^m$ und zur Definition des Maßes bzw. Inhaltes für solche Mengen.

Definition 7.17
Es sei $B \subseteq \mathbb{R}^m$ nichtleer und beschränkt, $I \subseteq \mathbb{R}^m$ sei ein m-dimensionales kompaktes Intervall mit $B \subseteq I$.
$f : B \to \mathbb{R}$ heißt <u>Riemann-integrierbar</u> über B (kurz: <u>R-integrierbar</u>) wenn die Funktion $f_B : \mathbb{R}^m \to \mathbb{R}$, wobei
$f_B(x) := \begin{cases} f(x) & , x \in B \\ 0 & , x \notin B \end{cases}$, *Riemann-integrierbar über I ist.*
In diesem Fall heißt $\int\limits_B f(x)dx := \int\limits_I f_B(x)dx$ <u>Riemann-Integral</u> (oder kurz: <u>R-Integral</u>) von f über B.
B heißt <u>Jordan-meßbar</u>, wenn die charakteristische Funktion χ_B von B, d.h.
$\chi_B(x) := \begin{cases} 1 & , x \in B \\ 0 & , x \notin B \end{cases}$, *über I R-integrierbar ist.*
$\mu(B) := \int_B \chi_B(x)dx =: \int\limits_B dx$ heißt dann <u>Jordan-Inhalt</u> (oder kurz: Maß) von B.
Für $m = 2$ bzw. $m = 3$ heißt $\mu(B)$ auch Flächeninhalt bzw. Volumen von B.

Bemerkung 7.18
Die Definition des R-Integrals über B ist von der Wahl des m-dimensionalen kompakten Invervalls $I \supseteq B$ unabhängig. Denn ist $\tilde{I}$ ein weiteres m-dimensionales kompaktes Intervall mit $\tilde{I} \supseteq B$, so gilt: $\int\limits_I f_B(x)dx = \int\limits_{\tilde{I}} f_B(x)dx$.

Beweis: *es wird gezeigt:* $\int\limits_I f_B(x)dx = \int\limits_{I\cap\tilde I} f_B(x)dx = \int\limits_{\tilde I} f_B(x)dx.$
o.B.d.A. zeigen wir es für $\tilde I$.
es existieren endlich viele Teilintervalle $\tilde I_j$, $j \in \mathcal{J}$, *von* $\tilde I$,
so daß $\tilde I = (I \cap \tilde I) \cup (\bigcup\limits_{j\in\mathcal{J}} \tilde I_j)$ *und das Innere der Intervalle auf der rechten Seite paarweise disjunkt ist. Aus Satz 7.14 folgt dann:*

$$\int\limits_{\tilde I} f_B(x)dx = \int\limits_{I\cap\tilde I} f_B(x)dx + \sum_{j\in\mathcal{J}} \underbrace{\int\limits_{\tilde I_j} f_B(x)dx}_{=0} = \int\limits_{I\cap\tilde I} f_B(x)dx. \square$$

Ist $B \subseteq \mathbb{R}^m$ *Jordan-meßbar, so gilt für den Jordan-Inhalt:*

$$\begin{aligned}\mu(B) &= \sup_{Z\in\mathcal{Z}(I)} \sum_{j=1}^{M} \inf_{x\in I_j} \chi_B(x)\mu(I_j) = \sup_{Z\in\mathcal{Z}(I)} \sum_{\substack{j=1\\ I_j\subseteq B}}^{M} \mu(I_j)\\ &= \inf_{Z\in\mathcal{Z}(I)} \sum_{j=1}^{M} \sup_{x\in I_j} \chi_B(x)\mu(I_j) = \inf_{Z\in\mathcal{Z}(I)} \sum_{\substack{j=1\\ I_j\cap B=\emptyset}}^{M} \mu(I_j)\end{aligned}$$

(wobei $I \supseteq B$ *ein* m*-dimensionales kompaktes Intervall ist).*
Die "anschauliche Vorstellung" des Jordan-Inhaltes ist das Supremum (bzw. Infinum) von Summen von Maßen vom Intervallen, die ganz in B *liegen bzw. mit* B *einen nichtleeren Durchschnitt haben.*

Satz 7.19
Eine beschränkte Menge $B \subseteq \mathbb{R}^m$ *ist Jordan-meßbar gdw. der Rand* $\operatorname{bd} B$ *von* B *eine Nullmenge ist.*

Beweis:
Wir wählen ein m-dimensionales kompaktes Intervall I so, daß $\overset{\circ}{I} \supseteq \operatorname{cl} B$. B ist Jordan-meßbar gdw. χ_B R-integrierbar über I ist (vgl. Def. 7.17). Die Menge der Unstetigkeitspunkte von χ_B in I ist gerade $\operatorname{bd} B \rightsquigarrow \chi_B$ ist R-integrierbar über B nach Satz 7.11 gdw. die Menge $\operatorname{bd} B$ eine Nullmenge ist. $\square$

Beispiel 7.20
Für $m := 1$ *betrachten wir* $B := [0,1] \cap \mathbb{Q}$, *d.h.* B *ist beschränkt und abzählbar, d.h., insbesondere eine Nullmenge nach 7.9. Es gilt* $\operatorname{bd} B = [0,1]$.
$\rightsquigarrow \operatorname{bd} B$ *ist keine Nullmenge (es gilt* $\mu(\operatorname{bd} B) = 1$) *und folglich (Satz 7.19) ist* B *keine Jordan-meßbare Menge.*
$\rightsquigarrow$ B *ist eine nicht Jordan-meßbare Nullmenge.*
Dies sind die Grenzen dieses Meßbarkeitsbegriffs !

Satz 7.21
Es sei $B \subseteq \mathbb{R}^m$ *Jordan-meßbar und* $f : B \to \mathbb{R}$ *sei beschränkt.*
f *ist* R*-integrierbar über* B *gdw.* $\exists$ *Nullmenge* $A \subseteq B$, *so daß* f *auf* $B \setminus A$ *stetig ist.*

Beweis:
Es sei $I \subseteq \mathbb{R}^m$ ein m-dimensionales kompaktes Intervall mit $B \subseteq I$.
($\Longrightarrow$) Sei f R-integrierbar über $B \rightsquigarrow f_B$ ist R-integrierbar über $I \rightsquigarrow \exists$ Nullmenge $A \subseteq I$, so daß f_B stetig auf $I \setminus A$ ist (7.11!)
$\rightsquigarrow f_B$ ist erst recht stetig auf $B \setminus A$. (betrachten $\tilde{A} := A \cap B$ anstelle von A)
$\rightsquigarrow f$ ist stetig auf $B \setminus A$.
($\Longleftarrow$) Es bezeichne $D_f(B)$ die Menge aller Unstetigkeitspunkte von f in B und $D_{f_B}(I)$ die Menge aller Unstetigkeitspunkte von f_B in I. Es sei nun $D_f(B)$ eine Nullmenge.
$\rightsquigarrow$ $D_f(B) \cup \operatorname{bd} B$ ist eine Nullmenge nach Satz 7.19.
$\rightsquigarrow$ $D_{f_B}(I) \subseteq D_f(B) \cup \operatorname{bd} B$ ist Nullmenge.
$\rightsquigarrow$ f_B ist R-integrierbar über I (Satz 7.11), d.h. f ist R-integrierbar über B. $\square$

Bemerkung 7.22
Satz 7.13 bleibt gültig für Riemann-Integrale über Jordan-meßbaren Mengen $B \subseteq \mathbb{R}^m$ (mit analogem Beweis und unter Verwendung von Satz 7.21).
Überdies gilt analog zu 7.4 für jede Jordan-meßbare Menge $B \subseteq \mathbb{R}^m$ und jede beschränkte R*-integrierbare Funktion $f : B \to \mathbb{R}$, daß*

$$\inf_{x \in B} f(x)\mu(B) \leq \int_B f(x)dx \leq \sup_{x \in B} f(x)\mu(B).$$

Ein wichtiges Hilfsmittel zur Berechnung von Riemann-Integralen über mehrdimensionalen Mengen ist die Rückführung auf die mehrmals hintereinander ausgeführte Integration über niedrigdimensionale Mengen. Wir beginnen mit dem Fall von R-Integralen über Intervallen.

Satz 7.23 *(Fubini)*
Es seien I_y und I_z kompakte p-dimensionale bzw. q-dimensionale Intervalle und I bezeichne das $m = (p+q)$-dimensionale Produktintervall $I := I_y \times I_z$.
Es sei $f : I \to \mathbb{R}$ beschränkt und R*-integrierbar über I, und für alle $z \in I_z$ sei $f(\cdot, z)$* R*-integrierbar über I_y.*
Dann ist die Funktion $g : I_z \to \mathbb{R}$, $g(z) := \int_{I_y} f(y,z)dy$, $\forall z \in I_z$, R*-integrierbar über I_z und es gilt*

$$\int_I f(x)dx = \int_{I_z} g(z)dz = \int_{I_z} (\int_{I_y} f(y,z)dy)dz.$$

Beweis:
Es sei $\varepsilon > 0$ beliebig und $Z \in \mathcal{Z}(I)$ so gewählt, daß

$$S(Z; f) - s(Z; f) < \varepsilon.$$

Nach Definition hat Z die Gestalt $Z = Z_y \times Z_z$, wobei $Z_y \in \mathcal{Z}(I_y)$ und $Z_z \in \mathcal{Z}(I_z)$. Es seien nun I_j, $j = 1, \ldots, M$, die p-dimensionalen Teilintervalle von Z_y und $\tilde{I}_i$, $i =$

$1, \ldots, \tilde{M}$, die q-dimensionalen Teilintervalle von Z_z.
Dann gilt für die Darboux'sche Untersumme $s(Z; f)$:

$$\begin{aligned} s(Z;f) &= \sum_{j=1}^{M}\sum_{i=1}^{\tilde{M}} \inf_{(y,z)\in I_j\times\tilde{I}_i} f(y,z)\mu(I_j\times\tilde{I}_i) \\ &= \sum_{i=1}^{\tilde{M}}(\sum_{j=1}^{M} \inf_{(y,z)\in I_j\times\tilde{I}_i} f(y,z)\mu(I_j))\mu(\tilde{I}_i). \end{aligned}$$

Für alle $z \in \tilde{I}_i$ und $i = 1, \ldots, \tilde{M}$ gilt:

$$g_i := \sum_{j=1}^{M} \inf_{(y,z)\in I_j\times\tilde{I}_i} f(y,z)\mu(I_j) \leq \sum_{j=1}^{M} \inf_{y\in I_j} f(y,z)\mu(I_j) \leq g(z).$$

Also: $g_i \leq \inf\limits_{z\in\tilde{I}_i} g(z)$, $\forall i = 1, \ldots, \tilde{M}$.
$\rightsquigarrow s(Z;f) \leq \sum\limits_{i=1}^{\tilde{M}} g_i\mu(\tilde{I}_i) \leq \sum\limits_{i=1}^{\tilde{M}} \inf\limits_{z\in\tilde{I}_i} g(z)\mu(\tilde{I}_i) = s(Z_z; g)$.
Analog zeigt man auch: $S(Z_z; g) \leq S(Z; f)$.
Daraus folgt insgesamt:

$$0 \leq S(Z_z; g) - s(Z_z; g) \leq S(Z; f) - s(Z; f) < \varepsilon$$

d.h. g ist R-integrierbar über I_z und es gilt

$$\int\limits_I f(x)dx = \sup_{Z\in\mathcal{Z}(I)} s(Z;f) \leq \int\limits_{I_z} g(z)dz \leq \inf_{Z\in\mathcal{Z}(I)} S(Z;f) = \int\limits_I f(x)dx. \qquad \square$$

Folgerung 7.24
Gilt zusätzlich zu den Voraussetzungen von Satz 7.23 auch, daß für jedes $y \in I_y$ die Funktion $f(y, \cdot)$ R-integrierbar über I_z ist, so gilt:

$$\int\limits_I f(x)dx = \int\limits_{I_y}(\int\limits_{I_z} f(y,z)dz)dy = \int\limits_{I_z}(\int\limits_{I_y} f(y,z)dy)dz$$

(Vertauschung der Integrationsreihenfolge).

Beweis:
Ein Teil folgt direkt aus Satz 7.23, der andere Teil analog zu dessen Beweis durch anderes Aufspalten der Untersumme (Doppelsumme). $\square$

Beispiel 7.25
Aus der R-Integrierbarkeit von f über $I := I_y \times I_z$ folgt i.a. nicht, daß $f(\cdot, z)$ R-integrierbar ist über I_y, $\forall Z \in I_z$!
$m := 2$, $I := [0,1] \times [0,1]$, $f : I \to \mathbb{R}$,

$$f(y,z) := \begin{cases} 1 & ,(y,z)\in I, y \neq \frac{1}{2} \\ 1 & ,y = \frac{1}{2}, z \in [0,1]\setminus\mathbb{Q}, \\ 0 & ,y = \frac{1}{2}, z\in[0,1]\cap\mathbb{Q}. \end{cases}$$

$\rightsquigarrow$ *$f(\frac{1}{2},\cdot):[0,1]\to\mathbb{R}$ ist <u>nicht</u> R-integrierbar über $[0,1]$ (7.6!)*
aber die Menge aller Unstetigkeitsstellen von f in I ist gerade $\{(\frac{1}{2},z):z\in[0,1]\}$ und damit nach Bsp. 7.10 Nullmenge.
Folglich ist f R-integrierbar über I nach Satz 7.11!

Folgerung 7.26
Es sei $I:=\times_{i=1}^{m}[a_i,b_i]$ ein m-dimensionales kompaktes Intervall und es gelte $f\in C(I)$. Dann gilt:

$$\int_I f(x)dx=\int_{a_1}^{b_1}(\int_{a_2}^{b_2}\cdots(\int_{a_m}^{b_m}f(x_1,\ldots,x_m)dx_m)dx_{m-1}\cdots)dx_2)dx_1,$$

wobei alle auftretenden R-Integrale existieren und die Integrationsreihenfolge beliebig vertauscht werden darf.

<u>Beweis:</u>
folgt durch wiederholte Anwendung von 7.23 bzw. 7.24 und wegen $f\in C(I)$:
Zunächst mit: $I_y=[a_1,b_1], I_z=\times_{i=2}^{m}[a_i,b_i], y=y_1, z=(x_2,\ldots,x_m)$
$\rightsquigarrow \int\limits_I f(x)dx=\int\limits_{a_1}^{b_1}(\int\limits_{I_z}f(x_1,z)dz)dx_1$ und danach sukzessive fortgesetzt.
Satz 7.23 bzw. Folgerung 7.24 sind anwendbar, da wegen der Stetigkeit von f alle R-Integrierbarkeitsvoraussetzungen nach 7.7 bzw. 7.11 erfüllt sind. Die Vertauschung der Integrationsreihenfolge ist möglich nach 7.24 bzw. 7.23. $\square$

Satz 7.27 *(allgemeiner Satz von Fubini)*
Es sei $B\subseteq\mathbb{R}^m$ Jordan-meßbar und $f:B\to\mathbb{R}$ sei beschränkt und R-integrierbar. Für fixierte $p,q\in\mathbb{N}$ mit $p+q=m$ setzen wir $P(B):=\{y\in\mathbb{R}^p:\exists z\in\mathbb{R}^q \text{ mit } (y,z)\in B\}$ ("Projektion von B auf $\mathbb{R}^p$") und $B_y:=\{z\in\mathbb{R}^q:(y,z)\in B\}$, $\forall y\in P(B)$. Für alle $y\in P(B)$ sei ferner das R-Integral

$$g(y):=\int\limits_{B_y}f(y,z)dz \qquad \text{definiert.}$$

Dann existiert das R-Integral $\int\limits_{P(B)}g(y)dy$ und es gilt

$$\int_B f(x)dx=\int_{P(B)}(\int_{B_y}f(y,z)dz)dy.$$

<u>Beweis:</u>
Da B beschränkt ist, existieren kompakte Intervalle $I\subseteq\mathbb{R}^p$ und $\tilde I\subseteq\mathbb{R}^q$ so daß $B\subseteq I\times\tilde I$. Dann gilt:

$$P(B)\subseteq I \quad \text{und} \quad B_y\subseteq\tilde I, \quad \forall y\in P(B).$$

Nach Voraussetzung und nach Def. 7.17 folgt dann:

$$\begin{aligned}\int_B f(x)dx &= \int_{I\times\tilde I}f_B(x)dx=\int_{I\times\tilde I}f_B(y,z)d(y,z)\\ g(y) &= \int_{B_y}f(y,z)dz=\int_{\tilde I}f_B(y,z)dz, \quad \forall y\in P(B).\end{aligned}$$

Das letzte Gleichheitszeichen ergibt sich dabei daraus, daß für bel. $y \in P(B)$ und alle $z \in \tilde{I}$ gilt

$$\begin{aligned} f_B(y,z) &= \begin{cases} f(y,z) & ,(y,z) \in B \\ 0 & ,\text{sonst} \end{cases} \\ &= \begin{cases} f(y,z) & ,z \in B_y \\ 0 & ,z \in \tilde{I} \backslash B_y \end{cases} . \end{aligned}$$

Da nach Vor. f_B R-integrierbar über $I \times \tilde{I}$ und $f_B(y, \cdot)$ für jedes $y \in P(B)$ R-integrierbar über $\tilde{I}$ ist, folgt aus Folg. 7.24, daß auch $g_{P(B)}$ R-integrierbar über I ist und daß gilt

$$\int\limits_{I \times \tilde{I}} f_B(x)dx = \int\limits_I (\int\limits_{\tilde{I}} f_B(y,z)dz)dy = \int\limits_I g_{P(B)}(y)dy = \int\limits_{P(B)} g(y)dy. \qquad \square$$

Folgerung 7.28 *(Satz von Cavalieri)*
Es sei $B \subseteq \mathbb{R}^m$ Jordan-meßbar und für alle $y \in P(B)$ sei auch B_y Jordan-meßbar. Dann gilt:

$$\mu(B) = \int\limits_{P(B)} \mu(B_y)dy.$$

(Dabei seien $P(B) \subseteq \mathbb{R}^p$, $B_y \subseteq \mathbb{R}^q$, $p + q = m$ wie in 7.27 definiert.)

Beweis:
Die Aussage resultiert sofort aus Satz 7.27 mit $f := \chi_B$, denn es gilt

$$\mu(B) = \int\limits_I \chi_B(x)dx = \int\limits_B dx = \int\limits_{P(B)} (\int\limits_{B_y} \chi_B(y,z)dz)dy = \int\limits_{P(B)} \mu(B_y)dy$$

(χ_B ist R-integrierbar über B, da bd B, und damit die Menge aller Unstetigkeitsstellen von χ_B eine Nullmenge ist (7.19). Ferner ist $\forall y \in P(B)\ \chi_B(y;\cdot)$ R-integrierbar über B_y, da bd B (und damit die Menge aller Unstetigkeitsstellen von $\chi_B(y;\cdot)$) eine Nullmenge ist. Damit sind alle Voraussetzungen von 7.27 erfüllt. $\square$

Im folgenden soll nun eine Klasse von Mengen diskutiert werden, die einerseits hinreichend allgemein ist und auf die andererseits Satz 7.27 und Folgerung 7.28 angewendet werden können (u.a. zur Berechnung von Jordan-Inhalten).

Definition 7.29
Eine nichtleere Menge $B \subseteq \mathbb{R}^m$ heißt Zylindermenge (oder: Normalbereich), falls Zahlen $a, b \in \mathbb{R}$ und Funktionen $\varphi_j,\ \psi_j \in C(\mathbb{R}^{m-j})$, $j = 1, \ldots, m-1$, existieren, so daß

$$B = \{(x_1, \ldots, x_m) \in \mathbb{R}^m : a \le x_m \le b, \varphi_j(x_{j+1}, \ldots, x_m) \le x_j \le \psi_j(x_{j+1}, \ldots, x_m), 1 \le j < m\}.$$

Beispiel:
$m = 2,\ B = \{(x_1, x_2) : a \le x_2 \le b,\ \varphi_1(x_2) \le x_1 \le \psi_1(x_2)\}$
Für den Rand bd B der Menge B gilt:

$$\begin{aligned} \text{bd}\, B = &\{(x_1, a) : x_1 \in [\varphi_1(a), \psi_1(a)]\} \cup \{(x_1, b) : x_1 \in [\varphi_1(b), \psi_1(b)]\} \\ &\cup \{(x_1, x_2) : x_1 = \varphi_1(x_2), x_2 \in [a, b]\} \cup \{(x_1, x_2) : x_1 = \psi_1(x_2), x_2 \in [a, b]\} \end{aligned}$$

Satz 7.30
Jede Zylindermenge $B \subseteq \mathbb{R}^m$ ist kompakt und Jordan-meßbar.

Beweis:
Wegen der Stetigkeit der Funktionen φ_i, ψ_j, $j = 1, \ldots, m-1$, in der Def. der Zylindermenge B ist diese abgeschlossenen und wegen $x_m \in [a, b]$ auch beschränkt. In der Tat ist B in folgendem kompakten Intervall I enthalten:

$$I := (\times_{i=1}^{m-1}[\bar{\varphi}_i, \bar{\psi}_i]) \times [a, b]$$

wobei $\bar{\varphi}_{m-1} := \inf\limits_{x_m \in [a,b]} \varphi_{m-1}(x_m)$, $\quad \bar{\psi}_{m-1} := \sup\limits_{x_m \in [a,b]} \psi_{m-1}(x_m)$ usw.
$\bar{\varphi}_1 := \inf\{\varphi_1(x_2, \ldots, x_m) : x_m \in [a, b], x_i \in [\bar{\varphi}_i, \bar{\psi}_i], i = 2, \ldots, m-1$
und $\bar{\psi}_1$ analog definiert.

Also ist B kompakt. Auf den Nachweis der Jordan-Meßbarkeit verzichten wir. □

Folgerung 7.31
Es sei $B \subseteq \mathbb{R}^m$ eine Zylindermenge und $f : B \mapsto \mathbb{R}$ sei stetig. Dann gilt:

$$\int_B f(x)dx = \int_a^b \Bigg(\int_{\varphi_{m-1}(x_m)}^{\psi_{m-1}(x_m)} \Bigg(\cdots \Bigg(\int_{\varphi_1(x_2,\ldots,x_m)}^{\psi_1(x_2,\ldots,x_m)} f(x_1, \ldots, x_m)dx_1 \Bigg) \cdots \Bigg) dx_{m-1} \Bigg) dx_m.$$

Beweis:
Wir wählen $I := \times_{i=1}^{m}[a_i, b_i]$ derart, daß $a_m := a$, $b_m := b$ und $a_{m-j} \leq \bar{\varphi}_{m-j}$, $\bar{\psi}_{m-j} \leq b_{m-j}$, $j = 1, \ldots, m-1$ (vgl. Beweis von 7.30).
Dann gilt $B \subseteq I$. Da B nach Lemma 7.30 Jordan-meßbar ist, und folglich f_B R-integrierbar über B, folgt nach Definition des R-Integrals und nach Satz 7.23 bzw. Folg. 7.24 sukzessive:

$$\int_B f(x)dx = \int_I f_B(x)dx = \int_{a_m}^{b_m} \Bigg(\int_{a_{m-1}}^{b_{m-1}} \Bigg(\cdots \Bigg(\int_{a_1}^{b_1} f_B(x_1, \ldots, x_m)dx_1 \Bigg) \cdots \Bigg) dx_{m-1} \Bigg) dx.$$

Ebenso sukzessive gilt aber nach Definition von B bzw. f_B:

$$\int_{a_1}^{b_1} f_B(x_1, \ldots, x_m)dx_1 = \int_{\psi_1(x_2,\ldots,x_m)}^{\psi_1(x_2,\ldots,x_m)} f_B(x_1, \ldots, x_m)dx_1,$$

$$\int_{a_2}^{b_2} (\int_{a_1}^{b_1} f_B(x_1, \ldots, x_m)dx_1)dx_2 = \int_{\varphi_2(\ldots)}^{\psi_2(\ldots)} (\int_{\varphi_1(\ldots)}^{\psi_1(\ldots)} f_B(x_1, \ldots, x_m)dx_1)dx_2 \quad \text{usw.}$$

wobei man für die Gleichheit mit dem Mittelwertsatz argumentiert. □

Beispiel 7.32 *(Volumen der Euklidischen Einheitskugel im $\mathbb{R}^m$, Teil I)*
Wir betrachten Euklidische Kugeln $B_m(0,r)$ um den Nullpunkt mit Radius r und speziell die Euklidische Einheitskugel $B_m = B_m(0,1)$ im $\mathbb{R}^m$, d.h.,

$$B_m = \{x \in \mathbb{R}^m : \|x\|_2 \leq 1\} = \{x \in \mathbb{R}^m : \sum_{i=1}^m x_i^2 \leq 1\}.$$

B_m ist eine Zylindermenge, denn sie kann wie folgt dargestellt werden:

$$B_m = \{x \in \mathbb{R}^m : -\sqrt{1 - \sum_{i=j+1}^m x_i^2} \leq x_j \leq \sqrt{1 - \sum_{i=j+1}^m x_i^2}, j = 1, \ldots, m\}.$$

Also ist B_m Jordan-meßbar und die Anwendung von Folgerung 7.28 möglich. Dazu setzen wir $p = 1$ und $q = m - 1$. Es gilt

$$\begin{aligned} \mu(B_m) &= \int_{-1}^1 \mu(\{z \in \mathbb{R}^{m-1} : (y,z) \in B_m\})dy = \int_{-1}^1 \mu(\{z \in \mathbb{R}^{m-1} : y^2 + \|z\|_2^2 \leq 1\})dy \\ &= \int_{-1}^1 \mu(B_{m-1}(0, \sqrt{1-y^2}))dy = \mu(B_{m-1}) \int_{-1}^1 (1-y^2)^{\frac{m-1}{2}} dy. \end{aligned}$$

Hierbei wurde die Formel $\mu(B_k(0,r)) = r^k \mu(B_k)$ verwendet. Diese ist eine Konsequenz aus einer allgemeinen Substitutionsregel für R*-Integrale (vgl. Kap. 205 in Heuser 2). Für $m = 1$ gilt $B_1 = [-1,1]$ und $\mu(B_1) = 2$, und für $m = 2$ nach obiger Formel $\mu(B_2) = 2\int_{-1}^1 (1-y^2)^{\frac{1}{2}} dy$. Auf die Berechnung des letzteren Integrals kommen wir in Kapitel 7.2 und speziell in Beispiel 7.42 zurück.*

7.2 Stammfunktion und Riemann-Integral

Wir wenden uns nun der eingangs von Kapitel 7 angekündigten Fragestellung der "Umkehrung" der Differentiation zu.

Definition 7.33
Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f, F : I \to \mathbb{R}$.
F heißt Stammfunktion von f (oder: unbestimmtes Integral) von f, falls F stetig und auf $\mathring{I}$ differenzierbar ist mit $F' = f$ auf $\mathring{I}$.
Bezeichnung: $F = \int f(x)dx$.

Bemerkung 7.34
Sei $I \subseteq \mathbb{R}$ ein Intervall, $f, F : I \to \mathbb{R}$ und F sei Stammfunktion von f.
Beh.: $F_0 : I \to \mathbb{R}$ ist Stammfunktion von f gdw. $\exists C \in \mathbb{R}$ ("Integrationskonstante") mit $F_0(x) = F(x) + C$, $\forall x \in I$,
Bew.: ($\leftarrow$) F_0 ist dann stetig und auf $\mathring{I}$ differenzierbar mit $F_0' = F' = f$
($\rightarrow$) $F - F_0$, ist stetig auf I und differenzierbar auf $\mathring{I}$ mit
$(F - F_0)' = f - f = 0$ auf $\mathring{I}$.
Deshalb gilt für fixiertes $x_0 \in I$ und beliebiges $x \in I$ nach Satz 6.10 (Mittelwertsatz): $\exists \tilde{x} \in]x_0, x[$ mit
$(F - F_0)(x) - (F - F_0)(x_0) = (F - F_0)'(\tilde{x})(x - x_o) = 0$.
$\rightsquigarrow$ Wir definieren $C := -(F - F_0)(x_0)$ und erhalten
$(F - F_0)(x) = -C$, $\forall x \in I$. $\square$

Stammfunktionen von f sind also bis auf eine additive Konstante eindeutig bestimmt. Sie sind demnach eindeutig, wenn man einen Funktionswert festlegt (z.B. $F(x_0) = 0$ für fixiertes $x_0 \in I$). Direkt aus der Definition folgt auch die Regel:
Sind $F, G : I \to \mathbb{R}$ Stammfunktionen von $f, g : I \to \mathbb{R}$, so ist $\alpha F + \beta G$ Stammfunktion von $\alpha f + \beta g$, $\forall \alpha, \beta \in \mathbb{R}$.
Wir besitzen zunächst keine Methode zur Konstruktion von Stammfunktionen, können aber natürlich in Beispielen durch Differentiation verifizieren.

Beispiel 7.35
(Alle folgenden Beispiele lassen sich durch Differentiation verifizieren; die jeweiligen Gültigkeitsintervalle I werden angegeben.)
a) $F = \int x^k dx \rightsquigarrow F(x) = \frac{x^{k+1}}{k+1}$, $\forall x \in I := \mathbb{R} \quad (\forall k \in \mathbb{N}_0)$
(vgl. Bsp. 6.7 a));
b) $F = \int \exp(x) dx \rightsquigarrow F(x) = \exp(x)$, $\forall x \in I = \mathbb{R}$ *(Bsp. 6.7 b));*

c)
$$\begin{array}{lll} F = \int \cos x \, dx & \rightsquigarrow F(x) = \sin x & \textit{Bsp. 6.7 b)} \\ F = \int \sin x \, dx & \rightsquigarrow F(x) = -\cos x \quad \forall x \in I = \mathbb{R} & \\ F = \int \cosh x \, dx & \rightsquigarrow F(x) = \sinh x & \\ F = \int \sinh x \, dx & \rightsquigarrow F(x) = -\cosh x & \end{array}$$

(Die beiden letzten Stammfunktionen ergeben sich sofort aus der Definition von sinh *und* cosh *in 3.51 und aus b))*

Wir wenden uns nun der Konstruktion von Stammfunktionen mit Hilfe des Riemann-Integrals zu. Es sei dazu $f : I \to \mathbb{R}$ beschränkt und R-integrierbar, $I \subseteq \mathbb{R}$ sei ein Intervall und $x_0 \in I$. Dann definieren wir

$$F_R : I \to \mathbb{R}, \; F_R(x) := \int_{x_0}^{x} f(t)dt, \quad \forall x \in I.$$

(Dabei setzen wir $\int_{x_0}^{x_0} f(t)dt := 0$ und $\int_{x_0}^{x} f(t)dt := -\int_{x}^{x_0} f(t)dt$, falls $x < x_0$.)
Klar ist, daß all diese Integrale korrekt definiert sind (vgl. Satz 7.11).

Satz 7.36
Es sei $f : I \to \mathbb{R}$ beschränkt und R-integrierbar, und F_R sei wie oben definiert.
Dann ist $F_R : I \to \mathbb{R}$ Lipschitzstetig und auf der Menge $D := \{x \in \overset{\circ}{I} : f \text{ ist stetig in } x\}$ differenzierbar mit $F_R'(x) = f(x)$, $\forall x \in D$.

Beweis:
Es seien $x, y \in I$ mit o.B.d.A. $x < y \rightsquigarrow |F_R(x) - F_R(y)| = |\int_{x_o}^{x} f(t)dt + \int_{x_o}^{y} f(t)dt|$.

1. Fall: $\quad x_o \le y : |F_R(x) - F_R(y)| = |\int_{x_o}^{x} f(t)dt - \int_{x_o}^{x} f(t)dt + \int_{x}^{y} f(t)dt|$

2. Fall: $x < x_o \le y : |F_R(x) - F_R(y)| = |-\int\limits_x^{x_o} f(t)dt - \int\limits_{x_o}^{y} f(t)dt|$

3. Fall $\quad y \le x_o : |F_R(x) - F_R(y)| = |-\int\limits_x^{x_o} f(t)dt - (-\int\limits_y^{x_o} f(t)dt)|$

$$= |\int\limits_y^{x_o} f(t)dt - \int\limits_x^{y} f(t)dt - \int\limits_y^{x_o} f(t)dt|$$

D.h. es gilt in allen Fällen

$$\begin{aligned} |F_R(x) - F_R(y)| &= |\int\limits_x^{y} f(t)dt| \le \int\limits_x^{y} |f(t)|dt \qquad &\text{(Satz 7.13 c)!)} \\ &\le |x-y| \sup_{t\in I} |f(t)| &\text{(Lemma 7.4).} \end{aligned}$$

Also ist F_R Lipschitzstetig (auf I).
Es sei nun $x \in D$ und $h \ne 0$ (o.B.d.A. $h > 0$) derart, daß $x + h \in I$. Dann gilt

$$|\frac{1}{h}(F_R(x+h) - F_R(x)) - f(x)| = |\frac{1}{h}\int\limits_x^{x+h} f(t)dt - f(x)| =$$

$$= \frac{1}{h}|\int\limits_x^{x+h} (f(t) - f(x))dt| \le \sup_{t\in[x,x+h]} |f(t) - f(x)| \underset{h\to 0}{\longrightarrow} 0$$

wegen der Stetigkeit von f in x. Für $h < 0$ läuft der Beweis analog. D.h. F_R ist differenzierbar in x mit der Ableitung $F'_R(x) = f(x)$. □

Satz 7.37 *(Hauptsatz der Differential- und Integralrechnung)*
Ist $f : [a, b] \mapsto \mathbb{R}$ und $F : [a, b] \to \mathbb{R}$ eine Stammfunktion von f auf $[a, b]$, so gilt:

$$\int\limits_a^b f(x)dx = F(b) - F(a).$$

<u>Beweis:</u>
Nach Satz 7.36 und Bemerkung 7.35 existiert eine Konstante $C \in \mathbb{R}$ mit
$F(x) = F_R(x) + C$, wobei $F_R(x) = \int\limits_a^x f(t)dt$, $\forall x \in [a, b]$.

$\rightsquigarrow F(b) - F(a) = F_R(b) - \underbrace{F_R(a)}_{=0} = \int\limits_a^b f(x)dx.$ □

Bemerkung 7.38
Eine andere Schreibweise für die Aussage des Hauptsatzes 7.37 ist:

$$\int\limits_a^b f(x)dx = F(x)|_a^b = \int f(x)dx|_a^b.$$

Dies rechtfertigt also die in Definition 7.34 eingeführte Bezeichnung für die Stammfunktion. Eine alternative Formulierung von 7.39 ist die folgende: Ist $f : [a,b] \to \mathbb{R}$ stetig und auf $]a,b[$ stetig differenzierbar, so gilt $\int\limits_a^b f'(x)dx = f(b) - f(a)$.

Satz 7.39
Es sei $f : [a,b] \to \mathbb{R}$ stetig.

a) *Es sei $g : [a,b] \to \mathbb{R}$ stetig und auf $]a,b[$ differenzierbar, F eine Stammfunktion von f und Fg' besitze eine Stammfunktion.*
Dann gilt die folgende Regel der partiellen Integration:

$$\int_a^b f(x)g(x)dx = F(b)g(b) - F(a)g(a) - \int\limits_a^b F(x)g'(x)dx$$

b) *Sei $g : [c,d] \to [a,b]$ stetig und auf $]c,d[$ differenzierbar; ferner gelte $g(]c,d[) =]a,b[$ und $g(c) = a,\ g(d) = b$.*
Dann gilt: $\int\limits_a^b f(x)dx = \int\limits_c^d f(g(t))g'(t)dt$ (Substitutionsregel)

Beweis:

a) Wir betrachten die Funktion $h := Fg - \int F(x)g'(x)dx$. Nach Definition und nach Voraussetzung ist h auf $[a,b]$ stetig und auf $]a,b[$ differenzierbar.
Es gilt: $h' = (Fg)' - Fg' = F'g = fg$ auf $]a,b[$.
Also ist h eine Stammfunktion von fg und es gilt nach Satz 7.39:

$$\begin{aligned}
\int\limits_a^b h'(x)dx &= h(b) - h(a) \quad \text{und damit} \\
\int\limits_a^b f(x)g(x)dx &= F(b)g(b) - F(a)g(a) - \int F(x)g'(x)dx\Big|_a^b \\
&= F(b)g(b) - F(a)g(a) - \int\limits_a^b F(x)g'(x)dx
\end{aligned}$$

b) Es sei $F : [a,b] \to \mathbb{R}$ eine Stammfunktion von f und wir betrachten die Funktion $F \circ g : [c,d] \to \mathbb{R}$. Diese ist als Zusammensetzung stetiger Funktionen stetig und nach Kettenregel (Satz 6.36) auf $]c,d[$ differenzierbar mit

$$\begin{aligned}
(F \circ g)'(t) &= F'(g(t))g'(t),\ \forall t \in]c,d[, \\
&= (f \circ g)(t)g'(t),\ \forall t \in]c,d[.
\end{aligned}$$

also ist $F \circ g$ eine Stammfunktion von $(f \circ g)g'$ und es folgt aus Satz 7.39:

$$\int_c^d f(g(t))g'(t)dt = (F \circ g)(d) - (F \circ g)(c) = F(b) - F(a) = \int_a^b f(x)dx. \quad \square$$

Die Regeln in Satz 7.39 können hilfreich zur Berechnung komplizierterer Integrale eingesetzt werden.

Beispiel 7.40

(a) $\int\limits_0^{2\pi} \cos x \sin x \, dx = (\sin x)^2|_0^{2\pi} - \int\limits_0^{2\pi} \sin x \cos x \, dx \quad \rightsquigarrow \int\limits_0^{2\pi} \cos x \sin x \, dx = 0$

(b) $\int\limits_a^b \exp(x)x^k dx = \exp(x)x^k|_a^b - k \int\limits_a^b \exp(x)x^{k-1}dx = \sum\limits_{j=0}^k (-1)^{k-j} \prod\limits_{i=k-j+1}^k i \exp(x)x^{k-j}|_a^b.$

(c) $\int\limits_{-1}^1 (1-x^2)^{\frac{m-1}{2}} dx = -\int\limits_0^\pi (\sin^2 t)^{\frac{m-1}{2}}(-\sin t)dt = \int\limits_0^\pi (\sin t)^m dt$
(nach Substitution $x = \cos t$).

(d) Sei $g : [a,b] \to]0,+\infty[$ stetig und auf $]a,b[$ stetig differenzierbar. Dann folgt mit der Funktion $f(x) := x^{-1}$, $x > 0$, aus 7.41 b):

$$\int_a^b \frac{g'(t)}{g(t)} dt = \int_a^b f(g(t))g'(t)dt = \int_{g(a)}^{g(b)} f(x)dx = \ln g(x)|_a^b.$$

Beispiel 7.41 *(Volumen der Euklidischen Einheitskugel im $\mathbb{R}^m$, Teil 2)*
Der Jordan-Inhalt der Euklidischen Einheitskugel $B_m := \{x \in \mathbb{R}^m : \|x\|_2 \le 1\}$ gilt

$$\mu(B_m) = \frac{\pi^{\frac{m}{2}}}{\Gamma(1 + \frac{m}{2})}.$$

wobei $\Gamma :]0,+\infty[\mapsto \mathbb{R}$ die sog. Gamma-Funktion ist, die die folgenden Eigenschaften besitzt: $\Gamma(1) = 1$, $\Gamma(x+1) = x\Gamma(x)$, $\forall x > 0$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.
(Insbesondere gilt: $\mu(B_1) = 2$, $\mu(B_2) = \pi$, $\mu(B_3) = \frac{4}{3}\pi$).
Beweis:
Nach Beispiel 7.32 und Beispiel 7.40 c) gilt die Rekursionsformel

$$\mu(B_m) = \mu(B_{m-1}) \int_0^\pi (\sin t)^m dt = \mu(B_{m-1})\sqrt{\pi}\frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2}+1)}$$

Dabei ist die letzte Identität in Heuser, Bd. 1, Kap. 94 bzw. Bd. 2, Kap. 150 bewiesen. Wir beweisen nun die Aussage mit Induktion über m:

$$m = 1 : \mu(B_1) = 2 = \frac{\sqrt{\pi}}{\frac{1}{2}\sqrt{\pi}} = \frac{\sqrt{\pi}}{\frac{1}{2}\Gamma(\frac{1}{2})} \rightsquigarrow \quad \textit{richtig für } m = 1.$$

Die Aussage sei nun für $m-1$ richtig und wir betrachten

$$\mu(B_m) = \sqrt{\pi}\frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2}+1)}\mu(B_{m-1}) = \sqrt{\pi}\pi^{\frac{m-1}{2}}\frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2}+1)\Gamma(1+\frac{m-1}{2})} = \frac{\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2}+1)}$$

Damit ist die Aussage vollständig bewiesen.
Die Spezialfälle $m=2$ bzw. $m=3$ ergeben sich wie folgt:

$$m=2: \qquad \mu(B_2) = \frac{\pi}{\Gamma(2)} = \pi$$

$$m=3: \qquad \mu(B_3) = \frac{\pi^{\frac{3}{2}}}{\Gamma(1+\frac{3}{2})} = \frac{\pi^{\frac{3}{2}}}{\frac{3}{2}\Gamma(1+\frac{1}{2})} = \frac{\pi^{\frac{3}{2}}}{\frac{3}{2}\frac{1}{2}\Gamma(\frac{1}{2})} = \frac{4}{3}\pi \qquad \square$$

Aus der eben bewiesenen Formel für den Jordan-Inhalt der Euklidischen Einheitskugel im $\mathbb{R}^m$ ergibt sich das folgende "Inhaltsparadoxon":

$$\mu(B_m) \underset{m\to\infty}{\longrightarrow} 0$$

d.h. der Jordan-Inhalt der Einheitskugel ist eine Nullfolge mit wachsender Dimension (für kleine m wächst er jedoch zunächst!).
Man sieht dies schnell ein für gerades $m \in I\!N$, d.h. $m = 2p$ bzw. $p = \frac{m}{2}$:

$$\rightsquigarrow \mu(B_m) = \frac{\pi^p}{\Gamma(1+p)} = \frac{\pi^p}{p!} \underset{m\to\infty}{\longrightarrow} 0!$$

da dies ja genau die Glieder der (konvergenten) Exponentialreihe von $\exp(\pi)$ *sind.*

7.3 Trigonometrische Fourierreihen

Wir betrachten die Menge $X = C([0,1])$ aller stetigen reellwertigen Funktionen auf dem Intervall $[0,1]$. Mit den üblichen Operationen (vgl. Beispiel 4.4) wird X zu einem linearen Raum über dem Körper $\mathbb{R}$. Wir definieren nun ein Skalarprodukt in X durch

$$\langle x,y\rangle := \int_0^1 x(t)y(t)dt \quad (\forall x,y \in X).$$

Die Eigenschaften eines Skalarproduktes (vgl. Definition 4.25)

$$\begin{aligned}
\langle \alpha x + \beta\tilde{x}, y\rangle &= \alpha\langle x,y\rangle + \beta\langle \tilde{x},y\rangle \quad (\forall \alpha,\beta \in \mathbb{R}, x,\tilde{x},y \in X),\\
\langle x,y\rangle &= \langle y,x\rangle \quad (\forall x,y \in X),\\
\langle x,x\rangle &> 0 \quad \forall x \in X, x \neq 0,
\end{aligned}$$

sind erfüllt. Hierbei folgt die dritte Eigenschaft aus den Sätzen 7.13b) und 7.16. Damit wird $(X, \langle\cdot,\cdot\rangle)$ zu einem unitären Raum, in dem durch $\|x\| = \langle x,x\rangle^{\frac{1}{2}}$ $(\forall x \in X)$ eine Norm definiert ist.
Wir betrachten die folgende Familie $(e_k)_{k\in I\!N_0}$ von Funktionen in X:

$$e_0(t) := 1, \;\; e_{2k-1}(t) := \sqrt{2}\cos(2k\pi t), \;\; e_{2k}(t) := \sqrt{2}\sin(2k\pi t), \;\; \forall k \in I\!N, t \in [0,1].$$

Lemma 7.42 *Die Familie $(e_k)_{k\in I\!N_0}$ von Funktionen in X ist ein Orthonormalsystem in $(X, \langle\cdot,\cdot\rangle)$.*

Beweis: Zu zeigen ist $\langle e_k, e_m\rangle = 0$ für alle $k, m \in I\!N_0$, $k \neq m$, und $\|e_k\|^2 = 1$ für jedes $k \in I\!N_0$.
Wir werden nur in einigen der Fälle ausführlich argumentieren. Es gilt für alle $k \in I\!N$:

$$\begin{aligned}
\|e_0\|^2 &= \int_0^1 dt = 1, \\
\langle e_0, e_{2k}\rangle &= \sqrt{2}\int_0^1 \sin(2k\pi t)dt = -\frac{\sqrt{2}}{2k\pi}\cos(2k\pi t)\Big|_0^1 = -\frac{\sqrt{2}}{2k\pi}(1-1) = 0, \\
\|e_{2k}\|^2 &= 2\int_0^1 (\sin(2k\pi t))^2 dt = -2\frac{\sin(4k\pi t) - 4k\pi t}{8k\pi t}\Big|_0^1 = -2(-\frac{1}{2}) = 1, \\
\langle e_{2k-1}, e_{2k}\rangle &= 2\int_0^1 \cos(2k\pi t)\sin(2k\pi t)dt = 2\frac{(\sin(2k\pi t))^2}{4k\pi t}|Big_0^1 = 0.
\end{aligned}$$

Bei den letzten beiden Integralen (die natürlich mit Hilfe der Regel der partiellen Integration (Satz 7.39) berechnet werden können, vgl. Beispiel 7.40) wurden die Tafel der Grundintegrale in Heuser, Analysis I, Kap. 76, verwendet. Die restlichen Orthonormalitäts-Beziehungen beweist man analog. □

Entsprechend Definition 4.34 in Kap. 4.1 kann man also die Fourierreihe

$$\sum_{k=0}^{\infty}\langle x, e_k\rangle e_k = \langle x, e_0\rangle + \sum_{k=1}^{\infty}(\langle x, e_{2k-1}\rangle e_{2k-1} + \langle x, e_{2k}\rangle e_{2k})$$

von $x \in X$ betrachten. Diese spezielle Fourierreihe heißt auch trigonometrische Fourierreihe und die Koeffizienten $\langle x, e_k\rangle$ trigonometrische Fourierkoeffizienten. Wann und in welchem Sinn konvergiert diese Fourierreihe gegen x ?

Satz 7.43 *Für jedes $x \in X = C([0,1])$ konvergiert die trigonometrische Fourierreihe von x gegen x im Sinne der obigen Norm $\|\cdot\|$ in X, d.h.*

$$\lim_{n\to\infty}\int_0^1 \Big(x - \Big(\langle x, e_0\rangle + \sum_{k=1}^{n}(\langle x, e_{2k-1}\rangle\cos(2k\pi t) + \langle x, e_{2k}\rangle\sin(2k\pi t))\Big)\Big)^2 dt = 0$$

("Konvergenz im quadratischen Mittel").

Beweisidee: Das Orthonormalsystem $(e_k)_{k\in I\!N_0}$ ist vollständig, da der kleinste abgeschlossene Unterraum U von X, der das Orthonormalsystem enthält, gerade $U = X$ ist. Deshalb folgt die Aussage aus Satz 4.35. □

Aus der Konvergenz einer Folge von Funktionen im quadratischen Mittel folgt jedoch nicht die punktweise Konvergenz der Folge (in X), wie das folgende Beispiel zeigt.
Beispiel: Wir betrachten die folgende Folge (x_n) von Funktionen in X:

$$x_n(t) := \begin{cases} nt & , t \in [0, \frac{1}{n}] \\ 1 & , t \in]\frac{1}{n}, 1]. \end{cases}$$

Für die Funktion $x(t) := 1$, $\forall t \in [0,1]$ gilt nun

$$\|x_n - x\|^2 = \int_0^1 (x_n(t) - x(t))^2 dt = \int_0^{\frac{1}{n}} (nt-1)^2 dt = \frac{1}{3n}(nt-1)^3\Big|_0^{\frac{1}{n}} = \frac{1}{3n}.$$

Also konvergiert (x_n) in X gegen x, aber $(x_n(0))$ konvergiert gegen 0.

Die folgende Aussage klärt nun, für welche Funktionen x die trigonometrische Fourierreihe von x auch punktweise konvergiert, und was in möglichen Unstetigkeitsstellen von x passiert.

Satz 7.44 *Ist $x \in X$ periodisch (d.h. $x(0) = x(1)$) und Lipschitzstetig, so konvergiert die trigonometrische Fourierreihe für jedes $t \in [0,1]$ punktweise gegen $x(t)$.*
Ist x eine periodische, stückweise Lipschitzstetige reellwertige Funktion auf $[0,1]$ (mit endlich vielen Unstetigkeitspunkten), so konvergiert die trigonometrische Fourierreihe für jedes $t \in [0,1]$ punktweise gegen $\frac{1}{2}(x(t+) + x(t-))$, wobei $x(t+) := \lim_{\tau \to t+} x(\tau)$ und $x(t-) := \lim_{\tau \to t-} x(\tau)$ die rechts- bzw. linksseitigen Grenzwerte von x in t sind.

(vgl. Heuser, Analysis II, Kapitel 136)

7.4 Uneigentliche Integrale

Bei der Definition des Riemann-Integrals in Kapitel 7.1 spielten zwei Beschränktheitsvoraussetzungen eine ganz wesentliche Rolle: der Integrationsbereich und der Integrand mußten beschränkt sein. Für eine ganze Reihe von Anwendungen reicht das jedoch nicht aus. Im folgenden soll deshalb eine Erweiterung der Integralrechnung (im eindimensionalen Fall) vorgenommen werden, die auch (gewisse) unbeschränkte Integranden und Bereiche zuläßt. Es liegt nahe, diese Erweiterung mittels Grenzübergängen (von Folgen nach Kapitel 7.1 definierter Integrale) vorzunehmen. Entscheidender Begriff ist dabei das sog. uneigentliche Integral.

Definition 7.45
Es seien $a \in \mathbb{R}$, $b \in]a, +\infty]$ und $f : [a, b[\to \mathbb{R}$.
Für alle $x \in [a, b[$ existiere das Integral $\int_a^x f(t)dt$, nicht jedoch für $x = b$.

Dann heißt $\int_a^b f(t)dt$ (bei b) uneigentliches Integral.

Das (bei b) uneigentliche Integral $\int_a^b f(t)dt$ heißt konvergent, wenn $\lim_{x \to b-} \int_a^x f(t)dt$ existiert; ansonsten heißt es divergent.
Im ersten Fall heißt $\int_a^b f(t)dt := \lim_{x \to b-} \int_a^x f(t)dt$ Wert des uneigentlichen Integrals.
Analog definiert man für $b \in \mathbb{R}$, $a \in [-\infty, b[$, $f :]a, b] \to \mathbb{R}$ ein (bei a) uneigentliches Integral und

$$\int_a^b f(t)dt = \lim_{x \to a+} \int_x^b f(t)dt, \quad \text{falls der Grenzwert existiert.}$$

Sind $a, b \in [-\infty, +\infty]$, $a < c < b$ und $\int_a^c f(t)dt$ bzw. $\int_c^b f(t)dt$ bei a bzw. bei b uneigentliche Integrale, so heißt $\int_a^b f(t)dt$ beidseitig uneigentliches Integral.

Falls die folgenden Grenzwerte existieren, heißt es konvergent und sein Wert ist

$$\int\limits_a^b f(t)dt := \lim_{\substack{x \to a+ \\ x \leq c}} \int\limits_x^c f(t)dt + \lim_{\substack{x \to b- \\ x \geq c}} \int\limits_c^x f(t)dt.$$

Beispiel 7.46

a) *Das (bei $+\infty$) uneigentliche Integral $\int\limits_1^{+\infty} t^{-\alpha}dt$ (mit $\alpha \geq 1$) ist konvergent (mit Wert $\frac{1}{\alpha-1}$), falls $\alpha > 1$.*
Beweis:

$$\forall x > 1 \text{ gilt: } \int\limits_1^x t^{-\alpha}dt = \begin{cases} \frac{t^{1-\alpha}}{1-\alpha}\Big|_1^x = \frac{1}{1-\alpha}(x^{1-\alpha} - 1) & ,\alpha \neq 1 \\ \ln t|_1^x = \ln x & ,\alpha = 1 \end{cases}$$

$\rightsquigarrow \lim\limits_{x \to +\infty} \int\limits_1^x t^{-\alpha}dt = \frac{1}{\alpha-1}$ *falls* $\alpha > 1$. □

b) *Das (bei 0) uneigentliche Integral $\int\limits_0^1 t^{-\alpha}dt$ (für $\alpha > 0$) ist konvergent (mit Wert $\frac{1}{1-\alpha}$) falls $\alpha < 1$.*
Beweis:
folgt analog zu a) wegen $\int\limits_x^1 t^{-\alpha}dt = \begin{cases} \frac{1}{1-\alpha}(1 - x^{1-\alpha}) & ,\alpha \neq 1 \\ -\ln x & ,\alpha = 1 \end{cases}$
$\rightsquigarrow \exists \lim\limits_{x \to 0+} \int_x^1 t^{-\alpha}dt = \frac{1}{1-\alpha}$ *gdw.* $\alpha \in]0,1[$. □

c) *Das beidseitig uneigentliche Integral $\int\limits_{-\infty}^{\infty} \sin t \, dt$ ist nicht konvergent.*

Mit Hilfe uneigentlicher Integrale wird die sog. Gamma-Funktion definiert, die bereits in Beispiel 7.41 verwendet wurde.

Satz 7.47
Das uneigentliche Integral $\Gamma(x) := \int\limits_0^{\infty} t^{x-1}\exp(-t)dt$ ist $\forall x > 0$ konvergent.
Für die so definierte Gamma-Funktion $\Gamma :]0, +\infty[\to \mathbb{R}$ gilt die Funktionalgleichung $\Gamma(x+1) = x\Gamma(x)$, $\forall x > 0$, sowie $\Gamma(1) = 1$ und $\Gamma(\frac{1}{2}) = \int\limits_{-\infty}^{\infty} \exp(-t^2)dt$.

(vgl. Heuser, Analysis II, Kapitel 150)

8 Gewöhnliche Differentialgleichungen

8.1 Aufgabenstellung und Beispiele

Differentialgleichung (DGL): Gleichung für eine gesuchte Funktion, wobei in der Gleichung die Funktion und einige ihrer Ableitungen auftreten.
Man unterscheidet gewöhnliche DGLen (die gesuchte Funktion hängt von einer reellen Variablen ab) und partielle DGLen (die gesuchte Funktion hängt von mehreren reellen Variablen ab, und es treten partielle Ableitungen in der Gleichung auf).

Allgemeine Form einer gewöhnlichen DGL:

$$F(x^{(k)}(t), x^{(k-1)}(t), \ldots, x'(t), x(t), t) = 0, \quad \forall t \in I, \tag{$*$}$$

wobei $F : \underbrace{\mathbb{R}^m \times \mathbb{R}^m \times \ldots \times \mathbb{R}^m \times \mathbb{R}^m}_{k+1} \times I \to \mathbb{R}^m$ und $I \subseteq \mathbb{R}$ ein Intervall ist.

Gesucht:
Eine Funktion $x : I \to \mathbb{R}^m$, die k-mal differenzierbar auf I ist und $(*)$ erfüllt.

Bemerkung 8.1
Die Gleichung $(*)$ kann stets in eine DGL äquivalent umgeformt werden, die nur Ableitungen erster Ordnung enthält.
Idee: *Man führt neue Variable durch $x_i(t) := x^{(i-1)}(t)$, $i = 1, \ldots, k$, $t \in I$, ein und betrachtet zusätzlich die folgenden Gleichungen:*
$x_i'(t) = x_{i+1}(t)$, $i = 1, \ldots, k-1$, $\forall t \in I$.

$$\rightsquigarrow \begin{array}{rcll} F(x_k'(t), x_k(t), x_{k-1}(t), \ldots, x_1(t), t) &=& 0 & \text{entsteht aus } (*) \\ x_{k-1}'(t) - x_k(t) &=& 0 & \\ &\vdots& & \\ x_1'(t) - x_2(t) &=& 0, & \forall t \in I. \end{array} \tag{$+$}$$

Dies ist eine DGL für die Funktion $(x_1(\cdot), \ldots, x_k(\cdot)) : I \to \mathbb{R}^{mk}$. Die Äquivalenz von $(*)$ und $(+)$ bedeutet dann folgendes: Für jede Lösung x von $(*)$ ist $(x(\cdot), x'(\cdot), \ldots, x^{(k-1)}(\cdot))$ Lösung von $(+)$, und für jede Lösung $(x_1(\cdot), \ldots, x_k(\cdot))$ von $(+)$ ist $x(\cdot) := x_1(\cdot)$ Lösung von $(*)$. Allerdings erhöht sich beim Übergang von $(*)$ zu $(+)$ die Dimension des Wertebereichs der gesuchten Funktion von m auf mk.
$\rightsquigarrow$ *Standardform für gewöhnliche DGLen:*

$$F(x'(t), x(t), t) = 0, \quad \forall t \in I \tag{2.0}$$

Definition 8.2
(2.0) heißt DGL (bzw. DGL-System, falls $m, n > 1$) 1. Ordnung.
(2.0) heißt explizit, wenn man (2.0) in der Form

$$x'(t) = f(x(t), t), \quad \forall t \in I,$$

schreiben kann, wobei $f : \mathbb{R}^m \times I \to \mathbb{R}^m$. In allen anderen Fällen heißt (2.0) implizit.
(2.0) heißt linear, wenn $F(y, x, t) = A(t)y + B(t)x + c(t)$, $\forall x, y \in \mathbb{R}^m$, $\forall t \in I$, wobei $A(t), B(t) \in \mathbb{R}^{m \times m}$ und $c(t) \in \mathbb{R}^m$, $\forall t \in I$, anderenfalls nichtlinear.

Wir befassen uns in diesem Kapitel mit expliziten gewöhnlichen DGLen 1. Ordnung, d.h. mit Gleichungen der Form

$$x'(t) = f(x(t), t), \quad \forall t \in I, \tag{2.1}$$

wobei $f : \mathbb{R}^m \times I \to \mathbb{R}^m$.

Definition 8.3
Eine Funktion $x : I \to \mathbb{R}^m$ heißt Lösung von (2.1), falls x differenzierbar auf I ist (d.h. einseitige Ableitungen in eventuellen Randpunkten sollen auch existieren), und die Gleichung (2.1) erfüllt.

Bemerkung 8.4
Mit Definition 8.3 kann keine Eindeutigkeit einer Lösung von (2.1) erwartet werden (z. B. $f \equiv 0$, d.h. unendlich viele Lösungen!). Deshalb benötigt man weitere Bedingungen an x, um eine eindeutige Lösung zu erhalten. Diese Bedingungen ergeben sich meist bei der Modellierung realer Prozesse "kanonisch", z. B.

- *Anfangsbedingung: $x(t_0) = x_0, t_0 \in I, x_0 \in \mathbb{R}^m$ ("Anfangswert").*
- *Randbedingung (auf $I := [a, b]$): $r(x(a), x(b)) = 0$, wobei $r : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$;*
  Bsp.: $x(a) = x(b)$ ("periodische Lösung").

Beispiel 8.5 *(Populationsdynamik, vgl. Kap. 0, M. Braun, Springer 1991, Kap. 1.5)*
Es beschreibe $p : [0, +\infty[\to \mathbb{R}$ den zeitlichen Verlauf der Population einer gewissen biologischen Art. Die Funktion p genügt der DGL:

$$p'(t) = ap(t) - bp(t)^2, \quad \forall t \in I = [0, +\infty[,$$

wobei $p'(t)$ die Populationsgeschwindigkeit ist und die Gleichung bedeutet, daß $p'(t)$ proportional zur Population verringert durch den Konflikt-/Konkurrenzterm $\frac{b}{a}p(t)^2$ ist. Die Proportionalitätskonstante ist gerade a, wobei in der Regel $a \gg b > 0$ gilt.
Später wird gezeigt, daß diese Gleichung der Populationsdynamik genau eine Lösung bei gegebenem Anfangswert besitzt. Dazu wird aber ein theoretischer Vorlauf benötigt.
Lösung zu Anfangswert $p(t_0)$:

$$p(t) = \frac{ap(t_0)}{bp(t_0) + (a - bp(t_0)) \exp(-a(t - t_0))}, \quad \forall t \in I, t_0 \in I \text{ fest.}$$

Man rechnet leicht nach, daß $p(\cdot)$ Lösung ist.
Aus der Lösungsgestalt folgt: $\lim\limits_{t\to\infty} p(t) = \frac{a}{b}$ (Limes existiert).
Weiterhin ist $p(\cdot)$ monoton wachsend, da $p'(t) \geq 0$, $\forall t \in I$. Insbesondere gilt deshalb $p(0) \leq p(t) \leq \frac{a}{b}$, $\forall t \in I$. Darüberhinaus gilt

$$\begin{aligned} p''(t) &= ap'(t) - 2bp(t)p'(t) = (a - 2bp(t))p'(t) \\ &= (a - 2bp(t))p(t)(a - bp(t)) \end{aligned}$$

$\rightsquigarrow p''(t) = 0$ gdw. $p(t) = \frac{a}{2b}$.
Für die menschliche Population ist $a = 0.029$, $b = 2.941 \cdot 10^{-12}$ und $\frac{a}{b} = 9.86 \cdot 10^9$.

8.2 Anfangswertaufgaben für lineare DGL-Systeme

Wir betrachten die Anfangswertaufgabe für eine lineare DGL

$$x'(t) = A(t)x(t) + b(t),\ \forall t \in I = [t_0, T],\ x(t_0) = x_0, \tag{2.2}$$

wobei $A : I \to \mathbb{R}^{m\times m}$, $b : I \to \mathbb{R}^m$, $x_0 \in \mathbb{R}^m$.
Die Ergebnisse von Kapitel 8.3 werden insbesondere implizieren, daß lineare DGL-Systeme der Form (2.2) eine eindeutig bestimmte Lösung besitzen.
Ziel: Berechnung der eindeutig bestimmten Lösung von (2.2).
Wir betrachten dazu zunächst die sog. "homogene" DGL

$$x'(t) = A(t)x(t), \quad \forall t \in [t_0, t_0 + T], \tag{2.3}$$

wobei die Matrixfunktion $A : I \to \mathbb{R}^{m\times m}$ als stetig vorausgesetzt wird.

Lemma 8.6
Es sei $A : I \to \mathbb{R}^{m\times m}$ stetig.
Die Menge $\mathcal{L} := \{x : I \to \mathbb{R}^m : x$ ist differenzierbar und Lösung von (2.3)$\}$ ist ein linearer Raum. Außer der Nullfunktion $x = \Theta$ verschwindet kein Element von $\mathcal{L}$ in einem Punkt von I.

Beweis:
Es seien $x, \tilde{x}$ Lösungen von (2.3). Dann gilt für beliebige $\alpha, \beta \in \mathbb{R}$:

$$(\alpha x + \beta\tilde{x})' = \alpha x' + \beta\tilde{x}' = \alpha A(\cdot)x + \beta A(\cdot)\tilde{x} = A(\cdot)(\alpha x + \beta\tilde{x})$$

Also ist $\mathcal{L}$ ein linearer Raum.
Annahme: $\exists \tilde{x} \in \mathcal{L}$, $\tilde{x} \neq \Theta$, $\exists \tilde{t} \in I : \tilde{x}(\tilde{t}) = 0$.
Nach Folgerung 8.19 hat die Aufgabe $x'(t) = A(t)x(t)$, $\forall t \in I$, $x(\tilde{t}) = 0$, genau eine Lösung (obwohl $\tilde{t}$ i.a. nicht linker Randpunkt von I ist, liefern die allgemeinen Resultate die Existenz und Eindeutigkeit). Offenbar ist aber auch die Nullfunktion $x = \Theta$ eine Lösung, und folglich müßte $\tilde{x} = \Theta$ gelten. $\rightsquigarrow$ Widerspruch. $\square$

Definition 8.7
Es sei $A : I \to \mathbb{R}^m$ stetig und $\mathcal{L}$ sei wie in 8.20 definiert.
Ein System $\varphi_1, \ldots, \varphi_m \in \mathcal{L}$ heißt Fundamentalsystem von (2.3), falls für alle $t \in I$, $\varphi_1(t), \ldots, \varphi_m(t) \in \mathbb{R}^m$ linear unabhängig sind.

Lemma 8.8
Es sei $A : I \to \mathbb{R}^m$ stetig. Dann gelten die folgenden Aussagen:

a) *Es existiert ein Fundamentalsystem $\varphi_1, \ldots, \varphi_m$ von (2.3) mit*

$$\varphi_j(t_0) = e_j = (0, \ldots, 0, 1, 0, \ldots, 0), \quad \forall j = 1, \ldots, m.$$

b) *Wir definieren für jedes $t \in I$ eine Matrix $Q(t) \in \mathbb{R}^{m\times m}$ durch*
$Q(t)y := \sum\limits_{j=1}^{m} y_j\varphi_j(t), \quad \forall y = (y_1,\dots,y_m) \in \mathbb{R}^m,$
wobei $\varphi_1,\dots,\varphi_m$ das Fundamentalsystem aus a) ist.
Die Matrixfunktion $Q(\cdot)$ hat folgende Eigenschaften:
(i) $Q(t_0)y = y, \quad \forall y \in \mathbb{R}^m$;
(ii) $Q(t) \in \mathbb{R}^{m\times m}$ ist regulär $\forall t \in I$;
(iii) für alle $x_0 \in \mathbb{R}^m$ ist $Q(\cdot)x_0$ Lösung von (2.3) mit Anfangswert x_0.

Beweis:

a) Nach Folgerung 8.19 existieren für alle $j = 1,\dots,m$ eindeutig bestimmte Lösungen φ_j der Aufgabe

$$x'(t) = A(t)x(t), \forall t \in I, x(t_0) = e_j.$$

Zu zeigen: $\varphi_1,\dots,\varphi_m \in \mathcal{L}$ ist ein Fundamentalsystem.
Annahme: $\exists \tilde{t} \in I$, so daß $\varphi_1(\tilde{t}),\dots,\varphi_m(\tilde{t})$ linear abhängig sind.
$\rightsquigarrow \exists c = (c_1,\dots,c_m) \neq 0 : \sum\limits_{i=1}^{m} c_i\varphi_i(\tilde{t}) = 0.$
Wir betrachten $\tilde{x}(t) := \sum\limits_{i=1}^{m} c_i\varphi_i(t), \forall (t) \in I. \rightsquigarrow \tilde{x} \in \mathcal{L}, \tilde{x}(\tilde{t}) = 0 \rightsquigarrow \tilde{x} = \Theta.$
$\rightsquigarrow \tilde{x}(t_0) = 0 \rightsquigarrow \sum\limits_{i=1}^{m} c_ie_i = 0 \rightsquigarrow$ Widerspruch!

b) Die $\varphi_1(t), \varphi_2(t),\dots,\varphi_m(t)$ sind nach Definition der Matrix $Q(t)$ gerade ihre Spalten. Deshalb gilt:
(i) $Q(t_0)y = \sum\limits_{j=1}^{m} y_je_j = y, \forall y \in \mathbb{R}^m$;
(ii) Die Spalten $\varphi_1(t),\dots,\varphi_m(t)$ sind für jedes $t \in I$ linear unabhängig.
$\rightsquigarrow Q(t)$ ist regulär für alle $t \in I$;
(iii) Sei $x_0 \in \mathbb{R}^m$. Dann folgt für $x(\cdot) := Q(\cdot)x_0 = \sum\limits_{j=1}^{m} x_{0j}\varphi_j(\cdot)$ aus (iii), daß $x(\cdot) \in \mathcal{L}$ und $x(t_0) = Q(t_0)x_0 = x_0$. □

Satz 8.9
Es seien $A : I \to \mathbb{R}^m$ und $b : I \to \mathbb{R}^m$ stetig, $x_0 \in \mathbb{R}^m$ und $Q(t), t \in I$, wie in 8.22 definiert. Dann ist die Funktion $x : I \to \mathbb{R}^m$ definiert durch

$$x(t) = Q(t)\left[x_0 + \int\limits_{t_0}^{t} Q^{-1}(s)b(s)ds\right], \ \forall t \in I,$$

die eindeutig bestimmte Lösung von (2.2).

Beweis:
Die inverse Matrix $Q^{-1}(t)$ existiert für jedes $t \in I$ nach Lemma 9.22 b). Aus der Stetigkeit von $Q(\cdot)$ und der Abschätzung

$$\|Q^{-1}(t) - Q^{-1}(s)\| = \|Q^{-1}(t)(Q(s) - Q(t))Q^{-1}(s)\| \leq \max_{t\in I} \|Q^{-1}(t)\|^2\|Q(s) - Q(t)\|$$

mit einer Matrixnorm $\|\cdot\|$ in $\mathbb{R}^{m\times m}$, folgt die Stetigkeit von $Q^{-1}(\cdot)$ als Abbildung von I in $\mathbb{R}^{m\times m}$. Deshalb ist das Integral $\int\limits_{t_0}^{t} Q^{-1}(s)b(s)ds$ als komponentenweises Integral über stetige Integranden wohl-definiert.
Es genügt zu zeigen: $x(\cdot)$ ist differenzierbar und Lösung von (2.2).
Die Eindeutigkeit dieser Lösung ist klar nach Folgerung 8.19.
Wir definieren $a(t) := x_0 + \int\limits_{t_0}^{t} Q^{-1}(s)b(s)ds, \quad \forall t \in I.$
Nach Satz 7.39 gilt $\frac{d}{dt}\int\limits_{t_0}^{t} Q^{-1}(s)b(s)ds = Q^{-1}(t)b(t).$
Folglich ist $x(\cdot)$ differenzierbar und es folgt aus 9.22, daß
$x'(t) = Q(t)a'(t) + A(t)Q(t)a(t) = Q(t)Q^{-1}(t)b(t) + A(t)x(t)$ und
$x(t_0) = Q(t_0)x_0 = x_0$, d.h. $x(\cdot)$ ist Lösung von (2.2). □

Beispiel 8.10 *(skalare lineare DGL 1. Ordnung)*
Es seien $I = [t_0, t_0+T]$ ein Intervall und $a, b : I \to \mathbb{R}$ stetig. Wir betrachten die skalare lineare DGL 1. Ordnung

$$x'(t) = a(t)x(t) + b(t), \quad \forall t \in I,\ x(t_0) = x_0.$$

Lösung der homogenen DGL: Ansatz: $\varphi(t) = \exp\left(\int\limits_{t_0}^{t} a(s)ds\right), \forall t \in I.$

$\rightsquigarrow \varphi'(t) = a(t)\exp\left(\int\limits_{t_0}^{t} a(s)ds\right) = a(t)\varphi(t),\ \varphi(t_0) = 1.$
Dann folgt aus Satz 8.23, daß

$$\begin{aligned} x(t) &= \varphi(t)\left[x_0 + \int\limits_{t_0}^{t} \frac{1}{\varphi(s)}b(s)ds\right] \\ &= \exp\left(\int\limits_{t_0}^{t} a(s)ds\right)\left[x_0 + \int\limits_{t_0}^{t} \exp\left(-\int\limits_{t_0}^{s} a(u)du\right)b(s)ds\right], \forall t \in I. \end{aligned}$$

einzige Lösung dieser DGL ist.

Als nächstes beschäftigen wir uns mit dem wichtigen Spezialfall von (2.2), daß $A(t) \equiv A \in \mathbb{R}^{m\times m}$, d.h. einem linearen DGL-System mit konstanten Koeffizienten. In diesem Fall ist es nun möglich, das Fundamentalsystem der homogenen DGL explizit anzugeben. Der Schlüssel dazu ist die Jordansche Normalform der Matrix A (vgl. Satz 4.74).

Satz 8.11
Es seien $\lambda_1, \ldots, \lambda_p$ die Eigenwerte der Matrix A, und η_i sei die Größe des größten zu λ_i gehörenden Jordan-Kästchens, $i = 1, \ldots, p$, wobei $\lambda_1, \bar\lambda_1, \ldots, \lambda_s, \bar\lambda_s$ komplexe und $\lambda_{2s+1}, \ldots, \lambda_p$ reelle Eigenwerte von A sind.
Die einzige Lösung von $x'(t) = Ax(t) + b(t)$, $\forall t \in I = [0, T]$, $x(0) = x_0$, hat die Form

$$x(t) = Q(t)x_0 + \int\limits_0^t Q(t-s)b(s)ds, \quad \forall t \in I,$$

wobei $Q(\cdot)$ die Fundamentalmatrix mit $Q(0) = I$ ist und die Gestalt

$$Q(t) := \sum_{j=1}^{s} (B_j(t)\cos\omega_j t + C_j(t)\sin\omega_j t)\exp(\mu_j t) + \sum_{j=2s+1}^{p} D_j(t)\exp(\lambda_j t),$$

besitzt, wobei $\lambda_j = \mu_j + i\,\omega_j$, $j = 1,\ldots,s$, und die $B_j(\cdot)$, $C_j(\cdot)$, $D_j(\cdot)$ (reelle) Polynom-Matrizen vom Grad kleiner als η_j sind.

Beweis:
Es sei J die Jordansche Normalform von A und X die zugehörige Transformationsmatrix, d.h. es gelte $J = X^{-1}AX$. Wir betrachten zunächst das homogene DGL-System $x'(t) = Ax(t)$, $t \in I$, führen die neue Variable $y(t) := X^{-1}x(t)$, $t \in I$, ein und erhalten

$$y'(t) = X^{-1}x'(t) = X^{-1}Ax(t) = X^{-1}AXX^{-1}x(t) = Jy(t), \forall t \in I.$$

Wir bezeichnen nun wie in Satz 4.74 mit $J_{\ell_{ij}}(\lambda_i) \in \mathbb{C}^{\ell_{ij}\times\ell_{ij}}$, $j = 1,\ldots,\gamma_i$, die Jordan-Kästchen, die zum Eigenwert $\lambda_i \in \mathbb{C}$, $i = 1,\ldots,p$, gehören. Wir bezeichnen mit $y_i(\cdot)$ den zum Eigenwert λ_i gehörigen Teil von $y(\cdot)$ mit $a_i = \sum\limits_{j=1}^{\gamma_i} \ell_{ij}$ Komponenten, wobei a_i die algebraische Vielfachheit von λ_i, $i = 1,\ldots,p$, ist und natürlich $\sum\limits_{i=1}^{p} a_i = m$ gilt. Bezeichnen wir mit $y_{ij}(\cdot)$, $j = 1,\ldots,\gamma_i$, die Teile von $y_i(\cdot)$, die den Jordan-Kästchen $J_{\ell_{ij}}(\lambda_i)$ zugeordnet sind, so ergibt sich für diese das folgende DGL-System

$$y'_{ij}(t) = J_{\ell_{ij}}(\lambda_i)y_{ij}(t) = \begin{pmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \ddots & \lambda_i & 1 \\ 0 & \cdots & 0 & \lambda_i \end{pmatrix} y_{ij}(t),\ \forall t \in I.$$

bzw. komponentenweise $\quad y'_{ij,k} = \lambda_i y_{ij,k} + y_{ij,k+1}$, $k = 1,\ldots,\ell_{ij}-1$, $y'_{ij,\ell_{ij}} = \lambda_i y_{ij,\ell_{ij}}$. Für die letzte Gleichung folgt sofort $y_{ij,\ell_{ij}}(t) = \exp(\lambda_i t)$ und für die anderen Gleichungen analog zu Beispiel 8.24 rekursiv

$$y_{ij,k}(t) = \exp(\lambda_i t)y_{ij,k}(0) + \int_0^t \exp(\lambda_i(t-s))y_{ij,k+1}(s)ds, k = \ell_{ij}-1,\ldots,1,$$

und damit $y_{ij,k}(t) = p_{ijk}(t)\exp(\lambda_i t)$, wobei $p_{ijk}(\cdot)$ ein Polynom vom Grad $\ell_{ij}-k$ ist. Auf diese Weise lassen sich ℓ_{ij} Elemente eines Fundamentalsystems für $y' = Jy$ erzeugen. Insgesamt setzt sich der zum Eigenwert λ_i gehörige Teil $y_i(\cdot)$ von $y(\cdot)$ aus den relevanten Teilen von $a_i = \sum\limits_{j=1}^{\gamma_i} \ell_{ij}$ Elementen eines Fundamentalsystems zusammen und hat die Form

$$y_i(t) = p_i(t)\exp(\lambda_i t),$$

wobei die $p_i(t)$ Vektoren in $\mathbb{C}^{a_i}$ sind, in deren Komponenten sich Polynome vom Grad kleiner als $\eta_i = \max_{j=1,\ldots,\gamma_i} \ell_{ij}$ befinden. Im Fall, daß λ_i reell ist, d.h. für $i = 2s+1,\ldots,p$

ist die Argumentation damit abgeschlossen. Im Fall $i \in \{1, \ldots, 2s\}$ verwenden wir die Darstellungen für y_i

$$y_i(t) = p_i(t)(\cos \omega_i t \pm i \sin \omega_i t) \exp(\mu_i t),$$

bestimmen davon Realteil und Imaginärteil und erhalten zwei reelle Polynom-Vektoren $\bar{p}_i(t)$, $\hat{p}_i(t)$ in $\mathbb{R}^{a_i}$, so daß die zu λ_i bzw. $\bar{\lambda}_i$ gehörigen reellen Funktionen

$$\bar{y}_i(t) = \bar{p}_i(t) \cos \omega_i t \exp(\mu_i t) \quad \text{bzw.} \quad \hat{y}_i(t) = \hat{p}_i(t) \sin \omega_i t \exp(\mu_i t)$$

ebenfalls Teile von Lösungen sind. Ebenso verfährt man mit den Elementen des Fundamentalsystems. Abschliessend transformieren wir die Lösungen bzw. das Fundamentalsystem durch Multiplikation mit der Matrix X zurück zu Lösungen von $x^{'}(\cdot) = Ax(\cdot)$. Wir betrachten nun die Matrix $Q(t)$, in deren Spalten gerade die Elemente des eben konstruierten (reellen) Fundamentalsystems stehen und die die Eigenschaft $Q(0) = I$ besitzt. Das homogene DGL-System $x^{'}(\cdot) = Ax(\cdot)$ mit konstanten Koeffizienten besitzt aber auf ganz $\mathbb{R}$ Lösungen und die Lösung $Q(t)x_0$ zum Anfangswert $x(0) = x_0$ ist wegen der Einzigkeit identisch mit derjenigen, bei der man zuerst von x_0 bis $t = s$ integriert und anschliessend von $t = s$ bis t, d.h. mit $Q(t-s)Q(s)x_0$. Also gilt $Q(t) = Q(t-s)Q(s)$ und damit $Q(t)Q^{-1}(s) = Q(t-s)$ für alle $t, s \in \mathbb{R}$. Die Darstellung der Lösung entnimmt man Satz 8.9. $\square$

Beispiel 8.12 *(mechanische Schwingungen)*
Wir untersuchen die Bewegung eines Massepunktes M mit Masse m, der an einer Feder befestigt ist, entlang der x-Achse. Befindet sich der Massepunkt M zur Zeit t in $x(t)$, so ist seine Beschleunigung $mx^{''}(t)$ gleich der Summe der Rückstellkraft $-kx(t)$, der Dämpfungskraft $-rx^{'}(t)$ und der äußeren Kraft $F(t)$. Es ergibt sich also die skalare DGL 2. Ordnung

$$mx^{''}(t) + rx^{'}(t) + kx(t) = F(t), \ t \in I = [0, T].$$

Dabei interessiert uns der Fall der periodischen Erregung, d.h. $F(t) = F_0 \cos \omega_0 t$, und der Fall "kleiner" Reibung, d.h. $r^2 < 4km$.

Ziel: Explizite Berechnung von $x(\cdot)$!
Wie in Bemerkung 8.1 transformieren wir die skalare DGL 2. Ordnung in DGL-System 1. Ordnung. Es entsteht

$$\frac{d}{dt}\begin{pmatrix} x \\ x^{'} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{pmatrix}}_{=:A} \begin{pmatrix} x \\ x^{'} \end{pmatrix} + \underbrace{\begin{pmatrix} 0 \\ \frac{F(t)}{m} \end{pmatrix}}_{=:b(t)}$$

Zunächst bestimmen wir die Eigenwerte von A:
$$\lambda I - A = \begin{pmatrix} \lambda & -1 \\ \frac{k}{m} & \lambda + \frac{r}{m} \end{pmatrix} \rightsquigarrow \det(\lambda I - A) = \lambda(\lambda + \tfrac{r}{m}) + \tfrac{k}{m}.$$
Wir erhalten also die folgenden Eigenwerte von A:

$$\begin{aligned} \lambda_{1/2} &= -\frac{r}{2m} \pm \sqrt{\left(\frac{r}{2m}\right)^2 - \frac{k}{m}} = -\frac{r}{2m} \pm \frac{1}{2m}\sqrt{r^2 - 4km} \\ &= \underbrace{-\frac{r}{2m}}_{=:\mu} \pm i \underbrace{\frac{1}{2m}\sqrt{4km - r^2}}_{=:\omega}. \end{aligned}$$

Aus Satz 8.11 resultiert die folgende Gestalt der dortigen Matrixfunktion $Q(t)$:

$$Q(t) = (C_1 \cos \omega t + C_2 \sin \omega t) \exp(\mu t),\ t \in I,$$

mit gewissen Matrizen $C_1, C_2 \in \mathbb{R}^{2\times 2}$.
Wir berechnen nun die Matrizen C_1, C_2 aus den Anfangsbedingungen $Q(0) = I$ und $\frac{d}{dt}Q(0) = AQ(0) = A$. Mit Hilfe von

$$\frac{d}{dt}Q(t) = \omega(-C_1 \sin \omega t + C_2 \cos \omega t) \exp(at) + a((C_1 \cos \omega t + C_2 \sin \omega t) \exp(at)$$

erhalten wir die Gleichungen $C_1 = Q(0) = I$ und $aC_1 + \omega C_2 = A$.
$\rightsquigarrow$ $C_1 = I$ *und* $C_2 = \frac{1}{\omega}(A - aI)$.

$$\begin{aligned} Q(t) &= \left(I \cos \omega t + \frac{1}{\omega}(A - aI) \sin \omega t\right) \exp(at) \\ &= \begin{pmatrix} \cos \omega t - \frac{a}{\omega} \sin \omega t & \frac{1}{\omega} \sin \omega t \\ -\frac{k}{m\omega} \sin \omega t & \cos \omega t - (\frac{r}{m\omega} + \frac{a}{\omega}) \sin \omega t \end{pmatrix} \exp(at) \end{aligned}$$

$$\int_0^t Q(t-s)b(s)ds = \int_0^t \begin{pmatrix} \frac{1}{\omega} \exp(a(t-s)) \sin \omega(t-s) \frac{F_0}{m} \cos \omega_0 s \\ \exp(a(t-s))(\cos \omega(t-s) - (\frac{r}{m\omega} + \frac{a}{\omega}) \sin \omega(t-s)) \frac{F_0}{m} \cos \omega_0 s \end{pmatrix} ds.$$

$$\begin{aligned} \rightsquigarrow x(t) &= \left[(\cos \omega t - \frac{a}{\omega} \sin \omega t) x(0) + \frac{1}{\omega} \sin \omega t\, x'(0)\right] \exp(at) \\ &\quad + \frac{F_0}{\omega m} \int_0^t \sin \omega(t-s) \exp(a(t-s)) \cos \omega_0 s\, ds\,,\ \forall t \in [0,T]. \end{aligned}$$

ist Lösung der DGL $\quad mx''(t) + rx'(t) + kx(t) = F_0 \cos \omega_0 t,\ \forall t \in [0,T]$.

Der Einfachheit halber betrachten wir ab jetzt den Spezialfall, daß die Dämpfung vernachlässigbar ist, d.h. $r = 0 \rightsquigarrow \mu = 0$, $\omega = \sqrt{\frac{k}{m}}$. Dann hat die Lösung die Form

$$x(t) = \cos \omega t\, x(0) + \frac{1}{\omega} \sin \omega t\, x'(0) + \frac{F_0}{\omega m} \int_0^t \sin \omega(t-s) \cos \omega_0 s\, ds\,,\ \forall t \in [0,T].$$

Es gilt: $\sin \omega(t-s) = \sin \omega t \cos \omega s - \cos \omega t \sin \omega s$ *(nach Additionstheorem).*

$\rightsquigarrow \int_0^t \sin \omega(t-s) \cos \omega_0 s\ ds = \sin \omega t \int_0^t \cos \omega s \cos \omega_0 s\ ds - \cos \omega t \int_0^t \sin \omega s \cos \omega_0 s\ ds.$

1. Fall: $\omega \neq \omega_0 \rightsquigarrow$ *die beiden obigen Integrale berechnen sich dann zu*

$$\left[\frac{\sin(\omega - \omega_0)s}{2(\omega - \omega_0)} + \frac{\sin(\omega - \omega_0)s}{2(\omega - \omega_0)}\right]_0^t \ \ bzw. \ \ -\left[\frac{\cos(\omega - \omega_0)s}{2(\omega - \omega_0)} + \frac{\cos(\omega - \omega_0)s}{2(\omega - \omega_0)}\right]_0^t.$$

$\rightsquigarrow$ *$x(\cdot)$ ist eine Funktion, die polynomial in $\sin \omega t, \cos \omega t, \sin \omega_0 t, \cos \omega_0 t$ ist.*
$\rightsquigarrow$ *$x(\cdot)$ ist oszillierend, aber mit fester Amplitude.*
2. Fall: $\omega = \omega_0$.

$$\begin{aligned} \rightsquigarrow x(t) &= \cos \omega_0 t \, x(0) - \frac{1}{\omega_0} \sin \omega_0 t \, x'(0) \\ &\quad + \frac{F_0}{\omega_0 m} \left[\sin \omega_0 t \int_0^t (\cos \omega_0 s)^2 ds - \cos \omega_0 t \int_0^t \sin \omega_0 s \cos \omega_0 s \; ds \right], \end{aligned}$$

und diese beiden Integrale berechnen sich dann zu

$$\left[\frac{1}{2} s \frac{1}{4\omega_0} \sin 2\omega_0 s \right]_0^t \quad \textit{bzw.} \quad \left[\frac{1}{2\omega_0} (\sin \omega_0 s)^2 \right]_0^t$$

$\rightsquigarrow$ *Die Lösung $x(\cdot)$ enthält den kritischen Term $\frac{F_0}{2\omega_0 m} t \sin \omega_0 t$, d.h. einen oszillierenden Term mit wachsender Amplitude!*
Bei vorhandener Dämpfung $r > 0$ entsteht der Term $\frac{F_0}{2\omega_0 m} t \exp\left(\frac{-r}{2m} t\right) \sin \omega_0 t$. Bei kleinem $\frac{r}{2m} > 0$ überwiegt zunächst das Wachsen von t, und nur asymptotisch konvergiert die Amplitude gegen 0.
Dies ist der sog. Resonanzeffekt bzw. die sog. Resonanzkatastrophe (d.h. die Frequenz der äußeren periodischen Erregung ist gleich der Eigenfrequenz ω!)

8.3 Existenz und Eindeutigkeit von Lösungen von Anfangswertaufgaben für gewöhnliche Differentialgleichungen

Problem: (1) $\quad x'(t) = f(x(t), t), \quad \forall t \in I, \; x(t_0) = x_0,$

wobei $I \subseteq \mathbb{R}$ ein Intervall, $t_0 \in I$, $x_0 \in \mathbb{R}^m$ und $f : \mathbb{R}^m \times I \to \mathbb{R}^m$.

Beispiel 8.13
Wir betrachten die Anfangswertaufgabe $\quad x'(t) = x^2(t), \; \forall t \in \mathbb{R}, \; x(1) = -1.$
Lösung auf $]0, +\infty[$: $\quad x(t) = \frac{1}{t}$. Diese Funktion ist nicht fortsetzbar auf $t = 0$!
I. a. kann man also nur die lokale Lösbarkeit erwarten:
$\rightsquigarrow$ *präzisierte Problemstellung: Existenz und Einzigkeit von Lösungen von (1) auf "kleinen" Intervallen um t_0.*

Satz 8.14 *(Cauchy-Peano)*
Es sei $I \subseteq \mathbb{R}$ ein Intervall, es seien $t_0 \in I$, $x_0 \in \mathbb{R}^m$, $r > 0$ und $f : \mathbb{R}^m \times I \to \mathbb{R}^m$ sei stetig auf $B(x_0, r) \times I$. Dann existieren ein Intervall $I_0 \subseteq I$ mit $t_0 \in I_0$ und eine differenzierbare Funktion $x : I_0 \to \mathbb{R}^m$ mit

$$x'(t) = f(x(t), t), \quad \forall t \in I_0, \; x(t_0) = x_0.$$

(ohne Beweis)

Frage: Ist unter den Voraussetzungen von Satz 8.14 die Lösung von (1) eindeutig ?

Beispiel 8.15
Wir betrachten die Anfangswertaufgabe $x'(t) = x(t)^{\frac{1}{3}}, \forall t \in [0, +\infty[, x(0) = 0.$
(Satz 8.7: $f(y,t) = y^{\frac{1}{3}}, \forall (y,t) \in \mathbb{R} \times [0, +\infty[\rightsquigarrow f$ *ist stetig* $\rightsquigarrow \exists$ *"lokale" Lösung.)*
Lösung: (Verifikation durch Nachrechnen)

$$x(t) := (\frac{2}{3}t)^{\frac{3}{2}}, \forall t \in [0, +\infty].$$

Es existieren aber unendlich viele Lösungen:

$$x_c(t) := \begin{cases} (\frac{2}{3}t)^{\frac{3}{2}}, & \forall t \in [c, +\infty[\\ 0, & \forall t \in [0, c[\end{cases} \qquad (\forall c \geq 0).$$

Ursache: Wir benötigen stärkere Stetigkeitseigenschaften von f*, um eindeutige Lösungen zu erhalten!*

Satz 8.16 *(Picard-Lindelöf)*
Es sei $I \subseteq \mathbb{R}$ *ein Intervall, es seien* $t_0 \in I, x_0 \in \mathbb{R}^m, r > 0,$ *und* $f : \mathbb{R}^m \times I \to \mathbb{R}^m$ *sei stetig auf* $B(x_0, r) \times I$*. Ferner existiere eine Konstante* $L > 0$ *mit* $\|f(y,t) - f(z,t)\| \leq L\|y - z\|, \forall (y,t), (z,t) \in \bar{B}(x_0, r) \times I.$
Dann existieren ein Intervall $I_0 \subseteq I$ *mit* $t_0 \in I_0$ *und eine eindeutig bestimmte differenzierbare Funktion* $x : I_0 \to \mathbb{R}^m$*, so daß*

$$x'(t) = f(x(t), t), \quad \forall t \in I_0, \, x(t_0) = x_0.$$

(ohne Beweis)
Insbesondere erfüllt die bei linearen DGLn auftretende Funktion

$$f(x,t) := A(t)x + b(t) \quad (x \in \mathbb{R}^m, \, t \in I)$$

mit stetigen $A : I \to \mathbb{R}^{m \times m}$ und $b : I \to \mathbb{R}^m$ die Vorausetzungen von Satz 8.16. Die Lipschitzstetigkeit von $f(\cdot, t)$ gilt auf jeder Kugel $B(x_0, r)$ und jedem kompakten Intervall I mit der Konstanten $L := \max_{t \in I} \|A(t)\|$.

Da wir bisher in den Sätzen 8.14 und 8.16 nur Lösungen auf kleinen Intervallen, die t_0 enthalten, erhalten haben, also "lokale" Lösungen, besteht unser nächstes Ziel in der Möglichkeit der Fortsetzung lokaler Lösungen. Unser nächstes Beispiel zeigt aber, daß nur Kleinigkeiten darüber entscheiden können, ob eine Fortsetzung zu globalen Lösungen, d.h., auf ganz I definierten Lösungen möglich ist.

Beispiel 8.17

a) $x'(t) = -2tx(t)^2, \forall t \in \mathbb{R}, x(0) = 1.$
Wir betrachten $x_*(t) = \frac{1}{1+t^2}, \forall t \in \mathbb{R}.$
$\rightsquigarrow (\mathbb{R}, x_*)$ *ist globale Lösung dieser Aufgabe (da* $x_*'(t) = -2t(1+t^2)^{-2}$*).*

b) $x'(t) = 2tx(t)^2, \forall t \in \mathbb{R}, x(0) = 1.$
Ansatz: $x_*(t) = (1 - t^2)^{-1}, \forall t \in]-1, 1[.$
$\rightsquigarrow (]-1, 1[, x_*)$ *ist maximale Lösung dieser Aufgabe (da* $x_*'(t) = 2t(1-t^2)^{-2}$*).*

Satz 8.18
Es sei $I \subseteq \mathbb{R}$ ein Intervall, $t_0 \in I$ sei linker Randpunkt von I, $x_0 \in \mathbb{R}^m$, und $f : \mathbb{R}^m \times I \to \mathbb{R}^m$ sei stetig. Dann existieren ein Intervall $\tilde{I} \subseteq I$ und eine maximale Lösung $\tilde{x} : \tilde{I} \to \mathbb{R}^m$ von (2.1), die nicht über $\tilde{I}$ hinaus fortgesetzt werden kann, und diese ist entweder eine globale Lösung von (2.1), d.h., $\tilde{I} = I$, oder $\tilde{I}$ hat die Form $[t_0, t_1[$ mit $t_1 \in I$, und $\tilde{x}$ ist nicht beschränkt auf $\tilde{I}$ (d.h. $\tilde{x}$ "explodiert").

(ohne Beweis)
Frage: Welche Bedingungen an die DGL (d.h. an die Funktion f) verhindern eine Explosion von maximalen Lösungen?

Lemma 8.19 *(Gronwall)*
Es seien w, g stetige Funktionen auf $[a, b]$, $c > 0$ und es gelte die "Integralungleichung"

$$0 \leq w(t) \leq g(t) + c \int_a^t w(s)ds, \quad t \in [a, b].$$

Dann gilt: $\quad w(t) \leq \max\limits_{t \in [a,b]} |g(t)| \exp(c(t-a)),\ \forall t \in [a, b].$

Beweis: Wir betrachten die Funktion $u(t) := \max_{t\in[a,b]} |g(t)| + c \int\limits_a^t w(s)ds$, $\forall t \in [a, b]$.
u ist stetig differenzierbar und es gilt $w(t) \leq u(t)$, $u'(t) = cw(t)$, $\forall t \in [a, b]$.
Die Funktion $v(t) := u(t) \exp(-c(t-a))$, $\forall t \in [a, b]$, ist monoton fallend, da
$v'(t) = (u'(t) - cu(t)) \exp(-c(t-a)) \leq 0$, $\forall t \in [a, b]$, gilt.
$\rightsquigarrow v(t) = u(t) \exp(-c(t-a)) \leq v(a) = u(a) = \max_{t\in[a,b]} |g(t)|$, $\forall t \in [a, b]$.
$\rightsquigarrow w(t) \leq u(t) \leq \max_{t\in[a,b]} |g(t)| \exp(c(t-a))$, $\forall t \in [a, b]$. □

Satz 8.20
Es sei $I := [t_0, T]$ $(T > 0)$, $f : \mathbb{R}^m \times I \to \mathbb{R}^m$ stetig, und es existiere ein Skalarprodukt $\langle \cdot, \cdot \rangle$ auf $\mathbb{R}^m$ mit zugehöriger Norm $\| \cdot \|$ und eine beschränkte Riemann-integrierbare Funktion $\ell : I \to \mathbb{R}^m$ so, daß

$$\langle f(y, t), y \rangle \leq \ell(t)(1 + \|y\|^2),\ \forall (y, t) \in \mathbb{R}^m \times I.$$

Dann existiert eine globale Lösung von (2.1).
Insbesondere gilt dies, falls f "lineares Wachstum" besitzt, d.h.
$\|f(y, t)\| \leq \ell(t)(1 + \|y\|)$, $\forall (y, t) \in \mathbb{R}^m \times I$, mit einer beliebigen Norm in $\mathbb{R}^m$, gilt.

Beweis:
Wir wenden Satz 8.18 an und zeigen, daß jede lokale Lösung von (2.1) beschränkt ist. Satz 8.18 besagt dann die Existenz einer maximalen Lösung, die globale Lösung sein muß. Es sei also $(\tilde{I}, x)$ eine lokale Lösung von (2.1).
Wir betrachten die Funktion $w : I \to \mathbb{R}$, $w(t) := \|x(t)\|^2$, $\forall t \in \tilde{I}$.
Diese ist differenzierbar und es gilt $w'(t) = 2\langle x'(t), x(t) \rangle$, $\forall t \in \tilde{I}$.

$\rightsquigarrow \frac{d}{dt} \|x(t)\|^2 = 2\langle f(x(t), t), x(t) \rangle \leq 2\ell(t)(1 + \|x(t)\|^2,\ \forall t \in \tilde{I}.$
$\rightsquigarrow \|x(t)\|^2 - \|x(t_0)\|^2 \leq 2 \int\limits_{t_0}^t \ell(s)(1 + \|x(s)\|^2)ds,\ \forall t \in \tilde{I}.$

$$\rightsquigarrow \underbrace{\|x(t)\|^2}_{=:w(t)} \leq \underbrace{\left[\|x(t_0)\|^2 + 2\int_{t_0}^{t} \ell(s)ds\right]}_{=:g(t)} + 2c\int_{t_0}^{t} \|x(s)\|^2 ds, \quad \forall t \in \tilde{I}.$$

wobei $c := \sup\limits_{s\in I} |\ell(s)|$. Aus Lemma 8.19 folgt dann
$w(t) = \|x(t)\|^2 \leq \sup\limits_{t\in I} |g(t)| \exp(2c(t-t_0)),\ \forall t \in \tilde{I}$,
d.h. die beliebig gewählte lokale Lösung ist beschränkt! Der Nachsatz folgt im Fall, daß die Normen identisch sind, aus der Ungleichung

$$\langle f(y,t), y\rangle \leq \|f(y,t)\|\|y\| \leq \ell(t)(1+\|y\|)\|y\| \leq 2\ell(t)(1+\|y\|^2),\ \forall y \in \mathbb{R}^m.$$

Sind die Normen verschieden, gehen zusätzlich die Normäquivalenz-Konstanten ein. $\square$

Beispiel 8.21

Gleichung der Populationsdynamik: $x'(t) = ax(t) - bx(t)^2$, $t \in [t_0, \infty[$, $x(t_0) = x_0$.
Die rechte Seite f *der Differentialgleichung hat die Form*

$$f(y,t) := \begin{cases} ay - by^2 & , y \in [0, \frac{a}{b}] \quad (a > b > 0) \\ 0 & , \text{sonst} \end{cases}$$

$\rightsquigarrow f : \mathbb{R} \times I \to \mathbb{R}$ *ist stetig und es gilt wegen* $\langle r, \tilde{r}\rangle := r\tilde{r}$, $\forall r, \tilde{r} \in \mathbb{R}$:

$$\rightsquigarrow f(y,t)y = \begin{cases} ay^2 - by^3 \leq ay^2 \leq a(1+y^2) & , \quad y \in [0, \frac{a}{b}] \\ 0 & , \quad \text{sonst} \end{cases}$$

8.20 $\rightsquigarrow$ *die DGL besitzt auf jedem Intervall der Form* $[t_0, T]$ *eine Lösung.*